Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.
Surface Proteins and Restriction-Modification Systems of Bovine Mastitis Strains of *Staphylococcus aureus*

A thesis submitted for the degree of Doctor in Philosophy

by

Tara Louise Sweeney

Moyne Institute of Preventive Medicine

Department of Microbiology

Trinity College Dublin

March 2013
Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work except where it is duly acknowledged in the text.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.
Acknowledgements

First and foremost I would like to express my profound gratitude to my dedicated supervisor Prof. Tim Foster for giving me the opportunity to join his research group. His constant guidance and encouragement throughout the past few years has been invaluable and I am forever indebted. A special thanks to Ian Monk for all his fantastic advice and encouragement, especially during my final year. Thanks also to the members of my thesis committee for their help and advice. For financial support during this project I thank both Trinity College Dublin and Science Foundation Ireland.

To all the members of the TJF lab “The Fosterettes” both past and present including; Dee, Joan, Helen, Rebecca, Fiona, Emma, Marta, Michelle, Niamh, Simon, Ian and Laura, thank you all for making each day that little bit brighter and of course for the never-ending supply of advice, moral support and cake, especially my “downstairs” comrades Marta, Emma and Simon for the chats and giggles. A big thank you to all the members of the other lab groups, especially Jimmy and Suzanne, for providing friendly faces and a fresh perspective. Thank you also to everybody in the Moyne in particular Jayne, Noreen and all the members of the prep room especially Ronan and Stephen. Stephen came to be a very important part of my daily routine as he was personally involved in getting a different song stuck in my head each day!

I would also like to thank all my friends in particular Róisin, Andrea and the Newpark “girlies” for all their support and tea over the years, my college buddies Jacqui, Jen, Jimmy and Leo for always making me laugh, my adopted family the Stellar Jets for putting a song in my heart, and to my beloved Keith, for his support, love and ability to make me smile everyday, thank you for being you!

Finally I thank my nearest and dearest. My brother and his family in Germany, who I miss constantly, my parents for shaping me into an independent and strong woman, especially my mother who kept me in constant supply of love, dinners and hugs throughout this thesis. A special thanks to my sister from another mister, my main damie Michelle, who not only was a wonderful member of my lab group but is my dearest and oldest friend, without who I would not be the person I am today. And finally, to my sister Carol, for her constant love, support and belief throughout my entire college education, who I could not have gotten through any of this without, this thesis is for you.
Summary

*Staphylococcus aureus* is frequently associated with bovine mastitis, one of the most economically important diseases in the dairy industry. The bovine *S. aureus* strain RF122 represents a geographically widespread clone that is particularly well adapted to the bovine host. Previous studies showed that extensive gene decay occurred in the genome of RF122, revealed by the level of pseudogenes detected during annotations. A logical approach for an improved means of controlling bovine mastitis is rational design of a staphylococcal vaccine. *S. aureus* is characterised by a wide array of virulence factors and antigenic diversity among different strains has hampered the development of an effective vaccine, therefore prevention of intramammary infection has not been demonstrated. Surface proteins singly and in combination have shown potential in combating diseases caused by *S. aureus* in humans. Analysis of the surface proteins of bovine strains may provide insights into the requirements for life in the mammary gland.

Bioinformatic analysis of the surface protein genes of RF122 revealed that many genes are either missing or exist as pseudogenes but the majority of *isd* genes are intact. This correlated with Western immunoblotting results which showed that several surface proteins are not expressed under the growth conditions tested. ClfB is expressed at very low levels by RF122 and cannot support adhesion, although the recombinantly expressed A domain can bind fibrinogen with the same affinity as the archetypal ClfB protein. Similarly the single fibronectin-binding protein is expressed at too low a level to promote adhesion. RF122 expresses bone sialoprotein-binding protein Bbp an allelic variant of SdrE. The majority of other bovine strains express SdrE.

To conclude, the results indicate that RF122 exhibits gene decay, which supports previous observations. Only a few surface protein genes have been retained and are expressed by RF122 (IsdA, IsdB, ClfB and Bbp) whereas the majority of bovine strains from other multilocus sequence types still express proteins associated with strains from human hosts (IsdH, ClfA, SdrE). The data implies that RF122's progenitor might have colonised the bovine earlier than some of the others and might be better adapted. IsdA may confer resistance to bactericidal lipids and protect against lactoferrin. IsdB may provide an alternative mechanism of invasion (to the major FnBP-dependent mechanism) into mammary gland epithelial cells. Bbp/SdrE bind complement factor H which may facilitate evasion of complement-mediated phagocytosis in the bovine host. Results in this thesis support the choice of IsdA, IsdB, ClfB and SdrE/Bbp as vaccine candidates that
will cover RF122 as well as the majority of other bovine STs. Recombinant proteins were expressed, purified and characterized and isogenic mutants of RF122 constructed to prepare for a possible vaccine trial. Difficulties in transforming plasmid DNA into RF122, to facilitate allelic exchange, led to the investigation of restriction-modification systems of bovine \textit{S. aureus} strains.

The major barrier to DNA transfer by the majority of \textit{S. aureus} strains of human origin is the type IV restriction enzyme SauUSI which recognises and degrades cytosine methylated DNA. Circumvention of this barrier allows the manipulation of many \textit{S. aureus} strains. However, RF122 cannot be transformed in this manner. A spontaneous derivative RF122t would accept plasmid DNA propagated in \textit{E. coli}. The role of restriction-modification systems in protecting RF122 from foreign DNA and the reason why RF122t would accept \textit{E. coli}-derived plasmid DNA were investigated. Southern blotting showed that the gene encoding SauUSI is present in the majority of bovine strains, but not in RF122. RF122 was found to employ a putative type II restriction system comprising an endonuclease (Sab2370c) and a modification cytosine methylase (Dcm). The \textit{dcm} and \textit{sb}ab2370c genes were located in the genome at the site occupied by \textit{sauUSI} in other strains. The \textit{sb}ab2370c gene was deleted by allele exchange and the mutant became transferable by plasmids from \textit{E. coli} XL-1 and \textit{S. aureus} RN4220. It was not possible to delete the \textit{dcm} gene most likely because such a mutation would be lethal. The \textit{sb}ab2370c gene from RF122t was found to have a premature stop codon which encodes a variant C terminus. The RF122 \textit{sb}ab2370c gene was cloned into an inducible expression vector in \textit{E. coli} Dcm'. Following induction the cells stopped growing and were rapidly killed. Co-expression of Dcm from RF122 with Sab2370c failed to allow growth but did allow survival of \textasciitilde1\% of the cells in the population.

The results imply a direct exchange of the type IV and type II restriction systems in RF122. The premature stop codon in the \textit{sb}ab2370c gene could be responsible for defective function of RF122t. These data are consistent with the RF122 \textit{dcm} gene encoding a modification methylase that protects DNA from cleavage by the endonuclease encoded by \textit{sb}ab2370c. Understanding of restriction systems that protect bovine \textit{S. aureus} strains will facilitate future genetic manipulation of these strains allowing identification of virulence factors that contribute to bovine mastitis.
# Contents

Declaration ....................................................................................................................... i  
Acknowledgements ......................................................................................................... ii 
Summary .......................................................................................................................... iii 
List of Tables ................................................................................................................... x  
List of Figures .................................................................................................................. xi 
Key to commonly used abbreviations ........................................................................ xiv 

Chapter 1 Introduction ................................................................................................. 1 
1.1 Biology of the staphylococci .................................................................................. 2  
  1.1.1 Classification and identification .................................................................... 2  
  1.1.2 Colonisation and disease ................................................................................ 3  
    1.1.2.1 Colonisation and disease in humans ................................................ 3  
    1.1.2.2 Colonisation and disease in animals ................................................ 4  
    1.1.2.3 Bovine mastitis ................................................................................... 5  
1.2 Genome stmcture of *Staphylococcus aureus* ......................................................... 6  
  1.2.1 Core genome .................................................................................................. 7  
  1.2.2 Accessory genomic elements ......................................................................... 7  
  1.2.3 Genotyping ...................................................................................................... 8  
  1.2.4 MLST typing .................................................................................................. 8  
  1.2.5 Population structure of 5. ..................................................................  9 
1.3 *S. ai/rew.v* cell wall-associated virulence factors .................................................... 11  
  1.3.1 Cell wall-associated surface proteins ............................................................. 12  
    1.3.1.1 Sortase-mediated anchoring of cell wall-associated proteins .......... 13  
    1.3.1.2 Protein A ............................................................................................. 13  
    1.3.1.3 Iron regulated surface determinants .................................................. 16  
      1.3.1.3.1 Haem transfer pathway ....................................................... 17  
      1.3.1.3.2 NEAT domains of lsd proteins ........................................... 17  
      1.3.1.3.3 Other functions of surface-exposed lsd proteins ............... 19  
    1.3.1.4 Clumping factor A .............................................................................. 20  
    1.3.1.5 Clumping factor B ............................................................................... 22  
    1.3.1.6 Sdr proteins ......................................................................................... 24  
    1.3.1.7 Fibronectin-binding surface proteins ................................................. 25
Chapter 2 Materials and Methods

2.1 Chemicals and reagents ................................................................. 38

2.2 Bioinformatic analysis ................................................................. 38
   2.2.1 Bioinformatic and phylogenetic analysis .................................. 38
   2.2.2 Blast search ............................................................................ 38

2.3 Bacterial strains and growth conditions ........................................ 39

2.4 Detecting expression of S. aureus cell-wall associated proteins .......... 39
   2.4.1 Preparation of solubilised proteins associated with the staphylococcal cell wall ......................................................... 39
   2.4.2 Antibodies to S. aureus Bbp antibodies .................................... 40
   2.4.3 Western immunoblotting .......................................................... 40
   2.4.4 Ligand-affinity blotting ............................................................. 41

2.5 Adherence of bacteria to immobilized fibronectin, cytokeratin 10 and fibrinogen 41

2.6 DNA manipulation ........................................................................ 42
   2.6.1 DNA manipulation ................................................................. 42
   2.6.2 Isolation of plasmid and genomic DNA .................................... 42
   2.6.3 Polymerase chain reaction ...................................................... 42
Chapter 3 Surface proteins of bovine mastitis strains of *Staphylococcus aureus* 51

3.1 Introduction ................................................................. 52
3.2 Results ........................................................................ 56
   3.2.1 Assessing the status of surface proteins of *S. aureus* RF122 ..... 56
   3.2.2 Surface-exposed Iron regulated surface determinant (Isd) proteins ..... 57
      3.2.2.1 IsdA ................................................................. 57
      3.2.2.2 IsdB ................................................................. 58
      3.2.2.3 IsdH ................................................................. 59
      3.2.2.4 Other *isd* proteins ........................................ 60
   3.2.3 SD repeat family proteins ........................................ 60
3.2.3.1 ClfA ................................................................. 60
3.2.3.2 ClfB ................................................................. 61
3.2.3.3 SdrE/Bbp .......................................................... 62

3.2.3.3.1 Investigating the distribution of sdrE or bhp genes in a number of bovine strains by cloning and sequencing the A domains ......................................................... 63
3.2.3.4 SdrC .................................................................. 63
3.2.4 Fibronectin-binding proteins ......................................................... 64
3.2.5 Sas proteins ............................................................................ 65
3.2.6 Protein A ............................................................................... 65
3.2.7 Phylogenetic analysis ................................................................. 66
3.3 Discussion ................................................................................... 67

Chapter 4 Further studies on bovine mastitis vaccine candidate antigens .... 73
4.1 Introduction .................................................................................. 74
4.2 Results .......................................................................................... 80

4.2.1 Cloning vaccine candidate genes into the pQE30 expression vector for the purification of recombinant proteins ................................................................. 80
4.2.1.1 Cloning isdB into the pQE30 expression vector for the purification of recombinant IsdB ................................................................. 80
4.2.1.2 Cloning clfB into the pQE30 expression vector for the purification of recombinant ClfB ................................................................. 81
4.2.1.3 Cloning bhp into the pQE30 expression vector for purification of recombinant Bbp ................................................................. 82
4.2.2 A multicomponent vaccine comprising IsdA, IsdB, ClfB and Bbp .......... 83
4.2.3 Deletion of isdA, isdB, clfB and bhp from the RF122 chromosome .......... 83
4.3 Discussion ................................................................................... 86
Chapter 5 The restriction-modification system of *Staphylococcus aureus* RF122

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>90</td>
</tr>
<tr>
<td>5.2 Results</td>
<td>94</td>
</tr>
<tr>
<td>5.2.1 Restriction-modification genes of RF122</td>
<td>94</td>
</tr>
<tr>
<td>5.2.2 The use of RF122t as an intermediate strain to facilitate genetic</td>
<td></td>
</tr>
<tr>
<td>manipulation of RF122</td>
<td>95</td>
</tr>
<tr>
<td>5.2.3 Determining the distribution of <em>sab2370c</em>, <em>dcm</em> and <em>sauUSI</em> by</td>
<td></td>
</tr>
<tr>
<td>Southern dot blot hybridisation</td>
<td>96</td>
</tr>
<tr>
<td>5.2.4 Allelic replacement mutagenesis of <em>sab2370c</em></td>
<td>97</td>
</tr>
<tr>
<td>5.2.5 Rendering RF122 electrocompetent</td>
<td>98</td>
</tr>
<tr>
<td>5.2.6 Transformation</td>
<td>99</td>
</tr>
<tr>
<td>5.2.7 Attempted deletion of the <em>dcm</em> gene in RF122</td>
<td>99</td>
</tr>
<tr>
<td>5.2.8 Cloning <em>sab2370c</em> into an expression vector in <em>E. coli</em></td>
<td>100</td>
</tr>
<tr>
<td>5.2.9 Sab2370c alters the growth of <em>E. coli</em> cells</td>
<td>100</td>
</tr>
<tr>
<td>5.2.10 Co-expression of Dcm and Sab2370c in <em>E. coli</em></td>
<td>101</td>
</tr>
<tr>
<td>5.3 Discussion</td>
<td>103</td>
</tr>
</tbody>
</table>

Chapter 6 Discussion

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Discussion</td>
<td>110</td>
</tr>
<tr>
<td>6.1.1 Niche adaptation</td>
<td>110</td>
</tr>
<tr>
<td>6.1.1.1 Niche adaptation; gene decay</td>
<td>111</td>
</tr>
<tr>
<td>6.1.1.2 Niche adaptation; gene retention</td>
<td>113</td>
</tr>
<tr>
<td>6.1.2 Vaccine</td>
<td>115</td>
</tr>
<tr>
<td>6.1.2.1 A proposed surface-protein vaccine for combatting bovine mastitis</td>
<td>115</td>
</tr>
<tr>
<td>6.1.2.2 Non-LPXTG vaccines/antigens</td>
<td>116</td>
</tr>
<tr>
<td>6.1.3 Restriction systems of bovine <em>S. aureus</em> strains</td>
<td>117</td>
</tr>
<tr>
<td>6.1.4 New approaches for genetically manipulating <em>S. aureus</em></td>
<td>118</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Bacterial strains</td>
<td>40</td>
</tr>
<tr>
<td>2.2</td>
<td>Antibodies</td>
<td>40</td>
</tr>
<tr>
<td>2.3</td>
<td>Plasmids</td>
<td>42</td>
</tr>
<tr>
<td>2.4</td>
<td>Primers</td>
<td>42</td>
</tr>
<tr>
<td>3.1</td>
<td>Intact genes possessed by <em>S. aureus</em> RF122</td>
<td>58</td>
</tr>
<tr>
<td>3.2</td>
<td>Summary of sequence analysis of <em>sdrE</em> A domains from different bovine <em>S. aureus</em> strains</td>
<td>64</td>
</tr>
<tr>
<td>3.3</td>
<td>Expression profile of bovine <em>S. aureus</em> strains</td>
<td>68</td>
</tr>
<tr>
<td>5.1</td>
<td>Viable counts</td>
<td>102</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Phylogenetic relatedness of *S. aureus* MLST genotypes ................................................. 10
1.2 Structural organisation of staphylococcal surface proteins ............................................. 12
1.3 Surface protein anchoring in *S. aureus* ............................................................................. 14
1.4 Genomic regions encoding the Isd proteins *S. aureus* ..................................................... 16
1.5 The cell-wall associated Isd proteins of *S. aureus* ............................................................ 18
1.6 Putative pathway of haem transfer between Isd proteins of *S. aureus* ......................... 18
1.7 Tandem β-zipper mechanism of fibronectin binding by *S. aureus* ................................. 26
1.8 The dock, lock and latch-binding mechanism ..................................................................... 28
1.9 Restriction modification system types I, II, III and IV ....................................................... 34
2.1 Gateway cloning using pKOR1 .......................................................................................... 44
2.2 Map of pKOR1 .................................................................................................................. 44
3.1 Neighbour-joining consensus tree based upon concatenated sequences of *S. aureus* MLST alleles .............................................................................................................. 52
3.2 Diagrammatic representation of the degree of protein truncation resulting from the introduction of premature stop codons at various points in the gene .......................... 56
3.3 (A) Amino acid sequence alignment of IsdA from *S. aureus* strains Newman and RF122 (B) Expression of IsdA by (i) ST151 and (ii) different bovine *S. aureus* strains (C) Bacterial adhesion to fibronectin and cytokeratin 10 ............................... 58
3.4 Amino acid sequence alignment of IsdB from *S. aureus* strains Newman and RF122 ........................................................................................................................................ 60
3.5 Expression of IsdB from (A) ST151 and (B) a panel of genetically distinct bovine *S. aureus* strains .................................................................................................................. 60
3.6 (A) Amino acid sequence alignment of IsdH from *S. aureus* strains Newman and RF122 (B) Lack of expression of IsdH by (i) ST151 and (ii) a panel of genetically distinct bovine *S. aureus* strains ........................................................................ 60
3.7 (A) Amino acid sequence alignment of ClfA from *S. aureus* strains Newman and RF122 (B) Lack of expression of ClfA by ST151 strains and a panel of genetically distinct bovine *S. aureus* strains (C) Bacterial adhesion to fibrinogen by bovine *S. aureus* strains .......................................................................................................................... 60
3.8 Amino acid sequence alignment of ClfB from *S. aureus* strains Newman and RF122 ................................................................................................................................. 62

3.9 (A) Expression of ClfB from (i) ST151 and (ii) a panel of genetically distinct bovine *S. aureus* strains (B) Bacterial adhesion to cytokeratin 10 .................................. 62

3.10 Amino acid sequence alignment of Bbp from *S. aureus* strains 024 and RF122 .. 62

3.11 (A) Bbp/SdrE expression from bovine *S. aureus* strains (B) sdrE amplimers from non-SdrE expressing strains .................................................................................. 64

3.12 DNA/amino acid sequence alignment of SdrC from *S. aureus* strains Newman and RF122 ............................................................................................................. 64

3.13 Amino acid sequence alignment of FnBPA from *S. aureus* strains 8325-4 and RF122 .................................................................................................................. 64

3.14 (A) Expression of FnBPs from *S. aureus* strains SH1000 and RF122 (B) Bacterial adhesion to fibronectin ...................................................................................... 64

3.15 Amino acid sequence alignment of protein A from *S. aureus* strains Newman and RF122 showing variants residues ........................................................................ 66

3.16 Neighbour-joining tree based upon concatenated sequences of *S. aureus* MLST alleles .................................................................................................................. 66

4.1 Cloning *isdB, clfBN2N3* and *bbpN2N3* DNA from RF122 for the expression of rIsdB, rClfBN2N3 and rBbp protein .................................................................................. 80

4.2 (A) Restriction analysis of the (i) pQE30::*isdB* construct, (ii) pQE30::*clfBN2N3* construct and (iii) pQE30::*bbpN23* (with ClfB-like cleavage site) (B) SDS-PAGE analysis of (i) rIsdB, (ii) rClfBN2N3 and (iii) rBbpN2N3 (C) Binding of (i) rIsdB to immobilised haemoglobin and (ii) rClfBN2N3 to immobilised fibrinogen ............................................................................................................. 80

4.3 3D model of ClfB N2N3 ....................................................................................... 82

4.4 Multiple amino acid sequence alignment of part of the A domains from Bbp, SdrE, ClfA and ClfB from *S. aureus* strains ........................................................................ 82

4.5 (A) Schematic representation of the *isdA, isdB, clfB* and *bbp* mutations (B) Validation of RF122Δ*isdA* (C) Validation of mutant strains ...................................................................... 84

5.1 Genomic organisation comparison of restriction genes of *S. aureus* RF122 and 8325-4 .................................................................................................................. 94

5.2 The use of intermediate *S. aureus* strains to deliver DNA into target *S. aureus* Newman or RF122 strains ............................................................................................... 96
5.3 (A) DNA sequence alignment of the 3’ end of sab2370c from S. aureus strains RF122 and RF122t (B) The predicted amino acid translation of the end of sab2370c from S. aureus (i) RF122 and (ii) RF122t ............................................. 96
5.4 Presence of sab2370c in S. aureus strains by Southern dot blot hybridization ..... 96
5.5 Presence of dcm in S. aureus strains by southern dot blot hybridization .......... 98
5.6 Presence of sauUS1 in S. aureus strains by southern dot blot hybridization .... 98
5.7 (A) Schematic representation of the RF122 sab2370c mutation (B) Validation of RF122Δsab2370c .................................................................................................. 98
5.8 Preparation of electrocompetent S. aureus RF122 cells .................................... 98
5.9 Transformation of RF122 and RF122Δsab2370c with pRMC2 ......................... 100
5.10 Integration of pKOR1Δdcm into the RF122 chromosome ................................ 100
5.11 (A) Cloning sab2370c DNA from S. aureus RF122 for the expression of Sab2370c (B) Restriction analysis of the pQE30::sab2370c construct ............... 100
5.12 Expression of recombinant Sab2370c stalls the growth E. coli XL-1 cells ...... 100
5.13 (A) Cloning dcm from RF122 into pIMK3 (B) Restriction analysis of the plMK3::dcm construct .......................................................... 102
5.14 Growth curve of E. coli DC10B strains .......................................................... 102
### Key to commonly used abbreviations

#### Single letter amino acid code

<table>
<thead>
<tr>
<th>Single letter</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Alanine</td>
</tr>
<tr>
<td>C</td>
<td>Cysteine</td>
</tr>
<tr>
<td>D</td>
<td>Aspartic acid</td>
</tr>
<tr>
<td>E</td>
<td>Glutamic acid</td>
</tr>
<tr>
<td>F</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>G</td>
<td>Glycine</td>
</tr>
<tr>
<td>H</td>
<td>Histidine</td>
</tr>
<tr>
<td>I</td>
<td>Isoleucine</td>
</tr>
<tr>
<td>K</td>
<td>Lysine</td>
</tr>
<tr>
<td>L</td>
<td>Leucine</td>
</tr>
<tr>
<td>M</td>
<td>Methionine</td>
</tr>
<tr>
<td>N</td>
<td>Asparagine</td>
</tr>
<tr>
<td>P</td>
<td>Proline</td>
</tr>
<tr>
<td>Q</td>
<td>Glutamine</td>
</tr>
<tr>
<td>R</td>
<td>Arginine</td>
</tr>
<tr>
<td>S</td>
<td>Serine</td>
</tr>
<tr>
<td>T</td>
<td>Threonine</td>
</tr>
<tr>
<td>V</td>
<td>Valine</td>
</tr>
<tr>
<td>W</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>Y</td>
<td>Tyrosine</td>
</tr>
</tbody>
</table>

#### Nucleotides

<table>
<thead>
<tr>
<th>Single letter</th>
<th>Nucleotide</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>T</td>
<td>Thymine</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>G</td>
<td>Guanine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>bp</td>
<td>base pair(s)</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxy nucleoside triphosphate</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme linked immunosorbent assay</td>
</tr>
<tr>
<td>Fg</td>
<td>fibrinogen</td>
</tr>
<tr>
<td>Fn</td>
<td>fibronectin</td>
</tr>
<tr>
<td>h</td>
<td>hour(s)</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pair</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>min</td>
<td>minute(s)</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>Phyre</td>
<td>Protein Homology/analogY Recognition Engine</td>
</tr>
<tr>
<td>PVDF</td>
<td>polyvinylidene fluoride</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>sec</td>
<td>second(s)</td>
</tr>
<tr>
<td>Tris</td>
<td>trishydroxymethylaminomethane</td>
</tr>
<tr>
<td>TSA</td>
<td>trypticase soy agar</td>
</tr>
<tr>
<td>TSB</td>
<td>trypticase soy broth</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
<tr>
<td>wt</td>
<td>wild-type</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction
1.1 Biology of the staphylococci

1.1.1 Classification and identification

Bacteria of the genus *Staphylococcus* are non-motile, facultative aerobic, Gram-positive cocci. They are approximately 1 μm in diameter and characteristically divide along multiple axes to form grape-like clusters which can be seen under the microscope. The name is from the Greek staphyle for ‘bunch of grapes’ and kokkus for ‘berry’.

Molecular typing and genetic analysis have placed the staphylococci in the *Bacillus-Lactobacillus-Streptococcus* cluster of *Micrococcaceae* (Ludwig *et al*., 1985; Stackebrandt & Teuber, 1988). They are most closely related to *Enterococcus*, *Bacillus* and *Listeria*. Their genomes contain DNA of a low G+C content (30-39%). The staphylococci are resistant to desiccation and are extremely halotolerant growing at up to 3.5 M NaCl.

There are at least 40 staphylococcal species comprising commensals and pathogens of both humans and animals, 9 of which also contain subdivisions with subspecies designations (Hauschild & Stepanovic, 2008). The main classification of staphylococci is based on the ability to produce coagulase, a zymogen that causes blood-clot formation. Only a few species are coagulase-positive, namely *Staphylococcus aureus* and the animal pathogens, *S. intermedius*, *S. schleiferi subs. coagulans*, *S. delphini* and some strains of *S. hyicus*. Coagulase negative staphylococci (CoNS) are considered to be less virulent. However some CoNS species, in particular *S. epidermidis*, *S. lugdunensis* and *S. haemolyticus* can cause serious infections in humans.

*S. aureus* is by far the most studied species of staphylococcus due to its prevalence as a human and animal pathogen. It forms smooth, raised colonies and is distinguished from other staphylococci by the presence of a golden pigment, the ability to ferment mannitol, and the expression of thermostable DNase and clumping factor (the cell wall-associated fibrinogen binding protein ClfA).
1.1.2 Colonisation and disease

1.1.2.1 Colonisation and disease in humans

Staphylococci can inhabit the skin, skin glands and mucous membranes of humans and warm-blooded animals. Humans are the natural reservoir of *S. aureus*. The primary ecological habitat of *S. aureus* in humans is the moist squamous epithelium of the anterior nares. Approximately 20% of the human population are permanently colonised while the remainder are colonised intermittently (van Belkum *et al.*, 2009). It is unclear why some individuals are persistent carriers of *S. aureus* and others are not. Colonisation of the nares depends on the ability of *S. aureus* to adhere to desquamated epithelial cells and to avoid the host immune response. Host factors may play a role in nasal colonisation. Polymorphisms in genes encoding components of the host immune system (e.g. interleukin 4, complement factor H, and C-reactive protein which have associated with nasal carriage of *S. aureus*) have been associated with increased or decreased carriage (van Belkum *et al.*, 2007; Emonts *et al.*, 2008).

In humans, staphylococcal infections are initiated when a breach of the skin or mucosal barrier allows bacteria access to adjoining tissues or the bloodstream. Colonisation has been identified as a risk factor for the development of *S. aureus* infections and individuals are usually infected by the isolate they carry (von Eiff *et al.*, 2001; Wertheim *et al.*, 2004). Whether an infection is contained or spreads depends on a complex interplay between *S. aureus* virulence determinants and host defence mechanisms. The most common types of *S. aureus* infection are superficial skin lesions such as abscesses, boils and impetigo. If the organism gains access to the bloodstream (bacteraemia), it can cause a wide variety of invasive infections resulting in high morbidity and mortality. Such infections include those of bone (osteomyelitis), joints (septic arthritis), lungs (pneumonia) and heart valves (infective endocarditis) (Lowy, 1998). While many instances of nosocomial *S. aureus* bacteraemia are attributable to an endogenous source (von Eiff *et al.*, 2001), it has been noted that bacteraemia-related death was significantly higher in infected non-carriers compared to infected carriers. This suggests that carriers could be immunologically adapted to the strain of *S. aureus* that they carry (Wertheim *et al.*, 2004). Treatment of invasive *S. aureus* infections relies heavily on the use of antimicrobial agents to which the organism is increasingly
developing resistance in forms such as small colony variants (Bhattacharyya et al., 2012) and bacterial persister cells (Lechner et al., 2012).

1.1.2.2 Colonisation and disease in animals

Although the majority of studies have been on the colonisation of humans, *S. aureus* is also a major veterinary pathogen which can colonise and infect a wide range of domesticated and wild animals, including sheep, rabbits, chickens and turkeys (Rodgers et al., 1999; Linares & Wigle, 2001; Hermans et al., 2003; Vimercati et al., 2006). MRSA ST398 has rapidly spread worldwide and is mainly associated with animal colonisation in livestock farming. Although MRSA ST398 has high prevalence rates in pigs and calves (and until recently did not cause frequent invasive disease in humans), a study by Price et al (2012) suggested that livestock-associated MRSA CC398 originated in humans as MSSA (Schijffelen et al., 2010; Price et al., 2012). Methicillin-resistant *S. aureus* strains have been isolated from cats, dogs, pigs and horses (Leonard & Markey, 2008).

*Staphylococcus aureus* is one of the most important etiologic agents in mastitis of cows, goats, and sheep. Moreover, it is probably the most infectious agent because it causes a chronic and deep infection in the mammary glands that is extremely difficult to cure. Mastitis may be clinical (severe) or subclinical (moderate). It constitutes a serious problem in dairy herds as it has considerable economic consequences, due mainly to a reduction in milk production and discarded milk (Aires-de-Sousa et al., 2007). The molecular mechanisms by which *S. aureus* causes mastitis remain poorly explained. Importantly, studies of the molecular epidemiology of *S. aureus* strongly suggest that a genetic subset of strains is particularly well adapted for causing infection in cattle (Herron-Olson et al., 2007). In addition to their economic importance, animal strains of *S. aureus* also pose a risk of human zoonoses and are therefore considered a potential threat to public health (Ben Zakour et al., 2008).

*S. aureus* is a common cause of broiler poultry infections which have major economic consequences in the poultry industry. In a study to investigate the evolutionary origin of *S. aureus* strains from poultry, a combination of population genetics and comparative genome sequencing revealed that the majority of isolates belonged to a single clonal complex (CC5) which was widespread among both healthy and diseased birds. Although the CC5 poultry isolates were widespread in countries in several different
continents, all isolates could be traced back to a single likely human to poultry host-jump which originated in or near to Poland (Lowder et al., 2009). This jump was followed by extensive genetic diversification including acquisition of novel mobile genetic elements and loss of virulence gene function. Identification of strains of livestock-associated MRSA, such as the CC398 clone found in pigs, cows, poultry and horses, which have also been associated with disease of humans, is worrisome (Cuny et al., 2010).

It was proposed that the globalized nature of the poultry industry facilitated the spread of a small number of broiler breeding lines (and their resident normal flora) from the poultry ST5 clade around the world. Comparison of the genomes of a representative poultry CC5 strain ED98 and the closely-related basal human strain MR1 allowed examination of the genetic basis for the adaptation of \textit{S. aureus} to an avian niche (Lowder et al., 2009). Since divergence from a common ancestor with the human strain MR1, mutations have occurred which have led to loss of gene function in the poultry strain ED98 including a nonsense mutation in the \textit{spa} gene which encodes Protein A (Protein A will be described in section 1.3.1.2). The ED98 genome contained several novel mobile genetic elements which were not found in human or ruminant strains. They were, however, widely distributed among poultry \textit{S. aureus} strains of distinct clonal origin, suggesting a fundamental role in avian host adaptation (Lowder et al., 2009). In addition to the common CC5 poultry subtype, less common genotypes of poultry \textit{S. aureus} (which are identical or closely related to extant human genotypes) contain mobile genetic elements (MGEs) unique for poultry strains. This suggests that other host switches may have occurred on numerous occasions and that the ST5 human to poultry host jump followed by host adaptation may not have been a unique event (Lowder & Fitzgerald, 2010).

1.1.2.3 Bovine mastitis

Bovine mastitis is a disease of worldwide importance and is associated with considerable economic consequences eg reported costs for the US dairy industry estimate at \( \sim 1.7 \) to 2 billion U.S. dollars each year (11% of total U.S. milk production) (Jones, 2010). It can be caused by a wide range of organisms, including gram-negative and gram-positive bacteria, mycoplasmas and algae. Whereas many common bovine mastitis-causing microbial species such as \textit{S. aureus}, \textit{E. coli}, \textit{Klebsiella pneumoniae}, and \textit{Streptococcus agalactiae} can also colonise humans as commensals or pathogens, others
such as *Strep. uberis*, *Strep. dysgalactiae subsp. dysgalactiae* or *S. chromogenes*, are almost exclusively found in animals (Zadoks *et al*., 2011).

*S. aureus* is primarily spread from cow-to-cow during the milking process either through contact with fomites or the milker’s hands (Rappouli & Bagnoli, 2011). The mammary gland is the main reservoir of infection. Although clinical disease occurs, the most common form of *S. aureus* mastitis in modern dairies is chronic subclinical infection which leads to elevations in milk somatic cell count (SCC) and decreased milk production. The quality of milk produced can also be reduced as a result of mammary epithelial cell and alveolar function impairment (Brouillette *et al*., 2003; Brouillette & Malouin, 2005).

Previous measures to control and prevent the disease included milking time hygiene, careful maintenance of milking equipment, culling of chronically infected cattle and routine use of intramammary antimicrobials during the non-lactating period (Dodd *et al*., 1969; Neave *et al*., 1969). Antimicrobial treatment is not an ideal method of controlling bovine mastitis as various factors affect the outcome and the success of treatment declines with increasing age of the cow, increasing milk SCC, chronicity of infection, increasing bacterial count prior to treatment and increasing number of quarters infected. Also, the extensive use of antibiotics presents a food safety concern as this may result in the emergence of drug-resistant bacteria in milk (Barkema *et al*., 2006).

Infection of the mammary gland initiates when *S. aureus* enters the teat orifice, traverses the keratinised teat canal, and moves into the mammary gland parenchyma (Sutra & Poutrel, 1994). Once in the gland, *S. aureus* can adhere to host cells or extracellular matrix proteins by surface proteins such as fibronectin binding protein, clumping factor and collagen binding protein, and teichoic acid (Shinefield & Black, 2005). Clearance of *S. aureus* from milk is impaired as neutrophils, which constitute the primary phagocyte responsible for clearing *S. aureus* from the mammary gland, have impaired phagocytic function due to the presence of fat globules (Kent & Newbould, 1969; Paape *et al*., 1975; Paape *et al*., 1981).

### 1.2 Genome structure of *Staphylococcus aureus*

*S. aureus* has a circular genome which contains approximately 2.8 Mb of AT-rich DNA (average G + C content of 33 %) and carries approximately 2600 genes in its
chromosome (Kuroda et al., 2001). Since the first whole genome sequence of a *S. aureus* strain, which was completed by random shot-gun sequencing, was made available in 2001, the entire genomes of several *S. aureus* isolates have been published. These include strains N315 and Mu50 which are closely related, hospital-acquired methicillin-resistant *S. aureus* (MRSA) isolates from Japan, strain MW2 (a community-acquired MRSA isolate from the USA) and the closely related, methicillin-sensitive *S. aureus* (MSSA) isolate MSSA476, an early MRSA isolate (strain COL) and a related MSSA isolate (strain NCTC 8325) and finally MRSA252, which is representative of the highly successful, epidemic EMRSA-16 clone (Kuroda et al., 2001; Baba et al., 2002; Holden et al., 2004; Gill et al., 2005). RF122 is the first bovine isolated-*S. aureus* strain to have its entire genome sequenced.

1.2.1 Core genome

DNA microarray analysis of diverse *S. aureus* isolates from different geographical locations indicates that approximately 75% of the *S. aureus* genome comprises a core component of genes present in all strains (Fitzgerald & Musser, 2001). It includes genes involved in metabolism and other house-keeping functions but also contains some genes not essential for growth, such as surface-expressed proteins and exoenzymes (Lindsay & Holden, 2004). Subtle differences in the conserved core genome have been exploited as a means of understanding the structure of the *S. aureus* population. Multilocus sequence typing (MLST) is the main method for analysing the clonality of *S. aureus* isolates (Feil & Enright, 2004) and will be described in section 1.2.4.

1.2.2 Accessory genomic elements

The accessory genome refers to dispensable genetic material that comprises up to 25% of any *S. aureus* genome and consists of mostly mobile (or once mobile) genetic elements that can be transferred horizontally between different clonal complexes (Fitzgerald & Musser, 2001; Lindsay & Holden, 2004). These elements include bacteriophages, pathogenicity islands, chromosomal cassettes, genomic islands, plasmids and transposons, many of which carry genes associated with virulence or resistance (Baba et al., 2002; Holden et al., 2004). Many of the genes that encode secreted virulence
factors, including superantigens, enterotoxins and leukocidins, are found on genomic islands, pathogenicity islands and prophages (Narita et al., 2001; Gill et al., 2005). Plasmids, transposons and chromosomal cassettes frequently contain genes encoding resistance to antimicrobials and heavy metals (Ito et al., 2003). Perhaps the most striking case of rapid evolution of virulent clones is the acquisition of the genes encoding the Panton-Valentine Leukocidin.

In conclusion, the *S. aureus* genome consists of a stable core and accessory elements that are strain specific. The strains that were used for genome sequencing only reflect a fraction of the diversity within the species as a whole. Further sequencing of the genomes of strains from major lineages would likely identify the full range of accessory genes and elements within the species *S. aureus*, and may provide insights on the success of certain clonal lineages.

1.2.3 Genotyping

Although MLST (described in section 1.2.4), is the main method for analyzing the clonality of *S. aureus* isolates (Feil & Enright, 2004) other methods for analyzing the population dynamics of *S. aureus* have been utilised. The “gold standard” for epidemiological analysis has been pulsed-field gel electrophoresis (PFGE) where genomic DNA is cleaved with endonucleases that recognize a small number of sites on the chromosome, yielding a distinct pattern of bands upon agarose gel electrophoresis (Trindade et al., 2003). Amplified fragment length polymorphism (AFLP) is a PCR-based method that scans for polymorphisms in selected restriction sites and the nucleotides bordering these sites (Melles et al., 2004). Other typing methods based upon the sequence of selected surface protein genes have been used in some studies eg *spa* typing (Shopsin et al., 1999; Robinson & Enright, 2003) and *dru* typing. Ideally a combination of methods should be used for efficient analysis.

1.2.4 MLST typing

Typing of *S. aureus* strains by MLST involves the precise sequencing of DNA fragments (approximately 450 bp) of seven slowly evolving metabolic genes: *arcC*, *aroE*,
The sequences obtained are compared to known alleles at each locus (http://www.mlst.net). The resulting seven-integer allelic profile defines a sequence type (ST) for each strain. A single nucleotide change in a sequence is sufficient to define a new allele. Therefore, *S. aureus* strains that share the same ST share identical alleles at all seven MLST loci. Strains that differ at only one allele out of seven are known as single-locus variants (SLV), and this indicates a close genetic relatedness.

Based on their MLST sequence types, isolates can be grouped into clonal complexes (CCs) to give an overall representation of the population structure of *S. aureus*. CCs are defined as groups of STs in which every ST shares at least five out of seven alleles with at least one other ST in the group (Feil et al., 2003; Robinson & Enright, 2004). Burst detection algorithms (Burst analysis) can be used to identify the most likely ancestral strain (ST) within a CC. For example, Enright et al. (2002) used BURST analysis to identify a large complex of related MSSA and MRSA clones, CC8. The ancestor of this group was predicted to be ST250. Significantly, almost all of the MRSA isolates from the 1960s were within CC8 and most of these were ST250 suggesting that methicillin resistance first arose within *S. aureus* ST250 (Enright et al., 2002).

