

**Structural and functional studies of a
Staphylococcus aureus surface protein**

Andrew Stephen Brentnall

PhD

University of York

Department of Biology

October 2012

Abstract

Staphylococcus aureus (*S. aureus*) is a common human pathogen that causes a variety of diseases including infective endocarditis, necrotizing pneumonia and sepsis. An important aspect of *S. aureus*' virulence is its ability to form biofilms, particularly following the implantation of indwelling medical devices. Biofilm infections are acutely difficult to treat due to the increased antimicrobial resistance rendered by this form of growth. Thus, there is a need to understand the molecular basis of biofilm formation to enable the development of new therapeutics.

The *S. aureus* fibronectin-binding protein FnBPA is a cell wall anchored adhesin also able to bind fibrinogen. Fibronectin binding is mediated by 11 disordered binding repeats (FnBRs) via the tandem β -zipper mechanism and fibrinogen binding by the N2N3 subdomains of FnBPA's A domain. Consequently, these regions of FnBPA are well characterised structurally. The A domain contains a third subdomain N1, which to date is poorly characterised and has no attributed function. The A domain of FnBPA is necessary for an FnBP-mediated mechanism of *S. aureus* biofilm formation. However, the molecular basis of FnBP-mediated biofilm formation is not understood. The main foci of this work are to determine the structure and function of the N1 subdomain and establish the molecular basis of FnBP-mediated biofilm formation.

Nuclear magnetic resonance (NMR) spectroscopy was used to demonstrate the N1 subdomain is intrinsically disordered, but exhibits secondary structure propensity in the C-terminal region. A range of pull-down experiments identified a novel function of FnBPA's A domain in the adherence to human endothelial cells. Adherence to host vasculature is potentially an important step in *S. aureus* bacteraemia and the involvement of N1 is the first evidence of functionality in this domain.

A proposed mechanism of protein-mediated biofilm formation involves the chelation of Zn^{2+} to form intercellular protein dimers. The ability of FnBPA's A domain to dimerise in a Zn^{2+} -dependent manner was investigated and despite forming dimers in conditions mimicking those known to induce FnBP-mediated biofilm formation, the concentration of Zn^{2+} required far exceeded physiological concentrations. Therefore,

interactions between the A domain and other biofilm matrix components were investigated. It was found that N1 interacts with wall teichoic acids, representing new insight into protein-mediated biofilm formation mechanisms and a novel function of the N1 subdomain.

Table of Contents

Abstract	2
List of Figures	8
List of Tables	11
Acknowledgements	12
Author's Declaration	13
1 Introduction	14
1.1 Staphylococcus aureus	14
1.2 <i>S. aureus</i> Surface Protein Adhesins	15
1.2.1 The Fibronectin Binding Proteins	16
1.3 Bacterial Biofilms	20
1.3.1 Staphylococcal Biofilm Formation Mechanisms	21
1.3.2 Biofilm Maturation and the Biofilm Matrix.....	23
1.3.3 Role of the FnBPs in Biofilm Formation	24
1.4 Intrinsically Disordered Proteins	25
1.5 Nuclear Magnetic Resonance Spectroscopy	27
1.5.1 The NMR experiment	29
1.5.2 Three-dimensional experiments and Resonance Assignment.....	31
1.5.3 Secondary Structure Determination	37
1.6 Aims.....	39
2 Materials and Methods	40
2.1 Materials	40
2.1.1 Bacterial Strains	40
2.1.2 Bacterial Culture Media	40
2.1.3 Protein Expression Vectors.....	42
2.1.4 HUVECs	42
2.1.5 Buffer Solutions	42
2.1.6 Native Proteins.....	43
2.1.7 Fibrinogen From Human Plasma	43
2.1.8 Fibronectin From Human Plasma	44
2.2 Methods in Molecular Biology	44
2.2.1 Preparation of Chemically Competent Cells.....	44
2.2.2 Transformation of Competent Cells.....	44
2.2.3 Preparation of Plasmid DNA	45
2.2.4 Agarose Gel Electrophoresis.....	45

2.2.5	Polymerase Chain Reaction (PCR)	45
2.2.6	Restriction Digest.....	46
2.2.7	Ligation	47
2.2.8	Construct Validation	47
2.3	Protein Expression and Purification.....	47
2.3.1	Expression of Unlabelled Proteins	47
2.3.2	Expression of ¹⁵ N and ¹³ C, ¹⁵ N Labelled Proteins.....	47
2.3.3	Expression of ² H, ¹³ C, ¹⁵ N N1	48
2.3.4	Expression of Selectively Unlabelled ² H, ¹³ C, ¹⁵ N N1.....	48
2.3.5	<i>E. coli</i> Cell Lysis.....	49
2.3.6	Purification of His ₆ -tagged Proteins.....	49
2.3.7	Purification of GST-tagged Proteins.....	50
2.3.8	Isolation of Wall Teichoic Acids	51
2.4	Biochemical Methods	52
2.4.1	SDS-PAGE	52
2.4.2	Native-PAGE	52
2.4.3	Determination of Protein Concentration	53
2.4.4	GST Pull-down Assays	53
2.4.5	His ₆ -tag Pull-down Assay	54
2.4.6	Affinity Column Pull-down Assay	54
2.4.7	HUVECs Adhesion Assay	54
2.4.8	Western Blot Analysis of HUVECs Extracts.....	55
2.4.9	Gel Electromobility shift Assay (EMSA)	55
2.5	Biophysical Methods	55
2.5.1	Circular Dichroism (CD)	55
2.5.2	Size-exclusion Chromatography Multi Angle Laser Light Scattering (SEC-MALLS) 56	
2.6	Methods in NMR Spectroscopy.....	56
2.6.1	Sample Preparation	56
2.6.2	Data Acquisition	57
2.6.3	Data Processing.....	59
2.6.4	Spectral Referencing.....	59
2.6.5	Resonance Assignment	60
3	Expression and Purification of Recombinant FnBPA Fragments.....	61
3.1	Introduction.....	61
3.2	Aims.....	61

3.3	Recombinant FnBPA Fragments	62
3.4	Expression of His ₆ -tagged rFnBPA Proteins	63
3.4.1	Unlabelled Protein Expression.....	63
3.4.2	Uniformly-Labelled rFnBPA Expression	64
3.5	Purification of His ₆ -tagged rFnBPA Proteins	64
3.5.1	Expression and Purification of ² H, ¹³ C, ¹⁵ N N1	68
3.5.2	Expression of Selectively Unlabelled ² H, ¹³ C, ¹⁵ N N1.....	68
3.6	Expression and Purification of GST-Fusion Proteins	72
3.6.1	Molecular Biology of GST-rFnBPA Constructs.....	72
3.6.2	Expression and Purification of GST-rFnBPA Proteins.....	72
3.7	Summary	73
4	Structural Characterisation of N1.....	75
4.1	Introduction.....	75
4.2	Aims.....	76
4.3	Sequence Analysis of N1	76
4.4	Preliminary NMR Analysis of N1	78
4.4.1	pH Does Not Affect the Conformation of N1	81
4.4.2	N1 is Disordered in the A Domain.....	82
4.4.3	Circular Dichroism of the A Domain.....	87
4.5	Resonance Assignment of N1	89
4.5.1	Deuterated N1 Spectra	90
4.5.2	Selective Unlabelling of N1	93
4.5.3	Resonance Assignment Procedure	100
4.6	Secondary Chemical Shift Analysis.....	105
4.6.1	Secondary Structure Propensity of N1	109
4.7	Summary	110
5	Interactions Between N1 and Host Ligands.....	114
5.1	Introduction.....	114
5.2	Aims.....	114
5.3	Pull-down Experiments.....	115
5.3.1	FnBPA A Domain Appears to be Proteolytically Cleaved in Plasma.....	116
5.3.2	His-tag Pull-down Assays.....	119
5.3.3	Recombinant Proteins are Active in Plasma	120
5.4	The Role of Surface Proteins in Host Colonisation	123
5.4.1	fA Binds to HUVECs.....	124
5.5	Summary	125

6	The Role of FnBPA in <i>S. aureus</i> Biofilm Formation.....	129
6.1	Introduction.....	129
6.2	Aims.....	131
6.3	fA Dimerises in a pH and Zn ²⁺ Dependent Manner.....	131
6.3.1	The Physiological Zn ²⁺ Concentration is Too Low to Induce Dimerisation	136
6.4	The Structural Role of eDNA in <i>S. aureus</i> Biofilms	136
6.5	Teichoic Acids	140
6.5.1	Purification of <i>S. aureus</i> WTAs.....	142
6.5.2	Preliminary Study of N1 and WTA Interaction	143
6.5.3	Further WTA Purification and Characterisation	145
6.5.4	N1 Interacts with WTA in a Residue Specific Manner.....	148
6.6	Summary	153
7	Discussion.....	157
7.1	Structural Characterisation of N1	157
7.1.1	Resonance Assignment of N1	160
7.1.2	N-terminal Disorder in <i>S. aureus</i> Surface Proteins.....	163
7.2	A Domain Interactions with Host Factors.....	166
7.3	FnBPA-mediated Biofilms.....	169
7.4	Role of Disorder in N1 Function.....	172
	References.....	173

List of Figures

Figure 1 Structural schematics of FnBPA and FnBPB	17
Figure 2 Fibronectin binding by FnBPA.....	18
Figure 3 Crystal structures of FnBPA ligand interactions.....	29
Figure 4 The four stages of staphylococcal biofilm formation.....	21
Figure 5 Mechanisms of biofilm primary attachment.....	22
Figure 6 The orientation of spins in a magnetic field.....	28
Figure 7 The basic NMR experiment.....	30
Figure 8 The basic HSQC experiment.....	31
Figure 9 Sequential assignment of protein spectra using 3D spectra	33
Figure 10 Effect of deuteration on the ^1H , ^{15}N spectrum of a 36 kDa protein.....	36
Figure 11 Recombinant FnBPA fragments.....	62
Figure 12 <i>E. coli</i> growth curves during the expression of rFnBPA proteins.....	64
Figure 13 Purification of His ₆ -N1.	65
Figure 14 Anion exchange chromatography of AF1.....	67
Figure 15 Size-exclusion chromatography of N2N3.....	67
Figure 16 Purification of selectively unlabelled ^2H , ^{13}C , ^{15}N N1 proteins.	69
Figure 17 Purification of GST-N1N2.....	73
Figure 18 Schematic Representation of FnBPA.....	75
Figure 19 Presence of Order and Disorder Promoting Residues in N1.....	76
Figure 20 Predicted Disorder in N1.....	77
Figure 21 NMR spectra of β -sheet and disordered proteins.....	79
Figure 22 1D ^1H NMR Spectrum of N1.....	79
Figure 23 ^1H , ^{15}N HSQC spectra of N1.....	81
Figure 24 ^1H , ^{15}N HSQC spectra of N1 over a range of pH's.....	82
Figure 25 Comparison of the ^1H , ^{15}N HSQC spectra of N1 and N1N2.....	85
Figure 26 ^1H , ^{15}N HSQC spectra of fA and N2N3.....	86
Figure 27 CD spectra of FnBPA A domain constructs.....	88
Figure 28 Resonance overlap in NMR spectra of ^{13}C , ^{15}N N1.....	90
Figure 29 1D ^1H spectra of ^{13}C , ^{15}N N1 and ^2H , ^{13}C , ^{15}N N1.....	92
Figure 30 ^1H , ^{15}N HSQC spectra of ^{13}C , ^{15}N N1 and ^2H , ^{13}C , ^{15}N N1.....	92
Figure 31 Improvement in the HNC0 spectrum following deuteration.....	92
Figure 32 2D ^1H , ^{15}N HNC0 plane of $^{\text{u}}\text{Ile}$	97

Figure 33 2D ^1H , ^{15}N HNC0 plane of $^{\text{u}}\text{Val/Leu}$	98
Figure 34 2D ^1H , ^{15}N HNC0 plane of $^{\text{u}}\text{Pro}$	99
Figure 35 Sequential assignment of N1.....	101
Figure 36 Sequence corrected secondary chemical shifts of N1.....	108
Figure 37 Secondary structure propensity scores for residues in N1.....	110
Figure 38 Schematic representation of a GST pull-down assay.....	116
Figure 39 SDS-PAGE analysis of FnBPA A domain GST pull-down assay.	117
Figure 40 SDS-PAGE analysis of the second A domain GST pull-down assay.....	118
Figure 41 SDS-PAGE analysis of A domain His ₆ -tag pull-down.....	120
Figure 42 SDS-PAGE analysis of affinity column pull-downs from Fn/Fg solutions	122
Figure 43 SDS-PAGE analysis of affinity column plasma pull-downs.....	123
Figure 44 Western blot analysis of the HUVECs FnBPA pull-down experiment.	125
Figure 45 The zinc-zipper model for Aap and SasG mediated biofilm formation.	130
Figure 46 SEC-MALLS analysis of oligomeric states of FnBPA A domain proteins.	133
Figure 47 SEC-MALLS analysis of N2N3 at different pHs.....	134
Figure 48 Cell accumulation in an FnBPA mediated biofilm facilitated by Zn ²⁺ - dependent fA dimerisation.....	135
Figure 49 Gel electromobility shift assay for N1 0.4 μM	137
Figure 50 ^1H , ^{15}N HSQC spectrum of N1 in the presence and absence of DNA....	138
Figure 51 ^1H , ^{15}N HSQC spectrum of N1 in the presence and absence of DNA and Zn ²⁺	139
Figure 52 Chemical structures of teichoic acids from <i>S. aureus</i>	141
Figure 53 Native-PAGE of crude WTA extract.....	143
Figure 54 ^1H , ^{15}N -HSQC of ^{15}N N1 in the absence and presence of WTAs.....	144
Figure 55 Size exclusion chromatography purification of crude WTA extracts.....	146
Figure 56 NMR spectra of purified WTAs.....	147
Figure 57 SEC-MALLS analysis of purified WTAs.	148
Figure 58 ^1H , ^{15}N HSQC spectra of N1 with increasing concentration of WTAs,..	150
Figure 59 ^1H , ^{15}N HSQC spectra of N1 with increasing concentration of WTAs.	151
Figure 60 N1 chemical shift perturbation map following the addition of WTAs.	152
Figure 61 FnBPA/eTA-mediated model of <i>S. aureus</i> biofilm formation.....	153

Figure 62 Secondary structure prediction of N2N3.....	158
Figure 63 The domain organisation of FnBPA and other cell-wall anchored <i>S. aureus</i> surface proteins.....	164
Figure 64 Sequence alignment of N1 subdomains.....	165
Figure 65 1D ^1H NMR spectra of disordered N-terminal subdomains.....	166

List of Tables

Table 1 Commonly used three-dimensional triple resonance experiments	35
Table 2 Bacterial strains.....	40
Table 3 Composition of <i>E coli</i> and <i>S. aureus</i> culture media.....	41
Table 4 Protein expression vectors.....	42
Table 5 Composition of buffers.....	42
Table 6 Composition of PCR.....	45
Table 7 PCR primers.....	46
Table 8 PCR cycling program.....	46
Table 9 Growth media supplements to selectively unlabel amino acids.....	49
Table 10 Pulse-program and basic parameters for two- and three- dimensional experiments recorded for this work.....	58
Table 11 Summary of rFnBPA properties and isotope incorporation.....	71
Table 12 Deconvolution of FnBPA CD spectra.....	87
Table 13 Metabolic precursors for specifically unlabelling amino acids.....	96
Table 14 Chemical shift assignments of N1	102
Table 15 Chemical shift index ranges	106

Acknowledgements

Firstly, I wish to thank my supervisor Prof. Jennifer Potts for allowing me to undertake my PhD under her supervision and for the continual advice, support and encouragement. I thank my training advisory panel, Dr. Gavin Thomas and Dr. Christoph Baumann, for their useful advice and suggestions throughout the course of my PhD. I would also like to thank the British Heart Foundation for their financial support.

I would like to thank the members of the Potts group Dr. Richard Jones, Dr. Gemma Harris, Dr. Fiona Whelan, Judith Hawkhead and Adrian Speakman. I would especially like to thank Dr. Dominika Gruzka for getting me started in the early days and many useful conversations ever since, Dr. Kate Atkin for having the patience to trawl through some terrible early drafts and Dr. Vaclav Stemberk for the motivational chats during his time squatting in my house. I thank Dr. David Williamson for his help in NMR spectroscopy and Dr. Michael Plevin for advice in assigning unstructured proteins. For general help with biophysical techniques, I would like to thank Dr. Andrew Leech. I thank Dr. Joan Geoghegan for her *in vitro* experiments and useful comments and advice. I would also like to thank past and current members of the Barilla and Kleanthous groups for their help and advice.

I would like to thank my friends, particularly Robin Hulme, Gemma Ormerod and Jonathan Briggs, and the entire of Wasteful Pedro FC for entertainment and distractions when I needed them, and a good number when I didn't.

Finally, I would like to thank my parents, Nicola and Stephen, and my brothers George and William, to whom I am immeasurably indebted for all their support throughout my studies and I dedicate this thesis to them.

Author's Declaration

I declare that all of the work presented in this thesis is my own, original work unless explicitly stated in the text.

Andrew Stephen Brentnall, October 2012

1 Introduction

1.1 *Staphylococcus aureus*

Staphylococcus aureus (*S. aureus*) is a Gram positive bacterium and an important organism due to its interactions with humans. Staphylococcal growth is characterised by the formation of grape-like cellular clusters, and this distinct architecture is responsible for the name of the genus. Cell division in staphylococci occurs in alternating perpendicular planes, with irregular or incomplete separation of daughter cells resulting in the observed clumps of cells (Tzagoloff and Novick, 1977). *S. aureus* is distinguishable from other staphylococci by a yellow pigmentation.

A commensal pathogen, *S. aureus* persistently colonises the anterior nares, pharynx or damaged skin of approximately 20% of the population, with a further 60% being intermittent carriers (Noble et al., 1967, Rippon and Vogelsang, 1956). The risk of infection is minimal in non-carrying healthy individuals, whereas persons colonised by *S. aureus* are more susceptible to infections by their carried strain (Foster, 2004, von Eiff et al., 2001). Being an opportunistic pathogen, *S. aureus* takes advantage of impaired host immune responses causing a variety of conditions ranging in severity from minor skin lesions to life threatening illnesses such as toxic shock syndrome, necrotising pneumonia and infective endocarditis (Wenzel and Perl, 1995, Lowy, 1998).

Resistance to antimicrobial agents is an important facet of *S. aureus*' virulence and a major clinical concern, particularly the rate at which resistance is acquired (Diekema et al., 2001, Yamazumi et al., 2001). In the United States, methicillin resistant *S. aureus* (MRSA) is a leading cause of death from infection by a single pathogen (Klevens et al., 2007). MRSA strains harbour genes encoding β -lactamases, rendering β -lactam antimicrobials obsolete as therapeutic agents (de Lencastre et al., 1991). Vancomycin was introduced to combat infections caused by MRSA and in a relatively short period vancomycin-resistant strains (VRSA) were isolated from post-operative patients that had developed infections (Hiramatsu et al., 1997). This highlights the ability of *S. aureus* to adapt quickly to novel antimicrobials.

S. aureus is a major cause of nosocomial infections, with the mortality rate of such infections, particularly those caused by MRSA, ranging from 20 to 30% (Mylotte et al., 1987, Julander, 1985, Bryan et al., 1984). Approximately one third of wound infections following cardiac surgery are attributed to MRSA (Jonkers et al., 2003, Bandyk, 2008). Such surgical site infections (SSIs) often cause complications during the implantation of prostheses, commonly resulting in the removal and re-implantation of the device. The outcome of this treatment is a lengthier hospitalisation period for patients and a significant financial burden for health-services (de Lissovoy et al., 2009). In recent times there has been a notable rise in the number of community associated MRSA (CA-MRSA) infections (Deleo et al., 2010, Hersh et al., 2008). Contrasting hospital acquired MRSA (HA-MRSA) infections that are traditionally linked to patients with established risk factors, CA-MRSA infections develop in otherwise healthy individuals (Naimi et al., 2003). Despite being deemed more transmissible and virulent than HA-MRSA, CA-MRSA is more susceptible to non- β -lactam antibiotics (Baba et al., 2002).

S. aureus possesses an arsenal of virulence factors that facilitate host colonisation and immune evasion (Chesney et al., 1984, Uhlen et al., 1984, Watanabe et al., 2007, O'Brien et al., 2002a). During the establishment of blood borne infections bacteria are in close contact with endothelial cells throughout the vasculature, and a number of surface components promote adherence to the host cell surface (O'Brien et al., 2002b, Roche et al., 2003, Weidenmaier et al., 2005, Massey et al., 2001). Working in tandem with adhesion molecules other virulence factors provide resistance to phagocytosis, enzymes promoting host cell lysis and bacteraemia, and other toxins damage host tissues (Claro et al., 2011, Fitzgerald et al., 2006, Peterson et al., 1977, Thurlow et al., 2011). Understanding the molecular mechanisms by which the aforementioned virulence factors are utilised could provide potential therapeutic targets.

1.2 *S. aureus* Surface Protein Adhesins

S. aureus expresses a number of protein adhesins that recognise host extracellular matrix (ECM) components (Massey et al., 2001, McDevitt et al., 1994, Patti et al., 1994b, Roche et al., 2004). Collectively these adhesins are referred to as

MSCRAMMs (microbial surface components recognising adhesive matrix molecules) (Patti et al., 1994a). These surface proteins do not exclusively recognise ECM components and also promote adhesion to endothelial cells. Clumping factor B (ClfB) (O'Brien et al., 2002b), iron-regulated surface determinant A (IsdA) (Clarke et al., 2004, Corrigan et al., 2009) and SasG (Roche et al., 2003) are utilised in the colonisation of the anterior nares. Though a large proportion of MSCRAMMs recognise only one ECM component, certain proteins are bi-functional, notably the fibronectin-binding proteins (FnBPs), which bind fibronectin (Fn) (Schwarz-Linek et al., 2003, Froman et al., 1987) and fibrinogen (Fg) (Keane et al., 2007b, Wann et al., 2000). Protein adhesins are often important in establishing infections (Arrecubieta et al., 2006, Edwards et al., 2010, Piroth et al., 2008). Expression of clumping factor A (ClfA) and fibronectin-binding protein A (FnBPA) on the otherwise non-pathogenic *Lactococcus lactis* (*L. lactis*) increased its capacity to colonise a host (Que et al., 2005). Synergistic binding of Fn and Fg by *L. lactis* led to a more efficient infection, suggesting that proteins recognising multiple targets are particularly effective virulence factors (Piroth et al., 2008).

1.2.1 The Fibronectin Binding Proteins

The FnBPs of *S. aureus* have been implicated in the pathogenesis of a number of diseases including infective endocarditis (Fitzgerald et al., 2006) and the development of sepsis (Edwards et al., 2010, Shinji et al., 2011). Many strains express two FnBPs, FnBPA and FnBPB, encoded by the closely linked genes *fnbA* and *fnbB* (Jonsson et al., 1991). Although some strains express only *fnbA*, the majority of clinical isolates also contained the gene *fnbB* encoding FnBPB (Peacock et al., 2000). FnBPA and FnBPB share very a similar domain organisation and have high sequence identity, 45% and 95% for the N-terminal Fg-binding and the Fn-binding regions respectively, (Figure 1) (Burke et al., 2010, Loughman et al., 2008). The N-terminus of FnBPA comprises a short signalling peptide, required to ensure expressed proteins are directed to the cell surface (DeDent et al., 2008), and the Fg-binding A domain (Wann et al., 2000, Keane et al., 2007a). Adjacent to the A domain is the Fn-binding region, containing 11 highly conserved Fn-binding repeats (FnBRs) (Schwarz-Linek et al., 2003, Meenan et al., 2007). The C-terminus is

composed of a proline rich repeat region, wall and membrane spanning domains, and a LPETG motif for sortase-catalysed cell wall anchoring (Novick, 2000).

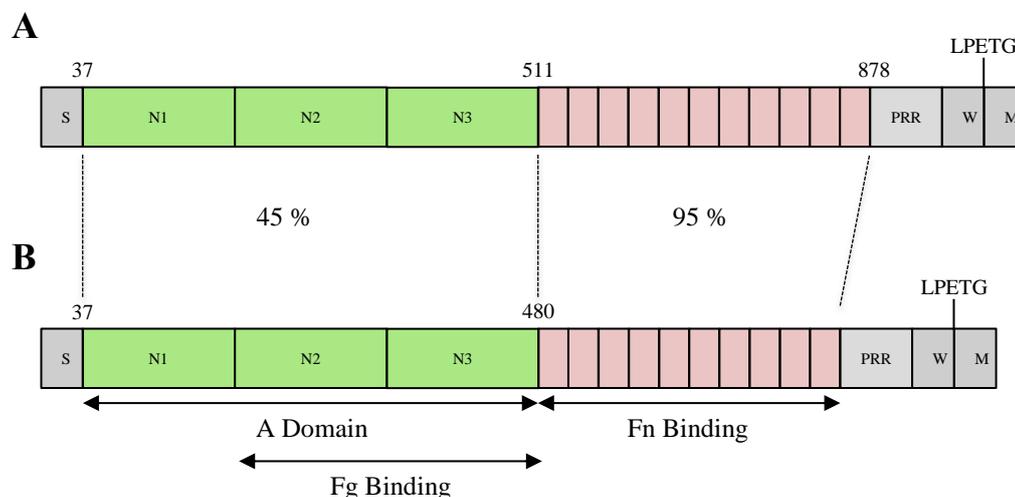


Figure 1 Structural schematics of FnBPA and FnBPB. FnBPA (A) and FnBPB (B) have very similar domain organisations. Both comprise a short signalling peptide (S) at the N-terminus, a Fg-binding A domain, multiple FnBRs, a proline rich repeat region (PRR), wall- (W) and membrane- (M) spanning domains. Covalent attachment to the cell surface is achieved via a LPETG motif. The percentage identity between regions of the FnBPs is indicated. Residue numbers correspond to UniProt entries P14738 (FnBPA) and Q53682 (FnBPB).

Interactions between FnBPA and Fn and Fg play a pivotal role in a number of processes. The internalisation of *S. aureus* into host cells (Edwards et al., 2010), platelet activation (an important step in the development of infective endocarditis) (Fitzgerald et al., 2006) and the adhesion of *S. aureus* to implanted ventricular assist devices (Arrecubieta et al., 2006) are all mediated by FnBPA-Fn and FnBPA-Fg interactions. Upon binding of Fn, disordered regions of bacterial proteins form additional β -strands antiparallel to the Fn F1 modules E-strands, in a mechanism known as the tandem β -zipper (Schwarz-Linek et al., 2003, Meenan et al., 2007). Recently solved crystal structures of the $^2\text{F1}^3\text{F1}$ and $^4\text{F1}^5\text{F1}$ module pairs of Fn's N-terminal domain (NTD) with FnBR peptides, show that FnBPA exploits this particular mechanism of binding (Figure 2) (Bingham et al., 2008). A significant entropic penalty exists upon complex formation, following the transition from a disordered to an ordered conformation. A high percentage of the FnBR residues form

an extensive interface with F1 modules. Thus, unfavourable entropic changes are offset by a favourable enthalpic contribution.

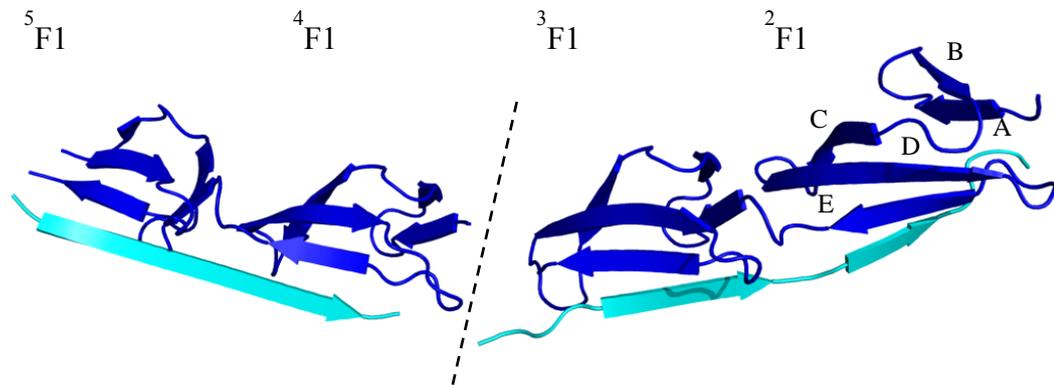


Figure 2 Fibronectin binding by FnBPA. On binding to Fn (navy) the disordered FnBRs (cyan) form an additional β -strand antiparallel to the F1 modules E-strand. Adapted from Bingham *et al.* (Bingham *et al.*, 2008).

The N2N3 subdomains of FnBPA's N-terminal A domain comprise the Fg-binding site, a region that has also been shown to bind elastin (Keane *et al.*, 2007b, Roche *et al.*, 2004, Keane *et al.*, 2007a). A recently solved crystal structure of N2N3 (Figure 3A) (Stemberk, manuscript in preparation) shows the independently folded subdomains adopt immunoglobulin type conformations, similar to other *S. aureus* Fg-binding surface proteins (ClfA and ClfB) (Ganesh *et al.*, 2011a, Deivanayagam *et al.*, 2002). The Fg-binding site is located in the mostly hydrophobic cleft between N2 and N3. Upon binding, the C-terminus of the Fg γ -chain forms an additional β -strand parallel to the G' strand of N3 (Figure 3B) (Stemberk, manuscript in preparation).

To date no function has been attributed to the N1 subdomain. In a pair-wise sequence alignment of FnBPA A domains from a number of *S. aureus* strains, N1 was found to be the most conserved subdomain (Loughman *et al.*, 2008, Keane *et al.*, 2007b). On average N1 is approximately 90% conserved, compared to 75% and 60% in N2 and N3, respectively. N1 appears to have evolved separately from N2N3 as high sequence identity in N2N3 does not necessarily correlate with similar identity in N1. The Fg-binding capacity of FnBPA is unaffected by N1 (Stemberk, manuscript in preparation), thus it is likely N1 has a role separate from N2 and N3. One hypothesis is that during the early stages of infection N1 may protect N2 and N3

from proteolysis (Burke et al., 2010). N1 is speculated to be less readily recognised by the host immune system rendering protection to N2 and N3, thus preserving FnBPA's Fg-binding capacity (Burke et al., 2010). A second proposed role is in FnBP-mediated *S. aureus* biofilm formation.

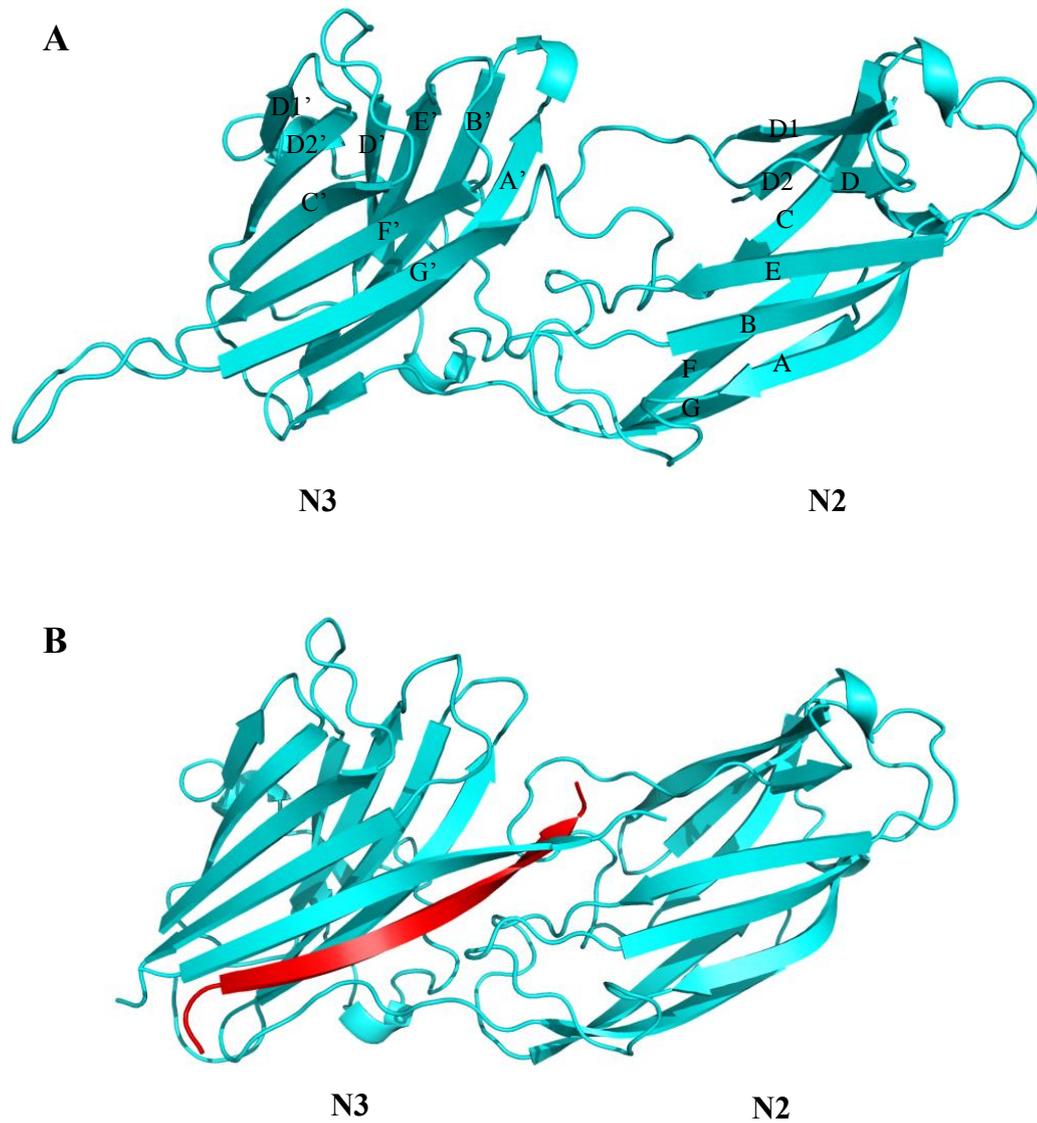


Figure 3 Crystal structures of FnBPA ligand interactions. Crystal structures of the N2N3 (cyan) subdomains from FnBPA in the *apo* form (A) and bound to Fg (red) (B).

1.3 Bacterial Biofilms

The current definition of a bacterial biofilm is a structured community of cells enclosed in a self-produced polymeric matrix and adhered to an inert or living surface (Costerton et al., 1995). Biofilms have been observed in nature for some time, often on moist surfaces as a slimy coating. By forming biofilms bacteria generate an environment that provides protection during growth in hostile conditions (Darouiche, 2004, Anwar et al., 1990). Recently, understanding the mechanisms of biofilm formation has become more important in a clinical setting as commensals such as *Staphylococcus epidermidis* (*S. epidermidis*) and *S. aureus* form biofilms following infection (Cha et al., 2010, Fey and Olson, 2010). The protective nature of a biofilm increases the difficulty of treating infections and this is attributed to two primary factors. Firstly, antimicrobial agents may not fully penetrate the biofilm, as the biofilm matrix is known to retard the diffusion and as a result bacteria situated deeper within the biofilm are less likely to be affected (Hoyle et al., 1992). Secondly, bacteria in a biofilm are in a semi-dormant state and in this mode of growth are less susceptible to antimicrobial agents (Kinniment and Wimpenny, 1992). An additional complication is that low levels of antibiotics can act to stimulate biofilm formation and growth (Hsu et al., 2011, Kaplan, 2011, Kaplan et al., 2011). Therefore, incomplete treatment or sub-lethal doses of antibiotics may be advantageous to the bacteria, even strains susceptible to a particular course of treatment. Typically *S. aureus* and *S. epidermidis* utilise polymeric intercellular adhesion (PIA), extracellular DNA (eDNA), proteins and teichoic acids as structural components in the biofilm matrix (Izano et al., 2008b, Conrady et al., 2008, Kaplan et al., 2011, Sadovskaya et al., 2004). A recent study has shown that despite being able to kill bacteria with antimicrobials such as lysostaphin, unless the matrix is also destroyed the biofilm will remain intact (Wu et al., 2003). *S. aureus* and *S. epidermidis* are adept at forming biofilms on indwelling medical devices, which can significantly hinder medical treatments (Mack et al., 2006, Donlan, 2001). Therefore, understanding the molecular mechanisms bacteria employ to form biofilms is an important area of research with significant clinical implications.

1.3.1 Staphylococcal Biofilm Formation Mechanisms

Bacterial biofilm formation is classically described as a four-stage process; adherence, accumulation, maturation and dispersal (Figure 4). Primary attachment by bacteria to a surface can be mediated by a number of factors. Perhaps the most intuitive mechanism employed by *S. aureus* involves utilising adhesins expressed on the cell surface that recognise extracellular matrix components. Following implantation, the surface of a prosthetic device is rapidly coated with host matrix proteins such as fibrinogen and fibronectin (Vaudaux et al., 1989). As a result the surfaces of prostheses become ideal environments for *S. aureus* to colonise via specific interactions between its surface adhesins (FnBPA, FnBPB and ClfA of *S. aureus*, for example) and the deposited host factors (Figure 5). A second method enables adherence to unconditioned surfaces and is mediated by DNA, released through the action of the major autolysin Atl, and dependent on electrostatic interactions (Heilmann et al., 1997, Qin et al., 2007). However, this method of primary attachment is less well understood and appears to be governed by the overall mechanism of biofilm formation (Houston et al., 2011).

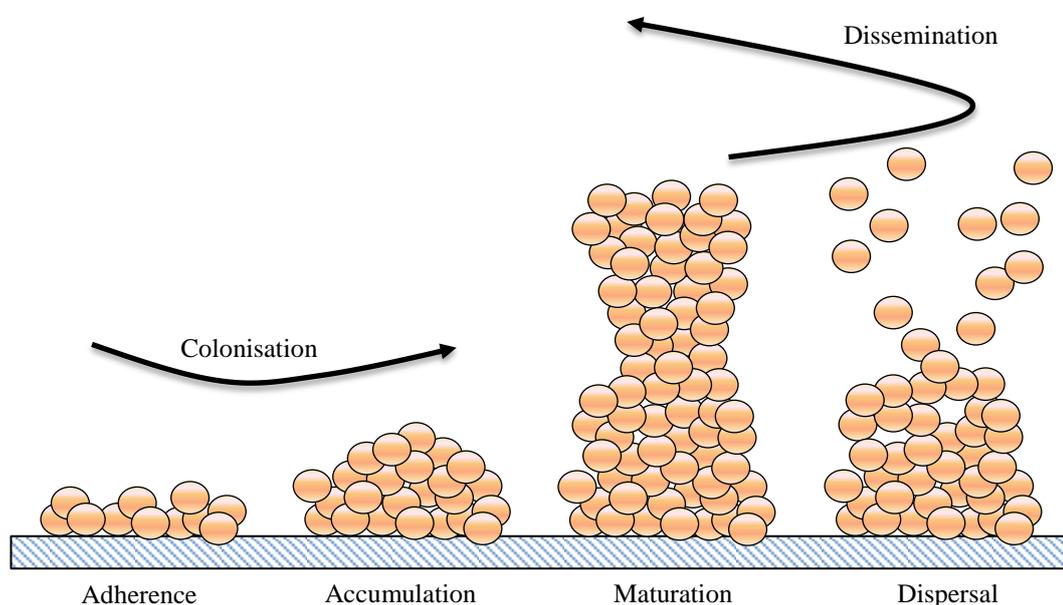


Figure 4 The four stages of staphylococcal biofilm formation. Biofilms growth is categorised into four stages; attachment, accumulation, maturation and dispersal. Adapted from (Otto, 2004).

