

A STRUCTURAL AND BIOCHEMICAL STUDY ON THE STAPHYLOCOCCAL SURFACE  
PROTEINS INTERACTIONS WITH HOST EXTRACELLULAR PROTEINS

A Dissertation

by

DHARMANAND RAVIRAJAN

Submitted to the Office of Graduate and Professional Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Magnus Hook
Co-Chair of Committee,	Vannakambadi K Ganesh
Committee Members,	Peter Davies
	David Reiner
	Sara Lawhon
Head of Department,	Warren Zimmer

December 2019

Major Subject: Medical Sciences

Copyright 2019 Dharmanand Ravirajan

## ABSTRACT

*Staphylococcus aureus* (*S. aureus*), a gram positive bacterium, expresses multiple surface proteins (MSCRAMMs) that target host proteins for its colonization on the host tissue and later for the pathogenesis. ClfA is a virulence factor that interacts with fibrinogen (Fg) and the interaction of ClfA with the Fg is crucial for the pathogenesis. The virulence potential of ClfA/Fg interactions in the pathogenesis of *S.aureus* had been proven in sepsis and endocarditis infection animal models. ClfB is a chief colonizing factor that interacts with multiple host proteins like loricrin, cytokeratin-10, dermokine and Fg. The binding mechanism and structure of the ClfB:Ln complex is crucial for designing inhibitors against ClfB that aid in preventing the host colonization of *S. aureus*. My thesis work includes two chapters (1) crystal structure of ClfA with a proteolytic fragment of Fg (FgD). (2) Crystal structure of ClfB in complex with Ln.

Chapter-1 results describe the structure of ClfA/FgD complex and identify the residues involved in the interactions of ClfA and Fg. The results of this chapter explain the structural basis for the Fg mediated virulence of ClfA and the mechanism of Fg interactions of ClfA of *S. aureus* and its functional homologs fbl of *S. lugdunensis* and SpsD of *S. pseudintermedius*. The ClfA interacting region in Fg overlaps with the M1 protein of *S. pyogenes* and Fg interacting site in ClfA overlaps with the ‘tefibazumab’ interacting site.

Chapter-2 results describe the structure of the ClfB/Ln complex and the mechanism of ClfB/Ln interactions. Comparison of ClfB/Ln structure with other ClfB ligand peptide complexes (Ck10, DK and Fg) shows that ClfB/Ln binding mechanism is unique and different from the previously known Dock, Lock and Latch mechanism of ClfB ligand binding. My study was the first structure to show that, the Ln peptide can induce the dimerization of ClfB which is not observed in the

interaction of ClfB with other known ligands. Furthermore, the crystal structure led to the discovery of an additional ligand binding site on top of the N3 subdomain in ClfB which is also present in other MSCRAMMS that forms the basis for chapter-1.

## CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a thesis committee consisting of Professor Dr. Magnus Hook (Advisor) and Dr. Vannakambadi K Ganesh (Co-Advisor) of the Department of Center for Infectious & Inflammatory Diseases, CIID, Texas A&M Health Sciences Center, Houston-Texas and Professor Dr. Sara Lawhon of the College of Veterinary Medicine, Texas A&M University, College Station-Texas.

The X-ray Diffraction data experiments of protein crystals used in this dissertation (Chapter-II & Chapter-III) were performed in collaboration with Dr. Nagarajan Venugopalan at Argonne National Laboratory, Lemont, Illinois, USA.

The Surface Plasmon Resonance (SPR) binding experiments and peptide inhibition experiments were performed with the help of Dr. Xiaowen Liang, CIID-TAMUHSC.

The recombinant protein MBP tagged Fg- $\gamma$  module was expressed and purified by Dr. Wen Liu of the protein core facility, CIID-TAMUHSC.

The A-domain of SpsD (strain ED99) used in Chapter-II this study was obtained from our collaborator Dr. Ross Fitzgerald, University of Edinburgh, Scotland, United Kingdom.

The GST-tagged Fg $\alpha$ , CK10 & Ln constructs used in Chapter-III of this study was obtained from our collaborator Dr. Tim J. Foster, Trinity College, Dublin, Ireland.

Analysis of the crystal structure data, Determination of binding sites in Chapter-II and Chapter-III were performed with the help of Dr. Vannakambadi K Ganesh of the Center for Infectious & Inflammatory Diseases, TAMUHSC, Houston-Texas.

This work was made possible in part by NIH under Grant Number A1020624.

## NOMENCLATURE

<i>S.aureus</i>	<i>Staphylococcus aureus</i>
MRSA	Methicillin resistant <i>Staphylococcus aureus</i>
<i>S.lugdunensis</i>	<i>Staphylococcus lugdunensis</i>
<i>S.pseudintermedius</i>	<i>Staphylococcus pseudintermedius</i>
<i>S.intermedius</i>	<i>Staphylococcus intermedius</i>
<i>S.epidermidis</i>	<i>Staphylococcus epidermidis</i>
SdrG	S.epidermidis surface protein G
<i>S.pneumoniae</i>	<i>Streptococcus pneumoniae</i>
GAS	Group A- <i>Streptococcus</i>
<i>E.faecium</i>	<i>Enterococcus faecium</i> , a gram-positive bacterium in the genus <i>Enterococcus</i>
CoNS	Coagulase Negative Staphylococci
Coa	Coagulase
vWbp	Von-Willebrand factor binding protein
MSCRAMMS	Microbial Surface Components Recognizing Adhesive Matrix Molecules
SERAMMS	Secretory Extracellular Recognizing Adhesive Matrix Molecules
DeV-IgG-fold	a DE variant of Immunoglobulin-G fold
ClfA	Clumping factor-A
ClfB	Clumping factor-B

Fbl	fibrinogen binding protein of <i>S. lugdunensis</i>
SpsD	<i>S. pseudintermedius</i> surface protein-D
SpsO	<i>S. pseudintermedius</i> surface protein-O
SpsL	<i>S. pseudintermedius</i> surface protein-L
M1	A virulence protein of GAS (M1-protein)
Fg	Fibrinogen
FgD	A proteolytic fragment of fibrinogen, D-fragment of fibrinogen (Plasmin)
Fib	Fibrin
D-dimer	A proteolytic fragment of fibrin (Thrombin)
Ca or Ca <sup>2+</sup>	A Calcium ion (divalent)
NAG	N-acetyl glucosamine, a disaccharide
NAM	N-acetyl muraminic acid, a disaccharide
TMAO	Trimethyl Ammonium Oxide
Fn	Fibronectin
Plg	plasminogen
Ln	Loricrin
CK10	Cytokeratin
DK	Dermokine
αMβ2	alpha-M,beta-2 integrin (neutrophil)
αIIbβ3	alpha-2b-beta-3 integrin (platelet)
DLL	Dock, Lock and Latch
ATCC	American tissue culture consortium

SSTI	Soft skin tissue infection
SSTS	Streptococcal septic toxic syndrome
SAB	<i>Staphylococcus aureus</i> bacteremia
Nosocomial	hospital associated infections
IE	infective endocarditis
CDC	center for disease control
FDA	food and drug administration agency
tRNA	Transfer ribonucleic acid
MGE	mobile genetic element
MDR	multi-drug resistant
SCC <i>mec</i>	A mobile genetic element present in MRSA
<i>mecA</i>	a gene inside SCC <i>mec</i> mobile genetic element that confers resistance to the beta lactam antibiotic methicillin
<i>PBP</i>	penicillin binding protein
<i>Tn</i>	transposon, a mobile genetic element
<i>Tn1546</i>	an MGE derived from <i>E. faecalis</i> present in <i>S. aureus</i> , confers resistant to vancomycin
Van	vancomycin
vanH	a dehydrogenase enzyme that reduces pyruvate to lactate
vanA	a ligase enzyme that attaches the lactate to depsipeptide D-Ala-D-Lac
D-Ala	D-enantiomeric form of alanine (right handed form)

D-Lac	A stereoisomer of lactate which rotates polarized light counterclockwise (-)
STRIVE	STaphylococcus aureus SuRgical Inpatient Vaccine Efficacy
<i>C.albicans</i>	<i>Candida albicans</i>
<i>Als3p</i>	<i>Candida albicans</i> surface protein-3
VAP	ventilator associated pneumonia
TPCK	Tosylamide-2-phenylethyl chloromethyl ketone
TLCK	Tosyl-L-lysine chloromethyl ketone hydrochloride
SDS	Sodium lauryl sulfate
PAGE	polyacrylamide gel electrophoresis
NR	non-reducing
GPRP	gly-pro-arg-pro amide peptide
GHRP	gly-his-arg-pro amide peptide
PEG	polyethylene glycol
MPD	2-methyl-2,4-pentanediol
HKL2000	software suite to process X-ray diffraction images
PHASER-MR	software suite for molecular replacement method based solving crystal structure data
Refmac	automated refinement tool for structure solution
COOT	manual refinement tool for structure solution
MR	molecular replacement method
IPTG	isopropyl 1-thio- $\beta$ -D-galactopyranoside

ITC	isothermal titration calorimetry
SPR	surface plasmon resonance
$\Delta G$	gibbs free energy change of a reaction, kcal mol <sup>-1</sup>
$\Delta^{\circ}G$	theoretical gibbs free energy change of reaction, kcal mol <sup>-1</sup>
$\Delta H$	enthalpy change of a reaction, kcal mol <sup>-1</sup>
$-T\Delta S$	entropy change of a reaction, kcal mol <sup>-1</sup>
$K_A$	Affinity constant
$K_D$	Dissociation constant
$\mu M$	micromolar
mM	millimolar
nM	nanomolar
Å	angstroms
K	kelvin
R	gas constant, 8.314 J/(mol.K)
T	temperature, degree Celsius
APS	Advanced photon source
$F_o$	observed structure factor
$F_c$	theoretically calculated structure factor
PDB	protein data bank
MODELLER	homology model and docking software
PyMOL	graphical molecular visualization tool

## TABLE OF CONTENTS

	Page
ABSTRACT.....	ii
CONTRIBUTORS AND FUNDING SOURCES .....	iv
NOMENCLATURE .....	v
TABLE OF CONTENTS.....	x
LIST OF FIGURES .....	xiv
LIST OF TABLES.....	xvi
CHAPTER I INTRODUCTION AND LITERATURE REVIEW .....	1
Background on <i>Staphylococci</i> .....	1
<i>Staphylococcus aureus</i> .....	2
<i>Staphylococcus lugdunensis</i> .....	2
<i>Staphylococcus pseudintermedius</i> .....	3
Emerging resistance of <i>Staphylococci</i> .....	3
Infections of <i>Staphylococci</i> .....	5
Staphylococcal vaccine strategies.....	6
CHAPTER II STRUCTURAL AND BIOCHEMICAL STUDIES ON FIBRINOGEN $\gamma$ -	
CHAIN BINDING MSCRAMMS OF STAPHYLOCOCCI .....	9
Introduction to Fibrinogen .....	10
Introduction to F $\gamma$ -chain binding MSCRAMMS .....	11
Clumping Factor-A .....	14
Fibrinogen binding protein of <i>S. lugdunensis</i> .....	15
<i>Staphylococcus pseudintermedius</i> surface protein-D .....	16
Fibrinogen binding mechanism of F $\gamma$ -chain binding MSCRAMMS.....	17
Significance of the Study .....	18
Preliminary Data and Hypothesis .....	19
Innovation .....	19
Materials and Methods.....	20
Cloning of ClfA construct.....	21

Cloning of Fbl construct .....	21
Cloning of SpsD construct .....	22
Site directed mutagenesis.....	23
Crystallization of ClfAcc/Fg complex .....	23
Generation and Purification of FgD.....	24
Purification of ClfAcc/FgD protein complex .....	24
Crystallization of ClfAcc/FgD protein complex.....	25
Cloning of Fgy module construct.....	26
Expression and Purification of Fgy module protein .....	27
Results of Chapter-II .....	28
Interaction between ClfAcc and Fibrinogen .....	28
Crystallization trials of ClfAcc/Fibrinogen protein complex .....	31
Crystallization trials of ClfAcc/FgD protein complex.....	32
Data Determination.....	33
Structural Analysis of ClfAcc/FgD Complex .....	34
Construction of the Fgy Module .....	38
Interaction of ClfAcc with Fgy Module.....	39
Fg Binding Site Analysis of Fbl and SpsD .....	41
Homology Modelling of FgD/Fbl and FgD/SpsD complex .....	42
Generation of FgD/Fbl structural model.....	43
Generation of FgD/SpsD structural model .....	45
Interaction between Fbl and Fibrinogen .....	46
Interaction between SpsD and Fibrinogen.....	47
Mapping of N3 binding site residues in ClfAcc, Fbl and SpsD that interacts with Fg.	49
Fibrinogen Interaction Studies of N3 site mutants of ClfAcc, Fbl and SpsD.....	50
Discussion of Chapter-II.....	50
ClfAcc:FgD lacks the electron density for the last C-terminus of the Fg $\gamma$ (393- 411)..	51
The Fgy Module studies with ClfAcc links the low and high affinity site in ClfAcc...	51
Binding Site analysis of ClfA/Tefibazumab in comparison with ClfAcc/FgD .....	51
ClfAcc, Fbl and SpsD share a conserved N3 binding site for FgD interaction .....	54
Biological Significance of ClfAcc/FgD interactions .....	54
ClfAcc inhibits Fg binding to platelet and neutrophil integrins .....	54

ClfAcc inhibits fibrin polymerization at the $\gamma$ - $\gamma$ chain interface of D-dimer.....	55
Analysis of ClfAcc/FgD structure at the crystal lattice contact interfaces .....	56
ClfAcc interacts at the coiled coil region of Fgy chain.....	57
M1 and ClfAcc targets an overlapping site in coiled coil region of Fgy chain ...	58
Future Directions for Chapter-II .....	61
Summary of Chapter-II.....	61

### CHAPTER III STRUCTURAL AND BIOCHEMICAL STUDIES ON MULTIPLE LIGAND

BINDING MECHANISM OF CLUMPING FACTOR-B of <i>Staphylococcus.aureus</i> .....	64
Introduction to Clumping factor-B of <i>S. aureus</i> .....	64
Introduction to host ligands of clumping factor-B of <i>S. aureus</i> .....	64
Multiple ligand binding mechanism of ClfB .....	65
Clumping factor-B is a key host colonization factor for <i>S. aureus</i> .....	65
Colonization is a risk factor for <i>S. aureus</i> infections.....	66
Mupirocin resistant <i>S. aureus</i> .....	66
Significance of the study.....	67
Hypothesis .....	67
Approach .....	68
Innovation .....	69
Materials and Methods.....	69
Bacteria and Growth Conditions.....	69
Cloning of ClfB and ClfB mutants by site directed mutagenesis .....	69
Expression and Purification of Recombinant Proteins .....	71
Isothermal Titration Calorimetry Assay .....	71
Solid Phase Assay.....	72
Surface Plasmon Resonance Spectroscopy.....	72
Peptide Inhibition Assay.....	73
Gel Permeation Chromatography .....	74
Native PAGE .....	74
X-ray Crystallography .....	75
Results of Chapter-III .....	75
(CG12) <sub>2</sub> peptide mimics the loop region of Loricrin.....	75
YY41 peptide binding studies with ClfB indicate additional binding site in ClfB .....	78
Interaction of Fg with ClfB displays tight binding than the other ligands of ClfB .....	79
YY41 and (CG14) <sub>2</sub> peptide inhibits ClfB across multiple ligands.....	80
Crystal structure of ClfB in complex with (CG12) <sub>2</sub> and (CG14) <sub>2</sub> peptides of Ln .....	82

Structural analysis of ClfB/Ln peptide complex.....	83
Determination of ligand induced dimerization of ClfB by Ln peptide (CG12) <sub>2</sub> .....	85
Discussion of Chapter-III.....	88
Comparison of ClfB:CG14 with ClfB:Fgα peptide complex .....	89
Comparison of ClfB:CG14 with ClfB:CK10 peptide complex .....	91
Structure based amino acid comparison of ClfB ligand peptides .....	93
Future Directions for Chapter-III.....	95
Summary of Chapter-III.....	95
CHAPTER IV CONCLUSION.....	100
REFERENCES .....	101

## LIST OF FIGURES

FIGURE		Page
1	Schematic structure of fibrinogen and fibrin polymerization .....	11
2	ClfA enables staphylococcal agglutination with fibrin cables in vitro .....	12
3	Structural organization of proteins in the MSCRAMM protein family.....	16
4	Dock, Lock and Latch Mechanism of ClfAcc binding to Fg $\gamma$ (392-411).....	17
5	Structure based sequence alignment of Fbl(205-530) and ClfA(229-545).....	22
6	Structure based sequence alignment of FnbpA(189-505) and SpsD(167-502) .....	23
7	ITC binding assay of ClfAcc with Fg and Fg fragments .....	30
8	Confirmation analysis of ClfAcc/Fg or ClfAcc/FgD complex protein crystals .....	33
9	Crystal Structure of ClfAcc/FgD complex .....	36
10	Structural Analysis of ClfAcc-N3/Fg $\gamma$ Binding interface.....	37
11	Comparison of FgD (Apo) and ClfAcc bound FgD complex (Composite Model) ...	38
12	The MBP-Fg $\gamma$ (144-411) module.....	39
13	ITC binding assay of ClfAcc with MBP-Fg $\gamma$ (144-411) module .....	42
14	Strucural Analysis of N3 binding site of the Fbl/Fg $\gamma$ structure (Homology Model).	46
15	Strucural Analysis of N3 binding site of the SpsD/Fg $\gamma$ structure (Homology Model).....	48
16	ITC binding assay of Fbl with Fg and Fg fragments .....	49
17	ITC binding assay of SpsD with Fg and Fg fragments .....	50
18	ITC binding assay of N3 site mutants of ClfAcc, Fbl and SpsD with Fg and Fg fragments.....	53
19	SPR binding assay of ClfA and ClfA variants with Tefibazumab and Fg.....	57
20	Structural comparison of ClfAcc:FgD with ClfA:Tefibazumab.....	58
21	N3-binding Site analysis of ClfAcc:FgD and ClfA:Tefibazumab.....	58
22	Schematic representation of knob-hole interactions of Fg targeted by ClfA.....	61
23	N2-binding site analysis of ClfAcc that interacts at the Fg $\gamma$ -coiled coil region .....	63
24	M1 and ClfAcc targets a similar overlapping site at the coiled coil region in Fg $\gamma$ ....	64
25	The EF hand motif (310-321) in ClfA overlaps with Fg $\gamma$ -coiled coil region .....	64
26	Schematic Structure of Loricrin and Cytokeratin-10.....	67

27	Overview of ClfB ligand interactions .....	78
28	ITC binding assays of ClfB with Ln peptides.....	79
29	ITC binding assays of ClfB with CK10 peptides.....	81
30	ITC binding assays of ClfB with Fg .....	82
31	Peptide inhibition assays of ClfB binding with various ligands by SPR.....	85
32	Crystallization and diffraction data of ClfB:Ln peptide complex .....	85
33	Crystal structure of ClfB:Ln peptide complex (asymmetric unit) .....	87
34	Ligand induced dimerization of ClfB:CG14 peptide complex .....	89
35	Determination of a Ligand induced dimerization of ClfB .....	89
36	Comparison of binding sites in ClfB-Ln and ClfB-Fg $\alpha$ complex structures .....	91
37	ClfB-N3 binding site mutants Interactions with Fg.....	92
38	Comparison of binding sites in ClfB-Ln and ClfB-CK10 complex structures.....	94
39	Comparison of ClfB:CG12 structure (dimer) over ClfB:CK10 structure.....	94
40	Structure based sequence alignment of ClfB:ligand peptide complexes and Fg $\alpha$ .....	95
41	Structure based sequence alignment of ClfB:ligand peptide complexes .....	96

## LIST OF TABLES

TABLE		Page
1	Vaccination attempts of <i>S. aureus</i> .....	7
2	Surface protein genes and their contribution to <i>S. aureus</i> sepsis.....	13
3	List of primers for amplification of SpsD(167-502).....	23
4	ITC thermodynamic parameters for ClfAcc binding to Fg.....	31
5	X-ray Data Collection and Refinement Statistics for ClfA:FgD. ....	40
6	Low affinity trench site binding analysis in Fbl and SpsD.....	43
7	Comparison of residues in Fbl and ClfAcc at the N3 binding site of FgD.....	46
8	Comparison of residues in SpsD and ClfAcc at the N3 binding site of FgD.....	48
9	ITC thermodynamic parameters for Fbl binding to Fg.....	48
10	ITC thermodynamic parameters for SpsD binding to Fg.....	51
11	Fg binding signatures of ClfAcc, Fbl and SpsD .....	52
12	Theoretical free energy ( $\Delta iG$ ) calculations of various binding interfaces of ClfA:FgD complex.....	63
13	List of primers for site directed mutagenesis of ClfB.....	73
14	Summary of ligand binding affinities of ClfB(212-542).....	82
15	Crystallization data refinement and statistics of ClfB:CG12 and ClfB:CG14 peptide complexes .....	88
16	Comparison of ITC thermodynamic parameters of Fibrinogen binding with ClfAcc and ClfB.....	98

## CHAPTER I

### INTRODUCTION AND LITERATURE REVIEW

#### **Background on staphylococci**

The origin of Staphylococci can be tracked to the University of Aberdeen in Scotland around 1880(5). Alexander Ogston along with Louis Pasteur and Joseph Lister had worked on pus from human abscesses and discovered a gram-positive spherical bacterium which Dr. Ogston named “micrococci”(5). He later renamed the bacterium “staphylococci” because of the appearance of “bunches of grapes” seen under a light microscope. When these isolated bacteria were injected into healthy guinea pigs and mice, they formed abscesses and hence staphylococci became an infectious pathogen although they seldom seen as commensal in humans and other hosts organisms.

There are at least 17 subspecies of Staphylococci, which can be broadly divided into two groups based on the agglutination response in the presence of plasma (rabbit plasma anticoagulated with EDTA). Agglutination refers to the formation of fibrin clot due to the action of staphylocoagulase (a protein found in staphylococci) with the prothrombin complex (staphylothrombin). Staphylothrombin converts soluble Fg to an insoluble fibrin. If plasma is agglutinated in response to the presence of staphylococci it is (1) coagulase positive and if not then (2) coagulase negative species (CoNS) of staphylococci. *S. aureus* and *S. intermedius* are well known species of Staphylococci which are coagulase positive while *S. epidermidis* and *S. lugdunensis* are well known species of staphylococci which are coagulase negative. My research work is based on the surface proteins of two coagulase positive species of staphylococci, (1) *S. aureus* and (2) *S. pseudintermedius* and one coagulase negative species of Staphylococci, *S. lugdunensis*. A brief overview of these species are described below.

### ***Staphylococcus aureus***

*Staphylococcus aureus* is a Gram positive bacterium that contains a cell wall composed of a network of peptidoglycans. The peptidoglycans have two parts, the disaccharides, primarily, the disaccharides N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) are cross linked to the lysine side chain of a pentapeptidyl chain (peptidyl l-alanine-d-iso-glutamine-l-lysine-d-alanine-d-alanine) segment. The neighboring peptidoglycans are connected to each other by a pentaglycyl bridge. The main function of this cell wall structure is to provide protection against osmotic pressure and keeping the cytoplasmic contents of the bacterium (6).

Disrupting the peptidoglycan cell wall will result in the death of the bacterium and remains a target for multiple antibiotics. For example, (1) The glycosidic bond between disaccharides can be cleaved by lysozyme and (2) The assembly of the interpeptide bridge is inhibited by beta lactam antibiotics like penicillin. Apart from the peptidoglycan structure, the cell wall of *S. aureus* also contains lipoteichoic acids which are polymers of glycerol or ribitol phosphates.

### ***Staphylococcus lugdunensis***

*Staphylococcus lugdunensis* is a coagulase-negative staphylococci (CoNS) which was originally described around 20 years ago (7). Based on phenotypic analysis, the closest species to *S. lugdunensis* (Type N860297, ATCC43809) is *S. hominis* but the unique production of ornithine carboxylase by *S. lugdunensis* can be used to differentiate the two. *Staphylococcus lugdunensis* differs from *S. aureus* by production of an antigenically different heat stable nuclease and lack of pigmentation. Although, it is CoNS, *S. lugdunensis* causes severe infections similar to *S. aureus* infections. For example, prosthetic and native valve infections leading to endocarditis, osteomyelitis, peritonitis, urinary tract infections. *S. lugdunensis* has also been reported to cause skin and soft tissue infections (SSTI), mainly in the moist areas such as groin and mammae (7).

### ***Staphylococcus pseudintermedius***

*Staphylococcus pseudintermedius* is a coagulase positive bacterium and present primarily in dogs. It can cause surgical wound infections and pyoderma in dogs. Although *S. pseudintermedius* is a canine pathogen, it is capable of cross-infecting humans. It was believed that these infections are caused by *S.intermedius*, however Devriese and others described a unique species that shares features of *S.intermedius* and *S.aureus* capable of infecting animals (8). A tRNA intergenic length polymorphic analysis study on clinical isolates obtained from four different host (lung tissue from a cat in 1999, skin lesion from a horse in 1999, ear lesion from a dog in 2001, and liver tissue from a parrot in 2003) displayed similar electrophoretic patterns (8). Another 16S rRNA sequencing study on these strains indicated a 100% identity among these four strains (9). This new strain is classified a new staphylococcal species and was named *S. pseudintermedius* (pseud-false) meaning a false intermedius (9). More recently, *S. pseudintermedius* causing zoonotic infections in humans and increased multidrug resistant strains of *S. pseudintermedius* had been reported (10). Hence, the need for alternative therapeutic strategy against *S. pseudintermedius* infections is required.

### **Emergence of multidrug resistance *S.aureus***

Earlier for the treatment of staphylococcal skin infections (abscesses) the topical application of phenols was recommended up to the year 1940, penicillin was discovered (11). Penicillin is naturally produced by fungi *Penicillium sp.*, Penicillin kills bacteria by irreversibly blocking bacterial cell wall synthesis. Upon binding to PBP, the  $\beta$ -lactam amide bond in penicillin is broken and hence penicillin makes a covalent bond with the catalytic serine residue at the PBPs active site. Some bacteria have developed resistance to beta-lactams. These bacteria contain beta-lactamases, a broad class of enzymes with a serine residue that cleaves the reactive beta lactam

ring through an acyl-enzyme intermediate and confers resistance to the bacterium that expresses beta-lactamase. Thus beta-lactamase expressing staphylococci emerged within 2 years of regular usage of penicillin treatment.

In the 1950's, semi-synthetic beta-lactam antibiotics like methicillin replaced penicillin but within 10 years, methicillin resistant strains of *S.aureus* (MRSA) appeared in hospital settings. The resistance to methicillin is partially due to the presence of *mecA* gene (SCC*mec*), which encodes PBP2a (12). PBP2a is a transpeptidase which has low affinity to all the beta-lactams including methicillin. Due to the low affinity of the drug, PBP2a functions as a transpeptidase that crosslink the dipeptide (D-Ala-Ala) and build peptidoglycan cell wall. So, the bacteria can survive and divide even in the presence of beta-lactam drugs like penicillin and methicillin (13, 14).

Vancomycin is a tricyclic glycosylated nonribosomal peptide produced by the soil bacterium *Amycolatopsis orientalis*. Vancomycin binds to the terminal D-ala-D-Ala of the growing N-acetyl muramic acid moieties of the peptidoglycan cell wall and inhibits further cell wall synthesis of gram positive bacterium like MRSA.

Vancomycin resistance had been reported in *S.aureus*. The acquisition of *vanA* operon carried by transposon *Tn1546*, from *Enterococcus faecalis* seems to be the reason for the high level vancomycin resistance (>250µg/ml). VanH protein is a dehydrogenase which converts pyruvate to lactate and VanA protein is a ligase that allow the formation of the D-Ala-D-Lac depsipeptide and this substitution decreases the affinity of the glycopeptides like vancomycin to the cell wall, Thus vancomycin cannot bind to this modified D-Ala-D-Lac moiety and the bacteria is resistant to the vancomycin treatment(15). Currently, for the treatment of *S.aureus* infections caused by strains which have acquired both *vanA* and *mecA*, a combination of both vancomycin and oxacillin is recommended.

The cost of antibiotic resistance creates a huge economic burden for our society; up to \$24 billion each year in the US alone (16). Staphylococci is capable of causing opportunistic infections in humans and are part of the human microflora they pose considerable risk factor (17) (18). *S. aureus*, *S. lugdenensis* and *S. pseudintermedius* are antibiotic resistance strains associated infections. Resistance to antibiotics develop as a result of acquiring drug resistance genes via horizontal gene transfer of mobile genetic elements (MGE's) like phages, plasmids etc. They can exchange these resistance genes to or from other pathogenic bacterium like *Streptococcus pneumoniae*, *Enterococcus faecium* etc. Currently, 30% of *S. pneumoniae* are penicillin resistant and 22% of *E. faecium* are resistant to vancomycin and almost every strains of *S. aureus* is resistant to methicillin (19).

### **Infections of *Staphylococci***

*S. aureus* causes a multitude of infections such as lung infections (pneumonia), skin infections (necrotizing fasciitis), bone infections (osteomyelitis), blood stream infections (bacteremia, sepsis), medical device associated infections (endocarditis) and specific toxin related diseases (toxic shock syndrome). *Staphylococcus aureus* is the leading cause of death caused by nosocomial infections in hospitalized patients, in particular *S. aureus* bacteremia (SAB) infections occur at a frequency of 10-20 cases per 100,000 populations per year of which 30% of the patients die from SAB (20). In the US about 100,000 individuals are affected by methicillin resistant *S. aureus* (MRSA) associated invasive infections resulting in about 20,000 deaths. Of these infections 85% are hospital associated invasive infections (21) the rest being community acquired. Infections acquired during hospital visits and hospital stays greatly increases the chances of colonization and infection by antibiotic resistant strains of staphylococci due to alteration in immune system, surgeries, overuse of antibiotics, dialysis, and diabetic conditions. Recently, the

CDC estimated one in five pathogens causing hospital infections are associated with multidrug resistant strain (MDR) (22).

With the routine usage of antibiotics in animal feeds and livestock, resistant strains emerge and can transmit across to humans causing infections. The probability of getting infected by these multidrug resistant strains is high and poses difficulties to treat even minor infections like strep throat or skin infections in the nearby future. Thus, we are well into the post antibiotic era and in the verge of losing the war against these antibiotic resistance bacterium associated infections. A new non-antibiotic based treatment strategies are in urgent need to combat against these multi drug resistant staphylococcal infections.

### **Staphylococcal Vaccine Efforts**

Due to the expanded multi antibiotic resistance in *S. aureus*, preventive and alternative treatment procedures are desperately needed. An effective staphylococcal vaccine would be very useful to prevent/treatment for staphylococcal infections in humans caused by various species of staphylococci such as *S. aureus*, *S. lugdunensis* and *S. pseudintermedius*. Recent advances in *S. aureus* research has provided valuable information on the molecular pathogenesis of *S. aureus* infections yet designing an effective vaccine for *S. aureus* has proven to be challenging and difficult. Partly, this may be due to the fact that *S. aureus* appears to use a different set of virulence factors ranging from soft skin tissue infections (SSTI) to blood stream infections. In addition, *S. aureus* infections often occur in individuals with reduced immune defense. Recently, several pharmaceutical companies have conducted active and passive immunization trails summarized in Table-1.

Table-1. Vaccination attempts of *S.aureus*. Adapted from (2)

Product	Sponsor	Composition	Status
<b>Active Immunization</b>			
SA4Ag	Phizer	ClfA/MntC/CP5/CP8 Conjugated to CRM <sub>197</sub>	Phase II
NDV3	NovaDigm Therapeutics	rAls3p-N ( <i>C.albicans</i> surface protein that cross reacts with <i>S. aureus</i> ) plus Alum	Phase I
<b>Passive Immunization</b>			
MEDI4893	MedImmune	mAb binding to <i>S.aureus</i> toxin	Phase II
Tefibazumab(Aurexis)	Inhibitex	mAb(Clfa)	Phase II

Under the active immunization strategy, Phizer developed SA4Ag, which contains capsular antigens and an Fg binding surface protein of *S. aureus* (ClfA). In phase I trial, the SA4Ag produced robust immune response against all the four microbial components (ClfA/MntC/CP5/CP8), and also elicited effective opsonophagocytic activity. Currently SA4Ag is in phase II stage placebo-controlled safety and efficacy study in adults undergoing elective spinal fusion surgery is underway (referred to as STRIVE: *S. aureus* SuRgical Inpatient Vaccine Efficacy)(11). Als3 is an agglutinating protein of *C. albicans* also showed protective immune response against *S.aureus* infection in mice. NDV-3 vaccine is a recombinant version of the N-terminal region (416 amino acids) of the *C. albicans* protein Als3p with Alum formulation that showed a robust anti-Als3 IgG production and stimulation of Th2 response with IL-17a production in a phase I trial(12).

Under the passive immunization strategy, Medimmune developed MEDI4893 which is an investigational monoclonal antibody that specifically binds to and neutralizes alpha-toxin, a key *S. aureus* virulence factor (13, 14). The impressive thing about MED14893 is the prolonged half-life (80-120 days) in serum and effective alpha toxin neutralizing activity. MED14893 is being tested in Phase II trials on ventilator associated pneumonia (VAP) (15, 16).

Most staphylococcal surface proteins are established virulence factors and are excellent targets candidates to develop these non-antibiotic based drugs and vaccines against staphylococci.

## CHAPTER II

### STRUCTURAL AND BIOCHEMICAL STUDIES ON FIBRINOGEN $\gamma$ -CHAIN BINDING

#### MSCRAMMS OF *STAPHYLOCOCCI*

*Staphylococcus aureus* causes infective endocarditis (IE) and high mortality 20-40% (17-19). Infective endocarditis is vegetative growth of *S. aureus* causing infection on the surface of heart valves leading to infection and organ failure that eventually causes endocarditis (19). Infective endocarditis of *S. aureus* primarily caused by resistance strains can become difficult to treat and are associated with high mortality rates (20). The formation of platelet-bacteria thrombi is crucial for the development of IE (21). The platelet-mediated aggregation caused by *S. aureus* involves both the host and microbial molecules primarily, the host molecule fib/ Fg), the microbial surface proteins called MSCRAMMs (Microbial Surface Component Recognizing Adhesive Matrix Molecules) and secreted proteins called SERAMS (Secreted Extracellular Recognizing Adhesive Matrix Molecules) of *S. aureus*.

The platelet-mediated thrombi of *S. aureus* was also associated with sepsis, another lethal infection caused by *S. aureus* in humans (22). There are at least 18 different types of MSCRAMMS that are expressed on the surface of *S. aureus*. These MSCRAMMS all carry a conserved 'LPXTG' motif at their C-termini which is recognized by the enzyme sortase and anchors the MSCRAMM to the cell wall by covalently attaching with the peptidoglycans in the cell wall (23). *S. aureus* secretes Coagulase (coa) and VonWillebrand factor binding protein (vWbp) which are two well-known SERAMS. Coa activates the zymogen prothrombin that upon activation converts the soluble Fg to insoluble Fibrin (Fib) cables (24, 25).