### 1.2.5 Population structure of *S. aureus*

MLST analysis of a large strain collection (n = 334), obtained from the Oxfordshire area in the U.K. revealed that the population structure of *S. aureus* is highly clonal (Feil et al., 2003). Grouping of isolates into clonal complexes based on their MLST genotype showed that 77 % of all tested isolates fall into eight major clonal complexes (Feil et al., 2003). Phylogenetic analysis revealed that the different clonal complexes within this population show significant divergence from each other. Analysis of clonal diversification within individual clonal complexes suggests that alleles are at least 15-fold more likely to change by point mutation rather than by recombination (Feil et al., 2003). This is in contrast to naturally transformable bacteria such as *Neisseria meningitidis* and *Streptococcus pneumonia* in which alleles change between 5- to 10-fold more frequently by recombination than by mutation (Feil & Spratt, 2001). Both *S. aureus* carriage and invasive isolates were evenly distributed among the clonal complexes suggesting there is no link between MLST genotype and the propensity to cause disease (Feil et al., 2003).
The population structure of a large number of *S. aureus* strains (n = 993), isolated either from healthy carriers or from patients with invasive disease in the Netherlands was recently studied using AFLP analysis (Melles *et al.*, 2004). A clonal population structure was identified comprising five major AFLP clusters, which match the major clonal complexes identified by MLST (Robinson & Enright, 2004, Feil *et al.*, 2003). This suggests that the same clonal clusters have spread successfully in both the U.K. and the Netherlands, and probably worldwide. It was found that carriage strains fell into the same main clusters as isolates from invasive disease, and that MRSA strains from international sources grouped in these same clusters (Melles *et al.*, 2004). Some AFLP sub-clusters contained proportionately more invasive isolates than carriage isolates (Melles *et al.*, 2004). This confirms the suggestion that essentially any *S. aureus* strain has the capacity to cause invasive disease (Feil *et al.*, 2003) but strains from some clonal lineages are more virulent than others (Melles *et al.*, 2004). Invasive disease encompasses a very wide range of disease symptoms which is associated with the wide variety of virulence factors expressed by *S. aureus* (see section 1.3). It is suggested that closely related isolates of the same ST may differ in their content of virulence genes and therefore differ in their capacity to cause disease. The presence of seven virulence factors in the strain collection from the U.K., including surface proteins and exotoxins, was associated with invasive disease (Peacock *et al.*, 2000). Some virulence factors or antibiotic-resistance determinants may be carried on mobile accessory genetic elements. This is discussed in section 1.2.2.

The percentage identity of the core genomes of the hospital-acquired isolates N315 and Mu50 suggest that these isolates are closely related (Kuroda *et al.*, 2001). These strains share identical MLST genotypes (ST-5; Figure 1.1). The closely related community acquired strains MW2 and MSSA476 both belong to ST-1. This sequence type represented the ancestral strain of clonal complex 1 in the Oxfordshire study (Feil *et al.*, 2003). Strains NCTC 8325 and COL (both isolated from the same geographical location in the U.K.) belong to different, but closely related, sequence types (ST-8 and ST-250, respectively) (Lindsay & Holden, 2004) (Figure 1.1). Strain MRSA252 is the most divergent of the sequenced *S. aureus* strains (Holden *et al.*, 2004). This is reflected in its MLST profile (ST-36) which clearly separates this strain from the other sequenced isolates based on phylogenetic analysis (Figure 1.1). This ST is well represented in the major CC30/39 clonal complex found in the U.K. (Feil *et al.*, 2003). The levels of
Figure 1.1 Phylogenetic relatedness of *S. aureus* MLST genotypes

A neighbour-joining tree of a representative sample of 30 sequence types (STs) from the *S. aureus* MLST database (http://www.mlst.net). The positions of the seven sequenced *S. aureus* strains are indicated. The sequences of each allele were concatenated to produce 3198-bp of sequence for each ST for construction of the neighbour joining tree.

Adapted from Lindsay and Holden, 2004.
relatedness inferred by MLST correlates well with the overall genomic divergence in the core genomes of the sequenced strains (Lindsay & Holden, 2004). This highlights the suitability of MLST as a means of understanding the evolution and population biology of *S. aureus*.

1.3 *S. aureus* cell wall-associated virulence factors

*S. aureus* produces an extensive range of virulence factors which allow it to colonise a number of different sites in the body and cause a wide range of diseases. Adhesins are ligand-binding surface proteins that are attached covalently to the cell wall peptidoglycan and are coordinately expressed to promote binding to the host extracellular matrix (ECM) and blood components such as fibrinogen. These proteins are called MSCRAMMs (microbial surface components recognizing adhesive matrix molecules) and their expression promotes adhesion to damaged tissue, to the surfaces of host cells and to implanted medical devices. Many *S. aureus* MSCRAMMS are multifunctional (e.g. Spa, IsdA which will be discussed below). Functional redundancy often occurs with different MSCRAMMs performing similar functions (e.g. Fg binding). Not all wall-associated proteins are MSCRAMMs.

Many MSCRAMMs have been shown to be critical in the initiation of infection. Evasins are proteins that help the bacterium to evade the host innate and adaptive immune responses. They may be associated with the bacterial cell wall or are secreted into the extracellular environment. Another group of virulence factors are the secreted enzymes, toxins, and proteases which are involved in the later stages of infection and facilitate tissue destruction and spreading. Small colony variants (SCVs) are metabolically inactive forms of *S. aureus* which contribute to intracellular survival by evasion of phagocytosis. This promotes the development of persistent staphylococcal infections and resistance to antibiotics. Several animal models have been developed to study the function of *S. aureus* virulence factors *in vivo*. However, pathogenicity is a complex phenotype and in most *S. aureus* infections no single factor is responsible.
1.3.1 Cell wall-associated surface proteins

Not only is the cell wall the point of contact between *S. aureus* and its surrounding environment but it acts as a physical barrier and provides a rigid exoskeleton which protects *S. aureus* from the external environment. The main component of the cell wall is peptidoglycan (approximately 60%) with the remainder comprising wall teichoic acids, lipoteichoic acids and small amounts of protein. Peptidoglycan consists of glycan strands made of repeating disaccharide units, N-acetylglucosamine and N-acetylmuramic acid (GlcNAc-\((\beta1-4)\)-MurNAc) (Ghuysen & Strominger, 1963). A rigid three-dimensional cell wall network (Ghuysen *et al*., 1965; Strominger & Tipper, 1965) results from the cross-linking of MurNAc moieties in the glycan chains by short tetrapeptides (L-Ala-D-Glu-L-Lys-D-Ala). A characteristic feature, which renders *S. aureus* peptidoglycan susceptible to cleavage by the endopeptidase lysostaphin, is the presence of pentaglycine interpeptide bridges that link the tetrapeptide units of neighbouring glycan chains (Schleifer & Kandler, 1972). The aforementioned MSCRAMMs, a diverse array of ligand-binding proteins which function in binding to components of the host's extracellular matrix, are produced on the *S. aureus* cell surface and are covalently anchored to the cell wall peptidoglycan.

MSCRAMMs have many features in common including an N-terminal signal sequence of ~ 40 amino acids that targets the proteins to the secretary (Sec) pathway. An 'AXA' motif at the end of the signal is recognised and cleaved by the membrane-anchored signal peptidase enzymes SpsA and SpsB during translocation across the cytoplasmic membrane (Cregg *et al*., 1996; Mazmanian *et al*., 2001). C-terminal sequences include a wall-spanning region and an LPXTG motif which facilitates targeted attachment to the cell wall where the cleaved peptide is covalently linked to the peptidoglycan layer by sortase enzymes in a process known as 'sorting' which will be discussed in section 1.3.1.1. The LPXTG motif is followed by a hydrophobic, membrane-spanning domain and a cytoplasmic tail composed of a number of positively charged residues (Fischetti *et al*., 1990; Navarre & Schneewind, 1999; Mazmanian *et al*., 2001). The structural organisation of staphylococcal surface proteins is illustrated in Figure 1.2.
**Figure 1.2 Structural organisation of staphylococcal surface proteins**

Schematic diagram of the structural organisation found in surface proteins from *S. aureus*. The signal sequence (S), ligand binding (N2, N3), repeat (R) and membrane spanning regions (M) are shown. The ligands bound by different surface proteins are indicated. The function of N1 is not known. The B repeats of SdrE and Bbp are shown.
1.3.1.1 Sortase-mediated anchoring of cell wall-associated proteins

As previously mentioned, *S. aureus* surface proteins are covalently linked to the peptidoglycan layer of the cell wall by a process referred to as ‘sorting’ which is represented diagrammatically in Figure 1.3. Sortases have also been shown to be involved in the biogenesis of Gram-positive pili (Kang & Baker, 2012). *S. aureus* has two sortase (Srt) enzymes, SrtA, which anchors the majority of *S. aureus* surface proteins and recognises the LPXTG motif, and SrtB, which anchors the iron-regulated surface determinant C (IsdC) and recognises the atypical C-terminal NPQTN motif (Mazmanian *et al.*, 2001; Pallen *et al.*, 2001; Mazmanian *et al.*, 2002). SrtA is a membrane-bound protein with a catalytic domain that is embedded in the peptidoglycan layer (Mazmanian *et al.*, 2001). SrtA plays a crucial role in the correct anchoring of surface proteins, many of which are virulence factors. *S. aureus* srtA mutants, which fail to correctly anchor LPXTG-motif surface proteins, are attenuated in animal infection models of septic arthritis and endocarditis (Tonsson *et al.*, 2002; Weiss *et al.*, 2004). Recognition of LPXTG is highly stringent and amino acid substitutions at position 1, 2, 4 or 5 prevents correct sorting of surface proteins (Kruger *et al.*, 2004). X-ray crystal structures of SrtA in complex with an LPETG peptide and NMR analysis of StrA in the presence or absence of ligand provided insight into the detailed action of sortase (Liew *et al.*, 2004; Zong *et al.*, 2004). SrtA cleaves the LPXTG motif between the threonine and glycine residues and subsequently attaches the cleaved protein covalently via threonine to glycine in the pentaglycine of the nascent peptidoglycan subunits (Ton-That *et al.*, 2000). The C-terminal membrane-spanning domain and positively charged tail is then released and degraded.

1.3.1.2 Protein A

The first cell wall-associated surface protein of *S. aureus* to be characterised was the *staphylococcal protein A* (Spa). Spa is best known for its ability to bind to the Fc region of mammalian IgG in a non-immune manner (Uhlen *et al.*, 1984; Moks *et al.*, 1986). It was later discovered to bind to the variable region of the Fab heavy chain of V_{H}3 subclass immunoglobulins (Hillson *et al.*, 1993). Other ligands for Spa include von Willebrand factor (vWF) domains A1 and D1-D3 and tumor necrosis factor receptor-1 (TNFR-1) (Hartleib *et al.*, 2000; Gomez *et al.*, 2004; O’Seaghdha *et al.*, 2006). The spa
gene is transcribed during the mid-exponential phase of growth but expression is repressed as cultures enter stationary phase (Vandenesch et al., 1991). Spa is expressed on the surface of over 95% of *S. aureus* strains (Forsgren & Nordstrom, 1974) and is a prime example of the multi-functionality shown by many *S. aureus* surface proteins. The surface-exposed region of Spa is composed of five homologous repeat domains (EDABC) of ~58 residues linked to the cell surface by a short polymorphic repeat region (Xr) containing a variable number of eight residue repeats and a conserved region (Xc) (Uhlen et al., 1984; Moks et al., 1986). Ligand-binding sites are found in each of the five repeat domains and residues important for ligand-binding are highly conserved between domains. Each domain comprises anti-parallel α-helices that pack together to form a compact three helical bundle (Gouda et al., 1992)(Gouda et al., 1992). Both the Spa-Fc and Spa-Fab complexes have been co-crystallized. The crystal structure of domain B of Spa in complex with an Fc fragment revealed that the Fc-γ binding site spans helices I and II on a single face of the molecule and comprises 11 residues (Deisenhofer, 1981).

The binding sites of von Willebrand factor and TNFR-1 have been localised by site-directed mutagenesis and both recognise a region in helices I and II that overlaps the IgG Fc binding site in Spa (Gomez et al., 2006; O'Seaghdha et al., 2006). The crystal structure of the D domain of Spa in complex with a V_{H3}-Fab shows that the V_{H3} binding site spans helices II and III and the residues from helix II that bind Fc differ from those that bind V_{H3} with the exception of glutamine 32 in Spa domain D (Graille et al., 2000). Spa residues that form contacts with Fab were confined to helices II and III and are distinct from the Fc-γ binding site. It is possible for a single Spa domain to bind Fc and V_{H3}-Fab simultaneously (Starovasnik et al., 1999; Graille et al., 2000). Spa has multiple functions in thwarting the host's immune response and is a virulence factor in murine infection models of subcutaneous infection, sepsis and septic arthritis and staphylococcal pneumonia (Patel et al., 1987; Gomez et al., 2004; Palmqvist et al., 2005b).

Spa also possesses immunosuppressive properties. The ability of Spa to bind to the Fab region on V_{H3} adjacent to the antigen-binding domain of IgM molecules exposed on the surface of B lymphocytes (Sasso et al., 1989; Sasano et al., 1993) causes their activation, proliferation and subsequent apoptotic destruction (Goodyear & Silverman, 2003). A reduction in antibody-secreting cells from the spleen and bone marrow may account for the immunosuppressive activity of Spa (Goodyear & Silverman, 2004).
Figure 1.3 Surface protein anchoring in *S. aureus*

(i) Export. Precursor proteins with an N-terminal signal peptide (SP) are directed into the secretory (Sec) pathway and the signal peptide is removed. (ii) Retention. The C-terminal sorting signal retains polypeptides within the secretory pathway. (iii) Cleavage. Sortase cleaves between the threonine and the glycine of the LPXTG motif, resulting in the formation of a thioester enzyme intermediate. (iv) Linkage. The acyl-enzyme intermediate is resolved following nucleophilic attack of the free amino group of lipid II by the thioester bond, synthesizing the amide bond between surface proteins and the pentaglycine cross-bridge and regenerating the active-site sulphydryl. (v) Cell wall incorporation. Lipid-linked surface protein is first incorporated into the cell wall by transglycosylation. The murein pentapeptide subunit with attached surface protein is then cross-linked to other cell wall peptides, generating the mature murein tetrapeptide.
Spa can also act as an immune effector by binding to and activating TNRF-1, the receptor for tumour necrosis factor α (TNFα). Spa mimics TNFα by activating the TNFR-1 proinflammatory signalling cascade and stimulating shedding of TNFR-1 by TNF-converting enzyme (TACE) (Gomez et al., 2004; Gomez et al., 2006). This is important in the pathogenesis of pneumonia as TNFR-1 is widely distributed on the airway epithelium. Consequently mice infected with wild-type *S. aureus* have a significantly higher incidence of pneumonia than mice infected with a Spa-deficient strain. Furthermore the Spa-deficient strain is attenuated similarly in TNFR-1 null and wild-type mice (Gomez et al., 2004). The Xr domain of Spa has been shown to induce type I INF signalling in airway epithelial cells, in particular INF-β transcription (Martin et al., 2009). INF-α/β receptor 1 null mice cannot respond to type I INF signalling and they exhibit increased survival in a lethal *S. aureus* pneumonia model compared to wild-type mice indicating that Spa-induced type I INF signalling is important in the pathogenesis of *S. aureus* pneumonia (Martin et al., 2009).

More recently the role of Spa in invasion of *S. aureus* across airway epithelial cells has been elucidated. Spa activates the RhoA/ROCK/MLC cascade that is involved in cytoskeleton rearrangement. ROCK is the effector molecule of RhoA and stimulates contraction of actin filaments by phosphorylation of MLC which increases myosin II ATPase activity. Inhibition of RhoA-MLC signalling with the ROCK inhibitor Y-27632 prevented transmigration of *S. aureus* across airway epithelial cells (Soong et al., 2011).

Spa was also shown to activate epidermal growth factor receptor (EGFR) signalling which in turn activates RhoA and stimulates protease expression. Inhibition of EGFR phosphorylation significantly reduced the transmigration of *S. aureus* across airway epithelial cells (Soong et al., 2011). A protein matrix spans the paracellular junction between adjacent epithelial cells which *S. aureus* must traverse during invasion. EGFR signalling activates TACE and the calcium-dependent cysteine proteases the m-calpains. Two of the paracellular junction proteins E-cadherin and occludin are substrates for the calpains. Cleavage products of both of these proteins were detected in epithelial cells incubated with wild-type *S. aureus* but not in cells exposed to the Spa-deficient strain. In addition, mice pretreated with a calpain inhibitor had significantly increased numbers of *S. aureus* in the bronchoalveolar lavage than PBS treated control mice (Soong et al., 2011). Taken together these results indicate that Spa mediates invasion across
airway epithelial cells by activating RhoA and subsequent proteolytic cleavage of paracellular junction proteins.

*S. aureus* is responsible for greater than 80% of osteomyelitis cases. Recent studies have indicated a role for Spa in adherence to osteoblasts through interaction with TNFR-1 expressed on the surface of osteoblasts. It has been shown that *S. aureus* adherence to osteoblasts via Spa leads to inhibition of proliferation and subsequent apoptosis of these cells whilst also increasing the production of the bone remodelling molecule RANKL (Claro *et al.*, 2011). Thus Spa is an important factor in causing destruction of bone in the pathogenesis of osteomyelitis.

### 1.3.1.3 Iron regulated surface determinants

In humans the majority of iron is sequestered by haem-containing proteins and is unavailable to invading pathogens. Iron is essential for bacterial survival and growth inside the human host. *S. aureus* expresses and secretes haemolysin toxins which lyse red blood cells and release haemoglobin. Free haemoglobin is toxic and, following its release, is rapidly cleared from plasma by associating with the haemoglobin carrier molecule haptoglobin. Free haem, which is also toxic, is found at very low concentrations and is cleared from circulation by haemopexin.

The expression of a subset of staphylococcal genes is induced under iron-limiting conditions, which are likely to resemble growth conditions in serum and within the host during infection. The iron regulated surface determinant (Isd) proteins expressed from the *isd* locus are thought to function together to capture and transport haem-iron across the cell envelope and into the cytoplasm of *S. aureus* where it can be degraded to release iron (Skaar & Schneewind, 2004). Recent studies have reported a pathway for transfer of haem between Isd proteins which will be discussed in section 1.3.1.3.1. The *isd* genes are located on five transcriptional units (*isdA, isdB, isdCDEFsrtBisdG, isdH* and *isdI*) (Figure 1.4) and expression of Isd proteins is controlled by the iron-dependent ferric-uptake-response (Fur) transcriptional repressor which inhibits the transcription of *isd* genes in the presence of Fe^{3+} (Dryla *et al.*, 2003; Mazmanian *et al.*, 2003).

*IsdA, IsdB* and *IsdH* (also known as HarA) are surface-located proteins each containing the C-terminal LPXTG sortase A sorting signal. *IsdC* is partially buried in the
Figure 1.4 Genomic regions encoding the Isd proteins *S. aureus*

The Isd system is encoded on five transcriptional units located in three distinct regions of the *S. aureus* chromosome. Genes represented in blue encode cell surface-exposed proteins that are anchored by sortase A. The *isdC* gene which encodes a sortase B substrate is shown in red. Genes that encode the haem transport system are shown in yellow and cytoplasmic haem-degrading enzyme genes in green. The genes shown in white are uncharacterized. Adapted from Skaar *et al.*, 2004.
cell wall and has an atypical cell wall sorting signal, NPQTN. It is cross-linked to an unknown component of the cell wall envelope by sortase B, a unique sortase encoded on the transcription unit isdCDEFsrtBisdG (Mazmanian et al., 2002). IsdA, IsdB, IsdC and IsdH contain different numbers of conserved NEAT (near iron transporter) domains which are involved in binding to haem and haem-containing proteins (Figure 1.5).

1.3.1.3.1 Haem transfer pathway

The main function of the Isd proteins collectively is to bind human haemoproteins, remove the haem molecule and transport haem through the cell wall and plasma membrane for accumulation in the bacterial cytoplasm (Skaar & Schneewind, 2004). IsdB binds to haemoglobin and haem, IsdH binds to haemoglobin and to haptoglobin-haemoglobin complexes while IsdC can bind to haemin (oxidised form of haem) (Torres et al., 2006; Dryla et al., 2007). Muryoi et al. demonstrated that haem can be passed in a unidirectional manner from IsdB NEAT domain 2 or IsdH NEAT domain 3 to the NEAT domain of IsdA (Muryoi et al., 2008) which then passes haem to IsdC, which acts as a central conduit and transfers haem into the cytoplasm to IsdE, a component of the ABC transporter composed of IsdD, IsdE and IsdF (Muryoi et al., 2008). Once inside the cell haemoxygenases IsdG and IsdI bind and degrade haem releasing biliverdin and free iron which is used as a nutrient source (Figure 1.6). IsdG and IsdI are differentially regulated by iron and haem and are required for growth of S. aureus when haem is the sole iron source available (Reniere & Skaar, 2008). Inactivation of components of the isd system does not eliminate the ability of S. aureus to utilise haem as an iron source, indicating that other pathways of haem acquisition exist (Skaar & Schneewind, 2004).

1.3.1.3.2 NEAT domains of Isd proteins

Genome searches revealed 125 residue domains located near to iron transporters which promote binding to haem and haem-containing proteins. These domains are possessed by IsdA, IsdB, IsdH and IsdC proteins and are referred to as NEAT (NEAr iron Transporter) domains (Andrade et al., 2002). IsdH has three NEAT domains, IsdB has two while IsdA and IsdC each have one NEAT domain. Based on primary sequences
NEAT domains can be divided into four groups and although the degree of sequence identity between NEAT domains can vary widely they were all predicted to form IgG-type folds.

The NEAT 1 and 2 domains of IsdH and NEAT 1 of IsdB are Type I NEAT domains that bind haemoglobin and haemoglobin-haptoglobin complexes. They do not bind haem. Type II NEAT domains include NEAT 3 of IsdH and NEAT 2 of IsdB. Type III and IV NEAT domains are found in IsdA and IsdC respectively (Figure 1.5). Type III and type IV NEAT domains have been resolved in complex with heme.

The crystal structure of the IsdA NEAT domain in complex with haem revealed that haem was bound in a hydrophobic pocket within a β-sandwich structure. The conserved tyrosine residues were crucial for haem-iron coordination (Grigg et al., 2007). The crystal structure of IsdC in complex with haem revealed that haem is bound near the hydrophobic core of the β-sandwich structure and rests against a prominent β-hairpin structure. Tyrosine residues located on the β-hairpin coordinate the iron atom of haem. Opposite to the β-hairpin a short helical peptide acts as a lip and is predicted to undergo conformational changes resulting in haem being locked in a closed conformation. The tyrosine residues which coordinate the haem-iron atom in IsdC are conserved in IsdA and the Type II NEAT domains of IsdB and IsdH. A non-conserved tryptophan residue in IsdC also contributes to haem binding (Sharp et al., 2007; Villareal et al., 2008).

The three-dimensional structure of the NEAT 1 domain of IsdH is structurally related to the immunoglobulin fold family and is organised into a β sandwich fold composed of two five-stranded antiparallel beta sheets. Aromatic residues in the loop connecting strands β1b to β2 are predicted to bind haemoglobin. Other surface loops that connect strands β7 to β8 and β3 to β4 are also predicted to contact haemoglobin. The negative charge of the putative binding surface is implicated in haemoglobin binding (Pilpa et al., 2006). The solution of the crystal structure of the NEAT domain 3 of IsdH in complex with haem revealed a similar method of interaction to that of the NEAT domain of IsdA (Watanabe et al., 2008). However, Wantanabe et al. predicted that IsdH NEAT 3 can bind multiple haem molecules and that haem binding involves protein multimerisation. This process is predicted to increase efficiency of haem transport.

The crystal structure of the IsdB N2 heme complex, which exists as an eight-strand β-sandwich fold, revealed that a conserved Tyr residue coordinates heme-iron. A Met residue is also involved in iron coordination, resulting in a novel Tyr-Met
Figure 1.5 The cell-wall associated Isd proteins of *S. aureus*

The surface exposed Isd proteins have variable numbers of NEAT domains. Type I NEAT domains are shown in blue, type II in pink, type III in yellow and type IV in green. NEAT domains 2 and 3 of IsdH are highly homologous to NEAT domain 1 and 2 of IsdB. The C terminal LPXTG SrtA recognition motifs are indicated for IsdH, IsdB and IsdA. IsdC has an atypical NPQTN motif which is recognised by SrtB.
Figure 1.6 Putative pathway of haem transfer between Isd proteins of \textit{S. aureus}

Lysis of erythrocytes by \textit{S. aureus} toxins releases haemoglobin which rapidly associates with haptoglobin (Hpt). IsdB and IsdH NEAT1 domains can bind to haemoglobin and haemoglobin-haptoglobin complexes, respectively. Haem (filled black circles) is extracted and transferred to the IsdA NEAT domain. IsdA transfers haem to IsdC. Haem is then transferred to IsdE a component of the ABC transporter which transports haem into the cytoplasm. Haem is then degraded by haem-monoxygenases IsdG and IsdI and free iron released. The cell wall (CM) and cytoplasmic membrane (CM) are indicated.
hexacoordinate heme-iron state (Gaudin et al., 2011). As well as being unprecedented in the Isd system of *S. aureus*, Tyr-Met heme-iron coordination has not been reported in the literature for any heme binding protein to date.

1.3.1.3.3 Other functions of surface-exposed Isd proteins

In addition to iron-containing compounds, Isd proteins have been shown to interact with other ligands. IsdA has two domains with distinct functions. The N-terminal NEAT domain of IsdA binds a broad spectrum of human extracellular matrix and serum proteins, including fibrinogen, fibronectin, feutin, loricrin, involucrin and cytokeratin 10. The Bβ and γ-chain of fibrinogen appear to be the binding site for IsdA (Clarke et al., 2004). Loricrin, involucrin and cytokeratin 10 are all associated with the outer envelope of desquamated nasal epithelial cells, and the interaction of IsdA with these ligands may account for the promotion of bacterial adhesion to desquamated nasal epithelial cells and colonisation of the cotton rat nares (Clarke et al., 2004; Clarke et al., 2009). IsdA helps to protect *S. aureus* against the bactericidal protein lactoferrin by binding and neutralising its antibacterial serine protease activity. IsdA mutants are more sensitive to killing by lactoferrin (Clarke & Foster, 2008). The C-terminal domain of IsdA decreases the cellular hydrophobicity of *S. aureus* and confers resistance to hydrophobic fatty acids and host antimicrobial peptides and thus aids survival on live human skin. IsdA is the first protein shown to have this function (Clarke et al., 2007; Clarke & Foster, 2008). Vaccination of cotton rats with IsdA and IsdH protected against nasal colonisation (Clarke et al., 2006).

Surface plasmon resonance demonstrated that recombinant IsdB interacts directly with the platelet integrin receptor GPIIb/IIIa and promotes *S. aureus* adhesion to and aggregation of platelets (Miajlovic et al., 2010). IsdH plays a role in the evasion of phagocytosis. An *S. aureus* IsdH-deficient strain was engulfed more readily by human neutrophils, survived poorly in fresh whole human blood and was less virulent in a mouse model of sepsis. IsdH binds and activates complement factor I, which enhances the conversion of the serum opsonin C3b to iC3b and C3d and promotes the ability of *S. aureus* to survive neutrophil uptake and killing (Visai et al., 2009).
Clumping factor A (ClfA) is a fibrinogen-binding protein that is located on the bacterial cell surface and is one of the best characterised MSCRAMMs of *S. aureus*. The *clfA* gene is expressed predominately in the stationary phase of growth from a sigma factor B-dependent promoter (Nicholas *et al.*, 1999; Bischoff *et al.*, 2004). Weaker expression of ClfA occurs in the exponential growth phase and is dependent on transcription from a sigma factor A-dependent promoter (Homerova *et al.*, 2004). The structural features of ClfA are similar to other Gram-positive bacterial surface-anchored proteins and are represented diagrammatically in Figure 1.2. A 40 residue N-terminal signal sequence is followed by the 520 residue A domain which is composed of three subunits, N1, N2 and N3 and contains a cell surface-exposed fibrinogen binding site. The A domain is susceptible to cleavage by aureolysin at a SLAAVA motif located between subdomains N1 and N2 (McDevitt *et al.*, 1995; McDevitt *et al.*, 1997; McAleese *et al.*, 2001; Perkins *et al.*, 2001).

C-terminal to the A domain is a serine-aspartate (SD) dipeptide repeat region which varies in length between strains (McDevitt & Foster, 1995). The function of the SD repeats is to project the A domain away from the cell surface allowing interaction with fibrinogen (Hartford *et al.*, 1997). Surface proteins with SD repeats are found in other staphylococci. Four other SD repeat containing proteins (ClfB, SdrC, SdrD, SdrE) are present in *S. aureus* and are described in subsequent sections. The C-terminal wall-spanning region, membrane-spanning region and LPDTG motif are involved in anchoring ClfA to the cell wall (McDevitt *et al.*, 1994).

ClfA promotes the formation of cell aggregates in soluble fibrinogen (clumping) and the adherence of bacterial cells to fibrinogen-coated surfaces and fibrin clots (McDevitt *et al.*, 1994). The minimum fibrinogen-binding domain of ClfA comprises subdomains N2 and N3. Crystallisation of the N2N3 subdomain of ClfA in the apo-form allowed identification of a hydrophobic trench between N2 and N3 that forms the binding site for the extreme C-terminus of the γ-chain of fibrinogen (Deivanayagam *et al.*, 2002). Substitution of residues around this trench resulted in proteins with markedly reduced affinities for fibrinogen. Solution of subdomain N2N3 of ClfA in complex with the fibrinogen γ-chain peptide indicated that binding occurs by a variation of the dynamic “dock-lock-latch” mechanism which will be discussed in section 1.3.1.9 (Ganesh *et al.*, 2002).
Recent studies have identified a second fibrinogen binding site in the N3 subdomain of ClfA. Amino-acid substitutions within the second binding site reduced fibrinogen-binding but did not affect binding to the γ-chain peptide (Geoghegan, 2008). The binding site in the γ-chain of fibrinogen is the same as that recognised by the GPIIb/IIIa integrin on platelets which is required for platelet aggregation (Farrell et al., 1992; Hettasch et al., 1992). ClfA has been shown to stimulate platelet activation by forming a fibrinogen bridge between the bacterium and GPIIb/IIIa. ClfA-specific immunoglobulin is also required to interact with the platelet immunoglobulin Fc receptor (FcγRIIa) for platelet activation to occur (Loughman et al., 2005). ClfA can also activate platelets by a fibrinogen-independent mechanism that requires complement and specific anti-ClfA immunoglobulin (Loughman et al., 2005). Fibrinogen binding by ClfA can be inhibited by millimolar concentrations of Ca\(^{2+}\) (O'Connell et al., 1998). The function and structure of the N1 subdomain of ClfA is not yet known. It is predicted to be composed mainly of β-sheets and may be organised into more than one subdomain (Deivanayagam et al., 1999).

ClfA is anti-phagocytic and protects *S. aureus* from phagocytosis by murine macrophages and by human neutrophils *in vitro* (Palmqvist et al., 2004; Higgins et al., 2006). The anti-phagocytic properties of ClfA are partially dependent on fibrinogen-binding (Higgins et al., 2006). The coating of bacteria with fibrinogen may impair deposition of or access of opsonins to the bacterial cell surface. However, bacteria expressing a non-fibrinogen binding derivative of ClfA (ClfAPY) were still partially protected from opsonophagocytosis (Higgins et al., 2006). Recent studies have indicated that immune evasion is due to the ability of ClfA to bind and activate complement regulatory protein factor I from human serum resulting in enhanced degradation of opsonin C3b to iC3b (Hair et al., 2008; Hair et al., 2010). This leads to loss of opsonins whilst also inhibiting C3 convertase formation and the alternative pathway amplification loop, as surface-bound iC3b can no longer be recognised by factor B. However, the ClfAPY mutant binds factor I differently to wild-type ClfA. ClfAPY sequesters inactive factor I on its surface preventing the release of factor I from the bacteria in its active form (Hair et al., 2010). ClfAPY-expressing *S. aureus* are more susceptible to complement-mediated phagocytosis than the wild-type strain suggesting that the substitutions at positions 336 and 338 of ClfA are important in modulating *S. aureus* susceptibility to phagocytosis (Hair et al., 2010).
ClfA has been shown to be a virulence factor in several infection models. *S. aureus* ClfA-deficient mutants were less infective in the endocarditis model than parental strains (Moreillon *et al*., 1995). ClfA is also a virulence factor in murine models of septic arthritis and sepsis (Josefsson *et al*., 2001; Palmqvist *et al*., 2005a). In experimental endocarditis, function-blocking antibodies to ClfA sterilize vegetations on heart valves when administered with vancomycin (Vernachio *et al*., 2003). Active immunization with recombinant ClfA or passive immunization with polyclonal anti-ClfA IgG protected mice from arthritis and sepsis-induced death (Josefsson *et al*., 2001). ClfA is one of the major factors involved in *S. aureus* adherence to ventricular assist devices (VADs), posing a threat to the survival of patients being treated for congestive heart failure (Arrecubieta *et al*., 2006).

ClfA is expressed by nearly all strains of *S. aureus* and has been shown to be expressed *in vivo* late in infection (Peacock *et al*., 2002; Dryla *et al*., 2005). It is the target of several novel strategies to combat *S. aureus* infections. Inhibitex Inc. (Georgia, USA) have developed Veronate, a hyperimmunoglobulin against ClfA and the fibrinogen-binding *S. epidermidis* SdrG protein (Vernachio *et al*., 2003). A humanised monoclonal antibody against the A domain of ClfA (Aurexis) has also been developed by Inhibitex (Patti, 2004). In combination with vancomycin, it reduced bacterial densities in infected vegetations, kidneys and spleens in an infective endocarditis model (Patti, 2004).

### 1.3.1.5 Clumping factor B

Unlike ClfA which is expressed predominantly in the stationary growth phase, ClfB is expressed exclusively during the early exponential phase of growth and its contribution to fibrinogen binding is masked by ClfA (Ni Eidhin *et al*., 1998; McAleese *et al*., 2001). Transcription of *clfB* ceases in the late exponential growth phase after which ClfB protein is degraded by the metallopeptase aureolysin. Degradation results in the loss of ClfB ligand binding ability. Cleavage occurs at the SLAVA motif at the end of the N1 domain of ClfB. Loss of ClfB in the post-exponential growth phase is also mediated by the dilution of existing protein as the cells grow and divide while transcription of *clfB* has ceased and by shedding of ClfB into the growth medium by autolysis (McAleese *et al*., 2001). Transcription of *clfB* is not directly affected by Agr or the staphylococcal accessory regulator, SarA. However, SarA represses expression of aureolysin which
cleaves ClfB. The regulator of toxins, Rot is a global regulator which activates certain cell wall-associated proteins expressed in the exponential growth phase including ClfB (McNamara et al., 2000; McAleese & Foster, 2003; Said-Salim et al., 2003).

In common with ClfA, the A domain of ClfB possesses fibrinogen-binding activity which can be inhibited by Ca$^{2+}$ (Ni Eidhin et al., 1998). The N-terminal A domain of ClfB shares 26 % sequence identity with the A domain of ClfA and is subdivided into three independently folded subdomains N1, N2 and N3. ClfB is also responsible for the clumping of *S. aureus* cells when suspended in soluble fibrinogen and for adherence to immobilised fibrinogen *in vitro* (Ni Eidhin et al., 1998). Unlike ClfA, the A domain of ClfB recognizes the repeat 5 of the αC-region of fibrinogen and can also bind to loricrin, cytokeratin 10 and cytokeratin 8 (Ni Eidhin et al., 1998; O'Brien et al., 2002b; Haim et al., 2010). The 3D structure of ClfB was solved in complex with Fg and K10 peptides which facilitated identification of a common 'motif' (GSSGXXG) which is found in many human proteins (Ganesh et al., 2011).

Cytokeratins belong to the family of intermediate filaments (IF) which provide mechanical strength to vertebrate epithelial cells. Cytokeratin 10 is expressed both within and on the surface of desquamated epithelial cells of the anterior nares and ClfB can promote adhesion to these cells *in vitro* (O'Brien et al., 2002b; Corrigan et al., 2007). The ClfB binding site in cytokeratin 10 has been localised to the tail region of cytokeratin 10 which is rich in glycine/serine repeats and is believed to form structures known as Ω loops (Walsh et al., 2004). *Lactococcus lactis* expressing ClfB with amino acid substitutions in the putative ligand-binding trench between subdomains N2 and N3 was defective in adherence to both immobilised fibrinogen and cytokeratin 10. This suggests that both ligands bind to the same or overlapping site(s) in the ClfB trench by a similar mechanism (Walsh et al., 2008).

ClfB is important in nasal colonization in rodents and humans (Schaffer et al., 2006; Wertheim et al., 2008). Studies using a murine model of nasal colonization demonstrated that immunization (both intranasally and systemically) with recombinant ClfB was protective (Schaffer et al., 2006). In addition, systemic administration of a monoclonal antibody directed against ClfB that inhibited *S. aureus* binding to mouse cytokeratin 10 protected against colonization in naïve mice. The titres of anti-ClfB antibodies have been observed to be higher in non-carriers than in individuals who have
S. aureus present in their nares (carriers) (Dryla et al., 2005) suggesting that anti-ClfB antibodies may help protect against nasal colonization. ClfB may prove to be a suitable candidate antigen for vaccination against nasal colonization by S. aureus as the clfB gene was present in all strains examined (Peacock et al., 2002). ClfB is capable of promoting platelet activation by a fibrinogen-dependent and also a complement-dependent mechanism that is similar to but slower than that of ClfA (Miajlovic et al., 2007).

1.3.1.6 Sdr proteins

Serine-aspartate repeat (Sdr) proteins C, D and E are LPXTG-anchored proteins. The sdrC, sdrD and sdrE genes are tandemly arrayed on the S. aureus chromosome although some strains do not contain all three (Josefsson et al., 1998; Sabat et al., 2006). The three proteins have a similar structure to ClfA and ClfB, but they contain B repeats between the N-terminal region A and the C-terminal region R. SdrC and SdrD promote adhesion of bacteria to an unknown ligand on desquamated nasal epithelial cells (Corrigan et al., 2009). SdrC binds β-neurexin 1 ectodomain and its expression increases adherence to cultured mammalian cells expressing β-neurexin on their surface (Barbu et al., 2010). Expression of SdrE by Lactococcus lactis promoted platelet aggregation in previous studies (O’Brien et al., 2002a). SdrD and SdrE are among the most immunogenic surface proteins (in terms of antibody titre). Immunization with recombinant SdrD and SdrE in combination with IsdA and IsdB provided a high level of protection against challenge in a murine kidney abscess infection model (Stranger-Jones et al., 2006). A correlation was observed between carriage/invasive strains and the presence of the sdrE gene (Peacock et al., 2002).

SdrE bears significant homology to the S. aureus bone sialoprotein-binding protein (Bbp) from S. aureus strain O24 which may be important in the localization of bacteria to bone tissue. Bbp is a 97-kDa protein which possesses an A and a B domain which share 76% and 96% identity, respectively, with the corresponding domains in SdrE (Tung et al., 2000). Bbp binds to the α-chain of Fg at a different region to ClfB (Vazquez et al., 2011). Detection of serum IgG directed against Bbp could serve as a marker of osteomyelitis (Persson et al., 2009).
1.3.1.7 Fibronectin-binding surface proteins

Fibronectin is a multifunctional, extracellular matrix glycoprotein found in blood plasma at high concentrations and is widely dispersed in the body, being present on the surface of cells, in the ECM and in all body fluids. There is only one copy of the human fibronectin gene and targeted inactivation of this gene leads to early embryonic death as a result of defects in fundamental tissue structures such as the mesoderm and neural tube (George et al., 1993). The fibronectin molecule is a dimer composed of two covalently linked 250 kDa subunits. It contributes to a number of cell processes including adhesion, migration, growth and differentiation (Pankov & Yamada, 2002). The ability to bind to immobilised fibronectin is a characteristic feature of many *S. aureus* strains (Peacock et al., 2000). *S. aureus* has two closely related cell wall-associated fibronectin-binding proteins (FnBP) A and B. FnBPA and FnBPB are encoded by two closely linked but independently transcribed genes fnhA and fnhB, respectively, and are expressed on the cell surface predominately in the exponential phase of growth (Signas et al., 1989; Jonsson et al., 1991; Saravia-Otten et al., 1997). All of the sequenced strains except strain MRSA252 and the bovine strain RF122 contain genes encoding both FnBPA and FnBPB. Strains MRSA252 and RF122 both encode the FnBPA protein only. The expression of either FnBPA or FnBPB on the surface of *S. aureus* is sufficient to promote bacterial adhesion to immobilised fibronectin (Greene et al., 1995).

The multifunctionality in ligand-binding demonstrated by *S. aureus* MSCRAMMs such as Spa, ClfA and ClfB is also apparent with FnBPA and FnBPB. FnBPA and FnBPB have structurally similar N-terminal A domains which, in addition to binding fibronectin, promote adhesion to immobilised fibrinogen and elastin (Wann et al., 2000; Roche et al., 2004). Each FnBP protein also promotes the accumulation of biofilm by some MRSA strains growing in glucose-containing media (O'Neill et al., 2008).