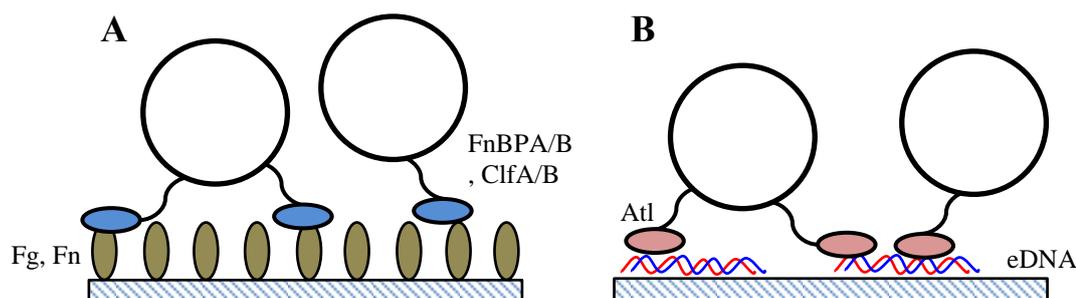


Figure 5 Mechanisms of biofilm primary attachment. *S. aureus* utilises specific interactions between surface proteins and deposited host factors (A) or non-specific electrostatic interactions (B) to mediate primary attachment.

The accumulation phase of biofilm formation is typically separated into two distinct mechanisms mediated by a polysaccharide intercellular adhesin (PIA) and surface proteins (Heilmann et al., 1996, O'Gara, 2007, Geoghegan et al., 2010). PIA, or poly-N-acetyl glucosamine (PNAG), is synthesised by the *ica* gene products (Mack et al., 1996a, Mack et al., 1996b). Initially identified in biofilm forming strains of *S. epidermidis*, the *ica* operon is composed of four genes (*icaA*, *icaB*, *icaC* and *icaD*). The presence of the *ica* locus was subsequently identified in *S. aureus*, with the gene products sharing high amino acid identity across strains (Cramton et al., 1999). The *ica* operon is conserved in a significant proportion of *S. aureus* clinical isolates and is one of seven factors associated with increased invasiveness and virulence (Kropec et al., 2005, Peacock et al., 2002). IcaA is a transmembrane N-acetylglucosaminyltransferase requiring IcaD for optimal activity. Coexpression of *icaAD* with *icaC* leads to the synthesis of long PNAG chains. Transfer of these polymers to the cell surface is likely achieved with IcaC involvement (Gerke et al., 1998). Partial deacetylation of PIA/PNAG is carried out by IcaB, generating a net positive charge throughout the polymer (Vuong et al., 2004). The partial deacetylation of PIA is crucial for the accumulation of bacteria, presumably to interact with negatively charged cell surfaces (Vuong et al., 2004).

The role proteins play in biofilms is more ambiguous and dependent on growth conditions and the bacterial strain. To date a number of proteins have been demonstrated to facilitate the accumulation phase of biofilm formation. *S. aureus* typically employs LPXTG-anchored surface proteins including Bap, SasG, SasC and

FnBPA/B to mediate cell biofilm formation (Cucarella et al., 2001, O'Neill et al., 2009, Schroeder et al., 2009, Corrigan et al., 2007). The involvement of proteins was identified by the dispersal of the biofilm following treatment with proteases or through genomic analysis and comparison with proteins known to mediate biofilm formation (Corrigan et al., 2007, O'Neill et al., 2007). Little is known about the molecular mechanisms by which proteins drive cell accumulation though the involvement of zinc has been confirmed in a number of cases (Conrady et al., 2008, Geoghegan et al., 2010). The addition of chelating agents resulted in the dispersal of cells, with the biofilm only re-established on the addition of Zn^{2+} . No other divalent cations had the same effect. A proposed mechanism involving Zn^{2+} is the “zinc-zipper”, wherein Zn^{2+} induces intercellular protein dimerisation and drives cell accumulation (Conrady et al., 2008). However, to date evidence to support this hypothesis has only been found for one protein from *S. epidermidis*.

The superstructure of biofilms also differs depending on the cell accumulation mechanism. In *S. epidermidis* PIA acts as a tether running through the biofilm matrix to which the bacteria adhere resulting in relatively dispersed cells (Schommer et al., 2011). Contrastingly cells growing in protein mediated biofilms form dense bacterial aggregates with little void space apparent between cells (Vergara-Irigaray et al., 2009). This may be indicative of interactions between surface bound molecules on neighbouring cells driving bacterial accumulation.

1.3.2 Biofilm Maturation and the Biofilm Matrix

Mature biofilms are complex environments consisting of cells in multiple metabolic states, and this contributes to the difficulty of clearing biofilm infections (Rani et al., 2007). An intricate response system exists within the biofilm influencing gene expression to optimise biofilm development in response to the bacteria's environment (Beenken et al., 2004). This multipart quorum-sensing system dictates levels of protein expression, cell lysis and shifts from aerobic to anaerobic respiratory mechanisms to mention but a few regulated processes (Marti et al., 2010, Mann et al., 2009, Rani et al., 2007). What is clear is that in mature biofilms the matrix is critical to the stability and survival of the biofilm. It allows intense interaction among cells in close proximity, permitting efficient response to various

stimuli (Flemming and Wingender, 2010). Maintaining the biofilm matrix relies on continual production of its constituents with DNA, teichoic acids, PIA and proteins seemingly ubiquitous throughout staphylococcal biofilms. The role played by each component differs significantly with eDNA a major structural component and PIA a minor one in *S. aureus* biofilms (Izano et al., 2008a). In *S. epidermidis* biofilms, PIA has a much more significant contribution (Izano et al., 2008a, Rohde et al., 2010). A commonality is that both polymers act as diffusion barriers to prevent access to embedded cells. Over time the contribution of a particular component can also change (Mann et al., 2009). For example, in certain *S. aureus* biofilms eDNA is known to be essential for primary attachment and in the initial stages of biofilm growth (Houston et al., 2011, Izano et al., 2008a). In due course the established biofilm becomes less susceptible to DNase activity and more reliant on the contribution of proteinaceous adhesins.

1.3.3 Role of the FnBPs in Biofilm Formation

An *ica*-independent biofilm was identified for clinically relevant MRSA strains and was induced by growth in glucose-supplemented media (O'Neill et al., 2008, Vergara-Irigaray et al., 2009). Biofilm formation was triggered by mild acid stress following the metabolism of glucose. The role of a LPXTG wall-anchored protein was determined through the deletion of sortase, required for cell-wall anchoring, and a subsequent six-fold reduction in biofilm density. Following the systematic deletion of cell wall anchored proteins, it was found that a *fnbAfnbB* double mutant severely disrupted MRSA's ability to form biofilms, implying FnBPA and FnBPB are necessary for biofilm viability. Bacteria were still able to adhere to a surface and consequently the role of the FnBPs is thought to be in cell accumulation. The A domains of FnBPA/B were found to be necessary and sufficient for biofilm development. Deletion of the A domain inhibited biofilm formation, whereas deletion of the FnBR region had no observable effect. The presence of either FnBPA or FnBPB was sufficient to establish biofilm at close to wild-type levels suggesting the proteins can functionally substitute for one another (O'Neill et al., 2008). Lowering the temperature from 37 to 30 °C resulted in reduced biofilm formation, and it has been suggested that this is due to attenuated levels of *fnb* gene expression

(O'Neill et al., 2009). To date there is no proposed molecular mechanism for FnBP-mediated biofilm formation.

1.4 Intrinsically Disordered Proteins

Ordered proteins are typically characterised as having a stable three-dimensional structure, governed by the primary sequence. For a long time it was thought a stable fold was a prerequisite for protein function. More recently there has been a significant shift away from this paradigm with the crucial role of disorder in protein interactions evident throughout nature. Proteins can typically be divided into four structural categories; tightly folded single domains, multi-domain proteins with flexible linker/disordered regions, compact but disordered molten globules and highly extended unstructured states (or random coil) (Uversky, 2003). There are no clearly defined boundaries to indicate whether a protein falls into a particular category; rather a continuum between states exists (Dyson and Wright, 2005). For simplicity intrinsically disordered proteins (IDPs) are generally described as proteins that lack stable secondary or tertiary structure under physiological conditions (Dunker et al., 2002). IDPs sample an ensemble of conformations that continually interconvert and thus have a rugged energy landscape and no clearly defined minimum (Rezaei-Ghaleh et al., 2012). However, proteins rarely behave as true random coils, showing rather a propensity to form local elements of secondary structure (Dyson and Wright, 2005). The regions with transient secondary structure are often stabilised on binding to a ligand.

As mentioned previously, proteins can be wholly or partially disordered. Thus, the classification of IDPs remains somewhat subjective. It is thought approximately 21%, 18% and 30% in bacteria, archaea and eukaryotes, respectively, are IDPs or contain intrinsically disordered regions (IDRs) (Pancsa and Tompa, 2012, Ward et al., 2004). There is an apparent correlation between the number of disordered proteins and complexity of the organism, with more complex organisms expressing a higher percentage of IDPs (Xue et al., 2012). However, this may be an oversimplification and structural disorder cannot simply be attributed to the phylogeny of an organism (Pancsa and Tompa, 2012).

IDPs exhibit low sequence complexity compared to folded proteins, a characteristic that enables the prediction of unstructured regions *in silico* (Bordoli et al., 2007, Oldfield et al., 2005). Typically IDPs contain relatively low numbers of what are commonly termed order-promoting residues such as Val, Leu, Ile, Phe, Trp and Tyr (Dosztanyi et al., 2005b, Garner et al., 1998). The lack of these residues prevents the formation of a stable hydrophobic core and consequently, a stable three-dimensional fold. Instead, IDPs are typically rich in polar and charged residues (Glu, Asp, Pro, Ser, Lys and Gln) (Romero et al., 2001, Midic and Obradovic, 2012). Interactions between the prevalent charged residues in IDPs can significantly impact the protein's dimensions (Marsh and Forman-Kay, 2010, Muller-Spath et al., 2010). Interestingly, the net charge, rather than the number of charges governs the hydrodynamic radius of an IDP, with a higher net charge resulting in increased compaction. Contrastingly, and somewhat counter intuitively, IDPs with a similar number of positive and negative charges are subject to pronounced expansion (Muller-Spath et al., 2010). In addition, the number of proline residues affects an IDPs hydrodynamic radius, with an increased number of proline residues resulting in increased expansion (Marsh and Forman-Kay, 2010).

The properties of IDPs lead to characteristic behaviour in a number of biophysical techniques. IDPs have a large hydrodynamic radius compared to folded proteins of similar size, a feature detectable using size exclusion chromatography. In addition IDPs are easily identifiable using techniques such as nuclear magnetic resonance (NMR) spectroscopy (Mittag and Forman-Kay, 2007) and circular dichroism (CD) (Sigalov et al., 2006), producing unique spectral features (Dyson and Wright, 2004, Tompa, 2002).

IDPs are known to play important roles in a number of processes, such as signalling and transcriptional regulation (Zeng et al., 2011, Kim et al., 2008). The recognition of multiple targets can be crucial to fulfilling such roles. For example, the Cyclin-dependent kinase inhibitor p21^{Cip1} is able to interact with a variety of kinases to regulate numerous aspects of the cell cycle (Kriwacki et al., 1996, Tanaka et al., 2002). Interactions involving IDPs have a significant entropic cost when undergoing a disordered to ordered transition. This is countered by IDPs exposing a large interface, relative to their overall size, for binding. Thus, there is a significant

enthalpic contribution to binding (Dyson and Wright, 2001). As a result, interactions involving IDPs tend to be highly specific but low affinity.

1.5 Nuclear Magnetic Resonance Spectroscopy

Solution state NMR spectroscopy is a powerful technique for probing protein structure and monitoring interactions. Along with crystallography and cryo-electron microscopy, NMR is one of able to elucidate three-dimensional chemical structures at high resolution. NMR can provide an insight into molecular dynamics and the kinetics of interactions. However, despite recent advances in methodology, including polarisation, pulsing and acquisition techniques, and spectrometer technologies NMR is limited by its inherent low sensitivity and requires significant amounts of material relative to other biophysical techniques.

Nuclei possess an intrinsic angular momentum known as spin, denoted with the spin quantum number (I) (Hore, 1995). The magnitude of I possessed by nuclei is dependent on the number of protons and neutrons they comprise, and increases in half integers from 0 to 4. When $I \neq 0$ the nuclei possess a magnetic dipole. Nuclei with I greater than $1/2$ possess electric quadrupole moments arising from non-spherical charge distribution. The lifetime of magnetic states arising from quadrupolar nuclei is typically short-lived and resonance lines broader (Levitt, 2001). As a result they are more difficult to study and subsequent discussions will focus on spin- $1/2$ nuclei, which are most widely used in chemical analysis. The spin angular momentum (I) of a nucleus has a number of permitted directions based on its spin quantum number. In the absence of an applied magnetic field all orientations have the same energy. When an external field (B_0) is applied, this degeneracy is removed and lower energy states become more populated. Spin- $1/2$ nuclei have two permitted orientations, in the same direction as the applied field, or in the opposite direction to the applied field (Figure 6). Spins oriented in the direction of the applied field have a lower energy. The energy gap between these two states (ΔE) depends on the magnitude of the nuclear magnetic moment (Equation 1) and the magnetic field strength. The application of radio frequency (RF) electromagnetic radiation excites the system inducing an inversion in the population of energy states. The frequency of

radiation required to achieve this at a given field strength is therefore known as the resonance frequency (Equation 2) (Keeler, 2005).

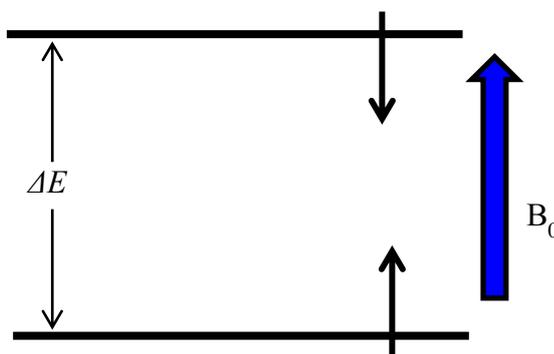


Figure 6 The orientation of spins in a magnetic field. Spin- $1/2$ nuclei have two permitted orientations when an external field is applied. A lower energy orientation in the direction of B_0 and a higher energy orientation in the opposite direction.

$$\mu = \gamma \cdot I$$

Equation 1 Calculating the magnetic moment of a nucleus. The magnetic moment (μ) of a nucleus is directly proportional to I , with the gyromagnetic ratio (γ) of the nucleus the proportionality constant.

$$\nu = \frac{\gamma B_0}{2\pi}$$

Equation 2 Resonance frequency of an isolated nucleus in an applied magnetic field. The resonance frequency (ν) of a nucleus is dependent on the gyromagnetic ratio of the nucleus (γ) and the magnetic field strength (B_0).

Each nucleus (e.g. ^1H , ^{13}C , ^{15}N , ^{19}F) has a characteristic resonance frequency based on its individual magnetic moment. The local chemical environment of a nucleus also has an effect on its resonance frequency within an external field. As a result, spin- $1/2$ nuclei can be used as probes to determine the environments of specific atoms and elucidate the structural characteristics of a particular molecule. By isotopically enriching proteins with the spin- $1/2$ nuclei ^{13}C and ^{15}N , along with using naturally occurring ^1H , NMR can be a powerful technique for probing protein structure and interactions by monitoring changes in the resonance frequency of particular nuclei.

The resonance frequency of a nucleus is primarily determined by its gyromagnetic ratio and the strength of the applied magnetic field (Hore, 1995). However, the local chemical environment, specifically the electron distribution, also contributes to a nucleus' resonance frequency. Application of an external magnetic field induces electron motion, which in turn generates small local magnetic field in the opposite direction to \mathbf{B}_0 . Nuclei situated within the field generated by electron motion are shielded from the external field and resonate at a slightly different frequency. Thus, Equation 3 more accurately describes resonance frequency of a nucleus, where \mathbf{B} is the magnetic field generated by electron motion. This effect, known as chemical shift, enables the identification of individual nuclei based on their unique chemical environment (Hore, 1995).

$$\nu = \frac{\gamma (\mathbf{B}_0 - \mathbf{B})}{2\pi}$$

Equation 3 Resonance frequency of nuclei in a magnetic field. The resonance frequency (ν) of a nucleus is dependent on the gyromagnetic ratio of the nucleus (γ), the applied magnetic field strength (\mathbf{B}_0) and magnetic field generated by electron motion (\mathbf{B}).

1.5.1 The NMR experiment

Magnetic nuclei within a sample are detected by generating and recording a free induction decay (FID) (Figure 7). Firstly, the nuclei are allowed to come to equilibrium within the magnetic field. Secondly, a short high-power RF pulse is applied at 90° to \mathbf{B}_0 generating the FID. The FID is recorded for a particular time (acquisition time) with Fourier transformation of the time-domain signal generating a frequency-domain NMR spectrum (Keeler, 2005).

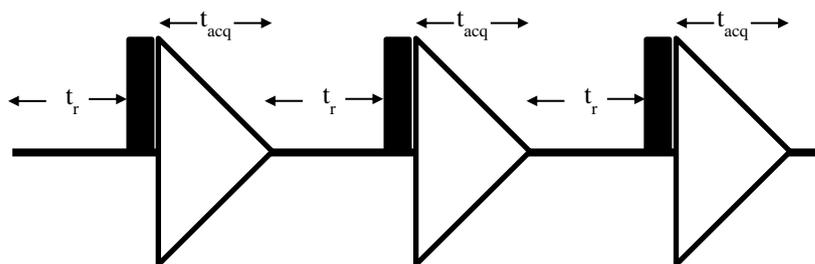


Figure 7 The basic NMR experiment. Following a relaxation period (t_r) a short high power RF pulse is applied and the resultant FID recorded for time t_{acq} . To improve the signal-to-noise ratio the experiment is repeated as necessary.

1.5.1.1 The HSQC Experiment

One-dimensional ^1H NMR spectra of proteins are generally very complicated due to the number of observable nuclei within the system and the resultant signal overlap. Acquiring two dimensional ^1H spectra (e.g. NOESY, ROESY and TOCSY) alleviates some of the spectral congestion, although acquisition of two-dimensional ^1H spectra and subsequent extraction of the requisite information (for example, inter-proton distance restraints) can be time consuming. These problems become particularly pronounced with high molecular weight species. Larger proteins typically have longer correlation times (τ_c), which is a measure of the rate of random molecular motion, both rotational and translational (Keeler, 2005). Peak linewidth is inversely proportional to the spin-spin (T_2) relaxation time. Slower molecular motion results in more efficient T_2 relaxation and correspondingly broader peaks. Thus, the slow tumbling properties of large proteins, and subsequently larger linewidths, coupled with the sheer number of protons present in the system results in a lack of sensitivity and signal degeneracy (Cavanagh, 1996). Labelling proteins with ^{13}C and ^{15}N isotopes allows correlations between these nuclei and ^1H nuclei to be recorded. Such two-dimensional heteronuclear spin correlation (or coherence) spectra can be recorded in a fraction of the time and are often simpler to interpret. Consequently, the ^1H ^{15}N heteronuclear single quantum coherence (HSQC) experiment (Figure 8) has become the most commonly used experiment to obtain simple secondary structure information as well as monitor protein interactions.

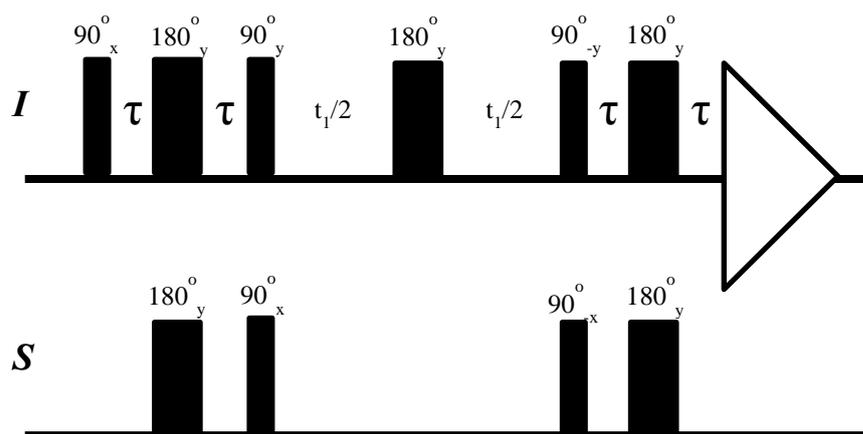


Figure 8 The basic HSQC experiment. First, insensitive nuclei enhanced by polarisation transfer (INEPT) sequence transfers magnetisation from the I spin to the S spin. Following an evolution period (t_1), during which a refocusing pulse is applied to I , a second INEPT sequence is used to transfer S encoded magnetisation back to I for detection.

Each peak in the ^1H , ^{15}N HSQC corresponds to a correlation between covalently linked ^1H and ^{15}N nuclei. As a result all amino acids in the sequence (except proline) give rise to a single resonance attributed to their backbone amide group. Certain side chains also produce observable signals, although this is dependent on the chemical exchange regime. Signal overlap in the ^1H dimension can often be resolved by the increased dimensionality as residues are less likely to have identical ^1H and ^{15}N chemical shifts.

1.5.2 Three-dimensional experiments and Resonance Assignment

Protein NMR spectroscopy was revolutionised by Wüthrich *et al.* following the introduction of techniques to complete the sequential resonance assignment of protein spectra (Wuthrich *et al.*, 1982, Strop *et al.*, 1983, Wagner and Wuthrich, 1982). Principally these techniques were introduced to provide a basis for protein structure determination by NMR and relied on through-bond rather than through-space correlations. Thus, spectra were far less congested and significantly easier to interpret. Sequential assignment techniques were developed further through the introduction of triple-resonance techniques (Ikura *et al.*, 1990). As a result the ^1H , ^{15}N and ^{13}C backbone resonances of fully ^{15}N and ^{13}C labelled proteins can be unequivocally assigned without relying on complex NOESY spectra (Pelton *et al.*,

1991, Powers et al., 1992). NOESY-based spectra are still required for determining inter-residue distance restraints, though the interpretation of such spectra is significantly easier following sequential resonance assignment. In addition to providing the basis for structure determination, resonance assignments are a powerful tool in obtaining sequence specific information from proteins unsuitable for standard structural characterisation, such as IDPs (Dyson and Wright, 2001). By assigning resonances in a spectrum to specific residues, the particular regions of a protein involved in an interaction can be mapped and kinetic parameters of the reaction determined.

Triple-resonance experiments normally correlate $^1\text{H}^{\text{N}}$, $^{15}\text{N}^{\text{H}}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$ and $^{13}\text{C}'$ shifts via one and two bond J -couplings. Typically these experiments give rise to three-dimensional datasets that are in essence a ^1H , ^{15}N HSQC spectrum moderated by a ^{13}C dimension at a frequency related to the ^{13}C nucleus of interest (Cavanagh, 1996). Triple-resonance experiments are often acquired in pairs; the first allows only one correlation to either the intra-residue ^{13}C or that of the preceding residue, the complementary experiment with intra- and inter-residue correlations permitted. The resultant spectra are analysed simultaneously and neighbouring residues can be determined (Figure 9). Acquisition of a number of these pairs of experiments can provide reliable, independent pathways to complete the sequential resonance assignment of a protein without knowledge of the amino acid type. Only proline residues interrupt connectivity's as they lack the amide proton necessary for many of the magnetisation pathways utilised.

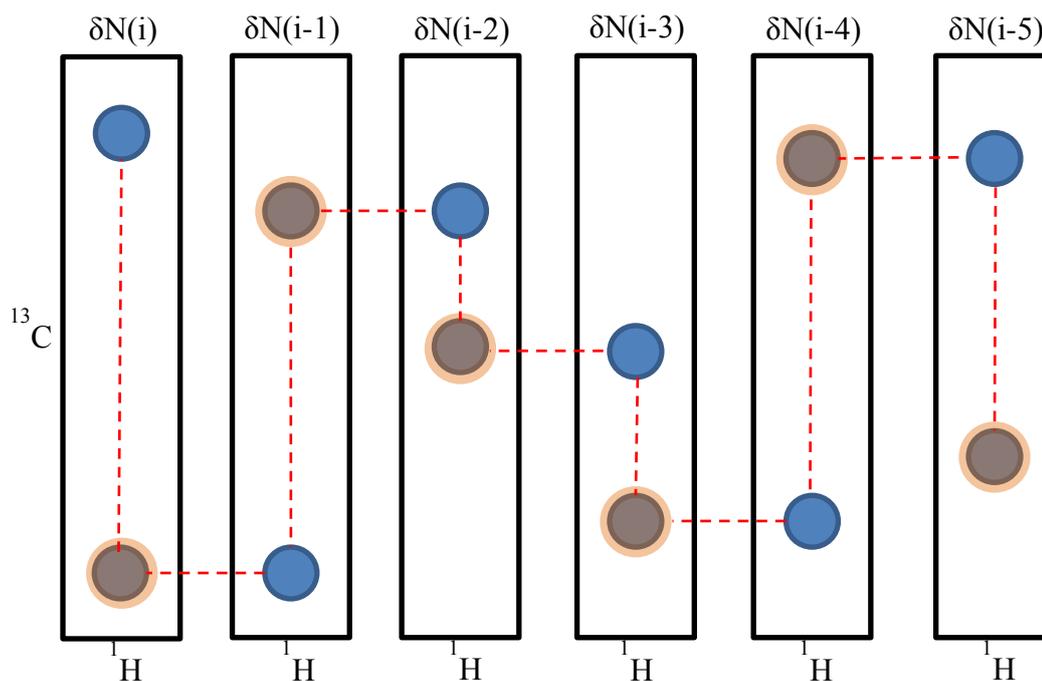
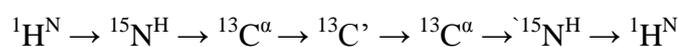


Figure 9 Sequential assignment of protein spectra using 3D spectra. Pairs of 3D triple resonance spectra are analysed simultaneously and the sequential connectivity's determined. A ^{15}N frequency is selected corresponding to a peak in the ^1H , ^{15}N HSQC and the ^1H , ^{13}C planes analysed. The blue spectrum contains peaks arising from connections to ^{13}C nuclei from the same and preceding residues, the orange only from the preceding residue. By overlaying spectra the spin system to which each peak belongs can be deduced, and by finding a resonance with the same ^{13}C frequency in a different ^{15}N plane the sequential assignment can be carried out. In the example above the ^{13}C resonance arising from the previous residue is used to determine connectivity's.

The nomenclature of triple-resonance experiments is derived from the magnetisation transfer pathway permitted in the pulse-sequence. For example, an HN(CA)CO experiment utilises the following coherence transfers:



Experiments that follow similar coherence pathways, that is initially exciting the amide proton and detecting via this nucleus also, are referred to as 'out and back' pulse sequences. The nomenclature only takes into account the 'out' pathway, omitting the 'back' transfer steps (Cavanagh, 1996). In the above example, amide group resonances are correlated with $^{13}\text{C}'$ nuclei and this forms the basis of the experiment name. The magnetisation transfer is allowed via $^{13}\text{C}^{\alpha}$ and this permitted

pathway is represented in brackets (in this case as '(CA)'). However, no chemical shift is measured for $^{13}\text{C}^{\alpha}$ nuclei. Examples of commonly used complementary triple-resonance experiments are shown in Table 1. Following the sequential backbone assignment of amide groups, specific residue types can be identified using spectra such as the HN-HSQC-TOCSY that correlate side-chain protons with backbone amide groups.

Table 1 Commonly used three-dimensional triple resonance experiments. Magnetisation transfer pathways are indicated by red boxes.

Experiment	Correlations Observed	Magnetisation Transfer
HNCO	C'_{i-1}	
HN(CA)CO	C'_{i-1}, C'_i	
HNCA	$C^{\alpha}_{i-1}, C^{\alpha}_i$	
HN(CO)CA	C^{α}_{i-1}	
HNCACB	$C^{\alpha}_{i-1}, C^{\alpha}_i, C^{\beta}_{i-1}, C^{\beta}_i$	
HN(CO)CACB	$C^{\alpha}_{i-1}, C^{\beta}_{i-1}$	

1.5.2.1 Deuteration of Proteins

As described previously (section 1.5.1.1), the slow tumbling rates of large proteins have adverse effects on their spin relaxation properties. Slower tumbling molecules have short transverse relaxation times, in turn leading to a rapid decay of signal and a poor signal to noise ratio. In practice this can be observed as broad peaks in a NMR spectrum. Early applications of deuteration, replacing ^1H nuclei with spin-1 ^2H nuclei, aimed to simplify spectral interpretation by reducing the peak congestion in ^1H spectra (Crespi and Katz, 1969, Markley et al., 1968). The incorporation of deuterium also greatly improved the relaxation properties of remaining ^1H nuclei resulting in narrower linewidths and improved signal to noise, as can be seen in Figure 10. A further application of deuteration is improving the resolution in spectra of disordered proteins. Typically the interpretation of IDP spectra suffers due to significant signal overlap. The narrower linewidths resulting from deuteration can dramatically improve spectral resolution, thus making them easier to analyse. Backbone amide resonances are still visible in ^1H , ^{15}N HSQC spectra when the protein is dissolved in H_2O due to proton exchange effects with aqueous solvents. Therefore the standard backbone resonance assignment experiments discussed in section 1.5.2 are still applicable to deuterated proteins.

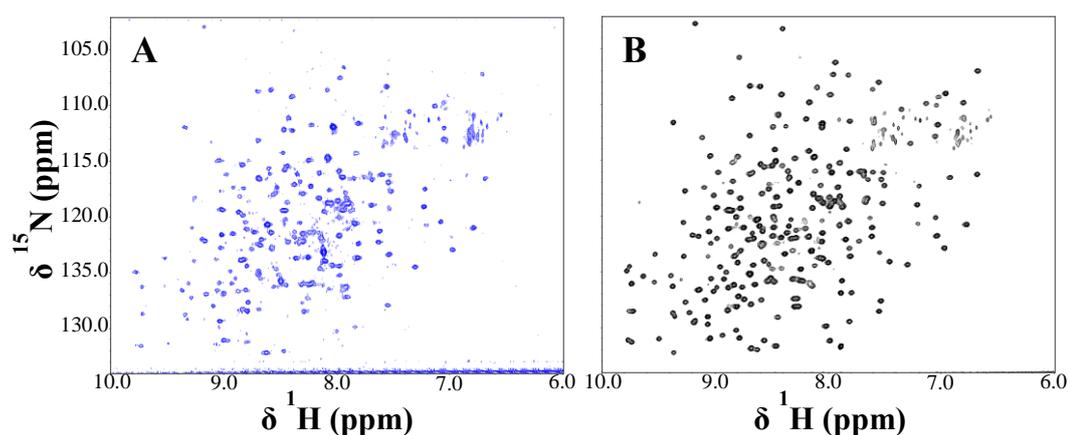


Figure 10 Effect of deuteration on the ^1H , ^{15}N spectrum of a 36 kDa protein. HSQC-TROSY spectra of ^{13}C , ^{15}N (A) and ^2H , ^{13}C , ^{15}N (B) N2N3 from FnBPA. Spectra were recorded under identical conditions.

1.5.3 Secondary Structure Determination

Secondary chemical shifts ($\Delta\delta$) can be used to determine regions of secondary structure within a protein. They are calculated by subtracting the experimentally measured chemical shift of a nucleus from the random coil chemical shift proposed for said nucleus. Random coil chemical shifts are defined as the chemical shift of an amino acid when it has no stable secondary or tertiary structure, thereby free to access all sterically allowed conformations (Wishart et al., 1995a). They are calculated experimentally from short peptides, or derived statistically from chemical shift databases where each residue in a known structure is classified as random coil, α -helix or β -strand (Wishart et al., 1995a, Wishart and Nip, 1998, Wishart et al., 1992). The advantage of using experimentally derived random coil shifts is that they are measured in defined conditions. However, the conditions used to determine these values may differ from those used for the protein of interest. In some cases denaturants were used and these are also known to influence random coil chemical shifts (Schwarzinger et al., 2000). Statistically-derived datasets incorporate values measured in a variety of conditions and referenced to different compounds. So while databases can be compiled from known structural elements, an inherent inaccuracy lies in the fact that conditions affecting random coil values vary throughout the sample set. Therefore, a common approach is to use multiple datasets in secondary chemical shift calculations (Mielke and Krishnan, 2004).

The magnitude and direction of secondary chemical shift values is indicative of α -helical, β -strand or intrinsically disordered regions. Wishart *et al.* proposed a method in which H^α nuclei are assigned a chemical shift index (CSI) based on their secondary chemical shifts (Wishart et al., 1992). Briefly, for an amino acid to be classified as random coil, its H^α chemical shift must be within ± 0.10 ppm of the defined random coil shift. Secondary chemical shifts that are more than 0.10 ppm and positive are assigned a CSI of 1, if they are more than 0.10 ppm and negative a CSI of -1 is assigned, and for those falling within the random coil shift range (± 0.10 ppm) a value of 0 is assigned. Clusters of the same CSI are indicative of particular secondary structure contributions; multiple 1's being a strand conformation, -1's helical and 0 disordered regions. Regions with less than 4 of the same CSIs assigned sequentially are also considered disordered (Wishart et al., 1992). Subsequent studies

found that secondary structure can be more accurately predicted if all assigned resonances are used in the analysis (Schwarzinger et al., 2001, Wishart et al., 1992). In addition to H^α shifts H^N , N^H , C^α , C^β and C' can also be used for CSI analyses, though the range whereby a CSI is applied differs between nuclei. Positive CSI values for C^α and C' , and negative values for H^α , H^N , N^H and C^β are indicative of helical secondary structure. Residues in a strand conformation have opposite CSI values for the respective nucleus. (Schwarzinger et al., 2001)

A number of factors can affect the random coil chemical shifts for amino acids in a protein. Therefore, a number of corrections need to be made prior to calculating secondary chemical shifts. Principally the random coil shifts must be corrected in a sequence-dependent manner. A number of studies have shown that the neighbouring residues can have a dramatic impact on the random coil shifts of a particular amino acid (Kjaergaard and Poulsen, 2011, Schwarzinger et al., 2001, Wang and Jardetzky, 2002, Wishart et al., 1995a). For example, residues preceding proline in the protein are subject to significant changes in random coil chemical shifts. These effects are attributed to the reduced conformational freedom of the proline ring and the absence of an amide proton (Schwarzinger et al., 2001). Aromatic residues can also cause changes in random coil shifts due to the action of the ring current generated by the aromatic rings. Nuclei that spend more time close to aromatic systems are correspondingly more susceptible to their effects. Many early studies were recorded at acidic pH and only one temperature. However, a recent investigation has shown that both parameters have non-negligible effects on certain nuclei. Therefore, coefficients correcting random coil chemical shifts values in commonly used datasets for temperature and pH variances are proposed to improve the reliability of secondary structure calculations (Kjaergaard et al., 2011). Unsurprisingly, inaccurate referencing of spectra has a significant impact on measured chemical shifts and therefore differences to reported random coil values (Wishart et al., 1995b). In turn this can impact calculated CSI values. Referencing to an internal standard, preferentially DSS (4,4-dimethyl-4-silopentane-1-sulphonic acid), is therefore crucial for accurately calculating secondary chemical shifts (Wishart et al., 1995b).

1.6 Aims

Despite the high level of sequence conservation of the N1 subdomain of FnBPA across strains of *S. aureus*, surprisingly little is known about its structure and function. N1 has no impact on the Fg-binding capacity of the A domain, and its absence does not affect the ability of N2 or N3 to fold. Therefore N1 is thought to have a role distinct from N2 and N3. FnBPA, and specifically the A domain, has a crucial role in the accumulation stage of *ica*-independent *S. aureus* biofilms. The specific aims of this work are to:

1. Structurally characterise the N1 subdomain of FnBPA by completing the resonance assignment of N1 and subsequently carrying out secondary chemical shift analysis.
2. Determine the involvement of N1 in possible ligand interactions. A series of pull-down experiments will be carried out with N1 as a bait protein in various target solutions to investigate potential unidentified ligands for FnBPA.
3. Investigate the role of FnBPA's A domain in adherence to host endothelial cells and subsequent colonisation, particularly on the role N1 plays any in such mechanisms.
4. Determine the molecular mechanism by which FnBPA mediates *S. aureus* biofilm formation and define the regions necessary for this function.
5. Deduce whether FnBP-mediated biofilm accumulation is facilitated exclusively by proteinaceous adhesins, or if other molecules in the biofilm matrix play a role.

2 Materials and Methods

2.1 Materials

2.1.1 Bacterial Strains

All *Escherichia coli* (*E. coli*) and *Staphylococcus aureus* (*S. aureus*) strains used in this work are listed in Table 2. *E. coli* DH5 α and BL21(DE3) Gold were used for molecular biology and protein production, respectively. Both strains were made chemically competent as described in section 2.2.1. *S. aureus* 8325-4 was kindly supplied by Dr. James Moir, University of York.

Table 2 Bacterial strains.

Organism	Strain	Supplier	Reference
<i>E. coli</i>	DH5 α	Invitrogen	(Meselson and Yuan, 1968)
<i>E. coli</i>	BL21(DE3) Gold	Stratagene	
<i>S. aureus</i>	8325-4		(O'Neill, 2010)

2.1.2 Bacterial Culture Media

The composition of *E. coli* and *S. aureus* culture media is described in Table 3. Selective media was supplemented with either 50 $\mu\text{g ml}^{-1}$ kanamycin or 100 $\mu\text{g ml}^{-1}$ ampicillin.

Table 3 Composition of *E. coli* and *S. aureus* culture media.

Medium	Application	Composition
LB	Molecular biology, protein production	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl
LB-agar	Molecular biology, protein production	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar
LB-glucose	<i>S. aureus</i> culture	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 0.3% (w/v) glucose
¹⁵ N M9 minimal medium	¹⁵ N labelled protein production	1% (w/v) ¹⁵ NH ₄ Cl, 0.68 % Na ₂ PO ₄ , 0.3% KH ₂ PO ₄ , 0.5% NaCl, 0.3% glucose, 2 mM MgSO ₄ , 0.2 mM CaCl ₂ , 5 μM FeCl ₃ , 1 μM MnCl ₂ , 1 μM ZnSO ₄ , 0.2 μM CoCl ₂ , 0.2 μM NiCl ₂ , 0.2 μM Na ₂ MoO ₄ , 0.2 μM Na ₂ SeO ₃ , 0.2 μM H ₃ BO ₃ , 1 μg ml ⁻¹ riboflavin, 1 μg ml ⁻¹ niacinamide, 1 μg ml ⁻¹ pyridoxine, 1 μg ml ⁻¹ thiamine monochloride
¹⁵ N, ¹³ C M9 minimal medium	¹³ C, ¹⁵ N labelled protein production	1% (w/v) ¹⁵ NH ₄ Cl, 0.68 % Na ₂ PO ₄ , 0.3% KH ₂ PO ₄ , 0.5% NaCl, 0.3% ¹³ C-glucose, 2 mM MgSO ₄ , 0.2 mM CaCl ₂ , 5 μM FeCl ₃ , 1 μM MnCl ₂ , 1 μM ZnSO ₄ , 0.2 μM CoCl ₂ , 0.2 μM NiCl ₂ , 0.2 μM Na ₂ MoO ₄ , 0.2 μM Na ₂ SeO ₃ , 0.2 μM H ₃ BO ₃ , 1 μg ml ⁻¹ riboflavin, 1 μg ml ⁻¹ niacinamide, 1 μg ml ⁻¹ pyridoxine, 1 μg ml ⁻¹ thiamine monochloride

2.1.3 Protein Expression Vectors

Vectors used for the cloning and over expression of protein are described in Table 4.

Table 4 Protein expression vectors.

Vector	Application	Supplier	Reference
pGEX-6P-1	Expression of N-terminally GST-tagged proteins	Novagen	(Norris et al., 2011)
pSKB2	Expression of N-terminally His ₆ -tagged proteins	Dr. James Brannigan, University of York	(Gruszka et al., 2012)

2.1.4 HUVECs

Human umbilical vein endothelial cells (HUVECs) were kindly donated by Dr. Dawn Coverley, University of York. Cell growth basal medium was prepared using the Lonza EGM-2 SingleQuot Kit (CC-4176), specifically designed to culture HUVECs.

2.1.5 Buffer Solutions

Table 5 Composition of buffers.

