1 **A short sequence within subdomain N1 of region A of the *Staphylococcus aureus***

2 **MSCRAMM clumping factor A is required for export and surface display.**

3 **Niamh McCormack, Timothy J. Foster[#] and Joan A. Geoghegan[#*]**

4 Department of Microbiology, Moyne Institute of Preventive Medicine, School of Genetics

5 and Microbiology, Trinity College Dublin, Dublin 2, Ireland.

6

7 Running Title: Export and surface display of ClfA.

8

9 Contents category: Cell and Molecular Biology of Microbes

10

11 Key words: *Staphylococcus aureus,* clumping factor A, MSCRAMM, protein secretion,

12 fibronectin binding protein B.

13

14 Word counts: Summary 248, main text 5108.

15

16

17 [#]These authors contributed equally

18 *Corresponding author: E-mail: geoghegj@tcd.ie. Telephone: +35318961188.

19 Fax: +35316799294

20   **Summary**

21   Clumping factor A (ClfA) is the archetypal fibrinogen-binding surface protein of

22   *Staphylococcus aureus* and a member of the Microbial Surface Component Recognising

23   Adhesive Matrix Molecules (MSCRAMM) family. An N-terminal signal sequence directs

24   export of the MSCRAMM by the Sec pathway and the C-terminal cell wall-anchoring

25   domain allows covalent attachment of ClfA to peptidoglycan by sortase. Region A of ClfA

26   comprises three independently folded subdomains N1, N2 and N3. Subdomains N2N3

27   comprise IgG-like folds and promote fibrinogen binding. Nothing is known about the

28   structure or function of subdomain N1. Here we demonstrate an unexpected role for N1 in

29   the export and surface localization of ClfA. Attempted expression of a ClfA variant lacking

30   subdomain N1 resulted in impaired growth of *S. aureus* and accumulation of ClfA protein in

31   the cytoplasm and cytoplasmic membrane. The presence of residues 211-228 of N1 was

32   required to allow display of ClfA on the bacterial surface. The importance of this region was

33   confirmed when a ClfA variant lacking residues 211-220 was also mislocalized to the

34   cytoplasm and cytoplasmic membrane. However these residues were not required for export

35   of ClfA lacking the Ser-Asp repeats that links region A to the wall-anchoring domain.

36   Similarly, subdomain N1 of a related MSCRAMM fibronectin binding protein B was

37   required for export and surface display of the full length protein but not a derivative lacking

38   fibronectin binding repeats. In summary we demonstrate that residues in the N1 subdomain

39   are required for export and cell wall localization of *S. aureus* MSCRAMMs.

40

41

42

Introduction

Staphylococcus aureus is a commensal bacterium that colonises the anterior nares of about 20% of the population (van Belkum *et al.*, 2009). The organism is an important opportunistic pathogen that can cause both superficial skin lesions and more serious invasive infections such as endocarditis and sepsis (Lowy, 1998).

The ability of *S. aureus* to colonise the host and cause infection is due in part to the expression of secreted virulence factors and cell wall-anchored surface proteins. The microbial surface component recognizing adhesive matrix molecules (MSCRAMM) family comprises proteins with related N-terminal ligand-binding A domains (Foster *et al.*, 2013). These contain two tandemly arrayed subdomains, N2 and N3, that are composed of IgG-like folds (Deivanayagam *et al.*, 2002) and the N-terminal N1 subdomain which, in the case of the related protein ClfB, is elongated and is not compact (Perkins *et al.*, 2001). The A domains are linked to the cell wall by an extended flexible region comprising repeats of the dipeptide Ser-Asp (ClfA, ClfB, SdrC, SdrD, SdrE, Bbp) or 10/11 fibronectin binding domains (FnBPA, FnBPB) (Hartford *et al.*, 1997; Schwarz-Linek *et al.*, 2003). MSCRAMMs are also found in coagulase-negative staphylococci and *S. pseudintermedius* (Bannoehr *et al.*, 2011; Foster *et al.*, 2013).

MSCRAMMs play a pivotal role during colonization and establishment of *S. aureus* infection yet very little is known about how these proteins are directed to the secretion apparatus (Schneewind & Missiakas, 2012; Sibbald *et al.*, 2006). Members of the MSCRAMM family possess an N-terminal signal sequence that is predicted to mediate their translocation across the cytoplasmic membrane by the general secretory (Sec) pathway (DeDent *et al.*, 2008; Sibbald *et al.*, 2006). The Sec system in *S. aureus* consists of homologues of the canonical SecYEG translocation channel and the SecA ATPase (Sibbald *et al.*, 2006). A 'YSIRK/GS' motif present in the signal sequence directs secretion of the

68   pre-proteins to sites of peptidoglycan biosynthesis at the region of septum formation (DeDent

69   *et al.*, 2008). However, the mechanisms involved in trafficking MSCRAMMs to this cellular

70   compartment are yet to be elucidated (DeDent *et al.*, 2008; Schneewind & Missiakas, 2012).

71   Surface display of *S. aureus* YSIRK/GS-motif containing surface proteins is reduced in

72   mutants lacking the *spdA, spdB or spdC* genes but the molecular basis of this is not

73   understood (Frankel *et al.*, 2010).

74   A sorting signal is located at the C-terminus of the protein. It comprises an LPXTG

75   motif, a hydrophobic membrane-spanning domain and at the extreme C-terminus, a stretch of

76   positively charged residues.   The last two elements delay secretion across the membrane and

77   facilitate recognition and cleavage of the LPXTG sequence by sortase A, a protein anchored

78   to the cytoplasmic membrane (Schneewind *et al.*, 1993). Sortase cleaves between threonine

79   and glycine forming an acyl-enzyme intermediate which allows transfer of the MSCRAMM

80   to the pentaglycine crossbridge of lipid II. Following transglycosylation the MSCRAMM

81   becomes covalently anchored to peptidoglycan (Ton-That *et al.*, 2000).

82   Clumping factor A is an important MSCRAMM that promotes virulence during

83   invasive infection (Josefsson *et al.*, 2001; McAdow *et al.*, 2011; Moreillon *et al.*, 1995) .

84   This involves both coating the bacterium with plasma fibrinogen and capturing and activating

85   the complement regulator factor I which in turn results in cleavage of the complement

86   opsonin C3b (Hair *et al.*, 2010; Higgins *et al.*, 2006). ClfA is also an adhesin which

87   promotes bacterial attachment to immobilized fibrinogen and to platelet-fibrin thrombi on

88   heart valves during the initiation of endocarditis (Moreillon *et al.*, 1995; Que *et al.*, 2001).

89   The A domain of ClfA binds to a peptide sequence comprising the extreme C-

90   terminal residues of the γ-chain of fibrinogen (McDevitt *et al.*, 1997). The X-ray crystal

91   structure of ClfA N2N3 has been solved both as an apo-protein and with the fibrinogen

92   peptide ligand bound (Ganesh *et al.*, 2008). The fibrinogen γ-chain peptide binds in a

93   hydrophobic trench formed between the separately folded N2 and N3 subdomains by the

94   'dock, lock, and latch' mechanism (Ponnuraj *et al.*, 2003). Docking of the peptide in the

95   binding trench induces a conformational change in the flexible extension of the C-terminus of

96   the N3 subdomain which is reoriented to cover the ligand peptide and "locks" it in place.

97   Latching is completed by formation of an extra β-strand which is complementary to a β-sheet

98   of the N2 domain. The MSCRAMMs ClfB, Bbp, FnBPA and FnBPB bind to fibrinogen by a

99   similar mechanism (Burke *et al.*, 2011; Ganesh *et al.*, 2011; Keane *et al.*, 2007; Vazquez *et*

100   *al.*, 2011). In all cases the ligand binding site appears to be confined to subdomains N2N3

101   (Burke *et al.*, 2011; Ganesh *et al.*, 2011; Keane *et al.*, 2007; Vazquez *et al.*, 2011).