FnBPA and FnBPB have considerable organizational and sequence similarity and are composed of a number of distinct domains. Both proteins contain a secretary signal sequence at the N-terminus and a C-terminal LPETG motif required for sortase-mediated anchoring of the proteins to the cell wall peptidoglycan. The N-terminal A domains of FnBPA and FnBPB are exposed on the cell surface and promote binding to fibrinogen and elastin. Located distal to the A domains of FnBPA and FnBPB are unfolded regions which contain multiple, tandemly arranged motifs that mediate binding to fibronectin.
The FnBPA and FnBPB A domains vary substantially between strains despite retaining ligand-binding functions. Antigenic variation of the A domain may be involved in evasion of the host immune response (Loughman et al., 2008; Burke et al., 2010). The A domains of FnBPA and FnBPB of 8325-4 are 45% identical and are predicted to consist of N1, N2 and N3 sub-domains (Jonsson et al., 1991). The N2N3 domains are predicted to bind to fibrinogen by the "dock lock latch" model described in section 1.3.1.9. FnBPA and FnBPB bind to the same region of fibrinogen as ClfA, the flexible C-terminus of the γ-chain (Wann et al., 2000). Elastin is also bound between the N2N3 domains and is thought to interact with FnBPA in a similar manner to fibrinogen (Keane et al., 2007).

Fibronectin contains five N-terminal type I modules which are the primary binding site for FnBPs (Sottile et al., 1991). A second binding site in fibronectin has been identified in the heparin-binding type III module 14 (Bozzini et al., 1992). Eleven fibronectin-binding repeats each with type I module-binding motifs have been identified in FnBPA. These repeats do not have an ordered structure until ligand binding takes place (House-Pompeo et al., 1996; Schwarz-Linek et al., 2003). FnBPs interact with fibronectin by a tandem β-zipper mechanism (Figure 1.7). Recently the crystal structure of two fibronectin binding repeats of FnBPA in complex with type I modules of fibronectin has been solved (Bingham et al., 2008). The fibronectin binding repeats form anti-parallel strands along four adjacent type I modules of fibronectin. The disordered nature of the fibronectin binding repeats is believed to facilitate the formation of large intermolecular interfaces allowing one FnBP molecule to bind up to 9 molecules of fibronectin (Matsuka et al., 2003; Bingham et al., 2008).

The ability of *S. aureus* to adhere to and to invade endothelial and epithelial cells is mediated by FnBPs (Ogawa et al., 1985; Dziewanowska et al., 1999). The FnBPs form a fibronectin-bridge between *S. aureus* and the α5β1 integrin on host cells (Sinha et al., 1999; Fowler et al., 2000). Invasion of cells may provide a means of evading the host immune response and of spreading to tissues such as bones and joints. The FnBPs are believed to be important in establishing infections by promoting adhesion of *S. aureus cells* to implanted devices coated in fibronectin (Vaudaux et al., 1993; Arrecubieta et al., 2006). Bacterial cells expressing FnBPA promoted heart valve colonization and invasion of surrounding endothelium in experimental endocarditis models (Que et al., 2001; Que et al., 2005; Que & Moreillon, 2011). Patients recovering from invasive *S. aureus*
Figure 1.7 Tandem β-zipper mechanism of fibronectin binding by *S. aureus*
(A) Schematic diagram of human fibronectin. Fibronectin monomers are linked at their C-termini by a disulphide bond to produce the mature dimer of approximately 500 kDa found in plasma. Each monomer is composed of three different types of proteins modules, F1 (blue), F2 (pink) and F3 (green). (B) The crystal structure of FnBR peptides from *S. aureus* FnBPA in complex with NTD F1 module pairs from Fn. The structure illustrates a tandem β-zipper mechanism for binding whereby the bacterial peptide contributes a fourth strand to the triple stranded β-sheet of sequential F1 modules. The N-terminal F1 modules of fibronectin each consist of a β-sandwich of two antiparallel β-sheets; a double-stranded sheet (strands A and B) and a triple-stranded sheet (Strands C, D and E). Taken from Bingham et al., 2008. (C) The fibronectin-binding domains of FnBPs contain repeated motifs forming β-strands (yellow).
infections have been shown to have higher anti-FnBPA antibody titres than healthy individuals (Dryla et al., 2005).

FnBPs are also involved in the intracellular accumulation phase of biofilm formation by MRSA strains. Methicillin-sensitive *S. aureus* strains mediate biofilm formation by a different mechanism (O'Neill et al., 2008). FnBPs play an important role in platelet aggregation by *S. aureus* cells from the exponential phase of growth (Fitzgerald et al., 2006).

### 1.3.1.8 Sas proteins

Analysis of *S. aureus* genome sequences identified ten putative LPXTG-proteins (Mazmanian et al., 2001; Roche et al., 2003a). These were named *S. aureus* surface (Sas) proteins. SasE, SasI and SasJ have since been renamed IsdA, IsdH and IsdB, respectively, and are described in previous sections. Antibodies against many Sas proteins have been detected in convalescent sera from patients with documented *S. aureus* infections indicating that expression of these proteins occurs during infection. SasG promotes bacterial adhesion to desquamated nasal epithelial cells and is also involved in the accumulation phase of biofilm formation (Roche et al., 2003b; Corrigan et al., 2007; Geoghegan et al., 2010). In addition, SasG can mask the ability of *S. aureus* surface proteins such as ClfA, ClfB and FnBPs to bind their ligands (Corrigan et al., 2007). SasG contains an N-terminal A domain followed by several homologous B-repeats. Full length SasG protein exposed on the cell surface is processed within the B domains resulting in cleaved proteins of various lengths being released into the extracellular environment. The length of SasG is crucial for its function. The binding of ligands was not masked in *S. aureus* strains expressing SasG proteins with fewer than 5 B-repeats. These strains showed reduced adherence to squamous cells and reduced biofilm formation (Corrigan et al., 2007). Recent studies have shown that it is the B repeats and not the A domain of SasG that are required for biofilm formation (Geoghegan et al., 2010). Recombinant B-repeat protein was found to bind to cells that have B domains anchored to their surface and to inhibit biofilm formation. In addition, the recombinant B-domains formed homodimers in *vitro* in a Zn$^{2+}$-dependent manner and it is proposed that homophilic association is necessary for biofilm formation (Geoghegan et al., 2010). This has been
revised. The structure of the B domains is known and is very unusual (Gruszka et al., 2012).

Sas A (also known as SraP) can promote *S. aureus* binding to platelets and is a virulence factor in the rabbit endocarditis model (Siboo et al., 2005). SasC was recently found to be involved in cell aggregation and biofilm formation (Schroeder et al., 2009). SasH (also known as AdsA) converts AMP to adenosine which is an important regulator of neutrophil activity (Thammavongsa et al., 2009). SasH/AdsA has recently been shown to have an important function in promoting survival in neutrophils (Kim et al., 2012). SasH is anti-phagocytic supporting the survival of *S. aureus* in whole human blood. The generation of adenosine by SasH promotes the survival of *S. aureus* within the neutrophil (Thammavongsa et al., 2009). No functions have been attributed to other Sas proteins (SasD, SasF and SasK).

1.3.1.9 The dock, lock and latch mechanism for ligand binding by surface proteins of staphylococci

The “dock-lock-latch” model was described for SdrG and has been proposed as a common mechanism for fibrinogen binding by staphylococcal surface-associated proteins. Docking of the fibrinogen peptide occurs in the hydrophobic trench that separates the N2 and N3 folded domains. Protein-protein interactions between residues in the trench and the ligand stabilise the docked peptide. Binding triggers structural rearrangements at the C-terminus of the N3 domain, also known as the latching peptide. Upon ligand docking, the latching peptide undergoes a directional change and crosses over the binding trench. The peptide becomes locked in place by the β-sheet which covers the binding trench and is secured by hydrogen bonding that takes place between the bound fibrinogen peptide and latching peptide/linker regions. The C-terminal latching peptide of the N3 domain then ‘latches’ in to the neighbouring N2 domain, where it inserts between strands which make up the latching cleft creating a new β-sheet in the N2 domain (G’’) (Figure 1.8). This β-strand complementation stabilizes the overall structure (Ponnuraj et al., 2003). A conserved motif (TYTFTDYVD) at the back of the latching cleft is likely to be involved in binding of the latching peptide to the latching cleft in domain N2 (Ponnuraj et al., 2003).
Figure 1.8 The dock, lock and latch-binding mechanism

**Schematic representation of the dock, lock and latch-binding mechanism.** The N2 and N3 subdomains are indicated. The flexible N3 latching peptide is represented by a red arrow. Fibrinogen docks in the trench located between N2 and N3, peptide docking redirects the N3 latching peptide which crosses over the binding trench and inserts into the N2 subdomain complementing a β-sheet.
The crystal structure of ClfA N2N3 in complex with the fibrinogen γ-chain peptide revealed that ClfA binds to the γ-chain by a variation of the dock, lock and latch mechanism (Ganesh et al., 2008). An open form of ClfA is not required because the fibrinogen γ-chain peptide was able to penetrate an artificially closed form of ClfA. Moreover, the closed form of ClfA had increased affinity for the γ-chain peptide compared to the apo-form. Another difference was the direction of the bound peptide; in SdrG the peptide formed an anti-parallel β-sheet with the C-terminal extension of subdomain N3 while in ClfA the bound peptide formed a parallel β-sheet (Ganesh et al., 2008).

1.3.1.10 Capsular polysaccharide

Capsular polysaccharide is produced by the majority of S. aureus clinical isolates (Roghmann et al., 2005, O'Riordan & Lee, 2004). Expression of capsular polysaccharide serotypes 5 and 8 reduces phagocytosis in vitro and is associated with increased virulence in animal infection models (Luong & Lee, 2002, Thakker et al., 1998). Capsule inhibits binding of antibodies to S. aureus cell surface components which hinders opsonisation (Thakker et al., 1998). It also obstructs access of phagocyte complement receptors to complement components assembled beneath the capsule layer (Cunnion et al., 2003). Reduced O-acetylation of capsular polysaccharide decreases antiphagocytic activity due to increased antibody penetration and cell surface recognition (Bhasin et al., 1998).

1.4 Regulation of virulence factors

Expression of cell surface adhesins from S. aureus is generally accepted to take place during the exponential phase of growth while secreted proteins and toxins are expressed as cells approach stationary phase (Chan & Foster, 1998; Novick, 2003). Expression of adhesins early in infection is believed to facilitate colonisation. Secreted exoproteins are produced later in infection and are involved in damage to host tissues, evading the host immune response and facilitating detachment and spreading throughout the body (McAleese et al., 2001; McGavin et al., 1997).
The regulatory network in *S. aureus* is complex and depends on the interplay of sigma factors, transcription factors and signal transduction via two-component regulatory systems. The interactions between regulators are likely to differ *in vivo* from those that have been described *in vitro.*

1.4.1 Two-component regulatory systems

Two-component regulatory systems consist of a sensor, which is often an integral membrane protein, and a cytoplasmic response regulator. In response to particular stimuli the enzymatic domain of the sensor becomes phosphorylated. This in turn activates the response regulator which acts on target genes, either directly or indirectly to initiate transcription (Stock *et al.*, 1989). Such systems include the accessory gene regulator (Agr) and the haem sensor system. The Agr system is a cell density-sensing system which is activated in the late-exponential growth phase and upregulates many secreted proteins of *S. aureus* while downregulating surface-associated proteins (Dunman *et al.*, 2001). As the accumulation of excess haem in the cytoplasm is potentially toxic, *S. aureus* has a haem sensor system (HssRS) which responds to excess haem in the cytoplasm (Torres *et al.*, 2007).

1.4.2 Transcription factors

Transcription factors interact with target gene DNA to increase or decrease the rate of transcription. In *S. aureus* transcription factors share sequence and structural homology with the staphylococcal accessory regulator (SarA) (Cheung *et al.*, 2004). The proteins in this family can be categorised on the basis of their structure: single-domain proteins that occur as homodimers and two-domain proteins that occur as monomers. SarA and its homologues have been proposed to hold promoter sequences in optimal or sub-optimal positions to increase or decrease transcription. SarA homologues include the repressor of toxins (Rot) and the multiple gene regulator (MgrA). The ferric uptake regulator (Fur) of *S. aureus* is homologous to Fur of *B. subtilis*.

SarA is a global regulator of *S. aureus* virulence factors which is predominantly expressed in late exponential growth phase (Manna & Cheung, 2001). It is encoded on
three overlapping transcripts from promoters P1, P2 and P3. Promoters P1 and P2 are dependent on the housekeeping sigma factor, SigA for transcription while P3 is dependent the stationary growth phase accessory sigma factor, SigB. SarA upregulates transcription of \textit{agr} by binding between the \textit{agr} P2 and P3 promoters. It also upregulates expression of FnBPs and haemolysins. Expression of protein A, aureolysin and the cysteine protease SspB are inhibited by SarA (Dunman \textit{et al.}, 2001). SarA activates its own expression and is downregulated by another single domain transcription factor, SarR (Bayer \textit{et al.}, 1996; Manna & Cheung, 2001).

Rot is a global regulator which repressed expression of secreted toxins and enzymes while also activating certain cell wall-associated proteins expressed in the exponential growth phase (Said-Salim \textit{et al.}, 2003). Rot therefore seems to be an antagonist of \textit{agr} regulation and is believed to be important in the early steps of infection. Many adhesins including ClfB are positively regulated by Rot. The two domain SarA homologue, SarS, is positively regulated by Rot and it in turn positively regulates Spa. Increased expression of adhesins facilitates colonisation early in infection. Rot negatively regulates factors associated with the later stages of infection such as lipase, \(\alpha\)- and \(\beta\)-haemolysins and the spl and ssp protease operons (Said-Salim \textit{et al.}, 2003).

The global regulator MgrA acts in a similar way to Agr, by upregulating exoproteins and downregulating surface-associated proteins as indicated by microarray studies and RT-PCR (Luong \textit{et al.}, 2006). When environmental iron levels are sufficiently high, Fur forms a complex with ferrous iron and acts as a transcriptional repressor. Genes regulated by Fur have a consensus sequence known as the Fur box located within their promoter regions. The Fur box sequence consists of a 19-bp inverted repeat which is highly homologous to the Fur box of \textit{E. coli} and \textit{B. subtilis} (Xiong \textit{et al.}, 2000). Fur regulates many systems involved in iron homeostasis. Expression of genes encoding the haem transport system (htsABC), the Isd proteins and \textit{S. aureus} siderophores are repressed by Fur.

### 1.5 Genetic manipulation of \textit{S. aureus}

The identification of \textit{S. aureus} virulence factors has relied on the ability of researchers to generate isogenic mutants by inactivating or deleting virulence factor gene
candidates and testing the resultant mutant strains for loss of virulence in infection models to fulfil Koch’s molecular postulates (Falkow, 1988). To validate a gene’s contribution, virulence must be restored to wildtype by complementation. Many advances in the tools available for the molecular analysis of *S. aureus* have been developed. They have focused mainly on the improvement of vectors for allelic replacement and avoiding the barriers to transformation in clinical isolates of *S. aureus*.

Electroporation is used to introduce plasmid DNA into *S. aureus*. Typically cells are grown to early logarithmic phase followed by washing with a hypertonic buffer (eg. 500 mM sucrose) to remove salts and to stabilize the cells (Augustin & Gotz, 1990; Oskouian & Stewart, 1990; Schenk & Laddaga, 1992; Lee, 1995). Once the cells are concentrated to 1-3 x 10^10 CFU/ml and purified plasmid DNA added, a defined electric pulse is discharged through the cells to facilitate the uptake of the DNA. The cells are then allowed to recover in broth for a short period of time prior to plating on media containing an antibiotic that selects for the plasmid-containing transformants. The electroporation conditions optimized for *Staphylococcus carnosus* have been adapted to *S. aureus* (Lofblom et al., 2007; Monk et al., 2012).

Genetic manipulation has not always been possible because of difficulties in transferring plasmids taken from *E. coli* into wildtype strains of *S. aureus*. Bacteria possess restriction barriers which prevent the transformation of plasmid DNA (See below).

### 1.5.1 Restriction-modification systems

Uptake of DNA may be potentially harmful or lethal to bacteria, such as bacteriophage which lyses and kills bacteria. In addition, the expression of proteins encoded by superfluous genes may result in metabolic demand which may compromise fitness of the bacteria. One way in which bacteria can control the uptake of foreign DNA is through the use of restriction-modification (R-M) systems. These systems are widespread in many types of bacteria. R-M enzymes can either exist as a single protein or a multicomponent complex and they can identify specific DNA sequences and modify them, usually by adding a methyl group. Recognition of specific sequences facilitates their ability to digest (restrict) DNA (Waldron & Lindsay, 2006). Four distinct types of restriction-modification systems have been identified, three of which are found in
staphylococci and will be discussed below. The distinguishing features of types I, II, III and IV are depicted in Figure 1.9.

1.5.1.1 Type I restriction-modification systems

Type I R-M systems are multifunctional enzymes that can catalyze both restriction and modification. The type I restriction system is composed of three subunits encoded by three closely linked genes, \( hsdR \), \( hsdM \), and \( hsdS \). R-M systems were originally referred to as host specificity systems which gave rise to the acronym \( hsd \) for “host specificity of DNA.” While \( hsdM \) and \( hsdS \) are transcribed from the same promoter, \( hsdR \) is transcribed from a separate promoter (Murray, 2000). In contrast, the type I RM genes of \( S. epidermidis \) and \( S. lugdunensis \) are clustered together (Monk & Foster, 2012).

When several genome sequences of \( S. aureus \) strains were analysed, conserved open reading frames with high homology to the type I R-M systems of other bacteria were found in all strains (Waldron & Lindsay, 2006). Functional type I R-M systems have been identified in \( Bacillus subtilis \), \( Citrobacter freundii \), \( Klebsiella pneumoniae \), \( L. lactis \), \( Mycoplasma pulmonis \), \( Pasteurella haemolytica \), \( Salmonella enterica \), and \( Staphylococcus aureus \), in addition to those originally identified in \( E. coli \) and have been shown to provide a barrier to phage infection (Murray, 2000).

\( HsdS \) is the specificity subunit which recognises specific DNA sequences via two target recognition domains (TRDs) and directs the targeted activities of both the restriction and modification subunits. \( HsdM \) is the modification subunit which serves to methylate host DNA protecting it from cleavage. It includes both the binding site for \( S\)-Adenosylmethionine (AdoMet), which is the cofactor and methyl donor for the methyltransferase activity, and the active site for DNA methylation. \( HsdR \) is the endonuclease which cleaves non-protected DNA. It includes the active site for ATP hydrolysis and other sequences essential for DNA translocation and endonuclease activity. The endonuclease activity of \( HsdR \) requires ATP, AdoMet, and Mg\( \text{2} \). \( HsdM \) and \( HsdS \) subunits are both necessary and sufficient for methyltransferase activity (Murray, 2000).

Type I enzymes recognize asymmetric nucleotide sequences that comprise two components (one of 3 or 4 bp and the other of 4 or 5 bp) that are separated by a nonspecific spacer of 6 to 8 bp and methylate one adenine residue in each component of
the target sequences on opposite strands. The methylation state of the target sequence determines whether the enzyme will behave as an endonuclease or a methyltransferase. The enzyme, while bound to its target site on unmodified DNA, pulls the DNA towards itself simultaneously in both directions in an ATP-dependent manner which allows enzymes bound to different sites on the same molecule to be brought together.

DNA cleavage by the endonuclease complex HsdS₁HsdM₂HsdR₂, which assembles on unmethylated DNA, occurs when translocation is impeded, either by collision with another translocating complex or by the topology of the DNA substrate. Cleavage is inhibited by the HsdS₁HsdM₂ complex recognising a specific DNA sequence and subsequently methylating hemi-methylated DNA (Murray, 2000).

1.5.1.2 Type IV restriction-modification systems

The type IV restriction system comprises a single protein which detects the methylation status of DNA and cleaves DNA that contains modified cytosine bases (Raleigh et al., 1988). It was shown to be the dominant barrier preventing the uptake of foreign DNA by S. aureus (Corvaglia et al., 2010). The biochemical properties of this endonuclease, SauUSI, were recently characterized, it recognizes cytosine methylated DNA in the motif C/GmCNGC/G (Xu et al., 2011).

Homologues of sauUSI occur in S. epidermidis and S. pseudintermedius and also in some enterococci, bacilli and lactobacilli. Examples from E. coli are mcrA, mcrBC which recognize 5-hydroxymethylcytosine and N-4-methylcytosine, respectively, while and mrr recognizes N-6-methyladenine as foreign (Kelleher and Raleigh, 1991) (Monk & Foster, 2012). SauUSI is highly conserved in S. aureus and degrades plasmids isolated from E. coli K-12 strains DH5α, TOP10, XL1-Blue and DH10B which methylate both adenine (dam) and cytosine (dcm) residues and are widely used for cloning.

The S. aureus strains N315 and Mu50 from CC5 contain an allele of sauUSI with a nonsense mutation within the middle of the gene which renders them permissive to transformation with plasmid DNA isolated from E. faecalis (Corvaglia et al., 2010). This may account for the enhanced spread of antibiotic resistance between these organisms in the hospital environment (Zhu et al., 2008). The inactivation of the sauUSI gene rendered clinical S. aureus strains hypersusceptible to the acquisition of foreign DNA, such as the vancomycin resistance gene from E. faecalis (Corvaglia et al., 2010). The sauUSI gene of
Figure 1.9 Restriction modification system types I, II, III and IV.

Schematic representation of the distinguishing features and organisation of the genetic determinants and subunits of Restriction-Modification system of types I, II, III and IV. ENase (Restriction endonuclease), MTase (Methyltransferase) activities are indicated. Adapted from Murray et al, 2000.
SA564 shares 98% identity to sauUSI of 8325. Disruption of this gene in SA564 allowed this strain to accept plasmid DNA from *E. coli* K-12 (Corvaglia et al., 2010).

### 1.5.1.3 Other restriction-modification systems

Type II restriction endonucleases, which comprise a site-specific methylase and a site-specific restriction endonuclease, are widely used as reagents by molecular biologists (Corvaglia et al., 2010). They typically recognise short (4-8 bp) palindromic sequences and cleave DNA within or in close proximity to the recognition site (Pingoud et al., 2005). Cleavage can be prevented by DNA methylation. The first *S. aureus* R-M system to be characterized in detail was a type II system (Sau3A) that specifically digests DNA at GATC sites. This system is rare in *S. aureus* isolates (Corvaglia et al., 2010).

The type III restriction system is also a multimeric enzyme (M$_2$R$_2$) encoding a methylase responsible for sequence specificity and an endonuclease responsible for cleavage of unmethylated DNA. DNA cleavage by the type III system, which is not found in any *S. aureus* strain, requires the binding of a pair of RM enzymes at two distant, inversely orientated recognition sequences followed by helicase-catalysed ATP hydrolysis and long-range communication (Toth et al., 2012).

### 1.6 Rationale for this study

*Staphylococcus aureus* is a highly prevalent, opportunistic, multifactorial pathogen and is the agent most frequently associated with contagious bovine mastitis which is one of the most economically important diseases in the dairy industry. To prevent the morbidity and mortality associated with bovine mastitis caused by *S. aureus*, and to decrease the associated economic costs, an improved means of controlling bovine mastitis is required. A logical approach is rational design of a staphylococcal vaccine. *S. aureus* is characterised by a wide array of virulence factors and although vaccination has been studied as a control measure for *S. aureus* mastitis for several years the antigenic diversity among different strains has hampered the development of an effective vaccine; therefore prevention of intramammary infection (IMI) has not been demonstrated (Hu et al., 2010; Scarpa et al., 2010). RF122 (alternately referred to as ET3-1) represents a bovine clone of *S. aureus* that is geographically widespread and is believed to be
particularly well adapted to the bovine host (Herron-Olson et al., 2007). Surface proteins singly and in combination have shown potential in combatting diseases caused by *S. aureus* in humans (Josefsson et al., 2001; Kuklin et al., 2006; Schaffer et al., 2006; Stranger-Jones et al., 2006; Zhou et al., 2006).

### 1.7 Aims and objectives

The initial aim of this project was to identify a protein, or proteins, that are ubiquitously expressed by mastitis-causing *S. aureus* strains in order to identify potential vaccine candidates to combat bovine mastitis. Vaccine candidates were to be investigated for their potential by (i) preparation of recombinant proteins for use as administrative antigens in a multicomponent vaccine and (ii) deletion of gene candidates from the RF122 chromosome to test the specificity of immune response to vaccination with recombinant antigens. Another objective was to investigate the role of RF122 proteins in survival in milk and interaction with the mammary gland epithelium. The establishment of a genetic system for manipulating RF122 led to work regarding the restriction systems employed by mastitis-causing strains to be investigated.
Chapter 2

Materials and Methods
2.1 Chemicals and reagents.

General laboratory chemicals, reagents and media were obtained from Sigma-Aldrich, unless otherwise stated.

2.2 Bioinformatic analysis

2.2.1 Bioinformatic and phylogenetic analysis.

Gene sequences from the RF122 genome were obtained from the GenBank database (http://www.ncbi.nlm.nih.gov/nuccore/NC_007622.1) and analysed by performing DNA and amino acid sequence alignments using the CLUSTALW program (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and compared to genes from well-characterised strains. Gene sequences that were not available in the GenBank database were identified by performing BLAST searches (described in section 2.2.2) using gene sequences of well-characterised strains and were analysed in a similar manner. The concatenated MLST allele sequences of *S. aureus* strains were downloaded from the MLST database (http://saureus.mlst.net/). A phylogenetic analysis of concatenated MLST allele sequences was conducted using the MEGA version 4 program (Tamura *et al.*, 2007).

2.2.2 The Basic Local Alignment Search Tool (BLAST)

The BLAST online tool was used to find regions of local similarity between DNA sequences. Using this program, DNA sequences were compared to sequence databases and the statistical significance of matches was calculated. The three algorithm options that were used were megablast which is optimized for highly similar sequences, discontiguous megablast which is optimized for more dissimilar sequences or blastn which is optimized for somewhat similar sequences (http://blast.ncbi.nlm.nih.gov/).
2.3 Bacterial strains and growth conditions

The bacterial strains used in this study are listed in Table 2.1. *Escherichia coli* was routinely grown on Luria (L) agar or in L-broth (Difco) whereas *S. aureus* was grown on Trypticase Soy agar (TSA, Oxoid) or broth (TSB) for liquid cultures. Broth cultures were grown in an orbital shaker at 200 r.p.m. at 37 °C. The following antibiotics were added to the media as required: ampicillin at 100 µg/ml, chloramphenicol at 10 µg/ml, erythromycin at 10 µg/ml, kanamycin at 50 µg/ml and tetracycline at 2 µg/ml.

To induce the expression of iron regulated surface determinant proteins, strains were grown in RPMI 1640. RPMI medium is an iron deficient medium which was originally designed for the cell culture of human leukocytes (Moore *et al.*, 1967). 5ml overnight cultures grown in RPMI were adjusted to an OD_{600nm} of 0.05 in a final volume of 25 ml and allowed to grow to stationary phase at 37 °C with shaking (200 r.p.m.).

Stocks of bacterial strains were made by supplementing broth cultures with 80 % (v/v) glycerol, snap-freezing in liquid nitrogen and were maintained at -70 °C. Bovine *S. aureus* strains were isolated from geographically distinct locations and had been characterised by multi-locus sequence typing (MLST). Each strain represents a different sequence type (ST).

Viable counts of bacteria were measured using the standard Miles and Misra procedure (Miles *et al.*, 1938).

2.4 Detecting expression of *S. aureus* cell wall-associated proteins

2.4.1 Preparation of solubilised proteins associated with the staphylococcal cell wall

Once bacteria were grown to a specified OD_{600nm}, cultures were harvested by centrifugation at 3000 x g and cells were washed twice in PBS. Cells were adjusted to an OD_{600nm} of 10 and following centrifugation at top speed in a bench top centrifuge for 2 min. were resuspended in 250 µl of digestion buffer (20 mM Tris-HCl, 10 mM MgCl₂, 30 % (w/v) raffinose, pH7.5) containing complete EDTA-free protease inhibitor cocktail (20 µl of 25X stock per 250 µl digestion buffer) and lysostaphin (5 µl of 10 mg/ml stock per 250 µl digestion buffer). Cells were incubated at 37 °C for 10 min to allow digestion of the cell wall which served to release any cell-surface associated proteins that were
covalently anchored to the cell wall. The lysate was spun at 5000 r.p.m. for 10 min at room temperature and the resulting supernatant containing the cell wall-associated proteins was analysed by SDS-PAGE (described in section 2.7.1). The supernatant could be stored at -20 °C for future analysis.

2.4.2 Antibodies to *S. aureus* Bbp antibodies

Antibodies to *S. aureus* Bbp were a gift from Prof. P. Speziale, University of Pavia, Italy. Polyclonal antibodies raised against the N2N3 domain of Bbp were generated by immunising mice with rBbp\textsubscript{53-384} cloned from *S. aureus* RF122. All antibodies and their working dilutions are listed in Table 2.2.

2.4.3 Western immunoblotting

Non-specific binding to PVDF membranes was blocked by incubation for 2-18 h in 10% (w/v) skimmed milk powder (Marvel) in TS buffer (10 mM Tris-HCl, 0.9% NaCl, pH 7.4). Primary antibodies (raised against Newman-derived proteins) diluted in 10% (w/v) Marvel/TS buffer were incubated with the membranes for 1 h at room temperature with shaking. The membranes were then washed three times with gentle agitation for 10 min each in TS-Tween (10 mM Tris-HCl, 0.9% NaCl, pH 7.4, 0.05% (v/v) Tween 20) to remove unbound antibodies. To detect bound antibodies membranes were incubated with HRP-conjugated protein A-peroxidase diluted 1:500 in 10% (w/v) Marvel/TS buffer for 1 h at room temperature with shaking. The membranes were then washed three times with gentle agitation for 10 min each in TS-Tween to remove unbound protein A-peroxidase. Reactive bands were visualised using the LumiGLO reagent and peroxide detection system (Cell Signalling Technology) as recommended by the manufacturer, exposed to X-Omat autoradiographic film (Kodak) and visualised using manual development with developer and fixer solutions (Kodak).
<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant Characteristics</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XL-1 Blue</td>
<td>Propagation of plasmids</td>
<td>Stratagene</td>
</tr>
<tr>
<td>DC10B</td>
<td>Dcm-deficient in the DH10B background; Dam methylation only</td>
<td>Monk <em>et al.</em>, 2012</td>
</tr>
<tr>
<td><strong>S. aureus</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newman</td>
<td>ST8; CC8</td>
<td>Duthie &amp; Lorenz, 1952</td>
</tr>
<tr>
<td>Newman <em>isdA</em></td>
<td>Newman derivative deficient in IsdA, ST8; CC8</td>
<td>Clarke <em>et al.</em>, 2007</td>
</tr>
<tr>
<td>Newman <em>isdB</em></td>
<td>Newman derivative deficient in IsdB, ST8; CC8</td>
<td>Helen Miajlovic, PhD thesis</td>
</tr>
<tr>
<td>Newman <em>isdH</em></td>
<td>Newman <em>isdH</em>:Em&lt;sup&gt;R&lt;/sup&gt; ST8; CC8</td>
<td>Visai <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>Newman <em>clfB</em></td>
<td>Newman <em>clfB</em>:lacZEm&lt;sup&gt;R&lt;/sup&gt; ST8; CC8</td>
<td>Fitzgerald <em>et al.</em>, 2006</td>
</tr>
<tr>
<td>Newman <em>sdrCDE</em></td>
<td>Δ<em>sdrCDE</em>:Tc&lt;sup&gt;R&lt;/sup&gt; ST8; CC8</td>
<td>O’Brien <em>et al.</em>, 2002</td>
</tr>
<tr>
<td>P1</td>
<td>Rabbit virulent strain. ST973</td>
<td>Sherertz <em>et al.</em>, 1993</td>
</tr>
<tr>
<td>P1 <em>/fnbAfnbB</em></td>
<td>P1 derivative deficient in FnBPA and FnBPB. <em>fnbA</em>:Tc&lt;sup&gt;R&lt;/sup&gt; <em>fnbB</em>:Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Roche <em>et al.</em>, 2004</td>
</tr>
<tr>
<td>SH1000</td>
<td>Functional <em>rsbU</em> derivative of 8325-4 CC8</td>
<td>Horsburgh <em>et al.</em>, 2002</td>
</tr>
<tr>
<td>SH1000 <em>/fnbAfnbBclfAclfB</em></td>
<td>SH1000 derivative, <em>fnbA</em>:Tc&lt;sup&gt;R&lt;/sup&gt; <em>fnbB</em>:Em&lt;sup&gt;R&lt;/sup&gt; <em>clfA</em>5 frameshift mutation <em>clfB</em>:Em&lt;sup&gt;R&lt;/sup&gt; CC8</td>
<td>Fiona Burke, PhD thesis</td>
</tr>
<tr>
<td>RF122 <em>isdA</em></td>
<td>RF122 derivative deficient in IsdA, unmarked mutation ST151</td>
<td>This study</td>
</tr>
<tr>
<td>RF122 <em>isdB</em></td>
<td>RF122 derivative deficient in IsdB, unmarked mutation ST151</td>
<td>This study</td>
</tr>
<tr>
<td>RF122 <em>clfB</em></td>
<td>RF122 derivative deficient in ClfB, unmarked mutation ST151</td>
<td>This study</td>
</tr>
<tr>
<td>Strain</td>
<td>Relevant Characteristics</td>
<td>Source/reference</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>RF122 bhp</td>
<td>RF122 derivative deficient in Bbp, unmarked mutation ST151</td>
<td>This study</td>
</tr>
<tr>
<td>RF122 sab2370c</td>
<td>RF122 derivative deficient in Sab2370c, unmarked mutation ST151</td>
<td>This study</td>
</tr>
<tr>
<td>RN4220</td>
<td>Restriction-deficient derivative of 8325-4 CC8</td>
<td>Kreiswirth et al., 1983</td>
</tr>
<tr>
<td>RF122t</td>
<td>Electrocompetent derivative of RF122 ST151</td>
<td>R. Fitzgerald, University of Edinburgh</td>
</tr>
<tr>
<td>MSA 1006</td>
<td>Bovine isolate ST8</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>DS37</td>
<td>Bovine mastitis isolate ST20</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>MSA 1363</td>
<td>Bovine mastitis isolate ST25</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF105</td>
<td>Bovine mastitis isolate ST71</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF26</td>
<td>Bovine mastitis isolate ST97</td>
<td>Fitzgerald et al., 2000</td>
</tr>
<tr>
<td>MSA915</td>
<td>Bovine mastitis isolate ST115</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>DS36</td>
<td>Bovine mastitis isolate ST126</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF287</td>
<td>Bovine mastitis isolate ST133</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF122</td>
<td>Bovine mastitis isolate ST151</td>
<td>Herron-Olson et al., 2007</td>
</tr>
<tr>
<td>MSA 1047</td>
<td>Bovine mastitis isolate ST350</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>MSA 1011</td>
<td>Bovine mastitis isolate ST352</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>DS42</td>
<td>Bovine mastitis isolate ST479</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>PSA 17.1</td>
<td>Bovine mastitis isolate ST693</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>MSA 1468</td>
<td>Bovine mastitis isolate ST694</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>PSA 13.1</td>
<td>Bovine mastitis isolate ST695</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>DS35</td>
<td>Bovine mastitis isolate ST696</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>Strain</td>
<td>Relevant Characteristics</td>
<td>Source/reference</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>DS70</td>
<td>Bovine mastitis isolate ST697</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>PSA 5</td>
<td>Bovine mastitis isolate ST698</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>MSA 1547</td>
<td>Bovine mastitis isolate ST699</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>MSA 1007</td>
<td>Bovine mastitis isolate ST708</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>DS41</td>
<td>Bovine mastitis isolate ST709</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF80</td>
<td>Bovine mastitis isolate ST151</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF113</td>
<td>Bovine mastitis isolate ST151</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF114</td>
<td>Bovine mastitis isolate ST151</td>
<td>Smyth et al., 2009</td>
</tr>
<tr>
<td>RF121</td>
<td>Bovine mastitis isolate ST151</td>
<td>Smyth et al., 2009</td>
</tr>
</tbody>
</table>

R Resistance
Tc Tetracycline
Em Erythromycin
Δ Deletion
<table>
<thead>
<tr>
<th>Antibody</th>
<th>Relevant features</th>
<th>Working dilution/concentration</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rabbit anti-IsdA</td>
<td>Raised against full length IsdA</td>
<td>1:2500</td>
<td>Corrigan et al., 2009</td>
</tr>
<tr>
<td>Rabbit anti-IsdB</td>
<td>Raised against full length IsdB</td>
<td>1:5000</td>
<td>Miajlovic et al., 2010</td>
</tr>
<tr>
<td>Rabbit anti-IsdH</td>
<td>Raised against recombinant IsdH&lt;sub&gt;41-321&lt;/sub&gt;</td>
<td>1:10000</td>
<td>P. Speziale</td>
</tr>
<tr>
<td>Rabbit anti-ClfA</td>
<td>Raised against recombinant ClfA&lt;sub&gt;221-559&lt;/sub&gt;</td>
<td>1:5000</td>
<td>P. Speziale</td>
</tr>
<tr>
<td>Rabbit anti-ClfB</td>
<td>Raised against recombinant ClfB&lt;sub&gt;N123&lt;/sub&gt;</td>
<td>1:5000</td>
<td>Ni Eidhin et al., 1998</td>
</tr>
<tr>
<td>Rabbit anti-SdrE</td>
<td>Raised against recombinant SdrE&lt;sub&gt;41-544&lt;/sub&gt;</td>
<td>1:10000</td>
<td>Josefsson et al., 1998</td>
</tr>
<tr>
<td>Polyclonal mouse anti-Bbp</td>
<td>Raised against recombinant Bbp&lt;sub&gt;53-382&lt;/sub&gt;</td>
<td>1:2500</td>
<td>P. Speziale</td>
</tr>
<tr>
<td>Polyclonal sheep alkaline phosphatase-conjugated anti-DIG antibody</td>
<td>Raised against digoxigenin</td>
<td>1:10000</td>
<td>Roche</td>
</tr>
</tbody>
</table>
2.4.4 Ligand-affinity blotting

Fibronectin-binding proteins in *S. aureus* cell wall extracts were detected by probing membranes with biotinylated fibronectin. Bovine fibronectin (0.5 mg/ml) was incubated with NHS-biotin (final concentration of 2 mg/ml) for 15 min at room temperature. The reaction was stopped by addition of 10 mM NH₄Cl and excess biotin was removed by dialysis against PBS for 16 h at 4 °C. Membranes were blocked with 10 % (w/v) skimmed milk in TS Tween buffer for 16 h at 4 °C. Biotinylated fibronectin, diluted to 30 μg/ml in TS Tween /10 % (w/v) skimmed milk, was incubated with the membranes for 1hr at room temperature and an additional 2 h at 4 °C. Excess fibronectin was removed by washing with TS Tween and bound fibronectin detected using peroxidase-conjugated streptavidin (Roche; diluted 1/2000 in 10 % skimmed milk / TS Tween) and developed by chemiluminescence.

2.5 Adherence of bacteria to immobilized fibronectin, cytokeratin 10 and fibrinogen

96-well flat bottomed microtitre plates were coated with 100 μl serial dilutions of bovine fibronectin (Sigma) in PBS, recombinant murine cytokeratin 10 (rMK10, 454-570) (Walsh et al., 2004) in carbonate buffer (15mM Na₂CO₃, 35mM NaHCO₃, pH 9.6) or bovine fibrinogen (Calbiochem) in PBS. Control wells contained sodium carbonate or PBS only. NUNC microtitre plates were used with cytokeratin/sodium carbonate and Sarstedt microtitre plates were used with fibrinogen/fibronectin/PBS. Following 16 h incubation at 4 °C, microtitre wells were washed three times with PBS. 100μl of Marvel (5% w/v) in PBS was added to each well and plates were incubated for 2 h at 37 °C. Wells were washed three times with PBS. Bacteria were grown to either exponential phase or stationary phase. Bacterial suspensions were adjusted to an OD₆₉₀ nm of 1 and 100 μl of bacteria was added to each well. Plates were incubated for a further 2 h at 37 °C. After three washes with PBS, adherent cells were fixed with formaldehyde (25 % v/v) for 15 min and then stained with 100 μl of crystal violet (0.5 % w/v) for 1 min. Excess stain was removed by washing with PBS. Cell-bound crystal violet was solubilised using 100 μl acetic acid (5 % v/v) and the absorbance at 570 nm was measured using an ELISA plate reader (Multiskan EX, Labsystems). Bacterial adherence binding graphs shown throughout this thesis are graphs of individual experiments that are representative of three
independent experiments. Each plot in such graphs represents the average of triplicate wells.

2.6 DNA manipulation

2.6.1 DNA manipulation

DNA manipulation techniques were performed using standard methods (Sambrook, 1989). Restriction endonucleases were purchased from Promega or Fermentas and were used according to the manufacturers' instructions. DNA ligase was purchased from Promega. Recombinant plasmids isolated from *E. coli* were established in *S. aureus* by electroporation. DNA cloning into the pKOR1 vector was performed using the BP Clonase II enzyme (Invitrogen) following manufacturer's protocol. All confirmatory DNA sequencing was performed by GATC Biotech or Eurofins.