Buffer/Solution	Composition
Tris-borate-EDTA (TBE)	890 mM Tris-HCl, 890 mM Boric acid, 20 mM EDTA, pH 8.3
Phosphate buffered saline	140 mM NaCl, 27 mM KCl, 100 mM Na ₂ PO ₄ , 18 mM KH ₂ PO ₄ , pH 7.4

SDS-PAGE

MES running buffer (20x)	1 M MES, 1 M Tris, 69.3 mM SDS, 20.5 mM EDTA
MOPS running buffer (20x)	1 M MOPS, 1 M Tris, 69.3 mM SDS, 20.5 mM EDTA

Native-PAGE

Tris-Tricine running buffer	0.1 M Tris-HCl, 0.1 M tricine, pH 8.2
-----------------------------	---------------------------------------

Nickel-affinity purification

Binding buffer	20 mM HEPES, 150 mM NaCl, 20 mM imidazole, pH 7.5
Elution buffer	20 mM HEPES, 150 mM NaCl, 0.5 M imidazole, pH 7.5

Anion Exchange purification

Binding buffer	20 mM Bis-Tris, pH 6.2
Elution buffer	20 mM Bis-Tris, 1 M NaCl, pH 6.2

Size exclusion

Running Buffer	20 mM HEPES, 150 mM NaCl, pH 7.5
----------------	----------------------------------

2.1.6 Native Proteins**2.1.7 Fibrinogen From Human Plasma**

Human plasma glycoprotein Fg was obtained as a lyophilised white powder of purified Fg in 20 mM sodium citrate-HCl, pH7.4 (Calbiochem/Merck). Lyophilised powder (1 g) was resuspended in 24 ml of pre-warmed sterilised water (37 °C) and incubated at 37 °C until the Fg was completely dissolved. The Fg solution was

subsequently aliquoted and stored at -80°C . The Fg was analysed by SDS-PAGE to confirm its purity.

2.1.8 Fibronectin From Human Plasma

Native human plasma glycoprotein Fn was obtained as a 1 mM solution in 50 mM TBS, pH 7.5 (Sigma Aldrich). The Fn was analysed by SDS-PAGE to confirm its purity. The addition of 0.02 % (v/v) NaN_3 rendered the sample stable for several weeks at 4°C .

2.2 Methods in Molecular Biology

2.2.1 Preparation of Chemically Competent Cells

DH5 α and BL21(DE3) Gold competent cells were prepared according to a modified version of Cohen's procedure (Cohen et al., 1972). Briefly, cells were grown to an OD_{600} of 0.5 in LB medium, harvested by centrifugation (4500 xg, 10 mins) and resuspended in half the original volume of ice-cold 20 mM Tris-HCl, 50 mM CaCl_2 , pH 8.0. Following a 1 hr incubation on ice, cells were harvested by centrifugation and resuspended in one quarter of the original volume of ice-cold 20 mM Tris-HCl, 50 mM CaCl_2 , 20% (w/v) glycerol, pH 8.0. 200 μL aliquots were frozen in liquid nitrogen at stored at -80°C .

2.2.2 Transformation of Competent Cells

Cryo-preserved competent cells were thawed on ice, mixed with 50 ng of plasmid DNA and incubated on ice for 30 mins. Cells were then heat-shocked by incubation at 42°C for 45 s. Following incubation on ice for 5 mins, pre-warmed LB medium was added to a final volume equal to 10 times the original cell suspension volume. Prior to plating on selective LB-agar medium, cells were incubated at 37°C for 45 mins with shaking at 200 rpm.

2.2.3 Preparation of Plasmid DNA

5 ml cultures of freshly transformed DH5 α cell were grown overnight in LB medium at 37 °C with shaking at 200 rpm. Plasmid DNA was isolated using QIAprep Spin Miniprepkit (Qiagen) according to the manufacturer's instructions. DNA concentration was determined by measuring the absorbance at 260 nm. The purity of DNA was estimated using the A_{260}/A_{280} ratio and samples with a ratio between 1.6 and 2.0 stored at -20 °C in MilliQ (Millipore) purified water.

2.2.4 Agarose Gel Electrophoresis

Agarose gels of between 0.8% and 1% (w/v) were prepared in TBE buffer (Table 5) supplemented with SYBRsafe DNA dye (Invitrogen). DNA was mixed with (6x) DNA loading dye (Fermentas) according to the manufacturers instructions and loaded into the gel. Separation was achieved following the application of a voltage of 100 V for 1 hr. DNA was visualised using a UV transilluminator.

2.2.5 Polymerase Chain Reaction (PCR)

The composition of a standard PCR is given in Table 6. Manually designed primers (Table 7) were synthesised by Eurofins MWG. PCRs were performed on an MJ research PTC-200 Gradient Thermal Cycler according to the cycling programme in Table 8.

Table 6 Composition of PCR.

Component	Supplier	Final Concentration
Phusion HF Buffer	Finnzymes	1 x
dNTPs	Finnzymes	200 μ M each
Forward primer	Eurofins MWG	1 μ M
Reverse primer	Eurofins MWG	1 μ M
Template DNA		0.5 ng ml ⁻¹
Phusion Hot Start II DNA Polymerase	Finnzymes	0.02 U μ l
H ₂ O		

Table 7 PCR Primers. BamH1 (GGATCC) and Xho1 (CTCGAG) restriction sites are underlined

Construct	Forward Primer 5' – 3' Reverse Primer 5' – 3'
pGEX-N1	TCAGCT <u>GGATCC</u> GCATCAGAACAAAAGACAACACTACAG TATACT <u>CTCGAG</u> TCAACCCGTTTCCACTTTTCGC
pGEX-N1N2	TCAGCT <u>GGATCC</u> GCATCAGAACAAAAGACAACACTACAG TGACGT <u>CTCGAG</u> TCAATATTTAACATCTAATTCCTTTGAAG
pGEX-fA	TCAGCT <u>GGATCC</u> GCATCAGAACAAAAGACAACACTACAG TGATGACT <u>CTCGAG</u> TGATGATGAATCATATTCCTCTTCAACAGTAGTT AC

Table 8 PCR cycling program.

Cycle Step	Temperature	Time	No. Cycles
Initial denaturation	98 °C	3 mins	1
Denaturation	98 °C	10 s	35
Annealing	65 °C	30 s	
Extension	72 °C	30 s	
Final extension	72 °C	5 mins	1
Hold	4 °C	∞	1

2.2.6 Restriction Digest

Restriction digests of PCR products and plasmids were performed using BamH1 and Xho1 restriction enzymes purchased from New England Biolabs (NEB). Reactions were conducted according to the manufactures recommendations. Reaction mixtures were typically 50 µl in volume and contained 2 – 3 µg of DNA and 10 – 20 U of each restriction enzyme in the appropriate buffer (supplied by NEB). Reactions were incubated at 37 °C for at least 3 hrs. Digested DNA was separated by agarose gel electrophoresis and extracted from the gel and purified using QIAquick Gel Extraction Kit (Qiagen).

2.2.7 Ligation

Ligation mixtures constituted approximately 10 ng of linearised vector, a 5-fold molar excess of digested PCR product and 20 U of T4 DNA ligase (Fermentas) in 20 μ L of the supplied reaction buffer. Ligation reactions were incubated at room temperature for 3 hrs. 10 μ l of each reaction was then used to transform 100 μ l of competent DH5 α cells.

2.2.8 Construct Validation

The correct composition of newly generated DNA constructs was verified by an in-house DNA sequencing service (Technology Facility, University of York).

2.3 Protein Expression and Purification

2.3.1 Expression of Unlabelled Proteins

BL21 (DE3) Gold cell were transformed with appropriate plasmid DNA according to the procedure outlined above (section 2.2.2). A single colony was selected based on resistance to kanamycin and used to inoculate 50 ml of LB medium containing kanamycin (LB-kan) (Table 3). After overnight incubation at 37 °C with shaking (220 rpm), 5 ml of the starter culture was inoculated into 500 ml of fresh LB-kan medium and incubated at 37 °C with shaking at 200 rpm until the optical density at 600 nm (OD_{600}) reached 0.6. Isopropyl- β -thiogalactopyranoside (IPTG, Melford) was added to a final concentration of 0.5 mM and cultures incubated for 4 hrs at 37 °C with shaking. Cells were harvested by centrifugation (5000 rpm, 20 mins, 4 °C) and resuspended in nickel affinity binding buffer (Table 5). Cell pellets were used immediately for protein purification or stored at -20 °C.

2.3.2 Expression of ^{15}N and ^{13}C , ^{15}N Labelled Proteins

Expression of isotopically enriched proteins was performed ^{15}N M9 or ^{13}C , ^{15}N M9 minimal medium (Table 3) according to the protocol described in section 2.3.1 with two exceptions. Firstly, 5 ml of the starter culture was pelleted by centrifugation

(5000 rpm, 4 °C, 15 mins) and resuspended in 5 ml of M9 medium prior to inoculating fresh expression medium. Secondly, due slower bacterial growth in minimal medium, following induction with IPTG, cultures were incubated with shaking at 30 °C overnight.

2.3.3 Expression of ^2H , ^{13}C , ^{15}N N1

BL21 (DE3) Gold cells were transformed with plasmid DNA encoding His₆-N1 according to the procedure outlined above (section 2.2.2). A single colony was selected based on resistance to kanamycin and used to inoculate 5 ml LB-kan. Following overnight incubation with shaking at 37 °C, 200 µl was inoculated into 5 ml M9 minimal medium with kanamycin, which in turn was used to inoculate 5 ml 10% $^2\text{H}_2\text{O}$ M9 minimal medium. Subsequent 5 ml cultures of 20%, 40%, 60% and 80% $^2\text{H}_2\text{O}$ M9 minimal medium were inoculated with 200 µl of the preceding $^2\text{H}_2\text{O}$ concentration culture following overnight growth at 37 °C with shaking. Finally 200 µl of 80% $^2\text{H}_2\text{O}$ M9 minimal medium culture was inoculated into a 50 ml 100% $^2\text{H}_2\text{O}$ M9 minimal medium overnight starter culture. 5 ml of overnight starter culture was inoculated into 500 ml of 100% $^2\text{H}_2\text{O}$ ^{13}C , ^{15}N M9 minimal medium and incubated with shaking at 37 °C until the culture reached an OD₆₀₀ of 0.6. Protein expression was induced through the addition of IPTG to a final concentration of 0.5 mM and cultures were incubated for a further 18 hrs with shaking at 30 °C. Cells were harvested by centrifugation (5000 rpm, 4 °C, 30 mins) and resuspended in nickel affinity binding buffer (Table 5). Cell pellets were used immediately for protein purification.

2.3.4 Expression of Selectively Unlabelled ^2H , ^{13}C , ^{15}N N1

Expression of ^2H , ^{13}C , ^{15}N N1 was performed according to the protocol described in section 2.3.4 with one exception. 1 hr prior to inducing protein expression, growth media supplements were added to selectively unlabel a specific amino acid (Rasia et al., 2012). The supplements, corresponding unlabelled amino acid and requisite concentrations are detailed in Table 9.

Table 9 Growth media supplements to selective unlabel amino acids.

Growth Medium Supplement	Residue(s) Unlabelled	Final Conc. (mg l⁻¹)	Supplier
L-proline	Proline	250	Sigma-Aldrich
4-hydroxyphenylpyruvic acid	Tyrosine	400	Sigma-Aldrich
Sodium 3-methyl-oxobutyrate	Leucine/Valine	120	Sigma-Aldrich
Sodium 2-oxobutyrate	Isoleuline	70	Sigma-Aldrich

2.3.5 *E. coli* Cell Lysis

Frozen cell pellets were thawed at room temperature and homogenised using a Pasteur pipette. Cells were lysed by sonication using a Misonix Sonicator 300 equipped with a $\frac{3}{8}$ inch stud probe. A standard sonication cycle involved 60 3 s pulses at 70 W with 7 s intervals between pulses. During sonication the cell suspension was cooled on ice. The crude extract was clarified by centrifugation (48000 xg, 4 °C, 30 mins). The supernatant was filtered through a 0.22 µm PVDF membrane (Millipore) prior to further purification.

2.3.6 Purification of His₆-tagged Proteins

Clarified cell extract was loaded at a flow rate 3 ml min⁻¹ onto a 5 ml HisTrap HP column (GE Healthcare) equilibrated with binding buffer (Table 5). Unbound species were removed from the column in a subsequent washing step with binding buffer until a stable baseline of absorbance at 280 nm was reached. His₆-tagged proteins were eluted at a flow rate of 3 ml min⁻¹ with a linear gradient of imidazole from 20 – 250 mM over 25 column volumes (nickel affinity elution buffer, Table 5). Fractions were collected and analysed by SDS-PAGE. Fractions containing protein were pooled and dialysed at 4 °C overnight against 20 mM HEPES, 150 mM NaCl, pH 7.5 using 5 kDa molecular weight cut-off dialysis tubing (Spectrum).

To cleave the N-terminal His₆-tag, purified recombinant protein was incubated with human rhinovirus (HRV) 3C protease (Novagen), a protease that recognises the cleavage site Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro and cleaves between Gln and Gly. Cleavage reactions were typically performed at a protease:protein ratio of 1:300 (w/w) and incubated at 4 °C for 6 hrs. The His₆-tag and protease were removed from the target protein by passage over a HisTrap HP column at 5 ml min⁻¹ with the flow-through containing recombinant protein collected. The purity of recombinant proteins was assessed by SDS-PAGE. In cases of insufficient purity (less than 95% pure), namely N2N3 and AF1, further purification was required.

N2N3 was concentrated to a volume of 1 ml by centrifugation (4500 xg, 4 °C) using VivaSpin20 5 kDa MWCO concentrators (Sartorius). The concentrated protein sample was loaded onto a HiLoad 16/60 Superdex 75 size exclusion column (GE Healthcare) and eluted with 1 column volume (120 ml) of 20 mM HEPES, 150 mM NaCl, pH 7.5 at a flow rate of 1 ml min⁻¹. Elution peaks were analysed by SDS-PAGE and fractions containing protein were pooled and concentrated using VivaSpin20 5 kDa MWCO concentrators.

Further purification of AF1 was achieved by anion exchange chromatography. The flow-through from the nickel affinity purification was dialysed against anion exchange binding buffer (Table 5) at 4 °C overnight using 5 kDa MWCO dialysis tubing. AF1 was loaded at 5 ml min⁻¹ onto a 5 ml HiTrap Q sepharose column (GE Healthcare) equilibrated with binding buffer (Table 5). Unbound species were removed from the column in a subsequent washing step with binding buffer until a stable baseline of absorbance at 280 nm was reached. AF1 was eluted at a flow rate of 5 ml min⁻¹ with a linear gradient of NaCl from 0 – 1 M NaCl over 30 column volumes. Identity and purity of the resulting peaks was determined by SDS-PAGE analysis.

2.3.7 Purification of GST-tagged Proteins

Clarified cell lysate was passed over a 5 ml GSTrap FF column (GE Healthcare), equilibrated with PBS, 3 times at 3 ml min⁻¹. Unbound species were removed from the column in a subsequent washing step with binding buffer until a stable baseline

of absorbance at 280 nm was reached. Proteins were eluted from the column at 3 ml min⁻¹ by the addition of elution buffer (Table 5) applied for 6 column volumes. Collected fractions were analysed using SDS-PAGE. Fractions that contained pure recombinant protein were pooled and dialysed (using 3 kDa MWCO tubing) against PBS and concentrated using VivaSpin20 10 kDa MWCO concentrators.

2.3.8 Isolation of Wall Teichoic Acids

WTAs were isolated essentially as described by Peschel *et al.* (Peschel *et al.*, 1999b). *S. aureus* was grown overnight at 37 °C with shaking (180 rpm) in 500 ml LB-glucose (Table 3). Cells were harvested by centrifugation (5000 rpm, 4 °C, 30 mins), washed with 100 ml 20 mM sodium acetate (NaOAc), pH 5.8 and resuspended in 50 ml of the same buffer. Cells were lysed by treatment with 2.5 mg of lysostaphin (Sigma-Aldrich) at 37 °C with shaking for 2 hrs. Crude cell extracts were diluted 4 times in 20 mM NaOAc, pH 4.6 and sonicated for 10 mins (150 W, 25 s on, 5 s off) using a Misonix Sonicator 300 equipped with an ethylene glycol-cooled 4 inch cup-horn probe. Cell extracts were then incubated at 60 °C for 1 hr with vigorous shaking. Cells walls were pelleted by centrifugation (48000 xg, 18 °C, 30 mins) and washed extensively with 20 mM NaOAc, pH 4.6. Teichoic acids were released by diluting cell wall extracts 4 times in 5% (w/v) trichloroacetic acid and incubating at 60 °C for 4 hrs. Peptidoglycan was removed by centrifugation (48000 xg, 18 °C, 45 mins), the crude WTA extract extensively dialysed against mildly acidic H₂O and lyophilised. WTAs were resuspended in 1 ml deionised H₂O and loaded on to a Superdex 75 HR 10/30 size exclusion column (GE Healthcare) and eluted at a flow rate of 0.5 ml min⁻¹ with 1 column volume (25 ml) of 20 mM HEPES, 150 mM NaCl, pH 5.5. Fractions containing WTA were pooled and analysed by SDS-PAGE. Pure samples were extensively dialysed against mildly acidic H₂O, lyophilised and stored at -20 °C.

2.4 Biochemical Methods

2.4.1 SDS-PAGE

The NuPAGE Novex Bis-Tris protein gel electrophoresis system (Invitrogen), XCell SureLock Mini-Cell gel tanks (Invitrogen) and a Bio-Rad power-pack were used for all SDS-PAGE analyses. Pre-cast NuPAGE Novex 4 – 12% Bis-Tris gels (Invitrogen) were used to separate protein species. Prior to loading onto a gel, samples were denatured in NuPAGE LDS Sample Buffer (Invitrogen), boiled for 5 mins and centrifuged (13000 xg, room temperature, 1 min). Mark12 unstained protein standard (Invitrogen) was used to estimate the molecular weight of protein samples. Electrophoresis was performed in MES- or MOPS- SDS running buffer (Table 5) at a constant voltage of 200 V for 35 – 40 mins. Proteins were visualised by staining with 0.2% (w/v) Coomassie brilliant blue R250 10% (v/v) acetic acid, 50% (v/v) ethanol, followed by destaining in 10% (v/v) acetic acid, 10% (v/v) ethanol.

2.4.2 Native-PAGE

The Bio-Rad Mini-PROTEAN Tetra Electrophoresis System and a Bio-Rad power-pack were used for all native-PAGE analyses. Teichoic acids were separated using NativePAGE Novex 15% Bis-Tris pre-cast gels (Invitrogen). Samples were prepared in 20 mM HEPES, 150 mM NaCl, 50% (v/v) glycerol, 0.5% bromophenol red, pH 7.5. Electrophoresis was performed in native-PAGE running buffer (Table 5) at a constant voltage of 100 V for 90 mins. Teichoic acids were visualised according to the alcian blue-silver staining procedure developed by Wolters *et al.* (Wolters et al., 1990). Briefly, gels were soaked in 1% (w/v) alcian blue 8GX (Sigma-Aldrich), 3% (v/v) acetic acid for 30 mins followed by destaining in 1% (v/v) acetic acid, 45% methanol (v/v) overnight. Gels were treated with 3.4 mM Na₂Cr₂O₇, 3.2 mM HNO₃ for 7 mins and washed extensively with H₂O. Gels were soaked with gentle agitation under a lamp in 12 mM silver nitrate for 30 mins and washed several times with H₂O. Gels were developed by 3 short treatments with 0.28 M sodium carbonate, 6 mM formaldehyde. As soon as the background darkened, the gel was flooded with 0.1 M acetic acid to quench the development process.

2.4.3 Determination of Protein Concentration

Protein concentration was determined from the absorbance measurement at 280 nm according to the Beer-Lambert law:

$$c = \frac{A}{\varepsilon \cdot l}$$

where c is the molar concentration of the protein, A is the absorbance at 280 nm, ε is the molar extinction coefficient at 280 nm and l is the path length. The theoretical molar extinction coefficient ($\varepsilon_{protein}$) was calculated using the ProtParam program according to the following equation (Gasteiger et al., 2003):

$$\varepsilon_{protein} = N_{Tyr} \cdot \varepsilon_{Tyr} + N_{Trp} \cdot \varepsilon_{Trp} + N_{Cys} \cdot \varepsilon_{Cys}$$

where N_{Tyr} , N_{Trp} and N_{Cys} is the number of Tyr, Trp and Cys residues in the protein, respectively, and ε_{Tyr} , ε_{Trp} and ε_{Cys} are the molar extinction coefficients of Tyr (1490 M⁻¹), Trp (5500 M⁻¹) and Cys (125 M⁻¹), respectively.

Due to the presence of only 1 Tyr residue, and lack of Trp and Cys residues, N1 has a correspondingly low molar extinction coefficient, which may lead to significant errors in concentration estimates based on absorbance measurements at 280 nm. To verify the reliability of concentrations obtained by spectrophotometric analysis, the protein concentration was also estimated by amino acid analysis (Alta Bioscience, University of Birmingham). The concentration values obtained from both techniques were identical confirming the suitability of absorbance-based concentration measurements for N1.

2.4.4 GST Pull-down Assays

20 μ l aliquots of MagneGST Glutathione Particles (Promega) were washed 5 times with 1 ml PBS and resuspended in 500 μ l of PBS containing bait protein (GST negative control, GST-N1, GST-N1N2, GST- fA) at a concentration of 4.4 μ M and incubated at 4°C for 90 mins with gentle agitation. The beads were washed 5 times with 1 ml PBS, resuspended in 750 μ l of plasma and incubated for 2 h at 4°C with

gentle agitation. The beads were then washed 5 times with 1 ml PBS before re-suspension in 50 μ l PBS. Protein was eluted from beads by boiling for 5 mins, and the supernatant was analysed by SDS-PAGE.

2.4.5 His₆-tag Pull-down Assay

62.5 μ l (2.5 mg) aliquots of Dynabeads (Invitrogen) for His₆-tagged protein isolation were washed 5 times with 1 ml PBS. The beads were then incubated with 4.4 μ M solution of bait protein (His-N1, His-N1N2, His-fA) in PBS for 1 h at 4°C with gentle agitation. The beads were washed 5 times with 500 μ l PBS before adding 500 μ l blood plasma and incubating for 2 h at room temperature. Supernatant was removed and the beads washed 5 times with 500 μ l PBS before re-suspension in 50 μ l PBS. Protein was eluted from beads by boiling for 5 mins, and the supernatant was analysed by SDS-PAGE.

2.4.6 Affinity Column Pull-down Assay

FnBPA sepharose columns were prepared using AminoLink Plus Immobilisation Kit (Thermo Scientific) according to the manufacturers protocol. Briefly, 20 mg purified FnBPA A domain constructs N1, N2N3 and AF1 in 3 ml PBS were incubated with beaded agarose resin and 40 μ l NaCNBH₃ mixing end over end at 4 °C for 18 hrs. Buffer and unbound protein was removed by centrifugation (1,000 xg, 1 min). Remaining binding sites on the resin blocked by incubation with quenching buffer and NaCNBH₃. One column was left uncharged as a negative control. Prior to applying 3 ml analyte solution (Fn/Fg control solutions, plasma), columns were equilibrated in PBS. The column was washed 5 times with 3 ml PBS, and bound protein was eluted with 3 washes of 3 ml 0.1 M Sodium citrate, 150 mM NaCl buffer, pH 3.0. Eluted fractions were analysed by SDS-PAGE.

2.4.7 HUVECs Adhesion Assay

HUVECs were cultured in basal medium (section 2.1.4) to approximately 95% confluency, liberated with trypsin-EDTA and resuspended in fresh basal medium. Cells were added to 24-well plates and incubated at 37 °C for 48 hrs. Following the

addition of proteins to final concentrations of 25, 50 and 70 μM , the HUVECs were incubated for a further 30 mins at 37 $^{\circ}\text{C}$. Following extensive washing with PBS cells were liberated from the 24-well plate by treatment with LDS and dithiothreitol (0.1 M) SDS-PAGE sample buffer. Cell extracts were boiled for 15 mins, centrifuged (13000 rpm, room temperature, 5 mins) and the supernatant subjected to SDS-PAGE (section 2.4.1).

2.4.8 Western Blot Analysis of HUVECs Extracts

Following SDS-PAGE, separated proteins were transferred onto a nitrocellulose membrane using an Invitrogen iBlot Blotting System. Membranes were blocked by incubation with PBST, 5% Marvel milk powder at 4 $^{\circ}\text{C}$ with gentle agitation for 1 hr and probed with mouse anti-His₆ antibodies for 1 hr at room temperature with gentle agitation. Membranes were washed 5 times with PBST (Table 5) and incubated with HRP-conjugated sheep anti-mouse antibodies for 1 hr with gentle agitation. Membranes were washed 5 times with PBST and reactive bands detected by chemiluminescence and visualised with photographic film.

2.4.9 Gel Electromobility shift Assay (EMSA)

Band shift assays were carried out using 1% agarose gels. TBE (Table 5) was used for gel preparation and running buffer. Gel wells were thoroughly flushed with running buffer prior to loading samples. DNA ligand at a concentration of 0.4 μM was incubated with a range of protein concentrations (0.0, 0.2, 0.4, 0.8, 1.2 and 1.6 μM) in 20 μL 20 mM HEPES, 150 mM NaCl, 50 μM ZnAc₂, pH 5.5 for 15 mins prior to electrophoresis. Gel electrophoresis was conducted for 1 hr at 60 V and 4 $^{\circ}\text{C}$. DNA was stained with SYBRsafe and visualised using a UV transilluminator.

2.5 Biophysical Methods

2.5.1 Circular Dichroism (CD)

Recombinant protein was dialysed overnight against 10 mM Na/K⁺ PO₄⁻, pH 7.4 using 2,000 Da MWCO Slide-a-lyzer dialysis cassettes (Thermo-Fisher) and then

diluted to a concentration of 0.2 mg ml^{-1} . 200 μl of protein was loaded into a quartz cuvette with a path-length of 0.1 mm. Spectra were recorded at $25 \text{ }^\circ\text{C}$ from 300 – 184 nm on a Jasco J810 CD spectrophotometer. Data was processed using Dichroweb software (Whitmore and Wallace, 2004).

2.5.2 Size-exclusion Chromatography Multi Angle Laser Light Scattering (SEC-MALLS)

Recombinant protein was dialysed overnight using 2,000 Da MWCO Slide-a-lyzer dialysis cassettes (Thermo-Fisher) against 50 mM HEPES, 150 mM running buffer at either pH 5.5 or 7.5 containing either 2 mM EDTA or 5 mM ZnAc_2 . Samples were then diluted to a concentration of 1 mg ml^{-1} . SEC-MALLS experiments were performed using a Superdex 200 HR10/30 column (GE Healthcare) on a Shimadzu HPLC system. 100 μl protein samples were loaded onto the size exclusion column and eluted at 0.5 ml min^{-1} with 1 column volume (24 ml) of an appropriate running buffer. Protein elution was monitored using a Wyatt Technologies Dawn HELEOS-II light-scattering detector and an Optilab rEX refractive index monitor. Recorded data were analysed using Astra 5.0 software (Wyatt Technologies).

2.6 Methods in NMR Spectroscopy

2.6.1 Sample Preparation

Prior to sample preparation, proteins were dialysed against the appropriate buffer at $4 \text{ }^\circ\text{C}$ overnight using 2000 MWCO Slide-a-lyzer dialysis cassettes. Typically, samples were 700 μl in volume comprising 630 μl of protein solution, 70 μl D_2O (10% (v/v)) and 1.2 μl of 10% (v/v) NaN_3 (final concentration of 0.02% (v/v)). Heteronuclear single quantum coherence (HSQC) NMR spectra of ^{15}N labelled were generally acquired with a protein concentration of 0.2 mM, while three-dimensional triple-resonance experiments were recorded for samples containing 1.0 mM ^{13}C , ^{15}N or ^2H , ^{13}C , ^{15}N labelled protein.

2.6.2 Data Acquisition

All NMR spectra were recorded on a Bruker Avance 700 MHz spectrometer equipped with a triple-resonance inverse-detection probe and z-axis pulsed field gradients. 1D ^1H spectra (pulse program, zgesgp) were acquired with a spectral window of 9000 – 12000 Hz, 64 scans and 8192 points. The basic parameters of all NMR experiments used throughout this work are listed in Table 10. The HN-HSQC-TOCSY experiment was acquired with a TOCSY mixing time of 40 ms.

Table 10 Pulse programs and basic parameters for two- and three- dimensional experiments recorded for this work.

Experiment	Pulse Program	Scans	Number of Points			Spectral Width Hz ppm		
			¹ H	¹⁵ N	¹³ C	¹ H	¹⁵ N	¹³ C
HSQC	hsqcetf3gpsi	8	2048	256	-	10000 14.28	1560.9 22.00	-
BEST-HSQC	b_hsqcetf3gpsi	8	2048	256	-	10000 14.28	22.001 560.9	-
TROSY-HSQC	trosetf3gpsi	16	2048	256	-	10000 14.28	32.00	-
BEST-HNCO	b_hncogp3d	8	2048	64	128	10000 14.28	1560.9 22.00	1760.8 10.00
BEST-HN(CA)CO	b_hncacogp3d	8	2048	64	128	10000 14.28	1560.9 22.00	1760.8 10.00
BEST-HNCA	b_hncagp3d	8	2048	64	128	10000 14.28	1560.9 22.00	3521.2 20.00
BEST-HN(CO)CA	b_hncocagp3d	8	2048	64	128	10000 14.28	1560.9 22.00	3521.2 20.00
BEST-HNCAi	b_hncaigp3d	8	2048	64	128	10000 14.28	1560.9 22.00	3521.2 20.00
BEST-HNCACB	b_hncacbgp3d	8	2048	64	128	10000 14.28	1560.9 22.00	10563.3 60.00
BEST-HN(CO)CACB	b_hncocacbgp3d	8	2048	64	128	10000 14.28	1560.9 22.00	10563.3 60.00
			Number of Points			Spectral Width Hz ppm		
			¹ H	¹⁵ N	¹⁵ N	¹ H	¹⁵ N	¹⁵ N
HNCANNH	hncannhgp3d	16	2048	128	128	10000 14.28	1560.9 22.00	1560.9 22.00
			Number of Points			Spectral Width Hz ppm		
			¹ H	¹⁵ N	¹ H	¹ H	¹⁵ N	¹ H
HN-HSQC-TOCSY	dipsihsqcf3gpsi3d	8	2048	80	300	10000 14.28	1560.9 22.00	6301.1 9.00

2.6.3 Data Processing

Spectra were processed with scripts generated in NMRPipe (Delaglio et al., 1995) that were optimised for each dataset. Scripts included apodization functions applied to the FID prior to Fourier transformation to improve signal-to-noise and spectral resolution. Typically a Gaussian function was applied in the first ^1H dimension, whereas squared sinebell functions were applied to the second and third dimensions or appropriate datasets. Zero-filling was applied in all dimensions to increase the digital resolution of the spectra and in some cases linear prediction was applied to dimensions in which relatively few points had been collected to correct for FID truncation and improve resolution.

2.6.4 Spectral Referencing

Following processing, spectra were referenced to 4,4-dimethyl-4-silapentane-1-sulphonic acid (DSS). DSS produces a signal in the 1D ^1H spectrum at 0 ppm. Therefore, all ^1H frequencies were corrected according to the formula:

$$\nu_{cor,0}^H = \nu_{obs,0}^H - \nu_{DSS}^H$$

where $\nu_{cor,0}^H$ is the corrected ^1H frequency at 0 ppm, $\nu_{obs,0}^H$ the observed ^1H frequency at 0 ppm and ν_{DSS}^H the observed DSS signal. For the indirect ^{13}C and ^{15}N dimensions, the reference frequencies were calculated using IUPAC accepted frequency-referencing ratios (Ξ^N and Ξ^C for ^{15}N and ^{13}C , respectively) and the corrected ^1H reference frequency ($\nu_{cor,0}^H$), using the formula;

$$\nu_{cor,0}^X = \Xi^X \cdot \nu_{cor,0}^H$$

where X is either ^{13}C or ^{15}N and $\nu_{cor,0}^X$ is the corrected reference frequency for nucleus X. Ξ^N and Ξ^C are 0.101329118 and 0.251449530, respectively (Yamazumi et al., 2001). Chemical shifts were then corrected according to the following formula:

$$\delta_{cor}^X = \delta_{obs}^X - \left(\frac{\nu_{obs,0}^X - \nu_{cor,0}^X}{\nu_{cor,0}^X} \cdot 10^6 \right)$$

where δ_{cor}^X is the corrected chemical for nucleus X (^1H , ^{13}C , ^{15}N), δ_{obs}^X is the observed chemical shift for nucleus X and $\nu_{obs,0}^X$ is the observed frequency of X at 0 ppm in the respective dimension.

2.6.5 Resonance Assignment

Processed and referenced spectra were analysed using CcpNmr Analysis version 2.1.5 (Vranken et al., 2005, Fogh et al., 2005). The procedure for resonance assignment is outlined in detail in section 4.5.3. Sequence-specific triple-resonance assignment of H^{N} , N^{H} , C^{α} , C^{β} and C' nuclei was based on HNCO/HN(CA)CO, HNCA/HN(CO)CA and HNCACB/HN(CO)CACB spectra.

3 Expression and Purification of Recombinant FnBPA Fragments

3.1 Introduction

FnBPA has been identified as a major *S. aureus* virulence factor associated with a variety of life-threatening conditions such as osteomyelitis and infective endocarditis (Shinji et al., 2011, Que et al., 2001). The contribution of FnBPA in the development of these conditions is largely attributed to its recognition of multiple ECM components such as fibronectin (Fn), fibrinogen (Fg) and elastin (Wann et al., 2000, Keane et al., 2007b, Froman et al., 1987). The Fn- and Fg-binding regions of FnBPA, the Fn binding repeats (FnBRs) and N2N3 subdomains, respectively, have been investigated in great detail and the interactions they facilitate are well understood (Bingham et al., 2008, Meenan et al., 2007)(Stemberk et al., manuscript in preparation). Upon binding to Fn, the FnBRs form additional β -strands antiparallel to the E-strands of Fn F1 modules in the N-terminal domain (Schwarz-Linek et al., 2003). Fg binding is mediated by the A domain of FnBPA, with the C-terminus of the Fg γ -chain binding in the hydrophobic cleft between the N2 and N3 subdomains (Stemberk et al., manuscript in preparation). Contrastingly, the N1 subdomain is poorly characterised and has no involvement in known FnBPA interactions.

3.2 Aims

The aim of this work was to express and purify a set of recombinant FnBPA (rFnBPA) fragments covering different regions of the A domain to enable the characterisation of the N1 subdomain. Primarily, poly-histidine (His₆)-tagged proteins were produced via well-established procedures (Stemberk et al., manuscript in preparation). Isotopically enriched proteins were expressed to facilitate the structural analysis of N1 by NMR spectroscopy. To aid the resonance assignment of N1 (section 4.5), a series of proteins were expressed using a recently developed 'unlabelling' strategy (Rasia et al., 2012). DNA constructs encoding GST-fusion proteins of N1, N1N2 and fA were produced to enable to expression of proteins for use in ligand binding studies (section 5.3).

3.3 Recombinant FnBPA Fragments

For the investigations presented in this work a number of different rFnBPA constructs were used. Primarily these were A domain variants encompassing the N1 subdomain and the Fg-binding N2N3 region. However, for work presented in Chapter 5, a protein containing the first FnBR was used and this will be also discussed here. The various proteins used are summarised in Figure 11.

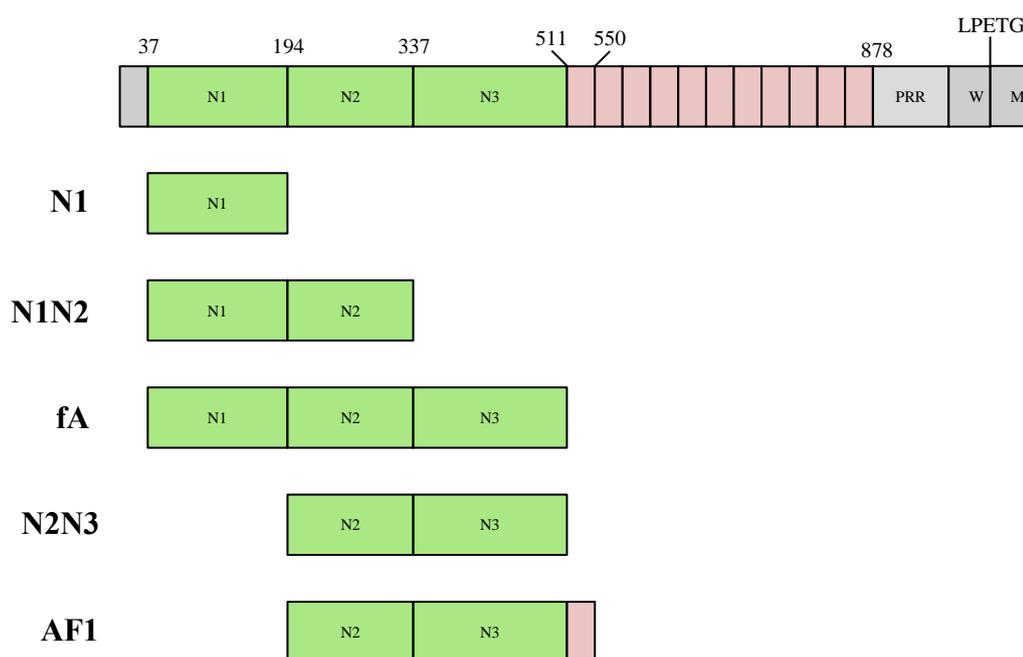


Figure 11 Recombinant FnBPA fragments.

The A domain of FnBPA is preceded by a short secretory signalling peptide (residues 1 – 36), responsible for ensuring FnBPA is directed to the cell surface (DeDent et al., 2008). The A domain comprises three subdomains; N1, N2 and N3. N1 is highly conserved with approximately 90% sequence identity between strains of *S. aureus*, compared to approximately 75% and 60% sequence identity in N2 and N3, respectively (Loughman et al., 2008). The N2N3 subdomains represent the minimum region required to bind Fg, with N1 having no involvement (Stemberk et al., manuscript in preparation). Elastin and tropoelastin binding by FnBPA is also achieved by N2N3 (Keane et al., 2007a, Keane et al., 2007b). No functional role has yet been established for N1. Structural characterisation of N2 and N3 revealed each subdomain adopts an immunoglobulin type fold (Stemberk et al., manuscript in

preparation). Proteins containing the independently folded N2 and N3 subdomains (N1N2 and fA) were used to investigate N1's structure in the context of larger A constructs (section 4.4.1). N1N2 is unable to bind Fg as it lacks the hydrophobic cleft between the N2 and N3 subdomains necessary for this interaction (Stemberk et al., manuscript in preparation). AF1 comprises the Fg binding region (N2N3) and the first disordered FnBR; as a result AF1 is able to bind both Fg and Fn (Stemberk et al., manuscript in preparation).

3.4 Expression of His₆-tagged rFnBPA Proteins

3.4.1 Unlabelled Protein Expression

The procedure for the optimal expression of unlabelled rFnBPA A domain proteins developed by Vaclav Stemberk was used here (Stemberk et al., manuscript in preparation). Briefly, plasmid DNA containing the rFnBPA fragments was transformed into the *E. coli* expression strain BL21 (DE3) Gold, as described in section 2.2.2. Positive transformants were selected by resistance to kanamycin and used to inoculate overnight starter cultures. Baffled flasks containing 500 ml LB with kanamycin were inoculated with the starter cultures, and following growth for 2 – 3 hours to an OD₆₀₀ of approximately 0.7, protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM. The growth curves associated with the expression of each rFnBPA fragment are shown in Figure 12. Overexpression of recombinant protein was confirmed by SDS-PAGE analysis prior to protein purification (data not shown).