102   The boundaries of the N1, N2 and N3 subdomains were originally defined for ClfB

103   before structures of any of the proteins were solved. The *S. aureus* zinc-metalloprotease

104   aureolysin was found to cleave ClfB at a 'SLAVA' motif and the cleavage site was

105   designated as the boundary between N1 and N2 (McAleese *et al.*, 2001). Virtually nothing is

106   known about the structure or function of subdomain N1 of the MSCRAMM family of surface

107   proteins. The objective of this study was to determine if subdomain N1 of ClfA plays a role

108   in export of the protein and its subsequent elaboration on the cell surface.

109

110

111

112

113

114

115

Methods

Bacterial culture conditions and reagents

Strains are listed in Table 1. *S. aureus* was grown on trypticase soy agar or broth (TSA, TSB, Oxoid) with shaking (200 rpm) at 37 $^{\circ}$C. For expression studies stationary phase cultures were diluted 1:100 in brain heart infusion broth (BHI, Difco), grown to an OD$_{600}$ of 0.5 and induced with anhydrotetracycline (ATc) for 2 h at 37 $^{\circ}$C. *E. coli* strains were cultured on Luria agar or broth (Difco) at 37°C. Antibiotics were incorporated into media where appropriate at the following concentrations: 100 µg ampicillin (Ap) ml^{-1}, 10 µg chloramphenicol (Cm) ml^{-1}, 10 µg erythromycin (Em) ml^{-1}, 50 µg kanamycin (Ka) ml^{-1}, 100 – 1200 ng anhydrotetracycline (ATc) ml^{-1}. Unless otherwise stated all reagents were obtained from Sigma.

Plasmid and strain construction

Bacterial strains and plasmids used in this study are listed in Table 1. Strain NM1 was constructed by transduction of *spa*::Ka^{R} into strain Newman *clfA*5 *clfB*::Em^{R}. Deletion of *clfB*::Em^{R} from strain NM1 to yield NM2 was achieved by allelic replacement using pIMAYΔ*clfB* (Table 1) as previously described (Monk *et al.*, 2012). Strain BH1CC Δ*fnbAfnbB spa*::Ka^{R} was constructed by transduction of *spa*::Ka^{R} into strain BH1CC Δ*fnbAfnbB*. The complete *clfA* gene from strain Newman was amplified from genomic DNA by PCR (primers study are listed in Table S1) and cloned between EcoRI and BglII sites of pRMC2 or between the EcoRI and KpnI sites of pALC2073 to generate pRMC2*clfA* and pALC2073*clfA,* respectively. The *fnbB* gene from strain 8325 was amplified from genomic DNA by PCR and cloned between the EcoRI and BglII sites of pRMC2 to generate pRMC2*fnbB*. Inverse PCR was used to generate derivatives of pRMC2*clfA*, pALC2073*clfA* and pRMC2*fnbB*. Alanine substitutions and FLAG epitopes were incorporated by primer

141   extension during inverse PCR. PCR products were treated with DpnI to eliminate parental

142   template DNA and following blunt-end ligation plasmids were transformed into *E. coli*

143   DC10B. Plasmids were purified from DC10B and transformed into *S. aureus* made

144   electrocompetent as previously described (Lofblom *et al.*, 2007; Monk *et al.*, 2012). Plasmid

145   DNA (5 µg) was introduced into *S. aureus* by electroporation and transformants were

146   recovered by plating on TSA containing Cm.

147

148   **Fractionation of *S. aureus* and Western immunoblotting.**

149   *S. aureus* cultures were grown as described above, washed in phosphate buffered saline

150   (PBS) and adjusted to an OD_{600} of 10. Cell wall-associated proteins were solubilised by

151   resuspending bacteria in 250 µl digestion buffer (50 mM Tris-HCl, 20 mM $MgCl_2$, 30 %

152   (w/v) raffinose, pH 7.5) containing protease inhibitors (Roche) and lysostaphin (AMBI

153   Products LLC, 200 µg ml^{-1}) at 37 $^{\circ}$C for 10 min. Protoplasts were removed by centrifugation

154   at 12,000 x *g* for 10 min and the supernatant containing solubilised cell wall proteins was

155   transferred to a clean tube and boiled for 5 min in Laemmli sample buffer. Protoplast pellets

156   were washed with digestion buffer and resuspended in ice-cold Tris-HCl (50 mM, pH 7.5)

157   containing protease inhibitors (Roche) and DNase (80 µg ml^{-1}). Protoplasts were lysed on ice

158   by vortexing. Complete lysis was confirmed by phase contrast microscopy. Lysed cell

159   fractions were separated by centrifugation at 44,500 x *g* for 1 h at 4 $^{\circ}$C. The supernatant was

160   retained as the cytoplasmic fraction. The pellet was washed with ice-cold Tris-HCl (50 mM,

161   pH 7.5) and resuspended in Tris-HCl (50 mM, pH 7.5). Fractions were separated by 7.5 %

162   sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, transferred onto PVDF

163   and blocked in 10 % (w/v) skimmed milk proteins. Blots were probed with ClfA mAb, a

164   murine monoclonal antibody raised against ClfA subdomains N2N3 (a kind gift from

165   Inhibitex Inc, 1:2000) and bound antibody was detected using horseradish peroxidise (HRP)-

conjugated rabbit anti-mouse IgG (1:3000, Dako). FnBPB blots were probed with biotinylated fibronectin and bound biotin was detected using horseradish peroxidise (HRP)-conjugated streptavadin. FLAG-tagged constructs were probed with murine monoclonal anti-FLAG IgG (1:2000, Pierce) and detected with horseradish peroxidise (HRP)-conjugated rabbit anti-mouse IgG (1:2000, Dako) Reactive bands were visualised using the LumiGLO reagent and peroxide detection system (Cell Signalling Technology). Antibodies were removed from blots by incubation at 70°C for 30 min in stripping buffer (2 % SDS, 100 mM β-mercaptoethanol and 50 mM Tris). Stripped blots were washed twice in PBS, blocked in 10 % (w/v) skimmed milk proteins and then probed with polyclonal rabbit anti-SdrE IgG (1:2000, O'Brien *et al.*, 2002) or rabbit anti-*Micrococcus luteus* F_1/F_0 ATPase serum (Downer *et al.*, 2002). Bound antibody was detected using HRP-conjugated goat anti-rabbit IgG (1:10,000, Dako) or protein A peroxidase (1:500).

Dot immunoblotting

Bacteria were washed twice in PBS and adjusted to an OD_{600} of 1.0. Doubling dilutions were spotted onto a nitrocellulose membrane (Protran) membrane and allowed to dry. For whole cell lysates bacteria were incubated in PBS containing protease inhibitors (Roche) and lysostaphin (AMBI Products LLC, 200 µg ml^{-1}) at 37°C for 15 mins. Lysates were incubated with DNase (80 µg/ml) for 10 min at room temperature and heated to 85°C for 5 min. Cellular debris was sedimented by centrifugation at 12,000 x *g*. Lysates were spotted onto a nitrocellulose membrane (Protran) membrane and allowed to dry. Membranes were blocked for 1 h in 10% (w/v) skimmed milk proteins and immunoblotting was performed as above.

191   **Competitive growth assay**

192   For competitive growth experiments stationary phase cultures were adjusted to an OD_{600} of

193   0.025 in BHI. Equal volumes of two *S. aureus* strains were mixed and grown in BHI

194   incorporating Cm to an OD_{600} of 0.5. Where appropriate ATc (1200 ng ml^{-1}) was added and

195   cultures were incubated at 37 $^{\circ}$C for a further 2 h before being plated on TSA incorporating

196   Cm. Colonies which were resistant to Cm were replica plated onto TSA containing Em and

197   the percentage of Em resistant bacteria was calculated.

198

199   **Statistical analysis**

200   Statistical analysis was performed using Prism Graphpad 5 software. P values were

201   calculated using Students t-test.

202

203

204

205

206

207

208

209

210

Results

Expression of ClfA variants lacking subdomain N1.