McAdow et al, studied these 18 MSCRAMMS along with the coa and vWbp for their role in sepsis (8). They found out that Clumping factor-A (ClfA), an MSCRAMM, as the single most important

contributor for the sepsis (Table-2). McAdow et al, also found out that Coa and vWbp also involved in agglutination since it is known these molecules are able to convert soluble Fg to insoluble Fibrin cables (through the activation of zymogen prothrombin) (8).

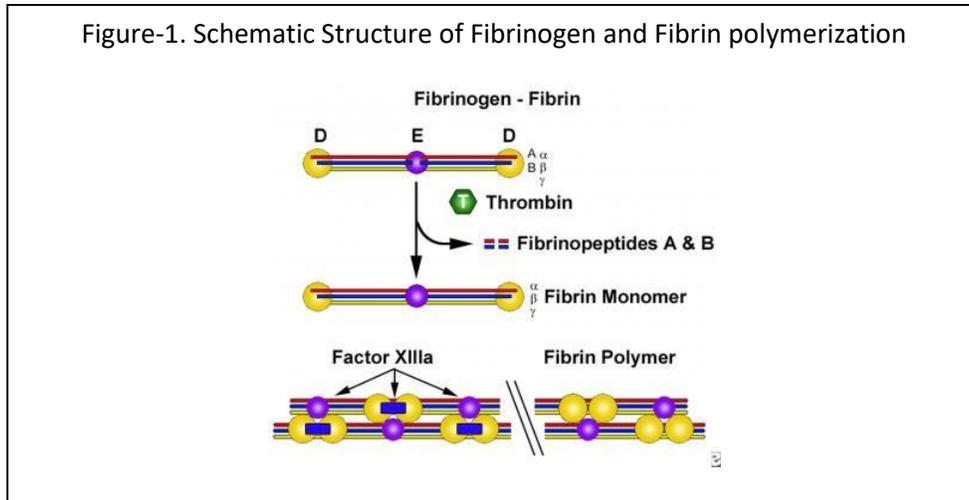
Collectively, it is clear the pathology of sepsis is a multifactorial process where, the MSCRAMM, ClfA was a key virulence factor and together with coa and vWbp causes the platelet/Fg/Fib mediated agglutination of *S. aureus*.

### **Introduction to Fibrinogen**

The adhesion of microorganisms to host tissues is the critical first step in the series of events that lead to clinically manifested infections. It has become evident that eukaryotic adhesive ECM components, that support adhesion of host cells, also serve as ligands for pathogenic microorganisms. Fg, the blood plasma coagulation protein, is also found in the ECM and plays important roles in wound-healing. During coagulation, Fg is proteolytically converted to fibrin which forms the structure of the blood clot. In addition, Fg is the major blood protein deposited on implanted biomaterials. *S. aureus* has long been known to form clumps in the presence of blood plasma. Hawiger et al, identified Fg as the plasma protein responsible for this phenomenon(26). *S. aureus*. In the 1950's Duthie clearly established Fg as the primary host factor responsible for staphylococcal clumping(27).

### **Fg/Fibrin**

Fg is a 340 kDa glycoprotein that is found in plasma at a 3mg/ml concentration in humans where it plays an important role in blood coagulation and wound healing. Fg is a hexamer that consists of three pairs of polypeptide chains ( $A\alpha B\beta\gamma$ ) which are inter linked covalently via disulfide bridges (28, 29). The molecule has three globular domains at the distal ends on both sides separated from the central 'E' domain. The central E domain is formed by the amino terminal ends of the chains



( $\alpha\beta\gamma$ ). During fibrin polymerization (fibrin clot), the amino terminal residues (fibrinopeptide A of  $\alpha$  and fibrinopeptide B of  $\beta$ ) are cleaved by thrombin which uncovers a site (knob) that is complementary to a preformed hole located in the globular domain of the  $\beta$  and  $\gamma$ - chains(30). This knob-hole interaction step allows molecules to move laterally, granting thickness to the fibrin fibers(31). The fibrin clot is stabilized via non covalent interactions between two neighboring distal globular  $\gamma$ -regions ( $\gamma$ - $\gamma$  dimer) (32) and additional iso-peptide crosslinking ( $\gamma$ -glutamyl-lysine) at the C-terminal of the  $\gamma$ -chains, catalyzed by Factor XIIIa (Figure-1) (33, 34).

Blood clotting needs to be precisely controlled to maintain a balance between hemorrhage and thrombosis(35). Fibrin has transient existence and fibrin clots are proteolytically processed by plasmin(36). Fg has three calcium binding sites, two of which are found within the distal domains and the third site is located in the central domain (31, 37). Calcium was found to stabilize Fg structure by limiting the course of plasmin digestion of Fg, and in absence of calcium, the  $\gamma$ -chain end of the D domain gets progressively shortened (38-40).

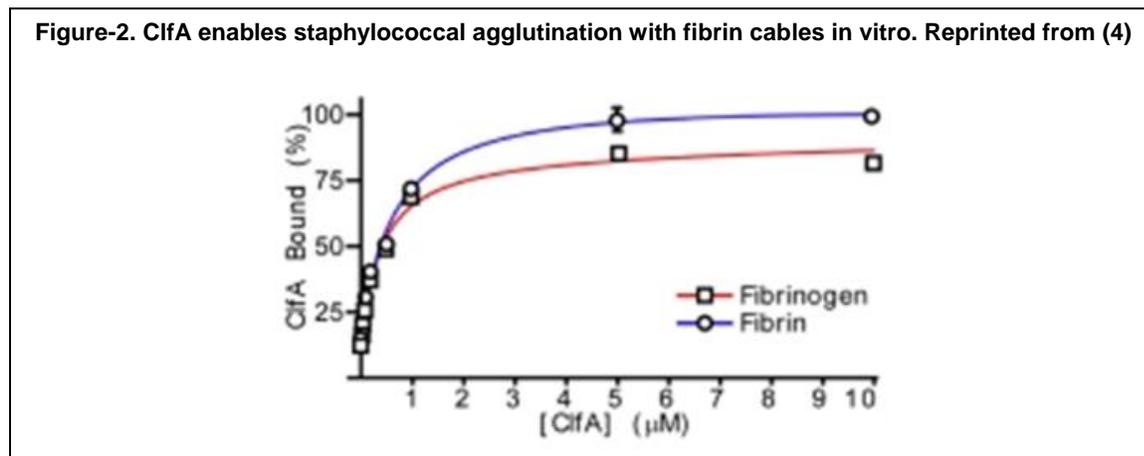
### Introduction to MSCRAMMS

In *S. aureus*, there are 18 known MSCRAMMs reported so far and most of them are known virulence factors. They all have a secretory signal sequence at their N-terminus and a ‘LPXTG’

cell wall anchor motif (sortase site) at the C-terminus, a hydrophobic cell wall- and membrane-spanning domain and a positively charged tail(6).

Clumping factor-A (ClfA) and Clumping factor-B (ClfB) are two well-known MSCRAMMs which are structurally related Fg binding proteins (41, 42). ClfA and ClfB share a common ‘R-region’, which is composed of stretch of dipeptide Ser-Asp repeats (SD repeats). This R-region connects the cell wall spanning domain to the ligand binding domain (A-domain) and act as a flexible stalk that projects the ligand binding domain into the extracellular matrix of the host (Figure-2).

In *S.aureus*, there are at least 3 other MSCRAMM proteins (SdrC, SdrE, and SdrD) that share the SD repeats containing R-region. In addition to the SD-repeats, the above three MSCRAMM’s also contain two (SdrC), three (SdrE) and five (SdrD) proline rich repeat regions (B-repeat) of around 100-110 residues/repeat (Figure-1) (43). The B-repeat is not present in both ClfA and ClfB.



**Table 2: Surface protein genes and their contribution to *S. aureus* sepsis. Reprinted from (4).**

Genotype	P values	Median survival time (hours ± SEM)	Genotype	P values	Median survival time (hours ± SEM)
wild-type	-	24 (1.6)	<i>isdB</i>	0.0243	30 (3.2)
<i>srtA</i>	<0.0001	>240	<i>sasA</i>	0.0004	36 (7.3)
<i>sasF</i>	1.000	24 (1.6)	<i>isdC</i>	<0.0001	36 (1.2)
<i>sdrC</i>	0.5416	24 (1.2)	<i>vwb</i>	<0.0001	36 (2.6)
<i>sdrD</i>	0.5416	24 (1.2)	<i>fnbpA</i>	0.0004	48 (5.5)
<i>sasD</i>	0.3415	24 (2.0)	<i>sasB</i>	<0.0001	48 (7.4)
<i>isdA</i>	0.3116	24 (1.8)	<i>sasC</i>	0.0011	54 (8.8)
<i>sasG</i>	0.1462	24 (0)	<i>fnbpB</i>	<0.0001	60 (8.0)
<i>clfB</i>	0.0888	24 (1.2)	<i>coa</i>	<0.0001	72 (12.5)
<i>sdrE</i>	0.0888	24 (4.8)	<i>adsA</i>	<0.0001	96(16.7)
<i>isdH</i>	0.0143	24 (2.0)	<b><i>clfA</i></b>	<b>&lt;0.0001</b>	<b>120 (15.3)</b>

Despite the similar structural organization of the A-domain among the Clf-Sdr protein family members, a local multiple protein sequence alignment of the ligand binding A-domain shows that Sdr-Clf proteins share a low level of ~25-27% sequence identity. ClfA and ClfB bind to Fg but on different chains/regions in Fg. In addition, ClfA can bind to Factor-I (44) and ClfB alone can bind to cytokeratin-10 (45, 46), loricrin (47) and dermokine (7). Thus, multiple ligand binding is a

characteristic feature of these MSCRAMMs and is the reason for their naming owing their adhesive properties against the extracellular molecules of the host.

### **Clumping factor-A of *S. aureus***

ClfA can binds to both soluble Fg and insoluble Fib (Fig-3) (8). ClfA targets a distinct site in Fg that serves multiple host function and it was known that ClfA inhibits Fg of its normal functions. The last four residues 'AGDV' at the C-terminus of the  $\gamma$ -chain of Fg (408-411) also bind to the platelet integrin receptor ( $\alpha$ IIb $\beta$ 3) (9, 48-51). This interaction is crucial for the host during normal wound healing where the association of Fg-platelet result in the formation of the 'platelet-plug' during initial clotting reaction. Interaction of ClfA at the platelet-integrin site of fg results in altered platelet aggregation and *S. aureus* exploits this interaction for the formation of thrombi during infection such as infective endocarditis (IE) (52). The genetic elimination of this site Fg- $\gamma$  (407-411) in Fg, predisposes mice to *S. aureus* infection and poses an impediment to microbial clearance (4).

Upstream to this platelet-integrin interaction site in Fg, the region in the  $\gamma$ -chain (390-396) is known to interact with the neutrophil(53) integrin ( $\alpha_M\beta_2$ ) that leads to neutrophil activation and phagocytosis (49, 54). In a peritonitis model in mice, Flick et.al., found that elimination of the  $\alpha_M\beta_2$  binding region Fg $\gamma$ (390-396) in Fg resulted in diminished capacity to clear *S. aureus* (49, 53-55).

Thus, *S. aureus* deploys ClfA to interact with Fg/Fib, thereby inhibiting the Fg dependent host function such as platelet aggregation and neutrophil activation (Phagocytosis).

Similar to ClfA of *S. aureus*, there are Fg binding MSCRAMMs in other species of staphylococci. Fbl is a Fg binding MSCRAMM of *S. lugdenensis* which is reported to cause invasive endocarditis infection in humans (56-60). SpsD is a Fg binding MSCRAMM of *S.pseudintermedius* which is

reported to cause infection in canine (61, 62). The Fbl and SpsD have a ligand binding A-domain and based on the similar structural domains (IgG like) of these proteins, it is expected that the Fbl and SpsD adopts a hydrophobic trench using their N2N3 subdomains to interact with Fg at the same region in Fgy(392-411) like ClfA (59, 60). The binding mechanism is expected to be Dock, Lock and Latch (DLL) and the structural information of Fbl/Fg and SpsD/Fg is not available.

### **Surface proteins of *S.lugdunensis***

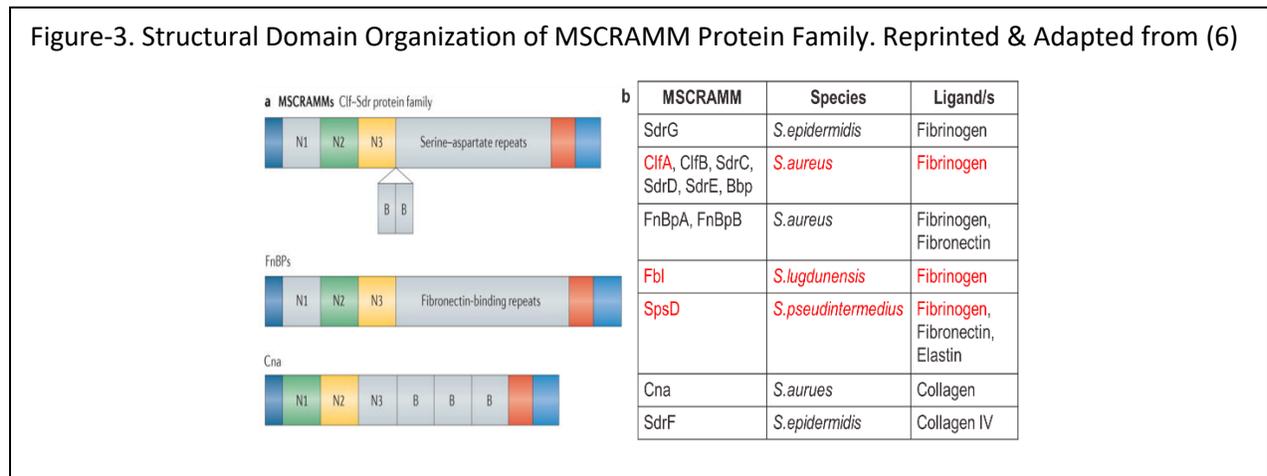
So far, the only known surface protein of *S. lugdunensis* is Fbl (Fg binding protein of *S. lugdunensis*). The Fbl is also a protein that has characteristic features of MSCRAMMs with a N-terminal signal sequence which is necessary for directing the protein from the cytoplasm to the surface via the Sec secretory pathway, A-domain containing repeated IgG-like folded domains, Ser-Asp repeats, and LPXTG anchoring motif followed by membrane spanning hydrophobic domain and a short positively charged cytoplasmic tail (59, 60).

The A-domain of Fbl shares ~60% sequence similarity with ClfA which is a known Fg binder and indeed the N2 and N3 subdomains (IgG folds) shown to bind to full length Fg. The binding mechanism is believed to be the established Dock, Lock and Latch mechanism of ligand binding (59, 60).

### **Surface protein of *S. pseudintermedius***

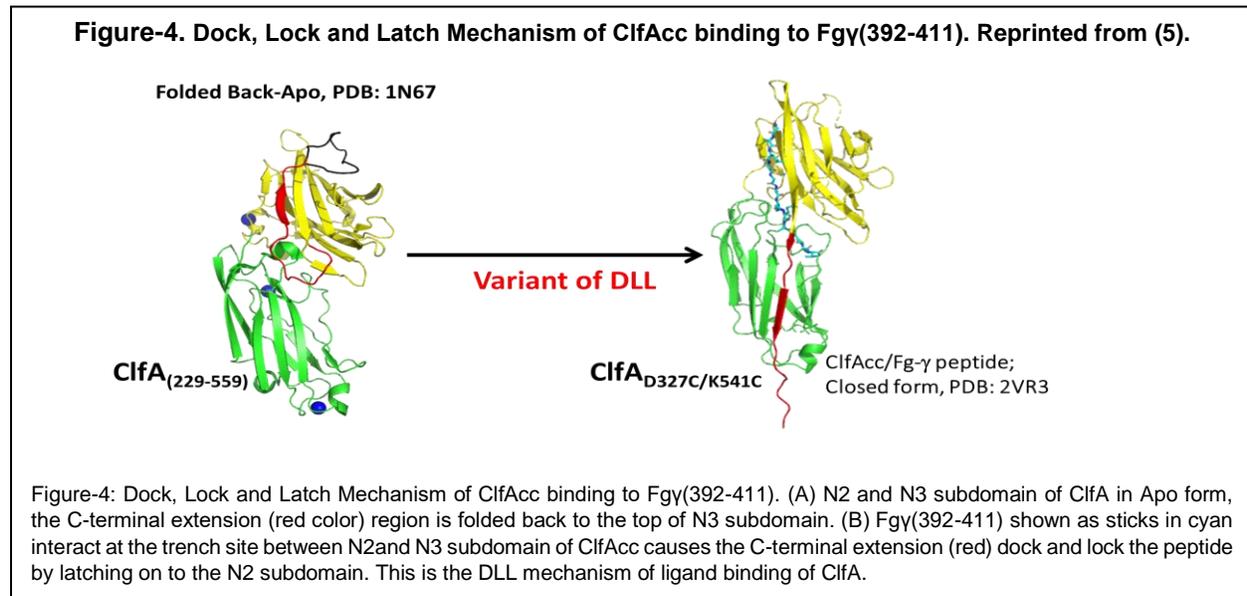
Currently, there is a lack of knowledge about the structure and function of *S. pseudintermedius* adhesins. Based on sequence analysis of the available strains, *S. pseudintermedius* encodes three putative MSCRAMMs: SpsD, SpsL, and SpsO. These are characterized as potential MSCRAMMs based on their genetic structure, which predicts that they contain the hallmarks of MSCRAMMs, N-terminal signal sequence which is necessary for

directing the protein from the cytoplasm to the surface via the Sec secretory pathway, A-domain containing repeated IgG-like folded domains, B repeats, and LPXTG anchoring motif followed by membrane spanning hydrophobic domain and a short positively charged cytoplasmic tail. The A-domain is believed to be involved with binding to Fg, similar to ClfA and the B-repeats had been shown to interact with Fibronectin (63, 64). The B-repeats of SpsD is known to bind fibronectin with a similar binding mechanism of FnbpA, an *S. aureus* MSCRAMM which also contain B-repeats and bind fibronectin via the extended tandem beta-zipper mechanism. Binding of FnbpA and SpsD to fibronectin is crucial for the endothelial cell invasion of *S. aureus* and *S. pseudintermedius* respectively (65).



Like some previously characterized MSCRAMMs, strain ED99 SpsD and SpsO confer adherence to canine corneocytes when heterologously expressed on the surface of *Lactotoccus lactis* subspecies *cremoris* strain MG1363. Additionally, SpsD and SpsL expressed on the surface of *L. lactis* confer attachment to Fg from multiple species, such as bovine fibronectin, and mouse cytokeratin 10. SpsD and SpsL-specific IgGs have been detected in sera from dogs with bacterial

pyoderma, suggesting that these proteins are expressed during *S. pseudintermedius* infection (62). Taken together, these data suggest an important role for these putative MSCRAMMs in *S. pseudintermedius* pathogenesis and warrant their further study.



### Fibrinogen binding mechanism of Fgy-chain binding MSCRAMMS

*S. aureus* has numerous Fg binding surface proteins such as ClfA, ClfB, FnbpA and FnbpB. The ligand binding region of these MSCRAMMs was narrowed to their A-region, more specifically the N2 and N3 subdomains which is composed of IgG-like domains (Figure-3). The general ligand binding mechanism involves the ligand peptide (of around 12 amino acid residue length) binding at a hydrophobic cleft formed in between the two IgG like domains of N2 and N3 (Figure-4).

The knowledge about the biological/virulence nature of Fg interactions by these staphylococcal surface proteins comes primarily through ClfA-Fg interaction studies and the crystal structure of a ClfA variant with double amino acid substitution ClfA<sub>D327C/K541C</sub> (ClfAcc) in complex with the

synthetic peptide (GV17) with sequence (NH-GEGQQHHLGGAKQAGDV-C00H), mimicking the carboxy-terminal of the  $\gamma$ -chain (392-411). The crystal structure shows that the GV17 peptide interacts with ClfAcc at a trench site in between its N2N3 (220-545) subdomains (Figure-4).

### **Significance of the Study**

The main factor for *S. aureus* associated clumping in blood is attributed to the virulence factor ClfA which has the ability to bind Fg. Fg can undergo polymerization upon converting to the insoluble fibrin and ClfA can bind to both of these forms of Fg. What we know about the interaction between Fg and ClfA is limited. The C-terminal of the Fg $\gamma$ -chain (390-411) binds in between the N2 and N3 subdomains of ClfA and recent biochemical and structural work on an anti-ClfA antibody (aurexis) (3) shows that this antibody could block the interaction of Fg at a different site (top of N3 subdomain) other than the currently known GV17 peptide binding site (trench site)(3, 9). Thus, it is clear that the N3 subdomain has some unknown additional Fg binding site.

With the finding of the novel N3 site in ClfA, suggests that the ClfA/Fg interaction mechanism becomes more complex that could involve multiple sites in both Fg and ClfA. Hence, there is a gap in the literature for answering the critical question, what are the series of binding events in ClfA/Fg interaction that leads to the clumping of *S. aureus* in plasma?

The above question cannot be explained with the currently available information of the ClfAcc/GV17 peptide structure and a thorough investigation on the structural aspects of ClfA-Fg complex is required.

The objective and outcome of my experimental goals will answer to some of the questions on this investigation such as. (1) How many binding sites are there in Fg and ClfA? (2) What are those

binding sites and which are the key residues that involve in the interactions? and (3) How does individual binding events lead to the clumping of Fg?

By answering the above questions, there would be better understanding of this complex interaction to pave the way for better MSCRAMM inhibitors to block Fg binding in future.

### **Preliminary Data and Hypothesis**

ClfA binds to both Fg and fibrin with tight nano molar affinity (Figure-2) (8). Earlier, the binding site in Fg where the ClfA interacts was narrowed down to the C-terminus of the Gamma chain (GV17) of Fg. The GV17 interacts at a trench site between the N2 and N3 subdomain of ClfA. However, ITC binding assays with ClfAcc, concluded that the affinity values of the GV17 peptide is at least 20 fold weaker ( $\mu\text{M}$  range) binding than intact Fg (Preliminary Data, Figure-7c, Table-4). Furthermore, I have observed that this ClfA binding region in Fg [(Fgy(392-411)], might not be available in fibrin (due to proteolysis by thrombin) but ClfA is known to bind to both Fg and fibrin.

In addition, our lab recently published a crystal structure of ClfA in complex with a Fg inhibiting anti-ClfA antibody that gave evidence of a potential Fg binding site located on top of the N3 subdomain in ClfA (Figure-4) (3).

Collectively, based on ITC binding studies with ClfA and previously reported findings on ClfA, I hypothesize that the N3 subdomain of ClfA interacts with the FgD of Fg and this interaction is crucial for ClfA to interact both Fg and Fibrin. I also extend the same hypothesis to other Fg binding surface proteins of staphylococci (SpsD) from *S. pseudintermedius* and, Fbl from *S. lugdunensis*) as well and propose that they also contain an additional Fg binding site in the N3 subdomain similar to ClfA.

## **Innovation**

Fg is an important extracellular protein and major component of plasma. It plays a vital role in the blood coagulation cascade during wound healing. Because of its crucial role in the host many microbes produce virulence factors against Fg. For example, the M1 protein (*Streptococcus*) binds at the coiled coil site of Fg (1). Several other proteins bind to the C-terminus of the  $\gamma$ -chain (392-411) including Clumping factor (ClfA), Fibronectin binding protein (FnbpA), Surface protein of *S. pseudintermedius*-D(SpsD), Fg binding protein (Fbl) of *S. lugdunensis*, Agglutinin-like sequence (als9) protein from the fungus *C. albicans* (9, 63, 66, 67).

Although there are number of crystal structures reported for the microbial proteins in complexes with synthetic Fg peptide ligands, no structures of the complex with a full-length Fg molecule had been reported so far (except the 3.3 Å resolution structure of the M1 protein in complex with a fragment of Fg (PDB ID: 2XNX) (1).

My approach of obtaining the high-resolution crystal structure of the ClfA in complex with Fg is of huge importance because it is going to reveal the multiple facets of Fg which all of the above target virulence factors could bind (Docking studies using experimental or predicted molecular models). The knowledge can even be further extended on the downstream events such as platelet activation by integrin ( $\alpha$ IIb $\beta$ 3) via Fg, inhibition of phagocytosis by targeting integrin ( $\alpha$ M $\beta$ 2).

## **Materials and Methods**

### **Cloning of ClfAcc Construct**

A disulfide bonded ClfA variant ClfA<sub>D327C/K541C</sub> (ClfA<sub>CC</sub>) in a plasmid pQE30 vector expressed from *S. aureus* strain Newman (9) used in this study.

### **Cloning of Fbl Construct**

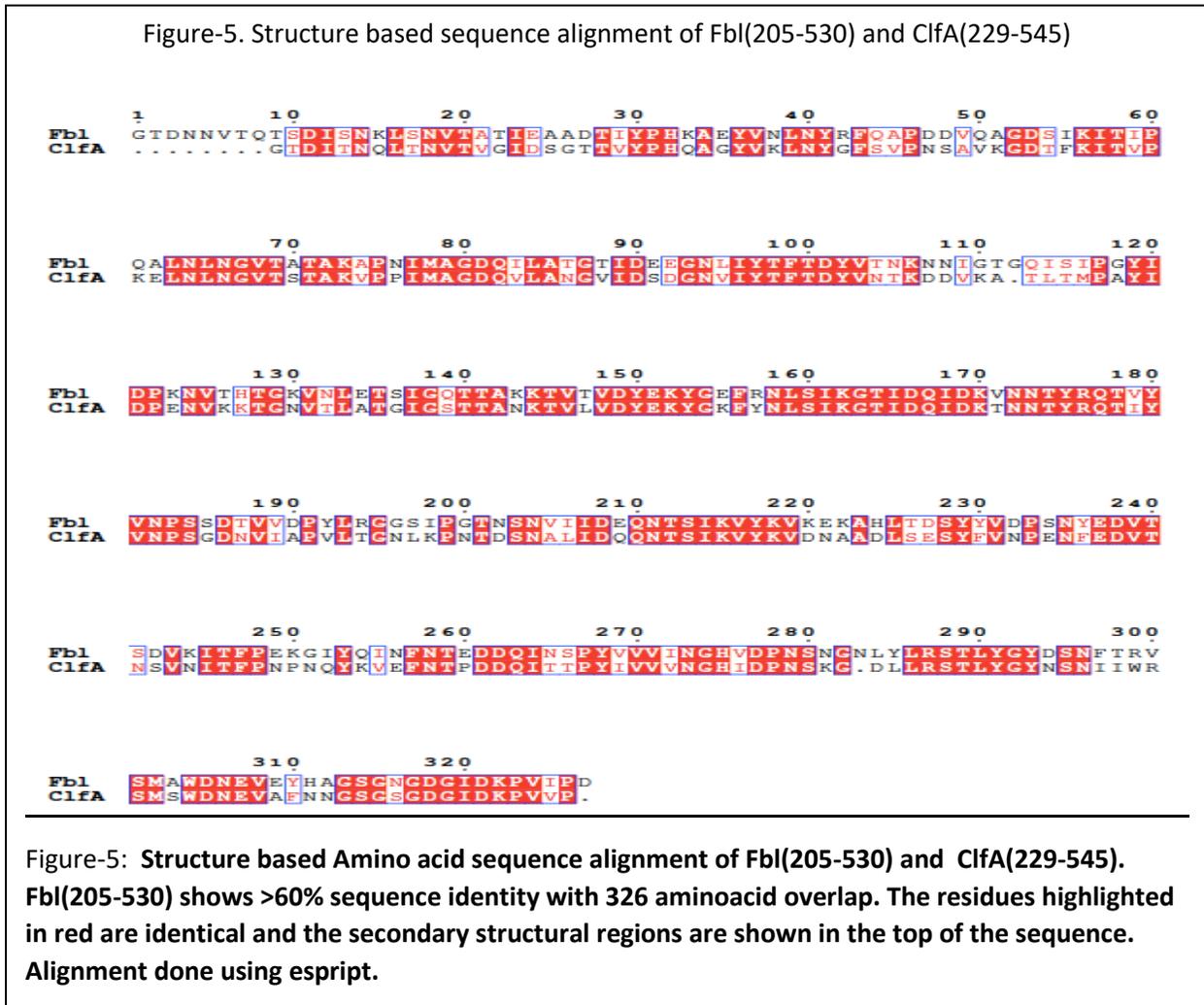
The Fg binding region of Fbl is already identified (60) as the segment corresponding to residues 205-530. This region shows 60% sequence identity with ClfA (Figure-5) and is the closest homolog to ClfA. The corresponding nucleotide regions coding of fbl(205-530) had been cloned into pQE30 vector, expressed in BL21(DE3) cells and purified by Nickel column affinity chromatography followed by Size Exclusion Chromatography.

### **Cloning of SpsD Construct**

The SpsD-N2N3 shows ~40% sequence identity with FnbpA (Figure-6) and ~25% sequence identity with ClfA (not shown). The minimal Fg binding region in SpsD was identified as residues 167-502. Using the full length SpsD (strain ED99) as template, the nucleotide coding regions for (167-502) of SpsD was amplified by PCR. The amplified regions using appropriate forward and reverse primers (Table-3) were ligated in pQE30 plasmid with a BamH1 and Pst1 at 5' and 3' ends and the plasmid transformed in DH5- $\alpha$  cells. The transformed DH5- $\alpha$  cells were positively selected by antibiotic selection (ampicillin resistance imparted by pQE plasmid). The plasmids were extracted from these cells and reconfirmed by colony PCR and sequences were verified by Sanger sequencing (GENEWIZ). These plasmids were retransformed into E-coli cells BL21(DE3) for the efficient expression of recombinant SpsD(167-502).

## Site-Directed Mutagenesis

Amino acid substitutions in the FgD binding site on top of N3 subdomain of ClfAcc, Fbl and SpsD were listed (Table-7, 8). A group of 7-8 amino acid substitutions were made for each MSCRAMM as a single mutant clone and were cloned separately with an N-terminal hexahistidine tag at the multiple cloning site (MCS) of the pQE30 expression vector purchased from GENEWIZ.



Primers used in study	Nucleotide Sequence	Restriction Enzyme
Forward primer	5-GCA <sup>ggatcc</sup> AATGTAGAAAACGGAACAGATGTGACA -3	BamH1
Reverse primer	5'-GCA <sup>ctgag</sup> TTAAACTGGTTTATCGTTCCTTC-3'	Pst1

**Table-3:** Nucleotide sequence of forward and reverse primers used for the amplification and ligation of the insert SpsD (167-502) into pQE plasmid vector. The restriction site BamH1 and Pst1 are colored in red.

**Figure-6. Structure based Amino acid sequence alignment of FnbpA(189-505) and SpsD(167-502).**



**Figure-6: Structure based Amino acid sequence alignment of FnbpA(189-505) and SpsD(167-502). SpsD (167-502) shows 39.6% sequence identity of 326 aminoacid overlap. The residues highlighted in red are identical and the secondary structural regions are shown in the top of the sequence. Alignment done using esprict.**

### **Crystallization of ClfAcc/Fg complex**

I subsequently tried crystallizing the complex using crystallizing conditions previously published for ClfA (9) and Fg (28) apo structures. These crystallization trials resulted in crystals (Figure-8a) but when the crystals were analyzed by SDS-PAGE, there is either ClfA or Fg alone but I could not detect any protein complex whatsoever. Subsequently, these crystals were subjected to X-ray diffraction at the high beam radiation facility (Synchrotron, APS-Chicago) and got low diffracting data ( $\sim 7\text{\AA}$  resolution). The diffraction data for the Fg crystals showed high anisotropy with P 1 21 1 symmetry, which is a hallmark for Fg and also contained similar unit cell dimensions (a $\sim 135\text{\AA}$ ; b $\sim 94\text{\AA}$ ; c $\sim 300\text{\AA}$ ) of previously published crystal structures of Fg (28).

### **Generation and Purification of FgD**

Human Fg-Pk1 (30-39ml, S200 DEAE Sepharose, GE Life Sciences) was digested with Hplasin (Enzyme Research) with 20mM CaCl<sub>2</sub> for 3hours @37C. Then the digest was concentrated to 4ml and the proteolytic inhibitors TPCK and TLCK added to a final concentration of 10 $\mu$ M. Then the digest was subjected to gel permeation chromatography (S200 DEAE Sepharose) and elute was collected as individual fractions (3ml fractions). The fractions were run in series in SDS-PAGE in non-reducing conditions and the fragment-D ran  $\sim 86\text{kDa}$  as expected (FgD molecular weight is 86kDa). The final purified preparation of FgD, which appeared homogeneous on SDS gels, was treated with benzamidine (final concentration of 50  $\mu$ M) – being flash-frozen in 0.5 mL aliquots and stored at  $-78^{\circ}\text{C}$ .

### **Purification of ClfAcc/FgD protein complex**

A molar ratio of 4 ClfA:1FgD was incubated in 1X TBS, pH7.4 for 2-3 hours in ice. After incubation, the sample solution was subjected to Ni<sup>2+</sup>- affinity chromatography. The unbound FgD was washed down with wash buffer (5mM Imidazole, 50mM NaCl, 50mM Tris (pH8.0)). The affinity column bound proteins eluted in elution buffer (500mM Imidazole, 50mM NaCl, 50mM Tris pH8.0) using gradient fractionation technique (5-100% gradient in 50ml elute volume). The individual fractions were subjected to SDS-PAGE in NR conditions and the fractions which showed two bands (86kDa and 36kDa) indicating the successful ClfAcc/FgD complex were pooled and subjected to crystallization.

### **Crystallization of ClfAcc/FgD protein complex**

Crystals were grown at room temperature in sitting drops made from: – equal volumes of a 9 mg/mL protein solution in 0.15M NaCl, 0.05M Tris (pH 7.0), and 1 mM CaCl<sub>2</sub>, containing 2 mM GPRPam and 2 mM GHRPam, – and well solutions composed of 0.05 M Tris (pH 7.0) containing 4% PEG-3350, 0.25-0.5 M TMAO, 1.0 mM CaCl<sub>2</sub>, and 2 mM sodium azide.

Typically, a precipitate formed during the first 24 h; crystals appeared in 1 week and continued to grow over~ the course of 1-2 months. At this point, the cryoprotectant MPD (methylpentanediol) was added gradually to a final concentration of 15%, after which the crystals were frozen and stored in liquid nitrogen. X-ray diffraction data of the complex with higher-resolution data were collected at the Advanced Light Source, Chicago. The X-ray diffraction Data were processed with HKL2000(68).

The structure was determined by molecular replacement using 'PHASER-MR'. Model reconstructions and refinement performed by RefMac (69) (real time refinement) and COOT (70) (manual refinement).

Molecular replacement (MR) is a method of solving the phase problem (loss in information) in X-ray crystallography. It relies upon the existence of a previously solved protein structure which is homologous to our unknown structure from which the diffraction data is derived. Molecular replacement was conducted with the equivalent of a single fragment D from human Fg (PDB: 1FZA) (33) and ClfAcc (PDB:2VR3)(9). The firmly packed globular regions of FgD are very well resolved and the N2N3 subdomains of ClfAcc are completely resolved were some of the loop regions in the N3 subdomain in ClfAcc which was not resolved in earlier structure (PDB ID:2VR3) including the C-terminal extension region (532-545) of ClfAcc (latch region)(9).