2.6.2 Isolation of plasmid and genomic DNA

All plasmids and derivatives are listed in Table 2.3. Plasmid DNA was extracted from bacteria using Promega Wizard SV Plus Minipreps DNA purification system as recommended by the supplier. To digest the cell wall peptidoglycan when extracting plasmid from *S. aureus* cells, 200 µg lysostaphin (AMBI, New York) was added and sucrose (500 mM) was added to stabilise the cells. After addition of lysostaphin and sucrose, cells were left at room temperature for 30 min. Preparation of genomic DNA was performed using the Genomic DNA purification kit (Edge Biosystems) according to the supplier's protocol, except that *S. aureus* cells were treated with 200 µg lysostaphin for 20 min at 37 °C.

2.6.3 Polymerase chain reaction

Polymerase chain reaction (PCR) amplification was carried out in a DNA thermal cycler (Techne). Plasmid DNA (10 ng) or *S. aureus* genomic DNA (20 ng) was used as a template. Primers were purchased from Integrated DNA Technologies and their sequences are listed in Table 2.4. DNA fragments for cloning PCR reactions were
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant Characteristics</th>
<th>Marker(s)</th>
<th>Source/References</th>
</tr>
</thead>
<tbody>
<tr>
<td>pQE30</td>
<td><em>E. coli</em> vector for the expression of hexa-histidine tagged recombinant proteins</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Qiagen</td>
</tr>
<tr>
<td>pQE30::rIsdB&lt;sub&gt;41-594&lt;/sub&gt;</td>
<td>pQE30 derivative encoding lsdB residues 41-594</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pQE30::rClfB&lt;sub&gt;201-541&lt;/sub&gt;</td>
<td>pQE30 derivative encoding the N2N3 subdomain of ClfBN2N3 (residues 201-541)</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pQE30::rBbp&lt;sub&gt;53-382&lt;/sub&gt;</td>
<td>pQE30 derivative encoding the N2N3 subdomain of Bbp (residues 53-382)</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pQE30::rSab2370c</td>
<td>pQE30 derivative encoding full length Sab2370c</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pIMK3</td>
<td>Shuttle vector with a high-level of IPTG-controlled gene expression</td>
<td>Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Monk <em>et al.</em>, 2008</td>
</tr>
<tr>
<td>pIMK3::rDcm</td>
<td>pIMK3 derivative encoding full length Dcm</td>
<td>Kan&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pKOR1</td>
<td>Temperature sensitive shuttle plasmid</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt; in <em>E. coli</em>, Cm&lt;sup&gt;R&lt;/sup&gt; in <em>S. aureus</em></td>
<td>Bae &amp; Schneewind, 2006</td>
</tr>
<tr>
<td>pKOR1&lt;sub&gt;ΔisdA&lt;/sub&gt;</td>
<td>pKOR1 derivative containing sequences for creating isdA mutation</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt; in <em>E. coli</em>, Cm&lt;sup&gt;R&lt;/sup&gt; in <em>S. aureus</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKOR1&lt;sub&gt;ΔisdB&lt;/sub&gt;</td>
<td>pKOR1 derivative containing sequences for creating isdB mutation</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt; in <em>E. coli</em>, Cm&lt;sup&gt;R&lt;/sup&gt; in <em>S. aureus</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKOR1&lt;sub&gt;ΔclfB&lt;/sub&gt;</td>
<td>pKOR1 derivative containing sequences for creating clfB mutation</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt; in <em>E. coli</em>, Cm&lt;sup&gt;R&lt;/sup&gt; in <em>S. aureus</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKOR1&lt;sub&gt;Δbhp&lt;/sub&gt;</td>
<td>pKOR1 derivative containing sequences for creating bhp mutation</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt; in <em>E. coli</em>, Cm&lt;sup&gt;R&lt;/sup&gt; in <em>S. aureus</em></td>
<td>This study</td>
</tr>
<tr>
<td>pKOR1&lt;sub&gt;Δsab2370c&lt;/sub&gt;</td>
<td>pKOR1 derivative containing sequences for creating <em>sab2370c</em> mutation</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt; in <em>E. coli</em>, Cm&lt;sup&gt;R&lt;/sup&gt; in <em>S. aureus</em></td>
<td>This study</td>
</tr>
</tbody>
</table>

<sup>R</sup> Resistance
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5'-3')$^{ab}$</th>
<th>5'Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>IsdB RP 5' F</td>
<td>GGGGGATCCGCTGAGGAAACAGGTGGTAC</td>
<td>BamHI</td>
</tr>
<tr>
<td>IsdB RP 3' R</td>
<td>GGGGCATGCTAATGATTTTGCTTTATTTTCT TGTG</td>
<td>Sphl</td>
</tr>
<tr>
<td>ClfB RP 5' F</td>
<td>GGGGGATCCGCTGAGGAAACAGGTGGTAC</td>
<td>BamHI</td>
</tr>
<tr>
<td>ClfB RP 3' R</td>
<td>GGGGAAGCTTTTTTACTGCTGAATCACCATC</td>
<td>HindIII</td>
</tr>
<tr>
<td>Bbp RP 5' F</td>
<td>GGGGGATCCGCTGAGGAAACAGGTGGTAC</td>
<td>BamHI</td>
</tr>
<tr>
<td>Bbp RP 3' R</td>
<td>GGGGAAGCTTTTTTACTGCTGAATCACCATC</td>
<td>HindIII</td>
</tr>
<tr>
<td>Sab RP 5' F</td>
<td>GGGGAAGCTCATGTTGGTCAAACCTTGAAGG ATG</td>
<td>SacI</td>
</tr>
<tr>
<td>Sab RP 3' R</td>
<td>GGGGAAGCTCATGTTGGTCAAACCTTGAAGG ATG</td>
<td>PstI</td>
</tr>
<tr>
<td>Dcm RP 5' F</td>
<td>GGGCCATGGCTGATTCTGTAAGTTCTAATG</td>
<td>NcoI</td>
</tr>
<tr>
<td>Dcm RP 3' R</td>
<td>GGGCTGCAAGTTATTTACAGGATATAATATTC TTTACCTACATTAC</td>
<td>PstI</td>
</tr>
<tr>
<td>Sab probe 5' F</td>
<td>ATGTGGTCAAACCTTGAAGGATGAAG</td>
<td></td>
</tr>
<tr>
<td>Sab probe 3' R</td>
<td>CTTATTTCCAAACCCAACCAATTACAC</td>
<td></td>
</tr>
<tr>
<td>Dcm probe 5' F</td>
<td>CAATTATATAAAACGAAGCAGGTAATATAGT</td>
<td></td>
</tr>
<tr>
<td>Dcm probe 3' R</td>
<td>TATTTTCAGAGTATAATTTCTTTACCTAC</td>
<td></td>
</tr>
<tr>
<td>SauUSI probe 5' F</td>
<td>TTATCTTTGGCTTTATCCCTAATAGC</td>
<td></td>
</tr>
<tr>
<td>SauUSI probe 3' R</td>
<td>ATGAAAAAGTTTCTTTCAACAATTATAGT</td>
<td></td>
</tr>
<tr>
<td>SdrE 5' F</td>
<td>CGGTGCTGAAAAACACTAGTACAG</td>
<td></td>
</tr>
<tr>
<td>SdrE 3' R</td>
<td>CTGTAGGGTGTATATTCTACAACAC</td>
<td></td>
</tr>
<tr>
<td>Primer</td>
<td>Sequence (5’-3’)</td>
<td>5’ Modifications</td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>IsdA △A</td>
<td>GGGG ACAAGTTTTTGTCACAAAAAAGCAGGCTTTGGCAGTGTTTTTACTAATAATTTTTC</td>
<td>attB1</td>
</tr>
<tr>
<td>IsdA △B</td>
<td>CATGTTGTTTTTCCTCCTAAGGATACAAA</td>
<td></td>
</tr>
<tr>
<td>IsdA △C</td>
<td>CATGTTGTTTTTCCTCCTAAGGATACAAAGAATCTAAATAATCATAATCACACTC</td>
<td></td>
</tr>
<tr>
<td>IsdA △D</td>
<td>GGGG ACCACTTTGTACAAAGAAGGCTGGGTGTTTTAATTGTCCAGAAGTAG</td>
<td>attB2</td>
</tr>
<tr>
<td>IsdB △A</td>
<td>GGGG ACAAGTTTTGTACAAAAAAGCAGGCTTTGCTAAACAGCAAGAACAAGATGATTTC</td>
<td>attB1</td>
</tr>
<tr>
<td>IsdB △B</td>
<td>TTAATCGTCCTTTATATTATTAATTATTAAC</td>
<td></td>
</tr>
<tr>
<td>IsdB △C</td>
<td>ATTAATCGTCCTTTATATTATTAATATTAAATTAATAATCTCCTAATTGTATA</td>
<td></td>
</tr>
<tr>
<td>IsdB △D</td>
<td>GGGG ACCACTTTGTACAAAAAAGGCTGGGTCAACAGTCAAAACAGCAACACACAGGCT</td>
<td>attB2</td>
</tr>
<tr>
<td>ClfB △A</td>
<td>GGGG ACAAGTTTTGTACAAAAAAGCAGGCTACTATTCAACTCAGGTAGTATGGA</td>
<td>attB1</td>
</tr>
<tr>
<td>ClfB △B</td>
<td>CAAAAATATTACTCCATTTCAATTCTA</td>
<td></td>
</tr>
<tr>
<td>ClfB △C</td>
<td>GAAATTGAAATGGAGTAATTTTTGTAAATACTTTTTAGGCCGAATATTTG</td>
<td></td>
</tr>
<tr>
<td>ClfB △D</td>
<td>GGGG ACCACTTTGTACAAAAAAGGCTGGGTAAATTGTGCCATATCCTCCCATAGAG</td>
<td>attB2</td>
</tr>
<tr>
<td>Bbp △A</td>
<td>GGGG ACAAGTTTTGTACAAAAAAGCAGGCTACAAGTGAGTACAAAGGCT</td>
<td>attB1</td>
</tr>
<tr>
<td>Bbp △B</td>
<td>CATATAAATCTCCTTTAAAAATATC</td>
<td></td>
</tr>
<tr>
<td>Primer</td>
<td>Sequence (5'-3')</td>
<td>5' Modifications</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Bbp ΔC</td>
<td>ATTTTAAAGGAGATTTATATGTAATAACAATA TGACCCAGGTC</td>
<td></td>
</tr>
<tr>
<td>Bbp ΔD</td>
<td><strong>GGGG ACCACTTTGTACAAGAAAGCTGGGT</strong> ATATCATACATATAATTGATG</td>
<td><strong>attB2</strong></td>
</tr>
<tr>
<td>Sab2370c ΔA</td>
<td><strong>GGGG ACAAGTTTGTACAAAAAAGCAGGCT</strong> ATGAATGACTATCGTTGGATTT</td>
<td><strong>attBI</strong></td>
</tr>
<tr>
<td>Sab2370c ΔB</td>
<td>CATTTCTATAACAACCACCCCTTGG</td>
<td></td>
</tr>
<tr>
<td>Sab2370c ΔC</td>
<td>AGGATGGTGTGGTATAGAATG TGACACATATGTACGTGGGAG</td>
<td></td>
</tr>
<tr>
<td>Sab2370c ΔD</td>
<td><strong>GGGG ACCACTTTGTACAAGAAAGCTGGGT</strong> AAAGCTTTCTTTAATTACTTTAAATGTG</td>
<td><strong>attB2</strong></td>
</tr>
<tr>
<td>Dcm ΔA</td>
<td><strong>GGGG ACAAGTTTGTACAAGAAAGCAGGCT</strong> AGACCTATTAGAGAAATGTATATTCC</td>
<td><strong>attBI</strong></td>
</tr>
<tr>
<td>Dcm ΔB</td>
<td>CATGTACATCTACCTACTTTCAT</td>
<td></td>
</tr>
<tr>
<td>Dcm ΔC</td>
<td>GAAAGTGGTTATGTACATGTAAATTTCC CTCTGTGGAATTC</td>
<td></td>
</tr>
<tr>
<td>Dcm ΔD</td>
<td><strong>GGGG ACCACTTTGTACAAGAAAGCAGGCT</strong> AATGTATTTGAACAAATGAATTTAATTG</td>
<td><strong>attB2</strong></td>
</tr>
<tr>
<td>Dcm out 5' F</td>
<td>AACAAAATTAAAAATTTTACTGATAATAATCATG</td>
<td></td>
</tr>
<tr>
<td>Dcm out 3' R</td>
<td>TGCCATATTGTACAAAAAACTTGG</td>
<td></td>
</tr>
</tbody>
</table>

*a* Restriction sites used for cloning are underlined  
*b* **att** sites for gateway cloning are underlined and nucleotides are indicted in red  
RP, Recombinant protein  
Δ, primers used for creating deletion mutations
typically carried out in 25 μl or 50 μl volumes using 1 U of the high fidelity Phusion™ Hot Start DNA polymerase in Phusion HF buffer (Finnzymes). Primers and dNTPs (Bioline) were used at final concentrations of 200 μM each.

The PCR cycling conditions for Phusion polymerase are as follows. Initial denaturation was carried out at 98 °C for 30 sec followed by 30 cycles of denaturation for 10 sec at 98 °C, 30 sec annealing (temperature dependent on primer used) and extension at 72 °C, allowing 30 sec/kb for genomic DNA and 15 sec/kb for plasmid DNA. A final extension step was carried out at 72 °C for 5 min.

2.6.4 Cloning using the Zero Blunt TOPO PCR Cloning kit.

Blunt ended fragments were cloned into TOPO vector using the Zero Blunt TOPO PCR Cloning kit according to the manufacturer’s instructions and subsequently sequenced using MyGATC (http://www.gate-biotech.com/en/sequencing).

2.6.5 Construction of pKOR1 derivatives for allele exchange

Primers were designed to amplify, from RF122 genomic DNA, fragments ~500 bp upstream (Primers ΔA and ΔB) and downstream (Primers ΔC and ΔB) of the genes of interest. At the extreme 5' and 3' ends, primers ΔA and ΔD had attB1 and attB2 sites, respectively, incorporated to allow recombination into pKOR1 by ‘gateway’ cloning. Primer ΔC had a 5' extension complementary to the 3' end of the upstream fragment which facilitates base pairing of the 3' end of the upstream fragment with the 5' end of the downstream fragment and production of a single amplimer by overlap primer PCR. 50 ng of each product were mixed for use as template in a subsequent PCR reaction. The resulting amplimer, represented diagrammatically in Figure 2.1, constitutes the upstream and downstream fragments flanked by attB1 and attB2, at the extreme 5' and 3' ends, respectively, whilst eliminating the gene of interest. The amplimer was purified and 300 ng of it was used in a BP clonase II reaction as recommended by the supplier (Invitrogen) with pKOR1 (200ng) made up to a final volume of 8 μl in TE buffer, pH8. pKOR1, which will be discussed below, is a temperature-sensitive plasmid that replicates at 28 °C whereas an upshift to a non-permissive temperature will select for cells where the plasmid
is integrated into the chromosome. The salient features of pKOR1 are represented diagrammatically in Figure 2.2.

The BP clonase reaction was incubated at 25 °C for 3hr. 5 µl of reaction mix was transformed, as described in section 2.8.3, into XL-1 Blue, plated onto L-agar containing Amp and incubated at 37 °C overnight. Plasmids isolated from resulting colonies were analysed by PCR for the presence of the \textit{attB} PCR product comprising upstream and downstream sequences with an in-frame deletion of the gene of interest. Plasmid DNA containing the insert was electroporated into RF122t or RF122\textit{Δsab2370c}, plated onto TSA containing Cm and incubated at 28 °C. Plasmid DNA was isolated and subsequently electroporated into the target strain (RF122) in an identical manner.

The pKOR1 \textit{E. coli}/\textit{S. aureus} shuttle vector contains a number of features which allow counter-selection for desired mutations. The ‘Gateway’ cloning system is used which avoids the use of restriction enzymes and ligases. \textit{attP} sequences, located on the recombination cassette, recombine with \textit{attB} sequences of DNA inserts in the presence of bacteriophage lambda integrase and \textit{E. coli} integration host factor (IHF). Another component of the recombination cassette \textit{ccdB}, encodes an \textit{E. coli} gyrase inhibitor protein which inhibits growth of cells transformed with non-recombinant plasmids. This results in enrichment of plasmids carrying the cloned mutational fragment. Plasmid transformation of staphylococci and growth at 43 °C, a non-permissive condition for pKOR1 replication, selects for homologous recombination and pKOR1 integration into the bacterial chromosome. SecY transports signal peptide bearing precursor proteins across the cytoplasmic membrane, a process which is essential for bacterial growth and survival. Expression of \textit{secY} antisense RNA inhibits growth on agar plates. Chromosomal excision and loss of plasmid is achieved using anhydrotetracycline-mediated induction of pKOR1-encoded \textit{secY} antisense transcripts via the Pxyl/tetO promoter (Bae & Schneewind, 2006).

2.6.6 Allelic replacement mutagenesis

RF122 cells containing either pKOR1\textit{ΔisdA}, pKOR1\textit{ΔisdB}, pKOR1\textit{ΔclfB}, pKOR1\textit{Δbbp} or pKOR1\textit{Δsab2370c} were grown at 28 °C in TSB containing Cm and subsequently subcultured and incubated at 42 °C. (To check for temperature sensitivity serial dilutions were plated in duplicate on TSA agar containing Cm. One set of plates
Figure 2.1 Gateway cloning using pKOR1

Primers were designed to amplify fragments ~500 bp upstream (Primers A and B) and downstream (Primers C and B) of the gene(s) of interest from RF122 genomic DNA. At the extreme 5' and 3' ends primers A and D had attB1 and attB2 sites, respectively, incorporated to allow recombination into pKOR1 by 'gateway' cloning. Primer C has a 5' extension complementary to the 3' end of the upstream fragment which facilitates base pairing of the 3' end of the upstream fragment with the 5' end of the downstream fragment and production of a single amplimer by overlap primer PCR. The resulting amplimer constitutes the upstream and downstream fragments flanked by attB1 and attB2, at the extreme 5' end and 3' ends, respectively, whilst eliminating the gene of interest.
Figure 2.2 Map of pKOR1

The lambda recombination cassette comprises phage attachment sites attP1 and attP2, ccdB (encodes an E. coli gyrase inhibitor protein) and cat(-) (encodes a chloramphenicol acetyltransferase which is functional in Gram-negative bacteria). The antisense secY expression cassette, which comprises secY570 and tetR, facilitates tetracycline inducible secY expression. Other features include bla (encodes β-lactamase), ori(-) (ColEl plasmid replication origin), cat(+) (encodes a chloramphenicol acetyltransferase which is functional in Gram-positive bacteria) and repF (replication gene of pE194ts). Figure adapted from (Bae & Schneewind, 2006).
was incubated at 42 °C and the other at 28 °C. Plates growing at 42 °C had 10^3-fold less colonies growing on them than those incubated at the lower temperature.) 10 μl of culture was streaked onto TSA agar containing Cm and grown at 42 °C until large and small colonies were observed. Large colonies putatively containing integrants were grown at 28 °C without Cm to encourage plasmid replication and excision. Cultures were diluted and plated on TSA containing anhydrotetracycline (1 μg) until large and small colonies were observed. Large colonies were toothpicked onto TSA with and without Cm. Colonies that had lost Cm resistance were screened using the PCR colony screen (as described in section 2.6.7) for successful mutation. Mutations were validated by PCR and Western blotting.

2.6.7 PCR colony screen

Bacteria was resuspended in 40 μl autoclaved TE buffer (1 mM EDTA/10 mM Tris-HCl, pH 7.8) and boiled for 10 min at 100 °C. Cell debris was pelleted following centrifugation at 6000 r.p.m. for 10 min. 2 μl of the resulting supernatant was used as template in a PCR reaction (25 μl total volume).

2.7 Electrophoresis

2.7.1 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins for SDS-PAGE analysis were adjusted to specific final concentrations in PBS and diluted 2-fold in final sample buffer (10 % (v/v) glycerol, 5 % (v/v) β-mercaptoethanol, 3 % (w/v) SDS, 0.01 % bromophenol blue in 62.5 mM Tris-HCl, pH 6.8). 20 μl volumes were separated by SDS-PAGE (Laemmli, 1970) using 4.5 % stacking and 7.5-12.5 % separating acrylamide gels. Prestained protein molecular weight markers were purchased from Fermentas. Protein samples were electrophoresed at 130 V. After separation, proteins were either visualised using Coomassie blue stain or electroblotted onto PVDF membranes (Roche) for 1 h at 100 V using a wet transfer cell (BioRad) for detection.
2.7.2 Agarose gel electrophoresis

Gels containing 0.5-2 % agarose which was dissolved by boiling in TAE buffer (Invitrogen) and cooled, were cast in mini trays (Life Technologies). DNA samples in loading buffer containing an electrophoretic dye were pipetted into wells alongside DNA size markers (Bioline). Electrophoresis of samples was routinely performed at 90-110 V. Gels were bathed in ethidium bromide for 10 min, washed and viewed under UV light. Gel images were analysed using Alpha Imager™ software.

2.8 Transformation

2.8.1 Preparation and transformation of electrocompetent *S. aureus*

1 ml of an overnight culture was subcultured into 100 ml fresh TSB in a 250 ml non-baffled flask and grown to an OD$_{540\text{nm}}$ of 0.2. The cells were chilled on ice for 10 min. Following centrifugation at 3000 x g for 10 min at 4 °C, cells were washed twice each in 100 ml ice-cold filter-sterilised sucrose solution (0.5 M sucrose). Following further centrifugation using the parameters stated above, cells were sequentially washed in 20 ml, 10 ml, 1 ml and finally 500 μl of ice-cold filter-sterilised sucrose (0.5 M sucrose). Cells were chilled on ice for 20 min and either used fresh or snap-frozen with liquid nitrogen and stored at -70 °C.

Plasmids were introduced into *S. aureus* by electroporation (Augustin & Gotz, 1990). Electrocompetent cells (50 μl) were mixed with 5 μg of plasmid DNA in a 0.1 mm electrode gap cuvette. Electroporation was carried out at 100 Ω resistance, 25 μF capacitance and 25k V. Cells were incubated in TSB containing 500 mM sucrose for 1 h at 37 °C (or 1.5 h at 28 °C for temperature-sensitive plasmids) with shaking (200 r.p.m.) prior to plating on antibiotic-containing media. Transformants were screened by PCR, restriction digest and Western immunoblotting.

2.8.2 Transformation of *E. coli* cells with plasmid DNA

Chemically competent *E. coli* cells were prepared by the CaCl$_2$ method (Sambrook, 1989). Competent cells (200 μl) were incubated on ice for 15 min with
plasmid or DNA ligation reactions. The mixture was then incubated at 42 °C for 2 min followed by a further 2 min on ice. Next, 1 ml L-broth was added and the mixture was incubated with shaking for 1 h at 37 °C. Cells were plated on L-agar plates containing antibiotic and incubated for 16-24 h. \textit{E. coli} transformants were screened for the presence of recombinant plasmids using the rapid colony screening procedure developed by Le Gouill and Dery (Le Gouill & Dery, 1991) or PCR colony screen as described in section 2.6.7.

2.9 Expression and purification of His-tagged recombinant proteins by immobilized metal chelate affinity chromatography

Recombinant domains of \textit{S. aureus} surface-associated proteins were expressed from pQE30 with an N-terminal hexa-histidine (His\(_6\)) affinity tag to allow purification by nickel affinity chromatography. The pQE30 vector is designed with sequences located 5' to the MCS that encodes for 6 x His residues. The vector also contains an IPTG-inducible promoter for controlled expression of recombinant proteins. Cultures were grown to OD\(_{600}\) nm of 0.5-0.6 and then induced with 1 mM IPTG for 3 h at 37 °C. Cells were harvested by centrifugation and resuspended in 30 ml PBS containing EDTA protease inhibitors (Roche). Cells were lysed in a French pressure cell and the lysate was cleared by high-speed centrifugation at 17 000 \(x\) \(g\) for 20 min. DNase was added to the supernatant which was filtered through a 0.45 \(\mu\)m Sartorius filter.

A HiTrap\textsuperscript{TM} Chelating HP column (5 ml; GE Healthcare) was charged with 150 mM Ni\(^{2+}\) and equilibrated with binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9). The cleared lysate was applied to the column and the column was then washed with binding buffer until the A\(_{280}\) nm of the eluate was < 0.001. Bound protein was eluted in fractions with a continuous linear gradient of imidazole (5-100 mM; total volume of 100 ml) in 0.5 M NaCl, 20 mM Tris-HCl (pH 7.9). Elution was monitored by measuring the A\(_{280}\) nm of the eluate. 10 \(\mu\)l of peak fractions were separated on 10 % acrylamide gels and visualized by Coomassie blue staining. Fractions containing protein of the correct molecular weight were pooled and dialysed against PBS for 16 h at 4 °C. Protein concentrations were determined by measuring the absorbance at 280 nm using a Nanodrop\textsuperscript{TM} 1000 spectrophotometer and by using the BCA assay kit (Pierce).
2.10 Enzyme linked immunosorbent assays (ELISA)

2.10.1 Binding of recombinant IsdB to haemoglobin

Recombinant IsdB protein at 2 μM in PBS was coated onto microtitre plates (Nunc) in sodium carbonate buffer (pH 9.6) for 18 h at 4 °C. The plates were washed three times with PBS and blocked with 5% (w/v) skimmed milk powder (Marvel) in PBS for 2 h at 37 °C. To detect recombinant IsdB, the plate was incubated with various concentrations of human haemoglobin in 2% Marvel for 2 h at 37 °C. After washing three times with PBS, bound haemoglobin was detected with polyclonal rabbit anti-human haemoglobin IgG (Dako) and incubated for 1 h at 37 °C. After three washes with PBS, to detect bound antibody the plate was incubated with HRP-conjugated goat anti-rabbit IgG for 1 h at 37 °C. After three further washes with PBS, bound HRP-conjugated antibodies were detected by the addition of 100 μl of 3,3',5, 5'-tetramethylbenzidine (TMB) (0.1 mg/ml) prepared in 0.05 M phosphate citrate buffer (pH5.0) containing 0.006% (v/v) hydrogen peroxide. Plates were then developed for 10 min. The reaction was stopped by the addition of 50 μl of 2 M H₂SO₄ to each well. The absorbance at 450 nm was measured with an ELISA plate reader (Multiskan EX, Labsystems). ELISA-type binding graphs are graphs of individual experiments that are representative of three independent experiments. Each plot in such graphs represents the average of triplicate wells.

2.10.2 Binding of recombinant ClfB N2N3 to fibrinogen.

Fibrinogen (100 μl) at 20 μg/ml was added to 96-well flat bottomed plates (Sarstedt) and incubated for 16 h at 4 °C. Wells were washed three times with PBS and incubated at 37 °C for 2 h with 5 % (w/v) Marvel/PBS buffer to block non-specific interactions. Wells were washed three times with PBS and serial dilutions of recombinant ClfB N2N3 were added to wells. Following incubation for 2 h at 37 °C, wells were washed three times with PBS. Anti-His-horseradish peroxidase (Roche 1:500) in 1 % (w/v) Marvel/PBS buffer was used to detect bound ClfB N2N3. Following incubation for 1 hr at room temperature, plates were washed with PBS. After washing, 100 μl of a chromogenic substrate (1 mg/ml tetramethylbenzidine and 0.006 % H₂O₂ in 0.05 M phosphate citrate buffer pH 5.0) was added to each well and plates were developed for 10-
15 min. The reaction was stopped by the addition of 2 M H$_2$SO$_4$ (50 µl/well), and the absorbance at 450 nm was read in an ELISA plate reader (Labsystems).

### 2.11 Molecular modelling of ClfBN2N3.

The Protein homology recognition engine (Phyre) service of the 3D-PSSM website (http://www.sbg.bio.ic.ac.uk/phyre/) provides a web-based tool which models the structure of a sequence based on the known crystal structures of similar proteins. The resulting hits are scored based on the estimated precision of the prediction. The X-ray crystal structure of the A domain of ClfB is known (Ganesh et al., 2011). The pdb file supplied by the Phyre service was downloaded and was viewed using the user-sponsored molecular visualization UCSF Chimera software (http://www.cgl.ucsf.edu/chimera/).

### 2.12 Southern hybridisation

#### 2.12.1 Southern dot hybridisation using dcm-, sab2370c- and sauUSI-specific probes

DIG-labelled probes were synthesised by PCR as described in section 2.6.3. Purified genomic DNA for analysis was quantified using the Qubit® 2.0 Fluorometer (Life Technologies) according to manufacturer’s instructions. 1000 ng of genomic DNA from laboratory strains or bovine isolates was spotted onto positively charged nylon membranes (Roche) and allowed to air-dry. Membranes were incubated for 5 min on blotting paper soaked in denaturation solution (1.5 M NaCl, 0.5 M NaOH), 5 min in neutralization solution (1.5 M NaCl, 1 M Tris-HCl, pH 7.4), and finally for 15 min on blotting paper soaked with 2x SSC solution (300 mM NaCl, 30 mM tri-sodium citrate). DNA was fixed onto the membranes by incubation at 120 °C for 30 min. Membranes were incubated for 2 h at 68 °C in pre-hybridization solution (5x SSC, 0.1 % w/v N-lauroylsarcosine, 0.02 % w/v SDS and 1x Blocking Reagent (Roche)). DIG-labelled probes were denatured by heating at 95 °C for 10 min, diluted in pre-hybridization solution and incubated with nylon membranes for 18 h at 68 °C. Following hybridization, the membranes were washed twice with 2x SSC / 0.1 % w/v SDS at room temperature followed by two washes with 0.5x SSC / 0.1 % w/v SDS at 68 °C for 20 min. Membranes were incubated for 30 min in blocking buffer at room temperature.
(comprising maleic acid buffer (100 mM maleic acid, 150 mM NaCl, pH 7.5) containing 1 X blocking reagent), and bound DIG-labelled probes were detected by incubation for 30 min with alkaline phosphatase-conjugated anti-DIG antibody (Roche) diluted 1:10,000 in blocking buffer. After washing twice with maleic acid buffer containing 0.3 % v/v Tween 20 and equilibrating for 5 minutes in buffer 3 (100 mM Tris-HCl, pH 9.5, 100 mM NaCl), membranes were incubated for 15 mins at 37 °C with the chemiluminescence substrate CSPD (Roche diagnostics, diluted 1:100 in buffer 3). Visualisation of “dots” was performed using the ImageQuant™ LAS 4000 biomolecular imager (GE healthcare BioSciences).

2.12.2 Probe construction

Fragments were amplified using PCR, as described in section 2.6.3, substituting regular dNTPs for a DIG-labelled dNTP mix (0.2 mM of each dNTP, 0.01 mM DIG-dUTP). Purified RF122 genomic DNA was used as template for amplifying dcm- and sab2370c-specific probes. Purified Newman genomic DNA was used as template for the sauUSI-specific probe. Amplified PCR product probes were visualised by standard agarose gel electrophoresis and purified using the Wizard purification kit (Promega) according to manufacturer’s instructions. 30 µl of each probe was added to 30 ml prehybridization solution.
Chapter 3

Surface proteins of bovine mastitis strains of *Staphylococcus aureus*
3.1 Introduction

*S. aureus* is a very important and versatile pathogen. It colonises and causes a variety of diseases in humans but it has also adapted to colonise various other hosts. It is capable of causing ruminant mastitis, skeletal infections of chickens and dermatitis and septicaemia in rabbits. This thesis focuses on bovine mastitis strains of *S. aureus*, which cause great economic losses in the dairy industry. One of the ways *S. aureus* has adapted to the bovine host is by modification of its genetic content. The genome organisation among different staphylococci is strongly conserved. In comparison to other pathogens, large scale recombination events play a limited role in speciation. Most *S. aureus* clones are specific to particular animal species which indicates host adaptation and restriction (Kapur *et al.*, 1995; Smith *et al.*, 2005). The genome can be categorised into three distinct regions, namely the core genome, the core variable genome and the accessory genome. The core genome, which comprises ~75% of total gene content, includes genes essential for cell maintenance, growth and survival. These genes are conserved in all strains.

The core variable genome, which constitutes ~10% of total gene content, includes genes which are predominantly conserved in strains sharing a similar evolutionary history. This region is characterised by a large number of species-specific surface proteins and virulence regulatory proteins which are involved in colonization and infection processes. Microbial surface components recognising adhesion matrix molecules (MSCRAMMs) are typically encoded by core or core variable regions (Lindsay *et al.*, 2006). Part of this thesis focuses on the MSCRAMMs of bovine mastitis strains of *S. aureus*. The profile of MSCRAMMs exhibited by each strain is governed by the requirements of the environmental niche inhabited. The accessory genome is characterised by mobile genetic elements (MGEs) which often encode virulence factors and resistance genes which are frequently exchanged within and between lineages (Lindsay *et al.*, 2006). This component is often the focus of bacterial niche adaptation and differences in virulence as it constitutes the most rapid and dramatic form of genetic adaptation (Hacker & Carniel, 2001). More long term adaptation is likely to be the result of a genome-wide combination of allelic diversification, lateral gene acquisition and loss of gene function (Ben Zakour *et al.*, 2008).

The majority of ruminant-associated sequence types belong to three major complexes, CC97, CC151 and CC133, which indicates a narrow distribution of ruminant-associated genotypes across the species tree (Figure 3.1, Guanine *et al.*, 2010). CC97
Figure 3.1 Neighbour-joining consensus tree based upon concatenated sequences of *S. aureus* MLST alleles.

Concatenated sequences of 33 *S. aureus* STs from ruminants and 97 representative STs from humans and other animals were used to construct a neighbor-joining consensus tree inferred from a bootstrap test (500 replicates). The major lineages associated with large (blue shading) and small (red shading) ruminants (cows and sheep/goats, respectively) are indicated. Black branches indicate STs of human, avian, or unknown host origin. Figure taken from (Guinane et al, 2010).
which is commonly isolated from cows is also detected among human and porcine hosts indicating broad host tropism (Feil et al., 2003; Smith et al., 2005; Guinane et al., 2008; Smyth et al., 2009). The evolutionary origin and molecular basis for host tropism remain unknown. There is evidence which suggests that the CC133 clone evolved as a result of human-to-ruminant host jump followed by adaptive genome diversification. Comparative whole-genome sequencing revealed the molecular evidence for host adaptation which included gene decay and diversification of proteins involved in host-pathogen interactions. Several novel MGEs encoding virulence proteins with attenuated or enhanced activity in ruminants were widely distributed in CC133 isolates, which suggests a key role in host-specific interactions. Numerous population genetic studies identified the existence of genotypes of *S. aureus* that are associated with cows, sheep and goats but are rarely isolated from humans, suggesting their specialization for ruminant hosts (Kapur et al., 1995; Fitzgerald et al., 1997; Jorgensen et al., 2005; Smyth et al., 2009). CC151 strains are predominantly associated with cows and have not been detected among humans previously (Guinane et al., 2008). The broad diversity of human lineages and narrow distribution of ruminant-associated lineages imply that *S. aureus* is predominantly human-adapted and has co-evolved with ruminant hosts. The capacity to colonise and infect ruminant host species is most likely the result of a small number of host jumps from *S. aureus* strains of human origin which were followed by genetic adaptation to ruminants through a combination of foreign DNA acquisition and gene decay, similar to the human-to-poultry host switch of the successful human CC5 lineage. Distinct bovine groups are often interspersed among human clonal groups (Herron-Olson et al., 2007; Lowder et al., 2009; Guinane et al., 2010).

RF122, which belongs to ST151, is a common bovine mastitis-isolated *S. aureus* clone with worldwide distribution which was originally isolated in Ireland (Fitzgerald et al., 2001; Herron-Olson et al., 2007). Gene decay, which is associated with the transition to intracellular lifestyle, was apparent in dozens of RF122 genes, including several well-characterised virulence factors which are conserved among human isolates. Certain genomic features distinguished clones of highly successful bovine-associated *S. aureus* from those that infect humans or are infrequently recovered from other bovine sources. The RF122 genome encodes several genomic elements which were not previously identified in *S. aureus*, including homologs of virulence factors from other Gram-positive pathogens. In comparison to human strains, allelic variation is high among virulence and
surface-associated genes involved in host colonisation, toxin production, iron metabolism, antibiotic resistance and gene regulation. Of note, well characterised virulence factors, including protein A and clumping factor A exist in the RF122 genome as pseudogenes (Herron-Olson et al., 2007). Pseudogenes are dysfunctional derivatives of genes that have lost their protein-coding ability or are otherwise no longer expressed in the cell (Vanin, 1985). Commonly, pseudogenes occur due to introduction of premature stop codons or frameshifts resulting from point mutations. The gene may simply not be transcribed or has lost the ability to encode RNA. The term, coined in 1977 by Jacq, et al., comprises the prefix pseudo, meaning false, and the root gene (Jacq et al., 1977).

The 2,742,531 base pair genome of RF122 contains a total of 2,589 open reading frames including 33 genes which are not found in other staphylococci but which encode products homologous to proteins from organisms other than \textit{S. aureus} or \textit{S. aureus}-associated phages. Eighty predicted proteins have no significant homology to any proteins in public databases, the majority of which are encoded by MGEs. The remainder of the genome closely resembles fully sequenced \textit{S. aureus} isolates from humans with respect to gene content and organization with considerable genetic variation within MGEs. It was observed that MGE content is similar among highly successful bovine \textit{S. aureus} clones that are closely related by multilocus sequence typing (MLST) (Herron-Olson et al., 2007).

The interaction between staphylococcal surface proteins and the host extracellular matrix (ECM) is important for host colonisation. In this study, to investigate how mastitis strains have adapted to life in the bovine host and to identify potential candidates for a surface protein vaccine, the repertoire of surface proteins expressed by a diverse collection of mastitis strains was analysed. The genome of \textit{S. aureus} RF122 is fully sequenced and the sequences of 18 common surface proteins of RF122 were compared to well-characterised sequences of the human-isolated strains Newman, 8325-4 and 024, and were assessed in detail using a number of online tools. For comparison with most staphylococcal surface protein amino acid sequences, Newman was used as a reference strain.

To validate bioinformatic analysis and to determine the repertoire expressed by this and other mastitis strains, the expression of a number of surface proteins by ST151 strains and a panel of genetically distinct bovine isolates was assessed by Western
immunoblotting using antibodies specific for each protein. Proteins for further analysis were selected based on their potential usefulness as vaccine candidates.
3.2 Results

3.2.1 Assessing the status of surface proteins of *S. aureus* RF122

When this project began in 2008 RF122 was the only bovine strain of *S. aureus* to have its genome fully sequenced and annotated (Herron-Olson *et al.*, 2007). In order to identify surface protein vaccine candidates for bovine mastitis it was decided initially to assess the array of LPXTG-anchored proteins which decorate the surface of RF122. This was achieved by comparing the sequences of RF122 surface protein genes, using online tools (specified in Chapter 2 section 2.2.1), to sequences from well-characterised *S. aureus* strains of human origin. Gene sequences of RF122 were deemed to be either intact, pseudogenes or missing.

Genes were deemed to be intact on the basis of a number of attributes. The open reading frame should contain specific features such as a signal sequence, start codon and ribosome binding site. The C-terminus should comprise the "stop secretion" and sorting signal viz LPXTG, a hydrophobic domain and positively charged residues at the extreme C-terminus. The sequence between the LPXTG motif and the hydrophobic domain should be of a correct distance. An online program was used to predict whether or not the gene sequence was likely to be translated into a functional protein. Intact genes did not contain point mutations (resulting in stop codons) that would lead to truncation of the protein. A frameshift also causes premature truncation because a stop codon soon appears in the other "frame".

Genes were defined as pseudogenes if they contained a premature stop codon (or codons) which would render the protein truncated in an otherwise intact open reading frame. The position of the initial stop codon can have different consequences. As illustrated in Figure 3.2, if the initial stop codon occurs at point A (just after the signal sequence) no functional protein is expressed. If it occurs at either point B (just after the ligand-binding domain) or point C (within the wall/membrane spanning region) then it is a possible that a truncated N-terminal fragment could be expressed and secreted into the growth medium, as happens with the FnBPs of strain Newman (Grandmeier *et al.*, 2004). A gene was deemed to be missing if no sequence could be identified. On this basis the following surface protein genes were deemed to be missing from RF122; *pls, sdrD, fnbB*, *sasG* and *cna*. The *isdI* gene encoding a cytoplasmic heme-degrading monooxygenase was also missing.
Figure 3.2  Diagrammatic representation of the degree of protein truncation resulting from the introduction of premature stop codons at various points in the gene

The signal sequence (SS), ligand binding A domain, repeat region, wall/membrane spanning region (WM) and the LPXTG motif for the covalent anchorage of the protein to the cell wall peptido glycan are indicated. A, B and C show the possible positions of stop codons in the gene.
To validate the results observed in bioinformatic analysis, Western immunoblotting was used to demonstrate that genes encoding intact proteins were expressed and conversely that pseudogenes were not. Surface proteins were released from the cell wall with lysostaphin during the formation of stabilised protoplasts and the expression was detected using specific antibodies. RF122 belongs to ST151 as determined by MLST (Herron-Olson et al., 2007).