The subdomain boundaries of the A region of ClfA were originally defined based on amino-acid sequence similarity with ClfB where the position of a cleavage site (SLAVA) for the *S. aureus* metalloprotease aureolysin defined the junction between subdomains N1 and N2 (Perkins *et al.*, 2001). ClfA contains a similar sequence (SLAAVA). Based on this motif the amino-acid coordinates of the N2N3 subdomains of ClfA were originally defined as residues 221-559 (McDevitt *et al.*, 1995; O'Connell *et al.*, 1998). More recently the N2N3 subdomains of ClfA have been crystallized (Deivanayagam *et al.*, 2002; Ganesh *et al.*, 2008) allowing the precise definition of boundaries of the N2N3 subdomains as comprising residues 229-545 (Fig. 1).

This study set out to determine if subdomain N1 of ClfA plays a role in protein secretion and elaboration on the cell surface. To facilitate this, the *clfA* gene of strain Newman was cloned into and expressed from the anhydrotetracycline (ATc)-inducible vector pRMC2 (Corrigan & Foster, 2009) to yield pRMC2*clfA*. Plasmid pRMC2*clfA* was manipulated to create derivatives lacking DNA encoding the N1 subdomain of ClfA (residues 40-220 or residues 40-228) yielding plasmids pRMC2*clfA* ΔN1$_{40-220}$ and pRMC2*clfA*ΔN1$_{40-228,}$ respectively (Fig. 1). Two pRMC2*clfA*ΔN1 constructs were generated to reflect the different definitions of the boundary of subdomain N1. ClfAΔN1$_{40-220}$ represents a deletion up to the SLAAVA motif while ClfAΔN1$_{40-228}$ removes the entire N1 sequence defined by the X-ray crystal structure of N2N3 (Fig. 1). Plasmids were transformed into strain NM1, a ClfA-, ClfB-, Spa-deficient derivative of *S. aureus* Newman (Table 1).

Whole cell dot immunoblotting was performed to measure the amount of ClfA on the bacterial surface. Serial dilutions of bacteria were applied to a nitrocellulose membrane and a

235 monoclonal antibody that recognizes an epitope in subdomains N2N3 of ClfA (ClfA mAb)

236 was used to detect the protein. Expression of ClfA by NM1 (pRMC2$clfA$) could not be

237 detected in the absence of inducer (Fig. 2a) in agreement with previous studies showing that

238 gene expression from pRMC2 is tightly repressed (Corrigan & Foster, 2009). Strain NM1

239 carrying the pRMC2 empty vector was used as a negative control. Addition of increasing

240 concentrations of ATc resulted in increasing amounts of ClfA on the surface of $S.$ $aureus$

241 (Fig. 2a). A high level of surface expression similar to that achieved by the single

242 chromosomal gene in strain Newman was achieved at inducer concentrations of 800 ng ml^{-1}

243 and 1.2 µg ml^{-1}. In contrast neither ClfA ΔN1$_{40\text{-}220}$ nor ClfAΔN1$_{40\text{-}228}$ could be detected on the

244 surface of $S.$ $aureus$ even at the highest concentration of inducer tested (Fig. 2a). All proteins

245 could be detected when whole cell lysates were probed with ClfA mAb (Fig. 2b). Thus the

246 failure to detect ClfA ΔN1$_{40\text{-}220}$ and ClfAΔN1$_{40\text{-}228}$ on the surface of $S.$ $aureus$ suggested that

247 N1 might have a role in protein export. As a control, the same bacteria were spotted onto a

248 nitrocellulose membrane and probed for expression of SdrE, a member of the MSCRAMM

249 family which is exposed on the surface of $S.$ $aureus$. All bacteria were recognized equally by

250 antibodies directed against SdrE (Fig. S1). This demonstrates that similar amounts of

251 bacteria were immobilised on the nitrocellulose membrane.

252 Variants of pRMC2 $clfA$ were constructed which retained differing amounts of N1 in

253 an attempt to determine the minimum sequence that supported surface display. A mutant

254 lacking residues 40-210 but retaining C-terminal residues of N1 (ClfAΔN1$_{40\text{-}210}$) was

255 displayed on the surface of $S.$ $aureus$, albeit at a slightly lower level than full-length ClfA

256 expressed from pRMC2$clfA$ (Fig. 2a). However when cultures of NM1 (pRMC2$clfA$) and

257 NM1 (pRMC2$clfA\Delta$N1$_{40\text{-}210}$) were induced with 800 ng and 1.2 µg ATc ml^{-1}, respectively,

258 the amount of ClfA and ClfAΔN1$_{40\text{-}210}$ on the surface of $S.$ $aureus$ was found to be the same.

259 These inducer concentrations were used for all further experiments.

Subdomain N1 is required for ClfA to be localized to the _S. aureus_ cell wall.

Whole cell immunoblotting indicated that ClfAΔN1$_{40-220}$ was not elaborated on the surface of _S. aureus_ following induction with ATc (Fig. 2a) but ClfAΔN1$_{40-220}$ could be detected in whole cell lysates (Fig. 2b). Therefore the subcellular location of ClfAΔN1$_{40-220}$ following induction was investigated. Fractions corresponding to the cytoplasm, the cytoplasmic membrane and proteins solubilised from the cell wall by lysostaphin treatment during protoplast formation (cell wall fraction) were analyzed by Western immunoblotting probing with the ClfA mAb. In the wild type ClfA sample a band of 150 kDa was detected only in the cell wall fraction. This is consistent with normal sorting. In contrast, ClfAΔN1$_{40-220}$ was not detected in the cell wall fraction but was instead located in the membrane and cytoplasmic fractions (Fig. 3a). This protein migrated more slowly than expected and may represent translation products that had not been cleaved by signal peptidase and sortase A. Similar results were obtained for ClfAΔN1$_{40-228}$ (Fig. S2). The failure to localize to the cell wall explains why ClfAΔN1$_{40-220}$ and ClfAΔN1$_{40-228}$ were not detected on the surface of intact bacterial cells. In contrast ClfAΔN1$_{40-210}$ was only detected in the cell wall fraction in a similar fashion to full-length ClfA indicating that this truncate had been correctly secreted and sorted. As a control for secretion and sorting activity in cells attempting to express the ΔN1 constructs, samples were also probed for the expression of SdrE, a member of the MSCRAMM family which is known to be sorted to the cell wall (O'Brien _et al._, 2002). This protein was detected only in the cell wall fraction indicating that it had been exported and sorted correctly in all strains (Fig. 3b). Fractions were also probed with antibodies recognizing the integral membrane protein F_1/F_0 ATPase as a control for purity of cytoplasmic membrane preparations. This protein was detected only in the membrane fraction of all strains (Fig. 3c). Cytoplasmic fractions were stained with Coomassie to show equal loading of protein (Fig. 3d). In summary, these data indicate that ClfAΔN1$_{40-220}$ is

285 mislocalised to the membrane and cytoplasmic fractions while ClfAΔN1$_{40\text{-}210}$ is only found in

286 the cell wall fraction in a similar fashion to full-length ClfA.

287

Identification of residues in subdomain N1 required for ClfA to be localized to the cell

289 **wall.**

290 The N1 truncate ClfAΔN1$_{40\text{-}210}$ was sorted efficiently to the cell wall of *S. aureus* while

291 ClfAΔN1$_{40\text{-}220}$ and ClfAΔN1$_{40\text{-}228}$ were found in the cytoplasm and membrane (Fig. 3a, Fig

292 S3) suggesting that residues 211 - 220 may be required for efficient translocation and cell

293 wall localization. In order to investigate this further, plasmid pRMC2*clfA* was manipulated

294 to create a derivative lacking DNA encoding amino-acids 211–220 of ClfA (plasmid

295 pRMC2*clfA*Δ$_{211\text{-}220}$).   ClfAΔ$_{211\text{-}220}$ was not detected in the cell wall but was present in the

296 membrane and cytoplasm fractions (Fig. 4a) indicating that residues 211-220 are necessary

297 for cell wall localization.

298   While ClfAΔN1$_{40\text{-}210}$ was detected in the cell wall fraction only (Fig. 3), a variant

299 where residues 211-220 were substituted with ten alanines (ClfAΔN1$_{40\text{-}220+10\ \text{Ala}}$) was not (Fig

300 5a). These data show that residues 211-220 are necessary for cell wall localization of ClfA.