The crystal packing exhibited a variety of unique intermolecular contacts, (a) a relatively open-face of  $\gamma$  chains that interacts with the D-D interface in the fibrin assembly interface chiefly interacting with the N3 subdomain of ClfAcc. (b) a novel antiparallel association of coiled coil region of FgD with the N2 subdomain of ClfAcc from neighboring molecules and (c) a lattice contact interface between  $\beta$ C domains of FgD with N2 subdomain of ClfAcc.

### **Cloning of the Fg $\gamma$ module construct**

Firstly the cDNA of Fg  $\gamma$ 143-411 module (from Homo sapiens) was PCR-amplified from a lab stored construct pQE30/Fg  $\gamma$ 27-411, and was then sub-cloned into a pET15b-based vector (Agilent Technologies) which was linearized with BamHI and XhoI by using In-Fusion HD Cloning kit (Takara Bio), thus producing a construct which can express 6HisMBP-Fg  $\gamma$  module under the driven of T7lac promoter. The integrity of the resulting plasmid was confirmed by DNA

sequencing. The expressed 6HisMBP-Fg  $\gamma$  module contains a TEV protease cleavage site between the 6HisMBP tag and Fg  $\gamma$  module protein, thus facilitating the 6HisMBP tag removal by using TEV protease digestion. Point mutations were introduced using the site-directed mutagenesis QuikChange kit (Agilent Technologies) with pET15b/6HisMBP-Fg  $\gamma$  as template, to generate the quadruple mutant, Fg $\gamma$ -Q239A/S240A/I242A/P243A. A StrepII-tag sequence (WSHPQFEK) was introduced at the C-terminus of Fg  $\gamma$  module to generate pET15b/6HisMBP-Fg  $\gamma$ -StrepII construct. The *E. coli* strain Rosetta-gami B(DE3) (Novagen) was transformed with the pET15b/6HisMBP-Fg  $\gamma$  module expression plasmid (or pET15b/6HisMBP-Fg  $\gamma$ -StrepII, or pET15b/6HisMBP-Fg $\gamma$ -Q239A/S240A/I242A/P243A) and grown overnight in the Terrific Broth containing ampicillin (150  $\mu$ g/mL) and chloramphenicol (34  $\mu$ g/mL) at 37 °C. The overnight culture was used to inoculate Terrific Broth containing ampicillin (150  $\mu$ g/mL) and incubated while being shaken at 37 °C until the A610nm reached 0.8. After the culture was chilled to 4 °C, 0.10 mM (final concentration) of isopropyl 1-thio- $\beta$ -D-galactopyranoside was added, and the culture was shaken for 36 hours at 18 °C. Cells were harvested by centrifugation at 5,000 $\times$ g for 15 min, 4 °C.

### **Expression and Purification of Fg $\gamma$ module**

To purify Fg  $\gamma$  module, cells were re-suspended in 50 mM Tris/HCl (pH 8.0) containing 250 mM NaCl, 20 mM imidazole, and 10% glycerol. The homogeneous suspension was lysed with three passes through a Microfluidics M-110P microfluidizer at 20,000 psi, and then centrifuged for 20 min at 100,000 $\times$ g and 4 °C. The supernatant containing recombinant 6HisMBP-Fg  $\gamma$  module was then applied onto a HisTrap HP/5-mL column (GE Healthcare) following the instructions the manufacture supplied, and after a thorough wash with 50xCV of wash buffer (50 mM Tris/HCl, 250 mM NaCl, 45 mM imidazole, 10% glycerol, pH 7.4) the target protein was then eluted with a

buffer containing 50 mM Tris/HCl (pH 7.4), 250 mM NaCl, and 500 mM imidazole, which (i.e. imidazole) was then removed by using a HiPrep 26/10 desalting column (GE Healthcare) pre-equilibrated with 50 mM Tris/HCl (pH 7.4), 250 mM NaCl. This process was conducted on the AKTApure (GE Healthcare) system. The purified 6HisMBP-Fg  $\gamma$  module was then loaded onto a MBPTrap HP/5-mL column (GE Healthcare) following the instructions the manufacture supplied, and after a thorough wash with 50xCV of wash buffer (50 mM Tris/HCl, 250 mM NaCl, pH 7.4) the target protein was then eluted with TBS buffer containing 15 mM maltose, pH 7.4. Half of the eluates was combined, concentrated and loaded onto a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) which was pre-equilibrated with TBS buffer for gel filtration on AKTApure system (GE Healthcare). The fractions containing 6HisMBP-Fg  $\gamma$  module were pooled and concentrated by using Amicon Ultra-15 Centrifugal Filters with 30,000 NMWL (Merck Millipore). The other half of the eluates were digested with 6His-Tagged TEV protease at a ratio of 6.67  $\mu$ g of TEV protease per mg of 6HisMBP-Fg  $\gamma$  module with 3 mM reduced L-glutathione and 0.3 mM oxidized glutathione for overnight in the cold room, and was then loaded onto a HisTrap HP/5-mL column (GE Healthcare); the flow-through (tag-free Fg  $\gamma$  module) was harvested and concentrated, and then was loaded onto a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) which was pre-equilibrated with TBS buffer for gel filtration on AKTApure system (GE Healthcare). The fractions containing tag-free Fg  $\gamma$  module were pooled and concentrated by using Amicon Ultra-15 Centrifugal Filters with 10,000 NMWL (Merck Millipore). (Note: TEV protease digestion/treatment usually will cause Fg  $\gamma$  module aggregation, and thus results in a very low recovery of tag-free target protein.)

The purities of the purified 6HisMBP- Fg  $\gamma$  module were checked via 4-20% sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue staining. The

protein concentration was determined by DC Protein Assay with a Bio-Rad (Hercules, CA) kit, using bovine serum albumin as the standard.

## **Results of Chapter-II**

### **Interaction between ClfAcc and Fibrinogen**

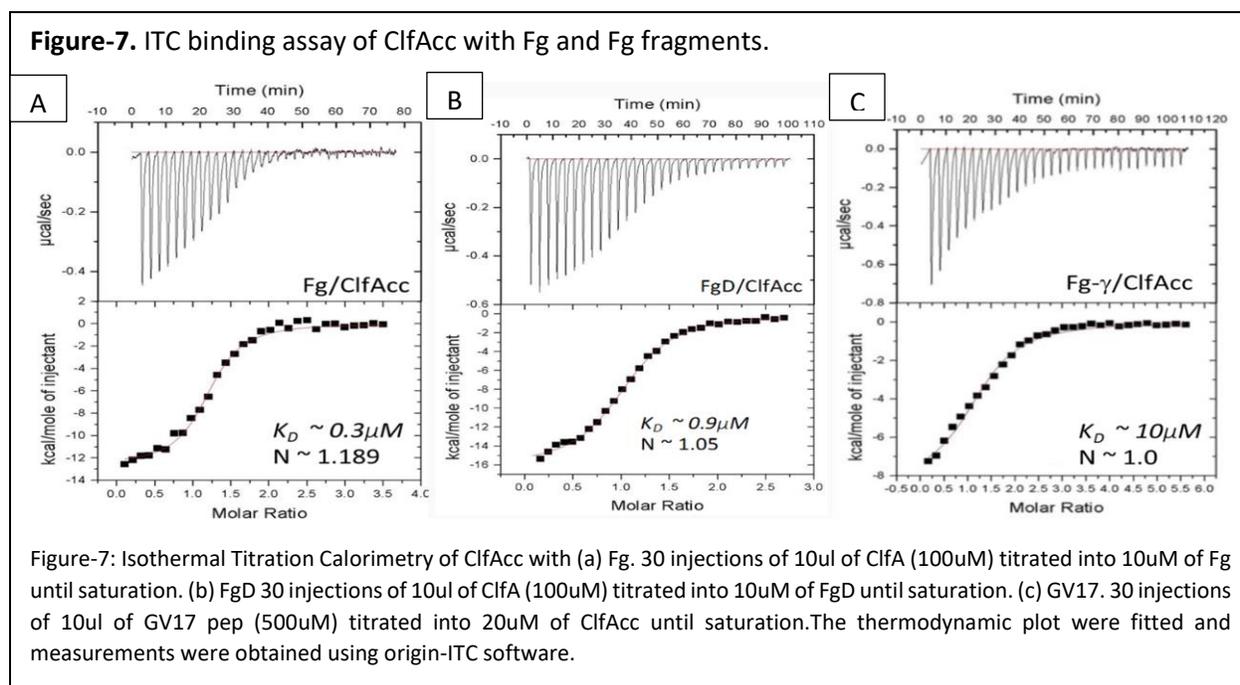
Earlier studies show that ClfA interact with Fg and that a linear segment of the C-terminal of the Fg  $\gamma$ -chain docks in a trench between the N2 and N3 domain of the MSCRAMM (9). Recent biochemical and structural studies of monoclonal antibody 'Aurexis' in complex with ClfA suggest that Fg also interact with some residues "on top of" the N3 domain (3). Taken together from these observations I hypothesize that residues outside the C-terminus of Fg  $\gamma$ -chain(392-411) in Fg interact with ClfA. To test this hypothesis I will conduct a series of isothermal calorimetry (ITC) studies using the previously published high affinity Fg binding disulfide variant form of ClfA [ClfAcc (ClfA<sub>D327C/K542C</sub>)] (3).

The three different forms of Fg were tested for evidence of binding. (1) An Intact soluble Fg dimer, (2) a synthetic peptide (GV17) that corresponds to the 17 C-terminal residues of the  $\gamma$ -chain (392-411), and (3) the FgD, which is a ~86kDa proteolytic fragment of Fg (See Methods section for the Preparation of FgD).

The FgD is a mimick of fibrin and the FgD doesn't have the region (392-396) of Fg $\gamma$  due to the proteolytic cleavage at position 388-392 of Fg $\gamma$  by plasmin.

From the ITC studies, the affinity values ( $K_a$  and  $K_d$ ), stoichiometry( $n$ ) and thermodynamic parameters ( $\Delta G, \Delta H, -T\Delta S$ ) of ClfAcc are obtained and listed in Table-2. The affinity  $K_d$  of ClfAcc with the peptide (GV17) is 9 $\mu$ M which is >20 fold weaker when compared to the affinity of ClfAcc with the intact Fg (0.3 $\mu$ M). The affinity of ClfAcc with the FgD is 1 $\mu$ M which is only ~3 fold weaker than the intact Fg but ~10 fold stronger than the ClfAcc/GV17 peptide affinity. ClfAcc

binding stoichiometry ( $n$ ) to all of the above mentioned ligands were around  $N=1.0$  (Figure-7 and Table-4).



The total gibbs free energy ( $\Delta G$ ) is calculated from the enthalpy and entropy change during the reaction ( $\Delta G = \Delta H - T\Delta S$ ). The  $\Delta G$  values of the ClfAcc binding to the peptide, Fg and FgD ( $-6.62 \text{ kcal mol}^{-1}$ ,  $-8.56 \text{ kcal mol}^{-1}$ , and  $-7.87 \text{ kcal mol}^{-1}$ ) respectively. While, there is not a significant change in the ( $\Delta G$ ) value among these ligands binding to ClfAcc, there was significant differences in both  $\Delta H$  and  $-T\Delta S$  values (Table-4).

The peptide binding of ClfAcc has an enthalpy ( $\Delta H$ ) of  $-7 \text{ kcal mol}^{-1}$  which is  $\sim 2$  fold decreased enthalpy ( $\Delta H$ ) value of intact Fg binding to ClfAcc ( $-14 \text{ kcal mol}^{-1}$ ). Also, the peptide is missing a positive entropy ( $-T\Delta S$ ) of  $\sim +4 \text{ kcal mol}^{-1}$  in comparison to intact Fg binding to ClfAcc (Table-4). Collectively, these differences point out that the peptide binding of ClfAcc is either (i) has a different binding mechanism as compared to intact Fg binding of ClfAcc or (ii) is missing a binding

site/conformational changes that contributes to the differences in ( $\Delta H$ ) and ( $\Delta S$ ) components when compared to ClfAcc/Fg interactions.

**Table-4. Summary of thermodynamic parameters of ClfAcc from the ITC binding studies**

	ClfAcc	n	KD ( $\mu M$ )	$\Delta G$ (kcal mol <sup>-1</sup> )	$\Delta H$ (kcal mol <sup>-1</sup> )	-T $\Delta S$ (kcal mol <sup>-1</sup> )
Peptide/0.5mM	20	1.04±0.26	<b>9.0</b>	-6.62	<b>-7.0±0.05</b>	<b>0.37</b>
Fg/0.01mM	100	1.18±0.02	0.3	-8.56	<b>-13.2±0.03</b>	<b>4.63</b>
FgD/0.01mM	100	1.05±0.12	1.0	-7.87	<b>-14.9±0.03</b>	<b>7.02</b>
MBP-Fgy	100	0.99±0.10	0.5	-8.27	<b>-10.4±0.04</b>	<b>2.13</b>

Primarily, enthalpic values are indicators for non-covalent interactions mainly hydrogen bond interactions and the entropy values represent hydrophobic interactions. From the crystal structure of ClfAcc with the GV17 peptide, we know that the peptide occupies a hydrophobic trench between the N2 and N3 subdomains of ClfAcc. Thus, the thermodynamic component observations of the peptide supports the findings of the ClfAcc-GV17 peptide structure.

In contrast, the FgD of Fg which is missing the last 17 residues (392-411) of the C-terminus of Fgy due to the protease labile site at positions between (388-382) bound to the ClfAcc with high affinity indicating that the FgD could be the binding region in Fg that interacts with the N3 site in ClfAcc. Furthermore, the FgD/ClfAcc interaction points out an enthalpy of 14.9 kcal mol<sup>-1</sup> and a positive entropy (-T $\Delta S$ ) of around 7 kcal mol<sup>-1</sup> which might be due to the result of those N3-site interactions in ClfAcc. The higher entropic values of FgD indicate that ClfAcc interacts with a hydrophobic surface located in the FgD. These findings support the presence of an additional Fg binding site at the top of the N3 subdomain in ClfA (from the ClfA-tefibazumab study) and support my hypothesis.

### **Crystallization trials of ClfAcc/Fibrinogen protein complex**

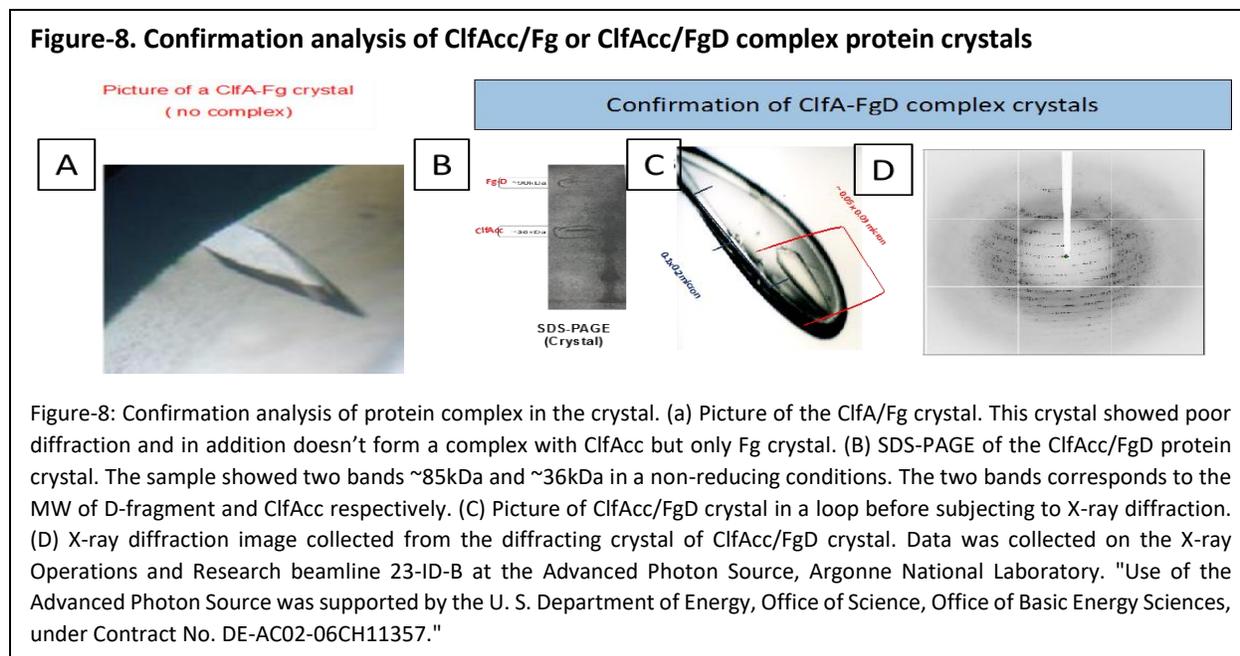
To further characterize the ClfA/Fg interaction I first tried to crystallize ClfA in complex with full length Fg. From the ITC studies (Figure-13) we know the binding molar ratio to be 2:1 (ClfA:Fg). Hence the molar ratio of 2:1 was used in my crystallization studies of a ClfA:Fg protein complex. I tried various crystallizing conditions using commercial crystal screen kits (Hampton Research), Wizard and Index Screens (Emerald Bioscience). None of the crystallizing conditions from screens yielded crystals of the ClfAcc/Fg complex.

Crystallizing the complex of Fg/ClfAcc has its inherent difficulties owing to several reasons, (a) the intrinsic floppiness of Fg has proved to be a formidable barrier to crystallization. (b) Flexible high-repeat regions that universally found in the  $\alpha$ -chains of mammalian Fg's. However, these difficulties can be overcome by adopting alternative strategy of crystallizing core fragments of proteolyzed Fg as numerous crystal structures have been reported for core fragments of proteolyzed Fg.

### **Crystallization trials of ClfAcc/FgD protein complex**

To increase the chances of obtaining useful crystal for X-ray diffraction of ClfA in complex with a segment of Fg I decided to use fragment D (Fg-D) which can be generated by plasmin digestion of intact Fg. Using FgD is advantageous over Fg because the FgD has a more defined and stable structure compared to full length Fg. The Fg-D was generated and purified as described in the Material and Method section. I have used a recombinant disulfide bonded variant of ClfA (ClfAcc) where two Cys residues (D327C/K541C) have been substituted. The presence of two Cys at these positions forces the N2 subdomain (D327C) disulfide bond to the residue K541C at the C-terminal extension of the N3 subdomain which is the latching region to pre-form the ligand binding trench

site. Purified monomeric ClfAcc and monomeric Fg-D were used in these experiments. First I incubated the two proteins with a ratio of two molecules of ClfAcc for every FgD segment to form a complex. Taking advantage of the presence of the N-terminal Histidine tag in ClfAcc, I have isolated the ClfAcc-FgD complex from the unbound ClfAcc by passing the complex through the nickel-sepharose chelating affinity chromatography column. The complex is subsequently further purified by Size Exclusion Chromatography using a S75(16/60) sepharose gel permeation column. The purified complex was subjected to crystallization experiments using the above-mentioned commercial screens. Crystal formation was observed under several conditions differing in morphology from quazi crystals to needles and small micro crystals. The conditions were reiterated, optimized and several additive reagents were tested such as chaotrophs and detergents to generate high quality crystals that diffracted up to 2.4Å resolution (Figure-8b, 8c and 8d).



## Data Determination

We processed the diffraction data using HKL2000 (Indexing, Integrating and Scaling) and using the Mathews coefficient, most likely there must be one molecule in the asymmetric unit. Furthermore the unit cell dimensions ( $a = \sim 180$ ,  $b = \sim 180$ ,  $c = \sim 113$ ) and space group (P 4<sub>2</sub> 2<sub>1</sub> 2) is different from the previously published (ref) crystal structure of apo FgD ( $a = \sim 108$ ,  $b = \sim 48$ ,  $c = \sim 167$ ) and space group (P 1 2<sub>1</sub> 1). This indicated conclusively that the crystal is indeed a protein complex. We used the previously published structures of ClfA (chain-A, PDB ID: 2VR3) and FgD (33) (chain-A, PDB ID: 1FZA) as ensemble to find the solution using PHASER by automated Molecular Replacement method by CCP4 and PhaserMR. The resulting solution had high LLG values indicating a definitive solution of the structure of the protein complex. The initial solution was further refined manually using COOT and PyMoL with several rounds of refinement cycles. The model coordinates of the protein complex were manually inspected using COOT and corrected for isotropic temperature factors of each atom, fitted side chains into the electron density and deleted side chains or residues that doesn't fit into the weighted electron density ( $2F_o - F_c$ ) map as well as manual addition of solvent/water molecules. The final structure has R-factor and R-free of 18.07% and 20.85% respectively (Table -5). The final coordinates are validated and are ready for submission in PDB.

## Structural Analysis of ClfAcc/FgD Complex

The crystal structure showed that the FgD is in fact bound to the ClfAcc molecule at the top of its N3 domain as anticipated based on previous studies (3).

The Fg  $\gamma$ -module, a 30 kDa globular domain of the FgD of Fg interacts at the top of the N3 and the  $\alpha$ - and the  $\beta$ -chains of Fg were not involved in the interactions (Figure-9A). The interaction

buried a total of 751 Å<sup>2</sup> (with the contributions from ClfAcc: 382 Å<sup>2</sup> and γ-module: 369 Å<sup>2</sup>) (Figure-9). The crystal structure revealed key residues important for the interaction in both Fg and the ClfA molecule (Figure-10).

The Fg γ-module (144-392) interacts with the N3 subdomain (380-530) of ClfAcc. Mainly, the ‘QSAIP’ motif comprising region (239-243) and a loop region connecting helix(C) (F389 and N390) of the γ-module constitutes a major continuous stretch of residues that make primary contact with the N3 subdomain of ClfA (Figure-10). The N3 subdomain residues in ClfAcc that interact with the Fgγ module (144-392) in the FgD is not located as a continuous stretch like the Fgγ, but extends across three loop regions that form the N3 subdomain of ClfAcc (Figure-10B). Mainly, the residues (406-411) that form the C-D loop, residues (477-487) that form the D’-E loop and the residues (511-521) that form the F-G loop participate in the interaction with Fgγ(144-392) in the FgD. The QSAIP motif docks in a shallow groove in ClfA formed by residues I408, D480, D481, Q482 and Y512. The dipeptide motif, F389 and N390 of the γ-module interacts with the A409, Y512 and I516 of ClfA-N3 subdomain of ClfAcc (Figure-10). In addition, residues K151, D155, N158, L172, and M264 of the γ-module minor contacts with V411, N463, E475, T478 and W518 of the ClfA-N3 subdomain (not shown).

**Figure-9. Crystal Structure of ClfAcc/FgD complex**

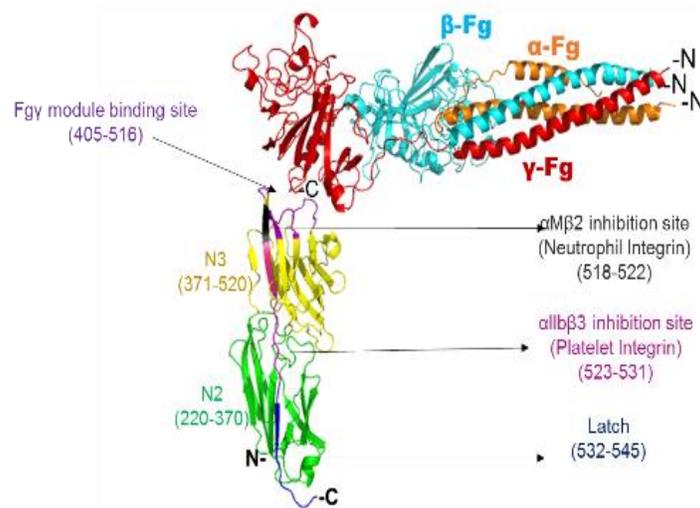
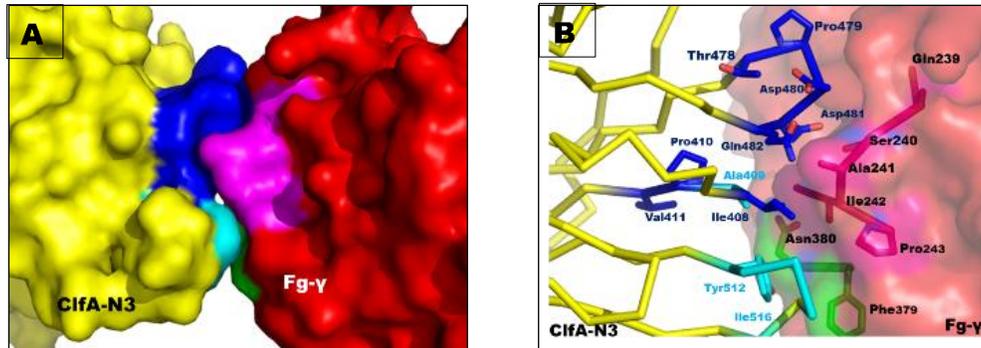


Figure-9: Crystal Structure of ClfAcc/FgD complex. The N2 and N3 subdomain of ClfAcc is shown as cartoon and colored in green and yellow respectively. The FgD is shown as cartoon and the three polypeptide chains of FgD is colored in orange( $\alpha$ -chain), cyan( $\beta$ -chain)and red( $\gamma$ -chain). The binding/inhibiting sites of ClfAcc towards various ligands of fibrinogen are shown in arrows. Region (532-545) of ClfAcc is the latch region that undergoes a conformational change upon ligand binding and reorients to the N2 subdomain of ClfAcc to form a  $\beta$ -sheet complementation with the G-strand of N2 subdomain of ClfAcc. Region (523-531) is the platelet integrin ( $\alpha$ IIb $\beta$ 3) inhibition site where the C-termini of Fgy (392-411) interacts with and region (518-522) is the neutrophil integrin ( $\alpha$ M $\beta$ 2) inhibition site by steric hindrance due to the binding of the Fg- $\gamma$  module binding of D-fragment. Region (405-516) of ClfAcc interacts with the Fgy(144-392) of D-fragment.

Crystal structure of ClfA/FgD complex reveals that FgD lacks the regions (393-411) that previous work has shown interact with ClfA at the N2N3 trench (9). To get a better understanding on the interaction of Fg with ClfA involving the two identified interaction sites, an overall structural model was reconstructed (Figure-11) by using the crystal structure of ClfA in complex with the GV17 peptide (PDB ID:2VR3), using previously reported data from our laboratory and the current ClfA:FgD structure. This gives an overall binding site information of Fg (Figure-11a, 11b).

Figure-10. Analysis of ClfAcc-N3/Fgy Binding interface



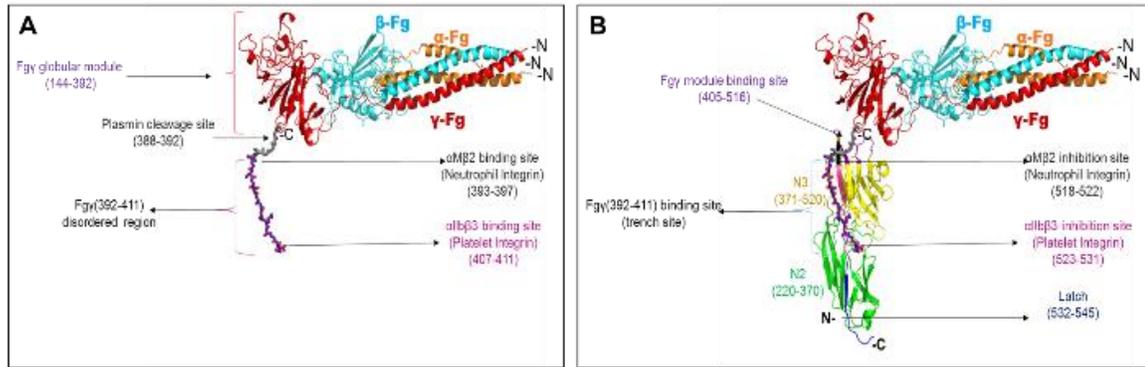
**Figure-10. Analysis of ClfAcc-N3/Fgy Binding interface. (A)** Closeup view showing the ClfA-N3/Fgy binding interface. The subdomain of ClfA-N3 (Yellow) and Fgy(Red) are only shown for clarity as surface. The binding interface QSAIP motif in Fgy colored in pink and this binding interface in ClfA colored in blue. The dipeptide helix(C) of Fgy is colored in green and this binding region in ClfA is colored in cyan. **(B)** Amino acid residues of the binding interface is shown in sticks and colored same as seen in Fig-7a. **Note:** The interacting residues in ClfAcc spread across three loop regions of ClfAcc-N3 subdomain, whereas the Fgy interacting residues is continuous stretch of (239-243) and (389-390).

### Cloning of Fgy module construct

To experimentally study a Fg fragment capable of interacting with ClfA at both identified interacting sites, we decided to seek to express the C-terminal segment of the  $\gamma$ -chain comprising residues 144-411. This segment called the Fgy(144-411) contains all the ClfA interacting residues that we can identified from the two structures. This construct is hypothesized to interact at both the N2N3 trench and on top on N3 subdomain of ClfA.

The Fg- $\gamma$  chain in Fg, especially the globular region (144-392) interacts with the  $\beta$ -chain of Fg and it is not clear to what extent a recombinant version of Fg- $\gamma$  (144-411) protein on its own is stable. Expression and purification of this domain might be challenge for solubility, yield and activity. We therefore explored two approaches. The objective of both of this strategy is to get a Fg- $\gamma$ (144-411) that can be expressed, purified as a stable protein without losing the ClfA binding activity.

**Figure-11. Comparison of FgD (Apo) and ClfAcc bound FgD complex (Composite Model).**

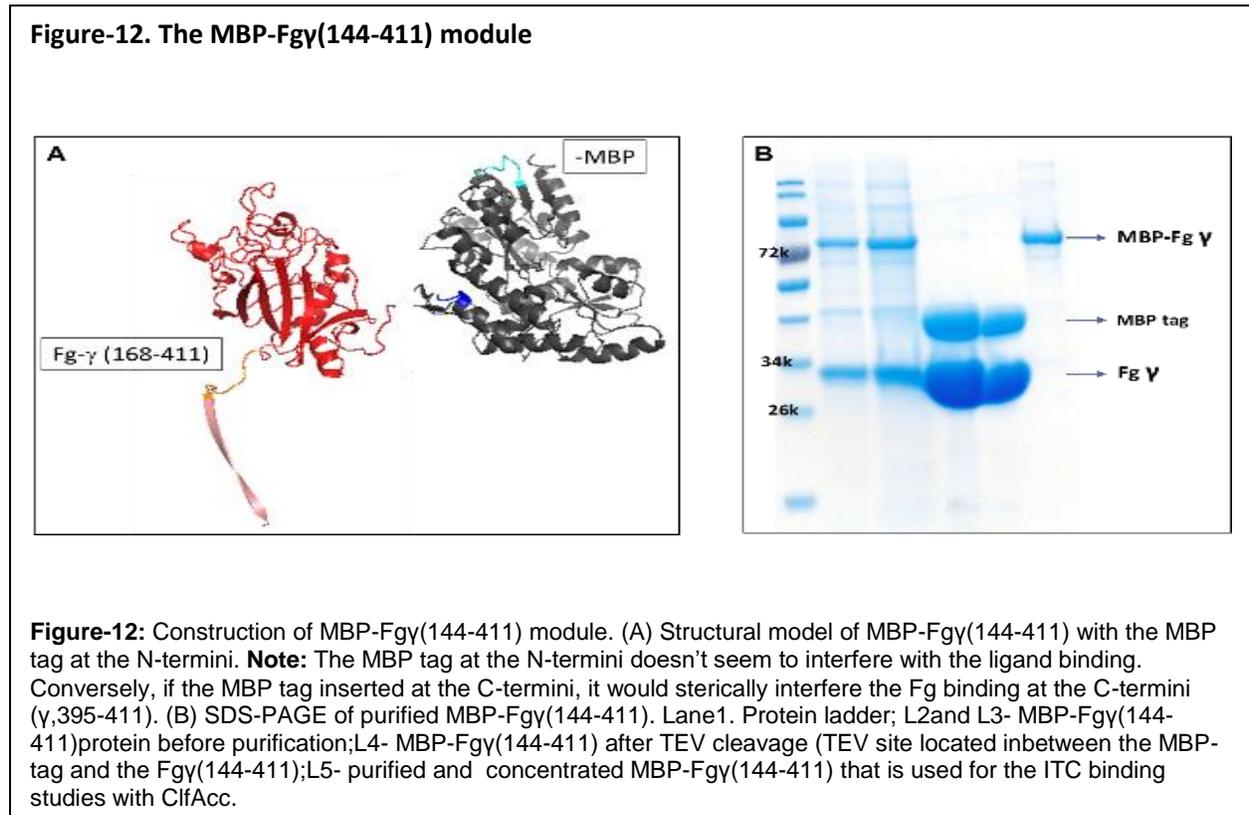


**Figure-11:** Comparison of FgD (Apo) and ClfAcc bound FgD complex (Composite Model). (A) Structural model of Apo-FgD. A structural model built for FgD connecting the globular region of Fgy(144-392) with the C-terminal disordered region of Fgy(397-411). The gap region comprising Fgy(393-397) is manually built by tracing using PyMol. The sites in FgD that are targeted by ClfAcc is color coded, label and arrowed. (B) Structural model of ClfAcc bound to both the ordered (globular) and disordered (C-termini) of the Fgy (D-fragment). Crystal structure of FgD/ClfAcc (current work) and structure of ClfAcc in complex with the Fgy(395-411) (PDB ID:2VR3) was used to build the model. The regions in ClfA that bind to the Fg that are of biological significance are highlighted (color coded, labeled and arrowed). ClfA N2 and N3 subdomains are shown in green and yellow respectively, while the latch region is colored blue. In FgD the three polypeptide chains are colored as orange( $\alpha$ -chain), cyan( $\beta$ -chain) and red( $\gamma$ -chain).

In the first approach we decided to make the following substitutions Ile144Thr, Phe167Ser, Leu178Asn, Ile198Thr, and Leu207Asn. These are mainly hydrophobic to hydrophilic changes of residues in the G-chain at the  $\beta$ -chain interface. We also decided to include Asp410-Ile which we expect will enhance the affinity at the peptide binding site. This mutated version of the His-Fgy-module was constructed with a hexa-histidine tag at the N-terminus to facilitate purification by Ni<sup>+</sup> chelating chromatography.

In the second approach we used a different fusion partner. Fusion partners like the Maltose Binding Protein (MBP) can serve as a chaperone and help the recombinant protein remain soluble. However, the assisted solubility of the recombinant protein could represent a non-native form and the size of the MBP (~25kDa) might interfere with the analysis and activity of the recombinant protein (Figure-12a). With these reservations we decided to express the wild type Fg- $\gamma$ (144-411)

with the MBP tag at the N-terminus. The objective of both of this strategy is to get a Fg- $\gamma$ (144–411) that can be expressed, purified without losing the ClfA binding activity.