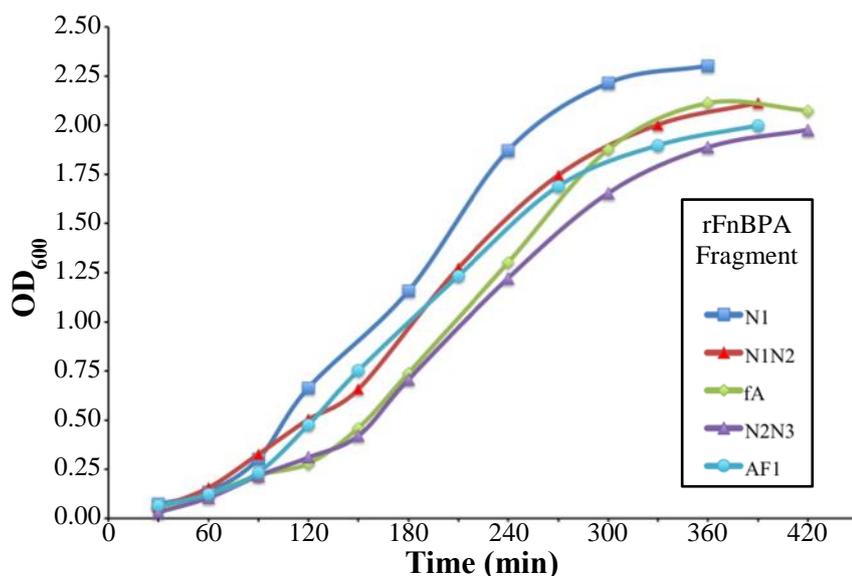


Figure 12 *E. coli* growth curves during the expression of rFnBPA proteins. The growth of *E. coli* BL21 (DE3) Gold was monitored by measuring the OD₆₀₀ of the culture. An OD₆₀₀ of ~0.7 was reached after approximately 180 minutes at which point protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM.

3.4.2 Uniformly-Labelled rFnBPA Expression

rFnBPA proteins uniformly ¹⁵N or ¹³C, ¹⁵N labelled were expressed in M9 minimal media (section 2.3.2). Supplementing the expression media with ¹⁵NH₄Cl and ¹³C-glucose (Marley et al., 2001), as the sole nitrogen and carbon sources, respectively, ensured sufficient isotope enrichment for NMR spectroscopy studies. Although cell cultures grew more slowly in minimal media, typically taking approximately 6 h to reach an OD₆₀₀ of around 0.7, following induction, overnight growth at 30 °C resulted in similar levels of protein expression compared to growth in LB. Following purification, the level of isotope enrichment was determined by comparison of the molecular mass determined by electrospray ionisation mass spectrometry (ESI-MS) analysis with the theoretical mass of the protein (Gasteiger et al., 2003) (Table 11).

3.5 Purification of His₆-tagged rFnBPA Proteins

After expression was induced, cells were typically grown for 4 h at 37 °C or overnight at 30 °C. Cells were harvested by centrifugation prior to lysis by sonication. His₆-tagged rFnBPA proteins were purified from the soluble cell extract

by nickel affinity chromatography. Proteins eluted from the column at between 75 and 150 mM imidazole (Figure 13A). Fractions containing protein were pooled and dialysed against 20 mM HEPES, 150 mM NaCl, pH 7.5 to remove imidazole. The N-terminal His₆-tag was cleaved using HRV 3C protease and removed, along with the enzyme, from the protein solution by a second nickel affinity purification step (Figure 13C). Following cleavage of the His₆-tag, N1 appears to increase in mass according to SDS-PAGE analysis. His₆-tags are known to affect the expansion properties of disordered proteins resulting in abnormal behaviour in a number of analytical techniques (Marsh and Forman-Kay, 2010), although the behaviour observed in this case has not previously been reported. Nonetheless, MS analysis confirmed the cleaved species was of the mass expected for recombinant N1. The volume of protein samples was reduced using VivaSpin20 5 kDa molecular weight cut-off concentrators (Vivaproducts) and the purity verified by SDS-PAGE analysis (Figure 13D).

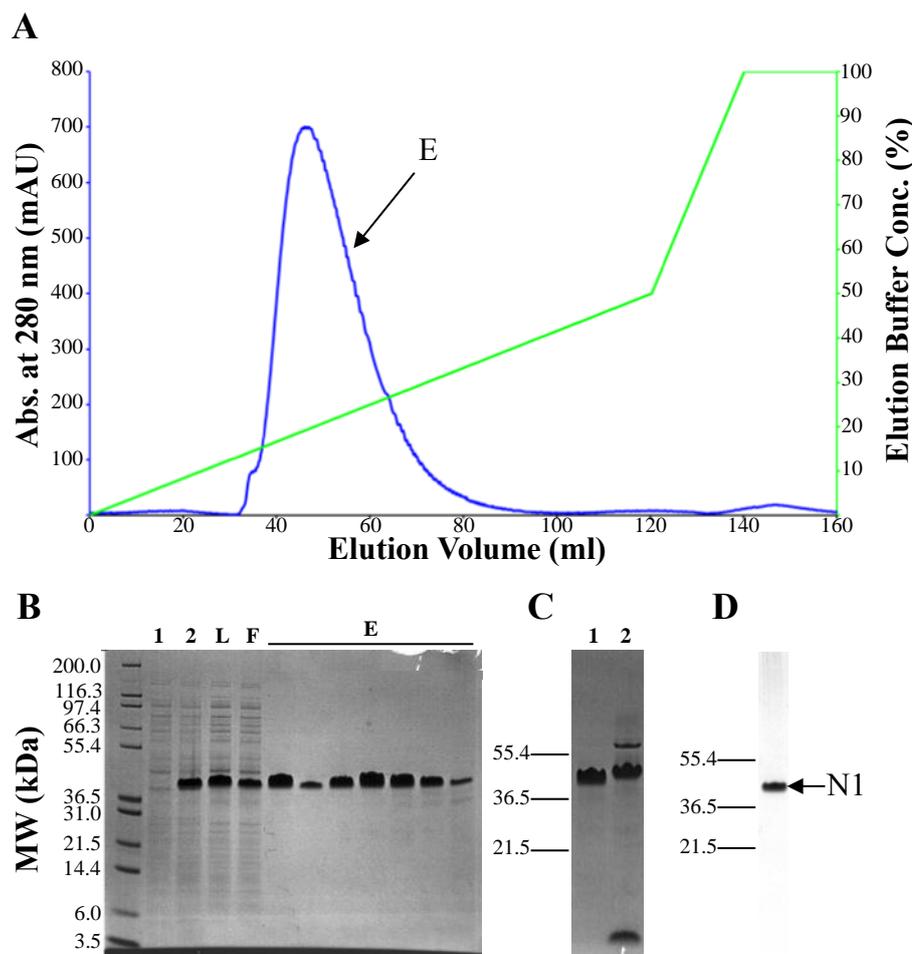


Figure 13 Purification of His₆-N1. (A) Nickel affinity chromatography elution profile of His₆-N1. Protein elution was monitored by absorbance at 280 nm (blue) relative to the concentration of the elution buffer (green). (B) SDS-PAGE analysis of the nickel affinity purification step containing pre- (1) and post-induction (2), soluble cell extract (L), flow-through (F) and elution peak (E). (C) SDS-PAGE analysis of the His₆-tag cleavage by HRV-3C protease. Before cleavage (1), post cleavage (2). (D) SDS-PAGE analysis of purified N1.

Further purification was required for AF1 and N2N3. AF1 is prone to C-terminal degradation and as proteins were expressed with a N-terminal His₆-tag, degradation products could not be separated by nickel affinity chromatography. Therefore, following cleavage of the His₆-tag and removal of the protease, intact AF1 was separated from degradation products by anion exchange chromatography (Figure 14). AF1 eluted at a NaCl concentration of approximately 180 mM. The final purification step applied to N2N3 was size-exclusion chromatography (SEC) to separate oligomeric and monomeric species (Figure 15). The equilibrium between the oligomeric states of N2N3 and AF1 is not dynamic (Stemberk et al., manuscript

in preparation), and a single SEC step was required to produce homogenous species. The purity of each rFnBPA protein produced in the aforementioned work was verified by SDS-PAGE as over 95%. ESI-MS (Table 11) verified the molecular weight along with the level of isotope incorporation of labelled species.

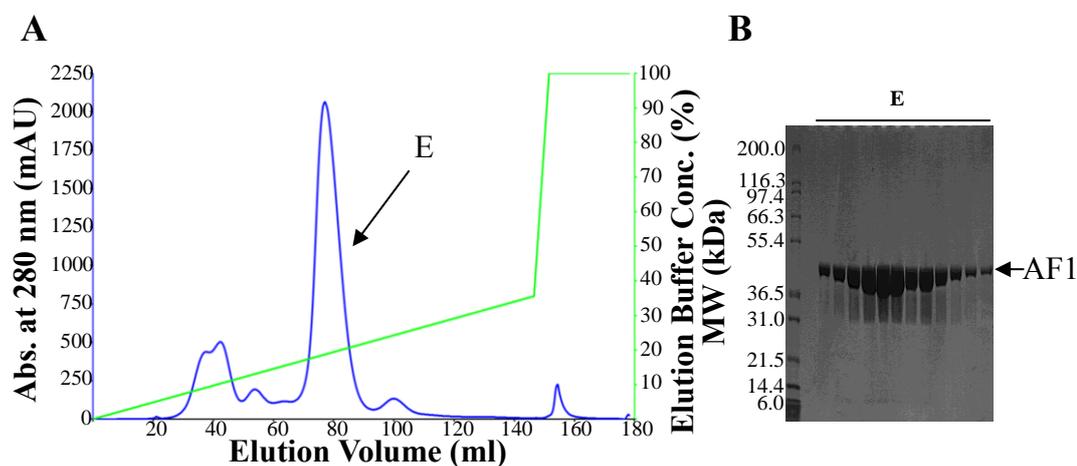


Figure 14 Anion exchange chromatography of AF1. (A) Anion exchange chromatography elution profile of AF1 monitored by absorbance at 280 nm (blue) relative to the concentration of elution buffer (green). (B) SDS-PAGE analysis the anion exchange step.

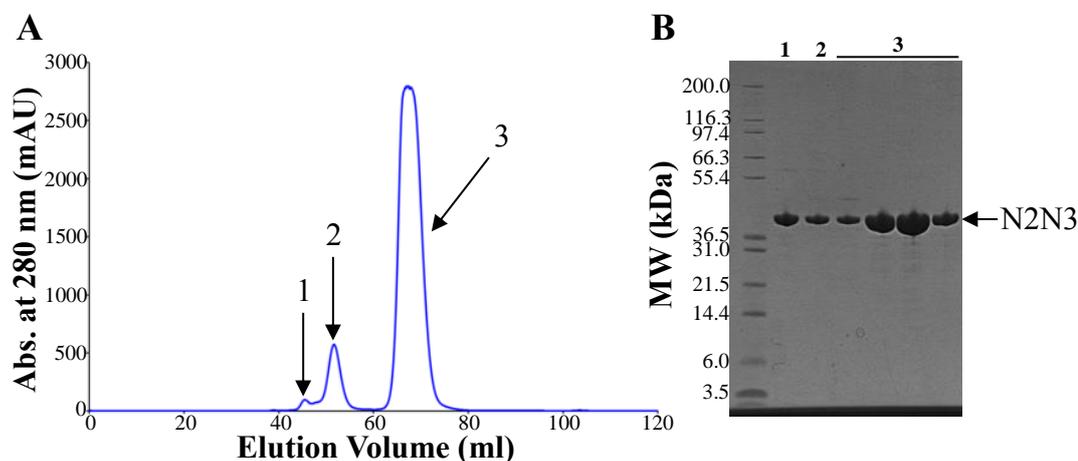


Figure 15 Size-exclusion chromatography of N2N3. (A) Size-exclusion chromatography elution profile of N2N3 monitored by absorbance at 280 nm. (B) SDS-PAGE analysis of the size exclusion step. Numbered lanes correspond to the three elution peaks.

3.5.1 Expression and Purification of ^2H , ^{13}C , ^{15}N N1

Substituting non-labile hydrogen atoms for deuterium introduces properties more favourable for the NMR analysis of proteins (Gardner and Kay, 1998). The specific benefits of deuterating proteins will be discussed in section 4.5.1. However, the expression of ^2H , ^{13}C , ^{15}N N1 differed from the procedure applied for non-deuterated N1 and those differences will be presented here. Prior to expression of N1 in 98% $^2\text{H}_2\text{O}$ M9 minimal media, the *E. coli* expression strain was acclimatised to growth in $^2\text{H}_2\text{O}$ by incrementally increasing the concentration in 5 ml cultures (section 2.3.3). Finally, a 50 ml 98% $^2\text{H}_2\text{O}$ M9 minimal media starter culture was inoculated into baffled flasks containing 500 ml 98% $^2\text{H}_2\text{O}$ M9 minimal media. Cell cultures grew more slowly in minimal media as described in section 3.4.2, and growth was even slower in 98% $^2\text{H}_2\text{O}$ minimal media. After approximately 8 h an OD_{600} of 0.7 was reached and protein expression was induced by the addition of 0.5 mM IPTG. Following overnight growth at 30 °C, cell cultures were harvested and proteins purified as described previously. The level of deuterium incorporation was determined by ESI-MS (Table 11).

3.5.2 Expression of Selectively Unlabelled ^2H , ^{13}C , ^{15}N N1

To aid the sequence specific resonance assignment of N1 (section 4.5), samples of ^2H , ^{13}C , ^{15}N N1 with particular amino acids ‘unlabelled’ were produced via the method developed by Rasia *et al.* (Goto *et al.*, 1999, Rasia *et al.*, 2012). Briefly, supplementing growth media with the metabolic precursor to a specific amino acid results in their incorporation into expressed proteins (section 4.5.2). The addition of non-isotopically enriched precursors to enriched expression media results in fully labelled proteins with specific amino acids unlabelled, namely leucine, valine, isoleucine, phenylalanine and tyrosine. Proline can also be unlabelled by the addition of L-proline. 4-hydroxyphenylpyruvic acid, sodium 3-methyl-oxobutyrate and sodium 2-oxobutyrate, precursors to tyrosine, leucine/valine and isoleucine, respectively, and L-proline were added 1 hr prior to the induction of protein expression according to the method developed Rasia *et al.* (Rasia *et al.*, 2012). All the unlabelled species, except tyrosine unlabelled N1, were successfully over-expressed. Supplementation of cell cultures with 4-hydroxyphenylpyruvate resulted

in cell death and no protein was expressed. It is uncertain why expression of tyrosine unlabelled N1 resulted in cell death as this behaviour has not been previously reported and the precursor has no known toxicity when used as a growth supplement (Rasia et al., 2012). However, N1 contains only one tyrosine residue and the impact of unlabelling this residue for the NMR analysis of N1 was likely to be minimal and the matter was not investigated further. Unlabelled ^2H , ^{13}C , ^{15}N N1 was purified as outlined in section 4.7 and an example of the procedure is presented in Figure 16. The successful incorporation of unlabelled residues was confirmed by NMR analysis (section 4.5.2).

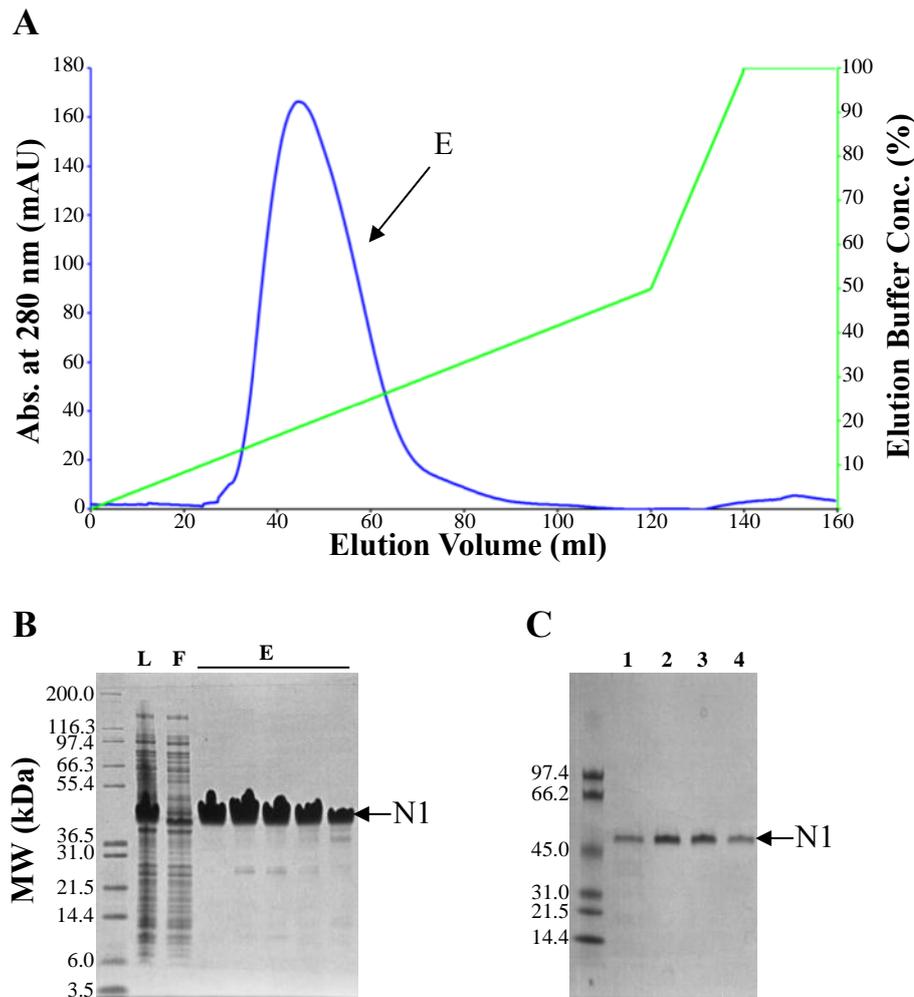


Figure 16 Purification of selectively unlabelled ^2H , ^{13}C , ^{15}N N1 proteins. (A) Nickel affinity chromatography elution profile of isoleucine unlabelled N1. Protein elution was monitored by absorbance at 280 nm (blue). (B) SDS-PAGE analysis of the nickel affinity purification step containing soluble cell extract (L), flow-through (F) and elution peak (E). (C) SDS-PAGE analysis of purified ^2H , ^{13}C , ^{15}N N1 (1), isoleucine unlabelled N1 (2), leucine/valine unlabelled N1 (3) and proline unlabelled N1 (4).

Table 11 Summary of rFnBPA properties and isotope incorporation. Isotope incorporation of labelled species was calculated using the theoretical weight (MW_t) and molecular weight determined by ESI-MS (MW). The extinction coefficient (ϵ) and theoretical pI of proteins are also included (Protparam)(Gasteiger et al., 2003).

rFnBPA	MW_t (Da)	MW (Da)	ϵ	Theoretical pI
N1	17136	17132	1490	4.66
N1N2	32947	32949	7450	4.86
fA	52929	52926	33810	5.09
N2N3	36551	36551	32320	5.59
AF1	40693	40690	36790	5.18
^{15}N N1	17350	17346 (98%)	1490	4.66
^{15}N N1N2	33351	33349 (99%)	7450	4.86
^{15}N fA	53575	53563 (98%)	33810	5.09
^{15}N N2N3	36992	36987 (99%)	32320	5.59
^{13}C , ^{15}N N1	18059	18024 (96%)	1490	4.66
^2H , ^{13}C , ^{15}N N1	19225	17721 (72%)	1490	4.66

3.6 Expression and Purification of GST-Fusion Proteins

3.6.1 Molecular Biology of GST-rFnBPA Constructs

The DNA constructs of N-terminally GST-tagged N1, N1N2 and fA were produced using standard molecular biology techniques as described in section 2.2. The pGEX-6P-1 vector, containing a PreScission Protease cleavable N-terminal GST-tag, was chosen as the expression vector. Primers used in PCR to generate the requisite constructs are listed in METHODS. The PCR products were ligated into the linearised expression vector using BamH1 and Xho1 restriction sites. The successful ligation of the inserts into the expression vector was confirmed using the restriction sites flanking the insert, or PCR using the primers used to generate the insert. The correct sequence of all GST-fusion constructs was confirmed by an in-house sequencing service (Technology Facility, University of York).

3.6.2 Expression and Purification of GST-rFnBPA Proteins

Various conditions, including temperature, time of induction and final concentration of IPTG, were screened to achieve optimal expression of the GST-rFnBPA proteins. Based on small-scale tests, optimum expression was achieved in *E. coli* BL21 (DE3) Gold at 37 °C with expression induced at a final IPTG concentration of 0.4 mM following growth to an OD₆₀₀ of between 0.5 and 0.6. Following induction, cell cultures were typically grown for 4 h and harvested by centrifugation prior to cell lysis (section 2.3.7).

Cell extracts were prepared by lysing cell cultures by sonication and separation of the soluble fraction from cell debris by centrifugation. rFnBPA proteins were isolated from the cell extracts by GST affinity chromatography. Cell extracts were passed over a 5 ml GST affinity column (GE Healthcare) 3 times to achieve maximum binding. Proteins were eluted from the column following the addition of 10 mM reduced glutathione (GSH). Fractions containing protein were pooled and dialysed against PBS to remove GSH. An example of the purification procedure is given in Figure 17.

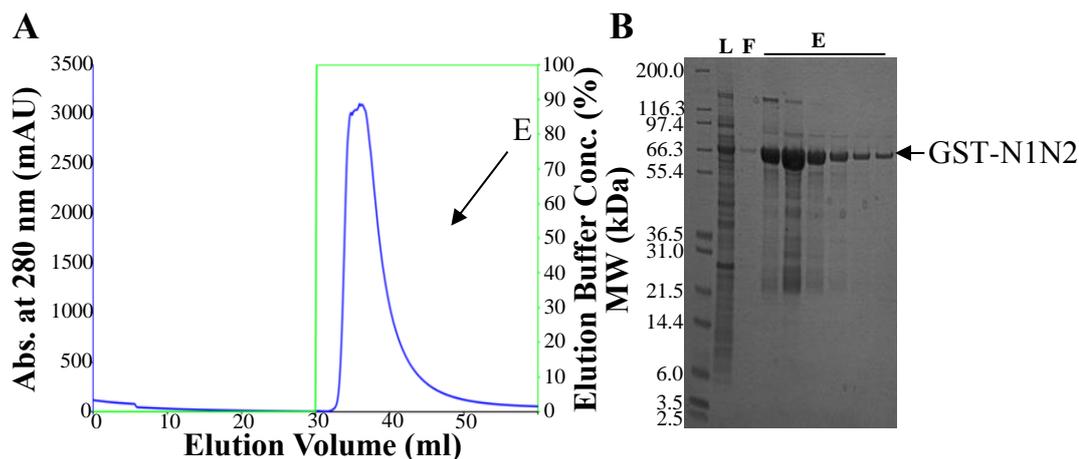


Figure 17 Purification of GST-N1N2. (A) GST affinity chromatography of GST-N1N2. Protein elution was monitored by absorbance at 280 nm (blue) relative to the concentration of elution buffer (green). (B) SDS-PAGE analysis of GST affinity purification step including soluble cell extract (L), flow through (F) and elution peak (E).

Although the expression and purification of GST-N1, N1N2 and fA was successful, all three proteins are susceptible to C-terminal degradation. In an attempt to reduce the degradation, protease inhibitors were added to the soluble cell extracts prior to GST affinity purification. However, the purified species still prone to degradation suggesting it is not the result of proteolysis. Despite the instability of the GST-rFnBPA proteins, a sufficient quantity of intact material was present in purified protein samples and used in subsequent studies (section 5.3.1).

3.7 Summary

A number of His₆-rFnBPA constructs covering various regions of the A domain (Figure 11) were successfully expressed and purified. The procedure for the optimum expression and purification of His₆-rFnBPA was developed previously (Stemberk et al., manuscript in preparation). The majority of His₆-rFnBPA proteins were also uniformly labelled with ¹⁵N for subsequent NMR studies. ¹³C, ¹⁵N N1 was expressed in M9 minimal media supplemented with ¹⁵NH₄Cl and ¹³C-glucose and high levels of isotope incorporation were achieved. Deuterated N1 was also successfully expressed and purified. Following an initial acclimatisation to growth in ²H₂O, deuterated N1 was expressed and purified following the same procedure employed for the purification of the His₆-rFnBPA proteins. Selectively unlabelled

variants of N1 were expressed via recently developed methodology (Rasia et al., 2012). Isoleucine, proline and leucine/valine residues were unlabelled following the supplementation of expression media with metabolic precursors. However, tyrosine unlabelling was unsuccessful due to unexplained cell death following the induction of protein expression. DNA constructs encoding GST-rFnBPA proteins were successfully produced via standard molecular biology techniques. The expression conditions were optimised on a small scale and implemented for large-scale protein production. GST-rFnBPA fusion proteins were purified by GST affinity chromatography and despite susceptibility to C-terminal degradation, sufficient amounts of intact recombinant protein was produced for subsequent investigations.

4 Structural Characterisation of N1

4.1 Introduction

FnBPA is a major *S. aureus* virulence factor associated with a number of life-threatening conditions (Fitzgerald et al., 2006, Edwards et al., 2011, Shinji et al., 2011, Arrecubieta et al., 2006). FnBPA's contribution to *S. aureus* virulence is mainly attributed to its recognition of multiple ECM proteins such as Fn, Fg and elastin (Bingham et al., 2008, Keane et al., 2007a, Roche et al., 2004, Wann et al., 2000). Interactions with these ligands are crucial for host colonisation and immune evasion (Massey et al., 2001, Edwards et al., 2010, Fitzgerald et al., 2006). FnBPA binds Fn through the well-characterised tandem β -zipper mechanism (Bingham et al., 2008, Schwarz-Linek et al., 2003). FnBPA interacts with Fg via the N-terminal A domain (Keane et al., 2007b). The binding site is located within the mostly hydrophobic cleft between the N2 and N3 subdomains (Stemberk et al., manuscript in preparation). The regions of FnBPA involved in ligand binding, the N2N3 subdomains and 11 FnBRs for Fg and Fn binding, respectively, are well characterised structurally. Contrastingly, the structure of the N1 subdomain has not been investigated despite being the most highly conserved region of the N-terminal A-domain (Loughman et al., 2008) and the domain boundaries being reasonably well-defined.

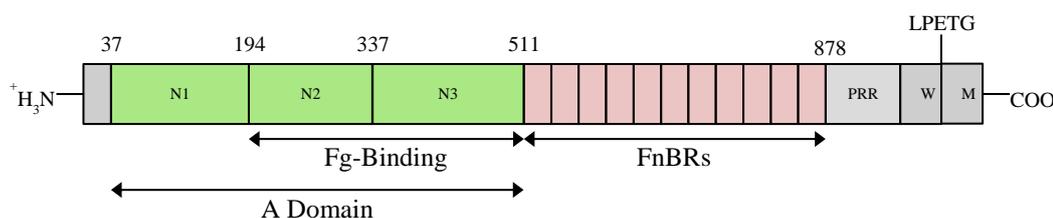


Figure 18 Schematic Representation of FnBPA. The N-terminal region of FnBPA comprises a short signalling and the A domain, consisting of the N1, N2 and N3 subdomains. The 11 highly conserved FnBRs, proline-rich repeat region (PRR), wall (W) and membrane (M) spanning domains constitute the C-terminus. FnBPA is anchored to the cell wall by an LPETG motif.

4.2 Aims

The aim of this work was to determine the structure of the N1 subdomain. The behaviour of N1 in the presence and absence of other regions of the A domain was investigated using NMR spectroscopy and circular dichroism (CD). The effect of temperature and pH on the conformation of N1 was also investigated. The resonance assignment of N1 was conducted using recently developed labelling strategies (Rasia et al., 2012). Secondary chemical shift analysis (Wishart and Sykes, 1994, Wishart et al., 1992) and structural propensity calculations (Marsh et al., 2006) were performed to identify residual secondary structure elements within N1.

4.3 Sequence Analysis of N1

IDPs typically contain few residues that would form the hydrophobic core of a folded protein (Val, Leu, Ile, Phe, Trp and Tyr) (Dyson and Wright, 2005, Midic and Obradovic, 2012). A high proportion of residues in IDPs are instead polar or charged ‘disorder-promoting residues’ (Glu, Pro, Ser, Lys and Gln). Simple sequence analysis of N1 reveals characteristics associated with IDPs (Figure 19). 46% of the amino acids in N1 promote disorder, with a further 20 alanine residues that can sometimes have similar effects. Contrastingly, only 20 of the 158 residues are considered order promoting, and these are sporadically positioned throughout the protein.

```

ASEQKTTTVE ENGNSATDNK TSETQTTATN VNHIEETQSY NATVTEQPSN
ATQVTTEEAP KAVQAPQTAQ PANIETVKEE VVKEEAKPQV KETTQSQDNS
GDQRQVDLTP KKATQNQVAE TQVEVAQPRT ASESKPRVTR SADVAEAKEA
SNAKVE TG

```

Figure 19 Presence of Order and Disorder Promoting Residues in N1. Order and disorder promoting amino acids are highlighted in blue and red, respectively. N1 contains 74 disorder-promoting residues compared to 20 that promote order.

A number of bioinformatics tools were used to analyse N1 and predict any structural elements it contained. These included Jpred (Cole et al., 2008), IUPred (Dosztanyi et al., 2005a), FoldIndex (Prilusky et al., 2005) and metaPrDOS (Ishida and Kinoshita,

2008). Each tool utilises slightly different methodology and algorithms to predict disorder on a residue-by-residue basis. Jpred uses multiple sequence alignments to calculate secondary structure contributions and scores suggested conformations in terms of likelihood (Cuff and Barton, 2000, Cole et al., 2008). IUPred estimates fold stabilising interactions based on amino acid composition and the sequential environment of residues (Dosztanyi et al., 2005a, Dosztanyi et al., 2005b). FoldIndex calculates the average hydrophobicity of residues and net charge of the protein to predict whether a given sequence is ordered or disordered (Prilusky et al., 2005). Using a different approach, metaPrDOS integrates results from other algorithms to predict disorder (Ishida and Kinoshita, 2008). Each of the aforementioned methods predicted a high level of disorder for N1 (Figure 20).

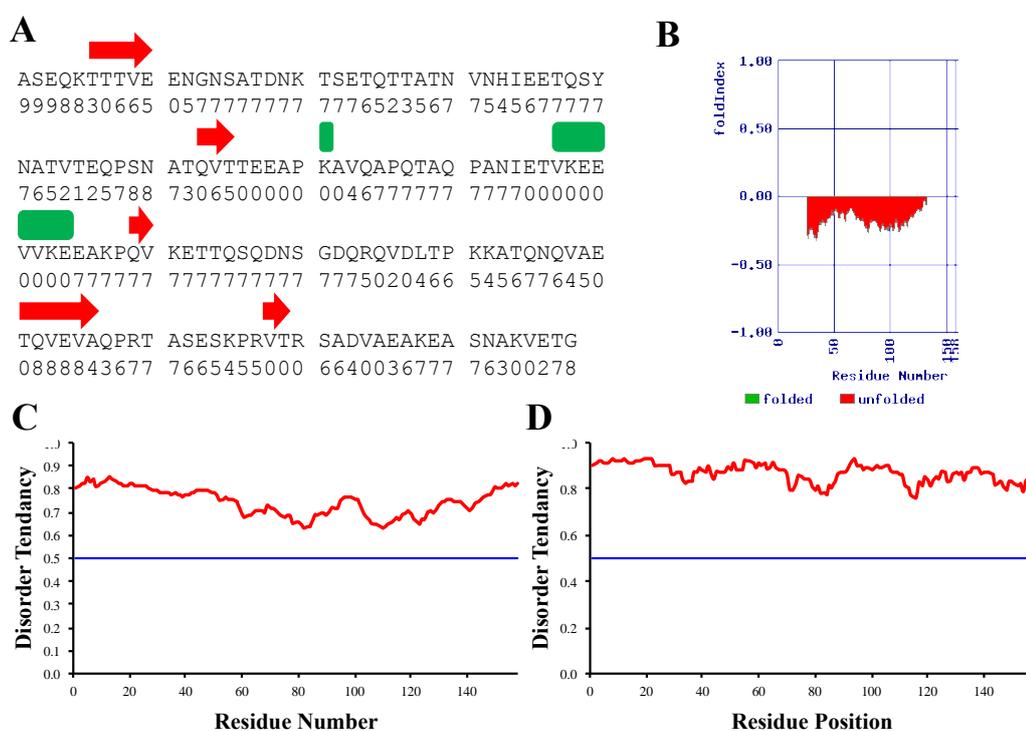


Figure 20 Predicted Disorder in N1. Secondary structure predictions of N1 using (A) Jpred, (B) FoldIndex, (C) metaPrDOS and (D) IUPred. The consensus of all four predictions is that N1 is an IDP. Small elements of secondary structure are predicted by Jpred, however the likelihood score assigned to these regions is very low.

4.4 Preliminary NMR Analysis of N1

NMR spectroscopy is a powerful tool for studying protein conformations. Chemical shifts are a sensitive indicator of structural elements within a protein. Therefore, even from complicated 1D ^1H spectra inferences can be made about a protein's structure. Typically β -sheet conformations give rise to a broad dispersion of resonances in the amide region (approximately 6 – 10 ppm) of the ^1H spectrum, and individual methyl groups are often resolved at around 0 ppm (Figure 21A). The narrow ^1H signal dispersion also results in characteristic ^1H , ^{15}N HSQC spectra (Figure 21C). IDPs exhibit poor signal dispersion, with peak overlap and signal degeneracy commonly observed throughout the spectrum (Figure 21B and D). Resonances arising from amide protons cluster around 8 ppm with a dispersion of approximately 0.5 ppm. Aliphatic resonances also exhibit high levels of degeneracy. Helical proteins generally give rise to features between the two aforementioned extremes. Although, signal dispersion in the amide region of the ^1H spectrum is typically narrow, the aliphatic region exhibits features similar to those observed for β -sheet proteins. An initial evaluation of a 1D ^1H spectrum for N1 suggests it is intrinsically disordered (Figure 22).

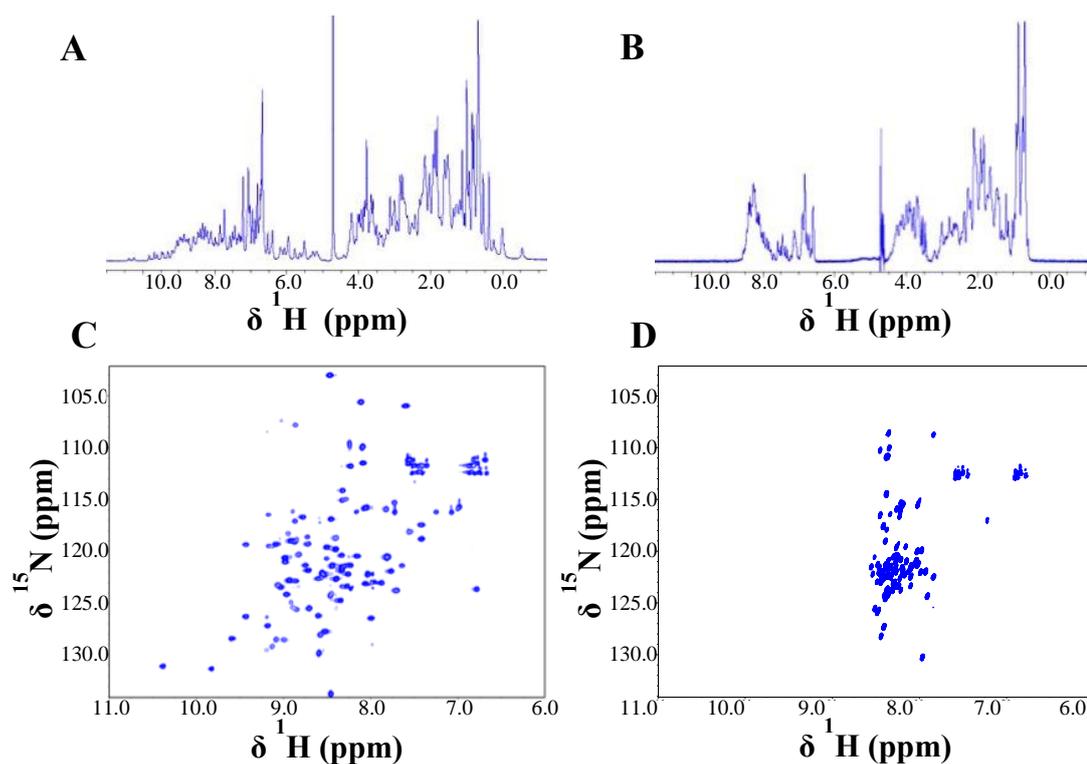


Figure 21 NMR spectra of β -sheet and disordered proteins. (A) 1D ^1H and (B) ^1H , ^{15}N HSQC NMR spectra of 8F19F1, an 11 kDa β -sheet module pair of Fn (Atkin et al., 2010). (C) 1D ^1H and (D) ^1H , ^{15}N HSQC spectra of a disordered region of BBK32, a surface protein from *Borrelia burgdorferi* (Harris, unpublished). Spectra were acquired at 298 K on a 700 MHz spectrometer.

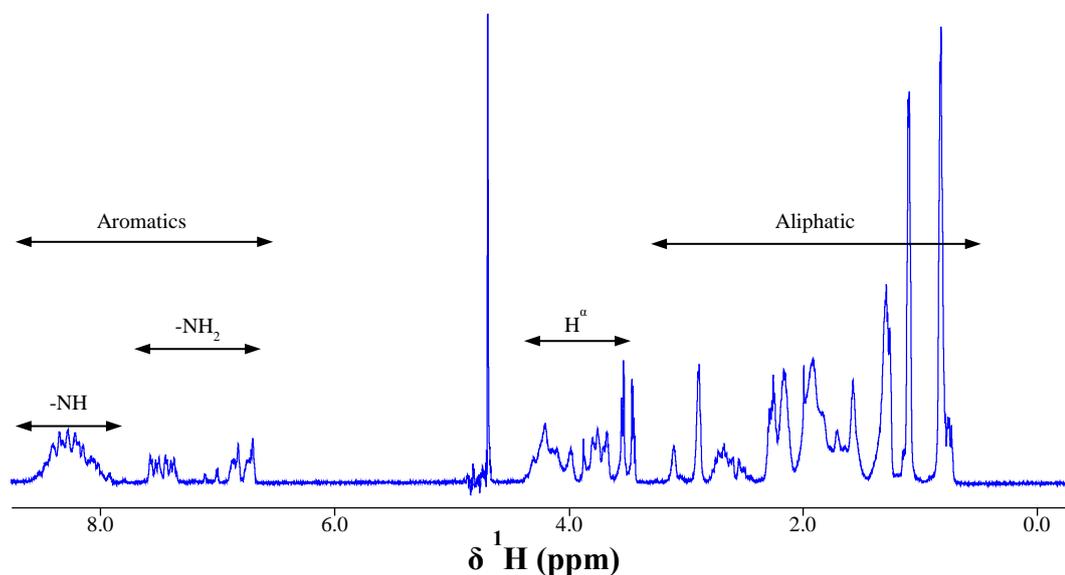


Figure 22 1D ^1H NMR Spectrum of N1. The spectrum was acquired at pH 6.0 and 298 K on a 700 MHz spectrometer.

Uniformly ^{15}N labelled N1 was expressed to acquire a ^1H , ^{15}N HSQC spectrum to confirm the perceived disorder in the ^1H spectrum. Resolution of degenerate proton chemical shifts via a second dimension enables a more accurate estimation of secondary structure to be made. ^1H , ^{15}N HSQC spectra show correlations between the ^1H and ^{15}N nuclei in amide groups. Thus, for every amino acid in a protein except proline, a peak is seen in the spectrum corresponding to a backbone amide group. Depending on the pH, certain side-chain amide groups are also visible. The ^1H , ^{15}N HSQC of N1 again exhibits the characteristic features of an IDP (Figure 23). In addition to degeneracy of amide ^1H shifts, the ^{15}N chemical shift dispersion is also low with the middle of the spectrum particularly congested. Therefore, the resolution of individual peaks is poor. The peaks have narrow linewidths, a feature indicative of IDPs due to the slower spin relaxation rates relative to globular proteins. One further feature symptomatic of IDPs is overlap of Gln and Asn side-chains amide resonances. In folded proteins the pairs of peaks observed for these $-\text{NH}_2$ groups are well resolved and easily identified. Despite N1 containing 28 Asn and Gln residues in total, there is significant signal overlap in this region and only a small number of peaks observed.