301 SdrE was detected only in the cell wall fraction indicating that it had been exported and

302 sorted correctly in all strains (Fig. 4b).

303

Expression of a ClfA variant lacking subdomain N1 inhibits growth of *S. aureus*.

305   Expression of ClfAΔN1$_{40\text{-}220}$ or ClfAΔN1$_{40\text{-}228}$ could not be detected on the surface of

306 *S. aureus* when NM1 (pRMC2*clfA*ΔN1$_{40\text{-}220}$) and NM1 (pRMC2*clfA*ΔN1$_{40\text{-}228}$) were grown in

307 the presence of a high concentration of inducer (Fig. 2a). To determine if expression of ClfA

308 $\Delta N1_{40-220}$ is deleterious, the *clfA* gene was cloned into pALC2073, the precursor of pRMC2

309 which allows constitutive expression in *S. aureus* due to weak TetR repression (Bateman *et*

310 *al.*, 2001; Corrigan & Foster, 2009). Plasmid pALC2073*clfA*$\Delta N1_{40-220}$ was constructed in

311 and isolated from *E. coli* DC10B. Equal amounts of the parental plasmid pALC2073*clfA* and

312 pALC2073*clfA*$\Delta N1_{40-220}$ DNA were used to transform *S. aureus* NM1. Transformation with

313 pALC2073*clfA* yielded ~2 x 10^3 Cm resistant colonies. However, pALC2073*clfA*$\Delta N1_{40-220}$

314 yielded fewer than 20 transformants. Both plasmids could be introduced into *E. coli* with

315 equal efficiency (~ 3 x 10^4 CFU per transformation, Fig. S3) confirming the integrity of the

316 DNA. The failure to establish transformants with pALC2073*clfA*$\Delta N1_{40-220}$ in *S. aureus*

317 suggested that this plasmid may be unstable or that constitutive expression of a variant of

318 ClfA lacking subdomain N1 inhibits growth.

319   This prompted us to investigate if inducing expression of ClfA$\Delta N1_{40-220}$ in *S. aureus*

320 would inhibit bacterial growth. Gene expression from pRMC2 is tightly repressed in the

321 absence of inducer. Plasmids pRMC2*clfA* and pRMC2*clfA*$\Delta N1_{40-220}$ transformed into *S.*

322 *aureus* with equal efficiency (data not shown). Competitive growth experiments were

323 performed with bacteria harbouring pRMC2*clfA* or pRMC2*clfA*$\Delta N1_{40-220}$. In order to

324 distinguish between cells harbouring the different plasmids, pRMC2*clfA*$\Delta N1_{40-220}$ was

325 introduced into strain NM2, a variant of NM1 where the *clfB*::Em^R mutation was replaced by

326 a deletion of the *clfB* gene ($\Delta clfB$) which is Em sensitive. Equal numbers of NM1

327 (pRMC2*clfA*) and NM2 (pRMC2*clfA*$\Delta N1_{40-220}$) were mixed and co-cultured either in the

328 presence or absence of the inducer ATc until the OD_{600} of the culture reached 2.0. Total

329 viable counts were obtained by plating on agar containing Cm. Replica plating onto agar

330 incorporating Em allowed the number of Em resistant bacteria NM1 (pRMC2*clfA*) to be

331 counted.

332   When ClfA expression was induced and the bacteria were allowed to grow for two

333   generations the number of NM2 (pRMC2*clfA*ΔN1$_{40-220}$) recovered was 1.5 times lower than

334   for NM1 (pRMC2*clfA*) (Fig. 5). This indicates that expression of ClfA lacking subdomain

335   N1 confers a growth disadvantage compared to expression of full-length ClfA. In the

336   absence of inducer when ClfA expression was repressed (Fig. 2a), equal numbers of NM1

337   (pRMC2*clfA*) and NM2 (pRMC2*clfA*ΔN1$_{40-220}$) were recovered (Fig. 5). This indicates that

338   the fitness disadvantage is due to the expression of ClfA ΔN1$_{40-220}$. To investigate if

339   expression of wild-type ClfA affected the fitness of *S. aureus*, control experiments were

340   conducted with NM1 (pRMC2*clfA*) and NM2 carrying the empty vector pRMC2. The

341   proportion of Em^{S} to Em^{R} bacteria was the same for cultures grown in ATc and the inducer-

342   free control (Fig. S4). These data indicate that expression of ClfAΔN1$_{40-220}$ in *S. aureus*

343   inhibits bacterial growth.

344

345   **Subdomain N1 is not required for cell wall localization of a ClfA variant lacking**

346   **residues comprising the serine-aspartate repeat region.**

347   The A domains of MSCRAMMs are linked to the cell wall by an extended flexible

348   region. In the case of ClfA this comprises Ser-Asp dipeptide repeats. We hypothesised that

349   the N1 subdomain might be required for the export of the flexible repeat region. To

350   investigate this, plasmids pRMC2*clfA* and pRMC2*clfA*ΔN1$_{40-220}$ were manipulated in order to

351   create variants lacking DNA encoding the Ser-Asp dipeptide repeat region (residues 559-875)

352   generating plasmid pRMC2*clfA*ΔSD$_{559-875}$ and plasmid pRMC2*clfA*ΔN1$_{40-220}$ΔSD$_{559-875}$,

353   respectively. Both ClfAΔSD$_{559-875}$ and ClfAΔN1$_{40-220}$ΔSD$_{559-875}$ were detected only in the cell

354   wall (Fig. 6). This demonstrates that the N1 subdomain is not required for export or cell wall

localization of ClfA lacking the flexible Ser-Asp repeat region and is only required for expression of the repeat containing protein.

Subdomain N1 is required for cell wall localization of another MSCRAMM with a long unstructured repeat region FnBPB. Region A of FnBPB shares ~ 25% amino-acid identity with the corresponding region of ClfA and is organised into N1, N2 and N3 subdomains (3). FnBPB does not have a Ser-Asp repeat region and instead harbours fibronectin binding repeats at the C-terminus of the protein. In order to determine if subdomain N1 was required for localisation of other *S. aureus* MSCRAMMs, we studied the subcellular localisation of FnBPB variants.

The *fnbB* gene from strain 8325 was cloned into pRMC2 to generate pRMC2*fnbB*. Plasmid pRMC2*fnbB* was manipulated to remove the entire N1 subdomain spanning residues 38-162 to yield pRMC2*fnbB*ΔN1$_{38\text{-}162}$. Plasmids were transformed into a strain of the MRSA isolate BH1CC lacking both FnBPA and FnBPB as well as protein A (BH1CC Δ*fnbAfnbB* *spa*::Ka^R, Table 1). This was chosen as the host strain since FnBPB is not subject to proteolytic processing and remains intact on the surface of BH1CC during all stages of growth (Geoghegan *et al.*, 2013).

Cell wall, membrane and cytoplasmic fractions were probed for FnBPB expression in a ligand affinity blot using biotinylated fibronectin. Similarly to ClfA, FnBPB is covalently anchored to the pentaglycine crossbridge of peptidoglycan and is released by lysostaphin treatment of *S. aureus*. FnBPBΔN1$_{38\text{-}162}$ was not detected in the cell wall fraction. Reactive bands with a molecular weight of ~100 kDa were observed in the membrane and cytoplasm fractions (Fig. 7a). This indicates that the N1 subdomain of FnBPB is required for cell wall localization of the protein.

379   To determine if the N1 subdomain is necessary for export of FnBPB in the absence of

380   the fibronectin binding repeat regions, pRMC2*fnbB*ΔN1$_{38-162}$ was manipulated to yield

381   pRMC2*fnbB*ΔN1$_{38-162}$ΔFnBR$_{481-811 \text{ FLAG}}$. A FLAG epitope tag was introduced by primer

382   extension during PCR. A reactive band with a molecular weight of ~ 65 kDa was observed in

383   the cell wall fraction (Fig. 7b). This demonstrates that the N1 subdomain is not required for

384   export or cell wall localization of FnBPB lacking the flexible fibronectin binding repeat

385   region.