**Table-5: Data Collection and Refinement Statistics for ClfA:FgD**

<b>Data Collection</b>	
Space group	P 42 21 2
Cell Dimensions	
a, b, c (Å)	180.6 180.6 113.69
$\alpha, \beta, \gamma$ (°)	90 90 90
Resolution(Å)	2.36
R <sub>merge</sub> <sup>*, †</sup>	0.055
Mean I/sigma(I)	1.33
Completeness (%) <sup>*</sup>	99.87
Redundancy <sup>*</sup>	2.8
<b>Refinement</b>	
No. of Unique reflections	77478
R <sub>work</sub> /R <sub>free</sub> (%)‡	18.07/20.85
No. of non- hydrogen atoms	
Protein residues	1340
Water	
Average B-factor (Å <sup>2</sup> )	59.98
Root mean square deviations	
Bond lengths (Å)	0.016
Bond angles (°)	1.92
<p>*Values in parentheses are for the last shell. †R<sub>merge</sub> = <math>\frac{\sum hkl \sum_i  I_i(hkl) - \langle I \rangle }{\sum hkl \sum_i I_i(hkl)}</math>, where I is the observed intensity, and <math>\langle I \rangle</math> is the mean value of I. ‡R<sub>work</sub>/R<sub>free</sub> = <math>\frac{\sum hkl   F_{obs}  -  F_{calc}  }{\sum hkl  F_{obs} }</math>, where R and R<sub>free</sub> are calculated using the test reflections, respectively. The test reflections (5%) were held aside and not used during the entire refinement process.</p>	

The results of the expression and purification of the two recombinant gamma modular version showed that the His-GM version did get expressed but unfortunately was not stable and precipitates out of solution at higher concentrations (Data not shown). Hence it cannot be used for the interactive studies with ClfA. Fortunately the MBP-Fgy (144-411) expressed well and remained soluble in solution during the purification process. This form could therefore be used for binding studies (Figure-12b).

### **Interaction of ClfAcc with Fgy Module**

The binding affinity of recombinant MBP-Fgy(144-411) to ClfAcc was determined by ITC technique. Briefly, a fixed concentration (10 $\mu$ M) of the soluble MBP-Fgy (144-411) was kept in the sample cell and soluble monomeric ClfAcc(100 $\mu$ M) is placed in the syringe in a matching buffer (TBS, pH7.4). Thirty injections of 10  $\mu$ l of ClfAcc was injected into the cell and the peaks of the heat released during individual titrations were recorded. The peak signal decreases as the FgD is saturated by ClfAcc binding. The resulting isotherm was presented in a wiseman plot and one site model fitted to yield the binding affinity ( $K_d$ ), stoichiometry(n) and thermodynamic values ( $\Delta H$  and  $\Delta S$ ) of the reaction (Figure-13). As expected, the MBP-Fgy (144-411) displayed an binding affinity ( $K_D$ ) of 0.5  $\mu$ M which is twofold higher than the binding affinity ( $K_D$ ) of the FgD/ClfAcc binding values (1.0 $\mu$ M). The thermodynamic values of MBP-Fgy/ClfAcc binding were similar to the thermodynamic values of intact Fg/ClfAcc binding (Table-4). This indicates that MBP-Fgy (144-411) binds to the two binding sites (low affinity trench site and high affinity N3 site) of ClfAcc. These results validate the findings and support the composite model (Figure-11) derived from the crystal structure of ClfA:FgD (current work) and ClfA:GV17 (PDB ID: 2VR3).

**Figure-13. ITC binding assay of ClfAcc with MBP-Fgy(144-411) module**

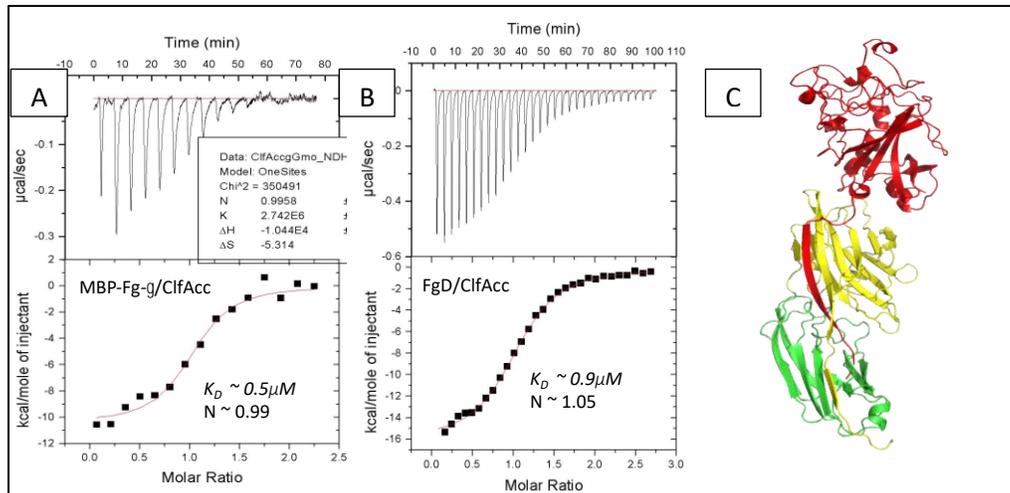


Figure-13: Isothermal Titration Calorimetry of ClfA with (a) MBP-Fgy(144-411) Fg-D. 30 injections of 10ul of ClfA (100uM) titrated into 10uM of FgD or MBP-Fgy(144-411). The thermodynamic plot were fitted and measurements were obtained using origin-ITC software. (c) Structural model of MBP-Fgy(144-411) interacting with ClfA (N2 (229-380) colored in green and N3 (381-545) colored in yellow respectively). The high affinity Fgy(144-392) site and the low affinity site Fgy(395-411) of MBP-Fgy(144-411) (colored in red).

### Fg Binding Site Analysis of Fbl and SpsD

Since ClfAcc exhibited a difference in both mechanism ( $\Delta G$  and  $-T\Delta S$ ) and binding affinity ( $K_d$ ) towards Fg, FgD and GV17 peptide, it is possible that other Fg binding MSCRAMMs (Fbl and SpsD) follow similar pattern of binding to these ligands. It is also possible that there might be amino acid residue differences in the binding site that contributes to the difference in the affinity values. In order to map and understand the amino acid differences and energetic contributions of each in the two binding sites (low affinity trench site and high affinity N3 site) in Fbl and SpsD, I mapped the amino acid differences in Fbl and SpsD at both of the Fg binding sites (Table-6,7,8). Low affinity site analysis: The ligand binding domain of ClfA comprised of N2 (230-366), a short linker

‘DYEK’ (367-370), N3 (371-520), and a C-terminal extension (521-545) of the N3 subdomain. The extreme C-terminal residues of the  $\gamma$ -chain (395-411) have been shown to interact at a hydrophobic trench in ClfA formed by residues of both N2 and N3 subdomain. Residues in the N2 domain (317-324) act as a lining trough of the trench and the incoming latch region (532-540) together forms an antiparallel  $\beta$  sheet meanwhile, the ligand  $\gamma$ -peptide docks to the ClfAcc in region (520-530). The residues at the N2-loop (336-340) locks the peptide in the groove by interacting with the Asn530 of ClfA, upstream of latch region. This complete interaction of ClfA binding to the  $\gamma$ -peptide of Fg is called Dock, Lock and Latch mechanism.

Table-6. Low affinity trench site binding analysis in Fbl and SpsD.

	N2 sequence	N3-latch	N2-Latch pair	N2-Loop
ClfA (221-545)	<b>IYFTDYVN</b> (317-324)	<b>GSGSGDG</b> (532-540)	<b>VKATLT</b> (330-335)	<b>PAYID</b> (336-340)
Fbl (205-530)	IYFTDYVT (300-308)	GSGNGDG (516-524)	ITGQIS (314-319)	PGYID (320-324)
SpsD (167-502)	RYRFMDYVNQ (267-275)	NNASGEG (491-497)	LKGKLA (361-366)	NLFIK (368-372)

The critical residues that actually interact are highlighted and listed in Table-6. In Fbl, only one of these residues (N520) differs from ClfA (S535) and the remaining residues are conserved. This indicates that sequence differences other than the peptide binding trench site (low affinity site), possibly at the D-domain binding site is responsible for the difference in the binding affinity with intact Fg when compared to ClfA/Fg interactions (Table-4).

In SpsD, although there are different residues in these regions, the critical residues (in bold) were mostly conserved and moreover these differences don't impact the Fg binding (SpsD binds to Fg with higher affinity than ClfA). This indicates that the trench site in SpsD, has little or no effect in the increased binding affinity towards Fg (Table-6).

Taken together these observations, it is unlikely that the amino acid differences in the N2N3 site (low affinity) as primary contributors to the observed differences in the binding affinity towards Fg in these MSCRAMMs.

### **Homology Modelling of FgD/Fbl and FgD/SpsD complex**

Using the structural information from the crystal structures of ClfA:FgD protein complex, a homology model of (1) Fbl:FgD and (2) SpsD:FgD protein complexes was built. The model of the protein complex allowed a refined alignment of the amino acid sequence of Fbl and ClfAcc, in addition identifying the gaps in the alignment and identifying the loops (if any) was possible. Utilizing the previous method, structure based sequence alignment is possible by narrowing down the critical residues that significantly impact the binding of Fg. It should be noted that there are certain loop regions that are missing in previously published crystal structures of ClfA (PDB ID:2VR3) and these loop regions are found to be participating in the interactions of FgD. Hence, we could not use the previously published ClfA structures of the peptide complex and I decided to use the ClfA:FgD complex structure (composite model) as template for modeling Fbl and SpsD structures. Since, there are no significant conformational changes or domain movements observed in the ClfA:FgD complex crystal structure, I decided to use the rigid body docking method for my docking studies of Fbl/FgD and SpsD/FgD complexes.

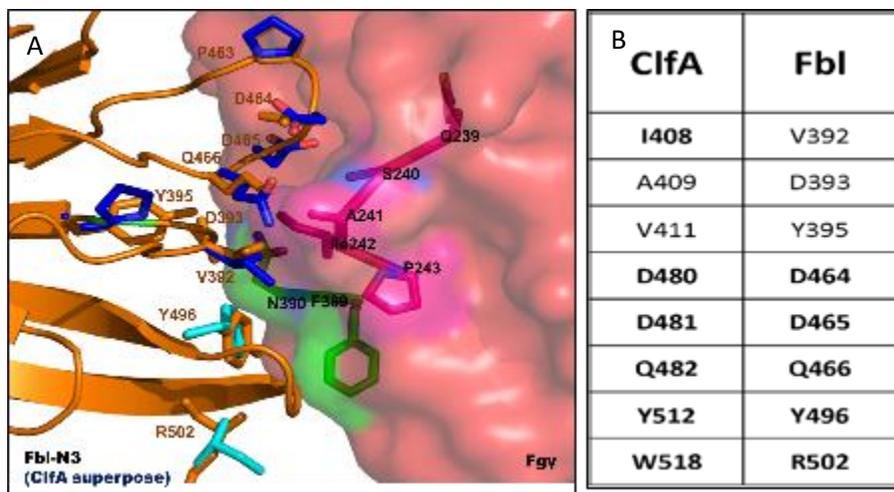
### **Generation of FgD/Fbl structural model**

Fbl(205-530) has high sequence homology (>60%) to ClfA. Hence modeling a Fbl-FgD complex is fairly straight forward without any gaps and loop regions also of similar length unlike SpsD (where two loops are different one shorter and one longer than ClfA, both are located at the N3 subdomain binding site of FgD). The model fits accurate with an RMSD of 0.25 angstroms for the Fg-D protein (Figure-14).

Based on my docking studies of Fbl/FgD complex model, The FgD interacting residues of Fbl are found to be V392, D393, Y395, D464, D465, Q466, Y496, R502 (Figure-14 and Table-6). All of these residues are located in the N3 subdomain of Fbl. Four out of the above listed eight residues are identical to ClfA (D464, D465, Q466 and Y496 is identical to D480, D481, Q483 and Y512 of ClfAcc). V392 and Y395 share similar hydrophobic side chains to the corresponding residues of ClfAcc (I408 and V411). D393 (acidic) and R502 (basic) differ from the corresponding residues of ClfAcc (A409 and W518) respectively.

Overall in Fbl, more than half (6/8) of the residues are favorable (V392, Y395, D464, D465, Q466 and K431) for the interaction with the FgD (Table-6). The differing side chain residues (D393 and R502) were not participating with the interaction of hydrophobic pocket (QSAIP motif) which comprises region (239-243) of the Fgy but seems to interact with the short helical region (F389 and N390) of the Fgy chain (Figure-14). Hence, based on my modeling/docking studies of Fbl/FgD, I propose that the N3 site of Fbl would interact with the FgD with similar affinity (KD) that of ClfAcc/FgD interaction.

**Figure-14.** Strucural Analysis of N3 binding site of the Fbl/Fgy structure (Homology Model).



**Figure-14. Analysis of Fbl-N3/Fgy Binding interface. (A)** Homology model of Fbl-Fgy complex. Fbl-N3 subdomain is shown as cartoon and colored in orange. The side chains of Fbl residues that potentially interact with Fgy is shown in sticks and the corresponding ClfA residues shown in blue and cyan (side chains only). The Fgy (Red) are only shown for clarity as surface. The binding interface QSAIP motif in Fgy colored in pink. The dipeptide helix(C) of Fgy (F389,N390) is colored in green and this binding region in ClfA is colored in cyan. **(B) Table-7.** Comparison of aminoacid residues in Fbl and ClfAcc at the N3 subdomain site that interacts with the D-fragment of Fg.

### Generation of FgD/SpsD structural model

I generated structural models SpsD(167-502) in complex with FgD using the ClfA structure from the ClfA:FgD complex (current work) using MODELLER tool. The docking studies helped to narrow down the list of residues located in the N3 site of SpsD that could interact with the FgD of Fg (Figure-15). Based on my docking studies of SpsD/FgD complex model, the FgD interacting residues of SpsD are T360, N361, T363, K431, D432, N433, and Y460 and Y468. All of these residues are located in the N3 subdomain of SpsD (Table-8). Two out of the above listed eight residues are identical to ClfA (D432 and Y460 is identical to D481 and Y512 of ClfAcc). N433 of

SpsD share similar polar side chain of Q482 of ClfAcc and Y468 of SpsD share similar hydrophobic side chains of W518 of ClfAcc. K431 has a positively charged side chain whereas D480 is negatively charged. T360, N361 and T363 have hydrophilic side chains but the corresponding residues of ClfA (I408, A409 and V411) have hydrophobic side chains. Overall in SpsD, half of the residues (T360-T363 and K431) that interact with the FgD are reversed to that of the respective residues of ClfAcc (I408-V411 and D480) (Table-8). Nevertheless, these residues only partly interact with the hydrophobic pocket (QSAIP motif) comprising region (239-243) of the Fg $\gamma$  (Figure-15). Hence, based on my modeling/docking studies of SpsD/FgD, I propose that the N3 site of SpsD would interact with the FgD with similar to (or) with a slightly weaker affinity ( $K_D$ ) than ClfAcc/FgD interaction.

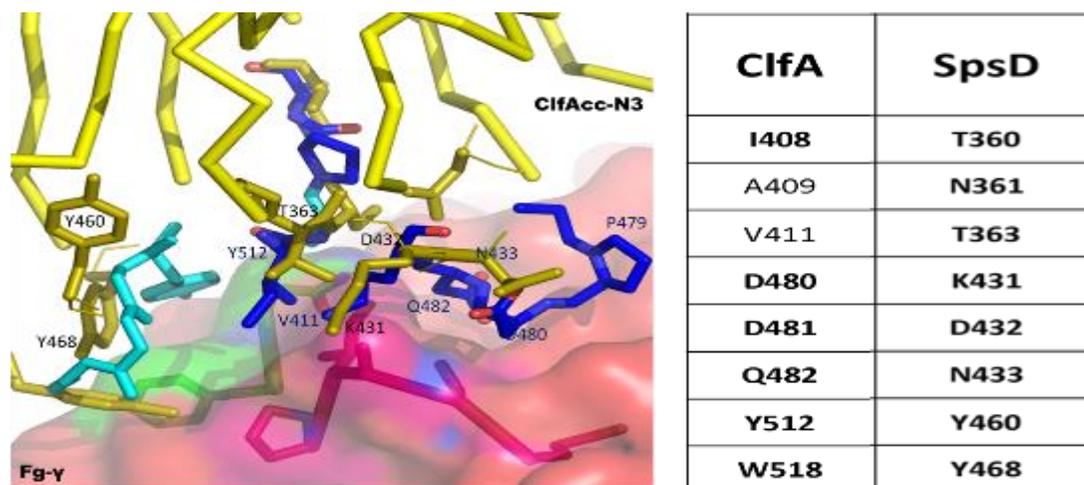
### **Interaction between Fbl and Fibrinogen**

To measure the binding affinities of Fbl towards Fg and its fragments (GV17 and FgD), I performed ITC studies using purified proteins of Fbl (205-530) with full length Fg, GV17 peptide and Fg-D (Figure-16). The binding affinities ( $K_{DS}$ ) were recorded and the respective affinities and the thermodynamic parameters were summarized in Table-9.

In my ITC studies, Fbl interaction with the GV17 peptide showed an enthalpy of -5.6 kcal mol<sup>-1</sup> and an entropy of -0.7 kcal mol<sup>-1</sup>. With the intact Fg, the interaction had similar enthalpic and entropic values of ClfAcc/Fg interaction. These results show that Fbl/peptide binding lacks the high affinity Fg interaction (FgD) at the N3 site and Fbl interacts with Fg in similar mechanism to that of ClfAcc/Fg interaction (Table-4 and Table-9).

As expected, Fbl interaction with FgD showed high affinity with a higher enthalpic and entropic values similar to the Fbl/Fg interaction. These values are equivalent to that of ClfAcc/Fg interaction indicating a similar ClfA like Fg binding mechanism for Fbl.

**Figure-15.** Strucural Analysis of N3 binding site of the SpsD/Fgy structure (Homology Model).



**Figure-15. Analysis of SpsD-N3/Fgy Binding interface. (A)** Homology model of SpsD-Fgy complex. SpsD-N3 subdomain is shown as cartoon and colored in olive green. The side chains of SpsD residues that potentially interact with Fgy is shown in sticks and the corresponding ClfA residues shown in blue and cyan (side chains only). The Fgy(Red) are only shown for clarity as surface. The binding interface QSAIP motif in Fgy colored in pink. The dipeptide helix(C) of Fgy (F389,N390) is colored in green and this binding region in ClfA is colored in cyan. **(B) Table-8.** Comparison of aminoacid residues in SpsD and ClfAcc at the N3 subdomain site that interacts with the D-fragment of Fg.

**Table-9.** ITC thermodynamic parameters for Fbl binding to Fg

	Fbl	$n$	KD ( $\mu$ M)	$\Delta G$ (kcal mol <sup>-1</sup> )	$\Delta H$ (kcal mol <sup>-1</sup> )	-T $\Delta S$ (kcal mol <sup>-1</sup> )
Peptide/0.5mM	20	1.05±0.31	30	-5.93	-5.6±0.04	-0.34
Fg/0.01mM	150	1.14±0.02	0.8	-8.0	-13.2±0.03	6.5
FgD/0.01mM	150	1.26±0.12	1.6	-7.6	-16.7±0.05	9.10

Figure-16. ITC binding assay of Fbl with Fg and Fg fragments

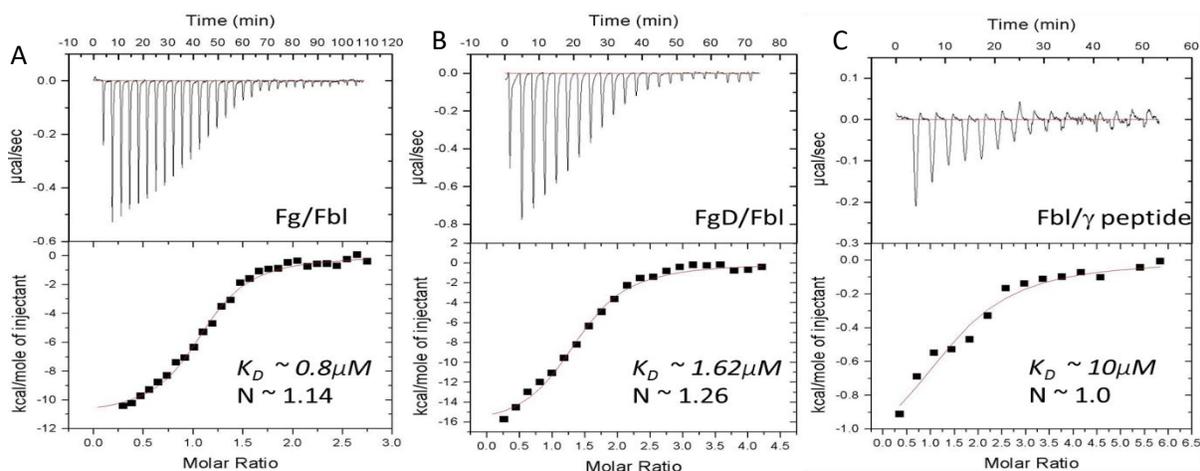


Figure-16. Isothermal Titration Calorimetry of Fbl(205-530) with (a) Fg. 30 injections of 10ul of Fbl (100uM) titrated into 10uM of Fg until saturation. (b) GV17. 30 injections of 10ul of GV17 pep (500uM) titrated into 20uM of Fbl until saturation. (c) FgD. 30 injections of 10ul of ClfA (100uM) titrated into 10uM of FgD until saturation. The thermodynamic plot were fitted and measurements were obtained using origin-ITC software.

### Interaction between SpsD and Fibrinogen

To measure the binding affinities of SpsD towards Fg and its fragments (GV17 and FgD), I performed isothermal titration calorimetry studies (ITC) using purified proteins of SpsD (167-502) with full length Fg, GV17 peptide and Fg-D (Figure-17). The binding affinities ( $K_{DS}$ ) and thermodynamic parameters were summarized in Table-10.

SpsD interacted with FgD showed high affinity with a higher enthalpic and entropic values. Surprisingly, this interaction of SpsD/FgD has a significantly higher enthalpy and entropy ( $-21.29$  and  $13.42 \text{ kcal mol}^{-1}$ ) than that of ClfAcc/Fg ( $-13.2$  and  $4.63 \text{ kcal mol}^{-1}$ ). However, the total Gibbs free energy of SpsD/D-frag and ClfAcc/D-frag remain almost similar ( $-8.56 \text{ kcal mol}^{-1}$  and  $-8.59 \text{ kcal mol}^{-1}$ ) (Table-4 and Table-10).

Figure-17. ITC binding assay of SpsD with Fg and Fg fragments

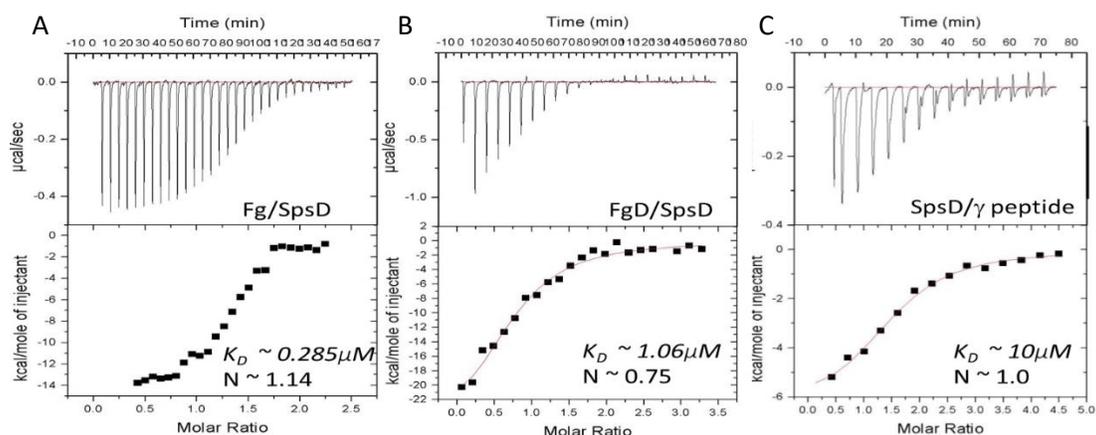


Figure-17. Isothermal Titration Calorimetry of SpsD(167-502) with (a) Fg. 30 injections of 10ul of SpsD (100uM) titrated into 10uM of Fg until saturation. (b) FgD. 30 injections of 10ul of SpsD (100uM) titrated into 10uM of FgD until saturation. (c) GV17. 30 injections of 10ul of GV17 pep (500uM) titrated into 20uM of SpsD until saturation. The thermodynamic plot were fitted and measurements were obtained using origin-ITC software.

The affinities of ClfA, Fbl and SpsD towards various fragments of Fg are listed in Table-11. ClfAcc, Fbl and SpsD all interacts with low affinity to the  $\gamma$ -peptide but all of them interact with FgD with high affinity. The affinity of FgD is at least ten fold higher than the corresponding affinities with the  $\gamma$ -peptide. Furthermore, the affinity of FgD is only two fold weaker than the corresponding affinities towards intact Fg. This indicates that ClfA, Fbl and SpsD all interact with an additional region in Fg other than the previously known C-terminus of the Fg- $\gamma$ (395-411). All of the Gamma chain binding MSCRAMM's exhibit similar binding affinities across various fragments of Fg indicating a general binding mechanism. This indicates that Fbl and SpsD also has additional binding sites (N3 site) that could potentially bind to the globular domain of Fg- $\gamma$ -regions comprising (144-392).

### Mapping of N3 binding site residues in ClfAcc, Fbl and SpsD that interacts with Fg

The Fg  $\gamma$ -module (144-392) interacts with the N3 subdomain (380-530) of ClfA. Mainly, The QSAIP motif comprising region (239-243) and a loop region connecting helix(C) (F389 and N390) of the  $\gamma$ -module constitutes a major continuous stretch of residues that make primary contact with the N3 subdomain of ClfA.

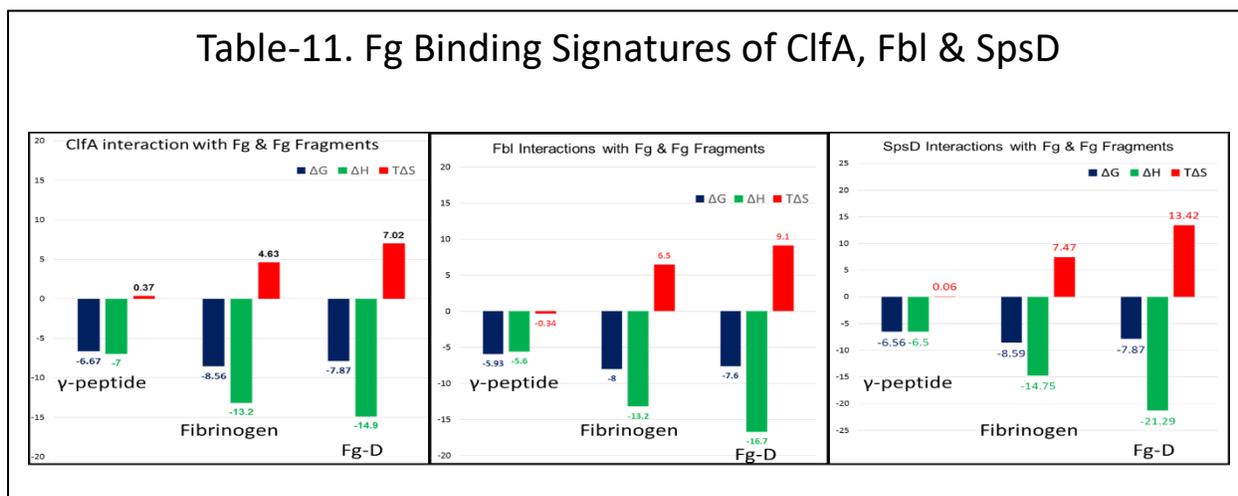
Table-10. ITC thermodynamic parameters for SpsD binding to Fg.

	SpsD	n	KD ( $\mu$ M)	$\Delta$ G (kcal mol <sup>-1</sup> )	$\Delta$ H (kcal mol <sup>-1</sup> )	-T $\Delta$ S (kcal mol <sup>-1</sup> )
Peptide/0.5mM	20	1.05 $\pm$ 0.31	<b>10</b>	-6.56	<b>-6.5<math>\pm</math>0.06</b>	<b>0.06</b>
Fg/0.01mM	150	1.14 $\pm$ 0.02	0.285	-8.59	<b>-14.75<math>\pm</math>0.03</b>	<b>7.47</b>
FgD/0.01mM	150	1.06 $\pm$ 0.12	<b>1.06</b>	-7.87	<b>-21.29<math>\pm</math>0.03</b>	<b>13.42</b>

The QSAIP motif docks in a shallow groove in ClfA formed by residues I408, D480, D481, Q482 and Y512. The dipeptide motif, F389 and N390 of the  $\gamma$ -module interacts with the A409, Y512 and I516 of ClfA-N3 subdomain (Figure-10) of ClfAcc. In addition, residues K151, D155, N158, L172, and M264 of the  $\gamma$ -module contact V411, N463, E475, T478 and W518 of the ClfA-N3 subdomain (Figure-10).

Using the structural information from the crystal structures of (1) ClfA:FgD protein complex, (2) homology model of Fbl:FgD and (3) SpsD:FgD protein complexes, we have narrowed down the critical residues that will significantly impact the binding of Fg. We have substituted these target residues all to alanine residues (alanine lacks a charged side chain). Hence, the side-chain contribution from any of these residues will be lost resulting in a significantly decreased binding affinity towards full length Fg.

Table-11. Fg Binding Signatures of ClfA, Fbl & SpsD



### Fibrinogen Interaction Studies of N3 site mutants of ClfAcc, Fbl and SpsD

The purified mutants of ClfAcc, Fbl and SpsD (Table-7 and 8) were tested for their binding affinity towards full length Fg and the binding affinities were compared to wild type affinities of ClfA, Fbl and SpsD (Figure-18a- c). As a positive control, we also tested these mutants binding to the GV17 peptide which is known to bind only at the trench site located between the N2-N3 subdomains and should not affect the GV17 peptide binding (Figure-18d-f) and comparable to the wild type affinities with Fg and Fg fragments summarized as in Table-11.

Figure-18. ITC binding assay of N3 site mutants of ClfAcc with Fg and Fg fragments

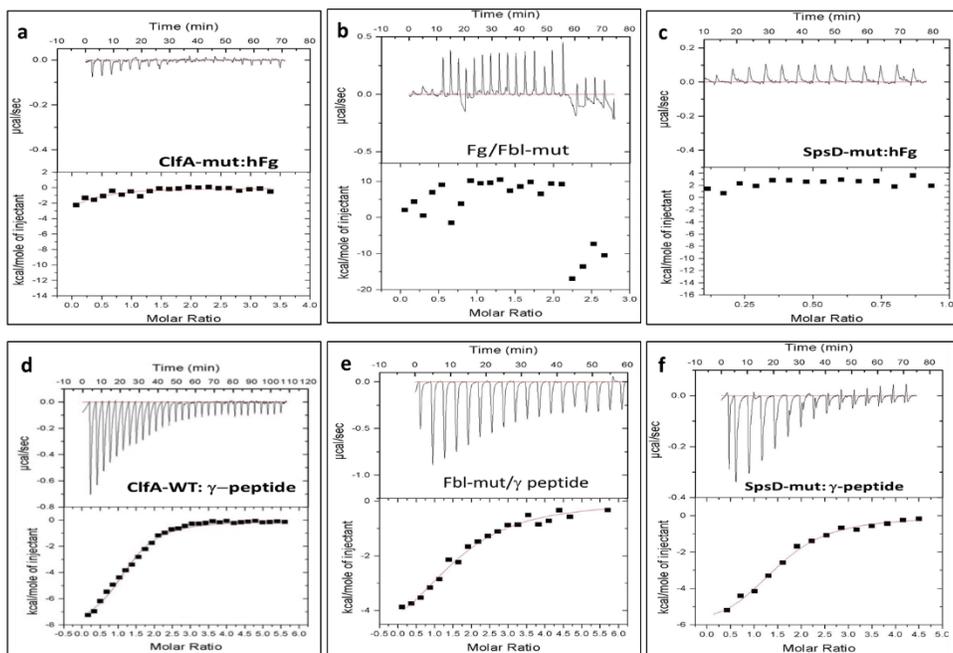


Figure-18. ITC binding assay of N3 site mutants of ClfAcc with Fg and Fg fragments. (a) ClfAcc-N3 mutant binding to Fg. (b) Fbl-N3 mutant binding to Fg. (c) SpsD-N3 mutant binding to Fg. (d) ClfAcc-N3 mutant binding to Fg- $\gamma$  peptide. (e) Fbl-N3 mutant binding to Fg- $\gamma$  peptide. (f) SpsD-N3 mutant binding to Fg- $\gamma$  peptide.

## Discussion of Chapter-II

### ClfAcc:FgD lacks the electron density for the last C-terminus of the Fg $\gamma$ (393- 411)

Electron density for the Fg- $\gamma$ (393-411) was missing in the crystal structure of ClfA:FgD protein complex. There are multiple reasons for the missing electron density at Fg- $\gamma$ (393-411), (1) the presence of a protease labile site at this position and plasmin digested this site. (2) High level of distortedness resulted in lack of sufficient electron density for mapping this site.

Hence, the structural details of the ClfAcc/Fg $\gamma$  (393-411) interactions could not be obtained directly from the crystal structure of our ClfAcc/FgD complex. However, an indirect approach was attempted by superimposing the crystal structures of ClfA (1) ClfA:FgD (current work) and (2) ClfA:GV17 (PDB ID:2VR3). There is a gap of residues in Fg- $\gamma$ (391-396) which was manually

built using PyMoL. The resulting composite model successfully used to map both of the binding sites of ClfAcc in Fg (Figure-11).

### **The Fgy Module studies with ClfAcc links the low and high affinity site in ClfAcc**

Although the ClfA:FgD crystal structure lacks the structural details of the second site in Fg (GV17), the binding studies on MBP-Fgy(144-411) with ClfA showed higher affinity (0.5uM) than FgD (1uM). This binding affinity measurements of ClfA with MBP-Fgy(144-411) were equivalent to that of full-length Fg (0.3uM) as evident from the thermodynamic parameters listed (Table-4 and Figure-11). Thus the binding studies supports the observations from the composite model and provides the missing details required for the understanding of ClfA:Fg interactions.

### **Binding Site analysis of ClfA/Tefibazumab in comparison with ClfAcc/FgD**

A mouse mAb (aurexis) raised against ClfA<sub>(229-545)</sub> (Newmann) showed great promise in protecting mice from *S. aureus* septicemia. The humanized version of this mouse mAb, called 'Tefibazumab' showed high affinity to ClfA<sub>(229-545)</sub> (Newmann) and its variants found in other strains of *S. aureus* (Figure-19a). Tefibazumab showed modest inhibition of ClfA (Newmann) to Fg (IC<sub>50</sub> =0.4μM) (Figure-19b) and tefibazumab was less impressive in a phase II human trial (71).