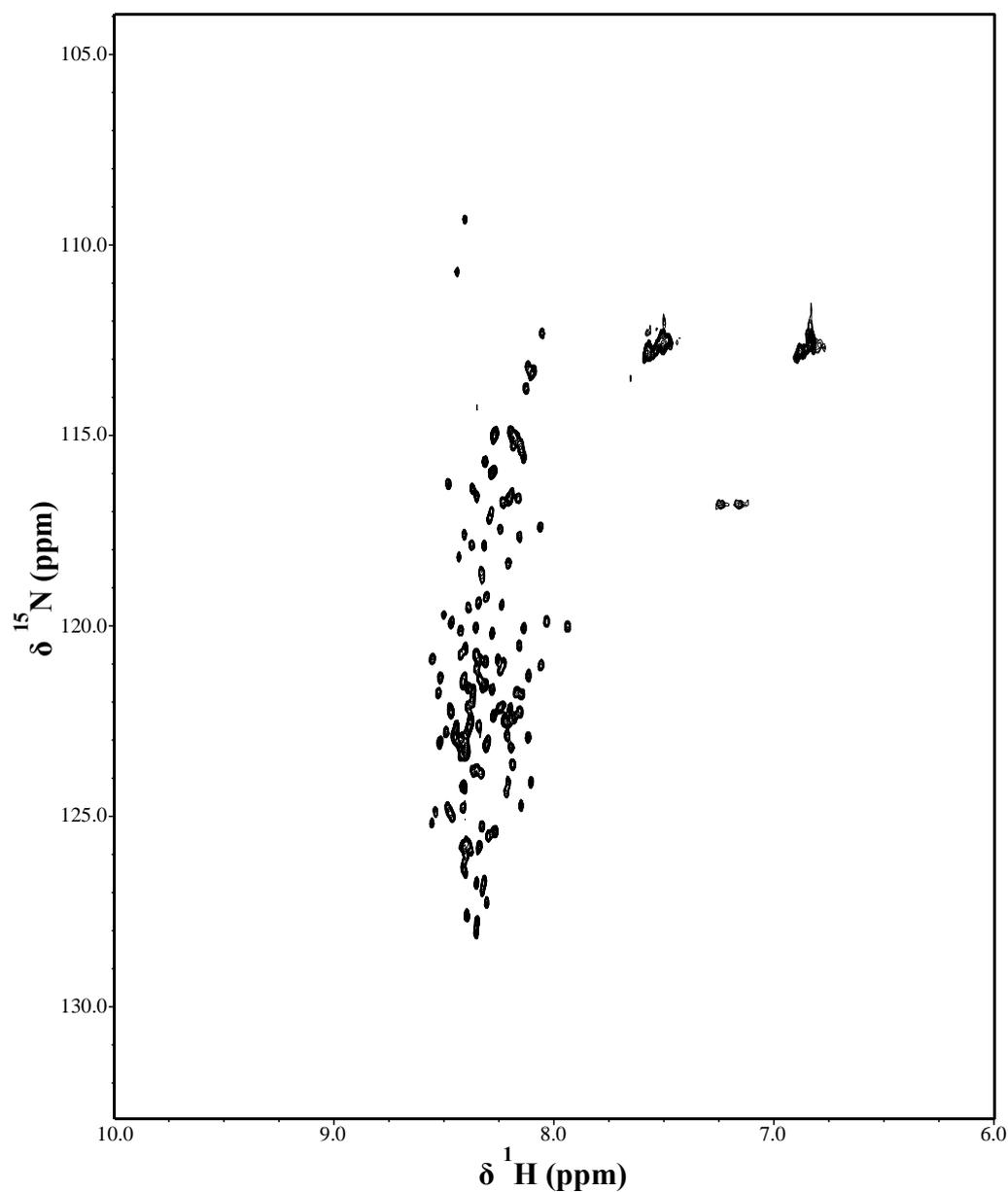


Figure 23 ^1H , ^{15}N HSQC spectrum of N1. The spectrum was recorded at pH 6.0 and 298 K on a 700 MHz spectrometer.

4.4.1 pH Does Not Affect the Conformation of N1

To assess the effect of pH on the conformation of N1, a series of ^1H , ^{15}N HSQC spectra were recorded over a range of pH's (Figure 24). The initial spectrum was acquired at pH 5.0, with subsequent spectra recorded at incrementally increasing pH's up to pH 8.0. Across the pH range it is clear that N1 remains disordered as

most resonances remain clustered in the middle of the spectrum. As the pH increases the rate of exchange between amide protons and the solvent also increases (Englander et al., 1996). The signals of labile protons in NMR spectra consequently decrease in intensity until they can no longer be observed. The vast majority of peaks in the N1 spectrum decrease in intensity or disappear as the pH increases, suggesting that most of N1 amide protons are freely accessible to the solvent. This is further evidence that N1 is an IDP.

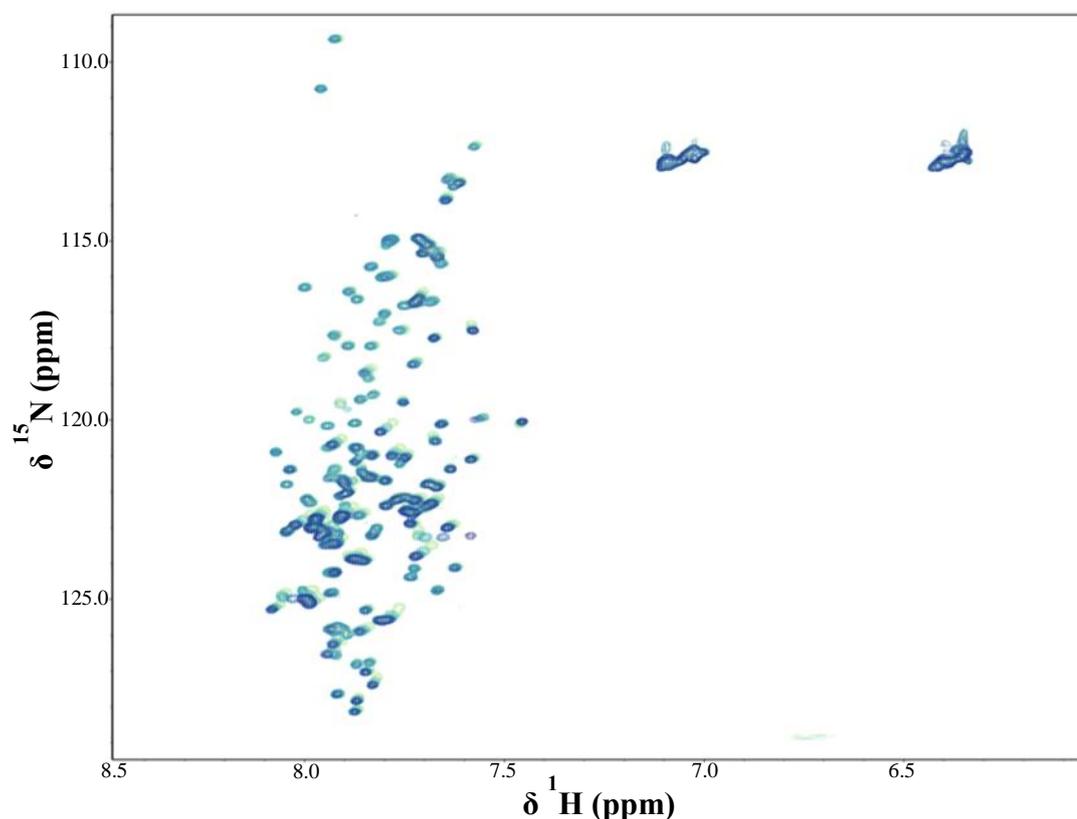


Figure 24 ^1H , ^{15}N HSQC spectra of N1 over a range of pH's. Data was acquired at 298 K on a 700 MHz spectrometer. Spectra are coloured from light to dark are the pH increases from pH 5 to 8.

4.4.2 N1 is Disordered in the A Domain

The disorder observed for N1 may not be indicative of its confirmation in the context of the full A domain. The disorder may arise from a C-terminal truncation if the domain boundaries between N1 and N2 were incorrectly defined. The subdomain boundaries within the A domain of FnBPA were originally proposed relative to *S. aureus* Clumping factor A (Deivanayagam et al., 2002). Structural studies found two

distinct domains, N2 and N3, comprised a recombinant ClfA fragment encompassing residues 221 to 559 (Deivanayagam et al., 2002). The N-terminal region of the A domain (residues 45 – 220), susceptible to protease activity, was termed N1. A subsequent model suggested the structure and organisation of the A domain of FnBPA was analogous to ClfA (Keane et al., 2007b). This was confirmed by a recently solved crystal structure of the independently folded N2 and N3 subdomains (Stemberk et al., manuscript in preparation). However, the precise boundary between the N1 and N2 subdomains has not been defined. Incorrectly defining the domain organisation of a protein can have a significant impact on subsequent structural and functional studies. For example, the revision of domain boundaries within the Fn binding region of FnBPA, identifying 11 highly conserved FnBRs, led to significantly increased understanding of Fn recognition and the specificity of particular motifs for certain regions of Fn (Schwarz-Linek et al., 2003, Meenan et al., 2007). Therefore, correctly identifying the C-terminal boundary of N1 is essential prior to further structural and functional studies.

An alternative explanation for the perceived disorder of N1 is that the presence of the N2 and N3 subdomains could act to stabilise a conformation in N1 that we do not observe when they are absent. Establishing the correct interfaces between domains can significantly impact a protein's conformation as evidenced in the B-repeat region of *S. aureus* surface protein SasG (Gruszka et al., 2012). In this instance, the E segment of the SasG B-repeats is disordered when preceded by a G5 domain, whereas a C-terminal G5 domain introduces a stable fold in the E segment. To address these questions, uniformly ^{15}N labelled N1N2, N2N3 and fA proteins were expressed and ^1H , ^{15}N HSQC spectra recorded.

A ^1H , ^{15}N HSQC of N1N2 suggests that the protein comprises two regions adopting different conformations (Figure 25). The difference in peak linewidths between folded proteins and disordered proteins reflects the relative correlation time (τ_c) of each component. In essence τ_c is a measure of the rate of random molecular motion of a molecule (Keeler, 2005). Thus, dynamic molecules such as disordered proteins typically have a shorter τ_c relative to globular proteins. Spin-spin, or T_2 , relaxation processes are more efficient when molecular motion is slower and peak linewidth is inversely proportional to T_2 (Levitt, 2001). Consequently, shorter τ_c corresponds to

an increased T_2 relaxation time and in turn narrower linewidths. Therefore, the broader, more dispersed resonances in the N1N2 spectrum indicate a folded β -sheet component, N2, whilst the narrower, centrally localised peaks could represent a disordered component, N1. By superimposing the N1 spectrum onto the ^1H , ^{15}N HSQC of N1N2 it is clear that the disordered features are attributable to N1, and the more dispersed peaks to the independently folded N2 subdomain. The ^1H , ^{15}N HSQC of fA is relatively poor due to the size of the protein and differing dynamics of the domains it contains (Figure 26). However, the spectrum exhibits similar features to the ^1H , ^{15}N HSQC of N1N2 with peaks attributable to a disordered domain and dispersed peaks indicative of folded regions being clearly visible. A ^1H , ^{15}N HSQC of N2N3 confirms that under the conditions used to record the N1N2 and fA spectra, both subdomains adopt stable β -strand conformations as expected (Figure 26). Therefore, under the conditions tested N1 appears to be an IDP and incorrect domain boundaries are not responsible for the disorder.

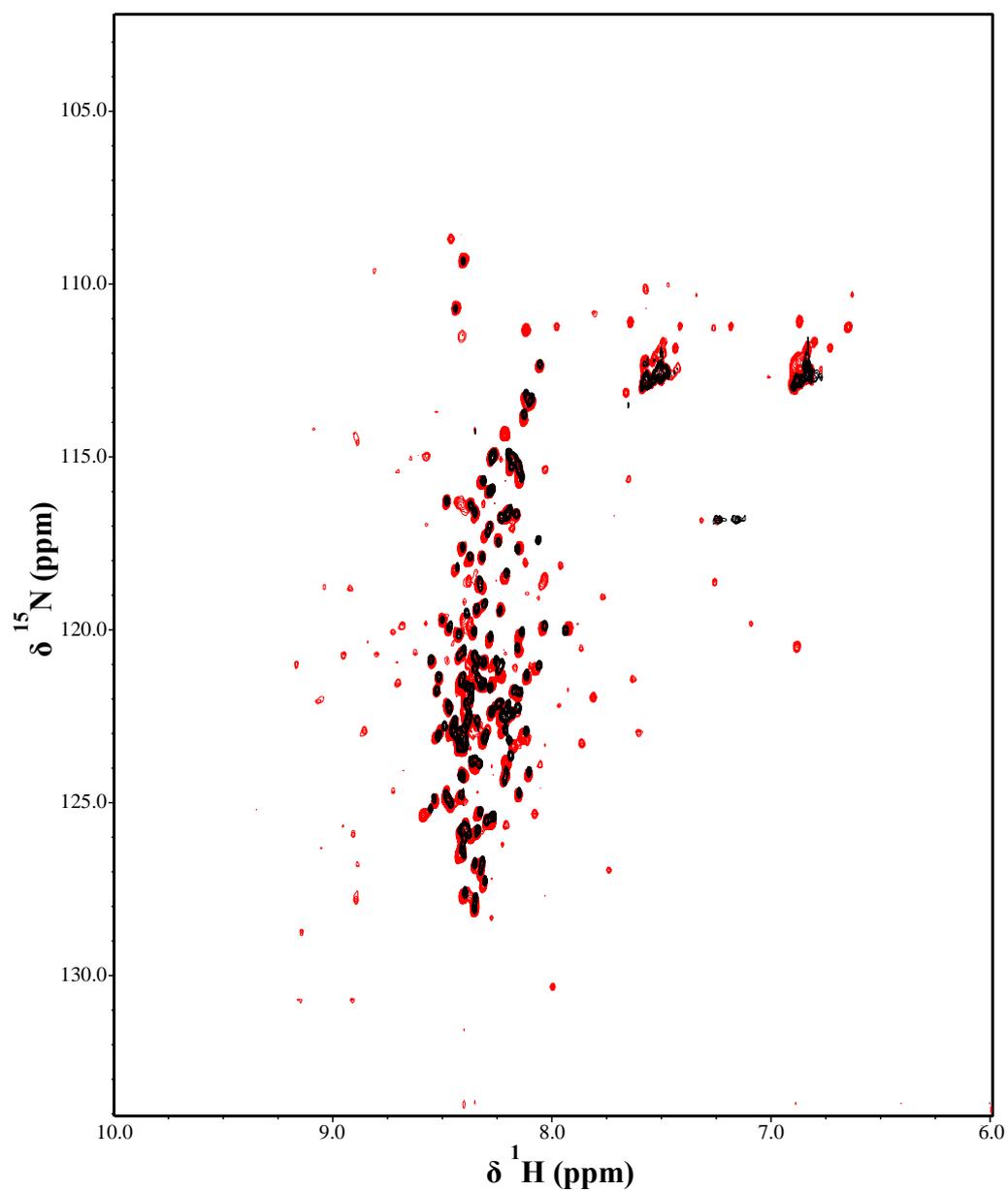


Figure 25 Comparison of the ^1H , ^{15}N HSQC spectra of N1 and N1N2. ^1H , ^{15}N HSQC spectrum of N1 (black) is superimposed on the ^1H , ^{15}N HSQC of N1N2 (red). Spectra were acquired at pH 6.0 and 298 K on a 700 MHz spectrometer.

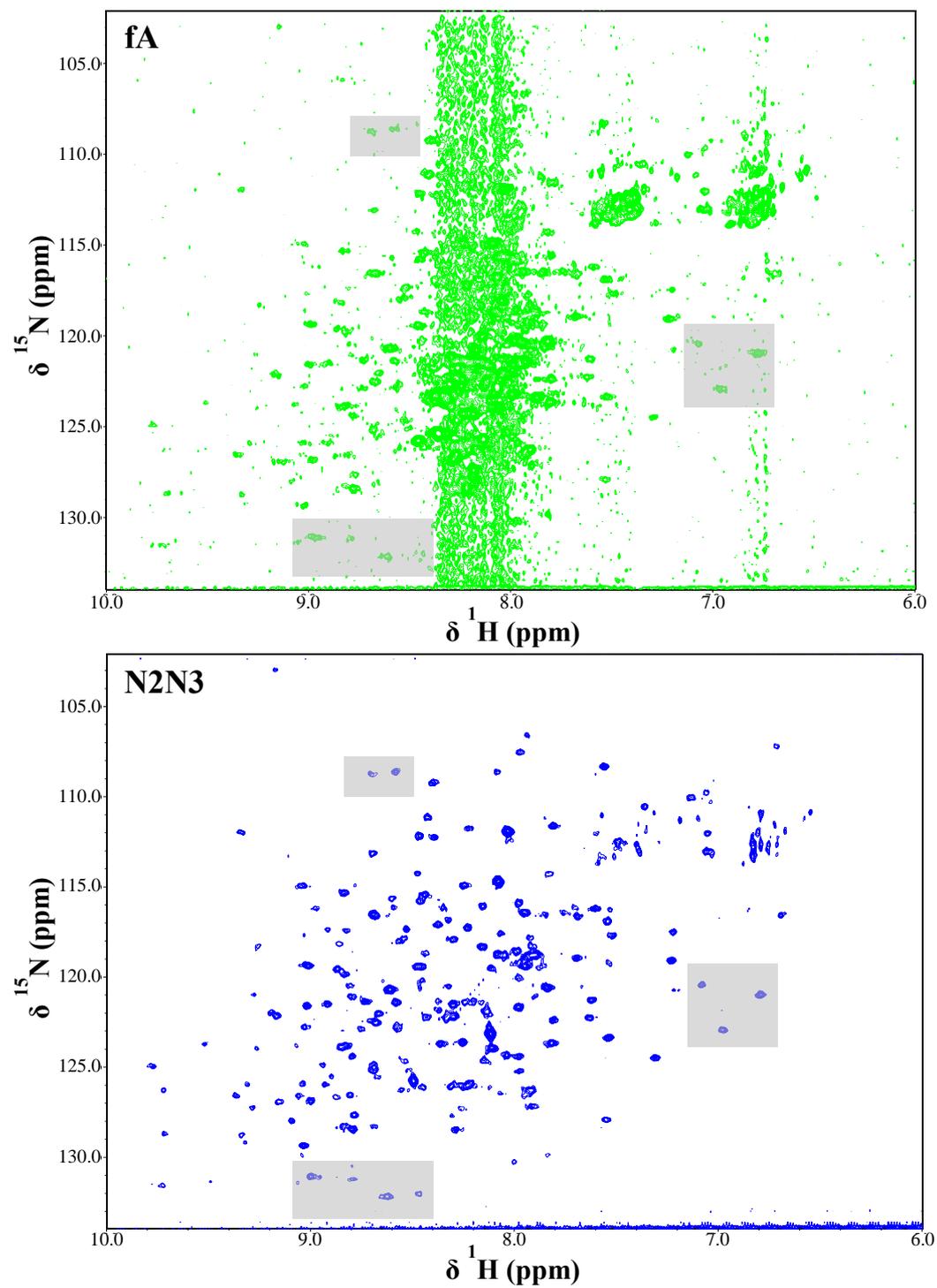


Figure 26 ^1H , ^{15}N TROSY-HSQC spectra of fA and N2N3. Both spectra were recorded at pH 7.4 and 298 K on a 700 MHz spectrometer. Examples of features common to both spectra are indicated.

4.4.3 Circular Dichroism of the A Domain

Despite the NMR spectra suggesting that N1 is disordered a small amount of ambiguity exists as helical proteins can exhibit similarly narrow amide chemical shift dispersions. Therefore, CD spectra of the various A domain proteins were recorded with the results corroborating the NMR data (Figure 27).

Helical proteins give rise to characteristic features in a CD spectrum, most notably double minima between approximately 205 and 220 nm (Nelson et al., 2008). IDPs produce a strong minimum at around 195 nm. β -sheet conformations are less distinctive in CD spectra, although contributions to overall protein structure can be estimated (Chen et al., 1974). N1 exhibits no features characteristic of α -helical conformations, with the minimum at approximately 198 nm indicative of disorder (Figure 27). The spectra of N1N2 and fA also show features attributable to disordered regions (Figure 27). Deconvolution of the CD spectra using DichroWeb algorithms (Whitmore and Wallace, 2004) predicts that N1 is almost completely disordered, with virtually no helical content (Table 12). The β -sheet contribution increases in N1N2 and fA, as expected due to the presence of folded subdomains. The contribution of β -strand in these proteins is surprisingly low given the known structure. However this is likely due to CD's unsuitability for studying β -sheet proteins. Taken with the NMR data it can be concluded that N1 is an IDP.

Table 12 Deconvolution of FnBPA CD spectra Deconvolution of the CD spectra of N1, N1N2 and fA using DichroWeb software (Whitmore and Wallace, 2004).

	Helix 1	Helix 2	Strand 1	Strand 2	Turns	Unordered	Total
N1	0.01	0.02	0.05	0.03	0.05	0.84	1.00
N1N2	0.00	0.03	0.11	0.07	0.12	0.67	1.00
fA	0.00	0.03	0.20	0.10	0.14	0.53	1.00

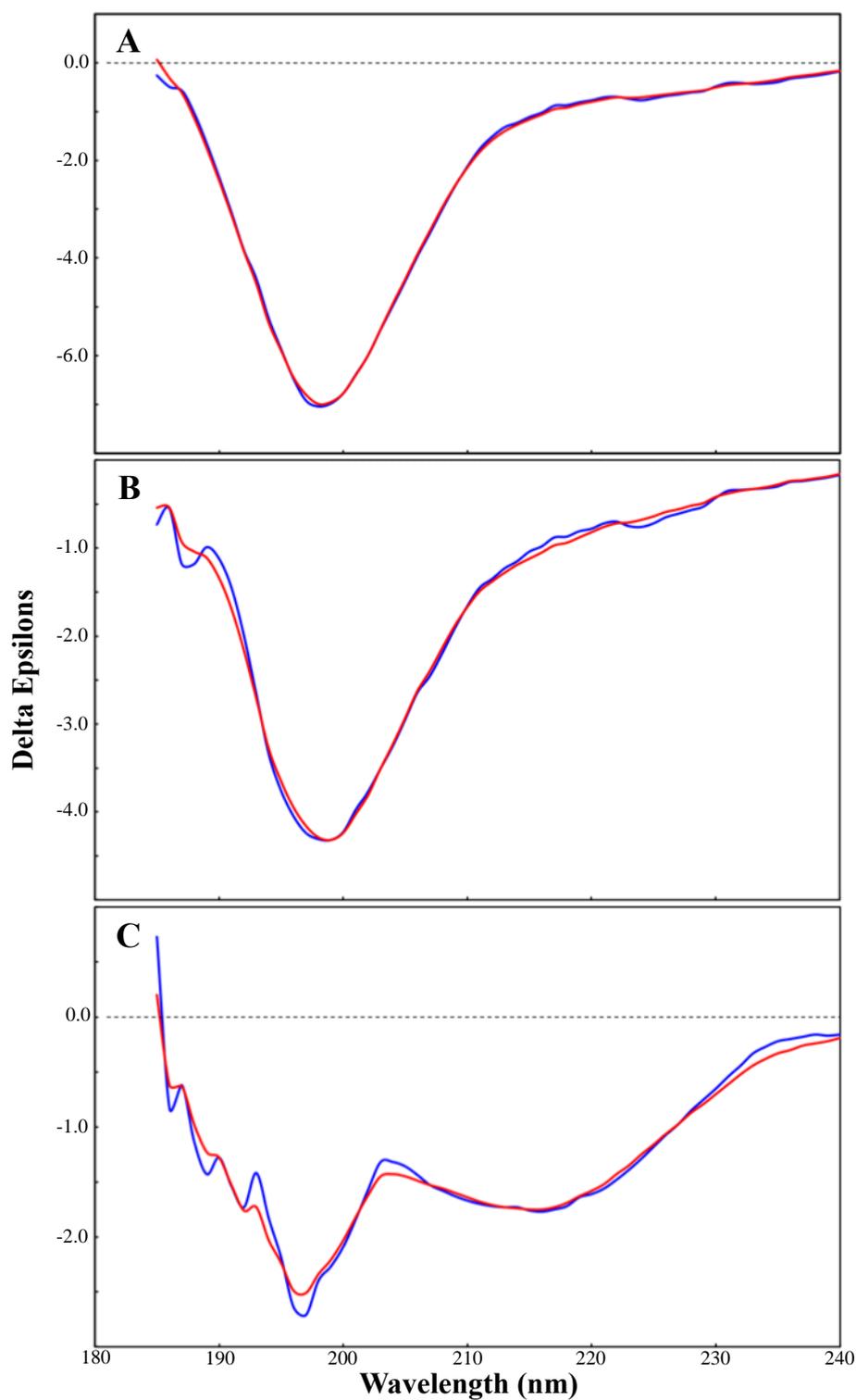


Figure 27 CD spectra of FnBPA A domain constructs. CD spectra of N1 (A), N1N2 (B) and fA (C) acquired at pH 7.4 and 298 K. Data was processed using DichroWeb algorithms (Whitmore and Wallace, 2004). Blue and red lines represent the experimental and reconstructed data, respectively.

4.5 Resonance Assignment of N1

The resonance assignment of proteins enables a detailed and quantitative chemical shift analysis to be performed. Such analyses are powerful tools in accurately describing the behaviour of IDPs. Transient secondary structure elements (Wishart and Sykes, 1994, Bozzi et al., 2003) and residues with propensities for secondary structure (Marsh et al., 2006) can be identified within an IDP by chemical shift analysis. This can be important in identifying regions likely to be involved in ligand binding activity (Norris et al., 2011, Zhang et al., 2008, Marsh et al., 2006). Completing the resonance assignment of a large unstructured protein is a far from trivial undertaking due to the inherent signal degeneracy associated with IDPs and apparent in N1 (Figure 23). Notwithstanding the recent advances in methodology and the role IDPs play in many of biological processes, there are relatively few examples of resonance assignments of proteins similar in size to N1 (Szalaine Agoston et al., 2011, Libich and Harauz, 2008, Csizmok et al., 2008, Mayer et al., 2012). However, the completed resonance assignments provided valuable insight into the protein of interest's function that were otherwise difficult to discern. To further probe the structural nuances of N1, the resonance assignment of the full subdomain (residues 37 -194, Figure 18) was carried out. The results described in Section 4.4.2 suggest the conformation observed for the recombinant fragment is representative of the conformation N1 adopts in the full A domain.

Uniformly ^{13}C , ^{15}N labelled N1 was expressed and purified as described (section 2.3.2) and mass spectrometry (MS) analysis revealed isotope incorporation of 96% (section 3.5.2). As discussed previously, there is significant overlap in the centre of the ^1H , ^{15}N HSQC spectrum of N1 and it is difficult to conclusively identify individual peaks (Figure 28). For disordered proteins it is not uncommon for this overlap to be observed in the ^{13}C dimensions used for sequential assignment (Dyson and Wright, 2001). However, ambiguities arising from signal overlap in a particular spectrum can often be solved using a dataset involving different nuclei (i.e. ambiguity in an HNC0 spectrum can be solved using a HNCA). The triple-resonance spectra recorded for N1 all had considerable signal overlap in the ^{13}C dimension (Figure 28). Despite the use of multiple experiments, a high degree of uncertainty remained in sequentially assigning resonances and an alternative

approach was required. Such difficulties are not uncommon in assigning the NMR spectra of disordered proteins. Strategies employed to enable the resonance assignment of IDPs include substituting non-labile ^1H nuclei with ^2H nuclei to improve spectral resolution (Gardner and Kay, 1998) or increasing the dimensionality of datasets used for sequential assignment (Zhang et al., 2008, Motackova et al., 2010). Despite the low sequence complexity of N1 there are few examples of sequence repetition (Section 4.3) and it was thought that simply improving the spectral resolution could solve signal degeneracy.

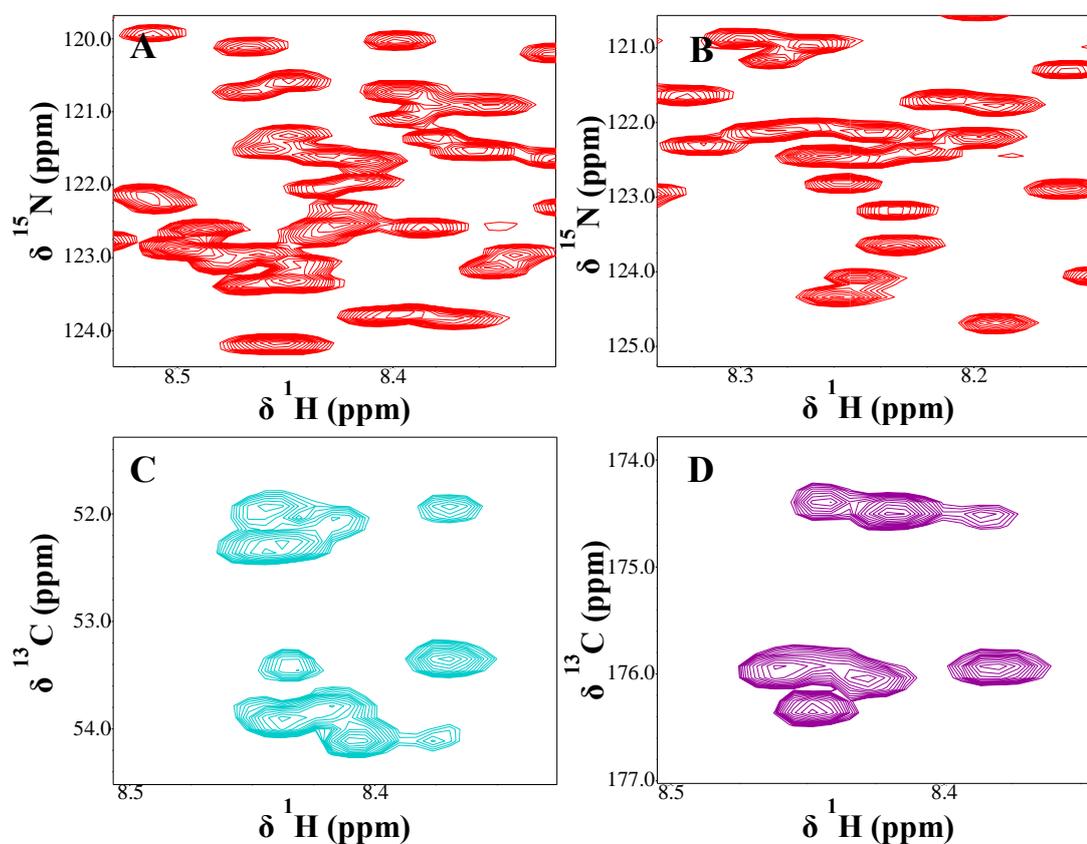


Figure 28 Resonance overlap in NMR spectra of ^{13}C , ^{15}N N1. Regions of the HSQC ((A) and (B)), HNCA (C) and HNCO (D) highlighting overlapping resonances in the spectra of N1. All spectra were recorded at pH 5.5 and 298 K on a 700 MHz spectrometer.

4.5.1 Deuterated N1 Spectra

As described previously (Section 4.4.2), peak linewidths in NMR spectra are inversely proportional to T_2 . Therefore, increased T_2 results in decreased linewidths and improved spectral resolution (Levitt, 2001). By deuterating proteins all non-

exchangeable hydrogen atoms are substituted with deuterium. Proton-proton dipolar relaxation mechanisms are removed, or significantly reduced, when typical deuteration levels of between 50-80% are achieved and the T_2 relaxation time increases (Agback et al., 1994, Markus et al., 1994). In non-deuterated aqueous solvents, amide groups that have exchangeable hydrogens are protonated and therefore observable in a ^1H , ^{15}N HSQC experiment. Deuteration also improves 3D triple-resonance spectra commonly used for sequential assignment and this is attributed to two factors. Firstly, the typical experiments implemented for resonance assignment record the amide proton chemical shift during acquisition. Therefore, narrower linewidths in the ^1H dimension translates to improved resolution in the indirect dimensions. Secondly, by substituting ^2H for ^1H in aliphatic groups ^{13}C - ^1H dipolar interactions are suppressed and the ^{13}C T_2 relaxation time increases (Venters et al., 1996). Thus, experiments involving such ^{13}C nuclei (C^α and C^β in the HNCA/HN(CO)CA and HNCACB/HN(CO)CACB experiments, for example) have narrower ^{13}C linewidths and improved resolution. ^2H , ^{13}C , ^{15}N labelled N1 was expressed as described previously (section 3.5.1) and MS analysis revealed a ^2H incorporation of approximately 70%.

The incorporation of non-labile ^2H nuclei is apparent by the reduced signal intensity in the aliphatic region of the ^1H spectrum compared to the non-deuterated material (Figure 29). As expected the ^1H , ^{15}N HSQC of ^2H , ^{13}C , ^{15}N N1 had significantly reduced linewidths and spectral resolution improved accordingly (Figure 30), enabling nearly all of the expected peaks to be identified. Despite the improved resolution in the triple resonance experiments (Figure 31), there were still congested regions of the spectra and not all assignment ambiguities were resolved.

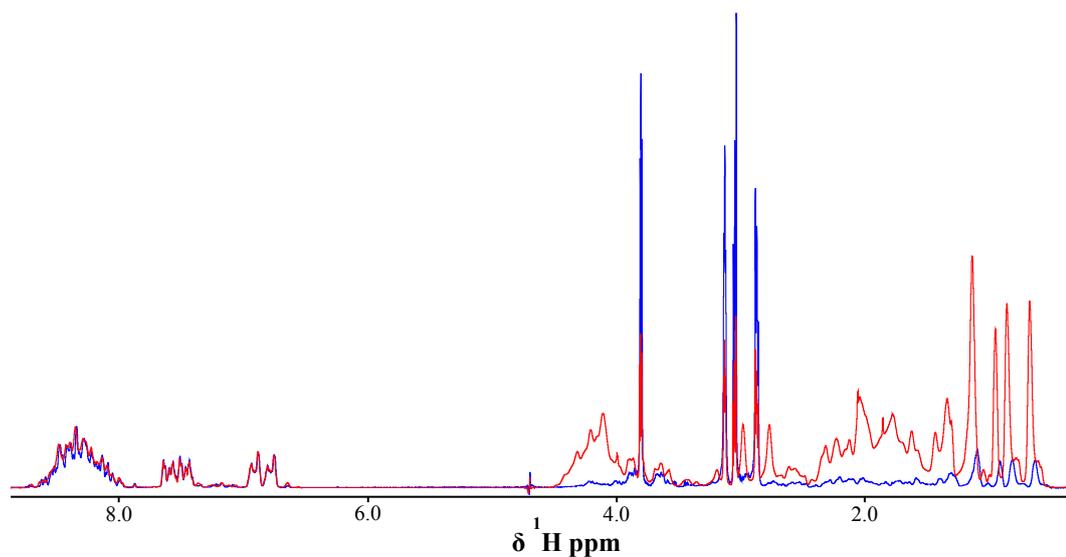


Figure 29 1D ^1H spectra of ^{13}C , ^{15}N N1 and ^2H , ^{13}C , ^{15}N N1. Overlaying the 1D ^1H spectra of deuterated (blue) and non-deuterated (red) N1 shows the extent of ^2H incorporation through the reduced signal intensity in the aliphatic region of the spectrum.

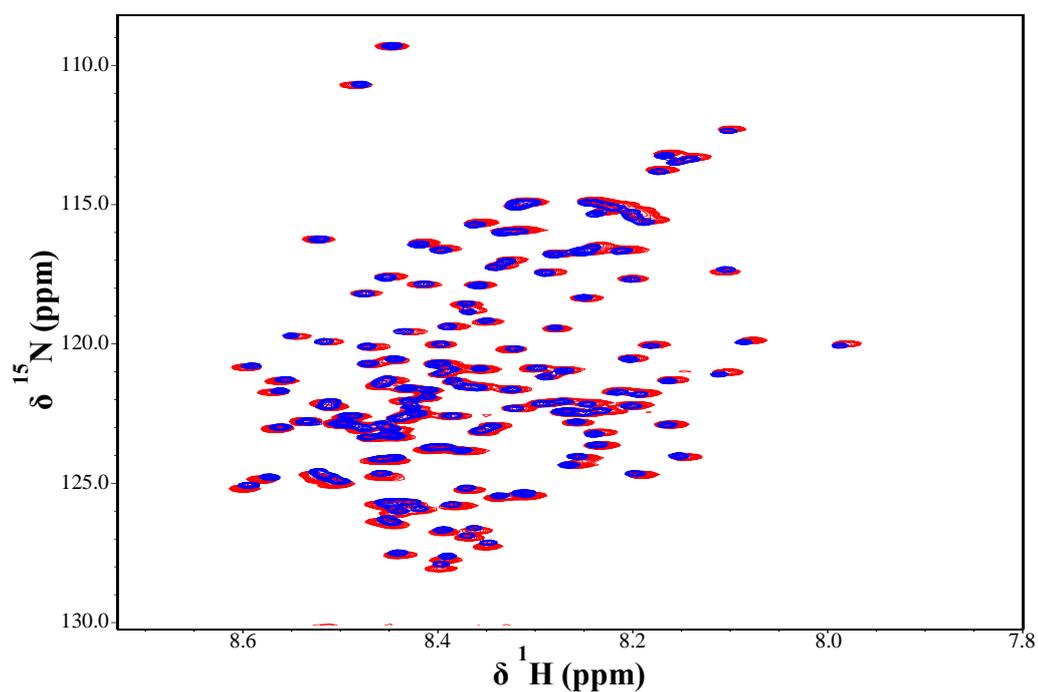


Figure 30 ^1H , ^{15}N HSQC spectra of ^{13}C , ^{15}N N1 and ^2H , ^{13}C , ^{15}N N1. Resolution in the ^1H , ^{15}N HSQC was improved through deuteration as can be seen by overlaying the spectra of ^{13}C , ^{15}N N1 (red) and ^2H , ^{13}C , ^{15}N N1 (blue).

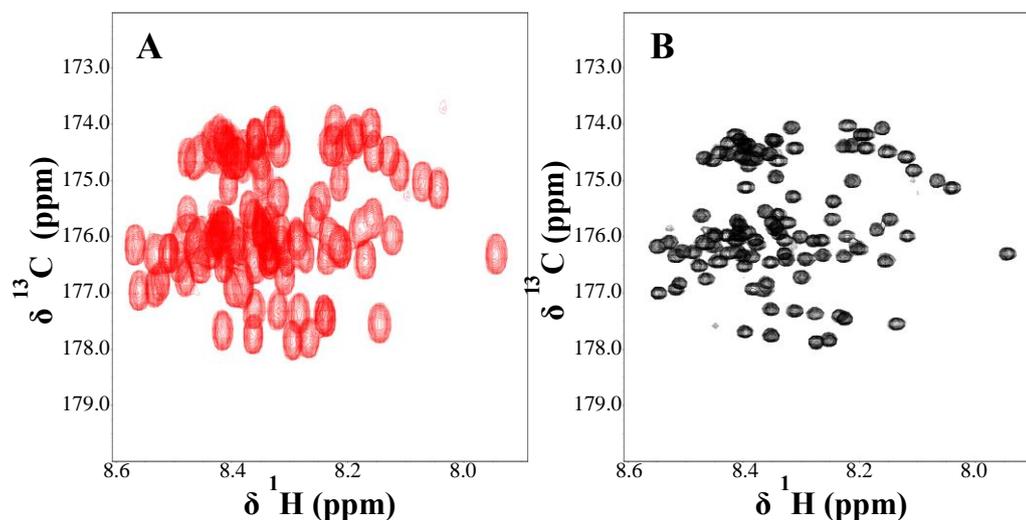


Figure 31 Improvement in the HNCOSY spectrum following deuteration. The first 32 of 64 ^{15}N planes are superimposed in both spectra. Increased resolution is achieved in deuterated (**B**) compared to non-deuterated (**A**) N1. Spectra were recorded at pH 5.5 and 298 K on a 700 MHz spectrometer.