386

387

388

389

390

391

392

393

394

395

396

397

398

Discussion

Clumping factor A is the archetypal member of the MSCRAMM family of wall-anchored proteins of *S. aureus*. The defining features of these proteins are (i) the A region and (ii) the flexible unfolded region linking the A region to the cell wall spanning/anchoring domain. The A domains comprise two separately folded subdomains N2 and N3. MSCRAMMs related to ClfA have an N-terminal subdomain N1 with no known function. This paper reports for the first time a crucial role for subdomain N1 in export and surface expression of MSCRAMMs. We demonstrate that residues 211-220 of N1 of ClfA are required for the protein to be translocated across the membrane so that it can be anchored by sortase to cell wall peptidoglycan. Deletion mutants lacking N1 were not exported and instead accumulated in the cytoplasm and the membrane fractions. This could only be detected in bacteria where ClfA expression was tightly repressed and was only expressed after addition of an inducer. Constitutive expression of an N1 deletion mutant following transformation into *S. aureus* resulted in a very low yield of transformants. Induction of the truncated *clfA* mutant in *S. aureus* caused a significant loss of fitness compared to cells expressing wild-type ClfA which shows that attempting to express the protein lacking N1 is causing damage to the cell. Dissection of the N1 subdomain of ClfA revealed that a stretch of 10 residues close to the boundary between N1 and N2 is necessary for surface expression, and that only this region plus the 8 residues that join it to the beginning of N2 as defined by the X-ray crystal structure is necessary. As yet no function has been identified for residues 40 – 210 of ClfA and the role of this region warrants further investigation.

We hypothesised that if this is a feature of the MSCRAMM family and not just a property specific to ClfA then the N1 subdomains of other MSCRAMMs would also be required for surface expression. Accordingly deletion of the N1 subdomain of FnBPB prevented surface expression of that protein. Further support comes from an

18

424 observation that a plasmid constitutively expressing an FnBPA N1 deletion mutant could not

425 be established in *S. aureus* (Geoghegan *et al.*, 2013). The level of amino acid sequence

426 similarity between the N1 subdomains of the MSCRAMMs is very low so that it is not

427 possible to determine if a similar 'motif' is present in this region. All MSCRAMMs with a

428 repeat region contain an N1 subdomain. MSCRAMMs are found in many species of

429 staphylococci including *S. epidermidis, S. pseudintermedius* and *S. lugdunensis* (Bannoehr *et*

430 *al.*, 2011; Heilbronner *et al.*, 2011; Ponnuraj *et al.*, 2003).

431   Previous studies have shown that MSCRAMM proteins are exported by the Sec

432 system with the other major secretory pathways (twin-arginine and accessory Sec system)

433 playing no role (DeDent *et al.*, 2008). However, virtually nothing is known about how

434 proteins of the MSCRAMM family are trafficked to the Sec apparatus (Schneewind &

435 Missiakas, 2012; Sibbald *et al.*, 2006). Previous studies identified a 'YSIRK-GS' motif in

436 the signal sequence of MSCRAMMs and certain other surface proteins (DeDent *et al.*, 2008).

437 This feature is required for the proteins to be targeted to the cross-wall of the cell but despite

438 the fact that this phenomenon has been described for streptococci and staphylococci the

439 molecular details of the mechanism are completely lacking (Carlsson *et al.*, 2006; DeDent *et*

440 *al.*, 2008; Schneewind & Missiakas, 2012). Similarly, we have herein identified a new

441 region of the MSCRAMM required for its export, but the mechanism involved is not yet

442 understood. In the absence of a fundamental understanding of pre-protein trafficking in *S.*

443 *aureus* it is difficult to speculate about the functions of N1. It is possible that the N1

444 subdomain acts as a cytoplasmic chaperone by directly binding to another region of the

445 protein. Alternatively, residues 211-220 could engage a molecular chaperone that guides the

446 long protein from the ribosome to the Sec translocation apparatus. Since truncates lacking

447 the repeat region are exported and cell wall anchored in the absence of N1 it is likely that this

448   region is required to stabilise the long flexible repetitive repeat regions and to assist with

449   transport.

450   In conclusion this paper provides new insights into the requirements for ClfA export.

451   Our findings suggest that secretion of large MSCRAMM proteins by Gram-positive bacteria

452   involves mechanisms more complex than previously appreciated.

453

454   **Acknowledgements**

457

458

459

460

461

462

463

464

465

466

467

468

469

470

References

Bannoehr, J., Ben Zakour, N. L., Reglinski, M., Inglis, N. F., Prabhakaran, S., Fossum, E., Smith, D. G., Wilson, G. J., Cartwright, R. A. & other authors (2011). Genomic and surface proteomic analysis of the canine pathogen *Staphylococcus pseudintermedius* reveals proteins that mediate adherence to the extracellular matrix. *Infect Immun* **79**, 3074-3086.

Bateman, B. T., Donegan, N. P., Jarry, T. M., Palma, M. & Cheung, A. L. (2001). Evaluation of a tetracycline-inducible promoter in Staphylococcus aureus in vitro and in vivo and its application in demonstrating the role of *sigB* in microcolony formation. *Infect Immun* **69**, 7851-7857.

Burke, F. M., Di Poto, A., Speziale, P. & Foster, T. J. (2011). The A domain of fibronectin-binding protein B of *Staphylococcus aureus* contains a novel fibronectin binding site. *FEBS J* **278**, 2359-2371.

Carlsson, F., Stalhammar-Carlemalm, M., Flardh, K., Sandin, C., Carlemalm, E. & Lindahl, G. (2006). Signal sequence directs localized secretion of bacterial surface proteins. *Nature* **442**, 943-946.

Corrigan, R. M. & Foster, T. J. (2009). An improved tetracycline-inducible expression vector for *Staphylococcus aureus*. *Plasmid* **61**, 126-129.

493   **DeDent, A., Bae, T., Missiakas, D. M. & Schneewind, O. (2008).** Signal peptides direct

494   surface proteins to two distinct envelope locations of *Staphylococcus aureus*. *The EMBO*

495   *journal* **27**, 2656-2668.

496

497   **Deivanayagam, C. C., Wann, E. R., Chen, W., Carson, M., Rajashankar, K. R., Hook,**

498   **M. & Narayana, S. V. (2002).** A novel variant of the immunoglobulin fold in surface

499   adhesins of *Staphylococcus aureus*: crystal structure of the fibrinogen-binding MSCRAMM,

500   clumping factor A. *The EMBO journal* **21**, 6660-6672.

501

502   **Downer, R., Roche, F., Park, P. W., Mecham, R. P. & Foster, T. J. (2002).** The elastin-

503   binding protein of *Staphylococcus aureus* (EbpS) is expressed at the cell surface as an

504   integral membrane protein and not as a cell wall-associated protein. *J Biol Chem* **277**, 243-

505   250.

506

507   **Duthie, E. S. & Lorenz, L. L. (1952).** Staphylococcal coagulase; mode of action and

508   antigenicity. *J Gen Microbiol* **6**, 95-107.

509

510   **Fitzgerald, J. R., Loughman, A., Keane, F., Brennan, M., Knobel, M., Higgins, J., Visai,**

511   **L., Speziale, P., Cox, D. & other authors (2006).** Fibronectin-binding proteins of

512   *Staphylococcus aureus* mediate activation of human platelets via fibrinogen and fibronectin

513   bridges to integrin GPIIb/IIIa and IgG binding to the FcgammaRIIa receptor. *Mol Microbiol*

514   **59**, 212-230.

515

516   **Foster, T. J., Geoghegan, J. A., Ganesh, V. K. & Hook, M. (2013).** Adhesion, invasion and

517   evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat Rev*

518   *Microbiol* **12**, 49-62.

519

520   **Frankel, M. B., Wojcik, B. M., DeDent, A. C., Missiakas, D. M. & Schneewind, O.**

521   **(2010).** ABI domain-containing proteins contribute to surface protein display and cell

522   division in *Staphylococcus aureus*. *Mol Microbiol* **78**, 238-252.