To investigate whether the surface protein expression profile of RF122 was unique to RF122/ST151 strains or was representative of other strains of bovine origin, 20 genetically distinct (defined by MLST) bovine S. aureus strains were selected for study and their expression profiles were analysed. Bovine strains were selected to include the three major clonal complexes found in bovine strains and isolates represented varied geographical distributions. Proteins that are ubiquitously expressed can be considered as vaccine candidates, although proteins expressed by the majority of strains could also be potentially considered. In each experiment, Newman (or P1 – FnBPs) and an isogenic mutant deficient in the particular protein being tested were included as positive and negative controls, respectively.

The interaction between staphylococcal surface proteins and the host ECM is important for host colonisation. A number of genes potentially encoding surface proteins were retained by RF122. These genes are likely to be important for life in the mammary gland. Retention or acquisition of such genes may indicate how this strain has adapted to this environment. Bacterial adhesion to specific ligands was used to analyse the phenotypic characteristics of some retained genes.

3.2.2 Surface-exposed iron-regulated surface determinant (Isd) proteins

3.2.2.1 IsdA

The Isd system contains a number of proteins including wall-anchored proteins, a lipoprotein, membrane transport proteins and heme-degrading monooxygenases located in the cytosol (See Figure 1.6 in Chapter 1). The *isdA* and *isdB* genes from RF122 are annotated incorrectly in the GenBank database with *isdA* called *isdB* and vice versa (gene locus tags and accession numbers are listed in Table 3.1). IsdA is annotated as an iron-regulated cell wall-anchored protein in the NCBI GenBank database. The *isdA* gene from
RF122 was deemed to be intact (Table 3.1). IsdA from *S. aureus* strains Newman and RF122 both comprise 350 amino acid residues and share 97.7% amino acid sequence identity. The coordinates for the NEAT 1 domain of IsdA are indicated in Figure 3.3A. The IsdA NEAT domain of RF122 is highly conserved compared to Newman (2 substitutions of functionally similar amino acids). Therefore it is reasonable to predict that IsdA is functional in RF122.

Surface proteins were released from the cell wall with lysostaphin and samples were probed with antibodies specific for IsdA (Figure 3.3B). Newman expressed an immunoreactive protein of 38 kDa. This band was missing in the Newman *isdA* mutant. RF122 and strains of ST151 expressed IsdA seemingly at the same level as Newman (Figure 3.3B(i)). The different bovine strains each expressed IsdA of a similar size (Figure 3.3B(ii)). This suggests that this protein is likely to be conserved in all bovine strains.

Clarke (2004) stated that IsdA expressed by SH1000 (A functional *rsbU* derivative of 8325-4) was responsible for adhesion to fibronectin (Clarke et al., 2004). Figure 3.3C(i) shows neither RPMl grown-RF122 nor RPMl grown-Newman, which express IsdA, adhere to fibronectin. P1 and an isogenic mutant lacking *fnbA* and *fnbB* were included as positive and negative controls, respectively, for adhesion to Fn (P1 expresses FnBPs which will be discussed later in the text). IsdA expressed by Newman or RF122 grown in RPMl does not appear to promote adhesion to immobilised Fn.

Clarke et al (2009) showed that adhesion to cytokeratin 10 by SH1000 is enhanced by the expression of IsdA (Clarke et al., 2009). Figure 3.3C(ii) shows that both Newman wildtype and the isogenic mutant lacking *isdA* adhered to keratin with similar affinity in a dose-dependent and saturable manner. IsdA expressed by Newman or RF122 grown in RPMl does not appear to promote adhesion to immobilised CK10. ClfB is expressed by Newman at very high levels and would likely “dominate” any low expression by IsdA. RF122 expresses ClfB at a very low level in TSB (see section 3.2.3.2).

### 3.2.2.2 IsdB

The *isdB* gene from RF122 is annotated in the GenBank database as a cell surface transferrin-binding protein and, in this study, was deemed to be intact (Table 3.1). The
<table>
<thead>
<tr>
<th>Protein</th>
<th>Locus tag</th>
<th>Gene coordinates</th>
<th>Location</th>
<th>No of conserved substitutions</th>
<th>Variant identities</th>
<th>Deleted conserved substitutions</th>
<th>% identity to Newman</th>
</tr>
</thead>
<tbody>
<tr>
<td>IsdA</td>
<td>SAB0995c</td>
<td>1073523-1075412</td>
<td>Cell surface</td>
<td>350</td>
<td>98</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>IsdB</td>
<td>SAB0994</td>
<td>1075615-1076667</td>
<td>Cell surface</td>
<td>629</td>
<td>94</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IsdC</td>
<td>SAB0999</td>
<td>1076876-1077559</td>
<td>Membrane-bound</td>
<td>227</td>
<td>99</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IsdD</td>
<td>SAB0996</td>
<td>1077559-1078635</td>
<td>Membrane-bound</td>
<td>358</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IsdE</td>
<td>SAB0997</td>
<td>1078632-1079510</td>
<td>Membrane-bound</td>
<td>292</td>
<td>96</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IsdG</td>
<td>SAB0998</td>
<td>1079520-1080488</td>
<td>Membrane-bound</td>
<td>322</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ClfB</td>
<td>SAB0100</td>
<td>1081303-1081626</td>
<td>Membrane-bound</td>
<td>107</td>
<td>98</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SigD</td>
<td>SAB0101</td>
<td>1084630-1084915</td>
<td>Membrane-bound</td>
<td>861</td>
<td>89</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SigF</td>
<td>SAB0102</td>
<td>1098401-1098430</td>
<td>Membrane-bound</td>
<td>899</td>
<td>89.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SasH</td>
<td>SAB0073</td>
<td>2670733-2672628</td>
<td>Membrane-bound</td>
<td>241</td>
<td>98</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SasJ</td>
<td>SAB0251c</td>
<td>2672629-2673303</td>
<td>Membrane-bound</td>
<td>631</td>
<td>95</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SasK</td>
<td>SAB0023</td>
<td>31231-33349</td>
<td>Extracellular</td>
<td>772</td>
<td>95</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SasL</td>
<td>SAB0846</td>
<td>919205-919359</td>
<td>Extracellular</td>
<td>145</td>
<td>99</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
A

**Signal sequence**

<table>
<thead>
<tr>
<th>Newman_IddA</th>
<th>RF122_IddA</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTKHYLNSKYQSERSSAMKITYTASILLGLSVYGADSOQVNAATEATNANNQSEEQ</td>
<td>MTKHYLNSKYQSERSSAMKITYTASILLGLSVYGADSOQVNAATEATNANNQSEEQ</td>
</tr>
<tr>
<td>VSQATOPSOFQKQDSSEKSHMDYMQHKVQLNKYYFCVLNNASFWHEKEYFYIN</td>
<td>VSQATOPSOFQKQDSSEKSHMDYMQHKVQLNKYYFCVLNNASFWHEKEYFYIN</td>
</tr>
<tr>
<td>ANQQEltATVNNKAKADRTTIAVAAVAPKSLTQTVHIVPQINNYHRTEFEKAI</td>
<td>ANQQEltATVNNKAKADRTTIAVAAVAPKSLTQTVHIVPQINNYHRTEFEKAI</td>
</tr>
<tr>
<td>PTLDAAKPNVKVPQKPAQQTPTEQTPQVPKVEKTPVTTSSKVEDHSKTVST</td>
<td>PTLDAAKPNVKVPQKPAQQTPTEQTPQVPKVEKTPVTTSSKVEDHSKTVST</td>
</tr>
<tr>
<td>DTDKQGQHTVTQTVTATQTAQEOFQKVQTVPKDAEAKSESNQUNSDNQSKQTNKTVK</td>
<td>DTDKQGQHTVTQTVTATQTAQEOFQKVQTVPKDAEAKSESNQUNSDNQSKQTNKTVK</td>
</tr>
<tr>
<td>NEPKQASKAELEPTGTSVNIFSTVAFATALLGSLSLLLFQKRESK</td>
<td>NEPKQASKAELEPTGTSVNIFSTVAFATALLGSLSLLLFQKRESK</td>
</tr>
</tbody>
</table>

---

**B**

(i) [Image of gel electrophoresis](image)

(ii) [Image of gel electrophoresis](image)

---

**C**

(i) RPMI-grown stationary phase

(ii) RPMI-grown stationary phase
Figure 3.3  A. Amino acid sequence alignment of IsdA from *S. aureus* strains Newman and RF122

Amino acid sequence alignment of IsdA from *S. aureus* strains Newman and RF122 showing variant residues. * indicates an identical residue while : (highlighted in magenta) and . (highlighted in green) indicate conserved and semi-conserved substitutions, respectively. The signal sequence is indicated by blue arrows, the cleavage site of which is highlighted in grey. The predicted boundaries of the NEAT 1 domain are indicated by black arrows. The LPKTG motif is highlighted in yellow. Hydrophobic residues are indicated by a red line and the positively-charged tail is indicated.

B. Expression of IsdA by (i) ST151 and (ii) different bovine *S. aureus* strains. *S. aureus* strains were grown in RPMI to the stationary phase. Cell-wall proteins were solubilised with lysostaphin, separated on SDS PAGE gels, electroblotted onto PVDF membranes and probed for the expression of IsdA using rabbit anti-isdA serum followed by HRP-conjugated protein A. *S. aureus* strain Newman and an isogenic mutant lacking *isdA* were included as positive and negative controls, respectively.

C. Bacterial adhesion to fibronectin and cytokeratin 10. Fibronectin (i) or cytokeratin 10 (ii) was immobilised on 96-well plates and incubated with bacteria grown to stationary phase in RPMI. Each graph is representative of experiments carried out in triplicate with similar results.
IsdB amino acid sequences from *S. aureus* Newman and RF122 share 94.42% sequence identity. As both the NEAT 1 and NEAT 2 ligand binding domains are intact it is likely that this protein is functional.

In the NEAT one domain there are two conserved substitutions and one variant residue (Figure 3.4). Although these residues have not been shown to be required for ligand binding they are localised close to residues comprising an aromatic loop that is crucial for interaction with human α-chain of haemoglobin (M. Zapotoczna, unpublished data). A 15 residue deletion was observed (residues 78-92) outside the NEAT 1 domain which could indirectly affect ligand binding by altering the folding of the protein.

In Western immunoblotting experiments when probed with anti-IsdB antibodies, Newman expressed a protein with an apparent molecular weight of 70kDa. This band was missing in the isogenic *isdB* mutant. Each ST151 strain produced a band of a similar size to RF122 with the exception of RF80 which expressed an immunoreactive band of a slightly smaller size (Figure 3.5A). This is likely to be due to an in-frame deletion. This could be confirmed by sequencing the *isdB* gene. Each bovine strain tested expressed a protein that corresponded to IsdB (Figure 3.5B). No immunoreactive band was seen from ST97 in Figure 3.5B but one was detected following longer exposure (data not shown). Although IsdB is expressed at different levels by bovine strains it seems that this protein is conserved amongst this group of organisms.

### 3.2.2.3 IsdH

IsdH proteins of Newman and RF122 have the potential to share 92.17% amino acid sequence identity. However, a point mutation at base 775 from a cytosine to a thymine (amino acid residue 259) results in a premature stop codon (Figure 3.6A(i)) which is likely to result in secretion of a truncated protein into the medium (Figure 3.6A(ii)). When the culture supernatant of RF122 was tested for the presence of secreted IsdH protein, a 28 kDa band was detected with anti-IsdH antibodies (data not shown). This corresponds to a truncated IsdH protein which contains the NEAT 1 domain. When probed with anti-IsdH antibodies a band corresponding to IsdH was detected in Newman and was missing in the isogenic mutant. As mentioned above the *isdH* gene exists as a pseudogene in RF122 and consistent with this an immunoreactive band was not detected in the cell wall fraction. Other ST151 strains did not express IsdH (Figure 3.6B(i)).
Interestingly, all of the other bovine strains, with the exception of one from ST350, expressed IsdH which shows that the gene is intact in the majority of bovine strains (Figure 3.6B(ii)). As shown in Figure 3.16, which will be discussed later, ST350 and ST151 are not phylogenetically closely related, which suggests an adaptive loss of IsdH function. It is not known if the ST350 strain carries the same or different mutations.

3.2.2.4 Other \textit{isd} proteins

Although the focus of this thesis is on surface proteins it is worth noting that other genes of the \textit{isd} system have been retained in the RF122 genome. RF122 possesses intact genes that encode the following Isd proteins (details of which are summarised in Table 3.1); the wall-associated SrtB-anchored IsdC, the membrane transport proteins IsdD, IsdE and IsdF and the heme monooxygenase IsdG. The gene encoding a second heme monooxygenase Isdl was found to be missing.

3.2.3 SD repeat family proteins

3.2.3.1 ClfA

The \textit{clfA} gene from RF122 is annotated as encoding a truncated clumping factor. When aligned with the \textit{clfA} sequence from \textit{S. aureus} strain Newman, the 5' part of the sequence provided in the GenBank database for \textit{clfA} from RF122 appears to be missing (Figure 3.7 A(i)). However, examination of DNA 5' to the indicated ORF shows that this region is actually present but a frameshift has occurred at bp 36 (Figure 3.7 A (ii)). The first stop codon 3' to the frameshift occurs after the 20\textsuperscript{th} codon just before the end of the region encoding the signal sequence (Figure 3.7A (iii)).

An immunoreactive band corresponding to ClfA was detectable in strain Newman and was missing in the isogenic mutant. None of the ST151 strains produced a band corresponding to ClfA (Figure 3.7B(i)) which suggests that, like RF122, these strains carry \textit{clfA} as a pseudogene. This could be confirmed by sequencing. However, the majority of other bovine strains tested expressed ClfA detectably, with the exception of ST350, ST479, ST699 and ST709 (Figure 3.7B(ii)).
Figure 3.4 Amino acid sequence alignment of IsdB from *S. aureus* strains Newman and RF122

Amino acid sequence alignment of IsdB from *S. aureus* strains Newman and RF122 showing variants residues. * indicates an identical residue while : (highlighted in magenta) and . (highlighted in green) indicate conserved and semi-conserved substitutions, respectively. - indicates the position of a residue that only occurs in the sequence to which it is compared to. Variant residues are highlighted in red. The NEAT 1 and NEAT 2 domains are indicated by arrows. The YSIRK and LPQTG motifs are highlighted in turquoise and yellow, respectively. The signal sequence is indicated by blue arrows, the cleavage site of which is highlighted in grey. C-terminal hydrophobic residues are indicated by a red line and the positively charged tail is indicated.
Figure 3.5 Expression of IsdB from (A) ST151 and (B) a panel of genetically distinct bovine *S. aureus* strains

*S. aureus* strains were grown in RPMI to the stationary phase. Cell wall associated proteins were solubilised with lysostaphin, separated on SDS PAGE gels, electroblotted onto PVDF membranes and probed for the expression of IsdB with rabbit anti-IsdB serum followed by HRP-conjugated protein A. *S. aureus* strain Newman and an isogenic *isdB* mutant were included as positive (+) and negative (-) controls, respectively.
Newman_isdH
RF122_isdH

AAATGACATTACAGTCTGCAATCTAAATTTTAGTTCTATCTATTTG
780
AAATGACATTACAGTCTGCAATCTAAATTTTAGTTCTATCTATTTG
780

Newman_isdH
RF122_isdH

ACATCAAGAGTCAACAATGCTATATCTAATCAACCCGAGCAGGACAAAATGATCAGCT
840
ACATCAAGAGTCAACAATGCTATATCTAATCAACCCGAGCAGGACAAAATGATCAGCT
840

Signal sequence

**N**

Newman_isdH
RF122_isdH

MKRSSRPGQJKLDRKQITPESLSYISQTEKNIAMSQHSLQKESDKDL
60
MKRSSRPGQJKLDRKQITPESLSYISQTEKNIAMSQHSLQKESDKDL
60

Newman_isdH
RF122_isdH

TOPRDRNONQGATQFAPNATAYPADESLKDAIKPAKENEDIGFREGQTVLLL
120
TOPRDRNONQGATQFAPNATAYPADESLKDAIKPAKENEDIGFREGQTVLLL
120

Newman_isdH
RF122_isdH

NETQYHFQSIKDAPVYTTKKEVEVLNATSWKKEYENQKLQPVLYSVYSYPFE
180
NETQYHFQSIKDAPVYTTKKEVEVLNATSWKKEYENQKLQPVLYSVYSYPFE
180

Newman_isdH
RF122_isdH

DYATRFVSDQGKLKVSTTLDYNTQDVLKAPFKYIYDPSLVKSDTNK
240
DYATRFVSDQGKLKVSTTLDYNTQDVLKAPFKYIYDPSLVKSDTNK
240

Newman_isdH
RF122_isdH

NDQSSERASQRTNTSTNTSNTSTNNANQCPQATNNMSQPAFQKSTNWADQASQPAHER
300
NDQSSERASQRTNTSTNTSNTSTNNANQCPQATNNMSQPAFQKSTNWADQASQPAHER
300

Newman_isdH
RF122_isdH

NSGNTNDKTNSQSNQSDVNQFYPADESLTQLAIKNPAIKPENEDIKADWPRIDFTMDK
359
NSGNTNDKTNSQSNQSDVNQFYPADESLTQLAIKNPAIKPENEDIKADWPRIDFTMDK
359

Newman_isdH
RF122_isdH

GQAYRIKPSNPSNPSKIVATTYKAEKPEKLSKQLYADNKQSYLQKESDKDL
400
GQAYRIKPSNPSNPSKIVATTYKAEKPEKLSKQLYADNKQSYLQKESDKDL
400

Newman_isdH
RF122_isdH

YAYIRFYVSNGPFEKTVVSIEYGENIHEDYDYTLMVFAQPITNNPDDYVDEETYNLQK
479
YAYIRFYVSNGPFEKTVVSIEYGENIHEDYDYTLMVFAQPITNNPDDYVDEETYNLQK
479

Newman_isdH
RF122_isdH

KRVVTKVSLKDPFKNNSRTLIFPYIPKAVYNAVYQNYNNTDKVTETENLTPT
540
KRVVTKVSLKDPFKNNSRTLIFPYIPKAVYNAVYQNYNNTDKVTETENLTPT
540

Newman_isdH
RF122_isdH

NDQLTDQEAHFWFNSNENKNSNNGDQEYFYFTPATLNGQKVVMKTDSDKYWKDLIVE
599
NDQLTDQEAHFWFNSNENKNSNNGDQEYFYFTPATLNGQKVVMKTDSDKYWKDLIVE
599

Newman_isdH
RF122_isdH

DKDVQHDVDHLSDMSGDYNFDKQLKADMTQIAKDTDRNVDKDAKN---KGVDNSVCTISSNVDTDKDS
775
DKDVQHDVDHLSDMSGDYNFDKQLKADMTQIAKDTDRNVDKDAKN---KGVDNSVCTISSNVDTDKDS
775

Newman_isdH
RF122_isdH

NKNKDKVIQLNDHADKNNKVNNTGTETNIDTMKYHPISTIKVTDKKTTEHLPSDIKTVDK
835
NKNKDKVIQLNDHADKNNKVNNTGTETNIDTMKYHPISTIKVTDKKTTEHLPSDIKTVDK
835

Newman_isdH
RF122_isdH

TVKTVKETGPTQSKNLSSMFLPKGETTSSQSSWAGYLYAGLMLAPIFRKEKSK
892
TVKTVKETGPTQSKNLSSMFLPKGETTSSQSSWAGYLYAGLMLAPIFRKEKSK
892
Figure 3.6 A. Amino acid sequence alignment of IsdH from *S. aureus* strains Newman and RF122

Amino acid sequence alignment of IsdH from *S. aureus* strains Newman and RF122 showing variant residues. * indicates an identical residue while : (highlighted in magenta) and . (highlighted in green) indicate conserved and semi-conserved substitutions, respectively. Variant residues are highlighted in red. (i) Shows a premature stop codon. (ii) shows the truncated IsdH. The NEAT1 domain as indicated by arrows. The signal sequence is indicated by blue arrows, the cleavage site of which is highlighted in grey. The YSIRK motif is highlighted in yellow.

B. Measurement of expression of IsdH by (i) ST151 and (ii) a panel of genetically distinct bovine *S. aureus* strains. *S. aureus* strains were grown in RPMI to the stationary phase. Cell-wall proteins were solubilised with lysostaphin, separated on SDS PAGE gels, electroblotted onto PVDF membranes and probed for the expression of IsdH with rabbit anti-IsdH serum followed by HRP-conjugated protein A. *S. aureus* strain Newman and an isogenic isdh mutant were included as positive (+) and negative (-) controls, respectively.
A

(i) Newman_clfA
RF122_clfA

ATGAGAAAAAGAAAAACACGCAATTCCGAAAAATACTGATTGGCGTGGCTTCAGTGCTT 60

GTCGCTTCAGTGCTT 15

(ii) Newman_clfA
RF122_clfA

ATGAGAAAAAGAAAAACACGCAATTCCGAAAAATACTGATTGGCGTGGCTTCAGTGCTT 60

GTCGCTTCAGTGCTT 15

(iii) Newman_clfA
RF122_clfA-

M KKKEKHAIRKKSIGVASVLV 21

MKKEKHAIRKNSLAWLQCL- 20

***********: . . *

B

(i) 176 kDa

119 kDa

(ii) 176 kDa

119 kDa

C
Figure 3.7  A. Amino acid sequence alignment of ClfA from *S. aureus* strains Newman and RF122

Amino acid sequence alignment of IsdH from *S. aureus* strains Newman and RF122 showing variants residues. * indicates an identical residue while : and . indicate conserved and semi-conserved substitutions, respectively. - indicates the position of a residue that only occurs in the sequence to which it is compared to. Variant residues are highlighted in red. (i) Shows the incorrect sequence provided on Pubmed. (ii) shows the correct 5' sequence. (iii) shows the protein remnant of ClfA.

B. Measurement of expression of ClfA by ST151 strains and a panel of genetically distinct bovine *S. aureus* strains

*S. aureus* strains were grown in TSB to the stationary phase. Cell-wall proteins were solubilised with lysostaphin, separated on SDS PAGE gels, electroblotted onto PVDF membranes and probed for the expression of ClfA with rabbit anti-ClfA serum and detected with HRP-conjugated protein A. *S. aureus* strain Newman and an isogenic mutant deficient in *clfA* were included as positive (+) and negative (-) controls, respectively.

C. Bacterial adhesion to fibrinogen by bovine *S. aureus* strains

Bovine fibrinogen was immobilised on 96-well plates and incubated with bacteria from the stationary phase of growth. *S. aureus* Newman and an isogenic mutant lacking *clfA* were included as positive (+) and negative (-) controls, respectively. The graphs show adhesion to bovine fibrinogen. Each graph is representative of experiments carried out in triplicate with similar results.
In some laboratory strains of *S. aureus*, ClfA is the only Fg-binding surface protein that is present on cells grown to stationary phase. Therefore it seemed reasonable to test bovine strains for adhesion to Fg in an attempt to attribute function to surface-expressed ClfA. All of the strains which expressed ClfA were able to adhere to bovine fibrinogen (ST126 adheres very weakly) apart from ST694 which did not adhere. Perhaps the ClfA protein of the ST694 strain is defective or is masked by expression of other surface proteins or polysaccharides. Molecular modelling of their ClfA proteins may elucidate why they have lost their ligand binding ability. Interestingly, ST709 does not express ClfA but adheres to Fg with levels comparable to the positive control which suggests that it expresses another Fg-binding protein in the stationary phase.

### 3.2.3.2 ClfB

The *clfB* gene of RF122 is intact (Table 3.1). The ClfB protein of *S. aureus* Newman is predicted to be composed of 913 amino acid residues whereas ClfB of RF122 consists of 861 residues. The N1, N2 and N3 subdomains of the ligand-binding A domain are indicated in Figure 3.8. The two proteins share 89.49% amino acid sequence identity with 96% sequence identity between the ligand-binding A domains. Residues Q235 and N526 have both been shown to be important for human fibrinogen and mouse cytokeratin 10 binding in Newman (Walsh *et al.*, 2008) and are conserved in RF122 (highlighted in red). Deletions in the proline rich and the SD repeat region which comprises the flexible stalk region were discernible (Figure 3.8). Figure 3.9A (i) shows that ClfB expressed by RF122 is smaller than that of Newman which is likely to be due to these deletions. ClfB was expressed weakly in comparison to ClfB expressed by Newman, whereas many of the other bovine strains expressed the protein at similar levels to Newman. Low immunoreactivity could be due to a lower affinity of the antibodies for the protein due to sequence variation. This warrants further investigation.

The other ST151 strains express a protein of a similar size to that of RF122 (Figure 3.9A(ii)). It is plausible to speculate that they have the same or similar deletions in their repeat regions. The expression of ClfB by all ST151 strains tested was weaker than that of Newman. ClfB was expressed by all of the bovine strains tested with the exception of strains from ST25, ST695, ST696 and ST697 (Figure 3.9A(ii)).
In lab strains Newman and 8325-4 ClfB is only found on the surface of exponential growth phase cells. In addition to its fibrinogen-binding properties (Ni Eidhin et al., 1998) ClfB also binds to cytokeratin 10 and is the only known \textit{S. aureus} protein to do so. Therefore adhesion to CK10 is an excellent phenotype for ClfB expression. It was reasonable to test bovine strains for adhesion to the ligand to correlate with ClfB detected by Western immunoblotting.

Some strains expressed an immunoreactive protein at similar levels to Newman but did not adhere to CK10. These strains included ST20, ST693 and ST694. Perhaps the ClfB protein in these strains is defective or it is masked by expression of other surface proteins or polysaccharides. ST71 and ST133 also expressed an immunoreactive protein, although at weaker levels than Newman, which may account for their lack of adhesion to CK10. Conversely, ST697 did not express ClfB detectably and ST709 expressed ClfB very weakly yet both adhered strongly to CK10 suggesting the presence of a second adhesin for the ligand. Alternately, this could be explained by a reduced antibody affinity due to allelic variation. This warrants further investigation. In the case of RF122 the failure to bind to CK10 is likely to be due to very low level of expression of the protein (Figure 3.9B).

3.2.3.3 SdrE/Bbp

In \textit{S. aureus} Newman the \textit{sdrC}, \textit{sdrD} and \textit{sdrE} genes occur in a tandem array and are closely linked. Some human strains lack one or more of these genes (8325 lacks \textit{sdrE}). The Bbp bone sialoprotein binding protein is an allelic variant of SdrE. RF122 carries an apparently intact \textit{bbp} gene (Table 3.1) and lacks \textit{sdrE}. The A domains of RF122 Bbp and Newman SdrE share 78\% amino acid sequence identity. In \textit{S. aureus} 024, Bbp comprises 1171 amino acid residues whereas Bbp from RF122 is slightly smaller, comprising 1113 residues due to shorter SD repeat region. They share 90.18\% sequence identity (93.4\% sequence identity in A domain) (Figure 3.10).

Initially only antibodies specific for the A domain of SdrE were available. These were used to probe cell wall extracts of RF122 and other ST151 strains. Immune cross-reactive protein could only be detected in RF122 and some ST151 strains when a higher concentration of antibody was used compared to that optimised for detection of SdrE in
Figure 3.8  Amino acid sequence alignment of ClfB from *S. aureus* strains Newman and RF122

Amino acid sequence alignment of ClfB from *S. aureus* strains Newman and RF122 showing variants residues. * indicates an identical residue while : (highlighted in magenta) and . (highlighted in green) indicate conserved and semi-conserved substitutions, respectively. - indicates the position of a residue that only occurs in the sequence to which it is compared to. Variant residues are highlighted in red. Residues known to be crucial for ligand binding are highlighted in red font. The N1, N2 and N3 subdomains are indicated. The signal sequence is indicated by blue arrows, the cleavage site of which is highlighted in grey. The YSIRR and LPETG motifs are highlighted in turquoise and yellow, respectively. The SLAVA motif, for recognition by the metalloprotease aureolysin, is highlighted in yellow.
Figure 3.9  

A. Expression of ClfB from (i) ST151 and (ii) a panel of genetically distinct bovine *S. aureus* strains

*S. aureus* strains were grown in TSB to the exponential phase. Cell-wall anchored proteins were solubilised with lysostaphin, separated on SDS PAGE gels, electroblotted onto PVDF membranes and probed for the expression of ClfB with rabbit anti-ClfB serum and detected with HRP-conjugated protein A. *S. aureus* strain Newman and an isogenic *clfB* mutant were included as positive (+) and negative (-) controls, respectively.

B. Bacterial adhesion to cytokeratin 10. Cytokeratin 10 was immobilised on 96-well plates and incubated with bacteria from the exponential phase of growth. Each graph is representative of experiments carried out in triplicate with similar results.
Figure 3.10  Amino acid sequence alignment of Bbp from *S. aureus* strains 024 and RF122

Amino acid sequence alignment of Bbp from *S. aureus* strains 024 and RF122 showing variants residues. * indicates an identical residue while : (highlighted in magenta) and . (highlighted in green) indicate conserved and semi-conserved substitutions, respectively. - indicates the position of a residue that only occurs in the sequence to which it is compared to. Variant residues are highlighted in red. The B repeats are indicated. The signal sequence is indicated by blue arrows, the cleavage site of which is highlighted in grey. The FSIRK and LPETG motifs are highlighted in turquoise and yellow, respectively.
strain Newman suggesting that they expressed Bbp not SdrE (Figure 3.11 A (i)). While the concentration of anti-SdrE optimised for detection of SdrE could not detect Bbp from RF122 or ST151 strains, it was sufficient to detect SdrE from the majority of genetically distinct strains (Figure 3.11A(ii)) which suggests that the majority of bovine strains tested here express SdrE not Bbp. The expression of Bbp by RF122 was confirmed using specific anti-Bbp antibodies (Figure 3.11 (iii)).

ST479, ST698 and ST350 did not express SdrE detectably. When primers directed against the sdrE region encoding the A domain were used, a fragment could be amplified from ST479 and ST698. These fragments were cloned into the TOPO vector and sequenced. Analysis showed that they possess DNA encoding the A domain of sdrE. A fragment could not be amplified from strain ST350 (Figure 3.11B) indicating that the gene is not present or it has undergone mutation in the region to which the primers are directed. This could be verified by more detailed analysis eg by Southern blotting.

3.2.3.3.1 Investigating the distribution of sdrE or bbp genes in a number of bovine strains by cloning and sequencing the A domains

In order to determine if the immunoreactive band detected with anti-SdrE serum was indeed SdrE or the variant Bbp, the part of the gene encoding the A domain from different strains was cloned into the TOPO vector and sequenced. The following STs which produced an immunoreactive band corresponding to SdrE were investigated; ST8, ST25, ST71, ST97, ST115, ST126, ST133 and ST352. As summarised in Table 3.2., the A domains of these STs were predicted to share over 95% amino acid sequence identity to the A domain of SdrE of strain Newman and less than 70% identity to Bbp from strain 024.

3.2.3.4 SdrC

The sdrC gene of RF122 is annotated in the GenBank database as a pseudogene. Analysis of the DNA sequence 5' to the sequence provided in the database revealed that the entire coding sequence is present but has suffered multiple point mutations leading to stop codons (Figure 3.12A), the first of which occurs at the 26th codon just before the end of the signal sequence (Figure 3.12B).
3.2.4 Fibronectin-binding proteins

As described in Chapter 1, one of the ways that *S. aureus* has adapted to its host is its ability to internalise into host epithelial cells. The major mechanism for this is provided by FnBPs A and B through their interaction with fibronectin. Because a premature stop codon in the *fnbA* gene of *S. aureus* Newman led to truncation of the protein at the end of the C domain (Grundmeier *et al.*, 2004) 8325-4 was selected as a reference strain for sequence comparison and analysis. RF122 possesses an intact *fnbA* gene (Table 3.1). Based on the divergence between FnBPA A domains of different strains, seven different isotype forms have been identified which share 66 to 76% amino acid sequence identity. Substantial divergence was observed here between 8325-4 (isotype I) and RF122 (Figure 3.13). As deduced by alignments RF122 FnBPA constitutes isotype VI (data not shown).

Out of the six high affinity Fn-binding sites in the ligand binding region of FnBPs, Fn-1, -4, -5, -9, -10 and -11 (Meenan *et al.*, 2007), binding sites 9, 10 and 11 from FnBPA of RF122 are intact with only a few conserved or semi-conserved substitutions when compared to FnBPA from 8325-4. There are several variant residues in Fn-binding site 1 and large parts of both binding sites 4 and 5 are missing.

To detect expression of FnBPA, it would be necessary to use antibodies specific for the specific A domain. The Fn binding domain is poorly immunogenic and antibodies to this region are not available. Another method of detecting FnBPs is to employ ligand affinity blotting using biotinylated fibronectin. P1 and an isogenic mutant deficient in *fnbA* and *fnbB* were used as positive and negative controls, respectively. None of the ST151 strains, including RF122 expressed detectable FnBPs using standard growth conditions (Figure 3.14A(i)).

To investigate if altered growth conditions might promote the expression of FnBPA the pH of the media was altered. The media was supplemented with glucose which, when metabolised, lowers the pH of the media (O'Neill *et al.*, 2008). Cultures of RF122 were grown in the presence and absence of glucose (1% w/v). As FnBPs are expressed exclusively in the exponential phase of growth cells were harvested at OD_{600nm} of 0.25, 0.3, 0.4 (to represent very early to mid-exponential phase) and 0.8 (to represent cells that have transitioned to stationary phase) (Figure 3.14 A(ii)). When a ligand affinity
Figure 3.11 A. Bbp/SdrE expression by bovine *S. aureus* strains

Shows expression of Bbp/SdrE by (i) ST151 strains using an increased amount of anti-SdrE antibodies and (ii) a panel of genetically distinct bovine isolates using a standard amount of anti-SdrE antibodies. *S. aureus* strain Newman and an isogenic mutant lacking *sdrCDE* were included as positive (+) and negative (-) controls, respectively for both (i) and (ii). (iii) shows expression of Bbp using anti-Bbp antibodies. Lane 1; RF122, lane 2; RF122Δ*bbp*, lane 3; Newman. *S. aureus* strains were grown in TSB to exponential phase. Cell-wall proteins were solubilised with lysostaphin, separated on SDS PAGE gels, electroblotted onto PVDF membranes and probed for the expression of SdrE/Bbp using rabbit anti-SdrE or anti-Bbp serum and detected with HRP-conjugated protein A.

**B. sdrE amplimers from non-SdrE expressing strains.** Lane 1; ST350, lane 2; ST479, lane 3; ST698. Although *sdrE* could not be amplified from ST350, a band corresponding to *sdrE* was amplified from both ST479 and ST698.
<table>
<thead>
<tr>
<th>Sequence type</th>
<th>Variant residues</th>
<th>Insertions</th>
<th>Deletions</th>
<th>Semi-conserved substitutions</th>
<th>Conserved substitutions</th>
<th>Identical residues</th>
<th>Sequence identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST71</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>13</td>
<td>480</td>
<td>5</td>
<td>95.20%</td>
</tr>
<tr>
<td>ST97</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>13</td>
<td>481</td>
<td>5</td>
<td>95.40%</td>
</tr>
<tr>
<td>ST352</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>13</td>
<td>481</td>
<td>5</td>
<td>95.40%</td>
</tr>
<tr>
<td>ST115</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>13</td>
<td>480</td>
<td>5</td>
<td>95.20%</td>
</tr>
<tr>
<td>ST126</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>13</td>
<td>489</td>
<td>2</td>
<td>97%</td>
</tr>
<tr>
<td>ST8</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>504</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>ST25</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>491</td>
<td>2</td>
<td>97.40%</td>
</tr>
<tr>
<td>ST33</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>7</td>
<td>488</td>
<td>3</td>
<td>96.80%</td>
</tr>
</tbody>
</table>
### Figure 3.12 DNA/amino acid sequence alignment of SdrC from *S. aureus* strains Newman and RF122

**A.** DNA sequence alignment showing the 5’ part of *sdrC* from *S. aureus* Newman and RF122. Variant residues are highlighted in red. - indicates the position of a base that only occurs in one of the compared sequences.

**B.** Amino acid sequence alignment of SdrC from *S. aureus* strains Newman and RF122 showing variants residues. * indicates an identical residue while : indicates a conserved substitution.
Figure 3.13 Amino acid sequence alignment of FnBPA from *S. aureus* strains 8325-4 and RF122

Amino acid sequence alignment of FnBPA from *S. aureus* strains 8325-4 and RF122 showing variants residues. * indicates an identical residue while : and . indicate conserved and semi-conserved substitutions, respectively. - indicates the position of a residue that only occurs in the sequence to which it is compared to. The Fn binding repeats are indicated (high affinity sites are highlighted in red). The signal sequence is indicated by blue arrows, the cleavage site of which is highlighted in grey. The LPETG motif is highlighted in turquoise.
A
(i) 
188 kDa
123 kDa
78 kDa

(ii) 
+ Glucose (1% w/v)  
- Glucose

<table>
<thead>
<tr>
<th>OD600</th>
<th>0.25</th>
<th>0.3</th>
<th>0.4</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>188 kDa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>123 kDa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78 kDa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(iii) 
188 kDa
123 kDa
78 kDa

B
(i) 

(ii) 

\[ \text{Ag/100mM} \]

\[ \text{P1} \]

P1, finbAB, S151, S153, S136, S167, S169, S170

RF122

RF122 + 1% glucose

Bovine fibronectin (μg/ml)
Figure 3.14 A. Expression of FnBPs from *S. aureus* strains P1 and RF122

*S. aureus* strains were grown in TSB. Cell-wall proteins were solubilised with lysostaphin, separated on SDS PAGE gels and electroblotted onto PVDF membranes. Biotinylated fibronectin was used to probe for the expression of FnBPs and was detected using HRP-conjugated streptavidin. (i) Shows lack of FnBPA expression of RF122 when grown using standard conditions. (ii) shows the expression of FnBPA from RF122 grown in TSB to indicated OD_{600nm} in the presence or absence of glucose. (iii) *S. aureus* strains were grown to early exponential phase in TSB. Lane 1; P1, lane 2; P1 fnbAfnbB, lane 3; RF122 grown in TSB supplemented with glucose (1% w/v).

B. **Bacterial adhesion to fibronectin.** Bovine fibronectin was immobilised on 96-well plates and incubated with bacteria grown to exponential phase in TSB, washed in PBS and adjusted to an OD_{600nm} of 1/ml. Bacterial adhesion was detected with crystal violet. (i) adhesion exhibited by a panel of genetically distinct bovine *S. aureus* strains. (ii) shows adhesion by RF122 grown in TSB supplemented with glucose. Each graph is representative of experiments carried out in triplicate with similar results.
blot was performed using these altered conditions FnBPA was expressed weakly by RF122 in the early exponential phase compared to P1 (Figure 3.14 A(iii)).

Bacterial adhesion to Fn was used to detect both the presence and functionality of any FnBPs present in the other bovine strains. The majority of the bovine strains tested here adhered to immobilised fibronectin which suggests that they express a functional fibronectin binding protein(s) (Figure 3.14B(i)). Using standard growth conditions (TSB without the addition of glucose), RF122 and DS37 did not adhere at all. The following strains bound weakly; ST25, ST71, ST97, ST115, ST126. This suggests that the majority of bovine strains have some fibronectin-binding activity. As Figure 3.14A shows, RF122 expressed FnBPA upon addition of glucose. However this did not promote detectable adhesion to fibronectin (3.14B(ii)).

Although the low affinity binding sites domains are mainly intact, there are several variant residues in the high affinity Fn-binding site 1 and deletions in both high affinity Fn-binding sites 4 and 5 of RF122 FnBPA compared to 8325-4 FnBPA, which may account for the lack of adhesion to Fn.

3.2.5 Sas proteins

RF122 has retained the genes that encode SasD, SasF and SasH. They are intact and are summarised in Table 3.1. SasD from RF122 was found to be annotated as a surface protein using the gene locus tag SAB0073. The gene locus tag for SasF of RF122 is SAB2521c protein of RF122. The gene locus tag SAB0023 of RF122 corresponds to SasH and is annotated in the GenBank database as a probable 5' nucleotidase.