4.5.2 Selective Unlabelling of N1

Despite the resolution improvement achieved through deuteration, the size and disorder of N1 meant that spectra were still congested and the ambiguity of sequential assignments was not completely resolved. Recent approaches employed to study large proteins by NMR involve selectively protonating side-chain methyl groups in perdeuterated proteins (Goto and Kay, 2000). Selective labelling is achieved using metabolic precursors of particular amino acids added 1 hr prior to inducing protein expression. Similar methods recently developed by Rasia *et al.* can be used to selectively ‘unlabel’ particular amino acids (Rasia *et al.*, 2012). Supplementing the growth media with metabolic precursors results in their incorporation, provided the precursor is from a suitable position in the biosynthetic pathway. The addition of non-isotopically enriched precursors to enriched expression medium results in a labelled protein with specific amino acids unlabelled (Rasia *et al.*, 2012). Leucine, valine, isoleucine, phenylalanine and tyrosine can be unlabelled in this way. Proline can also be unlabelled by the addition of L-proline prior to induction. The unlabelled residues are isotopically enriched in certain positions with ^{13}C and ^{15}N nuclei sequestered from the expression media (Table 13). However, the magnetisation pathways necessary for particular experiments are abolished rendering the amino acid ‘NMR invisible’. The precursors suitable for this strategy and the

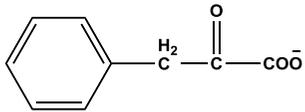
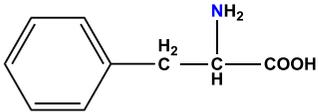
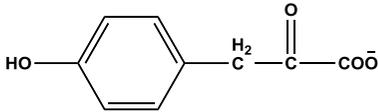
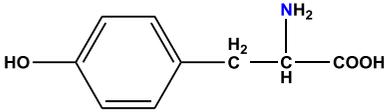
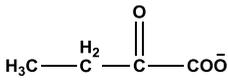
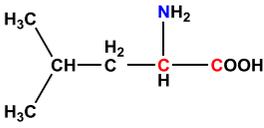
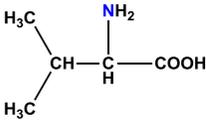
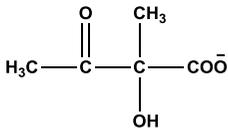
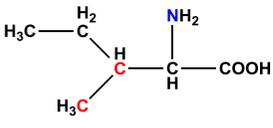
subsequently unlabelled amino acids are given in Table 13. The selective labelling strategy was originally developed to study large folded proteins otherwise outside the scope of NMR spectroscopy (Goto et al., 1999). Therefore, the precursors used to unlabel specific residues (Rasia et al., 2012) are not optimised for use in IDPs. In fact, the residues suitable for unlabelling, leucine, valine, isoleucine, tyrosine and phenylalanine, would form the typically hydrophobic core of a folded protein and are correspondingly sparse in IDPs (Romero et al., 2001). Preferentially residues such as glutamic acid and serine would be unlabelled to aid the resonant assignment of IDPs. However, the metabolic pathways of these residues are unsuitable for manipulation with precursors. Despite this drawback, N1 does contain 28 residues that can be selectively unlabelled and this strategy was pursued.

Of the residues able to be unlabelled, N1 contains leucine, valine, isoleucine, proline and tyrosine. As leucine and valine are unlabelled using the same precursor, four selectively unlabelled samples of ^2H , ^{13}C , ^{15}N N1 were expressed as described previously (section 3.5.2). The unlabelled species will be referred to as $^{\text{U}}\text{Val/Leu}$, $^{\text{U}}\text{Ile}$, $^{\text{U}}\text{Tyr}$ and $^{\text{U}}\text{Pro}$ in subsequent discussions. All proteins except $^{\text{U}}\text{Tyr}$ were expressed and purified in the same way as ^2H , ^{13}C , ^{15}N labelled N1. However, after induction the cells expressing $^{\text{U}}\text{Tyr}$ died and no protein was expressed. It is uncertain why expression of $^{\text{U}}\text{Tyr}$ resulted in cell death as this behaviour has not been previously reported, and 4-hydroxyphenylpyruvic acid has no known toxic effects when used as a growth medium supplement (Rasia et al., 2012). N1 contains only one tyrosine residue and the effect of unlabelling one residue was thought to have a minimal impact on easing spectral congestion. As such, the matter was not investigated further and $^{\text{U}}\text{Tyr}$ was not used to complete the resonance assignment of N1.

As can be seen in Table 13, backbone amide groups are sequestered from the expression media. Therefore, N1 was enriched with ^{15}N nuclei and the appearance of ^1H , ^{15}N HSQC spectrum was unaltered. In order to identify unlabelled residues HNCOC spectra were acquired. Carbon nuclei are less susceptible to isotopic scrambling (Tugarinov and Kay, 2004, Atreya and Chary, 2001), and the C' nuclei in unlabelled samples are preferentially incorporated from the precursors. Consequently, the magnetisation transfer pathway required for HNCOC spectra is no

longer allowed in unlabelled residues. Peaks corresponding to residues preceded by an unlabelled amino acid therefore disappear from the 2D ^1H , ^{15}N plane of HNCO datasets. Although leucine is unlabelled, the C' nucleus is sequestered from the expression media not the precursor, and residues preceded by leucine are still observed in the HNCO. Unlabelled leucine residues can be identified via spectra utilising a magnetisation pathway involving C^β nuclei, which are incorporated from the precursors. The 2D ^1H , ^{15}N plane of HNCO spectra acquired for each unlabelled N1 sample showed that the unlabelling was successful with the appropriate number of peaks missing from each spectrum; two from the $^{\text{U}}\text{Ile}$ HNCO (Figure 32), sixteen from the $^{\text{U}}\text{Val}$ HNCO (Figure 33) and eight from the $^{\text{U}}\text{Pro}$ HNCO (Figure 34). In some cases the ‘deleted’ peaks were in regions with signal overlap and the removal of one overlapping peak enabled the identification of the other. In this way the remaining resonances in the ^1H , ^{15}N HSQC that were previously unresolved were identified.

Table 13 Metabolic precursors for specifically unlabelling amino acids. The metabolic precursor used as a growth medium supplement and the subsequently unlabelled residue. Leucine and valine are produced via the same biosynthetic pathway and therefore are unlabelled by the same precursor. Unlabelled residues undergo some isotope enrichment with nuclei sequestered from expression media. Isotopically enriched carbon and nitrogen nuclei are indicated in red and blue, respectively.

Precursor	Structure	Residue(s) Unlabelled	Isotope Incorporation
Sodium phenylpyruvate		Phenylalanine	
4-hydroxyphenylpyruvic acid		Tyrosine	
Sodium 3-methyl-oxobutyrate		Leucine	
		Valine	
Sodium 2-oxobutyrate		Isoleucine	

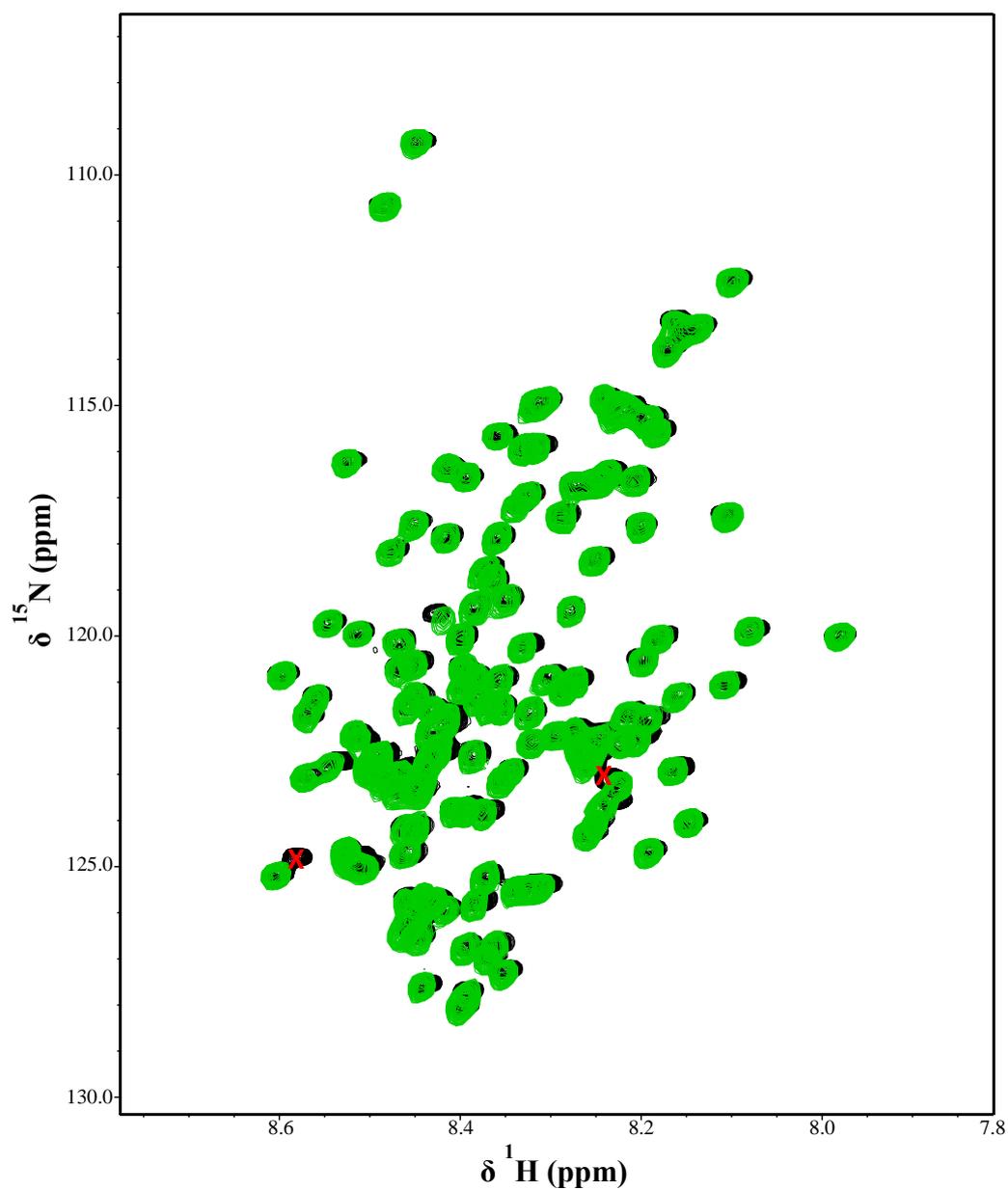


Figure 32 2D ^1H , ^{15}N HNCO plane of ^{13}C -Ile. Superimposing the ^{15}N -Ile spectrum (green) onto the fully labeled N1 spectrum (black) enables identification of the two residues preceded by Ile, highlighted with red crosses.

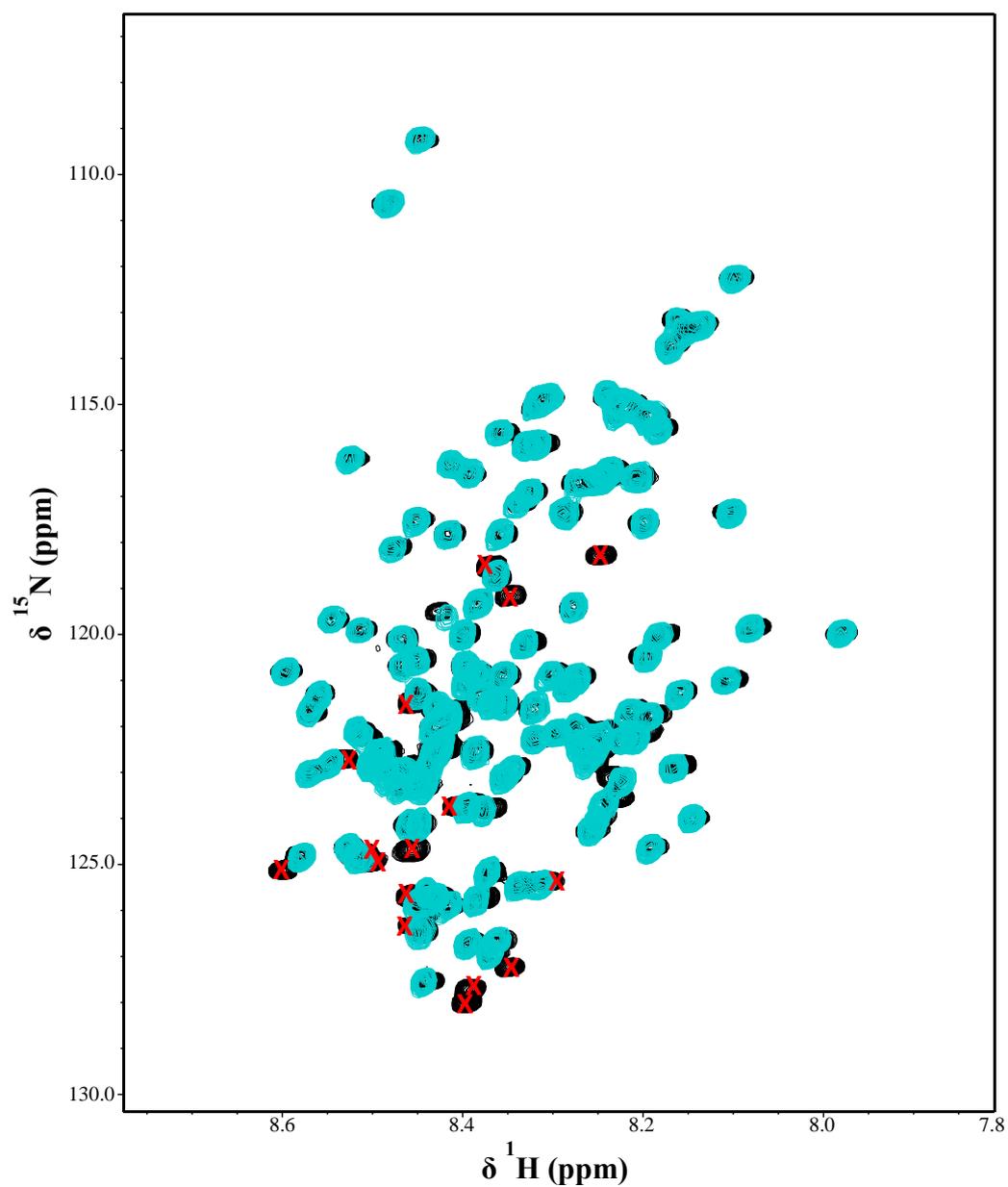


Figure 33 2D ^1H , ^{15}N HNCO plane of $^{15}\text{NVal/Leu}$. By superimposing the $^{15}\text{NVal/Leu}$ spectrum (cyan) onto the fully labelled N1 spectrum (black) the sixteen residues preceded by Val were identified and are highlighted with red crosses.

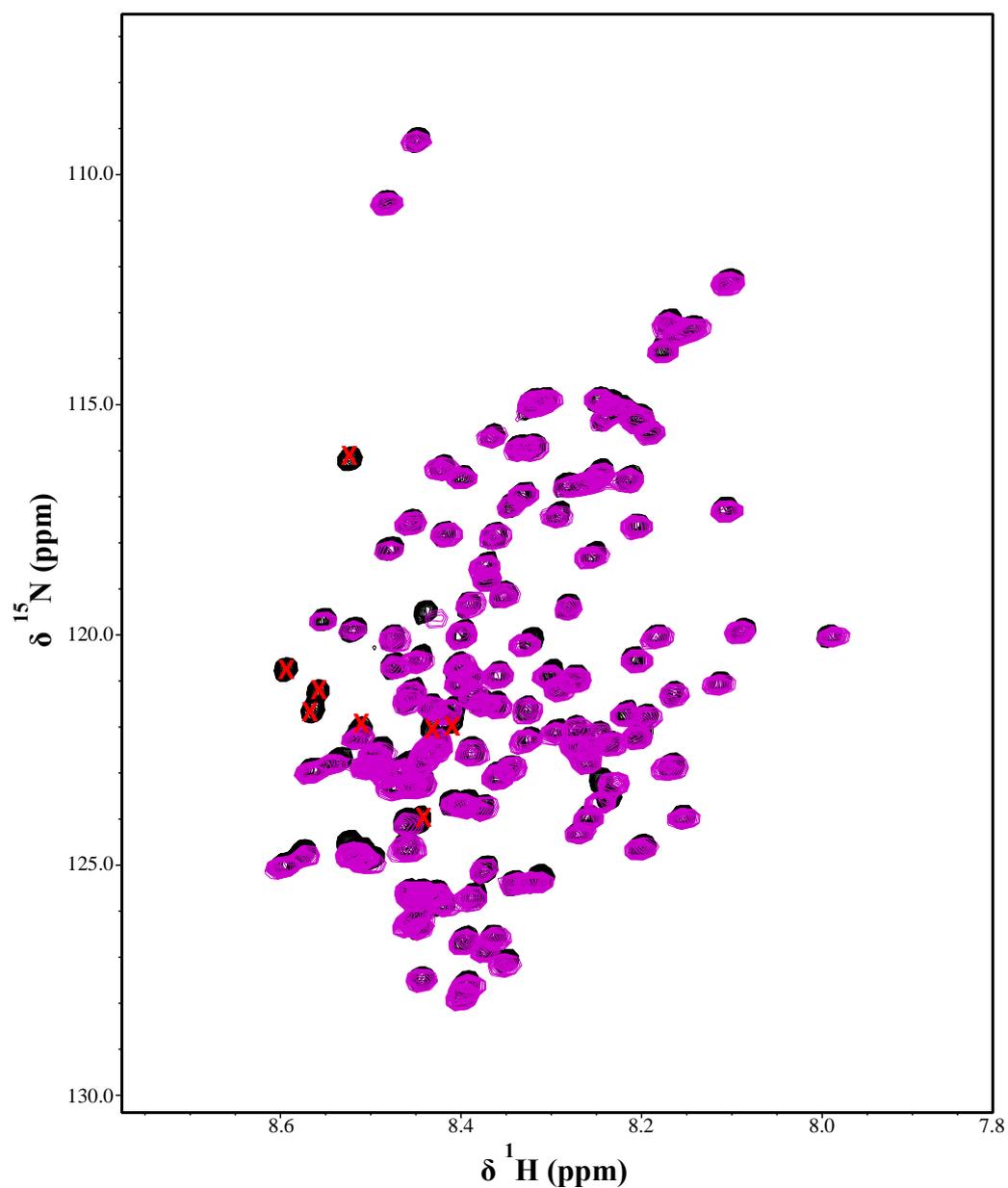


Figure 34 2D ^1H , ^{15}N HNCO plane of ^{15}N Pro. The eight amino preceded by proline residues in the sequence were identified by superimposing the ^{15}N Pro spectrum (purple) and fully labelled N1 spectrum (black) and are highlighted with red crosses.

4.5.3 Resonance Assignment Procedure

To complete the resonance assignment of N1, HNCO, HN(CA)CO, HNCA, HN(CO)CA, intra-HNCA, HNCACB and HN(CO)CACB triple resonance datasets were acquired. As was discussed previously (section 4.5), the size and disorder of N1 meant that there was significant signal overlap in all acquired spectra. Following the resolution improvements achieved via deuteration, the ambiguity of sequential assignments observed in certain spectra could be resolved using other datasets. Any further uncertainty in sequentially assigning residues was solved using the HN(CA)NNH experiment, which correlates the backbone amide group of a particular residue with those of its neighbours.

By knowing the position of peaks for residues preceded by the unlabelled amino acids, all unlabelled residue types can be also identified. For example, all peaks arising from residues preceded by valine were identifiable by their absence in the U Val/Leu spectrum. Using standard sequential assignment experiments correlating residue i to residue $i-1$ all valine residues were subsequently identified. The same method was applied to identify all isoleucine residues. Proline does not produce a peak in ^1H , ^{15}N -HSQC spectra so this process was redundant.

The identified residue types provide reference points or ‘anchors’ to the sequence, and the intervals between ‘anchors’ enable unambiguous resonance assignment. In addition to using the unlabelled residues as reference points, glycine, threonine and serine residues were also identifiable by their characteristic chemical shifts. Glycine residues typically have lower $^{15}\text{N}^{\text{H}}$ and $^{13}\text{C}^{\alpha}$ chemical shifts than other residues. In addition they have no C^{β} and produce fewer signals in the HNCACB spectrum, and on this basis were easily identifiable. The $^{13}\text{C}^{\beta}$ chemical shifts of threonine residues are typically greater than their $^{13}\text{C}^{\alpha}$ shifts, the opposite being true for other amino acids. Therefore, the threonine residues in N1 were apparent with a $^{13}\text{C}^{\beta}$ chemical shift of approximately 69 ppm. A ^1H , ^{15}N HSQC-TOCSY spectrum was acquired, and although its use was limited in many instances due to signal overlap, serine residues were identified by their distinctive H^{β} shifts at approximately 3.6 to 3.8 ppm. The ability to link residue i to residue $i-1$ via through-bond connectivity’s underpins the triple resonance assignment procedure. As was described previously

(section 1.5.3), collecting pairs of spectra (i.e. HNCO and HN(CA)CO) allows unambiguous assignment of a particular nucleus to a particular residue and enables identification of the residues neighbour. An example of this procedure as applied to N1 is given in Figure 35.

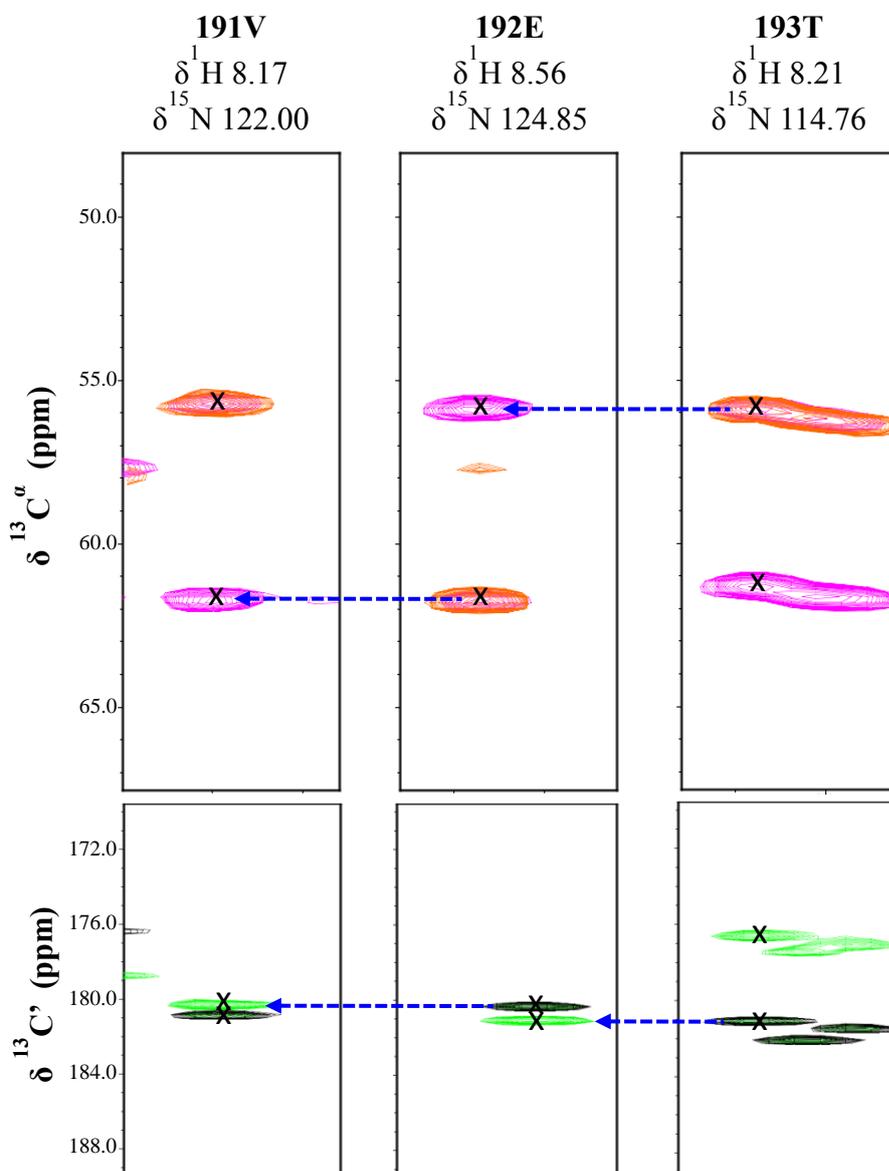


Figure 35 Sequential assignment of N1. Strips from the HNCA and HN(CO)CA spectra at particular ^1H and ^{15}N frequencies as annotated above. The HNCA peaks are magenta and the HN(CO)CA peaks are orange. Strips from the HNCO and HN(CA)CO spectra with peaks coloured black and green, respectively. Dotted lines indicate the connectivity used to complete the sequential assignment.

No residue exhibited anomalous or uncharacteristic shifts except the C-terminal glycine, which had a higher than expected $^{15}\text{N}^{\text{H}}$ shift. Of the 158 residues, 149 were assigned in the ^1H , ^{15}N -HSQC, with only Glu-39 of the observable peaks unassigned, the others being proline residues. The two foreign N-terminal residues derived from the expression vector were also unassigned in the ^1H , ^{15}N HSQC. Near complete assignment of $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$ and $^{13}\text{C}'$ nuclei was also achieved. The chemical shift values of all assigned nuclei are given in Table 14.

Table 14 Chemical shift assignments of N1. Reported values were obtained from the ^1H , ^{15}N HSQC, HNCO, HNCAi and HNCACB datasets acquired at pH 5.5 and 298 K. All spectra were referenced to an internal DSS signal.

Residue	$^1\text{H}^{\text{N}}$	$^{15}\text{N}^{\text{H}}$	$^{13}\text{C}'$	$^{13}\text{C}^{\alpha}$	$^{13}\text{C}^{\beta}$
37 Ala	-	124.43	177.75	52.07	17.88
38 Ser	8.37	119.84	176.05	54.84	-
39 Glu	-	-	176.31	56.15	-
40 Gln	8.34	121.34	175.79	55.30	28.29
41 Lys	8.42	123.05	176.72	55.78	31.86
42 Thr	8.29	115.80	174.61	61.39	69.09
43 Thr	8.25	116.57	174.50	61.31	69.16
44 Thr	8.26	117.25	174.39	61.43	69.18
45 Val	8.23	122.58	176.08	61.91	31.57
46 Glu	8.48	124.57	-	58.09	-
47 Glu	8.40	122.36	176.26	56.14	-
48 Asn	8.52	119.52	175.77	53.02	38.20
49 Gly	8.42	109.11	174.06	45.20	-
50 Asn	8.34	118.64	175.38	52.88	38.20
51 Ser	8.37	116.42	174.47	58.21	63.30
52 Ala	8.39	125.70	178.11	52.43	18.06
53 Thr	8.07	112.17	174.38	61.47	68.99
54 Asp	8.24	122.22	175.83	53.94	40.18
55 Asn	8.36	119.19	175.30	53.03	37.92
56 Lys	8.33	121.38	177.04	56.16	31.54
57 Thr	8.20	114.85	174.72	61.65	69.09
58 Ser	8.33	117.70	174.66	58.17	63.31
59 Glu	8.47	122.67	176.73	56.29	28.98
60 Thr	8.19	114.94	174.56	61.73	69.04
61 Gln	8.42	122.86	176.13	55.45	28.53
62 Thr	8.30	115.80	174.68	61.55	69.06
63 Thr	8.18	116.46	174.28	61.47	69.12
64 Ala	8.36	126.48	177.81	52.20	18.12
65 Thr	8.13	113.07	174.27	61.49	69.03
66 Asn	8.42	121.11	175.14	52.85	38.01
67 Val	8.06	119.76	175.72	61.94	31.34

68	Asn	8.43	121.26	174.74	52.81	38.02
69	His	8.41	119.36	174.19	55.13	28.04
70	Ile	8.21	123.02	176.10	60.88	37.40
71	Glu	8.54	124.58	176.44	56.05	28.84
72	Glu	8.46	122.38	176.72	56.23	28.93
73	Thr	8.21	115.12	174.60	61.76	69.08
74	Gln	8.41	122.47	175.86	55.43	28.39
75	Ser	8.30	116.85	174.18	57.93	63.31
76	Tyr	8.20	122.20	175.38	57.70	37.76
77	Asn	8.26	120.98	174.48	52.59	38.18
78	Ala	8.17	124.46	177.79	52.28	18.11
79	Thr	8.14	113.63	174.58	61.79	69.03
80	Val	8.13	122.70	176.19	61.85	31.56
81	Thr	8.22	118.11	174.29	61.44	69.10
82	Glu	8.36	123.49	175.95	55.85	29.11
83	Gln	8.47	122.67	174.00	53.10	-
84	Pro	-	-	177.07	-	-
85	Ser	8.49	116.04	174.54	58.05	63.29
86	Asn	8.44	120.52	175.01	52.83	38.05
87	Ala	8.23	123.84	177.89	52.40	18.07
88	Thr	8.11	113.19	174.43	61.65	68.96
89	Gln	8.33	122.90	175.69	55.26	28.38
90	Val	8.26	121.95	176.37	61.85	31.58
91	Thr	8.34	118.37	174.54	61.27	69.20
92	Thr	8.22	116.49	174.37	61.47	69.11
93	Glu	8.41	123.08	176.15	55.91	29.01
94	Glu	8.39	122.29	175.72	55.69	29.10
95	Ala	8.34	126.69	175.39	50.12	16.98
96	Pro	-	-	176.95		-
97	Lys	8.38	121.72	176.41	55.56	31.90
98	Ala	8.31	125.24	177.55	51.82	18.12
99	Val	8.15	119.88	176.06	61.69	31.63
100	Gln	8.43	124.45	175.12	54.90	28.51
101	Ala	8.41	127.32	175.48	50.10	16.99
102	Pro	-	-	177.01	62.68	-
103	Gln	8.56	120.61	176.20	55.38	28.43
104	Thr	8.16	115.44	174.06	61.35	69.24
105	Ala	8.33	126.41	177.28	51.87	18.15
106	Gln	8.37	120.85	173.99	53.06	27.98
107	Pro	-	-	176.52	62.67	-
108	Ala	8.41	123.86	177.44	51.99	18.16
109	Asn	8.38	117.65	175.00	52.79	37.97
110	Ile	8.08	120.89	176.11	60.71	37.55
111	Glu	8.48	124.57	176.49	56.02	29.05
112	Thr	8.21	116.31	174.43	61.61	69.07
113	Val	8.20	123.41	175.92	61.91	31.48
114	Lys	8.41	125.76	176.27	55.70	31.86
115	Glu	8.43	122.71	176.17	55.94	29.05
116	Glu	8.47	122.60	176.18	55.92	-

117	Val	8.22	122.32	176.08	61.92	31.54
118	Val	8.28	125.10	175.99	61.90	31.40
119	Lys	8.42	126.08	176.29	55.70	31.93
120	Glu	8.44	122.84	176.27	55.98	29.01
121	Glu	8.50	122.58	175.93	55.82	29.22
122	Ala	8.35	125.52	177.36	51.81	18.03
123	Lys	8.29	122.09	174.54	53.67	31.20
124	Pro	-	-	176.82	62.68	-
125	Gln	8.53	121.10	176.00	55.20	28.44
126	Val	8.24	122.23	175.96	61.75	31.61
127	Lys	8.42	125.44	176.34	55.74	31.89
128	Glu	8.53	122.80	176.61	56.11	29.06
129	Thr	8.33	115.51	174.74	61.38	69.12
130	Thr	8.22	116.50	174.50	61.49	69.08
131	Gln	8.46	122.71	175.98	55.42	28.31
132	Ser	8.42	117.41	174.61	58.10	63.31
133	Gln	8.48	122.08	175.61	55.41	28.33
134	Asp	8.35	121.13	175.97	53.98	40.23
135	Asn	8.44	119.92	175.60	52.83	38.02
136	Ser	8.39	116.22	175.32	58.97	63.28
137	Gly	8.45	110.49	174.07	45.23	-
138	Asp	8.17	120.37	176.46	54.06	40.16
139	Gln	8.37	120.52	176.08	55.46	-
140	Arg	8.29	121.48	176.35	55.89	29.38
141	Gln	8.40	121.42	175.98	55.40	28.11
142	Val	8.13	121.13	175.55	61.84	31.68
143	Asp	8.38	123.50	176.06	53.76	40.27
144	Leu	8.34	123.61	177.46	54.53	40.80
145	Thr	8.17	117.48	172.84	60.11	68.76
146	Pro	-	-	176.94	62.84	-
147	Lys	8.40	121.85	176.72	55.71	31.78
148	Lys	8.31	122.76	176.33	55.77	31.94
149	Ala	8.42	125.57	177.95	52.18	-
150	Thr	8.13	113.29	174.57	61.49	69.08
151	Gln	8.40	122.08	175.63	55.51	28.89
152	Asn	8.49	119.74	174.93	53.05	38.02
153	Gln	8.36	120.74	175.86	55.51	28.04
154	Val	8.18	121.54	175.87	61.82	31.55
155	Ala	8.36	127.45	177.68	51.96	18.09
156	Glu	8.41	120.36	176.65	56.19	29.01
157	Thr	8.17	115.26	174.33	61.49	69.18
158	Gln	8.44	123.11	175.69	55.24	28.37
159	Val	8.16	121.64	175.97	61.84	31.60
160	Glu	8.47	124.73	176.21	55.87	29.00
161	Val	8.22	122.02	175.72	61.63	31.58
162	Ala	8.37	127.71	177.32	51.87	18.14
163	Gln	8.33	120.72	174.01	53.06	27.83
164	Pro	-	-	176.95	62.74	-
165	Arg	8.53	121.49	176.65	55.72	29.54

166	Thr	8.17	115.13	174.39	61.32	69.20
167	Ala	8.42	126.18	177.93	52.35	18.06
168	Ser	8.29	114.83	174.65	58.20	63.25
169	Glu	8.35	122.38	176.39	56.09	29.11
170	Ser	8.31	117.05	174.02	57.96	63.27
171	Lys	8.23	124.13	174.30	53.79	31.34
172	Pro	-	-	176.75	62.65	-
173	Arg	8.48	121.90	176.34	55.67	29.56
174	Val	8.24	121.89	176.18	61.65	31.68
175	Thr	8.32	118.97	174.19	61.32	69.17
176	Arg	8.43	123.92	176.17	55.45	29.77
177	Ser	8.45	117.99	174.37	58.16	63.27
178	Ala	8.41	125.47	177.41	52.27	18.11
179	Asp	8.25	119.24	176.31	54.12	40.21
180	Val	7.96	119.89	176.01	61.83	31.52
181	Ala	8.32	126.95	177.88	52.27	18.00
182	Glu	8.29	119.98	176.33	56.13	28.97
183	Ala	8.28	125.14	177.80	52.16	17.86
184	Lys	8.27	120.67	176.83	56.01	31.77
185	Glu	8.38	121.47	176.41	56.15	28.82
186	Ala	8.34	124.96	177.89	52.25	18.01
187	Ser	8.28	114.77	174.54	58.16	63.20
188	Asn	8.37	120.58	174.83	52.91	38.02
189	Ala	8.12	123.79	177.45	52.18	18.09
190	Lys	8.24	120.79	176.42	55.73	31.76
191	Val	8.17	122.00	176.17	61.77	31.60
192	Glu	8.56	124.85	176.55	55.92	29.01
193	Thr	8.21	114.76	174.31	61.36	69.18
194	Gly	8.07	117.12	179.03	45.91	-

4.6 Secondary Chemical Shift Analysis

Secondary structure elements in proteins can be identified by analysing secondary chemical shifts ($\Delta\delta$) (Wishart and Sykes, 1994, Szalaine Agoston et al., 2011, Bozzi et al., 2003). $\Delta\delta$ values are calculated by subtracting experimentally measured chemical shifts from random coil values predicted for a given sequence. The chemical shifts of certain nuclei are highly sensitive to secondary structure and the magnitude of $\Delta\delta$ values indicative of the conformation an amino acid adopts. Interpretation of $\Delta\delta$ values is often simplified by assigning a chemical shift index (CSI) (Wishart and Sykes, 1994, Wishart et al., 1992). For a residue to be classified as random coil, $\Delta\delta$ values must be within a specified range (Table 15). Nuclei with $\Delta\delta$ in this range are assigned a CSI of 0. Nuclei with $\Delta\delta$ values outside this range are

assigned a CSI of ± 1 depending on the direction of the deviation. Clusters of the same CSI value are indicative of local secondary structure elements. Positive CSIs for $^1\text{H}^\alpha$, $^1\text{H}^N$, $^{15}\text{N}^H$ and $^{13}\text{C}^\beta$ nuclei and negative CSIs for $^{13}\text{C}^\alpha$ and $^{13}\text{C}'$ nuclei of four or more sequential residues suggests β -strand conformations (Wishart and Sykes, 1994, Wishart et al., 1992, Schwarzingner et al., 2001). Residues in helical conformation have the opposite CSI values for a given nucleus. Prior to calculating $\Delta\delta$ values, the random coil chemical shifts for a particular protein need to be determined. Random coil chemical shifts are dependent on the protein sequence and the temperature and pH at which they are measured (Kjaergaard et al., 2011, Kjaergaard and Poulsen, 2011, Schwarzingner et al., 2001). Therefore prior to $\Delta\delta$ values being calculated, corrections must be made to account for these factors. Referencing to an internal standard, typically DSS, is also essential when calculating $\Delta\delta$ values (Wishart et al., 1995b).

Table 15 Chemical shift index ranges. The range for each nucleus used to assign CSI values. Positive values outside of each range are assigned a +1 CSI, negative values -1 and values within the range 0.

Nucleus	$\Delta\delta$
$^1\text{H}^\alpha$	± 0.10 ppm
$^1\text{H}^N$	± 0.10 ppm
$^{15}\text{N}^H$	± 1.00 ppm
$^{13}\text{C}^\alpha$	± 0.70 ppm
$^{13}\text{C}^\beta$	± 0.70 ppm
$^{13}\text{C}'$	± 0.70 ppm

Prior to resonance assignment, all N1 spectra were referenced to DSS. A 1D ^1H spectrum of ^2H , ^{13}C , ^{15}N N1 containing 1 mM DSS was recorded and the reference signal measured at -0.091 ppm (data not shown). The ^{15}N and ^{13}C dimensions of each dataset were referenced indirectly based on the ^1H correction value. All spectra of N1 were recorded at 25 °C and pH 5.5. The sequence, temperature and pH corrected random coil shifts for N1 were predicted using an algorithm developed by Alexandr Maltsev of the National Institutes of Health based on coefficients determined by Kjaergaard *et al.* (Kjaergaard *et al.*, 2011, Kjaergaard and Poulsen, 2011). The calculated $\Delta\delta$ values of N1 are given in Figure 36. As near complete assignment of $^1\text{H}^{\text{N}}$, $^{15}\text{N}^{\text{H}}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$ and $^{13}\text{C}^{\gamma}$ nuclei in N1 was achieved, all were used in secondary chemical shift analysis. The vast majority of $\Delta\delta$ values for nuclei in all residues lie within the range indicative of random coil conformations. Therefore, analysis using the CSI method suggests N1 behaves as an IDP.