523

524   **Ganesh, V. K., Rivera, J. J., Smeds, E., Ko, Y. P., Bowden, M. G., Wann, E. R.,**

525   **Gurusiddappa, S., Fitzgerald, J. R. & Hook, M. (2008).** A structural model of the

526   *Staphylococcus aureus* ClfA-fibrinogen interaction opens new avenues for the design of anti-

527   staphylococcal therapeutics. *PLoS Pathog* **4**, e1000226.

528

529   **Ganesh, V. K., Barbu, E. M., Deivanayagam, C. C., Le, B., Anderson, A. S., Matsuka, Y.**

530   **V., Lin, S. L., Foster, T. J., Narayana, S. V. & other authors (2011).** Structural and

531   biochemical characterization of *Staphylococcus aureus* clumping factor B/ligand interactions.

532   *J Biol Chem* **286**, 25963-25972.

533

534   **Geoghegan, J. A., Monk, I. R., O'Gara, J. P. & Foster, T. J. (2013).** Subdomains N2N3 of

535   fibronectin binding protein A mediate *Staphylococcus aureus* biofilm formation and

536   adherence to fibrinogen using distinct mechanisms. *J Bacteriol* **195**, 2675-2683.

537

538 **Hair, P. S., Echague, C. G., Sholl, A. M., Watkins, J. A., Geoghegan, J. A., Foster, T. J.**

539 **& Cunnion, K. M. (2010).** Clumping factor A interaction with complement factor I increases

540 C3b cleavage on the bacterial surface of *Staphylococcus aureus* and decreases complement-

541 mediated phagocytosis. *Infect Immun* **78**, 1717-1727.

542

543 **Hartford, O., Francois, P., Vaudaux, P. & Foster, T. J. (1997).** The dipeptide repeat

544 region of the fibrinogen-binding protein (clumping factor) is required for functional

545 expression of the fibrinogen-binding domain on the *Staphylococcus aureus* cell surface. *Mol*

546 *Microbiol* **25**, 1065-1076.

547

548 **Heilbronner, S., Holden, M. T., van Tonder, A., Geoghegan, J. A., Foster, T. J.,**

549 **Parkhill, J. & Bentley, S. D. (2011).** Genome sequence of *Staphylococcus lugdunensis*

550 N920143 allows identification of putative colonization and virulence factors. *FEMS*

551 *Microbiol Lett* **322**, 60-67.

552

553 **Higgins, J., Loughman, A., van Kessel, K. P., van Strijp, J. A. & Foster, T. J. (2006).**

554 Clumping factor A of *Staphylococcus aureus* inhibits phagocytosis by human

555 polymorphonuclear leucocytes. *Fems Microbiol Lett* **258**, 290-296.

556

557 **Josefsson, E., Hartford, O., O'Brien, L., Patti, J. M. & Foster, T. (2001).** Protection

558 against experimental *Staphylococcus aureus* arthritis by vaccination with clumping factor A,

559 a novel virulence determinant. *J Infect Dis* **184**, 1572-1580.

560

561 **Keane, F. M., Loughman, A., Valtulina, V., Brennan, M., Speziale, P. & Foster, T. J.**

562 **(2007).** Fibrinogen and elastin bind to the same region within the A domain of fibronectin

563 binding protein A, an MSCRAMM of *Staphylococcus aureus*. *Mol Microbiol* **63**, 711-723.

564

565 **Lofblom, J., Kronqvist, N., Uhlen, M., Stahl, S. & Wernerus, H. (2007).** Optimization of

566 electroporation-mediated transformation: *Staphylococcus carnosus* as model organism. *J*

567 *Appl Microbiol* **102**, 736-747.

568

569 **Lowy, F. D. (1998).** *Staphylococcus aureus* infections. *N Engl J Med* **339**, 520-532.

570

571 **McAdow, M., Kim, H. K., Dedent, A. C., Hendrickx, A. P., Schneewind, O. &**

572 **Missiakas, D. M. (2011).** Preventing *Staphylococcus aureus* sepsis through the inhibition of

573 its agglutination in blood. *PLoS Pathog* **7**, e1002307.

574

575 **McAleese, F. M., Walsh, E. J., Sieprawska, M., Potempa, J. & Foster, T. J. (2001).** Loss

576 of clumping factor B fibrinogen binding activity by *Staphylococcus aureus* involves cessation

577 of transcription, shedding and cleavage by metalloprotease. *J Biol Chem* **276**, 29969-29978.

578

579 **McDevitt, D., Francois, P., Vaudaux, P. & Foster, T. J. (1995).** Identification of the

580 ligand-binding domain of the surface-located fibrinogen receptor (clumping factor) of

581 *Staphylococcus aureus*. *Mol Microbiol* **16**, 895-907.

582

583 **McDevitt, D., Nanavaty, T., House-Pompeo, K., Bell, E., Turner, N., McIntire, L.,**

584 **Foster, T. & Hook, M. (1997).** Characterization of the interaction between the

585 *Staphylococcus aureus* clumping factor (ClfA) and fibrinogen. *Eur J Biochem* **247**, 416-424.

586

587 **Monk, I. R., Shah, I. M., Xu, M., Tan, M. W. & Foster, T. J. (2012).** Transforming the

588 untransformable: application of direct transformation to manipulate genetically

589 *Staphylococcus aureus* and *Staphylococcus epidermidis*. *mBio* **3**.e00277-11.

590

591 **Moreillon, P., Entenza, J. M., Francioli, P., McDevitt, D., Foster, T. J., Francois, P. &**

592 **Vaudaux, P. (1995).** Role of *Staphylococcus aureus* coagulase and clumping factor in

593 pathogenesis of experimental endocarditis. *Infect Immun* **63**, 4738-4743.

594

595 **Mulcahy, M. E., Geoghegan, J. A., Monk, I. R., O'Keeffe, K. M., Walsh, E. J., Foster, T.**

596 **J. & McLoughlin, R. M. (2012).** Nasal Colonisation by *Staphylococcus aureus* Depends

597 upon Clumping Factor B Binding to the Squamous Epithelial Cell Envelope Protein Loricrin.

598 *PLoS Pathog* **8**, e1003092.

599

600 **O'Brien, L., Kerrigan, S. W., Kaw, G., Hogan, M., Penades, J., Litt, D., Fitzgerald, D.**

601 **J., Foster, T. J. & Cox, D. (2002).** Multiple mechanisms for the activation of human platelet

602 aggregation by *Staphylococcus aureus*: roles for the clumping factors ClfA and ClfB, the

603 serine-aspartate repeat protein SdrE and protein A. *Mol Microbiol* **44**, 1033-1044.

604

605 **O'Connell, D. P., Nanavaty, T., McDevitt, D., Gurusiddappa, S., Hook, M. & Foster, T.**

606 **J. (1998).** The fibrinogen-binding MSCRAMM (clumping factor) of *Staphylococcus aureus*

607 has a Ca2+-dependent inhibitory site. *J Biol Chem* **273**, 6821-6829.

608

609 **Perkins, S., Walsh, E. J., Deivanayagam, C. C., Narayana, S. V., Foster, T. J. & Hook,**

610 **M. (2001).** Structural organization of the fibrinogen-binding region of the clumping factor B

611 MSCRAMM of *Staphylococcus aureus*. *J Biol Chem* **276**, 44721-44728.

612

613 **Ponnuraj, K., Bowden, M. G., Davis, S., Gurusiddappa, S., Moore, D., Choe, D., Xu, Y.,**

614 **Hook, M. & Narayana, S. V. (2003).** A "dock, lock, and latch" structural model for a

615 staphylococcal adhesin binding to fibrinogen. *Cell* **115**, 217-228.

616

617 **Que, Y. A., Francois, P., Haefliger, J. A., Entenza, J. M., Vaudaux, P. & Moreillon, P.**

618 **(2001).** Reassessing the role of *Staphylococcus aureus* clumping factor and fibronectin-

619 binding protein by expression in *Lactococcus lactis*. *Infect Immun* **69**, 6296-6302.