A systematic analysis to discover the reasons of the failure had been published from our lab. (i) Tefibazumab did not inhibit the C-terminal γFg binding site in ClfA. (ii) Crystal structures of ClfA-Tefibazumab complex shows the binding site for Tefibazumab is located at the N3 subdomain. (iii) Tefibazumab inhibits a site in ClfA-N3 subdomain that also has a potential to bind an unknown site in Fg (3). However, the exact residues that contribute to the ClfA:Fg interface was not known. My current work on ClfA:FgD clearly pinpoints the interacting residues of both ClfA and FgD and also explains the difference in the tefibazumab and Fg binding interface in the ClfA-N3 subdomain. The results are discussed in the following section.

Structural superposition of the two structures of ClfA:FgD and ClfA:Tefibazumab, shows that both FgD and Tefibazumab indeed bind at the top of the N3 subdomain of ClfA (Figure-20). However, a detailed amino acid comparison on the N3 subdomain of ClfA indicates that while out of the 12 critical residues that interacts with FgD, tefibazumab binding to ClfA could inhibit only 2 out of 12 residues (Y512 and W518) (Figure-21). Mainly, these residues do not interact with the QSAIP motif but nevertheless crucial for the interaction of Fg. Thus ClfA:FgD crystal structure helps in giving a clear amino acid level information that reasons/agrees with the ClfA:Tefibazumab studies published earlier (3). Tefibazumab binding studies using the ClfA-mutants (Q429K, S461Q, N463R, N515R), which showed decreased affinity towards Tefibazumab but not towards Fg (Figure-19a, b and 21a). From the crystal structure of ClfA:FgD, we discovered that these residues of ClfA, although located at the N3 subdomain are not interacting with Fg (Figure-20 and 21b). However, P467(light chain) Y512,W518 (heavy chain) of ClfA that had been shown to bind to tefibazumab partially overlap with the ClfA binding site of Fg (Figure-21c). Thus, further work is warranted to improve the specificity and efficiency of a ClfA antibody that inhibits Fg-ClfA interaction.

### **ClfAcc, Fbl and SpsD share a conserved N3 binding site for FgD**

The N2N3 subdomains of wild type ClfA (229-545),Fbl(205-530) and SpsD(167-502) shows similar and high affinity to FgD as that of ClfA binding to the FgD (Figure-7,16,17). This affinity is twofold weaker than the full length Fg but 10 fold higher than the GV17 peptide which binds to the N2N3 binding site (trench) of ClfA, Fbl and SpsD (Figure-4,9,10). Collectively, this proves that the functional homologs of ClfA, Fbl and SpsD also bind to the FgD at their N3 subdomain.

Using the structural models of Fbl:FgD and SpsD:FgD complex (Figure-14,15), we have identified the critical residues that are important for the interaction of FgD. We have generated the

N3 site mutants of ClfA, Fbl and SpsD and tested the binding of these mutants against Fg and GV17 peptide (Figure-18). Since the residues mutated are not present in the trench site that binds to the peptide (GV17), all of these mutants showed positive and equivalent affinity to that of the WT, but unable to bind to the full-length Fg. This proves that ClfA, Fbl, SpsD all have a conserved N3 site that binds to the FgD of Fg.

### **Biological Significance of ClfAcc/FgD interactions**

Crystal structure of the ClfA:Fg complex not only revealed the atomic structural model of the key interaction, but also provided the structural basis for several biological observations.

### **ClfAcc inhibits Fg binding to platelet and neutrophil integrins**

Fg acts as a cross bridge that connects hemostatic factors and immune factors. Fg is known to interact with multiple cells including platelets, leukocytes and effect the changes in cell migration, phagocytosis, NF- $\kappa$ B-mediated transcription, production of chemokines and cytokines, degranulation, and other processes (72-74). Fg interacts with the integrin receptors present on the surface of these cells. Specifically, the last 5 residues of the Fg- $\gamma$ chain (407-411) had been known to be important for platelet recognition through  $\alpha_{IIb}\beta_3$ . It has already been known that the N2N3 site (trench site) of ClfA binds to Fg- $\gamma$ chain (407-411) and inhibits the platelet activation/aggregation. Also, residues Fg- $\gamma$ chain (390-396) is crucial for the  $\alpha_M\beta_2$  interaction in neutrophil activation of phagocytosis process (73). The genetic elimination of this site in Fg, predisposes mice to *S. aureus* infection and poses an impediment to microbial clearance. My composite model (figure-11, results section) shows that the neutrophil integrin site 'Fg- $\gamma$ chain (390-396) interacts to the N3 subdomain of ClfA.

Figure-19. SPR binding assay of ClfA and ClfA variants with Tefibazumab and Fg. Reprinted from (78)

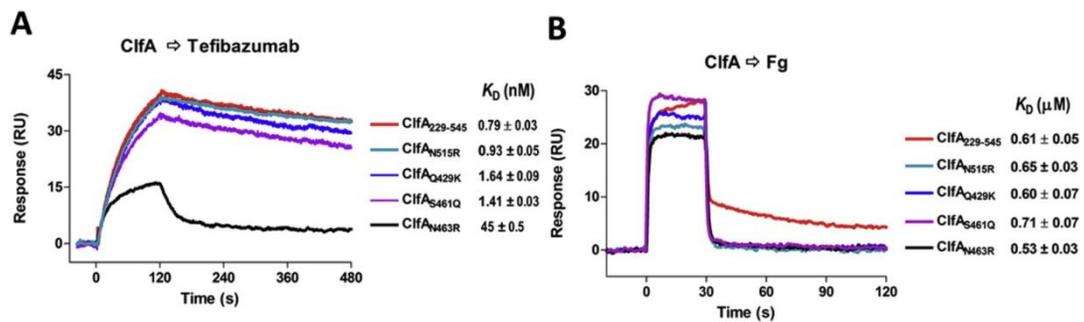


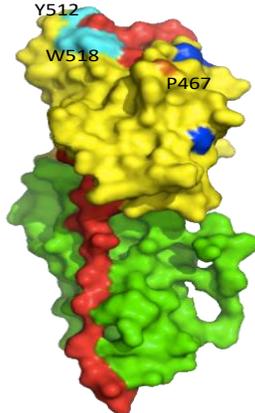
Figure-19: (A) Immobilized Tefibazumab binding to ClfA and ClfA mutants by SPR. (B) Immobilized Fibrinogen binding to ClfA and ClfA mutants by SPR (78). Reprinted from {Ganesh, 2016 #945}

### ClfAcc inhibits fibrin polymerization at the $\gamma$ - $\gamma$ chain interface of D-dimer

The  $\gamma$ -module contains another important biological hot-spot required for fibrin polymerization. During fibrin polymerization, two  $\gamma$ -modules from different Fg molecules form a  $\gamma$ - $\gamma$ -interface required for Fn assembly. Comparison of the  $\gamma$ - $\gamma$ -interface region in the fibrin structure and ClfA/FgD complex provided a structural basis for the ability of ClfA to adhere to both Fg and Fibrin filaments. The ClfA N3-subsite contact region in Fg  $\gamma$ -module is in a unique location of the Fg molecule that can influence several aspects of fibrin polymerization and the ability to adhere to different forms of fibrin(ogen) with high affinity, a binding feature that explain bacterial clumping (soluble Fg) and agglutination (fibrin cables). The end-to-end polymerization of the  $\gamma$ -module is an important step in fibrin assembly/polymerization associated with the knob-hole interaction (Figure-22a). A fibrin oligomer of three Fg molecules is a basic building block for protofibril

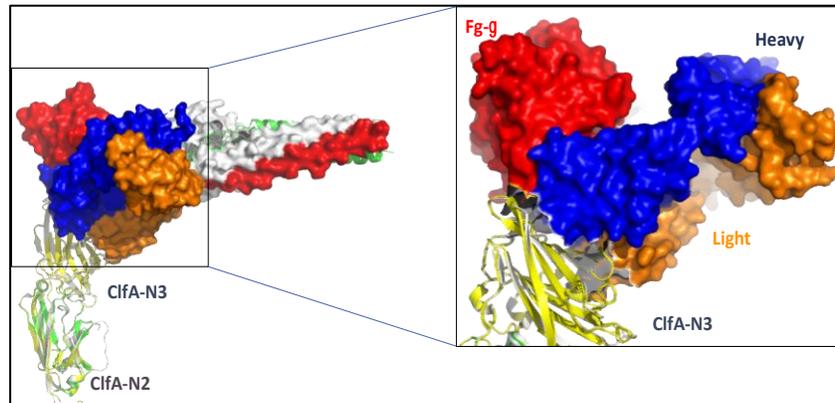
formation and eventually to fibrin polymerization(75). Fig. 22a shows the common and the unique regions in the two  $\gamma$ -modules that form a  $\gamma$ - $\gamma$ -interface.

**Figure-21. ClfA-N3 binding Site analysis of FgD Tefibazumab(3) and FgD.**

ClfA-N3 interacting residues (Tefibazumab binding site)				ClfA-N3 (Fg binding site)
	Affinity Values (tefibazumab)	Affinity Values (Fibrinogen)		
ClfA-WT	0.79nM	610nM		I408
ClfA-Q429K	1.64nM	600nM		A409
ClfA-S461Q	1.41nM	710nM		V411
ClfA-N463R	45nM	530nM		E475
ClfA-N515R	0.93nM	650nM		T478
ClfA- P467A,Y512A,W518A	No binding	No binding		P479
				D480
				D481
				Q482
				Y512
				I516
				W518

**Figure-21:** Structural comparison of ClfA:FgD with ClfA:Tefibazumab(3). (a) SPR binding affinities of ClfA-N3 site mutants with Tefibazumab and Fibrinogen. The triple mutant interacts with P467A (light chain) colored in orange, Y512A and W518A (heavy chain) colored in blue respectively. (b) surface representation of ClfA (N2 and N3 subdomains of ClfA are colored in green and yellow). The ClfA residues that interact with Fg is colored red and residues that overlap with the tefibazumab binding is colored in cyan (Y512,W518) and orange (P467) respectively. (c) list of ClfA residues that interact with FgD. TPDDQ(478-482) interacts with the QSAIP motif of FgD, Y512, W518 residues also bind to tefibazumab showing partial overlap/inhibition of Fg. Adapted from {Ganesh, 2016 #945}

**Figure-20. Structural comparison of ClfAcc:FgD with ClfA:Tefibazumab**



**Figure-20:** Structural comparison of ClfA:FgD with ClfA:Tefibazumab. (a) superposition of crystal structures of ClfA:Tefibazumab with ClfA:FgD. The N2 and N3 subdomains of ClfA are colored in green and yellow respectively. The FgD  $\alpha$ ,  $\beta$ ,  $\gamma$  chains are colored in green, grey and red respectively. The heavy and light chains of tefibazumab are colored in blue and orange respectively.

ClfA binds to a unique region, where it can interfere with the end-to-end polymerization  $\gamma$ - $\gamma$ -interface (Figure-22b-e) when bound to a in the buried in only one of the  $\gamma$  modules that form the  $\gamma$ - $\gamma$ -interface. Therefore, ClfA can also adhere to Fibrin fibrils by binding to one of the  $\gamma$ -module in the preformed  $\gamma$ - $\gamma$ -interface. This structural model is consistent with the earlier observations that ClfA has the ability to bind to Fg and delay clotting and affected crosslinking of FgDs and also can adhere to fibrin wherever possible. It is interesting to note that ClfA can interfere fibrin polymerization but the  $\gamma$ - $\gamma$  dependent polymerization event cannot completely compete out ClfA binding (sites). Modeling of the ClfA D-dimer, as observed in fibrin, indicates that additional weak cohesive contact can occur as a result of ClfA interacting close to the FgDs of the Fg-Fg interface (Figure-22).free Fg by competing for the  $\gamma$ - $\gamma$ -interface and imposing steric hindrance (Fig. 22E).

The inhibition would be more profound after reaching a threshold concentration of ClfA. The ClfA binding region in the buried in only one of the  $\gamma$  modules that form the  $\gamma$ - $\gamma$ -interface. Therefore, ClfA can also adhere to Fibrin fibrils by binding to one of the  $\gamma$ -module in the preformed  $\gamma$ - $\gamma$ -interface. This structural model is consistent with the earlier observations that ClfA has the ability to bind to Fg and delay clotting and affected crosslinking of FgDs and also can adhere to fibrin wherever possible. It is interesting to note that ClfA can interfere fibrin polymerization but the  $\gamma$ - $\gamma$  dependent polymerization event cannot completely compete out ClfA binding (sites). Modeling of the ClfA D-dimer, as observed in fibrin, indicates that additional weak cohesive contact can occur as a result of ClfA interacting close to the FgDs of the Fg-Fg interface (Figure-22).

### **Analysis of ClfAcc/FgD structure at the crystal lattice contact interfaces**

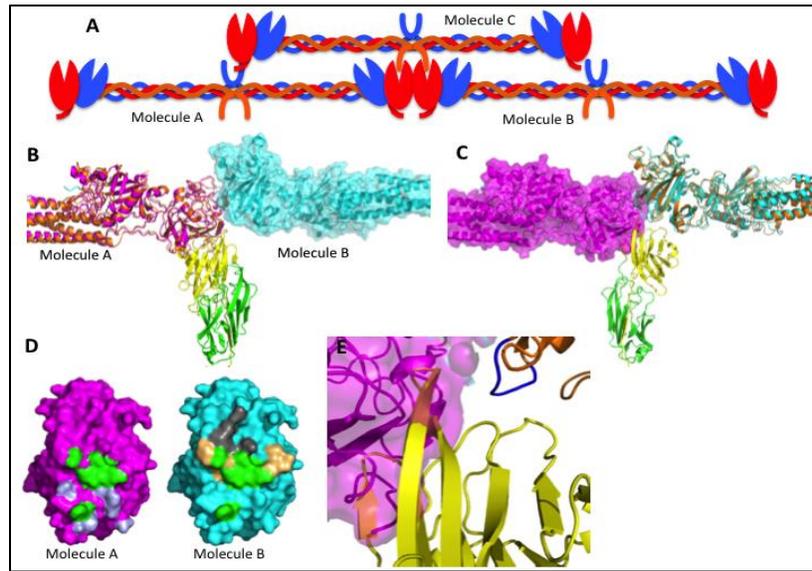
In a crystal structure, there is always more than one conformation of a protein or protein complex. These alternative conformations arise due to the nature of packing of the protein inside the protein crystal. Furthermore, depending on the asymmetric unit cell and space group dimensions, there may be one half of the protein complex in the unit cell (1/2 asymmetric molecule per unit cell) or two molecules in the unit cell (2 asymmetric molecules per unit cell). Based on the various biochemical binding data (ITC, SPR, ELISA), structural biologists use these information to constructs a structural model (biological assembly) that makes a biological sense. The following discussion is an attempt to describe one of those model that might have a potential biological sense for the ClfA:Fg interaction and pathogenesis of *S. aureus*.

Apart from the  $\gamma$ -module interface, various additional interfaces were observed inside the crystal structure of ClfA:FgD, namely (1) the coiled coil interface (Figure-23a) and (2) the beta chain interface (not shown). The critical residues of Fg that are involved in the binding interface are listed below (Figure-23b). The free energy calculations of the various binding interfaces are listed in Table-12.

### **M1 and ClfAcc targets an overlapping site in coiled coil region of Fgy chain**

The M protein is a surface protein of *streptococcus pyogenes* and is a virulence factor associated with streptococcal toxic syndrome (STSS)(76). M1 protein binds to Fg and this complex interacts with  $\beta$ 2 integrins of neutrophils and triggers the release of heparin binding protein (HBP), a vasodilator and marker for sepsis. Recently, the crystal structure of the M1 fragment the M1 fragment M1<sup>BC1</sup> (residues 132-263, ~17 kDa) and Fg fragment D was solved and published (1). Two molecules of M1 forms complex with four FgD in a cross like pattern and this M1-Fg network is essential for the triggering of neutrophils(1).

Figure-22. Schematic representation of knob-hole interactions of Fg targeted by ClfA



At the interface level, the M1 interacts with the coiled coil region of FgD, in contrast to ClfA which binds to the  $\gamma$  module of FgD. However, the analysis of all the many conformations of ClfA:FgD complex observed in the crystal contacts, it is revealed that there is a conformer of ClfA:FgD complex where the N2 subdomain interacts with the coiled coil region of FgD. This site overlaps with the M1-FgD complex interface of FgD (Figure-24). Thus, ClfA in addition to targeting the  $\gamma$ - $\gamma$  interface of Fg, also targets a coiled coil site which is also the binding site for M1 protein, a virulence factor of *Streptococcus pyogenes*.

The N2 subdomain region has a metal ion binding site called as ‘EF hand motif’ comprising regions (310-321) in ClfA (Figure-25). The coiled coil interacting residues of ClfA are D301, Q302, T325, K326, T327, D537, I539, and D540. These residues are located closer to the EF-hand motif (310-321) of ClfA(10).

## **Future Directions**

The findings of my work on ClfAcc/FgD clearly give the molecular details of ClfA interactions of Fg and identified the critical residues that has to be targeted both in ClfA and Fg. From this study, we can improve the design of Fg inhibiting antibodies like tefibazumab for vaccine development against *S. aureus* (work in progress)

From the comparative studies on Fbl/Fg and SpsD/Fg, it is clear that these proteins indeed have similar binding sites for Fg. Since, these are known virulence factors for *S. lugdunensis* and *S. pseudintermedius*, we could potentially develop a broad specificity antibodies and small molecules that can effectively neutralize/inhibit Fg interactions of ClfA, Fbl and SpsD (Future work). I have found out from the literature a number of natural variations of Fg that are present at the ClfAcc binding site (Fg $\gamma$ (144-411)) would be looked into and the binding affinities against ClfA, Fbl and SpsD will be characterized (Future work). A Fg $\gamma$  mutant at the QSAIP motif already been designed and expressed which is addressed in the methods section of chapter II.

Figure-23: Analysis of Fg-coiled coil interface interactions with ClfAcc

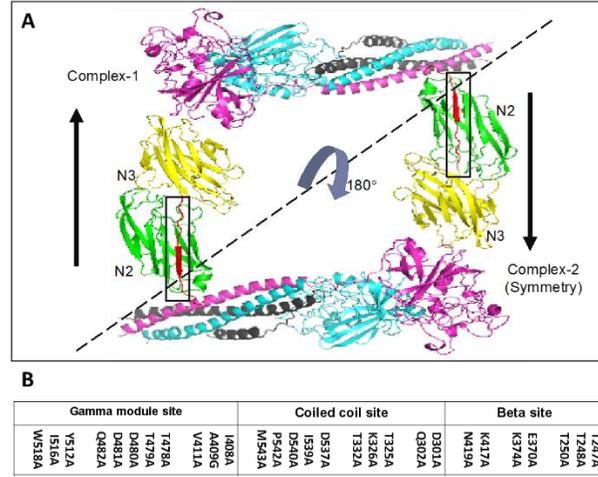
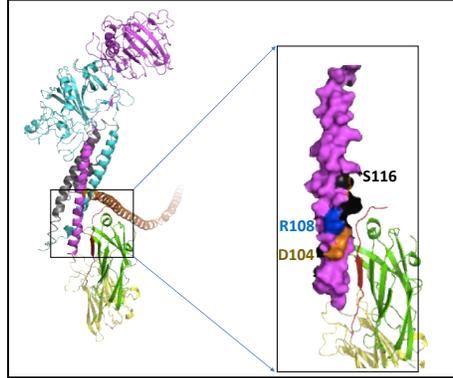
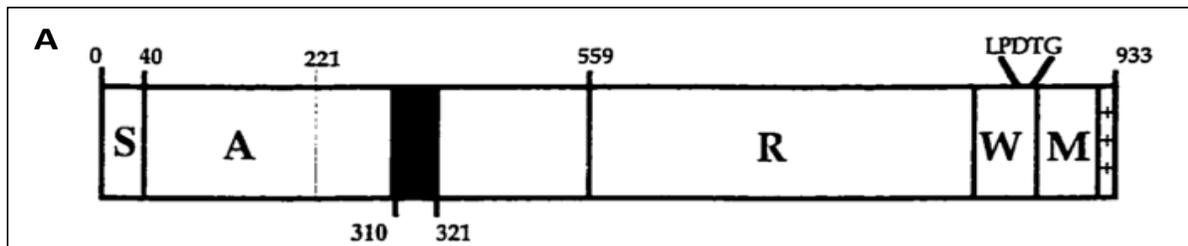


Figure-24. M1(1) and ClfAcc binding interface overlaps at the coiled coil region in Fgy



**Figure-24:** Structural Comparison of M1 and ClfA Binding Sites in FgD. (A) Overlay of ClfA:FgD structure over M1:FgD complex (PDB: 2XNX). FgD  $\alpha, \beta, \gamma$  chains are colored in grey, cyan and magenta respectively. The M1 (helix) colored in orange and ClfA N2 and N3 subdomain are colored in yellow and green respectively. The binding sites and overlapping residues are zoomed in (black box). The coiled coil region of  $\gamma$ -chain from FgD alone shown as surface for simplicity. The unique residues for interacting to the M1 colored in black (S116). The unique residues for interacting to the ClfA-N2 subdomain colored in orange (D104). The overlapping residue (R108) common to both M1 and ClfA is colored blue.

**Figure-25:** The EF hand motif (310-321) in ClfA overlaps with Fgy-coiled coil region.



**B** Proposed divalent cation-binding motif in ClfA

310-321 **D S D G N V I Y T F T D**

**Figure-25:** A) Schematic model of ClfA showing domain organization indicating the EF-hand location in the N2 subdomain. S. signal sequence, A. ligand binding IgG domains R. ser-asp repeat regions. W. cell wall anchoring region. LPDTG. sortase anchoring motif. M, membrane spanning domain. The putative metal-ion binding 'EF-hand motif' is colored black comprising regions between 310-321. (B) Amino acid sequences with critical residues in bold of the proposed divalent cation binding motif in ClfA(10) reprinted from {O'Connell, 1998 #618}.

## Summary of Chapter-2

ClfA, Fbl and SpsD are virulent surface proteins of *S. aureus*, *S. lugdenensis* and *S. pseudintermedius*. ClfA and Fbl share similar structural domain organization with a ligand binding IgG like (N1, N2 and N3) A-domain at the N-termini followed by a stretch of SD repeats at their C-termini. The SD repeat acts as a connecting bridge between the N-terminal ligand binding A-domain and the C-termini cell wall anchoring 'LPXTG'. The main difference between ClfA and Fbl is the lack of an aureolysin cleavage site (SLAVA) between the N1 and N2 domain that is present in ClfA. The aureolysin processing of ClfA exposes the N2 and N3 subdomains for ligand binding that interacts with Fg (77, 78).

SpsD, has a ligand binding A-domain at their N-termini that is similar to ClfA but in addition SpsD has an R-domain at the C-terminal side of the A-domain that is rich in proline residues. The R-domain is believed to interact with the host fibronectin, a pattern similar to another *S. aureus* protein called FnbpA (Fibronectin binding protein).

The interaction of these surface proteins of Staphylococci with Fg involves multiple regions in these proteins (1) the low affinity site at the N2N3 trench that targets the C-terminal region (392-411) of Fgy (2) the high affinity site at top of the N3 subdomain that targets the globular module (144-390) of Fgy and (3) a putative site at the bottom of the N2 subdomain that targets the coiled coil site of Fg. This coiled coil site has also reported to be targeted by yet another microbial virulence protein 'M1' from group-A streptococci (GAS) (79-81). The interaction of M1/Fg enables the GAS to evade phagocytosis by host cell macrophages. Similarly, the interaction of ClfA at the coiled coil site of Fg might benefit *S. aureus* to evade phagocytosis by host macrophages. The interaction of ClfA at the trench site and globular chain of Fgy together act in concert to prevent Fg interactions with platelet integrin ( $\alpha$ IIB $\beta$ 3) and neutrophil integrins ( $\alpha$ M $\beta$ 2).

## CHAPTER III

### STRUCTURAL AND BIOCHEMICAL STUDIES ON MULTIPLE LIGAND BINDING

#### MECHANISM OF CLUMPING FACTOR-B of *Staphylococcus aureus*

##### **Introduction to Clumping factor-B of *S. aureus***

*S. aureus* adapts cwa proteins for colonization. ClfB plays a key role in establishing human nasal colonization by binding to the human type I cytokeratin 10 (CK10) expressed on squamous epithelial cells. ClfB is a colonization factor known to interact with multiple host proteins other than ck10 (Figure-28a-b). ClfB is known to interact with Fg, dermokine (7) and loricrin(47) in addition to ck10(45). The interaction of these molecules with ClfB is key for the colonization/pathogenesis of *S. aureus* (82).

ClfB binding to Fg is significant in platelet activation, aggregation, and has been shown to contribute to the pathogenesis of experimental endocarditis in rats (83). Immunization of mice with ClfB reduces nasal colonization (82).

##### **Introduction to host ligands of clumping factor-B of *S.aureus***

In keratinocytes, clfB binds to cytokeratin10 and Loricrin and in blood it binds to fibrinogen and causes aggregation or clumps. Based on the crystal structure of the cytokeratin10: clfB and fibrinogen:clfB complexes, there is a sequence recognition motif “GSSGXG” which is thought essential for clfB binding to its host protein ligands. A genome mining strategy was deployed for “GSSGXG” motif and furthermore potential ligands for clfB has been discovered like engrailed protein, TCF20 and Dermokine.

Loricrin (Lor), a major cell envelope protein of keratinocyte, is a 312 residue containing ~28kDa molecule rich in gly-ser residues (GSR) known to form interchain crosslinks with one another by isoglutamyl lysine isopeptide bonds (K89:Q154, K89:Q216, Q212:K312, Q213:K312) to create



(L2v-GST). Also they showed that a *lor*<sup>-/-</sup> mice colonized weakly and was cleared in day 3 after introduction of staph intranasally when compared to wild type mice inoculated with staph the colonization increased with no reduction of bacterial burden till day 10 post inoculation of *S. aureus* Furthermore, they introduced *clfB* in a *Lactococcus lactis* and tested for its ability to bind loricrin *in vitro* and able to recover *clfB* containing *Lactococcus lactis* from the anterior nares of the wild type mice but a 80% reduced colonization of *clfB* containing *Lactococcus lactis* in *lor*<sup>-/-</sup> mice(47).

### **Colonization is a risk factor for *S. aureus* infections**

Infections can be soft skin tissue infections to invasive infections like endocarditis, septicemia, arthritis and osteomyelitis. The reason for disease attributed to cell wall associated surface proteins and secreted toxins. Colonization is a risk factor for infection caused by *S.aureus*. *S.aureus* colonizes multiple sites in humans including perineum, axilla and anterior nares and *S.aureus* cwa proteins ability to adhere multiple host ECM molecules reason for the successful colonization in multiple tissue types. *S. aureus* infections are primarily autologous, which means *S. aureus* strains which colonize superficial sites like skin and nares eventually migrate to deeper tissues through open wounds, during surgery, dialysis, catheters etc and cause invasive infections like septicemia, endocarditis.

### **Mupirocin resistant *S. aureus***

Currently, mupirocin is the only drug recommended as a topical agent to reduce the nasal carriage of *S. aureus* and resistance against mupirocin and resistance-associated infections has been reported (84). Mupirocin works by blocking native tRNA synthetase (IleS) of *S.aureus* and a low (8-256µg/ml) and high level (>512µg/ml) of mupirocin resistance has been reported (85). Low level resistance is associated with a V288F mutation in isoleucyl-tRNA synthetase (*ileS*)

whereas high level resistance through acquisition of plasmid borne *mupA* gene that codes for eukaryotic like tRNA synthetase which bears no affinity for mupirocin(86).

Routine use of mupirocin, disease conditions like atopic dermatitis (AD) as well as immunosuppression increase risk factors for the development of mupirocin resistance strains. With the high prevalence of MRSA strains in both the community and hospitals, transfer of plasmid borne mupirocin resistance to these MRSA strains will be deleterious and complicates the treatment regimen.

Thus, colonization is a risk factor for *S. aureus* pathogenesis and alternative strategies to reduce nasal colonization is important.

### **Significance of the study**

Bacterial colonization is a risk factor for both infection and the development of drug resistance. New classes of anti-bacterial agents, like virulence blockers, that do not function as conventional antibiotics may be less likely to lead to antibiotic resistance (100). Agents that block the adhesion/colonization of the pathogen to the host are attractive options for the eradication of nasal carriage and merit investigation as an alternative prevention strategy. We envision these adhesion blockers to be preferentially used in topical application, as their likelihood of success would be higher than when used systemically. In topically-applied drugs, the half-lives of adhesion blocking peptides and small non-peptide molecules would be dramatically longer with fewer possible toxic side effects. Thus, this system provides an almost unique opportunity to test the concept and therapeutic value of adhesion blocking drugs. Agents that reduce the incidence of *S. aureus* infections in high-risk patients would substantially reduce infection rates and associated costs.

## **Hypothesis**

Based on previous work on ClfA/FgD structure, and the homology model studies with other Fg binding MSCRAMMS defined in this study, we have convincingly proven that the Fg binding MSCRAMMS ClfA, Fbl and SpsD share a common Fg binding site on top of their N3 subdomain. Unlike ClfA, Clumping factor-B, is shown to bind to repeat 5 (NSGSSGTGSTGNQ) of the flexible region of its  $\alpha$  chain, yet retains the ability to bind multiple ligands. ClfB is known to bind with multiple host ligands like CK10, Ln, DK. The ligand binding region of ClfB had been narrowed down to the trench site between N2 and N3 subdomains, which is partly similar to the Fg binding region of ClfA. We currently don't know whether ClfB also contains a similar ligand binding site on top of the N3 subdomain (N3 site) that was found in ClfA. Furthermore, the N2 and N3 subdomains of ClfA and ClfB share ~27% sequence identity between each other. Based on these observations, I hypothesize that ClfB, also bears a ligand binding region on top of the N3 subdomain and this binding region (N3-site) in ClfB is capable of interacting with the multiple ligands of ClfB (Fg and Ln).

## **Approach**

The crystal structure of ClfB bound to ligands such as Fg, CK10 or Ln, has not been reported in the literature. So far, a synthetic peptide mimicking the ligand binding region had only been crystallized in complex with ClfB and remains the basis of our knowledge on the ligand binding interactions of ClfB. Furthermore, currently we don't know the structural information on the interactions of ClfB with Ln.

Loricrin has multiple binding sites in ClfB. I will determine the high affinity binding region in Loop-2 of Ln to ClfB and crystallize this fragment in complex with ClfB (212-542). I will solve the crystal structure of this complex and identify the complete interaction sites in both ClfB(212-

542) and Ln(peptide). The details of the interactions could provide clues on the putative ligand binding N3-site of ClfB.

## **Innovation**

Merging my research goals with a combination of biochemical studies by ITC and structural studies of the complex by single particle X-ray crystallographic techniques, I will be able to determine the thermodynamics of the interaction as well as detailed structural data for the complex. In detail: ITC studies gives the affinity and nature of the interactions via their thermodynamic profiles and XRD gives the atomic level resolution details on the binding interface that could fit the biochemical (affinity and thermodynamic parameters) observations of the ClfB/Ln complex. This information can further be extended to the interactions of other ClfB ligands like Fg, CK10 and also to other Ln binding MSCRAMMs like SpsD.

A complete analysis of the above mentioned approach would help in designing peptides/ drugs with high affinity and inhibition properties against ClfB. These peptides could be later instrumental to develop therapeutics against Staphylococcal infections.

## **Materials and Methods**

### **Bacteria and Growth Conditions**

*Escherichia coli* strains XL-1 Blue and TOPP3 (Stratagene, La Jolla, CA) were used as hosts for cloning and expression, respectively. Cells were grown in LB broth containing 100 µg/ml ampicillin.

## **Cloning of ClfB and ClfB mutants by site directed mutagenesis**

DNA restriction enzymes were purchased from New England Biolabs and used according to the manufacturer's instructions. DNA segments encoding the N2N3 domain of ClfB (accession number (WP\_000745871.1) or its corresponding truncated versions were amplified by PCR using *S. aureus* Newman chromosomal DNA as template. Enzyme restriction sites BamHI and HindIII were incorporated at the 5' ends of the primers to facilitate cloning in pQE30 expression plasmid (Qiagen, Chatsworth, CA). Site-directed mutagenesis was performed using the pQE30 ClfB(212–542) plasmid as template. The introduction of specific mutations was performed using primers listed in Table-13. The resulting PCR mixtures were digested using DpnI for 1 h at 37 °C to eliminate parental DNA and transformed into *E. coli* XL-1 Blue competent cells. The presence of appropriate DNA sequences was confirmed by Sanger sequencing using specific forward and reverse pQE30 sequencing primers at the Baylor College of Medicine DNA Core Facility, Houston, TX.

## **Expression and Purification of Recombinant Proteins**

Expression and purification of recombinant MSCRAMM proteins were performed as described previously for SdrC (87). Briefly, recombinant proteins were expressed from appropriate plasmids in *E. coli* TOPP3. Overnight starter cultures were diluted 1:50 in LB-containing ampicillin (100 µg/ml) and incubated with shaking until the culture reached  $A_{600}$  0.6–0.8. Recombinant protein expression was induced by addition of 0.1 mM isopropyl 1-thio- $\beta$ -D-galactopyranoside (final concentration) and continued during a 4-h incubation. Bacterial cells were harvested by centrifugation, resuspended in PBS, and frozen at –80 °C. Recombinant proteins were purified from cell lysates by Ni<sup>2+</sup> affinity and ion exchange chromatography.

Table-13. List of primers for site directed mutagenesis of ClfB

Mutant		NUCLEOTIDE SEQUENCE
ClfB <sup>N238A</sup>	Forward	5'- aca ttt atg gcg gca gct ttt aca gtg aca gat -3'
	Reverse	5'- atc tgt cac tgt aaa agc tgc cgc cat aaa tgt -3'
ClfB <sup>S384V</sup>	Forward	5'- ggt gta gat aca gct gta ggt caa aac aca tac -3'
	Reverse	5'- gta tgt gtt ttg acc tac agc tgt atc tac acc -3'
ClfB <sup>E227Ile/R529Ser</sup>	Forward-1	5'- tta ata aag act aca ttt gac cct aat caa agt -3'
	Reverse-2	5'- act ttg att agg gtc aaa tgt agt ctt tat taa -3'
	Forward-3	5'- gta agt tat ggt ggt gga agt gct gat ggt gat -3'
	Reverse-4	5'- atc acc atc agc act tcc acc acc ata act tac -3'
ClfB <sup>I505A/E507A</sup>	Forward	5'- aaa act cag gtt gct caa gca aat gtt gat cct -3'
	Reverse	5'- agg atc aac att tgc ttg agc aac ctg agt ttt -3'
ClfB <sup>W407A/Y409A</sup>	Forward	5'- tta ggt aat acg gcg gtg gct att aaa ggc tac -3'
	Reverse	5'- gta cgg ttt aat agc cac cgc cg tatt acc taa -3'

### Isothermal Titration Calorimetry Assay

ClfBN2N3 or its mutants were titrated against full length human Fg in the cylinder with a molar ratio of 1 Fg monomer: 30 rClfBN2N3 with TBS as buffer of choice. For the Peptide studies, the respective peptides (Lor peptides, CG12 and CG14; K10 peptides, YH15, GY26 and YY41; Fg peptide, GQ17) were titrated against ClfB(212-542) with a molar ratio of 1 ClfB(212-542):25-50 peptide with TBS. Heat of a reaction gives the enthalpy endothermic peaks were recorded. Repeated titrations of ligand extended until all the molecules in the cell gets saturated with the ligand in excess. The RAWITC plotted with molar ratio ( $\mu\text{M}$ ) in the X-axis and heat of reaction

( $\mu\text{Cal/sec}$ ) in the reaction to give a  $\Delta H$  plot and the curve fitted for one site-binding in MicroCal ITC. The resulting values for no of binding sites from the molar ratio (N value), Affinities of ligand (KdA) and entropy of the reaction ( $\Delta S$ ) were recorded for all experiments.