3.2.6 Protein A

Analysis of the DNA sequence 5' to the sequence provided in the GenBank database for the spa gene of RF122 indicates that the full sequence is present. A point mutation in the 326th codon led to a premature stop codon resulting in a truncated protein which lacks 12 residues of the Ig-binding domain C (Figure 3.15). A 36 kDa fragment was detected in the culture supernatant of RF122 using HRP-labelled rabbit IgG (data not shown). It is possible that this fragment is biologically active.
3.2.7 Phylogenetic analysis

A neighbour-joining phylogenetic tree was constructed based upon the concatenated sequences of the seven housekeeping genes and was used for MLST analysis. As MLST reflects the evolution of the stable core genome this tree describes the phylogenetic relatedness of the *S. aureus* strains studied here. Proteins expressed by each of these strains are indicated in red, or blue for RF122. As can be seen from the tree, phylogenetic relatedness is not reflective of the protein expression profiles of the strains, which may indicate that genes are being lost, or acquired independently of their sequence type (Figure 3.16).

IsdA, IsdB and ClfB are expressed by RF122 and all of or the majority of bovine *S. aureus* strains. RF122 expresses Bbp while the majority of bovine strains express its allelic SdrE or possess *sdrE* (apart from ST350). Based on these results IsdA, IsdB, ClfB and Bbp/SdrE have been proposed as candidates for a multicomponent vaccine to combat bovine mastitis.
Figure 3.15 Amino acid sequence alignment of protein A from *S. aureus* strains Newman and RF122 showing variants residues.

* indicates an identical residue while : (highlighted in magenta) indicate conserved substitutions. Variant residues are highlighted in red. - indicates the position of a residue that only occurs in the sequence to which it is compared to. The signal sequence is indicated by blue arrows, the cleavage site of which is highlighted in grey. The Ig-binding EDABC domains are indicated.
Figure 3.16 Neighbour-joining tree based upon concatenated sequences of S. aureus MLST alleles

MLST allele sequences representing each strain studied here were used to generate a neighbour joining tree using MEGA 4 (ref). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. *Did not produce SdrE-corresponding band but sequencing revealed the possession of sdrE.
3.3 Discussion

Detailed analysis of genes encoding LPXTG-anchored surface proteins of *S. aureus* RF122 has shown that very few are expressed under *in vitro* growth conditions. A number of proteins that are not expressed by RF122 are still expressed by the majority of bovine strains. This is consistent with a strain that has adapted to a novel niche. RF122 is most likely derived from a strain that originated in humans but has adapted for life in the bovine mammary gland where a number of genes are likely to be redundant (Herron-Olson et al., 2007). Genes that encode surface proteins are found in the core genome or the accessory genome (Baba et al., 2002; Holden et al., 2004). In RF122, known *S. aureus* surface protein genes could be categorised into three distinct groups those that are intact, are pseudogenes or are missing. Several genes encoding surface proteins of RF122 are poorly annotated. In the case of IsdA and IsdB the sequence present in the NCBI database does not correspond correctly to the named gene. Pseudogenes in particular are annotated poorly as some of the sequences provided in the database often appear to be missing 5' bases whereas analysis in this study of DNA upstream of the provided sequence shows that this region is present. The gene has been inactivated due to introduction of a stop codon and not the lack of N-terminal bases as inferred by the sequence provided (eg *isdH* and *spa*).

From the bioinformatic analysis described here it is predicted that all of the genes encoding the Isd system, apart from *isdH* which exists as a pseudogene and *isdI* which is missing, are intact in RF122. This disagrees with results reported by Herron-Olson et al where they stated that genes involved in iron acquisition and metabolism are shown to vary considerably relative to their homologs encoded by *S. aureus* recovered from humans. They stated that a number of polymorphisms occur in iron regulated genes including cell surface transferrin-binding protein *isdB* (SAB0994) and iron regulated surface protein *isdH* (SAB1590) and suggest that iron metabolism is considerably different from that previously described in *S. aureus* (Herron-Olson et al., 2007). However, it is apparent from multiple sequence alignments of amino acid sequences of RF122 and other well characterised *S. aureus* strains (data not shown) that IsdA (gene locus tag SAB0993c) and IsdB (gene locus tag SAB0994c) are annotated incorrectly. The amino acid sequence provided for IsdA corresponds to the IsdB sequence and *vice versa*. Therefore if the sequences of the incorrectly annotated proteins were compared to those of *S. aureus* recovered from humans they would appear to be considerably different.
When the correct sequences were analysed in this study, both *isdA* and *isdB* were predicted to be intact. This is consistent with results obtained by Western immunoblotting, which showed that all of the strains tested, including ST151 and a panel of genetically distinct bovine isolates, expressed both IsdA and IsdB. It is plausible to deduce that retention of both IsdA and IsdB is typical of all bovine strains. While heme is the preferred iron source for *S. aureus* during human infection (Herron-Olson *et al.*, 2007), a predilection for particular iron sources has not been extensively studied in bovine *S. aureus*.

The study by Clarke and Foster (2008) demonstrated that IsdA binds lactoferrin and protects the cell from its bactericidal effects (Clarke & Foster, 2008). This may be advantageous in the mammary gland as most iron is bound to lactoferrin or ferric citrate and the balance changes during different stages of lactation (Mazmanian *et al.*, 2003). Clarke, Wiltshire and Foster (2004), showed that SH1000 adhered to fibronectin and that an isogenic mutant deficient in *isdA* could not (Clarke *et al.*, 2004). However, the level of binding was very low (OD 0.1-0.2) which we would consider to be background. Figure 3.3 C(i) shows that neither RPMl grown-RF122 nor RPMI grown-Newman which express IsdA could adhere detectably to fibronectin. Adhesion to fibronectin could be attributed solely to the expression of FnBPs as P1 wildtype strongly adhered to Fn whereas the isogenic mutant lacking in both fnb genes did not adhere.

IsdB was expressed by all bovine strains here but at different levels. Perhaps adhesion to transferrin is not as important in the mammary gland as it would be during most human infections, where the pathogen is primarily exposed to iron-binding proteins such as haemoglobin and transferrin. Human-associated *S. aureus* isolates prefer heme as an iron source (Herron-Olson *et al.*, 2007). It is likely that the remaining Isd proteins that were predicted to be intact are also expressed suggesting that this system is still important for life in the mammary gland. Recombinant IsdB cloned from RF122 bound haemoglobin (see Chapter 4 section 4.2.1.1), demonstrating that this protein is functional. To test biological significance of this interaction it would be necessary to measure growth of wildtype and mutant RF122 in media that is iron-depleted.

The introduction of a premature stop codon has rendered the *isdH* gene of RF122 non-functional. IsdH facilitates an enhanced conversion of C3b to iC3b and C3d which promotes the ability of *S. aureus* Newman to survive neutrophil uptake and killing (Visai
Table 3.3 Expression profile of bovine *S. aureus* strains

<table>
<thead>
<tr>
<th>ST</th>
<th>IsdA</th>
<th>IsdB</th>
<th>IsdH</th>
<th>ClfA</th>
<th>ClfB</th>
<th>SdrE</th>
<th>Bbp</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>20</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>25</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>71</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>97</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>115</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>126</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>133</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>151</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>350</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>352</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>479</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+*</td>
<td>NA</td>
</tr>
<tr>
<td>693</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>694</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>695</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>696</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>697</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>698</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>699</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>708</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>709</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA Not applicable

*Did not produce SdrE-corresponding band but sequencing revealed the possession of *sdrE*
et al., 2009). As RF122 and a ST350 strain did not express IsdH this suggests that this property is not important for survival in the mammary gland. However, as it is still expressed by the majority of bovine strains it could be considered as a surface protein vaccine candidate.

The Isd sytem of human S. aureus strains has two heme-degrading monooxygenases, IsdI and IsdG. The isdl gene is missing but RF122 has retained isdG. Reniere and Skaar (2008) showed that both IsdG and IsdI heme monooxygenases are required for full pathogenesis. They are differentially regulated by iron and haemin and as both enzymes are required for maximal haemin catabolism they proposed that each enzyme has a distinct contribution. IsdI was expressed in abscess tissue in mice (Reniere & Skaar, 2008).

It seems reasonable to predict that the ability of S. aureus to colonise the bovine host required adhesion to keratinised epithelial surfaces. ClfB and IsdA have been shown to be required for adhesion to human squamous epithelial cells and for colonisation of the nares which might explain why they have been retained by RF122.

The fibrinogen-binding ClfB protein (Ni Eidhin et al., 1998) is important for binding to desquamated nasal epithelial cells through its interaction with cytokeratin 10 which is found in the differentiated layer of the human epidermis. Degradation by aureolysin results in the loss of ClfB ligand binding ability (McAleese et al., 2001). Bioinformatic analysis indicated that ClfB of RF122 is likely to be functional. Residues shown to be important for ligand-binding are retained in RF122 and the A domain shares 96% amino acid sequence identity with that of Newman. However, it is expressed very weakly by RF122 and other ST151 strains. The majority of bovine strains express ClfB, some with levels comparable to Newman. When the clfB gene was cloned from RF122 and expressed from the pQE30 expression vector, the resulting protein bound fibrinogen demonstrating that despite the lack of adhesion by RF122 this protein is functional. The weak expression levels exhibited by RF122 may not be sufficient to promote biologically significant adhesion or perhaps it is subjected to proteolytic degradation by metalloproteases. The expression of IsdA, which has also been shown to promote adhesion to desquamated epithelial cells, may compensate for the functional redundancy of ClfB to allow S. aureus to colonise the bovine host.
ClfA promotes bacterial resistance to phagocytosis by neutrophils. It was proposed that binding to fibrinogen also prevented the deposition or recognition of opsonins. However, phagocytosis experiments performed in the absence of fibrinogen demonstrated that expression of ClfA still had an antiphagocytic effect, suggesting that there is another mechanism. In addition to binding fibrinogen ClfA binds complement factor I and acts as a cofactor to trigger cleavage of C3b to iC3b (Hair et al., 2008; Hair et al., 2010).

The sdr genes occur in a tandem array in human strains – usually three but sometimes two (viz 8325 which lacks sdrE). The sdrC gene exists as a pseudogene in the chromosome of RF122 and sdrD is missing. The SdrC protein has been shown to bind β-neurexin 1 ectodomain (Barbu et al., 2010) and both SdrC and SdrD promote adhesion to an unknown ligand on desquamated nasal epithelial cells (Corrigan et al., 2009).

Bbp, an allelic variant of SdrE, is expressed by RF122. It has been shown to bind human Fg (Vazquez et al., 2011). RF122 did not adhere to Fg using standard conditions. Binding to bone sialoprotein has been difficult to validate (Hook, Foster personal communication). Whatever the true in vivo function of this protein it is likely to have a role to play in RF122.

The bbp gene of RF122 has an in-frame deletion in the SD repeat region compared to sdrE from Newman (Figure 3.10). This is common in repetitive DNA sequences where deletion and duplications occur more frequently. This deletion resulted in the expression of a Bbp protein that was smaller than SdrE from Newman. All of the other bovine strains tested here possess sdrE apart from ST350 from which neither the sdrE or hhp gene could be amplified.

The sas (Staphylococcus aureus surface) genes of RF122 have been poorly annotated but RF122 has retained the genes that encode SasD, SasF and SasH (Table 3.1). SasH (also known as AdsA) converts AMP to adenosine, an important regulator of neutrophil activity, which in turn promotes the survival of S. aureus within the neutrophil (Thammavongsa et al., 2009; Kim et al., 2012). As no functions have been attributed to SasD or SasF, it is difficult to speculate as to why the genes have been retained in the RF122 genome. RF122 does not carry sasG which promotes bacterial adhesion to desquamated nasal epithelial cells and is also involved in the accumulation phase of biofilm formation (Roche et al., 2003b; Corrigan et al., 2007; Geoghegan et al., 2010).
As discussed with regards to ClfB, IsdA may be responsible for adhesion to epithelial cells of the bovine host.

Uptake into the epithelial cells in the bovine mammary gland is regarded as important in pathogenesis of mastitis (Wesson et al., 1998). After internalization bacteria can (i) proliferate and activate apoptosis or cause necrosis thus ultimately destroying the cells or (ii) persist in a semi-dormant state in the form of small colony variants (SCVs) which are metabolically inactive forms of *S. aureus* which contribute to intracellular survival by evasion of phagocytosis. This promotes the development of persistent staphylococcal infections and resistance to antibiotics. Small colony variants have been shown to persist in bovine mammary epithelial cells (Atalla et al., 2010).

The best understood mechanism for *S. aureus* internalisation into human host epithelial cells is provided by FnBPs. Fibronectin acts as a bridging molecule which binds N-terminally to *S. aureus* FnBPs and via an RGD motif to the host cell integrin, α5β1 and signals uptake (Sinha et al., 1999; Fowler et al., 2000; Massey et al., 2001). RF122 possesses an intact fnbA gene but does not possess fnbB. Ligand affinity blotting showed that RF122 could not express FnBPA using standard growth conditions but the addition of glucose led to weak expression. However, it was not sufficient for significant ligand-binding. Several other bovine strains could adhere to fibronectin. Studies have shown that internalisation of *S. aureus* into mammary gland epithelial cells occurs and has been demonstrated to occur *in vivo* (Kintarak et al., 2004). Bacteria were identified within macrophages and bovine alveolar cells isolated from milk of naturally infected cows. In a murine model bacteria were found in the lumen of ducts and alveoli, in degenerating cells and in neutrophils (Brouillette et al., 2003).

Although FnBPs provide the major mechanism of internalisation, there is substantial evidence that shows internalisation can be mediated by both FnBP-dependent and -independent pathways. Kintarak et al reported that internalization by primary keratinocytes could occur through both FnBP-dependent and independent pathways, for which ClfB was not involved, whereas internalisation into immortalised skin keratinocytes requires FnBPs and is mediated by α5β1. Although adhesion of *S. aureus* to bovine mammary epithelial cells was reduced in FnBP-deficient mutants, an FnBP-deficient isogenic mutant was internalised by mammary epithelial cells in a murine mastitis model (Kintarak et al., 2004). Brouillette et al showed that FnBPs are not absolutely required for bacterial internalization into mammary epithelial cells *in vivo.*
However, the observed delay in internalisation of mutant bacteria confirms their importance in infection process (Brouillette et al., 2003).

Perhaps RF122 employs an alternate method of internalization. RF122 possesses the *cap* gene (Table 3.1). The *S. aureus* extracellular adherence protein Eap, has been shown to bind several plasma proteins including Fn, Fg and prothrombin and can enhance internalization into eukaryotic cells (Haggar et al., 2003). RF122 expresses IsdB (Figure 3.5). IsdB has been shown to bind the platelet integrin GPIIb/IIIa (Miajlovic et al., 2010) and to RGD-containing integrins. It promotes internalisation into epithelial cells independently of Fn (M. Zapotoczna, manuscript in preparation).

The *spa* gene exists in the RF122 genome as a pseudogene. The inactivation of this protein was previously observed in poultry where an identical *spa* mutation was apparent in all poultry strains of ST5. The avian equivalent of IgG is IgY, which has a structurally distinct Fc region that does not bind to Spa suggesting that protein A would not contribute to inhibition of opsonophagocytosis in birds (Lowder et al., 2009). However the secreted form might be biologically active as a proinflammatory molecule that binds TNFR-1 or acts as a B cell mitogen.

In this chapter, several genes of RF122 were deemed redundant and those that have been retained are also retained by the majority of bovine strains which has allowed us to speculate on how this successful strain has adapted to life in the bovine mammary gland. Retention of certain surface protein genes also has implications for development of a surface protein vaccine to combat bovine mastitis. Based on the results of this chapter the following MSCRAMMs are proposed as candidates for inclusion in a multicomponent vaccine to combat bovine mastitis; IsdA, IsdB, ClfB and Bbp/SdrE. As SasH/AdsA has recently been shown to have an important function in promoting survival in neutrophils it should be tested (Kim et al., 2012).
Chapter 4

Further studies on bovine mastitis vaccine candidate antigens
4.1 Introduction

*Staphylococcus aureus* is a highly prevalent, opportunistic, multifactorial pathogen and is the agent most frequently associated with contagious bovine mastitis which is one of the most economically important diseases in the dairy industry. In a study by Moser *et al.*, 2012, 78 mastitis-causing bovine *S. aureus* isolates were analysed for antimicrobial resistance. Resistance profiles were generated by disk diffusion. CC151 isolates (N = 40) were shown to be susceptible to all antimicrobial agents tested (ampicillin, amoxicillin with clavulanic acid, cephalothin, ceftiofur, erythromycin, cefoxitin, gentamicin, kanamycin, kanamycin-cefalexin, penicillin and penicillin-novobiocin). Out of 16 isolates belonging to CC97 one was resistant to both ampicillin and penicillin and three had intermediate sensitivity to kanamycin-cefalexin. One representative from each of CC7, CC8, CC9 and CC45 also showed resistance to ampicillin and penicillin. One representative each from CC8 and CC9 and three out of four isolates from CC20 exhibited intermediate sensitivity to kanamycin-cefalexin. However this study did not include any isolates from the CC133. Like CC151, isolates from CC479 and CC20 included no resistant isolates which is consistent with an extensive antimicrobial resistance study among Swiss and French bovine *Staph. aureus* isolates, which demonstrated that penicillin resistance rates among CC151 and CC20 were far lower than those among CC97 (Sakwinska *et al.*, 2011) (Moser, 2012).

To prevent the morbidity and mortality associated with bovine mastitis caused by *S. aureus*, and to decrease the associated economic costs, an improved means of controlling bovine mastitis is required. A logical approach is rational design of a staphylococcal vaccine. *S. aureus* is characterised by a wide array of virulence factors and although vaccination has been studied as a control measure for *S. aureus* mastitis for several years the antigenic diversity among different strains has hampered the development of an effective vaccine therefore prevention of intramammary infection (IMI) has not been demonstrated (Hu *et al.*, 2010; Scarpa *et al.*, 2010).

Many of the factors involved in *S. aureus* pathogenesis have been investigated as vaccine targets for both bovine mastitis and human infections. The simplest means of combatting *S. aureus* bovine mastitis is by using whole cell or whole cell lysates with or without adjuvants. For example, Lysigin is a polyvalent whole cell lysate vaccine containing five strains of *S. aureus* (Ma *et al.*, 2004). However, whole cell vaccines have met with varied success, and while some studies report reduced clinical severity of
disease and/or reduction in new IMI rates, none to date have completely prevented new IMI in dairy cattle.

An avirulent *S. aureus* mutant, generated by chemical mutagenesis, was tested for its potential as an immunogen in heifers following intramammary challenge with a homologous virulent strain. The avirulent strain was immunogenic but no significant differences were found for bacterial shedding in milk or milk SCCs (Pellegrino *et al.*, 2008). The same mutant increased LD$_{50}$ when challenged with a homologous strain but not heterologous strains (Bogni *et al.*, 1998).

Another strategy for vaccination is to use purified capsular polysaccharide antigens (Tuchscherr *et al.*, 2008). A large proportion of *S. aureus* strains express capsular polysaccharide (CP) in animal models of staphylococcal infection (CP serotype 5 (CP5) and CP8 produced by ~75% of human isolates). Loss of CP expression facilitates the internalisation of *S. aureus* into bovine epithelial cells and contributes to the persistence of *S. aureus* in infected mammary glands of mice. As CP antigens alone are poorly immunogenic they are often conjugated to proteins and such a vaccine may be more effective than either antigen alone for prevention of *S. aureus* infections. Following intramammary challenge with encapsulated *S. aureus* strains, mice immunized with antibodies to CP5 or CP8 or with ClfA had significantly reduced tissue bacterial burdens after 4 days. After several passages, increasing numbers of stable unencapsulated variants of *S. aureus* and small colony variants were recovered in the infected mammary glands of mice passively immunised with CP-specific antiserum. The unencapsulated *S. aureus* variants were internalised *in vitro* into MAC-T bovine cells in greater numbers than the corresponding encapsulated parental strains. Administration of antibodies to CP and ClfA abrogated the emergence of unencapsulated *S. aureus* and decreased the recovery of small colony variants from the infected mouse tissues. Antibodies to ClfA enhanced the protection against infection provided by antibodies to the CPs (Tuchscherr *et al.*, 2008).

Immunization with plasmid DNA encoding *S. aureus* antigens either singly or as chimeric products containing at least two antigens has been proposed as a novel strategy to prevent this costly disease. However DNA vaccines against ClfA and Cna did not provide protection against intraperitoneal challenge with *S. aureus* in mice (Brouillette *et al.*, 2002; Therrien *et al.*, 2007). DNA vaccines targeted against the mecA gene caused significant reductions in bacterial loads in kidneys of mice challenged with a sublethal dose of MRSA. However MRSA is infrequently isolated from bovine mammary glands so this target might not be appropriate. DNA vaccines targeting specific virulence factors
may be of use in future vaccine development (Ohwada et al., 1999; Senna et al., 2003; Roth et al., 2006).

A series of DNA vaccines directed against ClfA demonstrated a strong and specific antibody response in vaccinated dairy cattle groups relative to control cattle that received only the expression vector. In vitro studies showed that S. aureus preincubated with sera or milk from vaccinated cattle showed a reduced ability to adhere to MAC-T cells compared with preincubation with sera or milk from cattle injected with the expression vector only. However no in vivo efficacy data were reported (Nour El-Din et al., 2006).

One study used a chimeric protein composed of the surface-located GapB and GapC proteins of S. aureus, which were isolated from a bovine mastitis strain, and tested the effects of DNA vaccination with plasmids encoding the individual antigens as well as the GapC/B protein with or without a boost with the recombinant proteins. Immunization with plasmids encoding GapB and GapC or GapC/B did not elicit a significant humoral immune response and only elicited a low cell-mediated immune response after three immunizations in mice. This could be overcome by boosting with recombinant GapB and GapC proteins. However, it is more likely that the plasmid vaccines had no priming effect and the responses elicited were due to recombinant proteins alone (Kerro-Dego et al., 2006).

The protective immune responses of a multi-gene DNA vaccine against S. aureus in a murine model were evaluated. The multivalent polyprotein DNA vaccine was formulated against ClfA, FnBPA, and the enzyme sortase and was compared with monovalent DNA vaccines against each of the individual proteins and a plasmid vector control. When challenged with a virulent S. aureus strain 55 % of vaccinated animals survived compared to 15 % of control mice. However, all of the surviving mice developed arthritis (Gaudreau et al., 2007).

Although plasmid DNA-based vaccines encoding recombinant proteins can induce specific systemic and mucosal immune responses in mice, DNA may be integrated into the host cell chromosomes by insertional mutagenesis thus implicating safety concerns. A more global approach to vaccination is to target conserved antigens produced by multiple strains of S. aureus. Several surface proteins have been investigated for their potential as vaccine candidates. A study investigated the role of ClfB by mucosal immunization with killed S. aureus and both systemic and intranasal immunization with rClfB. The killed S. aureus cells used for mucosal immunization were capsule negative and had been grown
under conditions to optimally express ClfB. Mice were then challenged with *S. aureus* strain Newman. Results showed that the CFU of *S. aureus* isolated from the nose were significantly reduced compared to controls. As carriage rates did not differ much between treated and control groups it was suggested that although anti-ClfB antibodies may reduce nasal shedding by reducing intranasal bacterial loads they are not sufficient to prevent nasal colonisation (Schaffer et al., 2006).

A Cna-FnBP fusion protein was investigated for its immunogenicity against *S. aureus* in a mouse model. Results showed that it stimulated a stronger humoral immune response, enhanced *in vitro* opsonophagocytosis and decreased post-challenge mortality in vaccinates compared to controls. However, 3 out 12 vaccinates succumbed following challenge with *S. aureus* (Zhou et al., 2006).

Dairy cattle were vaccinated twice during lactation with a fusion protein comprising FnBP and Spa before intramammary challenge with *S. aureus*. No difference was observed between vaccinates and controls regarding bacterial shedding and milk somatic count. However, when the fusion protein was conjugated with immunostimulating complexes, higher antibody titres following immunization were observed. No cases of clinical mastitis post-intramammary-challenge were observed compared to 3 out of 5 cases in unvaccinated controls. As only the milk SCC immediately after challenge was reported and the numbers at day 17 post-challenge were not, the ability of the vaccine to prevent infection cannot be determined from the data presented (Rappouli & Bagnoli, 2011).

Some iron-regulated proteins were investigated for their potential as vaccine candidates. Seven genes from five different iron-acquisition systems which coded for surface-expressed proteins including the *isdH* gene were selected. Recombinant proteins were used to generate immune sera. Out of the seven proteins investigated, FeoB, SstD and IsdB showed potential as vaccine candidates. Although the antibody responses were moderate and no lymphoproliferation was detected, their sera labelled the bacteria and provided protection in a mouse mastitis model. Antibodies against SirA, FhuD2 and SrtB did not bind intact bacteria. IsdH seems to be a good vaccine candidate as strong and long-lasting immune responses were observed and the serum reduced mammary gland colonisation in the mouse model (Ster et al., 2010).

Purified IsdB with an aluminium hydroxyphosphate adjuvant proved to be highly immunogenic in mice and rhesus macaques. In a mouse sepsis model a significant reduction in mortality relative to control mice which received the adjuvant alone, was
observed. A multicomponent vaccine comprising recombinant IsdA, IsdB, SdrD and SdrE in a mouse challenge model was mixed with adjuvant and was administered twice 11 days apart. Mice were challenged with *S. aureus* 10 days after the second vaccination. The combined vaccine boasted significantly greater survival at 7 days post challenge compared with saline-treated controls. As murine models do not often correlate with human infection or bovine mastitis, results should be interpreted with caution. However, combining a series of conserved *S. aureus* antigens with other conserved *S. aureus* virulence factors may prove useful in future vaccine development (Kuklin *et al.*, 2006; Stranger-Jones *et al.*, 2006).

Another approach, used to combat other *S. aureus*-induced diseases, is the use of toxoids. *S. aureus* produces number of exotoxins which either directly cause disease or facilitate tissue penetration and immune cell recruitment. One such exotoxin is the *S. aureus* α-haemolysin (Hla). Its role in *S. aureus* pneumonia was investigated using murine and *in vitro* models and was shown to be required for pathogenesis. Following the active immunization of mice with a non-pore-forming Hla mutant, antigen-specific antibodies were generated and a significant decrease in mortality following challenge was observed compared to saline controls (Bubeck Wardenburg & Schneewind, 2008). α-toxin toxoid might be a good candidate for a multicomponent mastitis vaccine and will be discussed in Chapter 6 section 6.1.2.2.

A non-toxic mutant TSST-1 was assessed for its potential as a vaccine candidate. Mice were subsequently challenged with live *S. aureus*. Vaccinates had a higher survival rate and significantly lower bacterial counts than controls. Passive transfer of rabbit serum containing non-toxic mutant TSST-1-specific antibodies provided protection against *S. aureus* sepsis in mice. However, about 40% of actively immunized vaccinates succumbed to infection by day 11 post challenge and about 80% of mice passively immunized died by day 12 post challenge (Hu *et al.*, 2003).

A systematic review, which aimed to suggest the immunotherapeutic protocols which achieved the best and/or most promising results, assessed the efficacy of vaccines for bovine mastitis caused by *S. aureus* (Pereira *et al.*, 2011). It was concluded that vaccines employing new technologies (DNA and/or recombinant protein vaccines) and some long-standing bacterins achieved good results, which supports their use in the prevention and control of bovine mastitis caused by *S. aureus* (Pereira *et al.*, 2011). The complexity of this organism and the diversity of strains which cause infection have made the discovery of a pluripotent staphylococcal vaccine difficult, if not impossible. It is
likely that the best vaccine will need to possess an array of *S. aureus* virulence factors or block expression of major pathogenic determinants. The route of immunization and the ability of adjuvants to trigger both humoral and cell-mediated immune responses both to activate neutrophils and combat intracellular bacteria will be important.

The aim of the work described in this chapter is to propose and prepare potential candidates for a surface protein vaccine to combat bovine mastitis based on results discussed in Chapter 3. RF122 was used as the reference strain as it is widespread, successful and its genome is fully sequenced. Candidates were cloned into the pQE30 expression vector to allow purification of recombinant protein using a nickel column and recombinant proteins were tested for functionality. In order to demonstrate specificity in a murine mastitis model of infection it was necessary to delete the genes encoding these proteins in strain RF122. The original goal was to test these antigens as immunogens to protect against infection in the mouse mastitis infection model in the laboratory of A. Zecconi but this fell through.
4.2 Results

4.2.1 Cloning vaccine candidate genes into the pQE30 expression vector for the purification of recombinant proteins

4.2.1.1 Cloning isdB into the pQE30 expression vector for the purification of recombinant IsdB

In order to purify recombinant IsdB for use as an immunizing antigen as part of a multicomponent vaccine it was necessary to clone the coding region of *isdB* (amplified from RF122 genomic DNA) into the pQE30 expression vector which adds an N-terminal histidine tag allowing the recombinant protein to be purified on a nickel column. A study by Clarke et al (2004) described the cloning and purification of the IsdA protein from *S. aureus* (Clarke *et al.*, 2004). As they successfully purified a stable recombinant IsdA fragment it was reasonable to assume that an IsdB fragment could be cloned from RF122 genomic DNA and purified in a similar manner. Primers were designed using the *isdB* sequence from RF122 and a gene fragment encoding residues 41 to 594, which contain the putative ligand-binding NEAT 1 and NEAT 2 domains, was amplified. The 1659 bp PCR product was cloned into pQE30 between the *BamH*I and *Sph*I sites (Figure 4.1) and the construct was verified by restriction analysis (Figure 4.2A (i)). The construct was sequenced using primers which flank the multiple cloning site of pQE30. The fusion protein was expressed in *E. coli* strain XL-1 as a 6xHis fusion protein and purified from lysed cells by nickel-chelate affinity chromatography. The protein was analysed by SDS-PAGE and Coomassie blue staining following incubation at 4°C for 18 h. The protein migrated as a single band with an apparent molecular mass of approximately 80 kDa indicating that the recombinant IsdB protein was purified to homogeneity and was stable (Figure 4.2 B (i)). The NEAT 1 domain of IsdB is responsible for binding to haemoglobin and, as shown in Chapter 3 (Figure 3.4), the NEAT 1 domain of RF122 contains two conserved substitutions and one variant residue when compared to IsdB from *S. aureus* strain Newman. The 3D structure of the NEAT 2 domain could be modelled based on the most related hemin-binding NEAT domain (M. Zapotoczna, personal communication) and could have been used to investigate the functionality of RF122 IsdB. Recombinant IsdB from RF122 bound haemoglobin in a dose-dependent and saturable manner which indicates that the NEAT 1 domain is functional (Figure 4.2 C (i)).
Figure 4.1 Cloning *isdB*, *clfBN2N3* and *bbpN2N3* DNA from RF122 for the expression of rIsdB, rClfBN2N3 and rBbp protein

DNA encoding fragments of *isdB* (from after the signal sequence to the LPXTG motif), *clfBN2N3* (primers were designed to eliminate the metalloprotease aureolysin cleavage site so that a stable protein could be purified) and *bbpN2N3* (with a CltB-like cleavage site) were amplified from the genomic DNA of *S. aureus* strain RF122. PCR products were cloned into the *E. coli* expression vector pQE30 using the indicated restriction sites which were added to the PCR amplimers as 5' extensions to primers. pQE30 constructs were transformed into the *E. coli* strain XL-1 for the expression of recombinant proteins.
Figure 4.2 A. Restriction analysis of the (i) pQE30::isdB construct, (ii) pQE30::clfBN2N3 construct and (iii) pQE30::bbpN23 (with ClfB-like cleavage site)

(A) The pQE30 plasmid containing the coding region for (i) IsdB was verified by digestion with BamHI and SphI restriction enzymes, (ii) ClfBN2N3 was verified by digestion with BamHI and HindIII restriction enzymes, (iii) BbpN2N3 was verified by digestion with BamHI and HindIII restriction enzymes.

B. SDS-PAGE analysis of (i) rIsdB, (ii) rClfBN2N3 and (iii) rBbpN2N3. Recombinant proteins were incubated in PBS for 18h at 4°C and analysed by SDS-PAGE and Coomassie blue staining.

C. Binding of (i) rIsdB to immobilised haemoglobin and (ii) rClfBN2N3 to immobilised fibrinogen. (i) Microtitre plates were coated with 5μM rIsdB from RF122. Increasing concentrations of haemoglobin were added for 2h at 37°C. Anti-haemoglobin IgG was used to detect bound protein which was subsequently detected with protein A-peroxidase followed by development with TMB substrate. Graphs are representative of three separate experiments.

(ii) Microtitre plates were coated with 20 μg/ml human fibrinogen. Increasing concentrations of rClfBN2N3 were added and incubated for 2 h at 37°C. Bound protein was detected with HRP-conjugated anti 6XHis IgG followed by development with TMB substrate.
4.2.1.2 Cloning *clfB* into the pQE30 expression vector for the purification of recombinant ClfB

Chapter 3 showed that RF122 expressed ClfB very weakly compared to strain Newman which likely explains its inability to adhere to immobilised keratin 10 or fibrinogen. The crystal structure of the ClfB N2N3 subdomains has been solved. ClfB binds to CK10 and Fg by a variation of the Dock Latch Lock mechanism and residues that contact the ligand are known (Ganesh *et al.*, 2011). In order to determine if the lack of adhesion to its ligands by RF122 was due to loss of function of ClfB it was important to examine a molecular model of the N2N3 region of ClfB which was viewed using the Chimera software (http://www.cgl.ucsf.edu/chimera/).

Figure 4.3 shows the independently folded subdomains N2 and N3 which are composed largely of anti-parallel β-strands. Semi-conserved substitutions and variant residues, which were identified by comparing the amino acid sequences of the ClfB N2N3 subdomains from RF122 and Newman, are highlighted on the model. The residues lining the ligand-binding trench were conserved which suggests that ClfB of RF122 is functional but two variant residues were observed which could be involved in latching (Figure 4.3).

To investigate this further, a *clfBN2N3* gene fragment was amplified from RF122 and cloned into the expression vector pQE30 (Figure 4.1). Recombinant fragments comprising full length N-terminal A domains can show diminished stability. The use of N23 subdomains can circumvent this problem. The subdomains exhibited increased stability without compromising ligand-binding ability in comparison to the full length A domain. Based on the coordinates provided in a paper by Walsh *et al* (2004) primers were designed that eliminated the aureolysin cleavage site (Walsh *et al.*, 2004), which is located between the N2 and N3 subdomains, allowing the purification of a stable recombinant ClfB protein (Figure 4.2 B (ii)). The 1020 bp gene fragment encoded residues 201-541. The predicted molecular weight of rClfB is 38 kDa.

Adhesion to CK10 was investigated in Chapter 3. However, CK10 was not available during experimentation for work described in this chapter. Therefore fibrinogen, another ligand recognised by ClfB was used. rClfB from RF122 bound to fibrinogen in a dose-dependent manner with a similar affinity as ClfB from Newman which had been cloned and expressed using the same strategy, demonstrating that it is functional (Figure
4.2 C(ii)) and that the failure of TSB-grown RF122 to adhere to CK10 is due to poor expression.

4.2.1.3 Cloning bbp into the pQE30 expression vector for purification of recombinant Bbp

In order to purify recombinant Bbp for use as an immunizing antigen as part of a multicomponent vaccine, and to generate Bbp-specific mouse antibodies to specifically detect Bbp, it was necessary to clone the coding region of bbp (amplified from RF122 genomic DNA) into pQE30.

As previously mentioned, recombinant fragments comprising full length N-terminal A domains can show diminished stability. The precise coordinates for the start of the N2 subdomain of Bbp are not currently known. In an attempt to predict these coordinates the primary sequence of the A domain of Bbp was submitted to the Protein homology/analogy recognition engine (Phyre) from the Imperial College, London (http://www.sbg.bio.ic.ac.uk/~phyre/). This service predicts the 3D structure of the submitted sequence based on the crystal structures of similar proteins. The predictions are graded according to the precision of the predicted structure compared to that on which it is modelled. The highest recommendations for the structure of Bbp are based on SdrG, ClfA and ClfB. However, none of the models showed the first few residues of the N2 subdomain. For example the coordinates for the ClfBN2N3 subdomains are defined as residues 199-542 but only residues 212-528 are shown in the crystal structure. Therefore the coordinates for the start of the N2 subdomain of Bbp could not be predicted from the model.

Metalloproteases recognise the SLAAVA motif (in ClfA) and SLAVA (in ClfB) resulting in loss of N1. The SLAAVA/SLAVA cleavage sites mark the boundary between the N1 and N2 subdomains. When the A domain of RF122 Bbp was aligned with the A domains of ClfA, ClfB, SdrE from Newman and Bbp from 024, it was apparent that RF122 Bbp contains two putative cleavage sites. The first motif (RFAVA) is similar to the SLAAVA cleavage site motif of ClfA. The second (QPAAVA) is similar to the SLAAVA motif of ClfA (Figure 4.4).

To determine the coordinates of start of the N2 subdomain and to purify a protein that retains the HIS tag and is stable following purification, three pairs of primers were
A ligand binding trench forms between the N2 and N3 domains of ClfB. Residues that differ between the solved crystal structure of ClfB (Ganesh et al., 2011) and RF122 ClfB are indicated. Residues highlighted in red represent variant residues, while yellow represents semi-conserved residues.
Figure 4.4 Amino acid sequence alignment of part of the A domains from (A) ClfA and ClfB and (B) Bbp, SdrE from *S. aureus* strains

A. The A domains of ClfA and ClfB from *S. aureus* strain Newman were aligned. The SLAAVA/SLAVA metalloprotease cleavage sites are shown.

B. The A domains of Bbp from *S. aureus* strains RF122 and 024 were aligned with the A domain of SdrE from *S. aureus* strain Newman. Putative metalloprotease cleavage sites from Bbp and SdrE were predicted and are shown.
designed that would amplify a *bhp* N123 fragment, a *bhp* N23 fragment which eliminated the ClfB-like RFAVA putative cleavage site and a *bhp* N23 fragment which eliminated the ClfA-like QPAAVA putative cleavage site. Each fragment was cloned into the pQE30 expression vector but only when the RFAVA cleavage site was eliminated could a stable rBbpN2N3 protein be purified (Figure 4.2 B(iii)). The predicted size of rBbpN23 is ~37 kDa. The RFAVA cleavage site, marking the boundary between N1 and N2, is located between residues 247-251. After this work was performed it was reported that Bbp from 024 bound Fg (Vazquez *et al.*, 2011).

4.2.2 A multicomponent vaccine comprising *IsdA, IsdB, ClfB* and *Bbp/SdrE*

In Chapter 3 the IsdA, IsdB, ClfB and Bbp/SdrE proteins were shown to be expressed by RF122 and either all or the majority of bovine strains in a panel of genetically distinct isolates. This chapter has described the preparation of recombinant proteins for use as a multicomponent vaccine. IsdA was previously cloned and purified in our lab and was sent, along with recombinant IsdB, ClfBN2N3 and Bbp N2N3 to a lab in Milan. Unfortunately the work was not carried out by the intended collaborators as promised. Therefore a conclusion cannot be drawn as to the efficacy of the proposed vaccine.

4.2.3 Deletion of *isdA, isdB, clfB* and *bhp* from the RF122 chromosome

In order to demonstrate specificity of a multicomponent vaccine it is necessary to delete each gene involved and compare the protective efficacy of the vaccine against wild type and mutant bacteria using an animal model. To facilitate the generation of separate *isdA, isdB, clfB* and *bhp* mutations ‘deletion fragments’ were generated. To create a deletion fragment, regions up- and downstream of each gene of interest were amplified from RF122 genomic DNA using primer pairs “A and B”, and “C and D”, respectively. The upstream and downstream fragments are referred to as AB and CD fragments, respectively. Primer C had a 5' extension that complements the end of the AB fragment which facilitated joining of AB-CD fragments and amplification of the ABCD deletion fragment, using the AB and CD fragments as template, with the A and D primer pair. All of the deletion fragments were constructed in this manner. Deletion fragments could then
be cloned into pKOR1 (Figure 4.5 A) via gateway cloning as described in Chapter 2 section 2.6.5.

The *isdA* mutation deleted all but the start codon at the 5' end of the gene and 15 bp at the 3' end, creating a 1035 bp unmarked deletion. To construct the 1009 bp *isdA* deletion fragment, a 502 bp AB fragment located 5' to the *isdA* gene and a 507 bp CD fragment covering the 3' end of the gene were amplified by PCR. Primer C had a 25 bp extension that complemented the end of the AB fragment which facilitated amplification of the ABCD deletion fragment.

The *isdB* mutation deleted all but the start codon at the 5' end of the gene, creating a 1887 bp unmarked deletion. To construct the 1000 bp *isdB* deletion fragment, a 502 bp AB fragment located 5' to the *isdB* gene and 498 bp CD fragment which covered the 3' end of the gene were amplified by PCR. Primer C had a 32 bp extension that complemented the end of the AB fragment.

The *clfB* mutation deleted all but the start codon at the 5' end of the gene and the stop codon at the 3' end, creating a 2580 bp unmarked deletion. To construct the 999 bp *clfB* deletion fragment, a 499 bp fragment located 5' to the *clfB* gene and 500 bp covering the 3' end of the gene were amplified by PCR. Primer C had a 26 bp extension that complemented the end of the AB fragment.