620

621 **Schneewind, O., Mihaylova-Petkov, D. & Model, P. (1993).** Cell wall sorting signals in

622 surface proteins of gram-positive bacteria. *The EMBO journal* **12**, 4803-4811.

623

624 **Schneewind, O. & Missiakas, D. M. (2012).** Protein secretion and surface display in Gram-

625 positive bacteria. *Philosophical transactions of the Royal Society of London Series B,*

626 *Biological sciences* **367**, 1123-1139.

627

628 **Schwarz-Linek, U., Werner, J. M., Pickford, A. R., Gurusiddappa, S., Kim, J. H., Pilka,**

629 **E. S., Briggs, J. A., Gough, T. S., Hook, M. & other authors (2003).** Pathogenic bacteria

630 attach to human fibronectin through a tandem beta-zipper. *Nature* **423**, 177-181.

631

632 **Sibbald, M. J., Ziebandt, A. K., Engelmann, S., Hecker, M., de Jong, A., Harmsen, H.**

633 **J., Raangs, G. C., Stokroos, I., Arends, J. P. & other authors (2006).** Mapping the

634 pathways to staphylococcal pathogenesis by comparative secretomics. *Microbiol Mol Biol*

635 *Rev* **70**, 755-788.

636

637 **Ton-That, H., Mazmanian, S. K., Faull, K. F. & Schneewind, O. (2000).** Anchoring of

638 surface proteins to the cell wall of *Staphylococcus aureus*. Sortase catalyzed in vitro

639 transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. *J Biol Chem*

640 **275**, 9876-9881.

641

642 **van Belkum, A., Verkaik, N. J., de Vogel, C. P., Boelens, H. A., Verveer, J., Nouwen, J.**

643 **L., Verbrugh, H. A. & Wertheim, H. F. (2009).** Reclassification of *Staphylococcus aureus*

644 nasal carriage types. *J Infect Dis* **199**, 1820-1826.

645

646 **Vazquez, V., Liang, X., Horndahl, J. K., Ganesh, V. K., Smeds, E., Foster, T. J. & Hook,**

647 **M. (2011).** Fibrinogen is a ligand for the *Staphylococcus aureus* microbial surface

648 components recognizing adhesive matrix molecules (MSCRAMM) bone sialoprotein-binding

649 protein (Bbp). *J Biol Chem* **286**, 29797-29805.

650

Table 1

Strain	Relevant Features	Source
S. aureus		
Newman	Human clinical isolate, NCTC 8178	(Duthie & Lorenz, 1952)
Newman *clfA*5 *clfB*::Em^R	Derivative of strain Newman deficient in ClfA, ClfB. Frameshift mutation in the *clfA* gene, *clfB*::Em^R	(Fitzgerald *et al.*, 2006)
Newman *spa*	Derivative of strain Newman deficient in protein A. *spa*::Ka^R	(Higgins *et al.*, 2006)
NM1	Derivative of strain Newman deficient in ClfA, ClfB and protein A. Frameshift mutation in the *clfA* gene, *clfB*::Em^R *spa*::Ka^R	This study
NM2	Derivative of NM1. Replacement of *clfB*::Em^R with a deletion of the *clfB* gene isolated by allele exchange.	This study
BH1CC Δ*fnbAfnbB*	Derivative of MRSA strain BH1CC deficient in FnBPA and FnBPB	(Geoghegan *et al.*, 2013)
BH1CC Δ*fnbAfnbB* *spa*::Ka^R	Derivative of BH1CC Δ*fnbAfnbB* deficient in protein A. *spa*::Ka^R	This study
E. coli		
DC10B	*dam*^+Δ*dcm* Δ*hsdRMS endA1 recA1*	(Monk *et al.*, 2012)
Plasmids		
pALC2073	Vector allowing high level gene expression in	(Bateman *et al.*,

	the absence of inducer. $Ap^R Cm^R$	2001)
pALC2073*clfA*	pALC2073 containing full-length *clfA* gene.	A. Loughman, Trinity College Dublin. This study
pALC2073*clfA*ΔN1$_{40-220}$	pALC2073 *clfA* lacking region encoding residues 40-220 of ClfA	This study
pRMC2	Anhydrotetracycline-inducible expression vector with tight repression in the absence of inducer. $Ap^R Cm^R$	(Corrigan & Foster, 2009)
pRMC2*clfA*	pRMC2 containing full-length *clfA* gene from strain Newman	This study
pRMC2*clfA*ΔN1$_{40-220}$	pRMC2*clfA* lacking DNA encoding residues 40-220 of ClfA	This study
pRMC2*clfA*ΔN1$_{40-228}$	pRMC2*clfA* lacking DNA encoding residues 40-228 of ClfA	This study
pRMC2*clfA*ΔN1$_{40-210}$	pRMC2*clfA* lacking DNA encoding residues 40-210 of ClfA	This study
pRMC2*clfA*Δ$_{211-220}$	pRMC2*clfA* lacking region encoding residues 211-220 of ClfA	This study
pRMC2*clfA*ΔN1 $_{40-220}$ $_{+10Ala}$	pRMC2*clfA* lacking DNA encoding residues 40-220 and inserting DNA encoding 10 alanine codons.	This study
pRMC2*clfA*ΔSD$_{559-875}$	pRMC2*clfA* lacking DNA encoding residues 559-875	This study
pRMC2*clfA*ΔN1$_{40-220}$	pRMC2	This study

ΔSD$_{559\text{-}875}$	*clfA* lacking DNA encoding residues 40-220 and 559-875	
pRMC2*fnbB*	pRMC2 containing full-length *fnbB* gene from strain 8325-4	This study
pRMC2*fnbB*ΔN1$_{38\text{-}162}$	pRMC2*fnbB* lacking DNA encoding residues 38-167	This study
pRMC2*fnbB*ΔN1$_{38\text{-}162}\Delta$FnBR$_{481\text{-}811}$	pRMC2*fnbB* lacking DNA encoding residues 38-167 and 481-811	This study
pIMAYΔ*clfB*	Plasmid for creating *clfB* deletion of Newman by allelic exchange. Carries 500 bp fragment of DNA from upstream of and 500 bp downstream from *clfB*. Cm^R	(Mulcahy *et al.*, 2012)

652

653

Figure Legends.

655 **Fig. 1.** Schematic representation of ClfA derivatives. The N terminal signal sequence (SS) is

656 followed by three independently folded subdomains, N1, N2 and N3 comprising the A

657 domain. A flexible repeat region (R) links the A domain to the C-terminal Wall (W) and

658 Membrane (M) spanning region which contain the LPXTG motif that allows anchoring of the

659 protein to cell wall peptidoglycan by sortase A. The amino-acid coordinates of the signal

660 sequence (SS), N1, N2 and N3 subdomains, SD repeat region (R) and cell wall (W) and

661 membrane spanning (M) regions are indicated. The crystal structure of residues 229-545 of

662 subdomains N2N3 is shown (Protein Data Bank code 2VR3). Residues 211-229 comprising

663 the C- terminal region of subdomain N1 and the first residue of N2 as defined by the crystal

664 structure are highlighted. Amino acid coordinates of the predicted cleavage site of the

665 metalloprotease aureolysin are indicated.

666

667 **Fig. 2.** Expression of ClfA and ClfA derivatives. Dot immunoblot of intact bacteria (a) or

668 whole cell lysates (b) of *S. aureus* NM1 expressing ClfA, ClfA ΔN1 $_{40\text{-}220,}$ ClfA ΔN1 $_{40\text{-}228}$ or

669 ClfA ΔN1 $_{40\text{-}210}$ from pRMC2. Newman Δspa has been included as a positive control for

670 ClfA expression and NM1 carrying pRMC2 empty vector serves as a negative control. The

671 concentrations of inducer and the OD_{600} of the cells are indicated. Filters were probed with

672 ClfA mAb.

673

674 **Fig. 3**. Subcellular localization of ClfA, ClfAΔN1$_{40\text{-}220}$ and ClfAΔN1$_{40\text{-}210}$. Cellular fractions

675 from *S. aureus* expressing ClfA, ClfA ΔN1 $_{40\text{-}220}$ or ClfA ΔN1 $_{40\text{-}210}$ were probed with a) ClfA

676 mAb, b) anti-SdrE IgG or c) anti-ATPase IgG in a Western immunoblot. Fractions are

677 labelled cell wall (CW), membrane (CM) and cytoplasm (CP). Cytoplasm fractions were

678 stained with Coomassie brilliant blue (d).