### **Solid Phase Assay**

Immulon 4BH plates were coated with 1  $\mu\text{g}$  of Fg (obtained from ChromogenixAB) or recombinant K10 (produced as described previously (45)). Coated wells were blocked for 1 h at room temperature with 3% BSA in 0.5% Tween/TBS buffer. Increasing concentrations of ClfB recombinant proteins were added to the wells and incubated for 1 h at room temperature. The wells were washed, and bound ClfB protein was detected with a polyclonal antibody (1:3000) against ClfB<sub>N2N3</sub> (H.T.I. BioProducts) followed by an anti-rabbit HRP-labeled antibody (1:5000) (Bio-Rad).

HRP-dependent color development using Sigma*Fast* OPD (Sigma) as a substrate was monitored using a microtiter plate reader (Molecular Devices) at 450 nm. Data presented represent the mean  $\pm$  S.D. of three independent experiments performed in triplicate. The binding was analyzed by nonlinear regression for one binding site (GraphPad Prism). The relative binding was calculated as percent of the wild type protein binding which was considered 100%.

### **Surface Plasmon Resonance Spectroscopy**

Surface plasmon resonance-based binding experiments were performed at 25 °C on a Biacore 3000 (GE Healthcare/Biacore, Uppsala, Sweden). Phosphate buffered saline (PBS-T: 8.06 mM Na<sub>2</sub>HPO<sub>4</sub> and 1.94 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.4), 2.7 mM KCl, 137 mM NaCl, and 0.005% Tween-20) was used as running buffer for immobilization and binding experiments.

A flow rate of 5  $\mu\text{l}/\text{min}$  was used during immobilization and a higher rate of 30 or 50  $\mu\text{l}/\text{min}$  for binding experiments. The sensor surfaces of Fg, GST-Fg $\alpha$ (316-367), GST-Lor(152-230), GST-HK10(544-563) were prepared on different sensor chips (CM5 for high density, CM3 for medium density and C1 for low density ligand surfaces). The ligands were covalently coupled to the chips using standard amine-coupling chemistry. Frozen Fg stock (about 10 mg/ml in 20 mM sodium citrate-HCl, pH 7.4) were thawed in a 37 °C water bath without any agitation. After equilibrating to room temperature, Fg was diluted in 10 mM sodium acetate (pH 5.5) to 10 or 20  $\mu\text{g}/\text{ml}$  and injected into an EDC/NHS activated flow cell. The surface was deactivated with ethanolamine. A reference surface was made with activation and deactivation steps but with no protein coupled. For capturing the GST fusion protein, approximately 11,000 RU of goat anti-GST antibody (GE Healthcare/Biacore) was immobilized on a CM5 chip. The GST-tagged proteins was captured by the antibody and created a GST-protein ligand surface. Another flow cell with immobilized anti-GST antibody and captured GST was used a reference surface. To regenerate the ligand surfaces, bound proteins were removed by a 1 min injection of 1 M NaCl for the Fg surface, and 0.01% SDS for the GST-protein ligand surface.

### **Peptide Inhibition Studies**

For the peptide inhibition studies, the peptides were dissolved in the binding buffer (1XPBS) and about 500 $\mu\text{M}$  of each peptide incubated with ClfB before binding to the ligand and then the binding response was recorded. The difference in the normalized response of rClfB (212-542) in resonance units (RU) with and without the peptide towards each ligand was recorded and reported.

For the rClfB-N3 mutant binding studies with Fg, the rClfB(212-542), rClfB<sup>W407A,Y409A</sup> or rClfB<sup>I505A,E507A</sup> was captured using mouse anti-HIS antibody (GE Healthcare/Biacore) was immobilized on a CM5 chip. The surface was deactivated by ethanolamine. About 1.6 $\mu$ M of human Fg was allowed to bind to the surface and the response (RU) was recorded. The binding response of ClfB mutants was compared with rClfB(212-542) and reported.

All SPR responses were baseline corrected by subtracting the response generated from the corresponding reference surface. Double-referenced SPR response curves (with the buffer blank run further subtracted) were used for affinity determination. For steady-state interaction, the equilibrium response of each injection was collected and plotted against the concentration of injected protein. A one-site binding (hyperbola) model was fitted to the data (GraphPad Prism 4, GraphPad Software, Inc., La Jolla, CA, USA) to obtain the equilibrium dissociation constant  $K_D$ . Non-equilibrium data were globally fitted to a 1:1 Langmuir binding model using BIAevaluation software (Version 4.1). Association and dissociation rate constants  $k_{a1}$  and  $k_d$  were obtained from the fitting, and the dissociation constant  $K_D$  was derived ( $K_D = k_d / k_a$ ). Errors are reported as standard error of mean from two or more experiments.

### **Gel Permeation Chromatography**

The peptide is incubated with the ClfB for an hour in 1XTBS, pH 8.0 and the solution was injected into the sample loop of the HPLC machine. The complex was allowed to flow through a Sephadex S75 (16/60) gel permeation column which was preequilibrated with 1XTBS, pH 8.0. The Apo-ClfB in 1XTBS, pH8.0 also subjected to S75 column and the elution volume of the proteins compared to differentiate the ClfB/peptide complex from the Apo-ClfB using Native PAGE technique.

## **Native PAGE**

Native gel electrophoresis is used to assess the protein or protein complex in its native state, a monomer, a dimer or oligomers. Novex Bis-Tris gel system were purchased from Invitrogen Corporation (Carlsbad, CA). Pre-cast NativePAGE™ Novex 4–16% (v/v) Bis-Tris gels were run with near neutral pH at 90 V at 4 °C with stirring for 4.5 h. Protein samples (10 μL) were mixed with the sample buffer provided (2.5 μL) and 5% (w/v) Coomassie blue G-250 (0.3 μL). Gels were destained using simply blue safe stain (Invitrogen) for 2-3 h and destained in water overnight.

## **X-ray Crystallography**

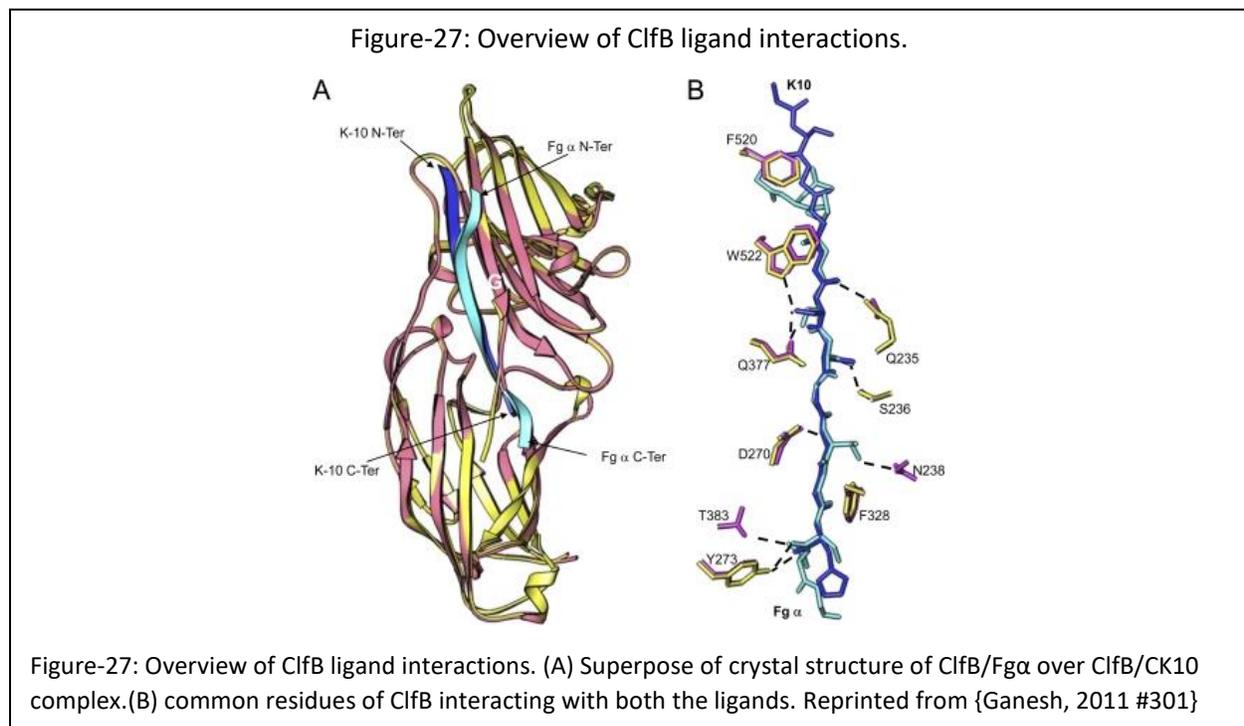
ClfB·(CG12) complex was prepared by mixing 36mg/ml of ClfB with (CG12)<sub>2</sub> to a final ClfB:CG12)2 molar ratio of 1:50. Crystallization attempts were made with co-crystallization ClfB:peptide using sitting drop vapor diffusion method. Diffraction quality crystals were obtained by mixing 2micro L of ClfB with 2 microL of reservoir solution containing PEG...Similar to ClfB:CG12)2, crystal of ClfB:(CG14)<sub>2</sub> was obtained under similar conditions.

Crystals were flash frozen using paratone-N as cryo protectant. X-ray diffraction data was collected in sector 23, APS, Chicago. The ClfB:CG12 and ClfB:CG14 complex crystals diffracted to about 2.5 and 2.1 Å respectively. A total of 180 degrees were collected with oscillation width of 1degree was collected for both ClfB:CG12 and ClfB:CG14. Diffraction data were processed using HKL2000. Both the complex structures were obtained by molecular replacement methods using N2N3 domains of ClfB structure (3ASW, 3ATO) as search model using PHASER. The ClfB:CG12 and ClfB:cg14 were refined to a final R factor of 19.4% and 23.3% respectively and a final Rfree of 27.4% and 28.5% respectively.

## Results

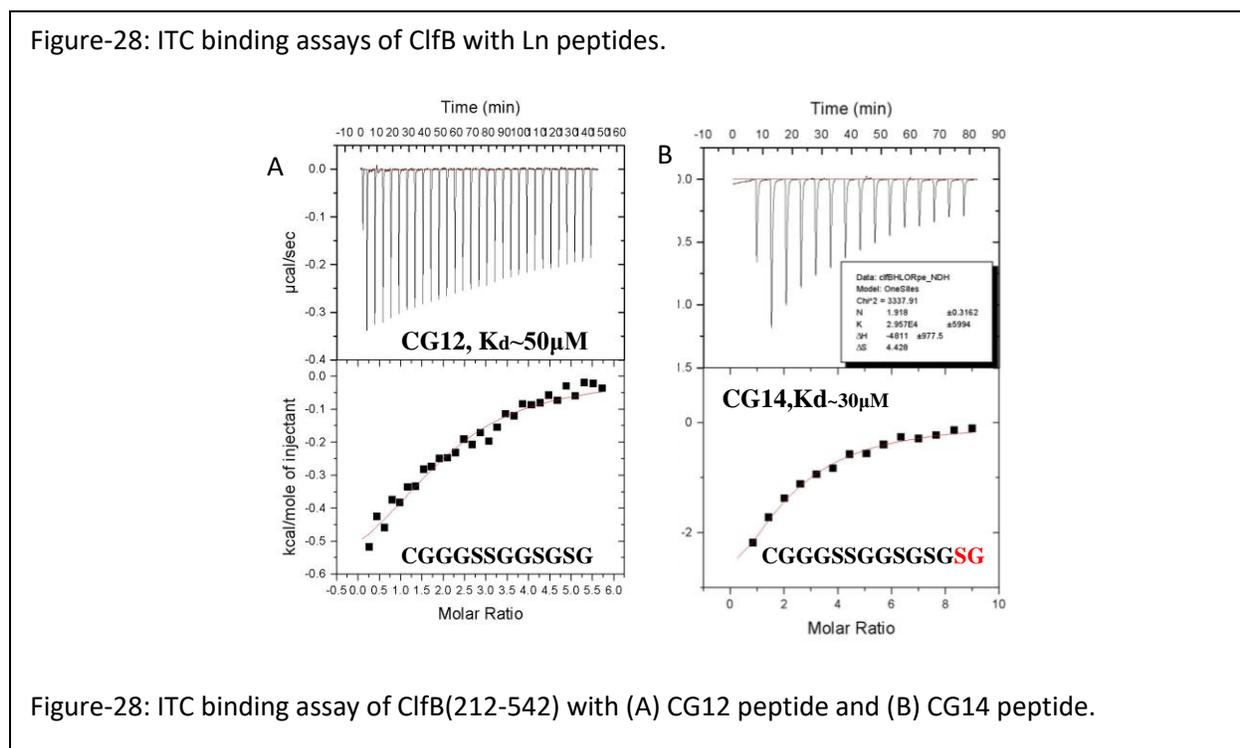
### (CG12)<sub>2</sub> peptide mimics the loop region of Loricrin

Earlier studies have shown that loricrin binds to ClfB with higher affinity than the rest of its other ligands. The trench site of ClfB interacts with ligands with a conserved ‘GSSGXG’ or a ‘GSSGXXG’ motif, which is found in the ClfB binding regions of  $\alpha$ -chain of Fg, CK10, DK. The structure of these complexes are superposed and shown in Figure-27. The ligand peptides dock at the trench site and ClfB interacts with these ligands through same residues although there are few exceptions (discussed in later sections).



Loricrin is a molecule rich in cysteines that interact with each other at the inter and intramolecular level. This level of interaction leads to the formation of disulfide bridges which subsequently lead

to the formation of loop regions (Figure-26a). For this reason, I chose a 12 mer sequence ‘CGGGSSGGSG’ which is naturally found in the loop-2 region of Ln. The presence of ‘Cys’ at the N-ter is crucial where it makes a disulfide bond with ‘Cys’ of another peptide, thus a disulfide bridge linked dimeric peptide (CG12)<sub>2</sub> mimicking the loop region of Ln. This (CG12)<sub>2</sub> peptide displayed high affinity values with ClfB(212-542) in my ITC binding studies (Figure-28a). I have also tested the affinity of ClfB(212-542) with a 14 mer sequence ‘CGGGSSGGSGSG’ (CG14)<sub>2</sub>, which is again a disulfide linked dimeric peptide of Ln. This peptide binds to ClfB with an even higher affinity than the (CG12)<sub>2</sub> (Figure-28b).



This proves that (1) the interaction of ClfB at the N-terminus of the peptide (where ‘Cys’ disulfide is located) doesn’t get affected due to increase in the length of the ligand peptide and (2) ClfB could potentially interact with a longer peptides containing the binding sites for that region of the ligand. This is important finding because earlier from the crystal structure of the peptide ligand,

the average length of the Fg $\alpha$ - chain peptide and CK10 peptide is ~12aa of which only 8-10 aa only showed positive electron density. Moreover, the affinity of both these peptides (CG12 dimer and CG14 dimer) is better than the affinities of ClfB with other ligand peptides of Fg and CK10 (Table-14) indicating that the C-terminus of these peptides make additional contacts with ClfB. To summarize, based on my ITC binding studies of ClfB(212-542) with (CG12)<sub>2</sub> peptide and comparison of affinities it is clear that (CG12)<sub>2</sub> peptide is a high affinity ligand of ClfB that mimics closely the loop-2 region of full length Ln.

### **YY41 peptide binding studies with ClfB indicate additional binding site in ClfB**

I have tested the binding affinity of three cytokeratin-10 peptides of different length with ClfB(212-542). The first peptide (YY41) is a 41mer of sequence (Figure-29d) comprising the tail region (474-514) of CK10. This peptide represents the Y-Y loop of CK10 where the presence of an aromatic 'Tyr' residue forms a loop with flexible gly-ser sequences in between them. The second peptide GY26 (489-514) represents the last 26 residues of the 41mer peptide. This peptide has a flexible gly on one end at the N-terminus and a tyr at the other end. The third peptide (YH15) comprising region (474-488) represents the first 15 residues from the N-ter of the 41mer peptide. Of these peptides, The YY41 peptide displayed highest affinity to ClfB(212-542) followed by GY26 and YH15 peptide. The binding stoichiometry of the YY41 peptide binding to ClfB(212-542) is 0.2:1 which indicates the involvement of more than one binding site in this interaction. The GY26 peptide has a binding ratio of 1.3:1, which is ideal molar ratio for a 1:1 complex. This peptide also has a length of 26 aa residues which is close to the length of a (CG12)<sub>2</sub> dimer peptide. I hypothesize that the peptide GY26 and CG12 dimeric peptide binds to both sites of ClfB (previously known N2N3-trench site and the putative N3 site). The YH15 peptide displayed the

lowest affinity compared to all of the peptides tested via ITC interactions to ClfB(212-542) possibly due to the lack of the interaction at the N3-site of ClfB (Figure-29).

### Interaction of Fg with ClfB displays tight binding than the other ligands of ClfB

The ClfB(212-542) binding region in Fg is the Fg $\alpha$ (316-367) which has been previously reported (7, 46). If this segment in Fg contains the only possible interacting binding site region to ClfB then the binding affinities of ClfB to intact Fg and Fg $\alpha$ (316-367) should be similar or of not much different from one to each other. To compare these binding affinities, I performed the ClfB binding assays to intact Fg and a recombinant GST-Fg $\alpha$ (316-367) by SPR and ITC techniques. SPR technique showed a modest affinity of  $\sim 1\mu\text{M}$  to intact Fg whereas for GST- Fg $\alpha$ (316-367) the affinity is  $5.25\mu\text{M}$  which is atleast 5 fold weaker than the intact Fg. The alpha chain peptide ‘GQ17’ which represents the Fg $\alpha$ -chain binding of region (as reported in crystal structure of ClfB:Fg $\alpha$  complex) affinity cannot be detected through ITC as it is very weak ( $>100\mu\text{M}$ ) and out of range of the detection limit through ITC or SPR (Figure-30). A summary of the various ligand binding affinities to the ClfB(212-542) are listed in Table-14.

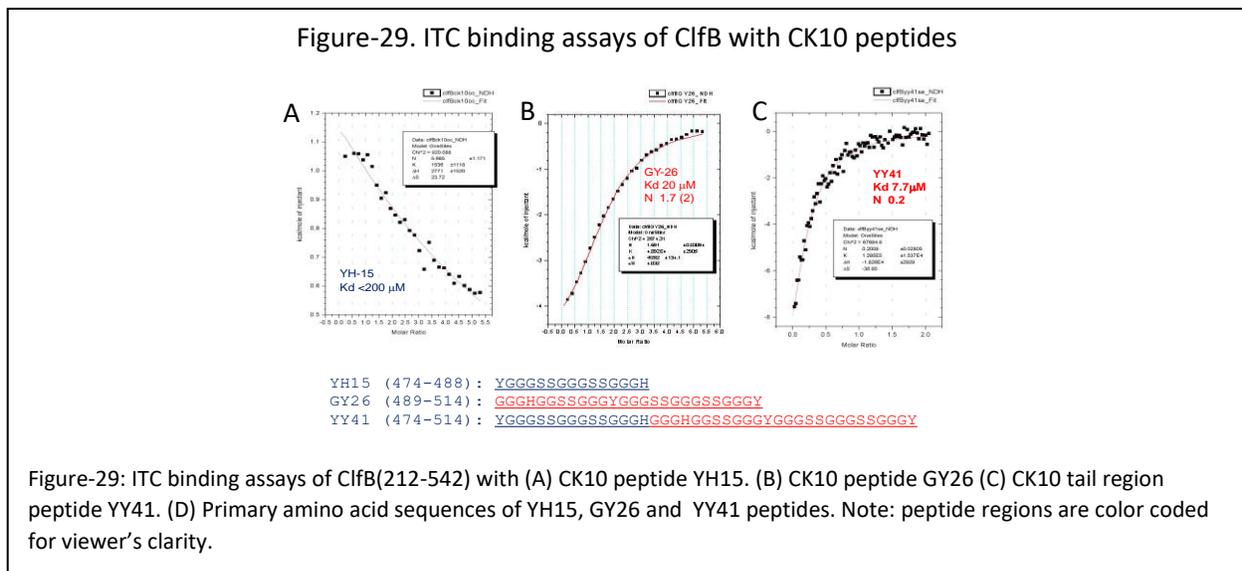


Figure-30. ITC binding assays of ClfB with Fg

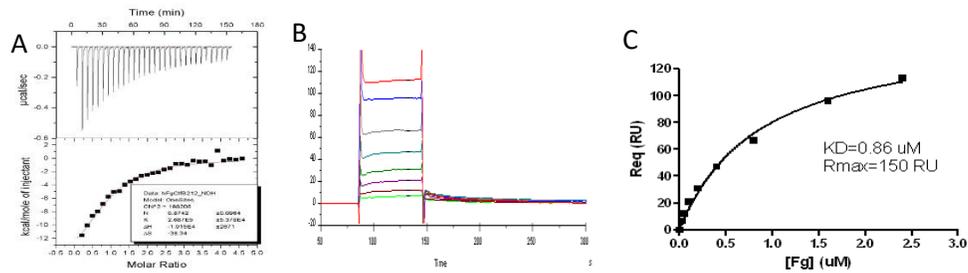


Figure-30: Fibrinogen binding assays of ClfB. (A) Fg binding assay of ClfB(212-542) using ITC technique. (B) immobilized ClfB(212-542) on a chip and intact Fg flown over the chip to detect the binding by SPR technique. (C) response units of ClfB(212-542) in Y-axis over increasing concentration of Fg( $\mu\text{M}$ ) in X-axis. The binding saturation with a  $K_D$  of  $0.86\mu\text{M}$ , which is similar to the  $K_D$  measurements of ClfB binding to intact Fg by ITC technique shown in Fig(30a).

Table-14. Summary of ligand affinities towards ClfB(212-542)

Ligand	Fragment	Affinity (kD)	
		SPR	ITC
Human Loricrin	CG12 (12 mer)	NA	~50 $\mu\text{M}$ (ITC) (unpublished)
	GST-Loop2 (152-230)	3.31 $\pm$ 0.81 $\mu\text{M}$	NA
	human loricrin	4.33 $\pm$ 1.10 $\mu\text{M}$	NA
Human Cytokeratin 10	GY-14	NA	> 10 <sup>-4</sup> M
	GY-26	NA	~20 $\mu\text{M}$
	YY-41	NA	~ 7 $\mu\text{M}$
	GST-HK10 (544-563)	7.89 $\pm$ 2.10 $\mu\text{M}$	working
Human Fibrinogen	Fg- $\alpha$ peptide	NA	~ 10 <sup>-4</sup> $\mu\text{M}$
	GST-Fg $\alpha$ (316-367)	5.25 $\pm$ 1.5 $\mu\text{M}$	NA
	Intact Fibrinogen	~1 $\mu\text{M}$	10 $\mu\text{M}$

### YY41 and (CG14)<sub>2</sub> peptide inhibits ClfB across multiple ligands

Ligands showing strongest affinity to its cognate receptors need not be the strongest inhibitors as well. For example, in our studies, the fg inhibiting monoclonal antibody ‘aurexis’ which binds to ClfA at the N3 subdomain shows nanomolar affinity towards ClfA but is a weak inhibitor. The weak inhibition is due to partial binding of aurexis to ClfA which inhibits fg

interaction to ClfA. Similarly, the ligand peptides which showed high binding affinity need not be a potent inhibitors as these peptides can bind at different sites in ClfB. Utilizing SPR techniques, I investigated the inhibiting potential of these ligand peptides to inhibit the ClfB binding towards its full length ligands. The results are shown below (Figure-31).

From the peptide inhibition studies, we can conclude that the inhibition potential is proportionate to the length of the peptide except for the 'GQ17' Fg $\alpha$  peptide which exhibited weak inhibition to Ln, Fg and Ck10. The YY41 peptide, which is the longest peptide (refer to table), exhibited a 80-~90% $\pm$ 10% inhibition of ClfB binding towards all three ligands (Ln, Fg and CK10). The Ln peptide CG14-p13S, showed inhibition but differs in the level of inhibition against different ligands with the greatest inhibition of ~80%  $\pm$ 10% against fg and CK10 but surprisingly average inhibition towards Ln. The Ln peptide CG14 and the CK10 peptide 'YH15' (of equivalent length, see table-14) displayed a trend similar to CG14-p13S but with comparatively lower levels of inhibition across different ligands. The Ln peptide 'CG12' did show inhibition against all the ligands with relatively low affinity (put data) when compared to the longer peptides but better than the  $\alpha$ -peptide 'GQ17'. The YG26 surprisingly didn't have much inhibition although it binds with high affinity to ClfB. YG26, which is a 26 residue fragment inhibits poorly (Figure-31).

In summary, based on the peptide inhibition it seems like YY41 and CG14-p13S inhibit ClfB binding with all its three different ligands making these peptides optimal inhibitors.

### **Crystal structure of ClfB in complex with (CG12)<sub>2</sub> and (CG14)<sub>2</sub> peptides of Ln**

The longer peptide YY41 and CG14 has the modest ligand inhibition and binding affinity towards ClfB. This observation led me to propose that both of these peptides bind to the N3 site of ClfB (in addition to binding at the trench site). To prove this hypothesis, I did crystallization trials on ClfB (212-542) in complex with either YY41, CG14 and CG12. ClfB·(CG12) complex was

prepared by mixing 36mg/ml of ClfB with (CG12)<sub>2</sub> to a final ClfB:(CG12)<sub>2</sub> molar ratio of 1:50. Crystallization attempts were made with co-crystallization ClfB:CG12 peptide using sitting drop vapor diffusion method. Diffraction quality crystals were obtained by mixing 2 $\mu$ l of ClfB with 2 $\mu$ l of reservoir solution to get high quality diffracting crystals. Similar to ClfB:(CG12)<sub>2</sub>, crystal of ClfB:(CG14)<sub>2</sub> was obtained under similar conditions (Figure-32).

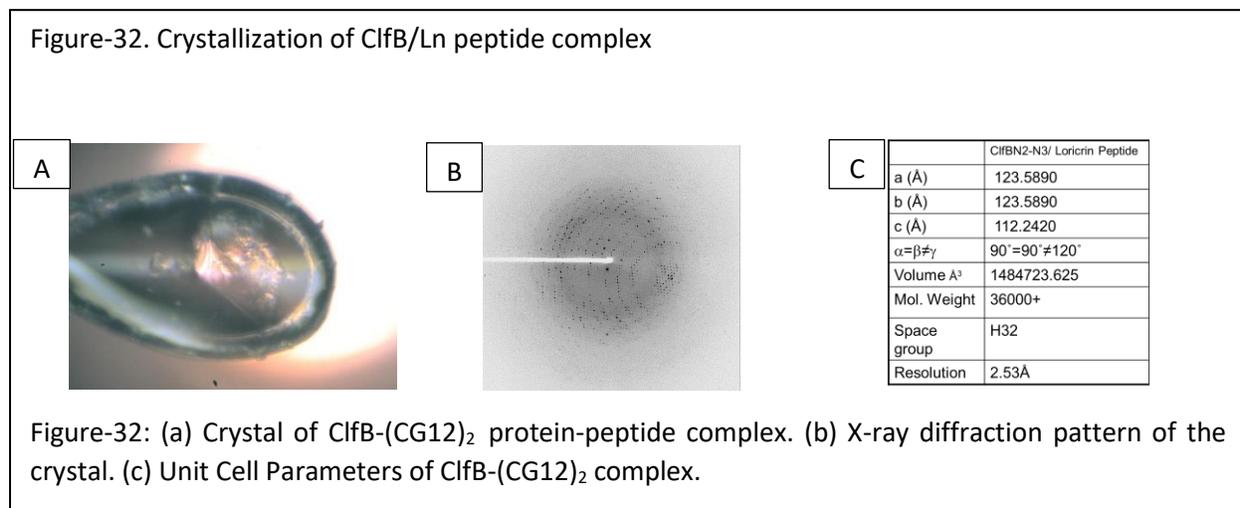
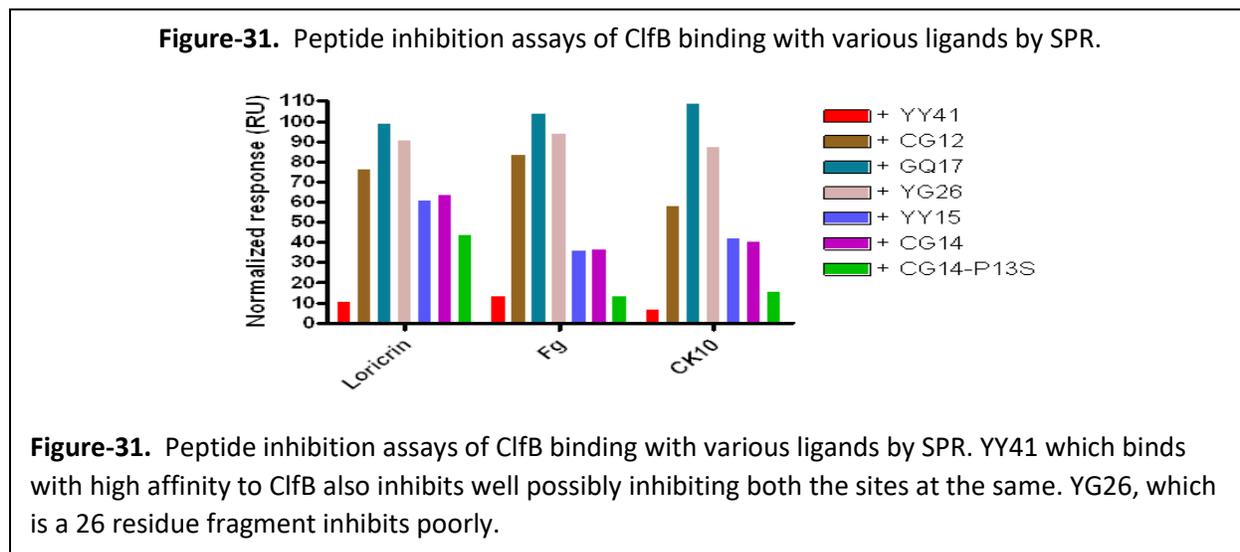
Unfortunately, crystallization of YY41 in complex with ClfB was not successful and the reason for this might be due to: (1) multiple ClfB molecules bind to single YY41 peptide that might interfere with the crystal packing or (2) the flexibility of the unbound region in the YY41 peptide interferes the crystal lattice formation.

### **Data Determination**

Crystals were flash frozen using paratone-N as cryo protectant. X-ray diffraction data was collected in sector 23, APS, Chicago. The ClfB:CG12 and ClfB:CG14 complex crystals diffracted to about 2.5 $\text{\AA}$  and 2.1 $\text{\AA}$  respectively. A total of 180 $^\circ$  degrees were collected with oscillation width of 1 $^\circ$  degree was collected for both ClfB:CG12 and ClfB:CG14. Diffraction data were processed using HKL2000. Crystal parameters and refinement statistics are given in Table 15.

Both the complex structures were obtained by molecular replacement methods using N2N3 domains of ClfB structures (7, 46) as search model using PHASER-MR. Electron density for the CG12 peptide was clearly visible in the initial model of the peptide. CG12 peptide was docked into the density by superposition method. Side chains for the Ser and Cys were then modeled to the density. The CG12 complex was used for the molecular replacement solution for CG12. In addition to the ligand peptide density between the N2N3 trench additional densities was observed at the top of the N2 domain in both the crystals structures. N-terminal Cys of the Ln peptide CG12 and up to three Glycine residues of the CG12 peptide could be modeled into the additional density

at the N3 site. The ClfB:CG12 and ClfB:cg14 were refined to a final  $R_{work}/R_{free}$  of 19.4/27.4 and 23.3/28.5 respectively (Table-15).



### Structural analysis of ClfB-Ln complex

The crystal structures of both ClfB:peptide complexes (CG12 and CG14) occupied exact binding regions in ClfB with no significant structural deviations in ClfB. The electron density at

the N2N3 trench is mapped to the peptide CG12 and CG14 and all the main and side chain atoms are fitted into electron density. Furthermore, additional electron density on top of the N3 subdomain of ClfB was observed, which was mapped to the first 4-5 amino acid residues from the N-terminus of peptide (CGGG). Most importantly, there is a continuous strong electron density past the side chain (-SH) of the N-terminal Cysteine, which indicates the disulfide bond of the cysteine of one peptide with the neighboring peptide (symmetry mate) of ClfB:peptide complex (Figure-33).

From the observations of crystal structure of ClfB/(CG12)<sub>2</sub> and ClfB/(CG14)<sub>2</sub>, the presence of 'cys' at the N-ter of the peptide mimics the loop region of Ln which allows the conformational flexibility to bring two molecules of ClfB in close proximity to each other in a head to head fashion. This ligand induced dimerization of ClfB was further stabilized by strong electrostatic interactions (salt bridge) between the residues R529-E227 of the neighboring ClfB molecules. The R529 is located at the latching strand of the N2 subdomain upon ligand binding. In the absence of ligand binding interactions, the R529 will be at the N3 subdomain as seen in the Apo-ClfB structure.

Thus, the Ln peptide (CG12)<sub>2</sub> binding to ClfB creates the following events (1) the N-ter 'cys' disulfide bonded peptides forces a ligand induced dimer formation of ClfB in a head to head fashion (2) re-orientation of the latching strand due to the DLL mechanism of binding at the trench site and (3) the re-orientation of the latching strand positions the R529 of the latching strand close to E227 of the dimer partner ClfB forming a salt bridge. Thus the ligand induced dimer complex of ClfB is stabilized (Figure-34).

Figure-33. Crystal Structure of ClfB/Ln peptide complex.

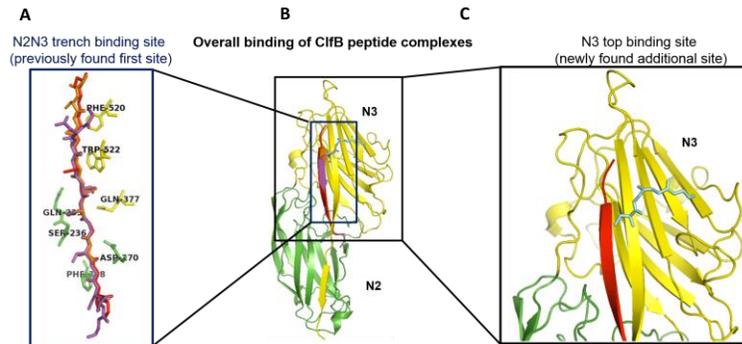


Figure-33: Crystal structure of the ClfB-CG12 peptide complex (asymmetric unit). (The cg12 peptide shown in red, overlaid with other ClfB ligands FG-alpha(purple and CK10(orange)). The CG12 interacting residues (side chains) of ClfB are shown in green (N2) and yellow (N3). (b) ClfB-CG12 complex superposed with CK10 (3ATO) and Fg-alpha(3ASW). The CG12 peptide that partially (first four residues) occupy the top of the N3 subdomain is shown in cyan. (c) Zoom in of the N3 subdomain of ClfB showing the second ligand binding site in ClfB.