The *hhp* mutation deleted all but the start codon at the 5' end of the gene and the stop codon at the 3' end, creating a 3336 bp unmarked deletion. To construct the 1006 bp *hhp* deletion fragment, a 504 bp fragment located 5' to the *hhp* gene and 502 bp covering the 3' end of the gene were amplified by PCR. Primer C had a 21 bp extension that complemented the end of the AB fragment.

Work carried out for this chapter indicated that RF122 will not accept DNA from the restriction-deficient *S. aureus* strain RN4220 (this is not the case, as will be discussed in Chapter 5 section 5.2.6). RN4220 is typically used as an intermediate strain to facilitate plasmid transfer to closely related wildtype strains eg SH1000, Newman and LAC (USA 300 CA-MRSA). However, different restriction-modification systems reduced transfer of unmodified DNA into less related strains. RF122 accepts DNA from a variant called RF122t which was used as an intermediate strain in this study to genetically manipulate RF122. The pKOR1Δ constructs were transferred into RF122t from *E. coli* K12 XL-1.
A  

Gene of interest

pKOR1

B

(i) 506bp 1053bp 501bp = 2060bp

506bp 501bp = 1007bp

(ii) 2 Kb 2060bp

1 Kb 1007bp

C

(i) RF122 ΔisdA

(ii) RF122 ΔisdB

(iii) RF122 ΔcfB

(iv) RF122 Δbpp
Figure 4.5  A. Schematic representation of the *isdA*, *isdB*, *clfB* and *bbp* mutations

A ~500bp fragment 5' to the each gene and 500bp fragment 3' to the end of the gene were cloned into pKOR1 via Gateway cloning.

B. Validation of RF122Δ*isdA* (i) Diagrammatic representation of PCR screening products to identify wildtype and mutant strains deficient in *isdA*. (ii) Diagnostic PCR with primers A and D which would amplify a 1007bp fragment from strains in which the *isdA* mutation was retained in the chromosome. Strains that retained the gene produced an amplimer of 2060bp. Primer positions are indicated in Figure A.

C. Validation of mutant strains. Western immunoblot showing the loss of expression by mutant strains RF122Δ*isdA*, RF122Δ*isdB*, RF122Δ*clfB* and RF122Δ*bbp* of (i) IsdA, (ii) IsdB, (iii) ClfB or (iv) Bbp, respectively. *S. aureus* strains were grown in RPMI or TSB to the stationary phase or exponential phase. Cell-wall proteins were solubilised with lysostaphin, separated on SDS PAGE gels, electroblotted onto PVDF membranes and probed for the expression of proteins using specific rabbit IgG serum followed by HRP-conjugated protein A.
selecting on Cm and growing at the permissive temperature of 28°C. DNA was isolated and validated by amplification of the ABCD fragment before transfer to RF122. Mutations in the 4 chromosomal genes were isolated by the standard allele exchange procedure described in detail in Chapter 2 section 2.6.6. The mutations were first identified by PCR analysis (Figure 4.5 B (ii)) and subsequently by Western blotting using specific antibodies (Figure 4.5 C). To investigate whether the mutant strains contained secondary mutations each mutant was checked for expression of IsdA, IsdB, ClfB and Bbp. With the exception of the specific deleted gene that was no longer expressed, each mutant expressed all of the other proteins and the mutants were compared to the parental wild-type strain for expression of δ-toxin as a surrogate for agr function. Spontaneous agr mutations can occur at high frequency under stress and these could affect expression of other virulence factors and alter fitness which could affect the rate of growth. δ-haemolysin has very weak activity on sheep blood agar but is strongly synergistic with β-haemolysin, producing a zone of clear haemolysis where they interact. RF122 and each mutant were cross-streaked against RN4220, which produces only β-haemolysin (Tegmark et al., 2000). A zone of clear haemolysis was observed for RF122 and each mutant indicating normal agr function in these strains (data not shown).

The in vivo work was to be performed by a collaborative lab but, due to circumstances beyond our control, they did not. However, a genetic system for the manipulation of the RF122 chromosome was established which led to work on restriction-modification systems of RF122 and will be discussed in Chapter 5.
4.3 Discussion

The highly prevalent, opportunistic, multifactorial pathogen *Staphylococcus aureus* is the agent most frequently associated with contagious bovine mastitis, which is one of the most economically important diseases in the dairy industry. The development of an effective vaccine against *S. aureus* mastitis is jeopardized in a number of ways including the low levels of secreted immunoglobulins in the milk and the complexity of *S. aureus* adhesion machinery. Although numerous attempts to develop effective vaccines have been made, including those containing killed bacteria or surface antigen virulence factors, none have proven to be ideal and an improved means of controlling bovine mastitis is required. Vaccination strategies to combat bovine mastitis caused by *S. aureus* which appear to be the most promising include those which employ new technologies such as DNA and/or recombinant protein vaccines.

*S. aureus* possesses on its cell surface a number of proteins, collectively called MSCRAMMs, which promote the binding of the organism to components of the host’s plasma and extracellular matrix, many of which have been assessed for their potential in recombinant protein vaccines or DNA vaccines. Although vaccination has been studied as control measure for *S. aureus* mastitis for several years the antigenic diversity among different strains has hampered the development of an effective vaccine therefore prevention of IMI has not been demonstrated (Hu *et al.*, 2010; Scarpa *et al.*, 2010).

Purified IsdB with an adjuvant proved to be highly immunogenic in mice and rhesus macaques and a multicomponent vaccine comprising recombinant IsdA, IsdB, SdrD and SdrE had significantly greater survival at 7 days post challenge compared with saline-treated controls in a murine model (Kuklin *et al.*, 2006; Stranger-Jones *et al.*, 2006). IsdH elicited strong and long-lasting immune responses and the serum reduced mammary gland colonisation in the mouse model (Ster *et al.*, 2010). Previous studies have shown that ClfB has potential as an attractive candidate for inclusion in a vaccine. Reduced nasal colonization compared to controls was observed in several studies including intranasal immunization of mice using killed *S. aureus* cells grown under conditions to optimally express ClfB (mutants were deficient in sortase A or ClfB), systemic or intranasal immunization of mice with recombinant ClfB A domain and passive immunization of mice with a ClfB monoclonal antibody (an isotype-matched antibody used as a control) (Schaffer *et al.*, 2006).
RF122 is the most common *S. aureus* clone derived from bovine mastitis worldwide. Since its transfer from man to the bovine the strain has undergone many changes as it has adapted to the new host. These include gene decay by mutations to create "pseudogenes". This project has identified those few remaining genes that are intact and are expressed *in vitro*. Genes retained by RF122 are likely to contribute to the success of this strain and therefore a number of the proteins encoded were considered as part of a multicomponent vaccine. Namely IsdA, IsdB, ClfB and Bbp.

The crystal structure of ClfBN2N3 subdomains has been solved and the residues involved in ligand-binding are known (Ganesh *et al.*, 2011). Comparison of RF122 ClfB residues that differ from the solved structure suggested that residues involved in ligand-binding by a variation of the dock, lock and latch mechanism are conserved which was consistent with the dose-dependent binding of recombinant ClfB protein from RF122 to fibrinogen observed here. This is inconsistent with results in Chapter 3 which suggested that ClfB expression in RF122 is too weak to promote bacterial adhesion. Perhaps ClfB expressed by RF122 is functional but is either being masked by a larger protein and therefore does not have access to its ligands or perhaps it is subjected to degradation by metalloproteases or it is expressed too weakly. This warrants further investigation. As shown here the Bbp protein of RF122 contains a ClfB-like metalloprotease cleavage site between residues 247-251. Elimination of this cleavage site allows the expression of a stable recombinant protein which retains a HIS-tag. Purification of recombinant Bbp also facilitated the generation of Bbp-specific antibodies which were used in Chapter 3 to detect the expression of Bbp more efficiently than relying on immunocrossreactivity of anti-SdrE.

In order to demonstrate specificity in the murine mastitis model of infection, which was to be carried out by a cooperating laboratory, it was necessary to delete the genes encoding *isdA, isdB, clfB* and *bbp*. To delete the aforementioned genes it was necessary to first establish a system for genetically manipulating RF122. This led to work regarding the restriction-modification systems employed by RF122 which will be discussed in detail in Chapter 5. Once a system for genetic manipulation was established the temperature-sensitive shuttle plasmid pKOR1 was used and deletion mutants were isolated by allelic replacement in the *isdA, isdB, clfB and bbp* genes of *S. aureus* strain RF122.
As the in vivo work in this study was not carried out as intended, one cannot draw any conclusion as to whether or not a multicomponent vaccine based on the antigens used would be efficacious. If vaccination with the recombinant proteins (with or without adjuvants) proposed here did not elicit an efficacious immune response, a DNA vaccine coding for the proteins could be used alternately and may generate more significant results. If this is the case, the recombinant proteins could be used as boosters. Non-LPXTG vaccine candidates have been investigated to combat staphylococcal infections in humans such as α-toxin, lipoprotein and secreted proteins (Adhikari et al., 2012; Jongerius et al., 2012; Mishra et al., 2012). These could be considered for vaccination to combat bovine mastitis and will be discussed in Chapter 6 section 6.1.2.2.

Based on vaccine candidates proposed in Chapter 3, this chapter described the preparation of recombinant IsdB, ClfB and Bbp proteins for use in a future multicomponent vaccine trial to combat bovine mastitis. Due to previous successes, it is reasonable to include these proteins along with IsdA and SdrE. It is possible that Bbp, the allelic variant of SdrE will make a promising candidate as part of a multicomponent vaccine. IsdH could be considered as a candidate although it would not protect against one of the main group of pathogens in ST151. The recombinant proteins could also be used to test if the vaccine candidates are expressed during natural infection which would indicate their importance in pathogenesis. This would be achieved by incubation of the recombinant proteins with sera from convalescing cows and testing for their recognition by antibodies within the sera. The isogenic mutants constructed in this chapter will be useful in testing specificity of immune responses in any future vaccine trial where RF122 is the challenge strain.
Chapter 5

The restriction-modification system of *Staphylococcus aureus* RF122
5.1 Introduction

*S. aureus* strain RN4220 is widely used as the gateway for introducing shuttle plasmid DNA that has been constructed in *E. coli* into (some) clinical *S. aureus* strains. It was isolated when strain 8325-4 (8325 cured of three prophages) was subjected to heavy chemical mutagenesis and then transformed by protoplast transformation with an *E. coli*- *S. aureus* shuttle plasmid. A single colony that grew was propagated, the plasmid was eliminated and the derivative (RN4220) was then shown to be transformable with shuttle plasmids isolated from *E. coli* (Kreiswirth et al., 1983).

Although RN4220 is used extensively to facilitate genetic manipulation of *S. aureus*, it only provides a gateway into a limited set of closely related strains. DNA isolated from RN4220 is modified by the HsdMS component of the type I restriction system and can be transferred to wild type strains that have the same HsdRMS system. However, many *S. aureus* strains carry a distinct type I system for which RN4220 DNA is not modified. Failure to transform *S. aureus* strains devoid of a functional type I restriction endonuclease with plasmids isolated directly from *E. coli* led to the identification of a novel type IV restriction endonuclease SauUS1 (which will be described below) in *S. aureus*. SauUS1 recognises DNA isolated from *E. coli* that is cytosine methylated and constitutes the main primary barrier to transfer of DNA from *E. coli* (Corvaglia et al., 2010; Xu et al., 2011).

Waldron and Lindsay (2006) showed that by complementing the RN4220 mutant *hsdR* allele with wild type *hsdR* expressed from a plasmid, transformation by electroporation with a shuttle plasmid isolated from *E. coli* K-12 could be prevented, transduction was inhibited and the frequency of conjugation of a plasmid from *Enterococcus faecalis* was reduced. However, deletion of the *hsdR* gene in 8325-4 and COL alone was not sufficient to overcome the restriction barrier while mild heat shocking electrocompetent cells prior to electroporation allowed transformation of 8325-4Δ*hsdR* but not 8325-4 (Waldron & Lindsay, 2006; Veiga & Pinho, 2009). This suggested the presence of an additional restriction system in *S. aureus*.

Strains that are defective in restricting foreign DNA but are capable of modifying the newly introduced DNA (R-M⁺) can take up foreign DNA and modify it so that it can be transferred to closely related wildtype strains. Restriction-deficient mutants have been
isolated in the clonal complex (CC) 8 strain 8325 (SA113) and the CC51 strain 879 (879R4) (Iordanescu & Surdeanu, 1976; Stobberingh & Winkler, 1977).

The sauUSI gene of RN4220 contains a nonsense mutation in the middle of the gene, which explains why RN4220 can be transformed with foreign DNA. When the mutation was complemented, transformation into RN4220 was reduced 100-fold (Corvaglia et al., 2010). This also demonstrates that SauUSI is an important barrier to horizontal gene transfer between bacterial species. Restoration of the sauUSI gene in the chromosome of RN4220 to wildtype by allelic exchange resulted in a $10^{-4}$-fold reduction in the transformation frequency in RN4220sauUSI + compared to RN4220 (Monk et al., 2012). SauUSI acts independently of the type I RM system. This was inferred when deletion of the type I RM specificity genes hsdS1 and hsdS2 in SA564 did not yield a transformable strain (Corvaglia et al., 2010). In order to bypass the type IV restriction barrier in S. aureus the plasmid must be isolated from an E. coli strain that is defective in cytosine methylation.

Plasmid DNA that is not cytosine-methylated can be transferred directly into many S. aureus strains as it avoids degradation by SauUSI. In addition, the sauUSI gene is highly conserved among S. aureus clinical isolates. Knowledge of this system and how to avoid it represents an important development in the tools available for molecular analysis of previously untransformable clinical strains. A 50-fold improvement to the competence of S. aureus strain Newman was achieved using the Lofblom S. carnosus protocol with minor modifications (Monk & Foster, 2012). This even allowed transformants to be obtained directly with DNA isolated from wildtype dcm+E. coli K-12 strains at a low frequency (between $10^1$-10² CFU/ 5 µg plasmid DNA). In strain Newman the restriction barriers cause a $10^{-4}$ reduction in the transformation efficiency when comparing uptake of plasmid DNA isolated from wild-type Newman with that isolated from E. coli K-12. Thus both the type IV and type I RM systems could be bypassed by improving the efficiency of electroporation.

A dam mutant of E. coli is not suitable for cloning as the loss of Dam methylation results in deregulated mismatch repair and an increased frequency of transition mutations (Wion & Casadesus, 2006), whereas the mutation rate of a dcm mutant is not enhanced (Palmer & Marinus, 1994). A dcm mutant of a high efficiency cloning strain of E. coli would be a useful host for constructing recombinant plasmids prior to direct
transformation into a wild type SauUSI proficient strain of *S. aureus*. Thus a dcm mutant of *E. coli* K-12 DH10B was constructed and has been distributed widely (Monk *et al.*, 2012).

Little is known about the restriction systems of bovine *S. aureus* strains. Although RF122 and the hospital-acquired epidemic *S. aureus* MRSA252 are phylogenetically distinct from each other, and from other *S. aureus* strains, comparison of their genomes showed that RF122 and MRSA252 uniquely share 14 different DNA regions which are from 144 to 4,950 bp in length and exhibit between 93% and 99% pairwise identity, with most having 97% or greater identity. The shared regions exhibit features of horizontal gene transfer events. As the flanking vector sequences were, in most cases, in MRSA252 and not in RF122, it was inferred that MRSA252 or one of its related clones was the recipient in the DNA exchange. These two phylogenetically distinct strains most likely have undergone multiple gene transfers either between themselves or with as yet unidentified additional bacteria. Their pathogenicity-associated genes have undergone significant genetic divergence (Brody *et al.*, 2008).

A number of ST151 strains including RF122 were shown to have an enhanced ability to accept antibiotic resistance genes from enterococci. There are two *sauIhsdS* genes required for SauI activity. Each strain had identical nonsense mutations, one in each of the two *sauIhsdS* gene copies (Sung & Lindsay, 2007). The mechanism of the spread of an antibiotic resistance gene from enterococci to *S. aureus* was first described by Clewell *et al.* in 1985 (Clewell *et al.*, 1985). *S. aureus* produces a lipoprotein signal which triggers conjugation between *S. aureus* and *E. faecalis*. Transposons, harboured by large pheromone-responsive conjugative plasmids of *E. faecalis*, can jump to the *S. aureus* chromosome. The plasmid is unable to replicate in *S. aureus* and is lost. The SauI system blocks uptake of DNA from *E. coli* and reduces uptake from enterococci. In addition, it prevents the transfer of DNA between the dominant lineages of *S. aureus*, which each have unique *sauIhsdS* gene variants which encode variant restriction systems that degrade DNA modified by isoforms (Sung & Lindsay, 2007).

Although they did not state the name of the gene, Xu *et al* mentioned that RF122 possesses a gene that has low homology to type II RM endonucleases. They stated that RF122 possesses a *dcm* gene which is likely to encode a sequence-specific DNA methylase. Dcm shares 62% amino acid sequence identity to Bsp6I of *Mycoplasma agalactiae*. Dcm methylates the internal cytosine residue in the sequence CC(T)GG to 5-methylcytosine. DNA modified in this way can potentially be cleaved by SauUSI.
modified plasmid DNA may be cleaved when transferred into *S. aureus* USA300 strain or other *Staphylococcus* species with active SauUSI-like restriction endonucleases (Xu *et al.*, 2011).

Observations in Chapter 4 indicated that RF122 cannot be manipulated genetically using conventional methods and that it does not accept DNA from RN4220. As it has been reported that ST151 strains accept DNA from enterococci this suggests that DNA from enterococci is modified for transfer into RF122 whereas DNA obtained from RN4220 is not, as it has a different modification system. It is likely that DNA from enterococci and RF122 exhibits the same methylation pattern. Therefore to manipulate RF122 genetically it is important to understand the restriction systems it employs and to overcome them. This chapter describes this investigation. To identify such barriers it was necessary to perform genome comparisons, to delete possible candidates using allelic exchange and to test the mutant strains for loss of restriction. To characterise the novel gene it was necessary to clone it into an expression vector in *E. coli* to allow expression of the recombinant protein.
5.2 Results

5.2.1 Restriction-modification genes of RF122

As previously documented, RF122 contains two genes which may be involved in restriction-modification, a putative cytosine methylase called \textit{dcm} which has homology to the modification methylase Bsp61 from \textit{Mycoplasma agalactiae} and a gene whose name was not stated, which has homology to a type II restriction endonuclease (Xu \textit{et al.}, 2011). To investigate this further, the genome sequences of RF122 and 8325-4 were compared.

Analysis of the RF122 gene sequences provided in the GenBank database revealed a gene called \textit{sab2370c} which is located immediately downstream of \textit{dcm} (annotated as \textit{sab2369c}). A blast search with the hypothetical protein gene sequence using the discontiguous megablast algorithm which is optimised for more dissimilar sequences (as detailed in Chapter 2 section 2.2.2) revealed that \textit{sab2370c} was not found in any other staphylococcal strain but had homology to genes that encode hypothetical proteins from a number of other species including \textit{Streptococcus salivarius}, \textit{Strep. gordonii} and \textit{Strep. suis}. Another blast search using \textit{dcm} from RF122 revealed that in addition to homology to the gene encoding the modification methylase Bsp61 from \textit{M. agalactiae}, it was homologous to genes encoding modification methylases of a number of other species including \textit{M. bovis}, \textit{Strep. salivarius} and \textit{Strep. gordonii}. Similarly to \textit{sab2370c}, the \textit{dcm} gene is not found in any other staphylococcal strain.

To investigate how \textit{dcm} and \textit{sab2370c} might have been acquired, as they are not found in other staphylococcal strains, the regions flanking these genes were analysed. The sequence immediately upstream of RF122 \textit{dcm} contains a pseudogene called \textit{sab2368c} (Figure 5.1). A Blast search with \textit{sab2368c} revealed homology to genes found in several other strains including 8325-4. In 8325-4 it is homologous to a gene which encodes a hypothetical protein with the gene locus tag SAOUHSC_02788. Analysis of the flanking regions in the 8325-4 genome revealed that this gene is located upstream to \textit{sauUSI} ( locus tag SAOUHSC_02790), separated by a gene encoding a conserved hypothetical protein (gene locus tag SAOUHSC_02789). The \textit{sauUSI} gene was not found in the RF122 genome.
Figure 5.1 Genomic organisation comparison of restriction genes of *S. aureus* RF122 and 8325-4

The restriction genes of *S. aureus* RF122 and 8325-4 found on the Pubmed database were compared. The dashed lines highlight genes that share homology as determined using the online Blast tool.
Analysis of the sequence immediately downstream of RF122 sab2370c revealed a gene called sab2371 which encodes a phosphoglucomutase. A Blast search with sab2371 revealed homology to genes of several other strains including a gene which encoded a conserved hypothetical protein (gene locus tag SAOUHSC_02793) found in the genome of 8325-4. This gene was located downstream of sauUSI, separated by a gene which encoded a putative pyrophosphohydrolase (gene locus tag SAOUHSC_02791). The results of this analysis are represented diagrammatically in Figure 5.1. Taken together this suggests that a fragment comprising three genes, including sauUSI was replaced with a fragment comprising dcm and sab2370c in the RF122 genome.

5.2.2 The use of RF122t as an intermediate strain to facilitate genetic manipulation of RF122

Until recently the conventional method of manipulating S. aureus strains genetically involved the use of a restriction deficient S. aureus strain called RN4220. DNA transformed from E. coli K-12 is modified and can be transferred to the target strain. Once plasmid constructs are transformed into the target strain the gene is subsequently deleted by allelic replacement. It was initially concluded, as mentioned in Chapter 4, that RF122 does not accept DNA from RN4220. Further investigation revealed that this is not the case and will be discussed in section 5.2.6. However, at the time this was the accepted observation and an electrocompetent derivative of RF122 referred to as RF122t can act as a gateway from E. coli K-12 and has been used to allow the transformation of plasmid DNA into RF122 (Figure 5.2), albeit at a low frequency.

The serendipitous discovery of RF122t occurred when pRN6684 was transduced into it in order to make a tst mutant and it was inadvertently selected for. The nature of the mutation resulting in its restriction-deficient phenotype is not known. It does not differ from RF122 in terms of growth rate/yield, exponential/stationary SDS PAGE profile, haemolytic activity or virulence in experimental bovine intramammary infection (Ross Fitzgerald, personal communication).

In an attempt to determine the nature of the mutation that allows RF122t to accept DNA from E. coli, the dcm and sab2370c genes were amplified from S. aureus strain RF122t, sequenced and compared to those from RF122. The dcm gene was intact and was identical to that found in RF122. This shows that the ability of RF122t to accept DNA
from *E. coli* is not due to a defect in *dcm*. The *sab2370c* gene from RF122t is similar to RF122. A 5 bp deletion at the 3' end causes a frameshift resulting in the C termini of the two proteins to differ (Figure 5.3 A). The bases at the 3' end of RF122 translate to the amino acids WVGK followed by a stop codon whereas the bases at the 3' end of RF122 translate to WEI followed by a stop codon (Figure 5.3 B). The stop codons differ because of an additional base substitution 7 nucleotides downstream from the deletion. Alignment of the DNA sequence downstream of the *sab2370c* gene from RF122 and RF122t show that apart from the aforementioned nucleotides, the sequences are identical (Figure 5.3 A). The *sab2370c* sequence from RF122t was determined using two separate sequencing companies.

### 5.2.3 Determining the distribution of *sab2370c*, *dcm* and *sauUSI* by Southern dot blot hybridisation

To determine if, like RF122, the majority of bovine strains possess *sab2370c* and *dcm*, their distribution in a panel of genetically distinct bovine strains was tested by Southern hybridisation using specific DIG-labelled probes. Specific probes were constructed as described in Chapter 2 section 2.12.2. To determine if other strains from ST151 carry *sab2370c* and *dcm*, four isolates from ST151 (RF80, RF113, RF114 and RF121) were included. Genomic DNA from RF122 and Newman were included as positive and negative controls, respectively. Equal amounts of DNA were spotted onto a positively charged membrane.

Probes were designed to eliminate cross-reaction with homologous genes. This was determined by performing BLAST searches using the online BLAST tool as mentioned in Chapter 2 section 2.2.2. There are three BLAST algorithms one can use to predict sequence homology based on highly similar sequences (megablast), more dissimilar sequences (discontiguous megablast) or somewhat similar sequences (blastn). Probes were designed taking into account more dissimilar and somewhat similar sequences.

Figure 5.4 shows that each ST151 strain tested possessed *sab2370c*. ST126 and ST694 also possessed *sab2370c*. The ST126 spot was slightly less intense compared to the other spots. This difference in intensity was observed in three separate experiments.
Figure 5.2 The use of intermediate *S. aureus* strains to deliver DNA into target *S. aureus* Newman or RF122 strains

DNA from *E. coli* strain XL-1 can be transformed into RN4220 and subsequently into Newman. RF122 will not accept DNA from RN4220. RF122t will accept DNA from *E. coli* strain XL-1 and this DNA can be subsequently accepted by RF122.
A

RF122_end  CCAATGATGAATTGTTGGGTTGGGAAATAGAACATATGTACGTATTCGAGATAGTGT  60
RF122t_end  CCAATGATGAATTGTTGG----------GAAATATAACATATGTACGTATTCGAGATAGTGT  55

RF122_end  GGTATAATCTTGAAATGATGACTGATTCTGTAAGT  120
RF122t_end  GGTATAATCTTGAAATGATGACTGATTCTGTAAGT  115

B

(i) RF122

ccaaatgatgtaattggttggttggaatatag
P N D V I G W V G K -

(ii) RF122t

ccaaatgatgtaattggttggttggaatatataa
P N D V I G W E I -

Figure 5.3  A. DNA sequence alignment of the 3' end of sab2370c from S. aureus strains RF122 and RF122t.

DNA sequence alignment of the end of sab2370c from S. aureus strains RF122 and RF122t.

* indicates an identical residue while - indicates the position of a nucleotide that only occurs in the sequence to which it is compared to.

B. The predicted amino acid translation of the end of sab2370c from S. aureus (i) RF122 and (ii) RF122t. – indicates a stop codon.
### Figure 5.4 Presence of \textit{sab2370c} in \textit{S. aureus} strains by Southern dot blot hybridization

Genomic DNA from \textit{S. aureus} isolates was purified and 1 μg was spotted onto nitrocellulose membranes and probed with a DIG-labelled \textit{sab2370c}-specific probe. The ST of each bovine \textit{S. aureus} strain tested is indicated. RF122 and Newman were included as positive and negative controls, respectively. In addition to the positive controls, the top line includes bovine \textit{S. aureus} strains from ST151 (RF80, RF113, RF114 and RF121). This is representative of 3 independent experiments.
and indicates sequence variation. This could be verified by sequencing the \textit{sab2370c} gene from ST126 and ST694. None of the other bovine strains possessed \textit{sab2370c}.

Figure 5.5 shows that all of the ST151 strains tested possessed the \textit{dcm} gene. ST126 and ST694 also carried \textit{dcm} while none of the remaining strains tested possessed it. The spot corresponding to \textit{dcm} from the ST151 \textit{S. aureus} strain RF121 was less intense than spots from other strains suggesting sequence variation which could be confirmed by sequencing the \textit{dcm} gene from that strain. Thus a strain either carried both \textit{sab2370c} and \textit{dcm} or did not carry either gene.

It seemed likely that the \textit{sauUSI} gene was replaced by \textit{dcm} and \textit{sab2370c}. Thus the strains were also probed for presence of \textit{sauUSI}. Genomic DNA from \textit{S. aureus} RF122 and Newman constituted the negative and positive controls, respectively. As Figure 5.6 shows, none of the other ST151 strains possessed \textit{sauUSI}. The majority of the other bovine strains possessed \textit{sauUSI}. The intensity of the hybridisation was similar to that of Newman. ST20, ST25, ST126, ST479 or ST694 did not possess \textit{sauUSI}. ST126 and ST694 did not possess \textit{sauUSI} and they possess both \textit{sab2370c} and \textit{dcm}. However, ST20, ST25 and ST479 do not possess any of the genes probed for here. Either they have undergone considerable sequence variation and cannot be detected using these conditions or they possess other restriction genes or none at all. This warrants further investigation.

5.2.4 \textbf{Allelic replacement mutagenesis of \textit{sab2370c}}

A deletion mutation was isolated in the \textit{sab2370c} gene of strain RF122 by allelic replacement. The mutation deleted all but the start codon at the 5' end of the gene and the stop codon at the 3' end, creating a 1587 bp unmarked deletion. A 503 bp upstream fragment (AB) located 5' to the \textit{sab2370c} gene and 503 bp downstream fragment (CD) covering the 3' end of the gene were amplified by PCR, using the primer pairs A + B and C + D (Figure 5.7 A (i)). Primer C had a 20 bp extension that complemented the end of the AB fragment which facilitated amplification of the ABCD deletion fragment with primers A + D (Figure 5.7 A (ii)). The ABCD deletion fragment was cloned into pKOR1 by gateway cloning, as described in Chapter 2, generating the plasmid pKOR1\textit{Δsab2370c}. To allow for allelic replacement mutagenesis of the \textit{sab2370c} gene, pKOR1\textit{Δsab2370c} was isolated first from \textit{E. coli} XL-1 Blue and then electroporated into the intermediate
strain *S. aureus* RF122t and then transferred into *S. aureus* strain RF122. The gene was then deleted by allelic replacement as described in Chapter 2 section 2.6.6.

Using the primer pair out F + out R, which hybridize to chromosomal DNA flanking *sab2370c*, two outcomes were possible when genomic DNA of putative mutants was amplified which either indicates restoration of the wildtype (Figure 5.7 B (i) - 3589 bp fragment) or a successful deletion mutation (Figure 5.7B (ii) – 1996 bp fragment). Amplification of a 3589 bp fragment showed that the gene was still present and therefore the wildtype *sab2370c* had been restored (Figure 5.7B (iii) lane 1) whereas amplification of a 1996 bp fragment indicated successful deletion of the *sab2370c* gene (Figure 5.7B (iii) lane 2).

### 5.2.5 Rendering RF122 electrocompetent

It was shown in our lab that a protocol for rendering *S. carnosus* cells electrocompetent yielded increased transformation frequencies in many *S. aureus* strains from different STs compared to the conventional protocol for preparing electrocompetent *S. aureus* cells (Monk & Foster, 2012). However, RF122 could not be made competent to take up plasmid pRMC2 from *E. coli* using this protocol. The original protocol for preparing electrocompetent cells using sucrose was then used to prepare electrocompetent RF122 cells but only yielded 10-15 CFU/µg plasmid DNA. It was hoped that combining the approaches from the two protocols would increase the transformation frequencies in RF122.

The combined protocol used is depicted in Figure 5.8 and detailed in Chapter 2 section 2.8.1. The protocol is largely based on the sucrose protocol with a few modifications. The volumes used for the washes reflect those used in the *S. carnosus* protocol. Using larger volumes of sucrose wash solution allows more efficient removal of salt. The final volume in which the cells are resuspended in also reflects that used in the *S. carnosus* protocol. Using a higher concentration of cells increases the transformation frequency. Using this combined approach, the competence of RF122 was increased to 100-200 CFU/µg (data not shown).
Figure 5.5 Presence of \textit{dcm} in \textit{S. aureus} strains by southern dot blot hybridization

Genomic DNA from \textit{S. aureus} isolates was purified and 1 µg genomic DNA was spotted onto nitrocellulose membranes and probed with a DIG-labelled \textit{dcm}-specific probe. The ST of each bovine \textit{S. aureus} strain tested is indicated. RF122 and Newman were included as positive and negative controls, respectively. In addition to the positive controls, the top line includes bovine \textit{S. aureus} strains from ST151 (RF80, RF113, RF114 and RF121). This is representative of 3 independent experiments.
<table>
<thead>
<tr>
<th>RF122</th>
<th>Newman</th>
<th>RF80</th>
<th>RF113</th>
<th>RF114</th>
<th>RF121</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST8</td>
<td>ST20</td>
<td>ST25</td>
<td>ST71</td>
<td>ST97</td>
<td>ST115</td>
</tr>
<tr>
<td>ST133</td>
<td>ST350</td>
<td>ST352</td>
<td>ST479</td>
<td>ST693</td>
<td>ST694</td>
</tr>
<tr>
<td>ST695</td>
<td>ST696</td>
<td>ST697</td>
<td>ST698</td>
<td>ST699</td>
<td>ST708</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ST709</td>
</tr>
</tbody>
</table>

**Figure 5.6  Presence of sauUS1 in *S. aureus* strains by southern dot blot hybridization**

Genomic DNA from *S. aureus* isolates was purified and 1 μg was spotted onto nitrocellulose membranes and probed with a DIG-labelled *sauUS1*-specific probe. The ST of each bovine *S. aureus* strain tested is indicated. RF122 and Newman were included as negative and positive controls, respectively. In addition to the positive controls, the top line includes bovine *S. aureus* strains from ST151 (RF80, RF113, RF114 and RF121). This is representative of 3 independent experiments.
Figure 5.7  A. Schematic representation of the RF122 sab2370c mutation
Generation of the ABCD deletion fragment. (i) Diagrammatic representation of construction of the sab2370c ABCD deletion fragment. A 503bp AB fragment 5’ to the sab2370c gene and a 503bp fragment at the 3’ end of the gene. (ii) Plasmid DNA was isolated and screened for the presence of the 1006 bp ABCD deletion fragment indicating successful cloning in pKOR1.

B. Validation of RF122Δsab2370c (i) Diagrammatic representation of PCR screening products to identify (i) wildtype or (ii) mutated strains deficient in sab2370c. (iii) Diagnostic PCR with primer pair out F and out R which are directed against chromosomal DNA flanking sab2370c. Lane 1; wildtype fragment, lane 2; mutant fragment.
Cells were grown overnight in 2 ml TSB in a test tube, subcultured into 100 ml fresh TSB in a 250 ml flask and grown to OD_540nm of 0.2. Cells were chilled on ice for 10 mins and harvested by centrifugation at 3000 x g. Cells were washed twice in ice-cold sterile sucrose (0.5 M). Cells were washed and resuspended in successively decreasing volumes, aliquoted and stored at -70°C.
5.2.6 Transformation

In order to determine if Sab2370c is responsible for restricting the uptake of DNA into RF122, pRMC2 DNA was isolated from *E. coli* XL-1 and *S. aureus* strains RN4220 and RF122. Although it was previously concluded in this study that RF122 does not accept DNA from *S. aureus* RN4220 and *E. coli*, this was found not to be the case. Figure 5.9 shows that using the modified protocol and an increased amount of DNA (5 μg instead of 1 μg) transformation of DNA from both *E. coli* XL-1 and *S. aureus* RN4220 occurred, albeit at a low frequency (~20 CFU/μg).

Plasmid DNA isolated from RF122 itself (ie modified for wildtype RF122) provided the maximum number of transformants and is the benchmark for comparison with DNA from other sources. RF122 lacking sab2370c can accept DNA from both *E. coli* XL-1 and *S. aureus* RN4220 at levels almost comparable to the levels shown with pRMC2 DNA isolated from wildtype RF122 demonstrating that Sab2370c is at least partly responsible for restricting the uptake of DNA into RF122.

5.2.7 Attempted deletion of the *dcm* gene in RF122

In an attempt to delete the *dcm* gene from RF122, a pKOR1Δdcm construct was generated. A 498 bp AB upstream fragment located 5' to the *dcm* gene and 500 bp CD downstream fragment covering the 3' end of the gene were amplified by PCR using the primer pairs A-B and C-D. Primer C had a 21 bp extension that complemented the end of the AB fragment which facilitated amplification of the ABCD deletion fragment. The ABCD deletion fragment was cloned into pKOR1, as described in Chapter 2 section 2.6.5, generating plasmid pKOR1Δdcm. After several attempts to delete the *dcm* gene by allelic replacement it was observed that the wildtype was restored in every case (40 candidates were screened). This suggests that *dcm* cannot be deleted by allelic replacement and that such a mutation might be lethal.

To investigate whether the plasmid containing the deletion fragment was merely failing to integrate into the chromosome, primers were designed that would allow a PCR-based detection system for integrants. Using the primer pairs “Out F + D” two possible fragment sizes would be amplified depending into which side of the gene the plasmid had integrated into. If plasmid integration occurred on the AB side a fragment of 1518 bp
would be amplified (Figure 5.10A). Alternately if the plasmid integrated on the CD side a fragment of 2727 bp would be amplified (Figure 5.10B). The two possible outcomes with primer pair “A + out R” were 2754 bp (AB side) or 1545 bp (CD side) (Figure 5.10A and B, respectively).

In each of the five cases shown these primer pairs amplified fragments of 2727 bp and 1545 bp, respectively, indicating that plasmid integration occurred on the CD side (Figure 5.10C). The pKOR1Δdcm plasmid was also tested for temperature sensitivity and was integrated into the chromosome at the same frequency as pKOR1Δsab2370c (data not shown). This further suggests that the inability to delete the dcm gene is because this would be lethal. This could be investigated by construction of a conditional mutant (which will be discussed in section 5.3).

5.2.8 Cloning sab2370c into an expression vector in E. coli

In order to investigate the phenotype associated with Sab2370c it was necessary to clone the sab2370c gene into an expression vector to facilitate the purification of recombinant Sab2370c protein. The 1593 bp sab2370c gene was cloned into pQE30 using the SacI and PstI restriction sites and successful cloning was verified by restriction analysis (Figure 5.11) and DNA sequencing.

5.2.9 Sab2370c alters the growth of E. coli cells.

In order to purify His-tagged recombinant protein from pQE30 it is necessary to induce expression from the IPTG-inducible promoter with IPTG. However, recombinant Sab2370c was not expressed as determined by probing lysates with Anti-His-horseradish peroxidase. It was observed that the turbidity of the cells in the induced sample was much less dense than its uninduced counterpart. To investigate this further growth was measured up to and after addition of IPTG. Samples were taken at one hour intervals following addition of IPTG. Figure 5.12 shows that when the expression of Sab2370c was induced growth of E. coli was stopped. It is possible that DNA in E. coli is being degraded by the induced Sab2370c.
Figure 5.9 Transformation of RF122 and RF122Δsab2370c with pRMC2

5 µg of pRMC2 DNA isolated from *E. coli* XL-1-Blue, RN4220 or RF122 was electroporated into either RF122 or RF122Δsab2370c. The number of transformants (CFU)/ 5µg is recorded.
Figure 5.10 Integration of pKOR1Δdcm into the RF122 chromosome

A. Integration of plasmid before the gene (AB side). Primer pair “out F and D” produces a small fragment and primer pair “A and out R” produces a large amplimer.

B. Integration of plasmid after the gene (CD side). Primer pair “out F and D” produces a large fragment and primer pair “A and out R” produce a small amplimer.

C. PCR colony screen testing for integration of pKOR1Δdcm into the RF122 chromosome. The blue bars represent fragments amplified using primer pair “out F and D” whereas magenta bars indicate fragments amplified using primer pair “A and out R”. Samples 1-5 are indicated.
Figure 5.11  A. Cloning sab2370c DNA from S. aureus RF122 for the expression of Sab2370c

DNA encoding sab2370c was amplified from the genomic DNA of S. aureus strain RF122. The PCR product was cloned into the E. coli expression vector pQE30 using SacI and PstI restriction sites. A pQE30 construct was transformed into the E. coli strain XL-1 for the expression of recombinant His-tagged Sab2370c protein.

B. Restriction analysis of the pQE30::sab2370c construct. The pQE30 plasmid containing the coding region for Sab2370c was cut with SacI and PstI restriction enzymes. The double digestion of plasmid DNA yielded two distinct bands corresponding to the pQE30 backbone (3.4 kb) and the Sab2370c-encoding insert (1593 bp).
Figure 5.12 Expression of recombinant Sab2370c stalls the growth 
_E. coli_ XL-1 cells

Cells were grown to exponential phase. The expression of Sab2370c was 
induced with the addition of IPTG. Induced and uninduced empty vector 
and uninduced pQE30::sab2370c samples were included as negative 
controls. The plasmid harboured by each strain is indicated.
5.2.10 Co-expression of Dcm and Sab2370c in *E. coli*

As shown in Figure 5.12 expression of rSab2370c stops the growth of *E. coli*. If DNA is degraded the gene is eliminated and no transcription or translation can occur. As cleavage by type II restriction endonucleases can be prevented by methylation, co-expression of Dcm may prevent DNA degradation by Sab2370c allowing growth of *E. coli* and expression of the gene, facilitating the purification of recombinant Sab2370c.

pIMK is a kanamycin-resistant, site-specific, integrative vector. pIMK was modified and three derivatives were constructed each which allow a different level of expression in *E. coli* (Monk *et al.*, 2008). pIMK2 allows constitutive overexpression, pIMK3 allows a high-level of IPTG-controlled gene expression and pIMK4 allows a low level of strictly IPTG-controlled gene expression (Monk *et al.*, 2008). Out of the three variants pIMK3 was selected for use in this study as it has a high level of IPTG-controlled gene expression. The coding region of the RF122 *dcm* methylase gene was cloned into the pIMK3 vector. It has a different plasmid replicon allowing co-existence with pQE30. Addition of IPTG would allow simultaneous induction of both Sab2370c expression from pQE30 and Dcm expression from pIMK3. Primers were designed using the *dcm* sequence from RF122 and a gene fragment encoding residues 1-403, was amplified. The forward primer was designed to incorporate the NcoI site into the 5' end of *dcm* so that the existing bases were changed from ACATGA to CCATGG. The 1209 bp PCR product was cloned into pIMK3 between the NcoI and PstI sites (Figure 5.13A) and the construct was verified by restriction analysis (Figure 5.13B). The construct was sequenced using primers which flank the multiple cloning site of pIMK3. Both pIMK3::*dcm* and pQE30::*hab2370c* were transformed simultaneously into *E. coli* DC10B, which possesses no intrinsic cytosine methylase activity, and grown on agar plates containing both Amp and Kan.