679

680 **Fig. 4.** Subcellular location of ClfA, ClfAΔN1$_{40-220 + 10 \text{ Ala}}$ and ClfAΔ$_{211-220}$. Cellular fractions

681 from *S. aureus* expressing ClfA induced with 800 ng ml^{-1} ATc, ClfAΔN1$_{40-220 + 10 \text{ Ala}}$ and

682 ClfAΔ$_{211-220}$ induced with ATc (1200 ng ml^{-1}) were probed with a) ClfA mAb or b) anti-SdrE

683 IgG in a Western immunoblot. Fractions are labelled cell wall (CW), membrane (CM) and

684 cytoplasm (CP).

685

686 **Fig. 5**. Growth inhibitory effects of expression of ClfAΔN1$_{40-220.}$

687 *S. aureus* NM1 (pRMC2*clfA*) and NM2 (pRMC2*clfA*ΔN1$_{40-220}$) cells were induced with ATc

688 (1200 ng ml^{-1}) and plated on agar incorporating Cm. Colonies resistant to Cm were replica

689 plated onto agar containing Em. Bars represent the mean percentage CFU based on the total

690 Cm resistant CFU. Error bars represent the standard deviation on the mean of three

691 independent experiments. Statistical analysis was performed using Students *t* test.

692

693 **Fig. 6.** Subcellular location of ClfA, ClfAΔSD$_{559-875}$ and ClfAΔN1$_{40-220}$ΔSD$_{559-875}$

694 Cellular fractions from *S. aureus* expressing ClfA, ClfAΔSD$_{559-875}$ or ClfAΔN1$_{40-220}$ ΔSD$_{559-}$

695 $_{875}$ probed with ClfA mAb. Fractions are labelled cell wall (CW), cytoplasmic membrane

696 (CM) and cytoplasm (CP).

697

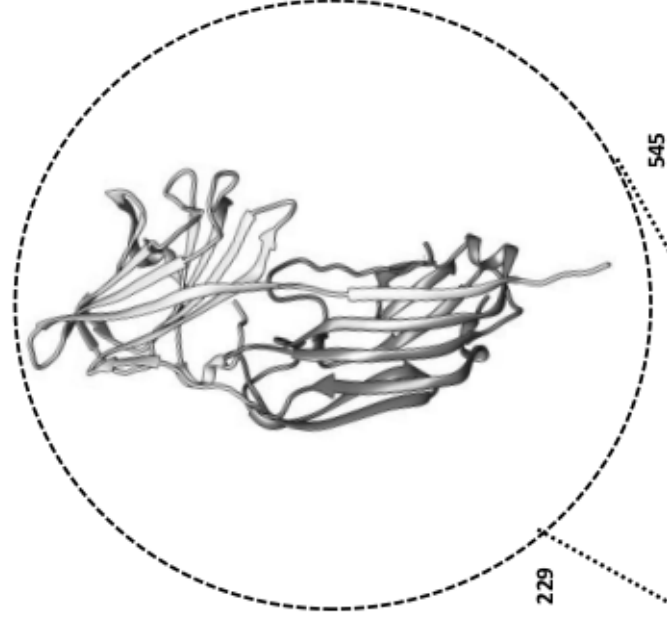698   **Fig. 7.** Subcellular location of FnBPB, FnBPBΔN1$_{38-162}$ and FnBPBΔN1$_{38-162}$ΔFnBR$_{481-811}$.

699   a) Cellular fractions from *S. aureus* expressing FnBPB or FnBPBΔN1$_{38-162}$ probed with

700   biotinylated fibronectin and detected with HRP conjugated streptavadin. Fractions are

701   labelled cell wall (CW), membrane (CM) and cytoplasm (CP).

702   b) Cellular fractions from *S. aureus* expressing FnBPBΔN1$_{38-162}$ΔFnBR$_{481-811}$ probed with

703   monoclonal rabbit anti-FLAG IgG. Fractions are labelled cell wall (CW), membrane (CM)

704   and cytoplasm (CP).

705

706

P$_{211}$RMRAFSLAA$_{220}$V$_{221}$AADAPAAG$_{229}$

ClfA

| SS | N1 | N2 | N3 | R | W | M |

1   39   221   368   560   876   933

LPDTG

ClfA ΔN1 $_{40-220}$

| SS | N2 | N3 | R | W | M |

1   39   221   368   560   876   933

LPDTG

ClfA ΔN1 $_{40-228}$

| SS | N2 | N3 | R | W | M |

1   39   229   368   560   876   933

LPDTG

ClfA ΔN1 $_{40-210}$

| SS | N2 | N3 | R | W | M |

1   39   211   368   560   876   933

LPDTG

229   545

(a)

Strain	ATc concentration (ng ml^{-1})	OD$_{600}$ 0.5	0.25	0.125	0.063
Newman *spa*	0	●	●	●	●
NM1 pRMC2	0				
NM1 pRMC2 *clfA*	0				
NM1 pRMC2 *clfA*	100				
NM1 pRMC2 *clfA*	200	○	○		
NM1 pRMC2 *clfA*	400	●	○		
NM1 pRMC2 *clfA*	800	●	●	●	
NM1 pRMC2 *clfA*	1200	●	●	●	●
NM1 pRMC2 *clfA* ΔN1 $_{40\text{-}220}$	0				
NM1 pRMC2 *clfA* ΔN1 $_{40\text{-}220}$	1200				
NM1 pRMC2 *clfA* ΔN1 $_{40\text{-}210}$	0				
NM1 pRMC2 *clfA* ΔN1 $_{40\text{-}210}$	1200	●	●	○	○
NM1 pRMC2 *clfA* ΔN1 $_{40\text{-}228}$	0				
NM1 pRMC2 *clfA* ΔN1 $_{40\text{-}228}$	1200				

(b)

Strain	ATc concentration (ng ml^{-1})	
Newman *spa*	0	●
NM1 pRMC2 1200	1200	
NM1 pRMC2 *clfA*	1200	●
NM1 pRMC2 *clfA*ΔN1$_{40\text{-}220}$	1200	●
NM1 pRMC2 *clfA*ΔN1$_{40\text{-}210}$	1200	●
NM1 pRMC2 *clfA*ΔN1$_{40\text{-}228}$	1200	●

(a)

	ClfA			ClfA ΔN1$_{40-220}$			ClfA ΔN1$_{40-210}$		
	CW	CM	CP	CW	CM	CP	CW	CM	CP

170 kDa →

130 kDa →

100 kDa →

ClfA mAb

(b)

170 kDa →

Anti-SdrE IgG

(c)

55 kDa →

40 kDa →

Anti-ATPase IgG

(d)

ClfA   ClfA ΔN1$_{40-220}$ ClfA ΔN1$_{40-210}$

130 kDa →
100 kDa →
70 kDa →
55 kDa →
40 kDa →
35 kDa →
25 kDa →

Cytoplasm

(a)

	ClfA			ClfA ΔN1$_{40-220+10\ Ala}$			ClfA Δ$_{211-220}$		
	CW	CM	CP	CW	CM	CP	CW	CM	CP

170kDa →

130kDa →

100kDa →

ClfA mAb

(b)

170kDa →

Anti-SdrE IgG

(a)

FnBPB

CW   CM   CP

FnBPB ΔN1 $_{38-162}$

CW   CM   CP

170 kDa

130 kDa

100 kDa

Biotinylated Fn

(b)

FnBPB ΔN1 $_{38-162}$ ΔFnBR

CW   CM   CP

170 kDa
130 kDa
100 kDa
70 kDa

50 kDa

Anti-FLAG
IgG