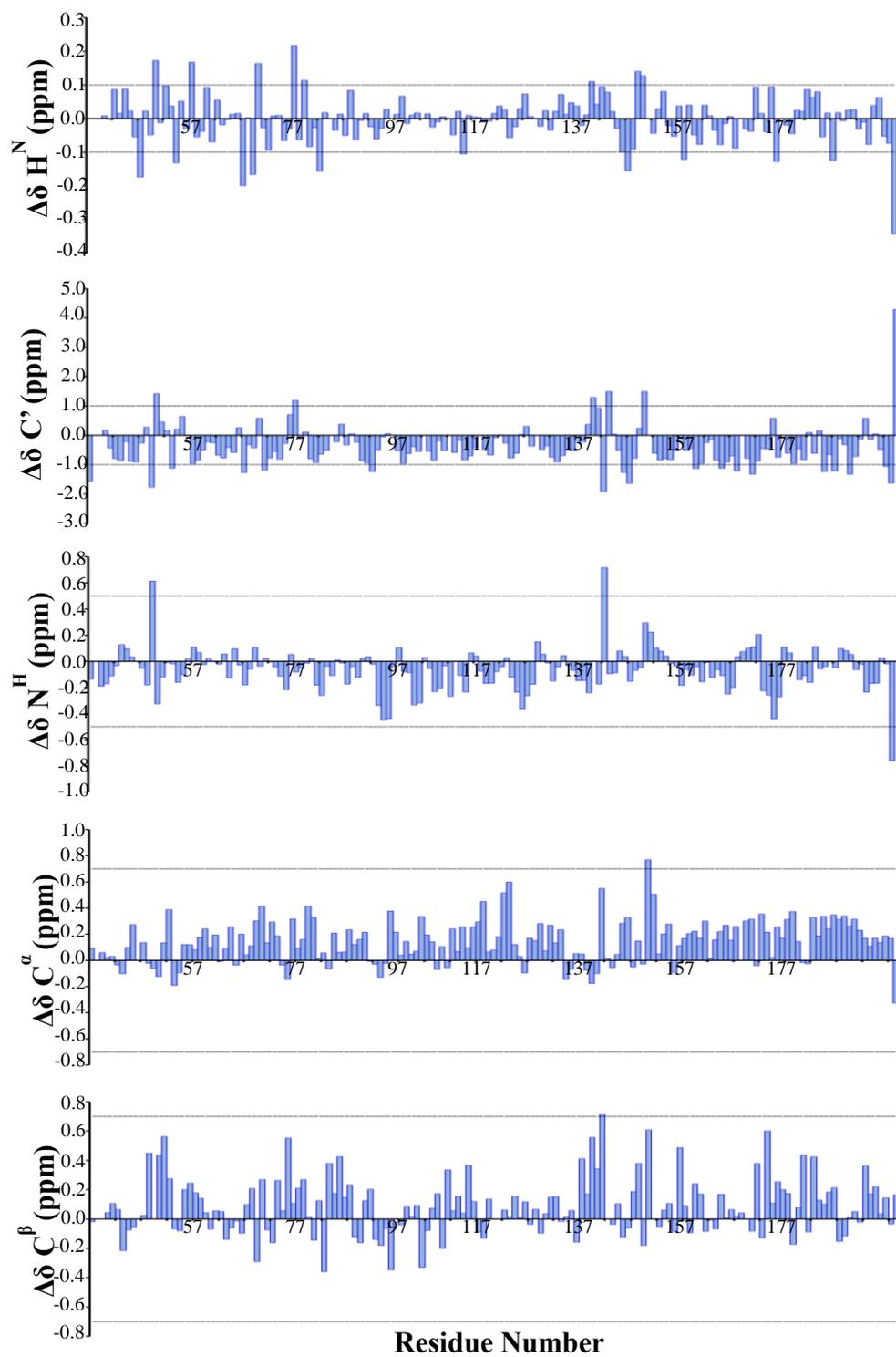


Figure 36 Sequence corrected secondary chemical shifts of N1. $\Delta\delta$ values were calculated using sequence, temperature and pH corrected random coil values. The CSI thresholds for each nucleus indicated with a dashed line.

4.6.1 Secondary Structure Propensity of N1

Despite lacking stable secondary structure it is possible that regions of N1 adopt transient conformations. $\Delta\delta$ values in such elements may not be over the CSI threshold but instead can be identified qualitatively. N1 contains no regions in which all nuclei have $\Delta\delta$ values suggestive of a particular conformation. The $\Delta\delta$ values of the C^α and N^H nuclei in most residues follow a trend suggestive of a helical conformation (Figure 36). However, the $\Delta\delta$ values of all other nuclei contradict this notion as there are no unidirectional clusters indicative of secondary structure propensity. A method proposed by Marsh *et al.* quantitatively describes the secondary structure propensity (SSP) of amino acids based on their chemical shifts (Marsh *et al.*, 2006). Each residue is assigned a SSP score representing the expected fraction of secondary structure. Values of 1 and -1 represent fully formed helical or strand structures, respectively. An advantage of this method is that the contributions of chemical shifts of each nucleus are weighted reflecting their usefulness in predicting secondary structure. SSP analysis of N1 suggests that different regions have a propensity for different conformations (Figure 37). Residues in the N-terminal region of N1 (residues 37 -90) are assigned scores that show no significant SSP. From residues 91 to 169 the SSP scores indicate a propensity for strand conformations. The C-terminal 18 residues exhibit a propensity for helical secondary structure. The propensity of these regions for a particular conformation may indicate that they are involved in interactions with ligands and that the suggested conformations are stabilised on binding.

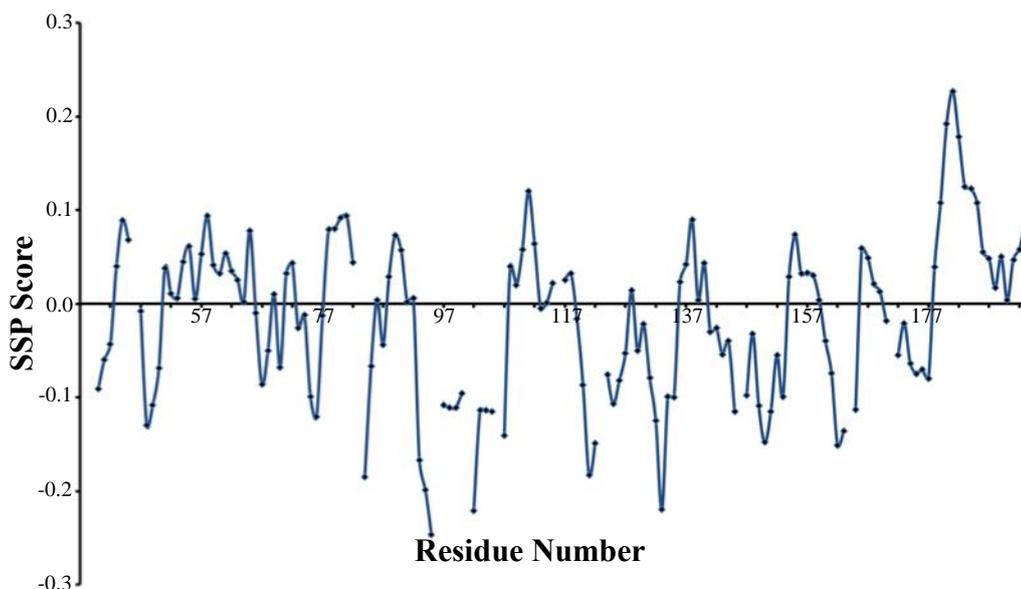


Figure 37 Secondary structure propensity scores for residues in N1. SSP scores were calculated based on all assigned resonances. Positive values indicate a propensity for helical secondary structure, negative for strand conformation.

4.7 Summary

FnBPA is a *S. aureus* surface protein and major virulence factor (Arrecubieta et al., 2006, Fitzgerald et al., 2006). Interactions involving FnBPA play crucial roles in host colonisation and the development of various life-threatening illnesses (Massey et al., 2001). FnBPA recognises a number of ECM components and the mechanisms of these known interactions are well described (Schwarz-Linek et al., 2003, Keane et al., 2007b). Correspondingly, the regions of FnBPA involved have been well-characterised structurally; Fn binding mediated by the highly conserved disordered FnBRs (Meenan et al., 2007), and Fg/elastin binding by the immunoglobulin-type folded N2N3 subdomains of the N-terminal A domain (Stemberk et al., manuscript in preparation). Conversely, the N-terminal N1 subdomain is poorly characterised structurally and has no attributed functions. Sequence conservation in N1 is much higher than in the N2 and N3 subdomains, with approximately 90% conservation compared to 75% and 60% for N2 and N3, respectively (Loughman et al., 2008). Variation within the N1 subdomain does not correlate with variation in the N2 and N3 subdomains, suggesting a divergent evolution. This hypothesis is reinforced by

the fact that N1 does not affect the Fg-binding capacity of the N2N3 subdomains (Stemberk et al., manuscript in preparation).

In silico sequence analysis of N1 (Figure 19) revealed it contains only 20 residues that would form the hydrophobic core of a folded protein. Instead, N1 is rich in disorder-promoting residues, a characteristic associated with IDPs (Romero et al., 2001). Unsurprisingly therefore, a number of tools that predict secondary structure based on the protein sequence suggest that N1 is an IDP (Figure 20). Each of the four tools used implements a different algorithm to predict secondary structure and agreement between each method is highly suggestive that N1 lacks stable secondary structure and is an IDP.

The 1D ^1H NMR spectrum of N1 contained low signal dispersion in the amide region of the spectrum and a high degree of signal degeneracy in the aliphatic region (Figure 22); features characteristic of intrinsic disorder. The acquisition of a ^1H , ^{15}N HSQC supported the hypothesis that N1 is an IDP (Figure 23). Peaks in the ^1H , ^{15}N HSQC had narrow linewidths, were poorly dispersed, and a high degree of degeneracy of side-chain resonances was observed. Again these features are typical of disordered proteins and suggest N1 is intrinsically disordered. To confirm the observed disorder observed in N1 was not due to incorrectly defined domain boundaries, particularly at the C-terminus, two dimensional NMR spectra of N1N2 and fA were acquired. Both spectra exhibit features characteristic of disordered and folded components, attributable to N1 and N2N3, respectively. A number of resonances for residues at the C-terminus of N1 are subject to minor chemical shift changes when N2 is present. Most are in close proximity to or overlap with the corresponding resonance in the spectrum N1 and have linewidths indicative of disorder. The C-terminal Gly-194 peak is no longer visible in the region of the spectrum it populates in N1. However, a distinct peak in a region of the spectrum typical for glycine residues is apparent. It is likely therefore, that the structure of recombinantly expressed N1 is representative of its behaviour in the intact A domain.

A small degree of ambiguity exists in NMR when distinguishing between helical and disordered proteins, as helical conformations can give rise to narrowly dispersed

resonances, though typically not as narrow as IDPs. CD spectra of N1, N1N2 and fA and their deconvolution (Whitmore and Wallace, 2004) revealed that N1 has no helical character (Figure 27). Taken together this provides strong evidence that N1 is an IDP.

Completing the resonance assignment of a protein enables chemical shift based structural characterisation and insight into potential functions. Therefore, the sequence specific resonance assignment of N1 was carried out. Uniformly ^{13}C , ^{15}N N1 was expressed, but substantial resonance overlap was observed in all acquired spectra. Deuterated N1 was expressed, and due to more favourable relaxation properties of deuterated proteins (Gardner and Kay, 1998), spectral resolution improved dramatically and nearly all peaks were identified in the ^1H , ^{15}N HSQC spectrum (Figure 30). Final ambiguities in the assignment procedure were solved by selectively unlabelling specific amino acids in N1 (Rasia et al., 2012). The addition of non-isotopically enriched metabolic precursors to enriched expression medium resulted in labelled N1 with leucine, valine, proline or isoleucine selectively unlabelled. HNCO spectra of the unlabelled proteins lack peaks attributed to residues preceded by the unlabelled amino acids (Figure 32, Figure 33, Figure 34). Regions with overlapping peaks therefore became less congested and individual peaks could be resolved. Further, identification of specific residue types provided reference points to the sequence and the interval between such points enabled unambiguous sequential assignment of N1. Of the 150 peaks observed in the ^1H , ^{15}N HSQC of N1, 149 were assigned.

Chemical shift analysis can give useful insight into the structure and dynamics of IDPs. Comparing experimentally measured chemical shifts to random coil values enables secondary structure elements within a protein to be identified (Wishart and Sykes, 1994, Wishart et al., 1992). Secondary chemical shift analysis of N1 suggests it behaves as a random coil and there are no regions of stable secondary structure (Figure 36). However, rather than adopting stable conformations, IDPs often form transient secondary structure elements or contain regions with a propensity for secondary structure (Dyson and Wright, 2005). Such elements can be identified in IDPs by carrying out secondary structure propensity (SSP) calculations based on chemical shifts (Marsh et al., 2006). SSP analysis of N1 revealed regions with

propensity for both strand and helical conformations (Figure 37). These were confined to the C-terminus of N1 with residues 91 – 169 and 177 - 194 exhibiting propensities for strand and helical confirmation, respectively. Identifying regions with a propensity for secondary structure can denote ligand-binding sites (Norris et al., 2011, Zhang et al., 2008), and the SSP observed in N1 may indicate ligand binding potential.

5 Interactions Between N1 and Host Ligands

5.1 Introduction

The ability of FnBPA to interact with multiple ECM components is thought to underpin its importance as a *S. aureus* virulence factor and the known interactions involving FnBPA have been described in detail (Stemberk et al., manuscript in preparation)(Bingham et al., 2008, Meenan et al., 2007, Roche et al., 2004). To date no function is attributed to the N1 subdomain. The Fg- and Fn-binding capacity of FnBPA is unaffected by N1 (Stemberk et al., manuscript in preparation). A study of the sequence diversity in the A domain of FnBPA across a number of *S. aureus* strains found that N1 is the most conserved subdomain (~90%), with N2 (~75%) and N3 (~60%) more divergent (Loughman et al., 2008). Taken together, these results suggest N1 fulfils a role detached from N2 and N3.

The work presented previously (Chapter 4) shows N1 is an IDP. A characteristic sometimes associated with IDPs is the recognition of multiple ligands (Dyson and Wright, 2002, Dunker et al., 2002). N1 is the most N-terminal subdomain of FnBPA. As the cell wall attachment site is at the C-terminus, N1 is likely to be projected away from the cell surface. The potential to recognise multiple ligands may be advantageous in host colonisation processes. As the known ligands of FnBPA are plasma components it is possible that N1 also recognises a plasma factor. The A domains of the *S. aureus* surface proteins clumping factor B (ClfB) and SasG are known to facilitate host colonisation via interactions with host endothelial cells (Corrigan et al., 2007, O'Brien et al., 2002b, Roche et al., 2003). ClfB recognises human type 1 cytokeatin 10 (O'Brien et al., 2002b), however the specific receptor recognised by SasG is not known. The functional overlap observed across *S. aureus* surface proteins suggests FnBPA's A domain could fulfil a similar role.

5.2 Aims

The aim of this work was to identify N1 ligands. FnBPA is known to bind plasma components, so pull-down experiments with various A domain constructs as 'bait' proteins in plasma were conducted. Affinity columns prepared using recombinant

FnBPA A domain fragments were used to screen for potential ligands. The role of FnBPA in host colonisation processes was investigated by screening A domain adherence to human umbilical vein endothelial cells (HUVECs).

5.3 Pull-down Experiments

As described previously (section 3.6), recombinant GST-fusion proteins of the FnBPA A domain are susceptible to C-terminal degradation. Despite this instability, a significant amount of intact material was still present and suitable for use in GST pull-down experiments. Pull-down experiments are commonly used to identify binding partners of previously uncharacterised proteins. Often, pull-downs involve using an affinity tag ‘anchor’ to isolate the ‘bait’ protein following a period of incubation with solutions containing potential ligands. Exploiting specific interactions involving anchor proteins enables efficient recovery of bait proteins, and as a result GST-fusion proteins are widely used in pull-down assays (Lee et al., 2008, Park et al., 2012). GST interacts specifically with reduced glutathione (GSH) with nanomolar affinity (Fabrini et al., 2009) and this interaction was used to purify the recombinant GST-fusion proteins (section 3.6.2). Therefore, beads or media coated with GSH will interact with GST-fusion proteins and enable simple isolation of potential complexes. Species isolated by GST pull-down assays are typically identified using tandem mass-spectrometry (MS/MS). To screen for potential ligands for N1, the A domain GST-fusion ‘bait’ proteins were anchored to magnetic beads via the GST/GSH interaction prior to incubation with blood plasma (Figure 38). The presence and identity of pulled-down ligands were determined by SDS-PAGE and mass spectrometry analysis, respectively.

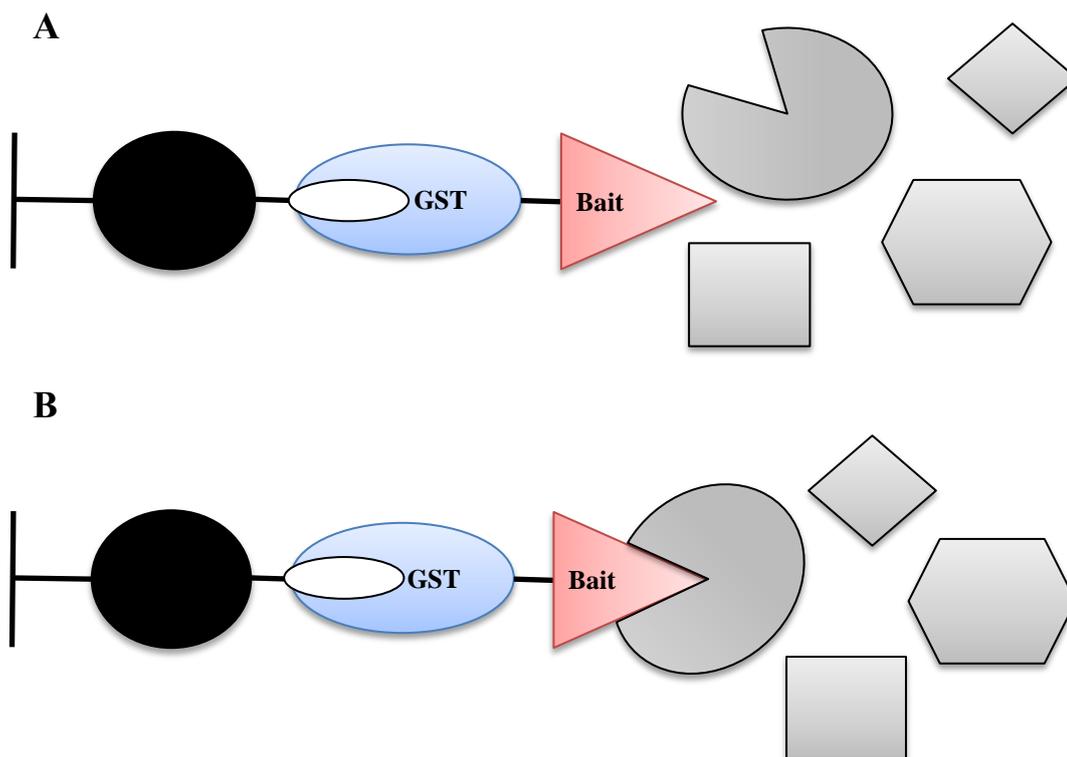


Figure 38 Schematic representation of a GST pull-down assay. The GST-fusion bait protein is anchored to the magnetic beads via GST binding to GSH (indicated by white ellipse). The beads are then mixed with a solution containing possible binding partners (**A**). Following a period of incubation complexes involving the bait protein are pulled-down using the magnetic beads (**B**).

5.3.1 FnBPA A Domain Appears to be Proteolytically Cleaved in Plasma

The GST pull-down experiments were conducted as described in section 2.4.4. SDS-PAGE analysis of proteins eluted from the beads following incubation with plasma revealed a large number of non-specific interactions with only trace amounts of intact GST-N1N2 recovered (Figure 39). GST-N1 and GST-fA were not present in the eluted samples. The inability to recover the bait proteins could be attributed to three factors. Firstly, GSH present in plasma could cause the GST-fusions to elute from the magnetic beads. However, the concentration of GSH in plasma is approximately 1.5 μM (Jones et al., 1998), and this was insufficient to elute the GST-fusions in control experiments (data not shown). Secondly, the non-specifically bound species may out-compete the GST-fusion proteins for binding to the magnetic beads. The third, and most likely, explanation is that the GST-fusion proteins are subject to proteolysis in the plasma. The recovery of free GST suggests the magnetic

beads retain their ligand binding capacity. Therefore, proteolytic cleavage within the N1 subdomain, separating GST and the A domain fragment, would prevent recovery of the intact bait protein. Consequently, any other interactions involving regions of FnBPA upstream of the cleavage site, such as the N2N3 interaction with Fg, are unlikely to be observed.

Two species pulled-down by GST-N1N2 were initially attributed to specific interactions with plasma components. These bands were cut out of the gel and analysed by in-house mass spectrometry services. Briefly, the samples were destained and digested with trypsin, with the resulting peptides analysed by matrix-assisted laser desorption/ionisation time-of-flight tandem mass spectrometry (MALDI-TOF MS/MS) and identified through Mascot database searches. The isolated proteins were identified as serum albumin, the most abundant human plasma protein (Koch-Weser and Sellers, 1976b, Koch-Weser and Sellers, 1976a, Owen et al., 1993), and GST. However, as neither species was pulled-down in subsequent assays the interactions are thought to be non-specific.

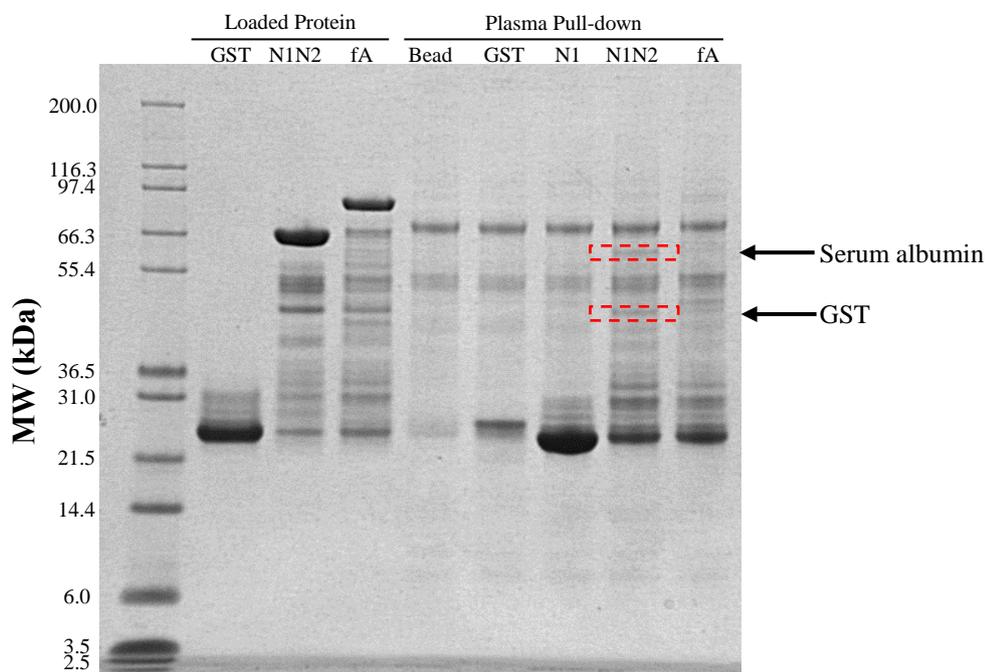


Figure 39 SDS-PAGE analysis of FnBPA A domain GST pull-down assay. The A domain GST-fusion proteins are susceptible to a significant amount of degradation. A high number of non-specific interactions are observed in each pull-down. Following incubation with plasma, no intact bait protein was recovered. Serum albumin and GST were pulled down by GST-N1N2.

To ensure that the inability to recover intact bait protein was not due to unsuitable elution conditions, protein coated beads were incubated with PBS rather than plasma. The bait proteins were then eluted under the same conditions used for the pull-down experiments (Figure 40). Although increased degradation was observed during the incubation period, near complete recovery of the loaded protein was achieved. Therefore, the elution conditions used in this assay are suitable. In an attempt to reduce the amount of non-specific interactions, a second pull-down was carried out with detergent added to the plasma during the incubation with the bait proteins (Figure 40). However, this had no observable effect with the high level of non-specific binding still apparent.

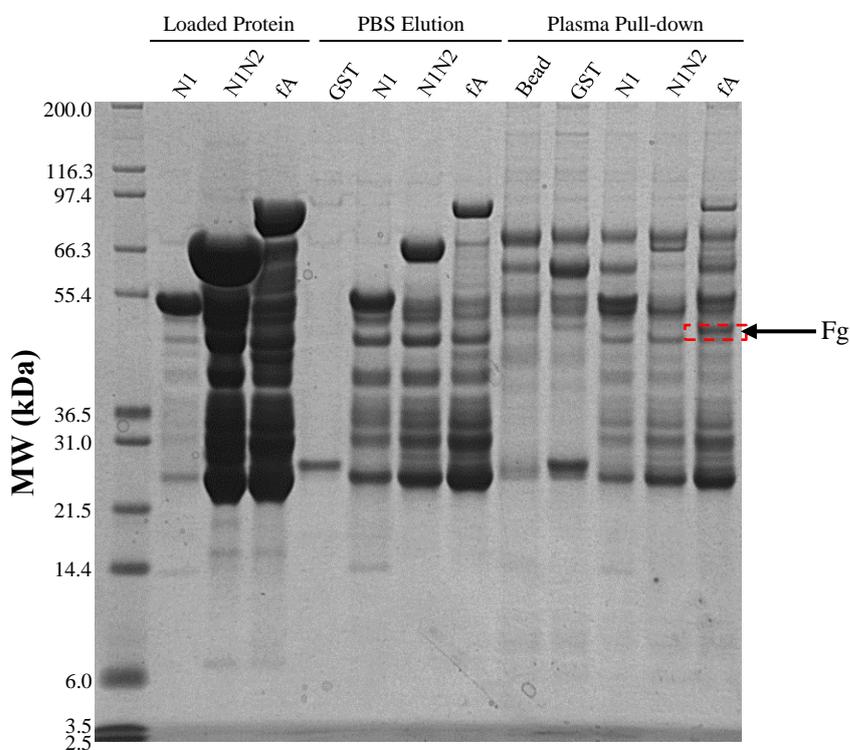


Figure 40 SDS-PAGE analysis of the A domain GST pull-down assay. Following incubation with PBS the loaded proteins were recovered, despite increased degradation. GST-fA pulled-down Fg from plasma, as expected, with a small amount of the bait protein also recovered.

In the second pull-down assay, GST-fA pulled down Fg as expected, a positive (albeit control) result. The ligand binding capacity of fA seemed to influence its susceptibility to proteolysis in plasma as intact GST-fA was recovered and eluted from the beads. Similar behaviour was observed in the previous pull-down when a

small amount of GST-N1N2 was recovered intact following the non-specific isolation of serum albumin (Figure 39). Therefore, it is likely that the major influence for the inability to recover protein is that the bait protein is cleaved from the GST-tag. In instances where a ligand is bound to the bait protein, regardless of the specificity, the cleavage site is obscured and the GST-fusion protein remains uncleaved. Despite the positive result of pulling-down Fg, the GST-fusion proteins used in this study are not ideal due to their high level of degradation. In addition, their apparent susceptibility to proteolysis in the plasma and the significant number non-specific interactions observed suggested a different approach might be more appropriate.

5.3.2 His-tag Pull-down Assays

Previous work has shown recombinant poly-histidine (His₆) tagged variants of the proteins used for the previously described pull-down experiments are less prone to degradation than the corresponding GST-fusion proteins (section 3.4). Therefore, similar pull-down experiments utilising the interaction between the His₆-tag and cobalt coated magnetic beads were carried out. The pull-down yielded results similar to those observed for the GST based assays. A large number of non-specific interactions are observed with little bait protein recovered (Figure 41). The elution conditions were again found to be suitable as intact protein was recovered from the beads when not subjected to plasma. Fg was isolated from plasma by His₆-fA, and again a small amount of intact bait protein was recovered. No novel bands were observed following the His₆-N1 or His₆-N1N2 pull-downs and no intact material was recovered. These results appear to corroborate those observed in the GST pull-down assays for fA, in that bait proteins forming complexes are less susceptible to proteolysis and can be recovered. However, competitive binding to the cobalt beads by other plasma components might also affect the ability to recover intact bait protein.

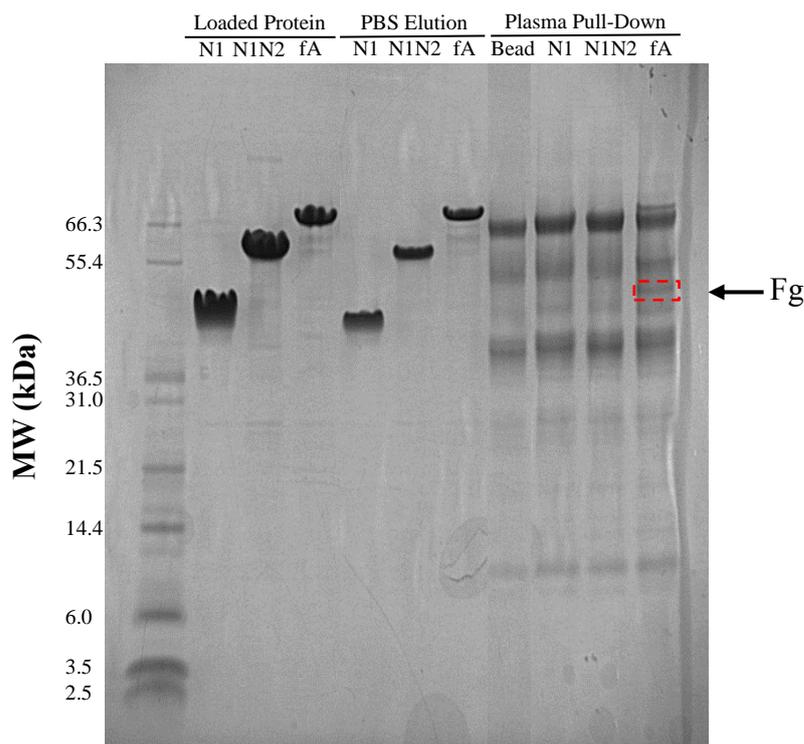


Figure 41 SDS-PAGE analysis of A domain His₆-tag pull-down. His₆-tagged variants of N1, N1N2 and fA are more stable than the corresponding GST-tagged proteins. Elution of intact bait proteins was achieved following incubation with PBS, confirming the suitability of the elution conditions. A significant amount of non-specific binding is observed when non-protein covered beads are incubated with plasma. No N1 or N1N2 was recovered. Fg is pulled-down by His₆-fA and a small amount of bait protein was recovered.

5.3.3 Recombinant Proteins are Active in Plasma

From the aforementioned affinity-tag pull-down assays, it is difficult to gauge the binding capacity of the proteins used. The apparent proteolysis of N1 and possible competitive binding to the beads renders complexes known to form *in vitro*, such as N2N3 and Fg, largely undetectable. To verify the activity of the various A domain fragments in plasma, significant amounts of intact bait protein, and consequently its binding partner, need to be recovered. Therefore, recombinant A domain proteins were covalently linked to beaded agarose and affinity columns were generated, over which analyte solutions could be passed. The method by which the affinity columns were prepared is given in section 2.4.6. In summary, the bait proteins are coupled to the beaded agarose via the primary amine groups of the N-terminal amino acid or lysine residues. N1, N2N3 and AF1 affinity columns were generated in this way,

along with a blank column to control for any non-specific interactions. Each protein contains multiple lysine residues allowing immobilisation in a number of different orientations, some of which could potentially obscure the ligand binding sites. However, a similar approach to protein immobilisation was employed in studies conducted previously (Stemberk et al., manuscript in preparation) and the ligand-binding capacity was unaffected.

Initially control solutions containing Fn or Fg at physiological concentrations were passed over the columns ($300 \mu\text{g ml}^{-1}$ and 3 mg ml^{-1} for Fn and Fg, respectively). Fn was pulled down by AF1, as expected, and Fg was pulled-down by AF1 and N2N3 (Figure 42). Lowering the pH to 3.0 eluted bound protein from the column, and this had no observable impact on the stability or binding capacity of the bait proteins. As predicted N1 did not interact with either Fn or Fg and there were no non-specific interactions with a blank column.

Following observed binding to Fn and Fg in control solutions, plasma was passed over the columns. Washing the blank column with PBS removed residual plasma components. Fn and Fg were isolated from plasma with the N2N3 and AF1 columns (Figure 43), demonstrating the physiological activity of the proteins and hence the suitability of this assay. However, no interactions involving N1 were identified.

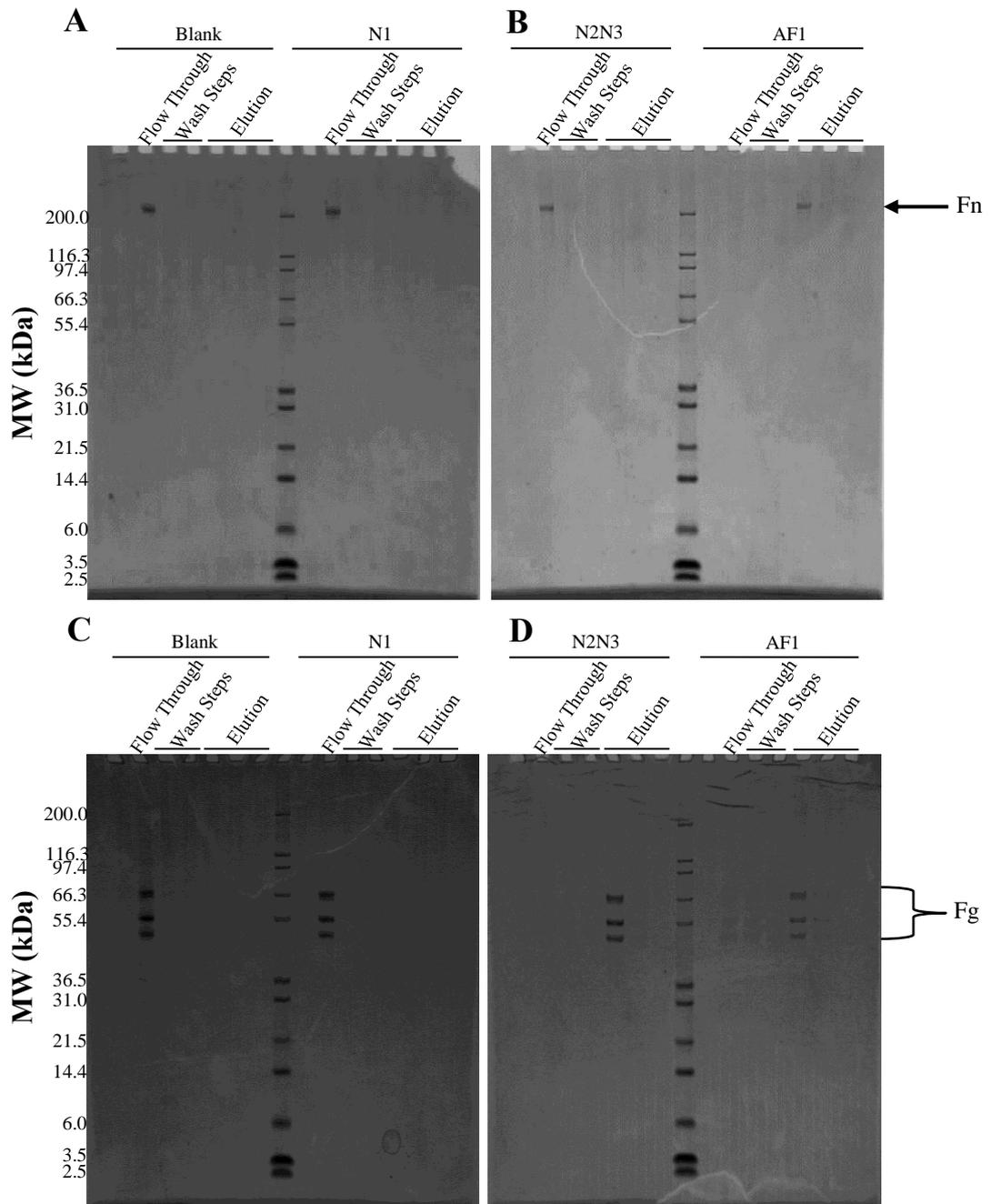


Figure 42 SDS-PAGE analysis of affinity column pull-downs from Fn/Fg solutions. Solutions containing physiological concentrations of Fn ((A) and (B)) and Fg ((C) and (D)) were passed over the prepared columns to confirm the binding activity of immobilised proteins. Bound species were eluted by lowering the pH to 3.0.

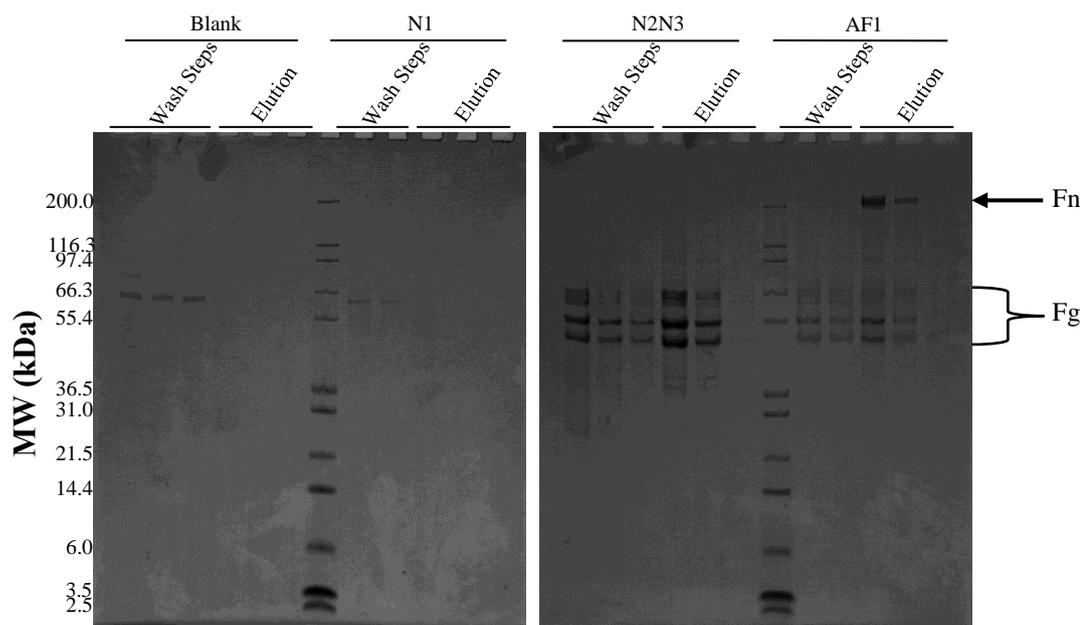


Figure 43 SDS-PAGE analysis of affinity column plasma pull-downs. No non-specific interactions were observed when plasma was flowed over the black column. N1 does not pull down any plasma components. N2N3 and AF1 pulled-down Fg as expected. Additionally, Fn was pulled-down by AF1.