Following induction with IPTG, expression of rSab2370c could not be visualised on a Coomassie-stained gel or detected by Western blotting using anti-His serum. To determine if the expression of Dcm provided any protective effect to counteract putative degradation of *E. coli* DNA by Sab2370c, a growth curve of induced and uninduced cultures was performed in conjunction with viable counting. DC10B carrying pQE30::*hab2370c* and pIMK3 with or without *dcm* were used as positive and negative controls, respectively. The cells were allowed to grow to mid-exponential phase, at which point IPTG was added then at 20 min intervals the OD$_{600\text{nm}}$ of samples was measured and
viable counts were determined. Viable counts are shown in Table 5.1. The data is representative of three independent experiments.

The following strains were also tested; DC10B carrying empty pQE30 and plMK3, pQE30::sab2370c and empty plMK3 or empty pQE30 and plMK3::dcm but the differences observed between the viable counts of both induced and uninduced samples were insignificant and are not shown here. As there is no methylase activity carried out by the E. coli strain, any protective effect caused by methylase activity can be attributed to dcm.

When the expression of Sab2370c was induced alone, growth of E. coli cells stalled compared to the uninduced controls (Figure 5.14). Measurement of viable counts during preliminary experiments indicated that after 20 mins no CFUs were recovered at $10^{2}$ dilution. A lawn of bacteria was observed at the same dilution in the uninduced control (data not shown). When both Sab2370c and Dcm were induced simultaneously the growth of E. coli cells stalled (Figure 5.14). The viable count dropped by 100 x after 20 min but some cells remained viable (Table 5.1) suggesting that Dcm protects some of the cells from death but does not allow replication. This explains why there was not enough cell growth to extract sufficient rSab2370c for analysis. Perhaps there is not enough expression of Dcm to completely protect the E. coli DNA from degradation by Sab2370c. It may be necessary to express more Dcm than Sab2370c using different vectors containing IPTG-inducible promoters which express different levels of protein.
Figure 5.13  A. Cloning dcm from RF122 into pIMK3

DNA encoding dcm was amplified from the genomic DNA of S. aureus strain RF122. The PCR product was cloned into the E.coli expression vector pIMK3 using Neol and PstI restriction sites. A pIMK3 construct and a pQE30 construct carrying sab2370c were transformed into the E. coli strain DC10b for the co-expression of the two recombinant proteins. Heterologous gene expression is enable through in-frame cloning of promoterless genes via a unique Neol site overlapping the start codon. Salient features include; KanR – resistance to kanamycin, p15A – origin of replication, Pp60 – Listerial P60 promoter, PSA integrase – Directs integration into tRNA^ARG locus, PSA attPP’ – Listerophage U153 integrase gene and attachment site, Phelp – promoter, lacOid sequence – LacI binding site, rnr B T1 terminators – transcription terminators. Adapted from Monk, 2008.

B. Restriction analysis of the plMK3::dcm construct. The plMK3 plasmid containing the coding region for Dcm was verified by digestion with Neol and PstI restriction enzymes. The double digestion of plasmid DNA yielded two distinct bands correspond to the plMK3 backbone (7.5 kb) and the Dcm-encoding insert (1209 bp) as indicated.
<table>
<thead>
<tr>
<th>Plasmids</th>
<th>0 min</th>
<th>20 min</th>
<th>40 min</th>
<th>60 min</th>
<th>24 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>sab^- dcm^+</td>
<td>8.6 x 10^8</td>
<td>8.66 x 10^8</td>
<td>5 x 10^8</td>
<td>1.6 x 10^9</td>
<td>8.33 x 10^9</td>
</tr>
<tr>
<td>sab^+ dcm^+ + IPTG</td>
<td>8.6 x 10^8</td>
<td>5.6 x 10^6</td>
<td>8.66 x 10^6</td>
<td>3.33 x 10^6</td>
<td>7 x 10^7</td>
</tr>
<tr>
<td>sab^+ dcm^-</td>
<td>4.3 x 10^8</td>
<td>5.6 x 10^8</td>
<td>7.92 x 10^8</td>
<td>1.6 x 10^9</td>
<td>8 x 10^8</td>
</tr>
<tr>
<td>sab^- dcm^- + IPTG</td>
<td>4.3 x 10^8</td>
<td>&lt; 1 x 10^2</td>
<td>&lt; 1 x 10^2</td>
<td>&lt; 1 x 10^2</td>
<td>&lt; 1 x 10^2</td>
</tr>
</tbody>
</table>

^a E. coli strain DC10B carrying sab2370c and/or dcm

^b CFU/ml at 0 m, 20 m, 40 m, 60 m or 24 h after addition of 1mM IPTG
Figure 5.14 Growth curve of *E. coli* DC10B strains

DC10B cells harbouring either pQE30 and pIMK3 (sab- dcm-), pQE30::~sab2370c and pIMK3 (sab+ dcm-), pQE30 and pIMK3::dcm (sab- dcm+) or pQE30::sab2370c and pIMK3::dcm (sab+ dcm+) were grown to exponential phase. The expression of proteins was induced with the addition of IPTG. Optical densities were measured after 1, 2 and 3 hours. Uninduced samples were included as negative controls. Data is representative of three independent experiments.
5.3 Discussion

In contrast to typical *S. aureus* strains, in which the major barrier to DNA uptake is SauUSI, RF122 employs a restriction system comprising two genes which were putatively acquired together, as indicated by comparing the genomes of RF122 and 8325-4. The putative cytosine methylase (*dcm*) is not found in any other *S. aureus* strain but has homology to the modification methylases from a number of other species. Similarly, *sab2370c* is not found in any other *S. aureus* strain but has homology to a type II restriction endonuclease (Xu *et al.*, 2011) and to hypothetical proteins from a number of other species. This suggests that RF122 has acquired *dcm* and *sab2370c* by horizontal transfer from another bacterial species.

The *dcm* and *sab2370c* genes are possessed by other strains of ST151 and two other bovine strains (ST126 and ST694), as shown by Southern hybridisation. With reference to Figure 3.16 in Chapter 3 which shows the phylogenetic relatedness of the bovine strains used in this study, ST126 and ST694 are next to each in the phylogenetic tree. This suggests that they came from a common ancestor and the acquisition of *sab2370c* and *dcm* did not occur independently. The majority of bovine STs do not possess either gene but do possess *sauUSI*. This is consistent with observations in Chapter 3 which suggested that RF122 has adapted to a novel niche. RF122 is most likely derived from a strain that originated in humans but has adapted to life in the bovine mammary gland where a number of genes are likely to be redundant (Herron-Olson *et al.*, 2007).

Although RF122 and other ST151 strains have lost the ability to degrade cytosine methylated DNA through the loss of *sauUSI*, they have evolved in an advantageous way through the acquisition of *sab2370c* and *dcm*, in that they can protect their DNA from degradation whilst degrading any foreign DNA. As neither of these genes are possessed by other *S. aureus* strains but are possessed by other species this may provide opportunities to acquire genes that are beneficial in the mammary gland from more diverse sources.

When the expression of *Sab2370c* was induced in *E. coli* the cells stopped growing and died. As *Sab2370c* has homology to a type II restriction endonuclease it is reasonable to suspect that it degrades DNA in the *E. coli* host. It was plausible to expect that this could be counteracted by co-expression of *Sab2370c* and Dcm as Dcm would
methylate cytosines and might protect the cell's DNA from degradation by Sab2370c thus allowing it to grow sufficiently so that Sab2370c could be purified.

When the expression of both proteins was induced the cells were only partially protected from death and growth was still stalled suggesting that there was insufficient Dcm to protect cells. As the level of Dcm expression was not measured one can only speculate as to why a low level of protection was observed. Perhaps in vivo they are expressed at different levels to allow sufficient protection from Sab2370c whilst allowing the cells to replicate or they are subjected to regulation by an unknown factor. It may be necessary to determine the relative levels of expression that occur in vivo and match them to achieve enough growth for extracting recombinant Sab2370c protein for analysis. Measuring the transcription/translation levels may elucidate this.

It is unlikely that sab2370c exists without dcm which would explain why, in RF122, sauUS1 is replaced by both genes. This is consistent with the fact that the sab2370c gene could be deleted whereas the dcm gene could not. Deletion of sab2370c resulted in a partial breaking of the RF122 restriction barrier thus allowing RF122 to accept DNA from E. coli XL-1 and S. aureus RN4220 at an increased frequency. This could have been confirmed by complementing the mutant with the sab2370c gene. RF122Δsab2370c is a useful strain for further genetic manipulation of RF122 and related strains.

To allow deletion of sab2370c, an electrocompetent derivative of RF122 which is referred to as RF122t was used. This strain accepts DNA from E. coli and modifies it for transfer into RF122. The nature of the mutation in RF122t was investigated. The dcm and sab2370c genes were sequenced in an attempt to understand the restriction systems that were inactivated in this strain. Whereas RF122t possesses an intact dcm gene, identical to that of RF122, a small deletion at the end of its sab2370c gene resulted in a frameshift which may provide an explanation as to why it can accept DNA. To elucidate whether RF122t sab2370c is functional despite its truncation the effects on E. coli growth by expression of sab27370c cloned from RF122t could be tested. Alternately, the RF122t sab2370c gene could be replaced in the chromosome with sab2370c from RF122 by allelic replacement and tested for loss of DNA uptake.

The dcm mutation is likely to be lethal as it was not possible to introduce it. This supports the hypothesis that cells require it to protect it from Sab2370c. To confirm
whether deleting *dcm* is a lethal mutation it would be necessary to create a conditional mutation. Dubail et al constructed a conditional mutant. During a previous attempt to delete *imo2537* they created a merodiploid strain designated EGDpG_imo2537-imo2537. It was created by integrating the plasmid pGlts-A-Km-B into *L. monocytogenes* EGD-e. The pGlts-A-Km-B plasmid constitutes a 416-bp upstream fragment of Imo2537 (designated A) and a 499-bp downstream fragment (designated B) cloned into the pGIts shuttle vector. The pG1ts shuttle vector is a thermosensitive plasmid that carries both the Em\(^\text{r}\) gene of pE194 and the promoterless and terminatorless kanamycin resistance cassette *aphA-3* (Km). Although plasmid excision could not be achieved the integrant strain was used to facilitate the construction of a conditional mutant (designated EGD_imo2537/pLIV-imo2537). This mutant carries a chromosomal deletion of *imo2537* and a plasmid-borne wild-type *imo2537* allele under IPTG-inducible promoter control. The procedure allowed the spontaneous excision of the pG- recombinant plasmid carrying the wild-type allele (via a unique crossover). In brief, they amplified the wild-type *imo2537* gene preceded by its ribosome-binding site by PCR from *L. monocytogenes* (EGD-e) genomic DNA. The fragment carrying the entire gene *imo2537* was cloned into the pLIV1 shuttle plasmid, located immediately downstream of the IPTG-inducible SPAC promoter to yield plasmid pLIV-imo2537. Plasmid pLIV-imo2537 was introduced into strain EGDpG_imo2537-imo2537 (carrying a chromosomally integrated copy of pGlts-A-Km-B) by electroporation and selection for the acquisition of pLIV1-encoded chloramphenicol resistance (Cm\(^\text{r}\)) at 30°C and in the presence of 1 mM IPTG (the recombinant clones were thus Em\(^\text{r}\), Km\(^\text{r}\), and Cm\(^\text{r}\)). Clones were subjected to repeated growth at 30°C in the presence of Km (to keep the mutated chromosomal allele), chloramphenicol, and 1 mM IPTG (to keep the pLIV recombinant plasmid and allow cell viability) but in the absence of Erm (to allow loss of pG1). Subsequently they screened for excision of the chromosomally integrated pG- recombinant plasmid, leading to substitution of the wild-type *imo2537* allele with the Km\(^\text{r}\) cassette (Dubail et al., 2006). This strategy may facilitate the creation of a *dcm* conditional mutant.

RF121 from ST151 possesses a *dcm* gene that was detected weakly compared to RF122 using a *dcm*-specific probe by Southern hybridisation. This was observed three times using the same DNA concentration as used for control strains (RF122 and Newman). This indicates sequence variation. However, as it possesses a *sah2370c* gene which is detected with a similar intensity to RF122 its *dcm* gene is likely to be functional.
DNA sequence variation could have occurred any of the strains tested. To definitively show this it would be necessary to sequence the genes as Dot blotting is subjective.

Without the protection provided by Dcm, degradation by Sab2370c may occur. To show this it would be necessary to amplify and sequence dcm from this strain. Or perhaps the sab2370c gene is non-functional. Perhaps there is another protein expressed by this strain that provides protection. This warrants further investigation. To summarize, it seems that both Dcm and Sab2370c are necessary as Sab2370c restricts the uptake of DNA thus protecting the cell from foreign DNA while Dcm protects the cell from degrading its own DNA.

As previously mentioned the saul system is inactive in ST151 strains due to the occurrence of stop codons in each of the two saulhsdS gene copies. ST151 isolates are 500 times more susceptible to vancomycin resistance transfer from E. faecalis than human S. aureus isolates (Sung & Lindsay, 2007). As the Saul system has been shown to block the uptake of DNA from E. coli and reduce uptake from enterococci, it is therefore likely that ST151 strains including RF122 will accept DNA from both E. coli and enterococci.

This is consistent with the observation that RF122 accepts DNA from E. coli at a low frequency as seen in Figure 5.9 using the optimised protocol for preparation of electrocompetent cells. This frequency was increased when sab2370c was deleted. It is important to use the optimised sucrose protocol and a high concentration of DNA (at least 5 µg). Initially, the original protocol was used with only 1 µg of DNA, which did not yield any colonies. Electrocompetent RF122t was prepared using the original sucrose protocol and only 1 µg of DNA was added. Therefore it is likely that the low transformation frequency observed would increase using the new parameters.

Restriction systems have been investigated in other animal S. aureus strains. MRSA ST398 represents a distinct lineage that has rapidly spread worldwide (Uhlemann et al., 2012). It is mainly associated with animal colonization in livestock farming, with high prevalence rates in pigs and calves and until recently did not cause frequent invasive disease in humans. This may be due to the absence of several common virulence factors. The importance of horizontal gene transfer in bacterial genome remodelling was implied when the genomes of MRSA ST398 strain S035 and several other S. aureus strains were compared (Schijffelen et al., 2010). The majority of differences were identified in unique
or novel allotypes of mobile genetic elements, many of which harbour determinants for virulence and antimicrobial resistance which may allow the bacterium to adapt to new niches. The acquisition of determinants which contribute to virulence in human infections could have resulted from the enhanced ability of these isolates to acquire mobile genetic elements. This may also lead to the uptake of mobile elements that encode virulence genes. The first Panton-Valentine Leukocidin (PVL)-positive MRSA ST398 isolates have been reported (Schijffelen et al., 2010; Uhlemann et al., 2012).

Allelic variants of vSaa and vSaβ islands are present in some *S. aureus* genomes with each vSa island containing a type I restriction-modification (R-M) system. Horizontal gene transfer into a genome is prevented by the presence of two RM systems. This also limits DNA transfer between *S. aureus* isolates to those of the same lineage. Although both types of vSa islands are present in the S0385 genome, only vSaa harbors a type I R-M system, which has a unique *hsdS* gene encoding the sequence specificity S subunit. Strains that harbour a dysfunctional type I R-M system are significantly more prone to accept foreign DNA via conjugation (Schijffelen et al., 2010).

Another factor that may influence the acquisition of foreign DNA in the S0385 genome is the presence of three Integrative Conjugative Elements (ICEs), which encode proteins that form a type IV secretion-like system. This multi-protein complex can hydrolyze the peptidoglycan layer of the cell-wall and transfer single-stranded DNA through both the cell-wall and cell-membrane (Schijffelen et al., 2010).

A ST398 methicillin-susceptible *S. aureus* (MSSA) strain, which could colonize humans, recently emerged in northern Manhattan and was readily transmitted within households, independently of animal contact. It was observed that ST398 MSSA genomes are smaller than that of ST398 strain S0385 due to fewer mobile genetic elements (Uhlemann et al., 2012).

Similarly to RF122 which possesses a putative type II endonuclease (*sab2370c*), which has yet to be characterised, and a putative cytosine-specific DNA methylase (*dcm*), the ST398 isolates and LA-ST398 strain S0385 harbour a putative type II RM system, encoding a type II endonuclease (*SsoII*) and a novel cytosine-specific DNA methylase (*EcoRII*). The characteristic Smal resistance of genomic DNA from the ST398 clonal lineage is likely to be due to the possession of this methylase which is unique among *S. aureus* genomes. The closest relative of this methylase is found in *S. pseudintermedius*, a
zoonotic pathogen which is known mainly for being a commensal and pathogen in dogs but is sometimes associated with disease in humans (Uhlemann et al., 2012).

Consistent with the observation in RF122 and studies of disease-causing poultry S. aureus isolates, where a number of genes involved in pathogenesis of human S. aureus infections, such as S. aureus protein A (Spa), were rendered non-functional (Lowder et al., 2009), the observed gene decay in strain S0385 suggests that genes encoding certain adhesion proteins are likely no longer contributing to virulence in swine, and further imply that the most recent common ancestor of the ST398 lineage was human associated (Uhlemann et al., 2012).

To conclude, in this chapter it was revealed that, unlike the majority of bovine S. aureus strains, ST151 strains and two other bovine S. aureus STs do not possess the major barrier to restriction that is found in typical S. aureus strains. Instead they possess a putative cytosine-methylase called dcm and a putative type II endonuclease called sab2370c. Although these genes are not found in other S. aureus strains they have homology to restriction-modification genes found in other species. This has allowed speculation about how this successful strain has lost redundant DNA upon its transition to the mammary gland and how it has acquired DNA from a more diverse pool of bacteria. Optimisation of the preparation of electrocompetent cells facilitated an increased frequency of transformation revealing that RF122 can accept DNA at a low frequency from E. coli XL-1 and S. aureus RN4220. This is consistent with previous results where an artificial breaking of the restriction-modification barrier was observed upon utilisation of an optimised protocol for preparation of electrocompetent S. aureus Newman cells. 5 μg of plasmid DNA was transformed. Presumably when the cell is saturated with DNA, additional DNA cannot be processed by the restriction barrier (Monk et al., 2012). Deletion of sab2370c increased this frequency further revealing that Sab2370c has a role in restricting the uptake of foreign DNA. The proposed ability of Sab2370c to degrade foreign DNA was further supported by the inhibition of E. coli growth upon the induction of Sab2370c. Although this could not be relieved by the co-expression of Dcm to allow growth of E. coli, Dcm provided some protection. Understanding the restriction systems of RF122 is important and may facilitate further genetic manipulation of RF122.
Chapter 6

Discussion
6.1 Discussion

*S. aureus* can colonise and infect humans and a wide range of domesticated and wild animals. For example MRSA ST398 has rapidly spread worldwide and is mainly associated with animal colonization in livestock farming, with high prevalence rates in pigs and calves. Until recently it did not cause frequent invasive disease in humans (Schijffelen *et al.*, 2010). *S. aureus* is one of the most important etiologic agents of mastitis of cows, goats, and sheep. Bovine mastitis, the focus of this thesis, constitutes a serious problem in dairy herds with considerable economic consequences (Aires-de-Sousa *et al.*, 2007). RF122 represents a clone of very successful mastitis-causing strains. The first step in the development of *S. aureus* infections is the adherence to host tissues, a process facilitated by cell surface adhesins called MSCRAMMs.

6.1.1 Niche adaptation

The possession of a repertoire of MSCRAMMs has facilitated this complex pathogen in adapting to a wide range of hosts and to survive in these hosts. The ligand-binding specificities and thus the profile of MSCRAMMs exhibited by each strain is governed by the requirements of the environmental niche inhabited. Analysis of strain-specific components of genomes can reveal the nature of genetic events that have contributed to the emergence of new virulent strains. Virulence factors and resistance genes which are frequently exchanged within and between lineages are encoded by mobile genetic elements which are typically found in the accessory genome (Lindsay *et al.*, 2006). The accessory genome constitutes the most rapid and dramatic form of genetic adaptation (Hacker & Carniel, 2001). More long term adaptation is likely to be the result of a genome-wide combination of allelic diversification, lateral gene acquisition and loss of gene function (Ben Zakour *et al.*, 2008). Large-scale recombination events play a limited role in staphylococcal speciation compared with other pathogens such as Listeria spp. and Bacillus spp (Glaser *et al.*, 2001). Studies on the population structure of *S. aureus* have shown that most *S. aureus* clones are specific to a particular animal species, indicating host adaptation and restriction (Kapur *et al.*, 1995; Smith *et al.*, 2005). A genome-wide study discovered that much of the genetic variation specific for CC133 strains in comparison with sequenced human strains was found in the core region of the
genome. This indicates a role for diversification of the core genome in host adaptation (Ben Zakour et al., 2008).

6.1.1.1 Niche adaptation; gene decay

The most robust examples of gene decay come from recently emerged pathogens that have changed lifestyle, usually to live in a simpler host-associated niche (Moran & Plague, 2004). MRSA 252 has the highest content of pseudogenes compared to other S. aureus genomes (Holden et al., 2004; Lindsay & Holden, 2004). RF122 is a close second (Herron-Olson et al., 2007). In RF122, allelic variation was discovered in genes encoding proteins involved in host colonization, toxin production, iron metabolism, antibiotic resistance and gene regulation, which correlates with adaptation to a non-human host (Herron-Olson et al., 2007). Several genes encoding proteins known to be important in human infection were no longer functional due to premature truncation, suggesting the redundancy of these proteins for survival in cows (Herron et al., 2002; Herron-Olson et al., 2007).

Results in this thesis are consistent with these observations. Bioinformatic analysis of the genes encoding surface proteins supported the observation that gene decay has occurred in the RF122 genome. This was confirmed by expression studies which showed that missing genes or those that exist as pseudogenes were not expressed by RF122. This further suggests that several genes are redundant in the RF122 genome and that RF122 has adapted to change its lifestyle.

This was expanded to include other bovine strains of a wide range of genetically distinct strains. The majority of bovine strains still express several proteins that are required for life in the human host. Genetic adaptation to new niches occurs through a combination of foreign DNA acquisition and gene decay, similar to the human-to-poultry host switch of the successful human CC5 lineage (Lowder et al., 2009; Guinane et al., 2010). As RF122 exhibits more extensive gene decay in comparison to other bovine hosts this suggests that RF122 “jumped” to the bovine mammary gland long before other bovine strains.

Gene decay has also been reported in some CC30 strains. S. aureus clonal complex CC30 has caused infectious epidemics for more than 60 years. Out of the three major branches that evolved from this common ancestor, Clade 3 strains, which included
contemporary hospital associated MRSA (HA-MRSA) and clinical MSSA strains (exhibiting attenuated virulence), appeared to be following an evolutionary blueprint toward niche-adaptation as indicated by analysis of pseudogenes, SaPI and IS elements, and gene deletion events.

One of the common traits of Clade 3 includes a premature stop codon within *isdH* (McGavin *et al*., 2012). This study shows that RF122 and ST151 strains do not express IsdH whereas the majority of other bovine strains do. IsdH has previously been implicated as a good vaccine candidate (Ster *et al*., 2010). Strong and long-lasting immune responses with IgG2 production were observed when cattle were immunized with IsdH. Reduced mammary gland colonisation was observed in the mastitis mouse model when *S. aureus* was preincubated with serum against IsdH (Ster *et al*., 2010).

IsdH could be a promising vaccine candidate to combat bovine mastitis caused by the majority of bovine strains but not RF122, ST151 strains or CC30 Clade 3 strains which are likely to be resistant to this vaccine strategy. It was mentioned in Chapter 3 that RF122 secretes a truncated IsdH which includes the NEAT one domain. Whether or not this would be effective as a vaccine candidate warrants further investigation. Vaccines will be discussed below.

Many *S. aureus* MSCRAMMs are multifunctional. Redundancy often occurs with different MSCRAMMs performing similar functions. Those which have ligand binding specificities that are important in causing diseases in humans may not be important in the mammary gland. In addition to its role in iron acquisition, IsdH facilitates an enhanced conversion of C3b to iC3b and C3d which promotes the ability of *S. aureus* to survive neutrophil uptake and killing (Visai *et al*., 2009). The latter function must not be important to RF122 in the mammary gland. RF122 can still acquire iron from haemoglobin and haem as it expresses intact IsdA and IsdB.

As previously mentioned, the NEAT one domain of IsdH is secreted by RF122. The NEAT one domain has been shown to be responsible for binding haemoglobin and haptoglobin (Pilpa *et al*., 2009). A source of heme in bovine milk has been identified (α-casein binds both iron and heme) (Shibuya *et al*., 2012). It has been shown that heme can be passed in a unidirectional manner from IsdB NEAT domain 2 or IsdH NEAT domain 3 to the NEAT domain of IsdA (Muryoi *et al*., 2008). Although RF122 does not express IsdH NEAT 3 it expresses IsdB.
As it is speculated that gene decay in RF122 is a result of redundancy of such genes in the mammary gland environment it is possible that the majority of bovine strains will eventually follow this pattern towards niche adaptation. This would involve the monitoring of genome changes in genetically distinct bovine strains.

6.1.1.2 Niche adaptation; gene retention

Despite the considerable gene decay in RF122 there are many novel genes of unknown function not identified previously among human isolates, suggesting a possible role in bovine-specific pathogenesis (Herron-Olson et al., 2007).

When surface protein genes were investigated in this study, several genes of the isd system (many of which encode proteins that are not surface-located) were amongst the genes retained by RF122. The main function of the Isd system is iron acquisition. As previously mentioned, a source of heme has been identified in bovine milk. $\alpha$-casein from bovine milk can bind both iron and heme (Shibuya et al., 2012). Perhaps the surface-exposed Isd proteins from RF122 can acquire heme from $\alpha$-casein-heme complexes.

Iron acquisition systems seem to be important in several other bacterial species, especially in relation to niche adaptation. The acquisition of iron has been shown to be crucial for aquatic organisms and the ability of utilizing iron through multiple systems was discussed to be important during growth of environmental Vibrio cholerae isolates. The biosynthesis of specific iron-chelators in addition to other unspecific siderophores may represent an advantage for the adaptation to a specific niche (Kahlke et al., 2012).

Iron acquisition was also shown to be important in Edwardsiella bacteria. Edwardsiella are leading fish pathogens which cause huge losses to aquaculture industries worldwide. E. tarda is a broad-host range pathogen that infects several species of fish and other animals including humans. Many iron-scavenging related genes were detected among the virulence genes under positive selection, showing strong signs of adaptive evolution in the E. tarda EdwGI lineages and fulfilled the criteria of key evolutionary factors likely facilitating the virulence evolution and adaptation to a broad range of hosts in the E. tarda EdwGI strains (Yang et al., 2012).
This is consistent with the putative adaptation of RF122 to the mammary gland. As all of the bovine strains tested here express both IsdA and IsdB it would be interesting to investigate if they express the remaining genes of this system. As it seems that the Isd system is important in the mammary gland it is unlikely that it will be lost.

In addition to their roles in iron acquisition, IsdA and IsdB have other functions which may be important in the mammary gland. Lactoferrin (LF) is synthesized in mammary epithelium and has bactericidal and bacteriostatic functions (Huang et al., 2012) and most of the iron in the mammary gland is bound to lactoferrin or ferric citrate (Mazmanian et al., 2003). IsdA binds lactoferrin and protects the cell from its bactericidal effects (Clarke & Foster, 2008). This could be important in the mammary gland to protect RF122.

It is unlikely that IsdA was retained for its putative fibronectin binding properties. The contribution of IsdA to fibronectin binding is not particularly convincing. Clarke, Wiltshire and Foster (Clarke et al., 2004), showed that SH1000 adhesion to fibronectin at levels which we would consider to be background. Work in this study showed that neither RF122 nor Newman grown under iron-starved conditions could adhere detectably to fibronectin.

FnBPs provide an important mechanism for invasion of mammalian cells via a Fn bridge and α5β1. Although RF122 has not been shown to adhere to fibronectin and thus cannot invade by this mechanism, this could be compensated by the expression of IsdB. IsdB has been shown to bind the platelet integrin GPⅠb/Ⅲa and to RGD-containing integrins (Miajlovic et al., 2010). It promotes internalisation into human epithelial cells independently of Fn using human cells (Zapotoczna et al., 2013).

To test whether IsdB of RF122 promotes internalisation it would be more biologically relevant to use bovine mammary epithelial (MAC-T) cells which have been used in other studies to investigate internalization (Kerro-Dego et al., 2012). It would be interesting to investigate the different integrin types displayed by MAC-T cells. This could aid identification of an integrin(s) that could be recognised by RF122 IsdB.

Bbp is expressed by RF122 but not by other bovine strains (the majority of which express SdrE instead). It is difficult to speculate as to why bbp has been retained/acquired. Although Bbp has previously been shown to bind human Fg (Vazquez et al., 2011), this study showed that RF122 does not adhere to Fg under in vitro growth
conditions. It is likely to have a role to play in RF122 and the true *in vivo* function warrants investigation.

The majority of bovine strains express ClfB, some with levels comparable to Newman. Recombinant ClfB (cloned from RF122 and expressed from the pQE30 expression vector) binds fibrinogen which shows that despite the lack of adhesion by RF122 this protein is functional. Whether or not ClfB has a role in the mammary gland warrants further investigation. ClfB has been shown to be important for binding to desquamated nasal epithelial cells (Corrigan *et al.*, 2009). If RF122 ClfB is functionally redundant this could be compensated by the expression of IsdA, which has also previously been shown to promote adhesion to desquamated epithelial cells (Clarke *et al.*, 2009) allowing *S. aureus* to colonise the bovine host. It would be interesting to investigate if RF122 ClfB is subjected to proteolytic degradation by metalloproteases. The biological significance of ClfB truncation by metalloproteases during the growth cycle is unclear. It was previously suggested that loss of an N-terminal domain from the surface-associated protein may release biologically active peptide fragments. The spread of infection within the host may be promoted by the cleavage of ClfB which may facilitate the detachment of bacterial cells from colonised sites (McAleese *et al.*, 2001).

### 6.1.2 Vaccine

#### 6.1.2.1 A proposed surface-protein vaccine for combatting bovine mastitis

Surface protein vaccines have shown potential for combatting diseases caused by *S. aureus* strains. Potential candidates for a multicomponent surface protein vaccine to combat bovine mastitis were investigated in this study. Although the protein expression screen included several proteins that have shown potential in combatting human infections, only a few candidates could be proposed to combat bovine mastitis (namely IsdA, IsdB, ClfB and Bbp/SdrE). Unfortunately this was not tested so one cannot draw a conclusion. Granting the considerable decay exhibited in the RF122 genome has provided insight as to how this successful mastitis-causing strain has adapted to the mammary gland, it has made the quest for a mastitis-combatting-vaccine difficult.

IsdA, IsdB and SdrE have previously shown potential in a vaccine together with SdrD in a murine model of abscess formation (Stranger-Jones *et al.*, 2006). The gene for
sdrD is missing in the RF122 genome and therefore was not considered in the mastitis vaccine. Although its allelic variant SdrE has shown potential, Bbp has not previously been recommended for a vaccine.

Although the requirement/redundancy of ClfB in the mammary gland is inconclusive, the fact that it is expressed by RF122 and the majority of bovine strains is promising for its use as part of a multicomponent vaccine to combat mastitis. ClfB has previously shown potential as a vaccine candidate in combatting human diseases. Mucosal immunization with killed *S. aureus* and both systemic and intranasal immunization with rClfB significantly reduced bacterial numbers in the nose compared to controls, following challenge infection with *S. aureus* Newman (Schaffer *et al.*, 2006).

### 6.1.2.2 Non-LPXTG vaccines/antigens

The profile of surface anchored proteins expressed by RF122, which only expressed a few, is not representative of those expressed by the majority of bovine strains. If the multicomponent vaccine proposed in this study is not efficacious it may be wise to investigate secreted proteins for use as vaccine candidates to combat mastitis caused by this atypical strain as well as other bovine strains. Amongst secreted vaccine candidates that have been investigated to combat staphylococcal infections in humans are α-toxin (Adhikari *et al.*, 2012), lipoprotein (Mishra *et al.*, 2012) and other secreted proteins (Jongerius *et al.*, 2012).

A recombinant subunit vaccine candidate for α-hemolysin (Hla - the candidate for which was denoted AT-62aa) provided significant protection along with a Glucopyranosyl Lipid Adjuvant-Stable Emulsion (GLA-SE) in pneumonia and bacteremia infection models using *S. aureus* strain Newman and the pandemic strain USA300 (LAC). Passive transfer of rabbit IgG against AT-62aa protected mice against intraperitoneal and intranasal challenge with USA300 and produced significant reduction in bacterial burden in blood, spleen, kidney, and lungs (Adhikari *et al.*, 2012). As RF122 has been shown to express high levels of α-toxin *in vitro* compared with human clinical *S. aureus* isolates, including MRSA WCUH29 and MRSA USA300 (Liang *et al.*, 2011) this shows promise for investigation of α-toxin as a candidate to combat bovine mastitis.
Vaccination with the surface-exposed lipoprotein FhuD2 which is involved in iron-hydroxamate uptake, generated protective immunity against diverse clinical *S. aureus* isolates in murine infection models. It was somewhat surprising that lipoprotein was surface-exposed but functional antibodies that were shown to mediate opsonophagocytosis were effective in provided protection in passive transfer experiments (Mishra *et al.*, 2012). Lipoproteins of RF122 have not been investigated but their potential as vaccine candidates for combatting bovine mastitis should now be considered.

Secreted proteins have also been investigated for their potential as vaccine candidates. *S. aureus* secretes several small proteins that block specific elements of the host innate immune system, but their role in bacterial pathogenicity is unknown due to human specificity. Targeted inactivation of the genes encoding Ecb and Efb strongly attenuates *S. aureus* virulence in a murine infection model. Ecb and Efb were shown to be essential to *S. aureus* virulence *in vivo* and could be attractive targets in future vaccine development efforts (Jongerius *et al.*, 2012). Analysing the secretome of RF122 could reveal potential candidates and the expression profile of proteins secreted by RF122 could be investigated.

### 6.1.3 Restriction systems of bovine *S. aureus* strains

The majority of bovine strains tested possess the *sauUSI* gene encoding the type IV restriction system SauUSI which specifically recognizes cytosine methylated DNA. SauUSI was identified as the major barrier to transformation with foreign DNA (Corvaglia *et al.*, 2010; Xu *et al.*, 2011). Direct transformation can be performed with plasmid DNA isolated from *E. coli* DC10B into *S. aureus* strains from the majority of clonal complexes. DC10B was created by deletion of *dcm* encoding a cytosine methylase in the high efficiency *E. coli* cloning strain DH10B. The absence of cytosine methylation allows plasmid DNA to bypass the type IV restriction barrier. The only strain that could not be transformed was from CC97. *S. epidermidis* contains an ortholog of SauUSI (termed McrR) which also recognizes cytosine methylation. McrR can also be bypassed with plasmid DNA isolated from DC10B (Monk & Foster, 2012; Monk *et al.*, 2012).

CC97 constitutes one of the major complexes which the majority of ruminant-associated sequence types belong to (Guinane *et al.*, 2010). As the *sauUSI* barrier could not be by-passed in a representative of CC97 (Monk & Foster, 2012), it is possible that
many bovine strains cannot be manipulated in this manner. Identification of other restriction systems employed by genetically distinct bovine strains may aid in the quest to understand how they can be manipulated. RF122 does not possess sauUSI which is the major barrier that restricts uptake of foreign DNA. Instead it possesses sab2370c, a putative type II endonuclease, and dcm, which encodes a putative cytosine methylase. Expression of Sab2370c stops E. coli cell growth and deletion of the sab2370c gene facilitates an increased frequency of transformation into RF122. The dcm gene could not be deleted as this mutation is likely to be lethal. Dcm provides partial protection from Sab2370c.

6.1.4 New approaches for genetically manipulating S. aureus.

To genetically manipulate RF122 in this thesis the temperature-sensitive pKOR1 plasmid was used. Utilisation of high temperatures during allelic exchange can result in secondary mutations eg mutations in sae genes which encode a two-component system can be selected during growth at high temperature in the presence of the antibiotic erythromycin or chloramphenicol (Sun et al., 2010). To circumvent this, the plIMAY vector, which avoids the use of high temperatures, can be used for allelic exchange in staphylococci which may aid the future genetic manipulation of bovine strains and thus identification of virulence factors.

The pVWO1ts replicon of plIMAY is functional in staphylococci at the permissive temperature (30°C) but the plasmid cannot replicate at the restrictive temperature of 37°C. Thus integrants can be selected for at the low temperature of 37°C (Monk et al., 2012). The strongly expressed chloramphenicol resistance (cat) gene marker allows efficient selection as a single copy when integrated into the staphylococcal chromosome. The replicon used to propagate the plasmid in E. coli is low copy number which should improve the stability of cloned staphylococcal DNA. The plasmid carries the inducible secY antisense counterselection determinant of pKOR1. The production of deletion constructs has been optimized by application, to plIMAY, of sequence- and ligation-independent cloning (Li & Elledge, 2007). This has facilitated an increase in cloning efficiency and a decrease in both costs and time (Monk & Foster, 2012).

To complete this study, major goals include performing vaccine trials using the RF122-derived recombinant IsdB, ClfB and Bbp along with recombinant IsdA and SdrE (cloned from strain Newman) singly and in combination, measuring the elicited immune
responses and analysing immunological memory upon challenge infection. The ClfA and IsdH proteins could be included to further broaden the strains combatted by this vaccine. Another major goal would be to generate a conditional mutant of \textit{dcm} to prove that its deletion would be lethal thus demonstrating its importance in bacterial survival and pathogenesis. As RF122 is a successful and widespread mastitis-causing bacteria it would be useful to characterise the Sab2370c protein which would aid the investigation of its role in protecting ST151 strains from foreign DNA.

The accumulated information obtained from this thesis has broadened our understanding of the array of surface proteins and restriction-modification systems employed by diverse bovine mastitis strains of \textit{S. aureus}. Completion of this project would determine whether a surface-protein based vaccine is a viable approach to combat bovine mastitis (if not it would be necessary to adopt an alternate approach) and further investigation of the restriction-modification systems employed by bovine \textit{S. aureus} strains would facilitate the future identification of \textit{S. aureus} virulence factors that cause bovine mastitis.
References


125


Staphylococcus aureus reveals insights into the origin and molecular basis of ruminant host adaptation. *Genome Biol Evol* 2, 454-466.


induces rapid antibody responses in rhesus macaques and specific increased survival in a murine S. aureus sepsis model. *Infect Immun* 74, 2215-2223.


Moreillon P, Entenza JM, Francioli P, McDevitt D, Foster TJ, Francois P & Vaudaux P 
(1995) Role of *Staphylococcus aureus* coagulase and clumping factor in 

Murray NE (2000) Type I restriction systems: sophisticated molecular machines (a legacy 

Demonstration of the iron-regulated surface determinant (Isd) heme transfer 

conversion of Panton-Valentine leukocidin in *Staphylococcus aureus*: molecular 

Navarre WW & Schneewind O (1999) Surface proteins of gram-positive bacteria and 

Neave FK, Dodd FH, Kingwill RG & Westgarth DR (1969) Control of mastitis in the 

factor B (ClfB), a new surface-located fibrinogen-binding adhesin of 

Nicholas RO, Li T, McDevitt D, Marra A, Sucoloski S, Demarsh PL & Gentry DR (1999) 
Isolation and characterization of a sigB deletion mutant of *Staphylococcus aureus*. 

Nour El-Din AN, Shkreta L, Talbot BG, Diarra MS & Lacasse P (2006) DNA 
immunization of dairy cows with the clumping factor A of *Staphylococcus aureus*. 

Cox D (2002a) Multiple mechanisms for the activation of human platelet 
aggregation by *Staphylococcus aureus*: roles for the clumping factors ClfA and 
1033-1044.