### Determination of ligand induced dimerization of ClfB by Ln peptide (CG12)<sub>2</sub>

To isolate the ligand induced dimer of a ClfB:(CG12)<sub>2</sub> complex and to show experimental evidence of the complex formation in physiological (non-crystallizing) conditions, I tried to purify the complex by gel permeation chromatographic technique. The Apo-ClfB(212-542) of 36mg/ml, was also run in the column with similar concentrations and conditions used for purifying the ClfB:peptide complex. The elution volume of the peaks of Apo-ClfB and ClfB:(CG12)<sub>2</sub> peptide was compared and used as an indicator of the dimer complex. The peak of the Apo-ClfB elution volume was ~95ml, whereas the ClfB in the presence of the (CG12)<sub>2</sub> peptide eluted at a volume of ~85ml, which was clearly an early elution profile indicating the higher molecular weight, possibly the ClfB dimer with the bound (CG12)<sub>2</sub> peptide (Figure-35a).

I have tried to visually prove the dimer by running a native PAGE of the complex but unfortunately the ClfB:CG12 peptide complex ran around ~36kDa, which was similar MW of the monomeric ClfB. The reason for the separation of the dimer during gel electrophoresis is not clear but may be

due to the weaker affinity of the peptide (30 $\mu$ M) which is not strong enough to keep the dimer stable in the running conditions of even a native PAGE (Figure-35b).

Table-15. Crystallization data refinement and statistics of ClfB:CG12 and ClfB:CG14 peptide complexes.

	ClfB/CG12	ClfB/CG14
Cell dimensions		
a,b,c (Å)	123.58,123.58,112.24	124.39,124.39,112.20
$\alpha,\beta,\gamma$ (°)	90.0,90.0,120.0	90.0,90.0,120.0
Space group	R32	R32
Max Resolution (Å)	2.54	2.15
Reflections unique	10351	17173
Completeness (%)	98.73	98.74
$R_{merge}$	4.9	5.1
No. of molecules in the ASU	1	1
$R_{work}/R_{free}$	19.4/27.4	23.3/28.5
Average B value (Å <sup>2</sup> )	49.18	33.59
No of non-hydrogen atoms	2574	2710
MSCRAMM	2200	2429
Loricrin	201	235
Water	120	130
RMSD from Ideal Values		
Bond lengths (Å)	0.011	0.014
Bond Angles (°)	1.492	1.674
* $R_{merge} = \sum  I_j - \langle I \rangle  / \sum I_j$ ; where $I_j$ is the measured and $\langle I \rangle$ is the mean intensity of reflection hkl		
+ $R_{free}$ is calculated over 5% of randomly selected reflections not included in the refinement.		

Figure-34. Ligand induced dimerization of ClfB:CG14 peptide complex

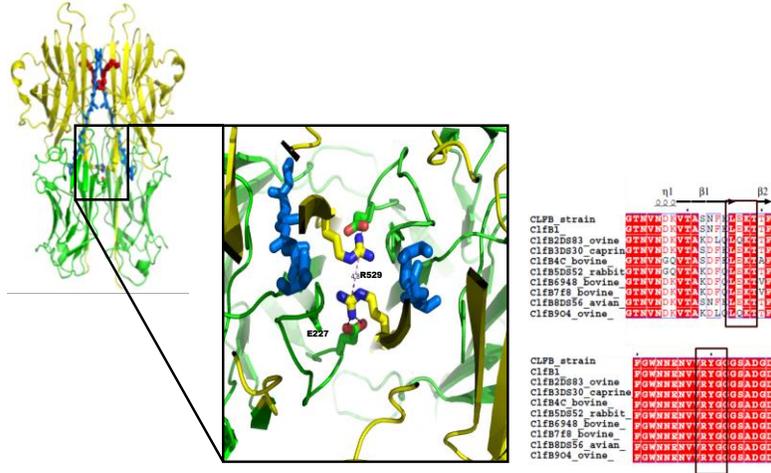


Figure-34: Ligand induced dimerization of ClfB. (A) 2 molecules of ClfB [green (N2) and yellow (N3)] brought together by the disulfide bonded CG12 ligand peptides (blue sticks). The salt bridge interactions of ClfB dimers are zoomed in with R529-E227 side chains involved in salt bridge interactions. (B) Sequence conservation of

Figure-35. Determination of a Ligand induced dimerization of ClfB

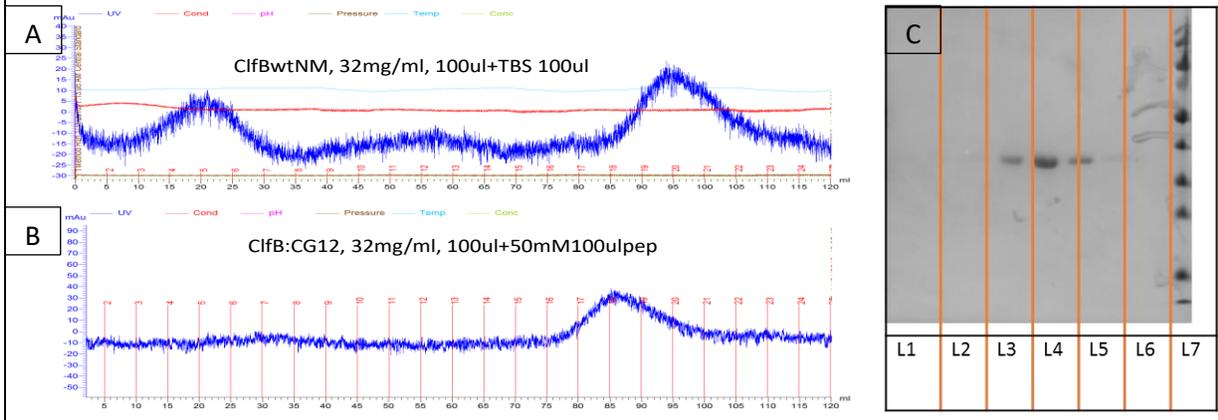


Figure-35. Determination of a Ligand induced dimerization of ClfB by Gel Permeation Chromatography. (A) Apo-ClfB (B) ClfB in the presence of (CG12)<sub>2</sub>. (C) SDS-PAGE of SEC fractions. L1: Apo-ClfB (80-85ml fraction17), L2: Apo-ClfB(85-90ml fraction 18), L3: Apo-ClfB (90-100ml pooled fraction-19-20), L4: ClfB-CG12 complex (80-85ml fraction-17), L5: ClfB-CG12 complex (85-90ml fraction-18). L6: empty, L7: Protein Ladder. Although the ClfB in complex elutes earlier than apo-ClfB, in the presence of native page (no SDS), ClfB in both the apo and complex runs similar MW of ~36kDa.

### Discussion of Chapter III

The ligand-binding N-terminal A-region of ClfB consists of three sub-domains, N1, N2 and N3 where the N2N3 sub-domains are sufficient for ligand binding. We determined the crystal

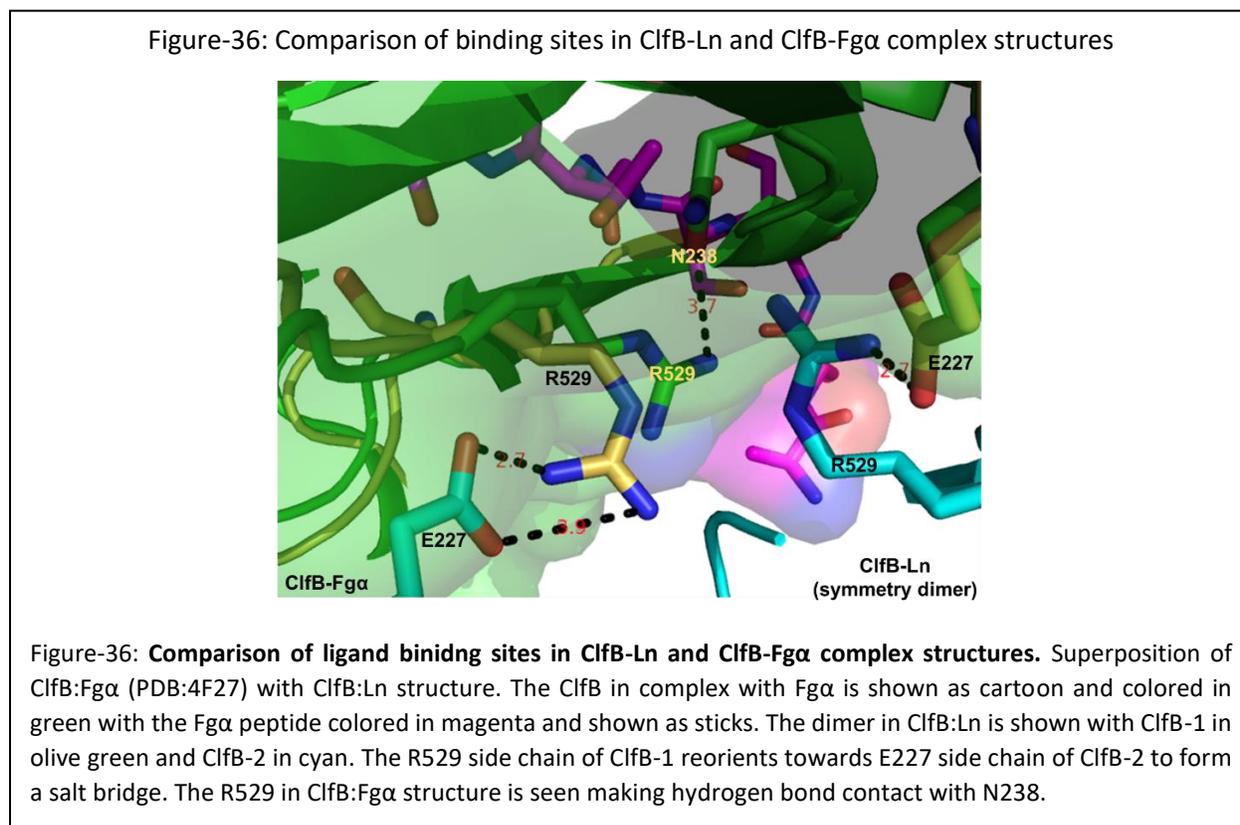
structures of the ligand binding domains of ClfB N2N3<sub>(212-542)</sub>, (ClfB) in complex with synthetic peptides corresponding to the Fg (Fg) and CK10 binding motifs at 2.5 Å and 2.6 Å resolution, respectively (88). Superimpositions of several ClfB:ligand complexes with several ligand molecules are shown in Figure-27. These peptides dock in a trench formed between the N2 and N3 sub-domains (N2N3 trench) by making a β-sheet complementation with the N3 domain and partially interact with the N2 domain, similar to the interactions of ligand peptides shown for other staphylococcal MSCRAMMs (9, 89). A short common sequence motif, GSSGXG, present in both peptide ligands was shown to engage the same residues within ClfB. This motif is found in numerous proteins, indicating that ClfB may recognize additional ligands, including dermokine, plakophilin and Ln, which are potential ligands for bacterial adhesins within the nares (88). Later, in support of our hypothesis, ClfB N2N3 was shown to bind Ln (90) and a peptide representing the putative binding site in dermokine (91). The crystal structure of the dermokine peptide/MSCRAMM complex was solved (91). However, preliminary data presented in this thesis demonstrate that the GSSGXG containing ligand peptides show a much lower affinity for ClfBN2N3 compared to that exhibited by intact ligand sub-domains suggesting that this interaction is more complex.

In our study we have used a 24 residue loricrin peptide with two 12 amino acid residue binding sequence linked by disulfide bridge (CG12)<sub>2</sub> and 28 aa residues loricrin peptide (CG14)<sub>2</sub> analog with 14 aa peptide ‘CGGGSSGGSGSGSG’ linked together by disulfide bridge. The rClfB binding affinities of the loricrin derived peptide were compared with cytokeratin derived peptide and Fg alpha chain derived peptides. Biochemical protein-peptide binding experiments of rClfB and lor peptides carried out using ITC assay. Peptide inhibition of rClfB and ligand binding

demonstrated by SPR assays using an immobilized ligand with increasing concentration of rClfB in the presence and absence of the peptides were tested.

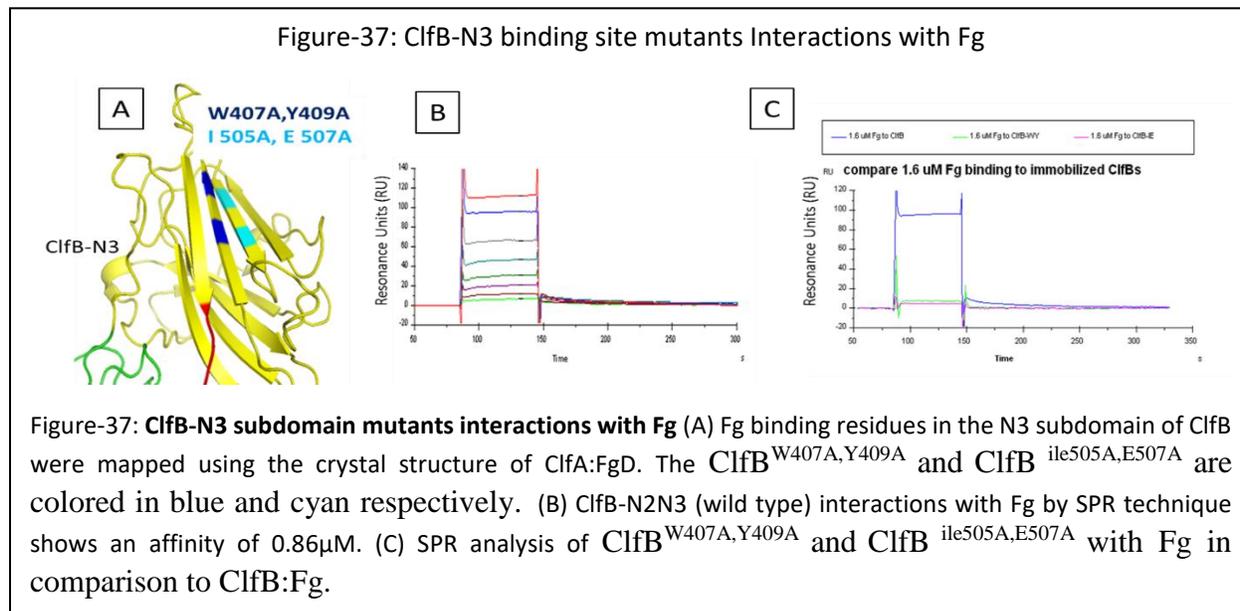
### Comparison of ClfB:CG14 with ClfB:Fga peptide complex

The crystal structure of ClfB in complex with the Fga<sub>(316-328)</sub> is published in PDB (PDB ID:3AT0 and 4F27). The ClfB:Fga<sub>(316-328)</sub> complex is a 1:1 monomeric interaction in both of these structures and the last residue R529 of the G-strand of N3 subdomain (latch) makes hydrogen bond with N238 of ClfB, which is located at the loop region of N2 subdomain. This R529-N238 hydrogen bond interactions seems critical for the locking of the Fga<sub>(316-328)</sub> peptide in the N2N3 trench site. However, in comparison of the ClfB:Fga<sub>(316-328)</sub> with our structure of ClfB:Ln peptide we found that the side chain of R529 significantly moves away from the N238 and instead makes ionic interactions with E227 residue of the dimeric ClfB. At neutral pH (7.0), the acidic side chains



of asp (E227) lose its proton and is negatively charged ( $\text{CH}_2\text{-CH}_2\text{-COO}^-$ ) and the basic side chain of arg (R529) retains its proton  $\text{-(CH}_2\text{)}_3\text{-NH-C}\begin{matrix} \text{NH} \\ \text{NH}_3^+ \end{matrix}$ , hence the side chains of Arg-Asp undergoes electrostatic interactions in the ClfB:Ln peptide complex. In comparison, R529 interacts to two different residues where it interacts with N238 of ClfB (monomer) in ClfB: Fg $\alpha_{(316-328)}$  and with E227 of ClfB (at the dimer partner) in ClfB:Ln structures (Figure-36).

From the crystal structure of ClfA:FgD complex, it was hypothesized that ClfB also has additional Fg binding regions on top of the N3 subdomain. Mainly, the residues Ile505, E507, W407 and Y409 in ClfB-N3 subdomain were found to be crucial for this interaction. I mutated these residues to alanine and got two mutants of ClfB, (a) ClfB<sup>(W407A, Y409A)</sup> (b) ClfB<sup>(ile505A, E507A)</sup>. The affinities of these mutants towards full length Fg were measured in SPR technique and compared to the wild type ClfB:Fg interactions (Figure-37).



### **Comparison of ClfB:CG14 with ClfB:CK10 peptide complex**

Crystal structure of ClfB in complex with the CK10<sub>(499-512)</sub> is published in PDB (PDB ID:3ASW and 4F1Z) (7, 46). There were structural differences at the D-D'loop of the N2 subdomain of ClfB in both of these structures. In the structure published by Dr. Xiang et.al (7) (PDB:4F1Z), the D-D' loop undergoes a rearrangement upon peptide docking into the groove of N2 subdomain and locks the peptide, whereas the structure published from our lab by Dr. Ganesh et.al (46), the D-D' loop doesn't undergo this rearrangement of D-D' loop. However, in both published crystallized structures, the G'strand of the last C-terminus of N3 subdomain that interact with the E strand of N2 subdomain is missing, which corresponds to the structure for the latching region (531-542). In the absence of this region, 'DLL' mechanism of binding for a ClfB:CK10 interaction is unlikely (Figure-38a). In comparison to these structures, our ClfB:CG12 structure showed complete interaction including the missing latch region (531-542) in the previously reported ClfB:CK10 complexes. In addition, the D-D' loop region of N2 subdomain had undergone similar structural rearrangement as reported in PDB:4F1Z (Figure-38b).

In the paper published by Dr. Xiang et.al., the author suggested that the hydrogen bonding interactions of the C-terminus 'R529' with the residue of N3 subdomain to residue N238 at the loop region of N2 subdomain is crucial for the 'DLL' mechanism of binding (7). However, in our ClfB:CG12 structure, this R529 moves away from the N238, and instead interacts with E227 of ClfB in a symmetrical dimer (Figure-39).

Figure-38: Comparison of binding sites in ClfB-Ln and ClfB-CK10 complex structures

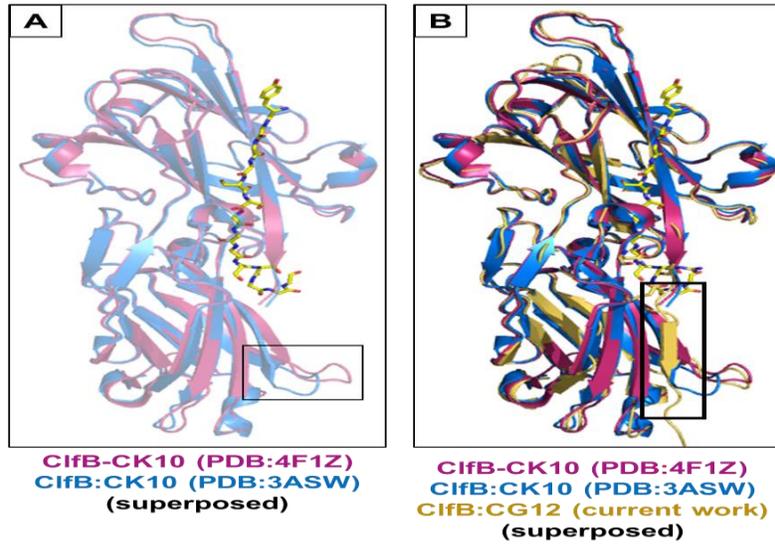


Figure-38: Comparison of ligand binding sites in ClfB-Ln and ClfB-CK10 complex structures. (A) Superpose of ClfB:CK10 complex. ClfB molecule from PDB 3asw and 4f1z structure is shown in blue and pink respectively. The CK10 peptide is shown in sticks. The D-D' loop rearrangement of ClfB-N2 subdomain is shown in black box. (B) superposition of ClfB:CG12 structure over ClfB:CK10 (both 3asw and 4f1z). The latch region (531-542) structure which is missing in ClfB:CK10 is available in ClfB:CG12 structure shown in black box. The D-D' loop rearrangement is also shown.

Figure-39: Comparison of ClfB:CG12 structure (dimer) over ClfB:CK10 structure

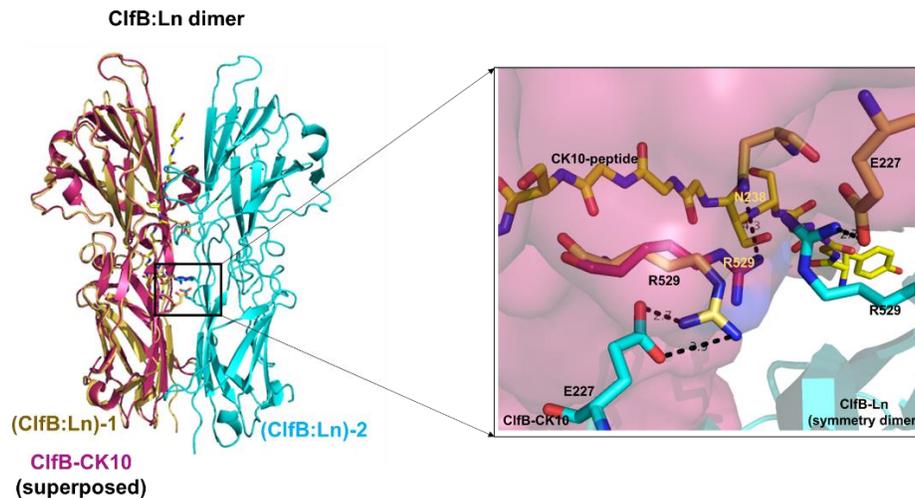
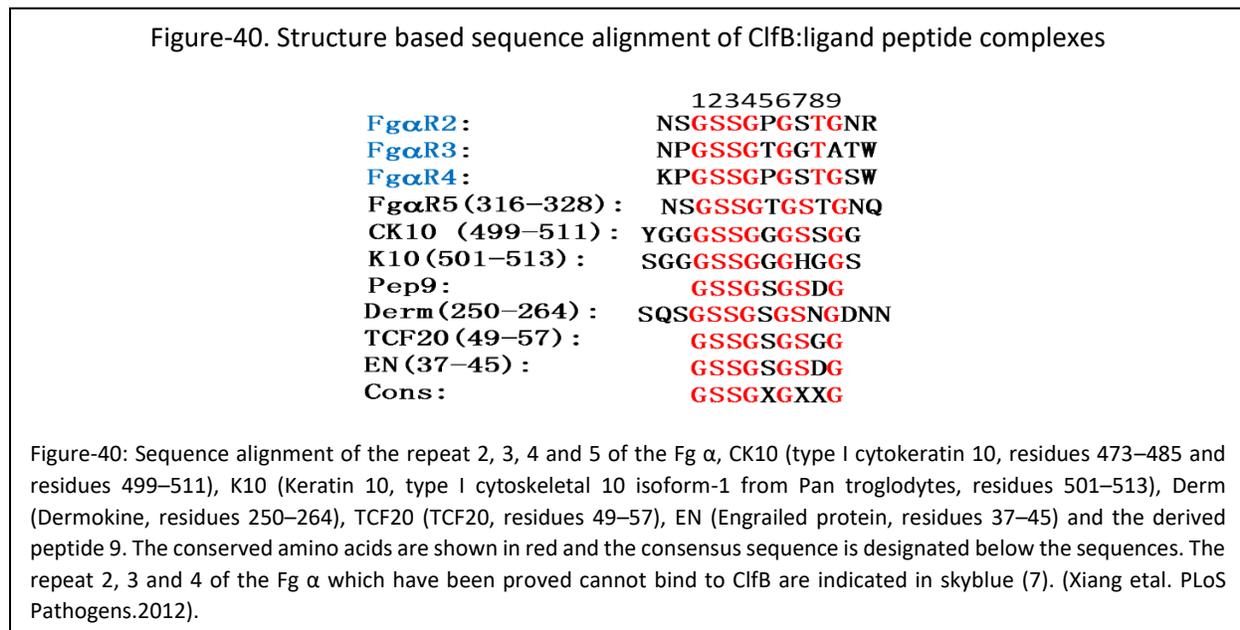


Figure-39: superposition of ClfB:CG12 structure (dimer) over ClfB:CK10 (both 3asw and 4f1z). The dimerization interface residue R529 interacts with E227 residue of another ClfB in the dimer (zoomed in). The R529 interacts with N238 in the ClfB:ck10 structure also shown for comparison.

## Structure based amino acid comparison of ClfB ligand peptides

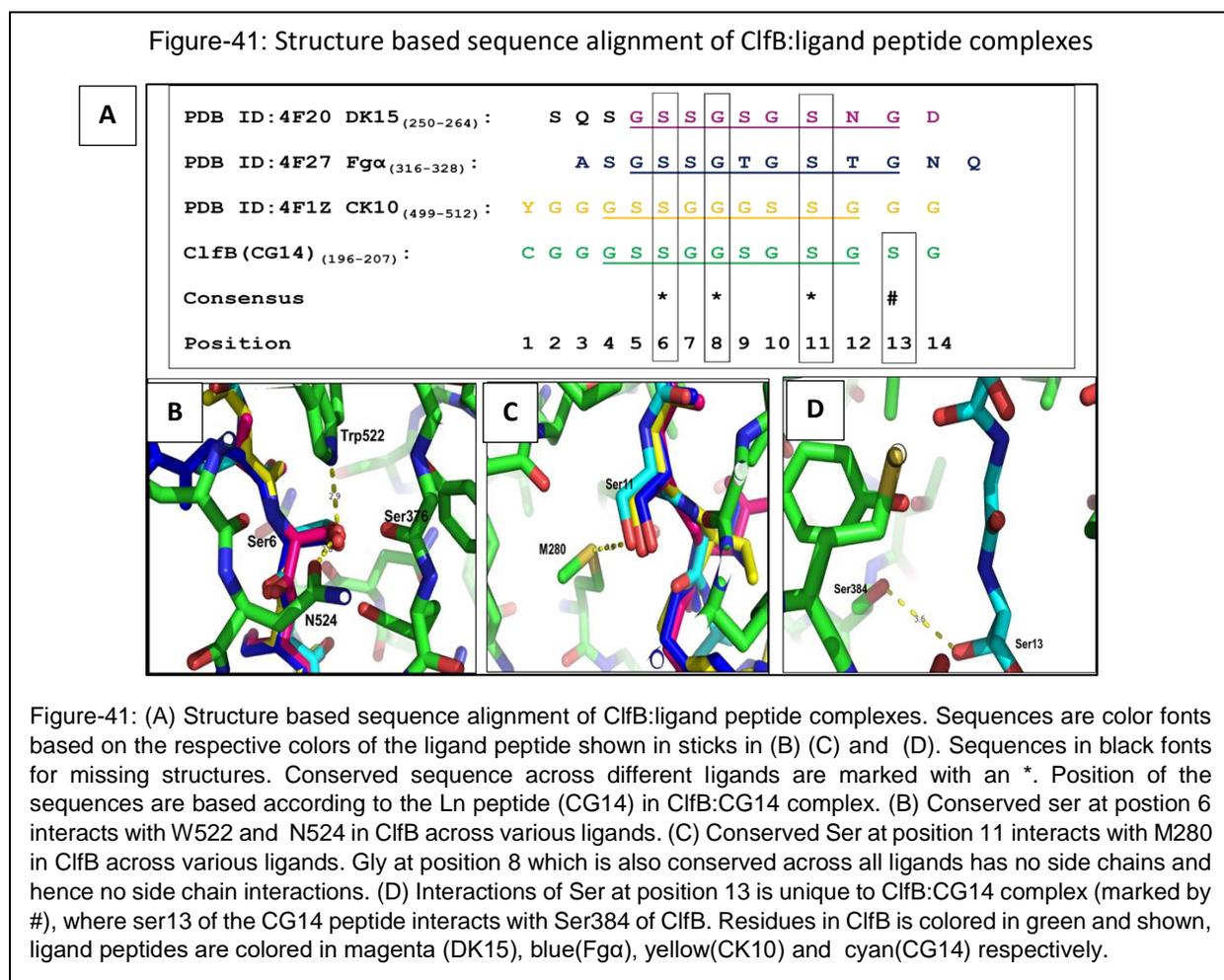
The Fg  $\alpha$  C-terminal domain (amino acids 221–610) of human Fg contains four to five repeats of approximately 13-residues regions of which up to eight of them are either glycines or serines (92). Despite these similarities, only repeat 5 (316–328) of Fg  $\alpha$ 5 was shown to be recognized by ClfB (45). The reason for this specificity is due to the presence of proline or arginine residues in the center of the putative  $\Omega$  loops in the other repeats (93). Sequence alignment of the repeats indicates that Fg  $\alpha$ 5 differs from the other repeats at the 5th, 7th and 9th positions (Figure-40) (7).



Earlier work reported that ClfB has a preference towards a ‘GSSGXG’ motif region in the ligands for optimal binding. However, sequence alignment analysis of ligand peptides of dermokine (DK), Fg $\alpha$ , and CK10 with the CG14 peptide of Loricrin (Ln) comprising region (196–207) shows that the requirement for a conserved ‘GSSGXG’ motif in the ligands is unlikely (Figure-40a).

Sequence alignment analysis shows that Ser at position 6 and Gly at position 8 and Ser at position 11 are the only three residues that are conserved across these different ligands (Figure-41a). Ser-6

interacts with Trp522 and Asn524 residues present in the C-terminus of the ClfB-N3 latch region (Figure 41b). Gly-8 has no side chains and the preference for a Gly at position 8 might be determined by the character of the neighboring residues of ClfB that cover this region due to steric hindrance. Ser-11 is interacting with the side chain of Met280 present in the N2 subdomain of ClfB (Figure 41c). Ser-13 is unique to ClfB:CG14 interactions where it is interacting with Ser384 of the ClfB-N3 subdomain. Position 13 is a Glycine in other ligand peptides (DK, Fg $\alpha$  and CK10) (Figure 41d).



## Future Directions

From the analysis of ClfB:CG14 peptide complexes discussed above (Figure-37,38,39,40), it is clear which residues are critical for the ClfB:CG14 interaction (see table or figure X). Mainly, N238 of ClfB seems to be involved in the locking of ligand peptides CK10 and Fg $\alpha$  yet doesn't participate in ClfB:CG14 interactions (Figure-38,39). Residues R529 and E227 of ClfB seem to participate in the ligand induced dimerization of ClfB:CG14 complex (Figure37). Residue Ser384 of ClfB interacts with the Ser13 of the CG14 peptide which is unique for ClfB:CG14 interactions. I will obtain three different sets of mutants to characterize the effects of these mutants binding to CG14 peptide. I will first mutate ClfB residue N238 to an alanine, which is predicted to interact with R529. Mutating N238A should knock down this interaction. Furthermore, I hypothesize that ClfB<sub>(N238A)</sub> mutant should lack/decrease affinity towards CK10 and Fg $\alpha$  peptide but still be able to interact with CG14 peptide without losing affinity. Secondly, the Ser384 of ClfB that is predicted to interact with the CG14 peptide will be mutated to a valine residue and hence a ClfB<sub>(S384V)</sub> mutant will lose the binding towards CG14 peptide. The last set of mutants to demonstrate the presence of this interaction will be ClfB residues R529 and E227 predicted to interact with each other during CG14 binding and hence a ClfB<sub>(R529Ser/E227Ile)</sub> double mutant should significantly lose the CG14 binding affinity due to lack of dimerization. I will also mutate the N3 site mutants of ClfB, (ClfB<sup>I509A/E511A</sup>, and ClfB<sup>W407A/409A</sup>) and characterize the binding of these mutants with Fg. Once I have clone and express these mutant proteins, I would be able to perform the binding assays using ITC or SPR technique (Future Directions).

### Summary of Chapter-3

The ligand binding regions (N2N3 subdomains) Clumping factor-A(220-545) and ClfB(212-542) share 26% sequence identity and has composed of two IgG like folds. The trench binding ligand peptides of ClfA is located at the C-termini (Fg $\gamma$ (392-411)) and ClfB is in the middle region (Fg $\alpha$ (316-328)). The affinities of these peptides to these MSCRAMMS were of weaker affinities more than 20 fold weaker than their affinities to intact Fg. This similarity led the hypothesis of an additional Fg binding regions in these MSCRAMM at the top of their N3 subdomain.

Upon comparison of the N-terminus of the ligand peptides, the upstream sequences in the Fg- $\alpha$  and Fg- $\gamma$  chains fold differently. Mainly, the upstream sequences in Fg- $\gamma$  (144-390) fold a globular module that is part of the FgD that is present in both Fg (soluble) and Fibrin (insoluble) and hence ClfA can bind to both Fg and Fibrin due to the binding at this region.

However, both the upstream (100-195) and downstream (221-610) sequences of Fg- $\alpha$  with respective to the ClfB binding region (196-207) in Fg $\alpha$  were highly disordered lacking any structural domains. This region is only found in intact Fg and gets cleaved during fibrin formation. Hence, ClfB can theoretically bind only to intact Fg but not to Fibrin due to the absence of this region.

**Table-16:** Comparison of ITC thermodynamic parameters of Fibrinogen binding with ClfA and ClfB.

Ligand	MSCRAMM	n	KD ( $\mu$ M)	$\Delta$ G (kcal mol <sup>-1</sup> )	$\Delta$ H (kcal mol <sup>-1</sup> )	-T $\Delta$ S (kcal mol <sup>-1</sup> )
Fibrinogen/0.01mM	ClfA(100)	1.18 $\pm$ 0.02	0.3	-8.56	-13.2 $\pm$ 0.03	<b>4.64</b>
Fibrinogen/0.01mM	ClfB(100)	0.88 $\pm$ 0.2	<b>3.7</b>	-7.12	-12.9 $\pm$ 0.02	<b>5.77</b>

The binding affinity of ClfB to intact Fg (1-4 $\mu$ M) is four fold weaker than ClfA binding to intact Fg (0.3 $\mu$ M) by ITC measurements. The thermodynamic analysis of ClfB/Fg interactions suggests a possible interaction between disordered region in Fg (possibly the  $\alpha$ -chain) interacting with the N3 subdomain of ClfB. The enthalpic ( $\Delta H$ ) contribution of Fg interaction towards ClfA (-13.2 kcal mol<sup>-1</sup>) and ClfB (-12.9 kcal mol<sup>-1</sup>) were pretty similar to each other and the free energies for the ClfA:Fg ( $\Delta G$ ) -8.56 kcal mol<sup>-1</sup> and for ClfB/Fg is -7.01kcal mol<sup>-1</sup>. (Table-16). Thus the ClfB could have additional Fg binding sites at the top of the N3 subdomain similar to ClfA:Fg interactions.

## CHAPTER IV

### CONCLUSION

Bacteria like staphylococci employs similar surface proteins that interact with host Fg molecule in a similar manner. The interaction of these surface proteins can be narrowed down to distinct sites in the,  $\alpha$ - (ClfB) and  $\gamma$ - chains (ClfA, Fbl and SpsD) of Fg. Interaction of these surface proteins interfere with the normal biological function of Fg and disrupt (1) fibrin assembly during wound healing (2) platelet retraction by interacting with platelet integrin ( $\alpha$ IIb $\beta$ 3) and (3) neutrophil activation by interacting with neutrophil integrin ( $\alpha$ M $\beta$ 2). My research work conclusively prove the structural and biochemical basis on the interaction of Fg by these surface proteins. In particular the Fg $\gamma$ -chain binding surface proteins of staphylococci might form the basis for infections like endocarditis and bacterial sepsis. The implications of a Fg $\alpha$ -chain binding surface protein (ClfB) with respect to Fg binding is not clear and the multiple ligand binding ability of ClfB to host molecules such as Fg, DK, CK10 and Ln enables the staphylococci to mount successful colonization in various host tissues make ClfB a virulent adhesion of *S.aureus*. The structural and biochemical studies carried over on ClfB, in particular loricrin interactions, and structural studies led me to compare the ClfB interaction with its other ligands DK, CK10 and Fg. The outcome of these studies are (1) ligand induced dimerization of ClfB with Ln (2) clues on the existence of an additional ligand binding site on top of N3 subdomain (3) violation of the conserved 'GSSGXG' ligand binding motif for ClfB interactions.

## REFERENCES

1. Macheboeuf P, Buffalo C, Fu CY, Zinkernagel AS, Cole JN, Johnson JE, Nizet V, Ghosh P. Streptococcal M1 protein constructs a pathological host fibrinogen network. *Nature*. 2011;472(7341):64-8. doi: 10.1038/nature09967. PubMed PMID: 21475196; PMCID: PMC3268815.
2. Clowry J, Irvine AD, McLoughlin RM. Next-generation anti-Staphylococcus aureus vaccines: A potential new therapeutic option for atopic dermatitis? *J Allergy Clin Immunol*. 2019;143(1):78-81. doi: 10.1016/j.jaci.2018.08.038. PubMed PMID: 30218675.
3. Ganesh VK, Liang X, Geoghegan JA, Cohen ALV, Venugopalan N, Foster TJ, Hook M. Lessons from the Crystal Structure of the *S. aureus* Surface Protein Clumping Factor A in Complex With Tefibazumab, an Inhibiting Monoclonal Antibody. *EBioMedicine*. 2016;13:328-38. doi: 10.1016/j.ebiom.2016.09.027. PubMed PMID: 27789272; PMCID: PMC5264652.
4. Flick MJ, Du X, Prasad JM, Raghu H, Palumbo JS, Smeds E, Hook M, Degen JL. Genetic elimination of the binding motif on fibrinogen for the *S. aureus* virulence factor ClfA improves host survival in septicemia. *Blood*. 2013;121(10):1783-94. doi: 10.1182/blood-2012-09-453894. PubMed PMID: 23299312; PMCID: PMC3591798.
5. Deivanayagam CC, Wann ER, Chen W, Carson M, Rajashankar KR, Hook M, Narayana SV. A novel variant of the immunoglobulin fold in surface adhesins of *Staphylococcus aureus*: crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *EMBO J*. 2002;21(24):6660-72. doi: 10.1093/emboj/cdf619. PubMed PMID: 12485987; PMCID: PMC139082.

6. Foster TJ, Geoghegan JA, Ganesh VK, Hook M. Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat Rev Microbiol*. 2014;12(1):49-62. doi: 10.1038/nrmicro3161. PubMed PMID: 24336184; PMCID: PMC5708296.
7. Xiang H, Feng Y, Wang J, Liu B, Chen Y, Liu L, Deng X, Yang M. Crystal structures reveal the multi-ligand binding mechanism of *Staphylococcus aureus* ClfB. *PLoS Pathog*. 2012;8(6):e1002751. doi: 10.1371/journal.ppat.1002751. PubMed PMID: 22719251; PMCID: PMC3375286.
8. McAdow M, Kim HK, Dedent AC, Hendrickx AP, Schneewind O, Missiakas DM. Preventing *Staphylococcus aureus* sepsis through the inhibition of its agglutination in blood. *PLoS Pathog*. 2011;7(10):e1002307. doi: 10.1371/journal.ppat.1002307. PubMed PMID: 22028651; PMCID: PMC3197598.
9. Ganesh VK, Rivera JJ, Smeds E, Ko YP, Bowden MG, Wann ER, Gurusiddappa S, Fitzgerald JR, Hook M. A structural model of the *Staphylococcus aureus* ClfA-fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics. *PLoS Pathog*. 2008;4(11):e1000226. doi: 10.1371/journal.ppat.1000226. PubMed PMID: 19043557; PMCID: PMC2582960.
10. O'Connell DP, Nanavaty T, McDevitt D, Gurusiddappa S, Hook M, Foster TJ. The fibrinogen-binding MSCRAMM (clumping factor) of *Staphylococcus aureus* has a Ca<sup>2+</sup>-dependent inhibitory site. *J Biol Chem*. 1998;273(12):6821-9. doi: 10.1074/jbc.273.12.6821. PubMed PMID: 9506984.
11. Anderson AS, Miller AA, Donald RG, Scully IL, Nanra JS, Cooper D, Jansen KU. Development of a multicomponent *Staphylococcus aureus* vaccine designed to counter multiple

bacterial virulence factors. *Hum Vaccin Immunother.* 2012;8(11):1585-94. doi: 10.4161/hv.21872. PubMed PMID: 22922765; PMCID: PMC3601133.

12. Ibrahim AS, Luo G, Gebremariam T, Lee H, Schmidt CS, Hennessey JP, Jr., French SW, Yeaman MR, Filler SG, Edwards JE, Jr. NDV-3 protects mice from vulvovaginal candidiasis through T- and B-cell immune response. *Vaccine.* 2013;31(47):5549-56. doi: 10.1016/j.vaccine.2013.09.016. PubMed PMID: 24063977; PMCID: PMC3866209.

13. Yu XQ, Robbie GJ, Wu Y, Esser MT, Jensen K, Schwartz HI, Bellamy T, Hernandez-Illas M, Jafri HS. Safety, Tolerability, and Pharmacokinetics of MEDI4893, an Investigational, Extended-Half-Life, Anti-Staphylococcus aureus Alpha-Toxin Human Monoclonal Antibody, in Healthy Adults. *Antimicrob Agents Chemother.* 2017;61(1). doi: 10.1128/AAC.01020-16. PubMed PMID: 27795368; PMCID: PMC5192133.

14. Tabor DE, Yu L, Mok H, Tkaczyk C, Sellman BR, Wu Y, Oganessian V, Slidel T, Jafri H, McCarthy M, Bradford P, Esser MT. Staphylococcus aureus Alpha-Toxin Is Conserved among Diverse Hospital Respiratory Isolates Collected from a Global Surveillance Study and Is Neutralized by Monoclonal Antibody MEDI4893. *Antimicrob Agents Chemother.* 2016;60(9):5312-21. doi: 10.1128/AAC.00357-16. PubMed PMID: 27324766; PMCID: PMC4997823.

15. Hua L, Cohen TS, Shi Y, Datta V, Hilliard JJ, Tkaczyk C, Suzich J, Stover CK, Sellman BR. MEDI4893\* Promotes Survival and Extends the Antibiotic Treatment Window in a Staphylococcus aureus Immunocompromised Pneumonia Model. *Antimicrob Agents Chemother.* 2015;59(8):4526-32. doi: 10.1128/AAC.00510-15. PubMed PMID: 25987629; PMCID: PMC4505239.

16. Hilliard JJ, Datta V, Tkaczyk C, Hamilton M, Sadowska A, Jones-Nelson O, O'Day T, Weiss WJ, Szarka S, Nguyen V, Prokai L, Suzich J, Stover CK, Sellman BR. Anti-alpha-toxin monoclonal antibody and antibiotic combination therapy improves disease outcome and accelerates healing in a *Staphylococcus aureus* dermonecrosis model. *Antimicrob Agents Chemother.* 2015;59(1):299-309. doi: 10.1128/AAC.03918-14. PubMed PMID: 25348518; PMCID: PMC4291365.
17. Chang FY. *Staphylococcus aureus* bacteremia and endocarditis. *J Microbiol Immunol Infect.* 2000;33(2):63-8. PubMed PMID: 10917874.
18. Lin JC, Wu JS, Chang FY. Community-acquired methicillin-resistant *Staphylococcus aureus* endocarditis with septic embolism of popliteal artery: a case report. *J Microbiol Immunol Infect.* 2000;33(1):57-9. PubMed PMID: 10806967.
19. Brandt CM, Rouse MS, Tallan BM, Laue NW, Wilson WR, Steckelberg JM. Effective treatment of cephalosporin-rifampin combinations against cryptic methicillin-resistant beta-lactamase-producing coagulase-negative staphylococcal experimental endocarditis. *Antimicrob Agents Chemother.* 1995;39(8):1815-9. doi: 10.1128/aac.39.8.1815. PubMed PMID: 7486924; PMCID: PMC162831.
20. Gentry CA, Rodvold KA, Novak RM, Hershov RC, Naderer OJ. Retrospective evaluation of therapies for *Staphylococcus aureus* endocarditis. *Pharmacotherapy.* 1997;17(5):990-7. PubMed PMID: 9324187.
21. Sullam PM, Bayer AS, Foss WM, Cheung AL. Diminished platelet binding in vitro by *Staphylococcus aureus* is associated with reduced virulence in a rabbit model of infective endocarditis. *Infect Immun.* 1996;64(12):4915-21. PubMed PMID: 8945526; PMCID: PMC174468.

22. Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S, Harrison LH, Lynfield R, Dumyati G, Townes JM, Craig AS, Zell ER, Fosheim GE, McDougal LK, Carey RB, Fridkin SK, Active Bacterial Core surveillance MI. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA*. 2007;298(15):1763-71. doi: 10.1001/jama.298.15.1763. PubMed PMID: 17940231.
23. Mazmanian SK, Liu G, Ton-That H, Schneewind O. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science*. 1999;285(5428):760-3. doi: 10.1126/science.285.5428.760. PubMed PMID: 10427003.
24. Friedrich R, Panizzi P, Fuentes-Prior P, Richter K, Verhamme I, Anderson PJ, Kawabata S, Huber R, Bode W, Bock PE. Staphylocoagulase is a prototype for the mechanism of cofactor-induced zymogen activation. *Nature*. 2003;425(6957):535-9. doi: 10.1038/nature01962. PubMed PMID: 14523451.
25. Ko YP, Kang M, Ganesh VK, Ravirajan D, Li B, Hook M. Coagulase and Efb of *Staphylococcus aureus* Have a Common Fibrinogen Binding Motif. *MBio*. 2016;7(1):e01885-15. doi: 10.1128/mBio.01885-15. PubMed PMID: 26733070; PMCID: PMC4725008.
26. Lipinski B, Hawiger J, Jeljaszewicz J. Staphylococcal clumping with soluble fibrin monomer complexes. *J Exp Med*. 1967;126(5):979-88. doi: 10.1084/jem.126.5.979. PubMed PMID: 4228857; PMCID: PMC2138412.
27. Duthie ES. The action of fibrinogen on certain pathogenic cocci. *J Gen Microbiol*. 1955;13(2):383-93. doi: 10.1099/00221287-13-2-383. PubMed PMID: 13278487.
28. Kollman JM, Pandi L, Sawaya MR, Riley M, Doolittle RF. Crystal structure of human fibrinogen. *Biochemistry*. 2009;48(18):3877-86. doi: 10.1021/bi802205g. PubMed PMID: 19296670.

29. Hoeprich PD, Jr., Doolittle RF. Dimeric half-molecules of human fibrinogen are joined through disulfide bonds in an antiparallel orientation. *Biochemistry*. 1983;22(9):2049-55. doi: 10.1021/bi00278a003. PubMed PMID: 6860649.
30. Doolittle RF. X-ray crystallographic studies on fibrinogen and fibrin. *J Thromb Haemost*. 2003;1(7):1559-65. PubMed PMID: 12871291.
31. Everse SJ, Spraggon G, Veerapandian L, Doolittle RF. Conformational changes in fragments D and double-D from human fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide. *Biochemistry*. 1999;38(10):2941-6. doi: 10.1021/bi982626w. PubMed PMID: 10074346.
32. Mosesson MW, Siebenlist KR, DiOrio JP, Matsuda M, Hainfeld JF, Wall JS. The role of fibrinogen D domain intermolecular association sites in the polymerization of fibrin and fibrinogen Tokyo II (gamma 275 Arg-->Cys). *J Clin Invest*. 1995;96(2):1053-8. doi: 10.1172/JCI118091. PubMed PMID: 7635941; PMCID: PMC185294.
33. Spraggon G, Everse SJ, Doolittle RF. Crystal structures of fragment D from human fibrinogen and its crosslinked counterpart from fibrin. *Nature*. 1997;389(6650):455-62. doi: 10.1038/38947. PubMed PMID: 9333233.
34. Yang Z, Pandi L, Doolittle RF. The crystal structure of fragment double-D from cross-linked lamprey fibrin reveals isopeptide linkages across an unexpected D-D interface. *Biochemistry*. 2002;41(52):15610-7. doi: 10.1021/bi026666i. PubMed PMID: 12501189.
35. Triantaphyllopoulos DC, Triantaphyllopoulos E. Solubility of Fibrin Clots in Solutions of Heparin. *Nature*. 1964;204:1096-8. doi: 10.1038/2041096b0. PubMed PMID: 14243392.

36. Latallo ZS, Budzynski AZ, Lipinski B, Kowalski E. Inhibition of Thrombin and of Fibrin Polymerization, Two Activities Derived from Plasmin-Digested Fibrinogen. *Nature*. 1964;203:1184-5. doi: 10.1038/2031184a0. PubMed PMID: 14217981.
37. Kostelansky MS, Betts L, Gorkun OV, Lord ST. 2.8 Å crystal structures of recombinant fibrinogen fragment D with and without two peptide ligands: GHRP binding to the "b" site disrupts its nearby calcium-binding site. *Biochemistry*. 2002;41(40):12124-32. doi: 10.1021/bi0261894. PubMed PMID: 12356313.
38. Soria J, Soria C, Samama M, Tabori S, Kehl M, Henschen A, Nieuwenhuizen W, Rimón A, Tatarski I. Fibrinogen Haifa: fibrinogen variant with absence of protective effect of calcium on plasmin degradation of gamma chains. *Thromb Haemost*. 1987;57(3):310-3. PubMed PMID: 2958955.
39. Yamazumi K, Doolittle RF. The synthetic peptide Gly-Pro-Arg-Pro-amide limits the plasminic digestion of fibrinogen in the same fashion as calcium ion. *Protein Sci*. 1992;1(12):1719-20. doi: 10.1002/pro.5560011220. PubMed PMID: 1304901; PMCID: PMC2142132.
40. Odrlićin TM, Rybarczyk BJ, Francis CW, Lawrence SO, Hamaguchi M, Simpson-Haidaris PJ. Calcium modulates plasmin cleavage of the fibrinogen D fragment gamma chain N-terminus: mapping of monoclonal antibody J88B to a plasmin sensitive domain of the gamma chain. *Biochim Biophys Acta*. 1996;1298(1):69-77. doi: 10.1016/s0167-4838(96)00090-8. PubMed PMID: 8948490.
41. McDevitt D, Francois P, Vaudaux P, Foster TJ. Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*. *Mol Microbiol*. 1994;11(2):237-48. doi: 10.1111/j.1365-2958.1994.tb00304.x. PubMed PMID: 8170386.

42. Ni Eidhin D, Perkins S, Francois P, Vaudaux P, Hook M, Foster TJ. Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus*. *Mol Microbiol.* 1998;30(2):245-57. doi: 10.1046/j.1365-2958.1998.01050.x. PubMed PMID: 9791170.
43. Josefsson E, McCrea KW, Ni Eidhin D, O'Connell D, Cox J, Hook M, Foster TJ. Three new members of the serine-aspartate repeat protein multigene family of *Staphylococcus aureus*. *Microbiology.* 1998;144 ( Pt 12):3387-95. doi: 10.1099/00221287-144-12-3387. PubMed PMID: 9884231.
44. Hair PS, Echague CG, Sholl AM, Watkins JA, Geoghegan JA, Foster TJ, Cunnion KM. Clumping factor A interaction with complement factor I increases C3b cleavage on the bacterial surface of *Staphylococcus aureus* and decreases complement-mediated phagocytosis. *Infect Immun.* 2010;78(4):1717-27. doi: 10.1128/IAI.01065-09. PubMed PMID: 20100856; PMCID: PMC2849425.
45. Walsh EJ, O'Brien LM, Liang X, Hook M, Foster TJ. Clumping factor B, a fibrinogen-binding MSCRAMM (microbial surface components recognizing adhesive matrix molecules) adhesin of *Staphylococcus aureus*, also binds to the tail region of type I cytokeratin 10. *J Biol Chem.* 2004;279(49):50691-9. doi: 10.1074/jbc.M408713200. PubMed PMID: 15385531.
46. Ganesh VK, Barbu EM, Deivanayagam CC, Le B, Anderson AS, Matsuka YV, Lin SL, Foster TJ, Narayana SV, Hook M. Structural and biochemical characterization of *Staphylococcus aureus* clumping factor B/ligand interactions. *J Biol Chem.* 2011;286(29):25963-72. doi: 10.1074/jbc.M110.217414. PubMed PMID: 21543319; PMCID: PMC3138276.
47. Mulcahy ME, Geoghegan JA, Monk IR, O'Keeffe KM, Walsh EJ, Foster TJ, McLoughlin RM. Nasal colonisation by *Staphylococcus aureus* depends upon clumping factor B binding to the

squamous epithelial cell envelope protein loricrin. *PLoS Pathog.* 2012;8(12):e1003092. doi: 10.1371/journal.ppat.1003092. PubMed PMID: 23300445; PMCID: PMC3531522.

48. Liu CZ, Shih MH, Tsai PJ. ClfA(221-550), a fibrinogen-binding segment of *Staphylococcus aureus* clumping factor A, disrupts fibrinogen function. *Thromb Haemost.* 2005;94(2):286-94. doi: 10.1160/TH05-03-0205. PubMed PMID: 16113817.

49. Flick MJ, Du X, Witte DP, Jirouskova M, Soloviev DA, Busuttill SJ, Plow EF, Degen JL. Leukocyte engagement of fibrin(ogen) via the integrin receptor alphaMbeta2/Mac-1 is critical for host inflammatory response in vivo. *J Clin Invest.* 2004;113(11):1596-606. doi: 10.1172/JCI20741. PubMed PMID: 15173886; PMCID: PMC419487.

50. Siboo IR, Cheung AL, Bayer AS, Sullam PM. Clumping factor A mediates binding of *Staphylococcus aureus* to human platelets. *Infect Immun.* 2001;69(5):3120-7. doi: 10.1128/IAI.69.5.3120-3127.2001. PubMed PMID: 11292731; PMCID: PMC98267.

51. McDevitt D, Nanavaty T, House-Pompeo K, Bell E, Turner N, McIntire L, Foster T, Hook M. Characterization of the interaction between the *Staphylococcus aureus* clumping factor (ClfA) and fibrinogen. *Eur J Biochem.* 1997;247(1):416-24. doi: 10.1111/j.1432-1033.1997.00416.x. PubMed PMID: 9249055.

52. Loughman A, Fitzgerald JR, Brennan MP, Higgins J, Downer R, Cox D, Foster TJ. Roles for fibrinogen, immunoglobulin and complement in platelet activation promoted by *Staphylococcus aureus* clumping factor A. *Mol Microbiol.* 2005;57(3):804-18. doi: 10.1111/j.1365-2958.2005.04731.x. PubMed PMID: 16045623.

53. Higgins J, Loughman A, van Kessel KP, van Strijp JA, Foster TJ. Clumping factor A of *Staphylococcus aureus* inhibits phagocytosis by human polymorphonuclear leucocytes. *FEMS*

Microbiol Lett. 2006;258(2):290-6. doi: 10.1111/j.1574-6968.2006.00229.x. PubMed PMID: 16640587.

54. Flick MJ, Du X, Degen JL. Fibrin(ogen)-alpha M beta 2 interactions regulate leukocyte function and innate immunity in vivo. *Exp Biol Med (Maywood)*. 2004;229(11):1105-10. doi: 10.1177/153537020422901104. PubMed PMID: 15564436.

55. Scully IL, Timofeyeva Y, Keeney D, Matsuka YV, Severina E, McNeil LK, Nanra J, Hu G, Liberator PA, Jansen KU, Anderson AS. Demonstration of the preclinical correlate of protection for Staphylococcus aureus clumping factor A in a murine model of infection. *Vaccine*. 2015;33(41):5452-7. doi: 10.1016/j.vaccine.2015.08.029. PubMed PMID: 26319743.

56. Heilbronner S, Hanses F, Monk IR, Speziale P, Foster TJ. Sortase A promotes virulence in experimental Staphylococcus lugdunensis endocarditis. *Microbiology*. 2013;159(Pt 10):2141-52. doi: 10.1099/mic.0.070292-0. PubMed PMID: 23943787.

57. Giormezis N, Kolonitsiou F, Makri A, Vogiatzi A, Christofidou M, Anastassiou ED, Spiliopoulou I. Virulence factors among Staphylococcus lugdunensis are associated with infection sites and clonal spread. *Eur J Clin Microbiol Infect Dis*. 2015;34(4):773-8. doi: 10.1007/s10096-014-2291-8. PubMed PMID: 25471196.

58. Geoghegan JA, Ganesh VK, Smeds E, Liang X, Hook M, Foster TJ. Molecular characterization of the interaction of staphylococcal microbial surface components recognizing adhesive matrix molecules (MSCRAMM) ClfA and Fbl with fibrinogen. *J Biol Chem*. 2010;285(9):6208-16. doi: 10.1074/jbc.M109.062208. PubMed PMID: 20007717; PMCID: PMC2825416.

59. Nilsson M, Bjerketorp J, Guss B, Frykberg L. A fibrinogen-binding protein of *Staphylococcus lugdunensis*. *FEMS Microbiol Lett.* 2004;241(1):87-93. doi: 10.1016/j.femsle.2004.10.008. PubMed PMID: 15556714.
60. Mitchell J, Tristan A, Foster TJ. Characterization of the fibrinogen-binding surface protein Fbl of *Staphylococcus lugdunensis*. *Microbiology.* 2004;150(Pt 11):3831-41. doi: 10.1099/mic.0.27337-0. PubMed PMID: 15528668.
61. Bannoehr J, Brown JK, Shaw DJ, Fitzgerald RJ, van den Broek AH, Thoday KL. *Staphylococcus pseudintermedius* surface proteins SpsD and SpsO mediate adherence to ex vivo canine corneocytes. *Vet Dermatol.* 2012;23(2):119-24, e26. doi: 10.1111/j.1365-3164.2011.01021.x. PubMed PMID: 22112246.
62. Bannoehr J, Ben Zakour NL, Reglinski M, Inglis NF, Prabhakaran S, Fossum E, Smith DG, Wilson GJ, Cartwright RA, Haas J, Hook M, van den Broek AH, Thoday KL, Fitzgerald JR. Genomic and surface proteomic analysis of the canine pathogen *Staphylococcus pseudintermedius* reveals proteins that mediate adherence to the extracellular matrix. *Infect Immun.* 2011;79(8):3074-86. doi: 10.1128/IAI.00137-11. PubMed PMID: 21576333; PMCID: PMC3147560.
63. Pietrocola G, Geoghegan JA, Rindi S, Di Poto A, Missineo A, Consalvi V, Foster TJ, Speziale P. Molecular Characterization of the Multiple Interactions of SpsD, a Surface Protein from *Staphylococcus pseudintermedius*, with Host Extracellular Matrix Proteins. *PLoS One.* 2013;8(6):e66901. doi: 10.1371/journal.pone.0066901. PubMed PMID: 23805283; PMCID: PMC3689669.
64. Pietrocola G, Gianotti V, Richards A, Nobile G, Geoghegan JA, Rindi S, Monk IR, Bordt AS, Foster TJ, Fitzgerald JR, Speziale P. Fibronectin Binding Proteins SpsD and SpsL Both

Support Invasion of Canine Epithelial Cells by *Staphylococcus pseudintermedius*. *Infect Immun*. 2015;83(10):4093-102. doi: 10.1128/IAI.00542-15. PubMed PMID: 26238710; PMCID: PMC4567651.

65. Peacock SJ, Foster TJ, Cameron BJ, Berendt AR. Bacterial fibronectin-binding proteins and endothelial cell surface fibronectin mediate adherence of *Staphylococcus aureus* to resting human endothelial cells. *Microbiology*. 1999;145 ( Pt 12):3477-86. doi: 10.1099/00221287-145-12-3477. PubMed PMID: 10627045.

66. Salgado PS, Yan R, Rowan F, Cota E. Expression, crystallization and preliminary X-ray data analysis of NT-Als9-2, a fungal adhesin from *Candida albicans*. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2011;67(Pt 4):467-70. doi: 10.1107/S1744309111003460. PubMed PMID: 21505243; PMCID: PMC3080152.

67. Stemberk V, Jones RP, Moroz O, Atkin KE, Edwards AM, Turkenburg JP, Leech AP, Massey RC, Potts JR. Evidence for steric regulation of fibrinogen binding to *Staphylococcus aureus* fibronectin-binding protein A (FnBPA). *J Biol Chem*. 2014;289(18):12842-51. doi: 10.1074/jbc.M113.543546. PubMed PMID: 24627488; PMCID: PMC4007472.

68. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr*. 2006;62(Pt 8):859-66. doi: 10.1107/S0907444906019949. PubMed PMID: 16855301.

69. Rinaldelli M, Ravera E, Calderone V, Parigi G, Murshudov GN, Luchinat C. Simultaneous use of solution NMR and X-ray data in REFMAC5 for joint refinement/detection of structural differences. *Acta Crystallogr D Biol Crystallogr*. 2014;70(Pt 4):958-67. doi: 10.1107/S1399004713034160. PubMed PMID: 24699641; PMCID: PMC4306559.

70. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 2004;60(Pt 12 Pt 1):2126-32. doi: 10.1107/S0907444904019158. PubMed PMID: 15572765.
71. Weems JJ, Jr., Steinberg JP, Filler S, Baddley JW, Corey GR, Sampathkumar P, Winston L, John JF, Kubin CJ, Talwani R, Moore T, Patti JM, Hetherington S, Texter M, Wenzel E, Kelley VA, Fowler VG, Jr. Phase II, randomized, double-blind, multicenter study comparing the safety and pharmacokinetics of tefibazumab to placebo for treatment of *Staphylococcus aureus* bacteremia. *Antimicrob Agents Chemother*. 2006;50(8):2751-5. doi: 10.1128/AAC.00096-06. PubMed PMID: 16870768; PMCID: PMC1538656.
72. Rubel C, Fernandez GC, Rosa FA, Gomez S, Bompadre MB, Coso OA, Isturiz MA, Palermo MS. Soluble fibrinogen modulates neutrophil functionality through the activation of an extracellular signal-regulated kinase-dependent pathway. *J Immunol*. 2002;168(7):3527-35. doi: 10.4049/jimmunol.168.7.3527. PubMed PMID: 11907115.
73. Altieri DC, Agbanyo FR, Plescia J, Ginsberg MH, Edgington TS, Plow EF. A unique recognition site mediates the interaction of fibrinogen with the leukocyte integrin Mac-1 (CD11b/CD18). *J Biol Chem*. 1990;265(21):12119-22. PubMed PMID: 1973686.
74. Kazura JW, Wenger JD, Salata RA, Budzynski AZ, Goldsmith GH. Modulation of polymorphonuclear leukocyte microbicidal activity and oxidative metabolism by fibrinogen degradation products D and E. *J Clin Invest*. 1989;83(6):1916-24. doi: 10.1172/JCI114098. PubMed PMID: 2542377; PMCID: PMC303912.
75. Weisel JW, Litvinov RI. Mechanisms of fibrin polymerization and clinical implications. *Blood*. 2013;121(10):1712-9. doi: 10.1182/blood-2012-09-306639. PubMed PMID: 23305734; PMCID: PMC3591795.

76. Macheboeuf P, Ghosh P. [Structural basis for streptococcal toxic shock syndrome]. *Med Sci (Paris)*. 2011;27(10):814-6. doi: 10.1051/medsci/20112710007. PubMed PMID: 22027416; PMCID: PMC5517076.
77. Beaufort N, Wojciechowski P, Sommerhoff CP, Szmyd G, Dubin G, Eick S, Kellermann J, Schmitt M, Potempa J, Magdolen V. The human fibrinolytic system is a target for the staphylococcal metalloprotease aureolysin. *Biochem J*. 2008;410(1):157-65. doi: 10.1042/BJ20070650. PubMed PMID: 17973626.
78. McAleese FM, Walsh EJ, Sieprawska M, Potempa J, Foster TJ. Loss of clumping factor B fibrinogen binding activity by *Staphylococcus aureus* involves cessation of transcription, shedding and cleavage by metalloprotease. *J Biol Chem*. 2001;276(32):29969-78. doi: 10.1074/jbc.M102389200. PubMed PMID: 11399757.
79. Valderrama JA, Riestra AM, Gao NJ, LaRock CN, Gupta N, Ali SR, Hoffman HM, Ghosh P, Nizet V. Group A streptococcal M protein activates the NLRP3 inflammasome. *Nat Microbiol*. 2017;2(10):1425-34. doi: 10.1038/s41564-017-0005-6. PubMed PMID: 28784982; PMCID: PMC5750061.
80. Dohrmann S, LaRock CN, Anderson EL, Cole JN, Ryali B, Stewart C, Nonejuie P, Pogliano J, Corriden R, Ghosh P, Nizet V. Group A Streptococcal M1 Protein Provides Resistance against the Antimicrobial Activity of Histones. *Sci Rep*. 2017;7:43039. doi: 10.1038/srep43039. PubMed PMID: 28220899; PMCID: PMC5318940.
81. Uchiyama S, Andreoni F, Zurcher C, Schilcher K, Ender M, Madon J, Matt U, Ghosh P, Nizet V, Schuepbach RA, Zinkernagel AS. Coiled-coil irregularities of the M1 protein structure promote M1-fibrinogen interaction and influence group A *Streptococcus* host cell interactions and

virulence. *J Mol Med (Berl)*. 2013;91(7):861-9. doi: 10.1007/s00109-013-1012-6. PubMed PMID: 23443671; PMCID: PMC3695690.

82. Schaffer AC, Solinga RM, Cocchiario J, Portoles M, Kiser KB, Risley A, Randall SM, Valtulina V, Speziale P, Walsh E, Foster T, Lee JC. Immunization with *Staphylococcus aureus* clumping factor B, a major determinant in nasal carriage, reduces nasal colonization in a murine model. *Infect Immun*. 2006;74(4):2145-53. doi: 10.1128/IAI.74.4.2145-2153.2006. PubMed PMID: 16552044; PMCID: PMC1418917.

83. Entenza JM, Foster TJ, Ni Eidhin D, Vaudaux P, Francioli P, Moreillon P. Contribution of clumping factor B to pathogenesis of experimental endocarditis due to *Staphylococcus aureus*. *Infect Immun*. 2000;68(9):5443-6. doi: 10.1128/iai.68.9.5443-5446.2000. PubMed PMID: 10948180; PMCID: PMC101814.

84. Lee AS, Gizard Y, Empel J, Bonetti EJ, Harbarth S, Francois P. Mupirocin-induced mutations in *ileS* in various genetic backgrounds of methicillin-resistant *Staphylococcus aureus*. *J Clin Microbiol*. 2014;52(10):3749-54. doi: 10.1128/JCM.01010-14. PubMed PMID: 25122856; PMCID: PMC4187773.

85. Hurdle JG, O'Neill AJ, Ingham E, Fishwick C, Chopra I. Analysis of mupirocin resistance and fitness in *Staphylococcus aureus* by molecular genetic and structural modeling techniques. *Antimicrob Agents Chemother*. 2004;48(11):4366-76. doi: 10.1128/AAC.48.11.4366-4376.2004. PubMed PMID: 15504866; PMCID: PMC525403.

86. Hurdle JG, O'Neill AJ, Chopra I. The isoleucyl-tRNA synthetase mutation V588F conferring mupirocin resistance in glycopeptide-intermediate *Staphylococcus aureus* is not associated with a significant fitness burden. *J Antimicrob Chemother*. 2004;53(1):102-4. doi: 10.1093/jac/dkh020. PubMed PMID: 14657089.

87. Barbu EM, Ganesh VK, Gurusiddappa S, Mackenzie RC, Foster TJ, Sudhof TC, Hook M. beta-Neurexin is a ligand for the *Staphylococcus aureus* MSCRAMM SdrC. *PLoS Pathog.* 2010;6(1):e1000726. doi: 10.1371/journal.ppat.1000726. PubMed PMID: 20090838; PMCID: PMC2800189.
88. Ganesh VK, Barbu EM, Deivanayagam CC, Le B, Anderson AS, Matsuka YV, Lin SL, Foster TJ, Narayana SV, Hook M. Structural and biochemical characterization of *Staphylococcus aureus* clumping factor B/ligand interactions. *The Journal of biological chemistry.* 2011;286(29):25963-72. Epub 2011/05/06. doi: 10.1074/jbc.M110.217414. PubMed PMID: 21543319; PMCID: 3138276.
89. Ponnuraj K, Bowden MG, Davis S, Gurusiddappa S, Moore D, Choe D, Xu Y, Höök M, Narayana SV. A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. *Cell.* 2003;115(2):217-28. PubMed PMID: 14567919.
90. Mulcahy ME, Geoghegan JA, Monk IR, O'Keefe KM, Walsh EJ, Foster TJ, McLoughlin RM. Nasal colonisation by *Staphylococcus aureus* depends upon clumping factor B binding to the squamous epithelial cell envelope protein loricrin. *PLoS pathogens.* 2012;8(12):e1003092. Epub 2013/01/10. doi: 10.1371/journal.ppat.1003092. PubMed PMID: 23300445; PMCID: 3531522.
91. Xiang H, Feng Y, Wang J, Liu B, Chen Y, Liu L, Deng X, Yang M. Crystal structures reveal the multi-ligand binding mechanism of *Staphylococcus aureus* ClfB. *PLoS pathogens.* 2012;8(6):e1002751. Epub 2012/06/22. doi: 10.1371/journal.ppat.1002751. PubMed PMID: 22719251; PMCID: 3375286.
92. Tsurupa G, Tsonev L, Medved L. Structural organization of the fibrin(ogen) alpha C-domain. *Biochemistry.* 2002;41(20):6449-59. doi: 10.1021/bi025584r. PubMed PMID: 12009908.

93. Walsh EJ, Miajlovic H, Gorkun OV, Foster TJ. Identification of the *Staphylococcus aureus* MSCRAMM clumping factor B (ClfB) binding site in the alphaC-domain of human fibrinogen. *Microbiology*. 2008;154(Pt 2):550-8. doi: 10.1099/mic.0.2007/010868-0. PubMed PMID: 18227259; PMCID: PMC2885624.