5.4 The Role of Surface Proteins in Host Colonisation

An important facet of *S. aureus*' virulence lies in its ability to persistently colonise a host (Cole et al., 2001, Farley et al., 2012). Nasal carriage of *S. aureus* is a key risk factor in the infections it causes (von Eiff et al., 2001). Approximately 20% of the population are persistent carriers, with a further 60% being intermittent carriers (Noble et al., 1967). The role of *S. aureus* surface proteins is not limited to the recognition of ECM components and a number promote adhesion to host endothelial cells facilitating colonisation (Edwards et al., 2010, O'Brien et al., 2002b, Roche et al., 2003, Edwards et al., 2012). ClfB has been shown to promote adherence to squamous nasal endothelial cells (O'Brien et al., 2002b). The otherwise non-adherent *Lactococcus lactis* (*L. lactis*) was able to adhere to squamous tissues following the expression of ClfB. Cellular adhesion mediated by ClfB involves binding to type 1 cytokeratin 10 via the N-terminal A domain, specifically the N2 and N3 subdomains. However, ClfB deficient *S. aureus* mutants were still able to adhere to nasal epithelia suggesting that other factors are involved in host colonisation. SasG has also been shown to promote bacterial adherence to squamous endothelial cells (Roche et al.,

2003, Corrigan et al., 2007). Adherence is facilitated by SasG's A domain, specifically the unique N-terminal region, although the receptor to which it binds has not been identified. FnBPA shares structural and functional similarities to both ClfB and SasG. The organisation of the ClfB A domain is analogous to FnBPA, comprising three subdomains N1, N2 and N3, with N2 and N3 adopting immunoglobulin type folds and mediating Fg binding (Ganesh et al., 2011b). The structure and function of N1 from ClfB are unknown. The unique N-terminal region of SasG's A domain is thought to be disordered (Gruszka, unpublished). The involvement of these regions of ClfB and SasG in host colonisation, coupled to the functional redundancy observed across a number of *S. aureus* surface proteins, suggests that a similar protein, such as FnBPA, could be involved host colonisation processes. However, there is no experimental evidence to support this hypothesis and FnBPA A domain-mediated host colonisation has not been investigated.

5.4.1 fA Binds to HUVECs

FnBPA A domain-mediated *S. aureus* adherence to endothelial cells was investigated by carrying out a series of cell pull-down experiments using various A domain fragments. Pull-down employing intact endothelial cells as 'bait' are commonly used to identify proteins that bind to cell surface-receptors (Valle et al., 2012, Edwards et al., 2011). Briefly, proteins implicated in surface receptor binding are incubated with the cell line of interest and species able to adhere to the cells are pulled-down from the analyte solution. The cells and analyte solution are separated and pulled-down proteins specifically detected by western blot analysis. Dr. Dawn Coverley, University of York, kindly supplied the HUVECs used in this study. His₆-tagged N1, N2N3 and fA were expressed and purified as described in section 2.3.6. 25, 50 and 75 μ M concentrations of the A domain proteins were incubated with approximately 3×10^5 cells in a 24-well plate for 10 minutes. HUVECs adhere to the surface of the wells and following supernatant removal and extensive washing with PBS, were released and harvested using SDS-PAGE loading buffer (section 2.4.7). Pulled-down proteins were detected by western blot analysis using a mouse-anti-His₆ primary antibody and a sheep-anti-mouse secondary antibody. fA was pulled-down at all concentrations by the HUVECs (Figure 44), but N1 and N2N3 were not detected. This is the first evidence that the A domain of FnBPA can adhere to

endothelial cells and potentially identifies a contribution of the N1 subdomain to FnBPA ligand binding capacity.

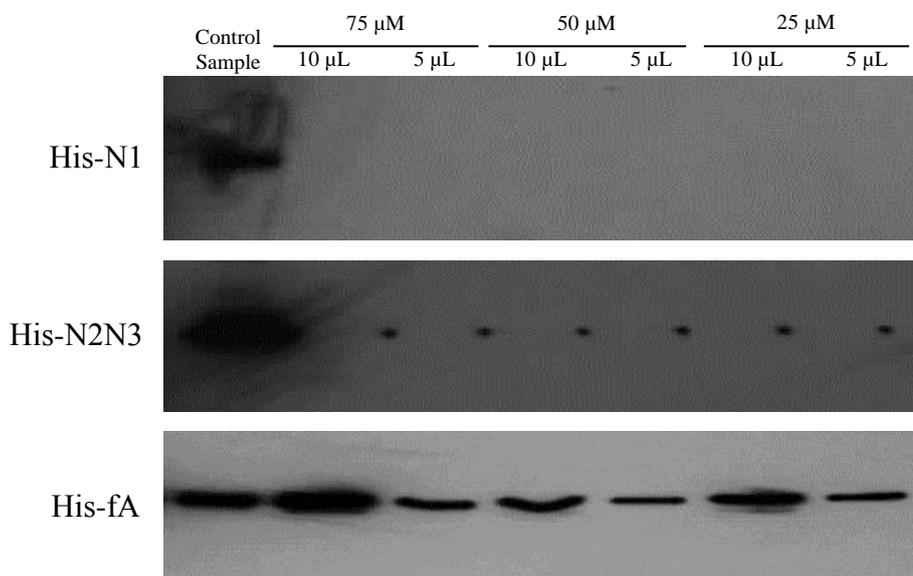


Figure 44 Western blot analysis of the HUVECs FnBPA pull-down experiment. Pull-downs were carried out in 24-well plates. FnBPA A domain proteins were added to final concentrations of 75, 50 and 25 μ M. 10 μ L and 5 μ L volumes of resuspended cell extracts were used for the blot. Pulled-down species were detected using a mouse anti-histidine tag primary antibody and a sheep anti-mouse secondary antibody.

5.5 Summary

S. aureus expresses a number of cell-wall anchored surface proteins that are implicated in the in aetiology of infective endocarditis, osteomyelitis and metastatic abscess formation (Claro et al., 2011, Edwards et al., 2010, Piroth et al., 2008). Collectively these proteins are referred to as MSCRAMMs, although some do not exclusively recognise ECM components (O'Brien et al., 2002b, Patti et al., 1994a). FnBPA is a relatively well-characterised surface protein that binds Fn and Fg via well-understood mechanisms (Stemberk et al., manuscript in preparation) (Bingham et al., 2008, Meenan et al., 2007). Fn binding is mediated by 11 Fn-binding repeats (FnBRs), while Fg-binding is achieved through the N2N3 subdomains of the N-terminal A domain. The A domain comprises a third subdomain, N1, which plays no role in Fn- and Fg-binding and to date has no attributed function. N1 is highly conserved with 90% sequence identity between strains of *S. aureus*, compared to

75% and 60% identity in N2 and N3, respectively (Loughman et al., 2008). High sequence identity in N1 does not necessarily correlate with similar identity in N2 and N3, suggesting a divergent evolution not linked to Fg-binding. The known ligands of FnBPA are both plasma components (Matsuda et al., 1978, Herrick et al., 1999). Thus, possible binding partners of N1 may also be present in plasma. N1 is likely to be the region of FnBPA projected furthest from the cell surface, as anchoring to the cell wall is achieved by a C-terminal LPETG motif (Novick, 2000). Therefore, potential interactions involving N1 might play a role in host colonisation events. Fully understanding the function of FnBPA is important from a clinical standpoint as expression of *fnbA* is attributed to increased virulence in *S. aureus* (Peacock et al., 2000).

To determine whether N1 interacted with a plasma component a series of pull-down experiments using various A domain fragments as bait proteins were carried out. Initially GST-fusion proteins were used as bait, anchored to magnetic beads via the GST/GSH interaction and incubated with plasma. A substantial number of non-specific interactions were observed by SDS-PAGE analysis and very little of the intact bait proteins were recovered. Two species were pulled-down by GST-N1N2, identified by MS/MS analysis as serum albumin, the most abundant plasma protein (Owen et al., 1993) and GST. However, these are likely to be non-specific interactions because they were not detected in subsequent experiments. Fg was pulled-down by GST-fA and, in this instance, a small amount of bait protein was recovered (Figure 40). The inability to recover the bait protein is likely due to proteolytic cleavage of FnBPA by a plasma component. Substantial amounts of GST were recovered confirming that competitive binding by plasma components to the magnetic beads does not impede binding to GST, and that the plasma GSH concentration is insufficient to elute bait proteins from the beads (Jones et al., 1998). Therefore, proteolytic cleavage within the N1 subdomain, generating the recoverable GST with the A domain fragment absent, is likely to prevent the retrieval of the GST-fusions. Fg binding appears to protect GST-fA enabling the intact bait protein to be recovered. Thus, ligand binding may obscure the cleavage site preventing proteolysis. However, the inherent instability of the GST-tagged proteins did not allow firm conclusions to be drawn.

His₆-tagged variants of the A domain proteins are not susceptible to the degradation observed in their GST-tagged counterparts (Chapter 3). Therefore, plasma pull-down experiments with His₆-N1, -N1N2 and -fA as bait proteins, anchored to cobalt coated magnetic beads were performed. Despite the increased stability, very little of the bait proteins were recovered. fA pulled-down Fg, as expected, and a small amount of intact bait protein was retrieved (Figure 41). These results appear to support the hypothesis that the A domain is susceptible to proteolysis in plasma and that ligand binding obscures the cleavage site. However, the high level of non-specific binding observed in all the pull-downs suggests competitive binding to the magnetic beads may also contribute to the inability to recover the bait proteins. To circumvent these problems recombinant N1, N2N3 and AF1 were covalently linked to beaded agarose to generate affinity columns via primary amine groups. Control solutions containing physiological concentrations of Fn and Fg were passed over the columns and pulled-down by the expected species (Fn by AF1, Fg by N2N3 and AF1) confirming the viability of the immobilised proteins. N1 did not interact with Fn or Fg, as expected (Figure 42). Fn and Fg were also pulled down from plasma, positive albeit control results. N1 did not pull down any species (Figure 43). That N1 did not isolate any plasma components could be attributed two factors. Firstly, the proteolytic cleavage to which N1 appears to be susceptible cleaves the region of the protein involved in ligand binding from the column. However, the amine coupling reaction is not site specific and proteins are covalently linked via any accessible lysine. Thus, it is likely that in some proportion regions of N1 involved in potential ligand binding interactions would remain on the column and any N1 ligands would be pulled down. As this was not the case, the second inference is that N1 is not involved in the binding of any plasma components.

Adherence to host tissues is a pivotal step in *S. aureus* infections, and subsequent immune evasion renders persistent infections difficult to treat (Clement et al., 2005, Sinha and Herrmann, 2005, Plouin-Gaudon et al., 2006). Surface protein adhesins have essential roles in a number of these colonisation and invasion processes. FnBPA plays a critical role in host cell internalisation, and expression of FnBPA on the surface of the otherwise non-invasive *L. lactis* enabled invasion of host cells (Sinha et al., 2000a). *S. aureus* uptake is initialised by Fn-mediated contacts between FnBPA and integrins on the surface of endothelial cells (Sinha et al., 1999). More

recently, Edwards *et al.* showed that the recognition of multiple Fn molecules enables more efficient bacterial uptake (Edwards *et al.*, 2011, Edwards *et al.*, 2010). The *S. aureus* surface proteins ClfB and SasG are able to promote *S. aureus* adherence to squamous host tissues (O'Brien *et al.*, 2002b, Roche *et al.*, 2003). Of these the role played by ClfB is best understood, with adherence mediated by binding to human type 1 cytokeratin 10 (O'Brien *et al.*, 2002b). The specific surface receptor recognised by SasG is unknown (Roche *et al.*, 2003, Corrigan *et al.*, 2007). Common to both cases is the involvement of the A domains of each protein, which share structural similarities to FnBPA. The ClfB A domain comprises three subdomains N1, N2 and N3 (Ganesh *et al.*, 2011b), with host cell adherence mediated by N2N3 (O'Brien *et al.*, 2002b). SasG-mediated adherence requires the entire A domain (Roche *et al.*, 2003), including the unique N-terminal subdomain, which is thought to be disordered (Gruszka, unpublished). Functional overlap is common in *S. aureus* surface proteins, thus it was postulated the FnPBA A domain might fulfil a role in host colonisation.

The ability of the A domain to promote endothelial cell adherence was investigated via cell pull-down experiments. HUVECs were used as bait and incubated with varying concentrations of N1, N2N3 and fA. Adherence was determined by Western blot analysis. N1 and N2N3 were not pulled down by the HUVECs. However, fA was pulled-down, representing the first evidence that the A domain interacts with an endothelial cell surface receptor (Figure 44). The role played by the individual subdomains is not yet understood, and further investigation is required to determine the minimum region of the A domain able to adhere to endothelial cells and the physiological relevance of such an interaction.

6 The Role of FnBPA in *S. aureus* Biofilm Formation

6.1 Introduction

Biofilms are multicellular bacterial communities encased in a self-produced extracellular matrix (Costerton et al., 1995). Growth in biofilms affords bacteria protection against antimicrobial agents (Nickel et al., 1985, Duguid et al., 1992) and the host immune system (Thurlow et al., 2011, Voyich et al., 2005). *S. aureus* biofilms represent a significant clinical problem as they typically form following the implantation of indwelling medical devices (Mack et al., 2006, Murga et al., 2001). The protective nature of biofilms renders such infections difficult to treat and in many cases necessitates additional surgery to remove and replace prostheses. Treatment of biofilm-associated infections is further complicated by the involvement of antibiotic resistant strains (Cha et al., 2010). Understanding the molecular basis of biofilm formation is of paramount importance to developing novel strategies for combatting staphylococcal infections.

Biofilm formation occurs as a four stage process; adherence, accumulation, maturation and dispersal (Otto, 2008). The surface of prostheses are rapidly coated with host ECM components soon after insertion (Vaudaux et al., 1989), and primary attachment is thought to be dominated by interactions between bacterial surface adhesins and deposited host factors. However, bacteria are also able to adhere to unconditioned surfaces, relying on electrostatic and hydrophobic interactions (Veenstra et al., 1996, Heilmann et al., 1997, Rohde et al., 2005). Cell accumulation, the next stage of biofilm formation, is generally categorised as two distinct mechanisms mediated by a polysaccharide intercellular adhesin (PIA) or surface proteins. PIA-mediated cell accumulation is currently the best understood, although the exact molecular mechanism is not known (Cramton et al., 1999, Heilmann et al., 1996, Izano et al., 2008b, Rohde et al., 2010). Conversely, protein mediated biofilms are less well studied and the molecular mechanisms that facilitate cell-to-cell adherence in protein-mediated biofilms are poorly characterised. The involvement of proteins in biofilm formation was identified through genomic analysis (Rohde et al., 2005, Fitzpatrick et al., 2005) or by the susceptibility of biofilm to protease

treatment (O'Neill et al., 2008). The homologous cell wall anchored surface proteins accumulation associated protein (Aap) and SasG of *S. epidermidis* and *S. aureus*, respectively, are known to mediate cell accumulation in biofilms (Conrady et al., 2008, Geoghegan et al., 2010). It has been suggested that cellular adhesion in Aap- and SasG-mediated biofilms is achieved via a zinc-induced (Zn^{2+} -induced) dimerisation of proteins on neighbouring cells (Conrady et al., 2008). Multiple association events between the repeat regions of opposing molecules leads to an extensive adhesive interface described as a “zinc-zipper” (Figure 45).

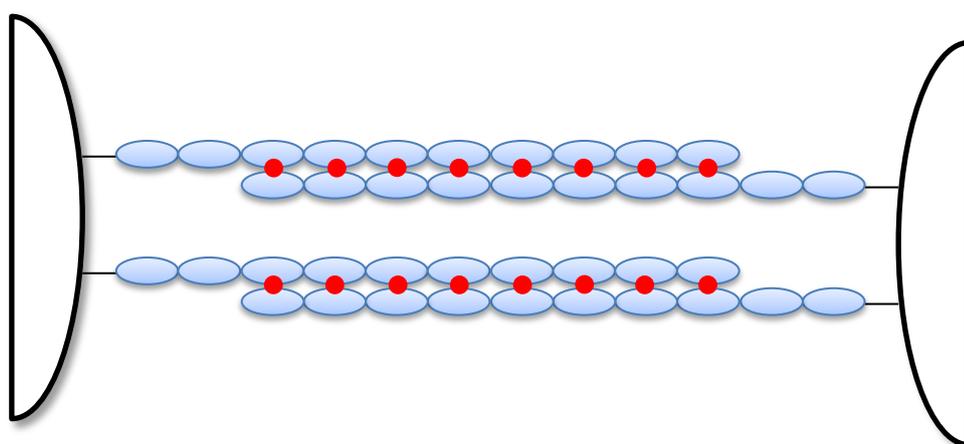


Figure 45 The zinc-zipper model for Aap and SasG mediated biofilm formation. Adapted from Conrady *et al.* (Conrady et al., 2008). Cell accumulation in *S. epidermidis* and *S. aureus ica*-independent biofilms is mediated by a Zn^{2+} induced dimerisation of the B-repeat regions of the surface proteins Aap and SasG. Coordination of multiple Zn^{2+} ions results in the formation of an extensive interface between proteins on neighbouring cells. Aap/SasG B-repeats are represented by blue ovals, Zn^{2+} ions by red circles.

A number of other *S. aureus* surface proteins, including SasC and biofilm-associated protein (Bap), have been implicated in the accumulation phase of biofilm formation (Schroeder et al., 2009, Cucarella et al., 2001). More recently, a FnBP-mediated biofilm was identified for the clinical *S. aureus* isolate BH1CC (O'Neill et al., 2009, O'Neill et al., 2008, Vergara-Irigaray et al., 2009). The formation of FnBP-mediated biofilms is triggered by mild acidic stress induced by glucose utilisation, to date a unique phenotype, and may suggest a novel molecular mechanism is utilised during biofilm development. Expression of either *fnbA* or *fnbB* was sufficient to restore the biofilm forming capacity in a double *fnbAfnbB* mutant of BH1CC (O'Neill et al.,

2008). Bacteria were still able to adhere to a surface following the deletion of the FnBPs suggesting their roles lie in the accumulation phase of biofilm formation. The A domains of FnBPA/B were found to be necessary for biofilm formation. Mutations of the FnBPA A domain in the Fg-binding region did not impair biofilm development (O'Neill et al., 2008), suggesting that the ability of FnBPA to mediate cell-accumulation is not related to its ligand binding capacity. Despite being commonly associated with methicillin resistant *S. aureus* (MRSA) (O'Neill et al., 2009, O'Neill et al., 2008) and the clinical relevance of MRSA infections (Klevens et al., 2007, Naimi et al., 2003), the molecular basis of FnBP-mediated biofilm formation is not understood.

6.2 Aims

The aim of this work was to investigate the molecular basis of FnBPA-mediated cell accumulation in *S. aureus* biofilms. The effect of zinc on protein oligomerisation was investigated as Zn²⁺-induced dimerisation is thought to be involved in protein-mediated biofilm formation. Possible interactions between FnBPA and other biofilm extracellular matrix components were also investigated.

6.3 fA Dimerises in a pH and Zn²⁺ Dependent Manner

SasG- and Aap-mediated biofilms of *S. aureus* and *S. epidermidis*, respectively, have been shown to be dependent on the presence and availability of divalent zinc cations (Conrady et al., 2008, Geoghegan et al., 2010). Chelation of Zn²⁺ resulted in dispersal of these biofilms (Geoghegan et al., 2010). FnBP-mediated biofilms exhibit similar behaviour (Geoghegan, unpublished). The addition of the chelator diethyltriaminepentaacetic acid (DTPA) reduced FnBPA mediated biofilm formation in a concentration dependent manner (Geoghegan, unpublished). Biofilm was only re-established following the addition of ZnCl₂, but not MgCl₂ or CaCl₂, showing that Zn²⁺ is crucial for FnBP-mediated *S. aureus* biofilm formation. Zn²⁺ is required for the activity of the major autolysin Atl (Zoll et al., 2010). In the early stages of biofilm development, Atl activity results in cell lysis and the release of DNA, necessary for primary attachment to hydrophobic and hydrophilic surfaces (Houston et al., 2011). ClfA/B can mediate primary attachment on Fg-coated surfaces in *atl*

deficient mutants; however, Zn^{2+} chelation still results in biofilm dispersal suggesting Zn^{2+} is necessary for cell accumulation. The proposed role for Zn^{2+} in Aap- and SasG- mediated biofilms is to induce protein dimerisation, which in turn drives cell accumulation (Conrady et al., 2008, Geoghegan et al., 2010). It has been suggested that FnBPA may mediate biofilm formation via a similar mechanism. Therefore, the ability of the A domain to dimerise under the conditions conducive to FnBP-mediated biofilm formation was explored using size exclusion chromatography multi-angle laser light scattering (SEC-MALLS).

FnBPA-mediated biofilms form in response to mild acid stress; biofilms formed at higher pHs are *ica*-dependent with no protein involvement (O'Neill et al., 2008, Vergara-Irigaray et al., 2009). Therefore, the ability of the A domain to dimerise at pH 7.5 and 5.5 was investigated. As indicated by the calculated masses of the eluted peaks, all three proteins were monomeric in the absence of Zn^{2+} regardless of the pH (Figure 46). N1 and N1N2 also remained monomeric with the addition of Zn^{2+} . However, at pH 5.5 and in the presence of Zn^{2+} fA dimerised.

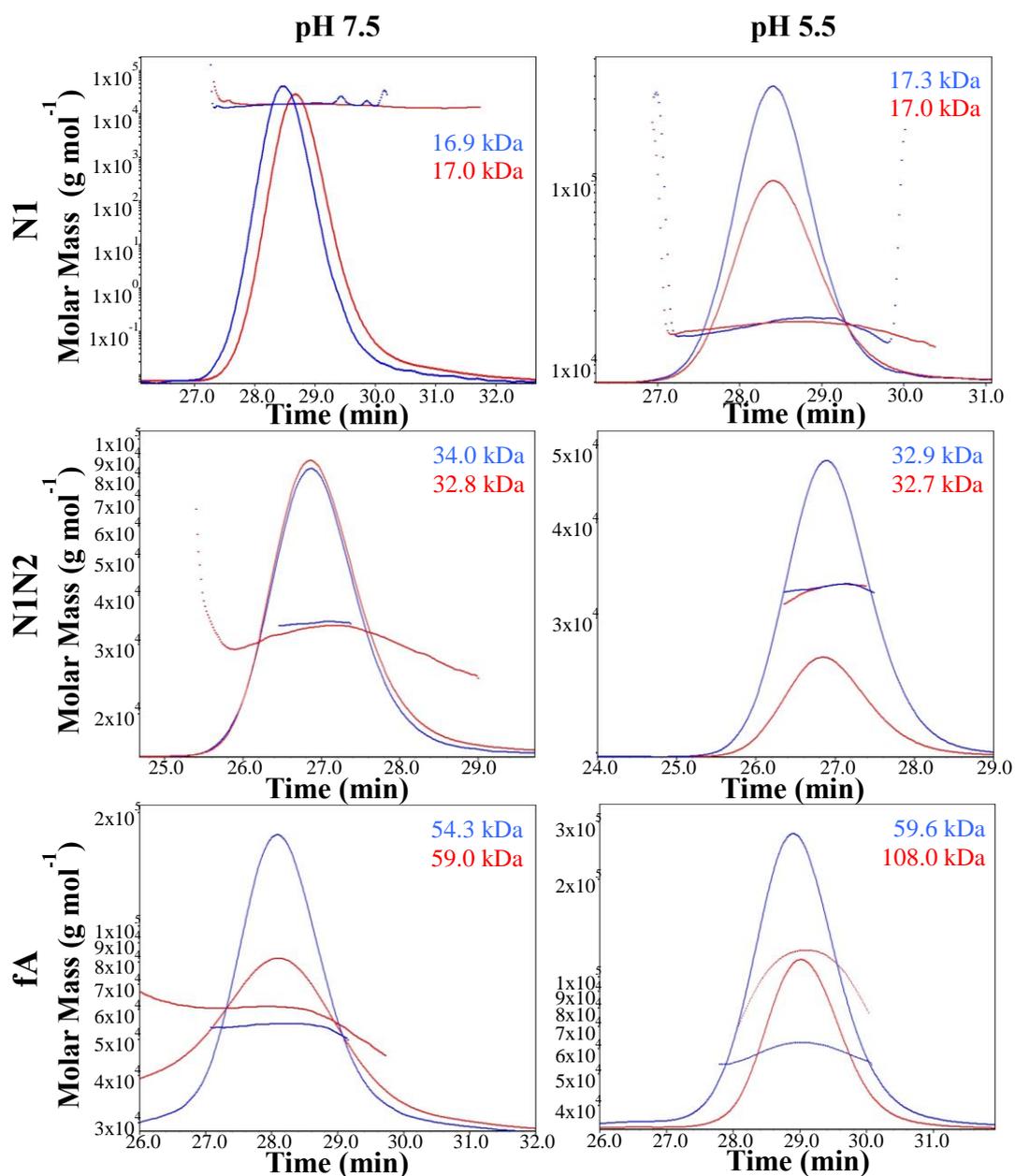


Figure 46 SEC-MALLS analysis of oligomeric states of FnBPA A domain proteins.

Chromatograms were acquired at pH 7.5 and 5.5 in the presence (red) and absence (blue) of Zn^{2+} . The estimated molar masses for each peak are indicated.

These data show that the full A domain is required for Zn^{2+} and pH induced dimerisation. The inability of N1 and N1N2 to dimerise suggests that the N3 subdomain may be the driving force behind fA dimerisation. However, N3 dimerisation could not be investigated due to the instability of the subdomain in isolation (data not shown). Therefore, the ability of N2N3 to dimerise was investigated. Regardless of the pH N2N3 is always monomeric (Figure 47).

Following the addition of Zn^{2+} N2N3 precipitated at both pH 7.5 and 5.5. However, N2N3 precipitation is not exclusively induced by Zn^{2+} , with Ni^{2+} and Mn^{2+} having the same effect (data not shown). The non-specific nature of the metal-induced precipitation suggests that the behaviour of the entire A domain, rather than the contribution of a single subdomain, is more likely to influence FnBPA's ability to facilitate biofilm formation via protein dimerisation. A model for the Zn^{2+} -induced fA dimerisation in cell accumulation is shown in Figure 48.

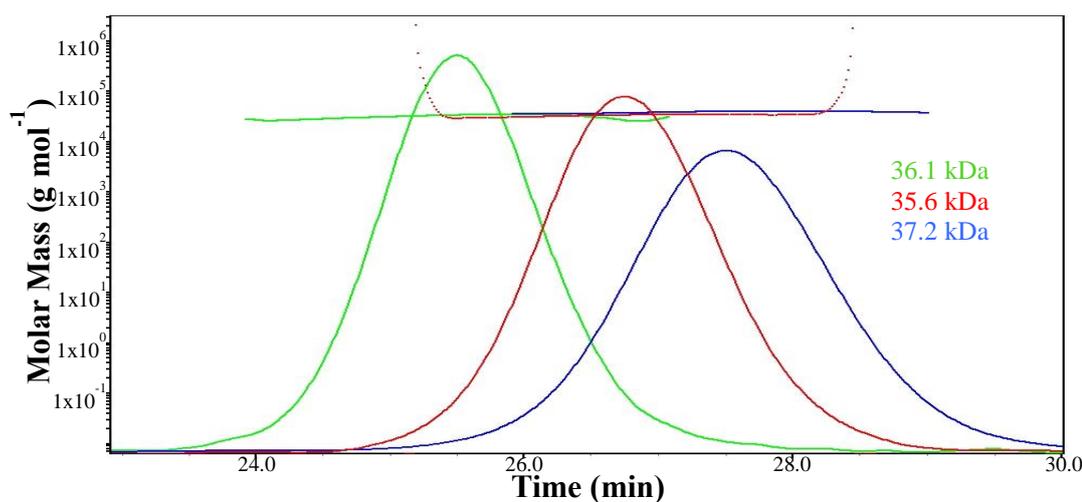


Figure 47 SEC-MALLS analysis of N2N3 at different pHs. Traces were recorded at pH 7.5 (green), pH 5.5 (red) and pH 4.7 (blue). The estimated molar mass of N2N3 from each peak is close to the measured value of 36.0 kDa.

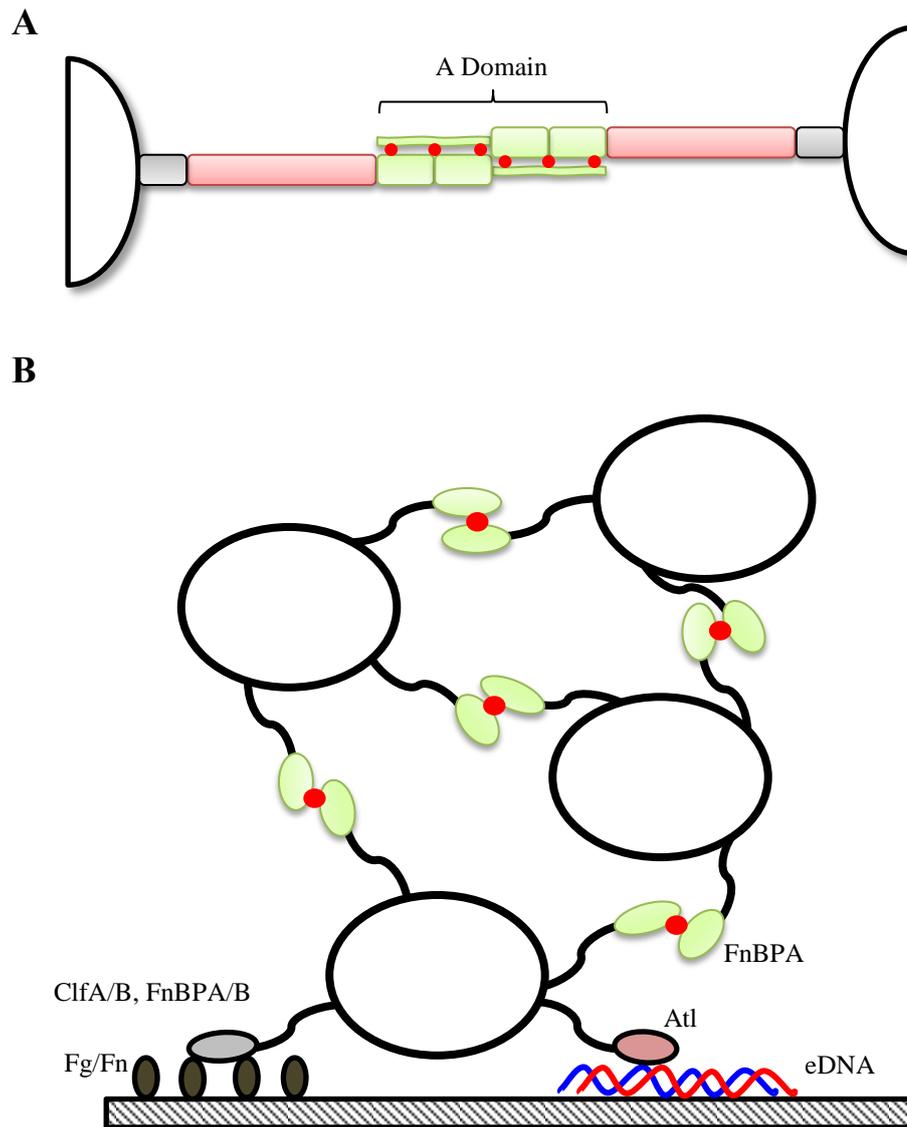


Figure 48 Cell accumulation in an FnBPA mediated biofilm facilitated by Zn²⁺-dependent fA dimerisation. Co-ordination of Zn²⁺, represented as red circles, may result in an extensive zinc-zipper type interface between the A domains of two FnBPA molecules on neighbouring cells (A). Following Fg- or DNA-mediated adherence to a surface, FnBPA dimerisation enables cell accumulation in the biofilm (B).

6.3.1 The Physiological Zn²⁺ Concentration is Too Low to Induce Dimerisation

The Zn²⁺-dependent dimerisation of FnBPA as observed using SEC-MALLS requires a Zn²⁺ concentration of 10 mM, compared to an approximate protein concentration of 20 μM, a 500-fold molar excess. When the concentration of zinc was lowered to 2 mM, the ability of FnBPA to dimerise was abolished. *In vivo* the concentration of zinc is estimated to be 10 – 15 μM, with free zinc concentration thought to be significantly lower (Outten and O'Halloran, 2001). Depending on the phase of growth, FnBPA expression is estimated to be between 1,000 and 15,000 molecules per cell (Mohamed et al., 2000). The ratio of free zinc to protein would therefore be insufficient to promote biofilm formation by the proposed mechanism.

It has been argued that the Zn²⁺ concentration in a biofilm promoting environment would be bolstered by free zinc present in the cytosol that is released on the lysis of cells. However, while the cytosolic zinc concentration is between 10⁻⁵ and 10⁻¹² M, free zinc is again thought to be significantly lower, in the femtomolar range (Outten and O'Halloran, 2001). Further, this concentration is not persistently maintained within the cell, with any free zinc not required for a specific process often non-specifically chelated by other molecules (Outten and O'Halloran, 2001). Consequently, even with the Zn²⁺ concentration supplemented by cytosolic zinc, it is unlikely to be sufficient to induce biofilm formation. Therefore, it is likely another molecule contributes to FnBP-mediated biofilms representing a novel mechanism of cell accumulation in *S. aureus* biofilms.

6.4 The Structural Role of eDNA in *S. aureus* Biofilms

The composition of the extracellular matrix is similar in all reported staphylococcal biofilms, with proteins (Shanks et al., 2008, Vergara-Irigaray et al., 2009), DNA (Izano et al., 2008a, Kaplan et al., 2011), teichoic acids (TA) (Sadovskaya et al., 2005, Sadovskaya et al., 2004) and PIA (Rohde et al., 2010, Vuong et al., 2004) consistently present. Although the relative quantities of each component can vary depending on the mechanism utilised by the bacteria to form biofilms, extracellular DNA (eDNA) is ever-present (Izano et al., 2008a, O'Neill et al., 2008). eDNA is released following Atl-mediated cell autolysis and is known to mediate primary

attachment to surfaces not coated by host ECM components (Houston et al., 2011). The susceptibility of *S. aureus* biofilms to DNase activity, demonstrated by their dispersal following DNase I treatment, suggests eDNA is also a major structural component of the biofilm matrix (Izano et al., 2008a). eDNA can promote cell-to-cell adhesion in other processes and it is reasonable to surmise a similar role may be adopted during biofilm formation (Petersen et al., 2005, Izano et al., 2008a). Interactions between proteins and DNA in a number of systems rely on the coordination of Zn^{2+} ions (Patzer and Hantke, 2000, Schmitt and McEntee, 1996). The elongated and poly-anionic nature of DNA potentially provides the requisite features to form an extensive interface with other molecules (Otwindowski et al., 1988, Richardson et al., 2009). Thus, the ability of N1 to interact with DNA, and the possible involvement of Zn^{2+} , was investigated using a gel electromobility shift assay (EMSA) (Figure 49) and NMR spectroscopy (Figure 50, Figure 51). The EMSA involved adding increasing protein concentrations, ranging from 0.1 to 1.6 μM , to a 0.4 μM solution of linearised plasmid DNA containing 100 μM zinc acetate (ZnAc_2). The samples were incubated at room temperature for 10 mins prior to electrophoresis. DNA was visualised by SYBRsafe staining. NMR analysis entailed the acquisition of ^1H , ^{15}N HSQC spectra of 0.1 mM uniformly ^{15}N labelled N1 in the presence and absence of DNA, added to achieve a molar ratio of 2:1 DNA to protein. ZnAc_2 was added to a final concentration of 200 μM . Both experiments indicate there is no interaction between N1 and DNA.

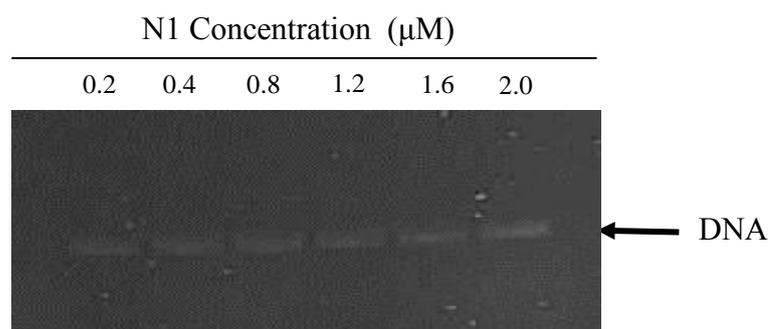


Figure 49 Gel electromobility shift assay for N1 0.4 μM . Linearised plasmid DNA was incubated with increasing concentrations of N1 and 50 μM ZnAc_2 for 30 mins prior to gel electrophoresis. DNA was visualised by SYBRsafe staining.

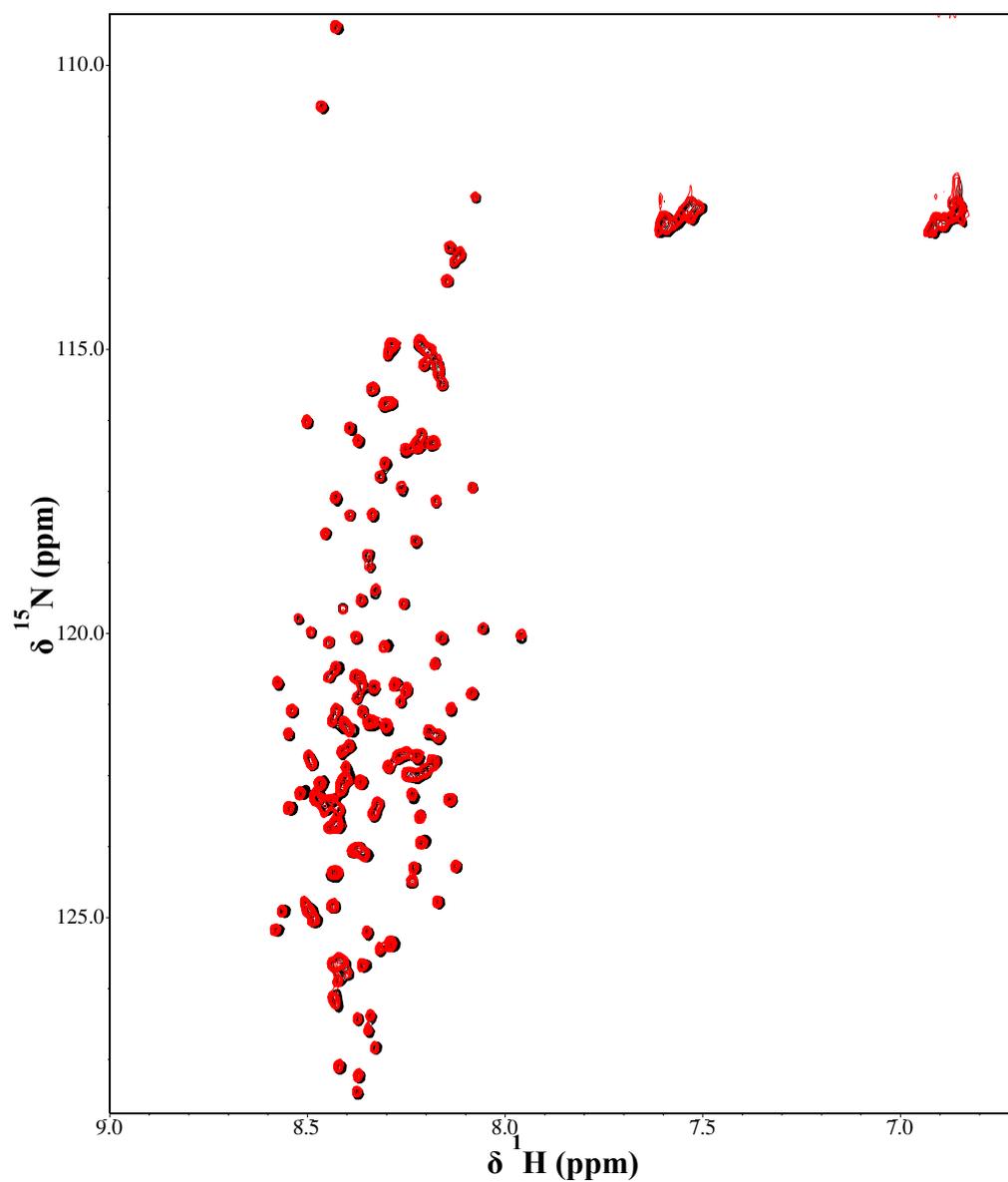


Figure 50 ^1H , ^{15}N HSQC spectrum of N1 in the presence (red) and absence (black) of DNA. Data was acquired at 298 K on a 700 MHz spectrometer at pH 5.5. DNA was added to achieve a 2:1 DNA:protein molar ratio.

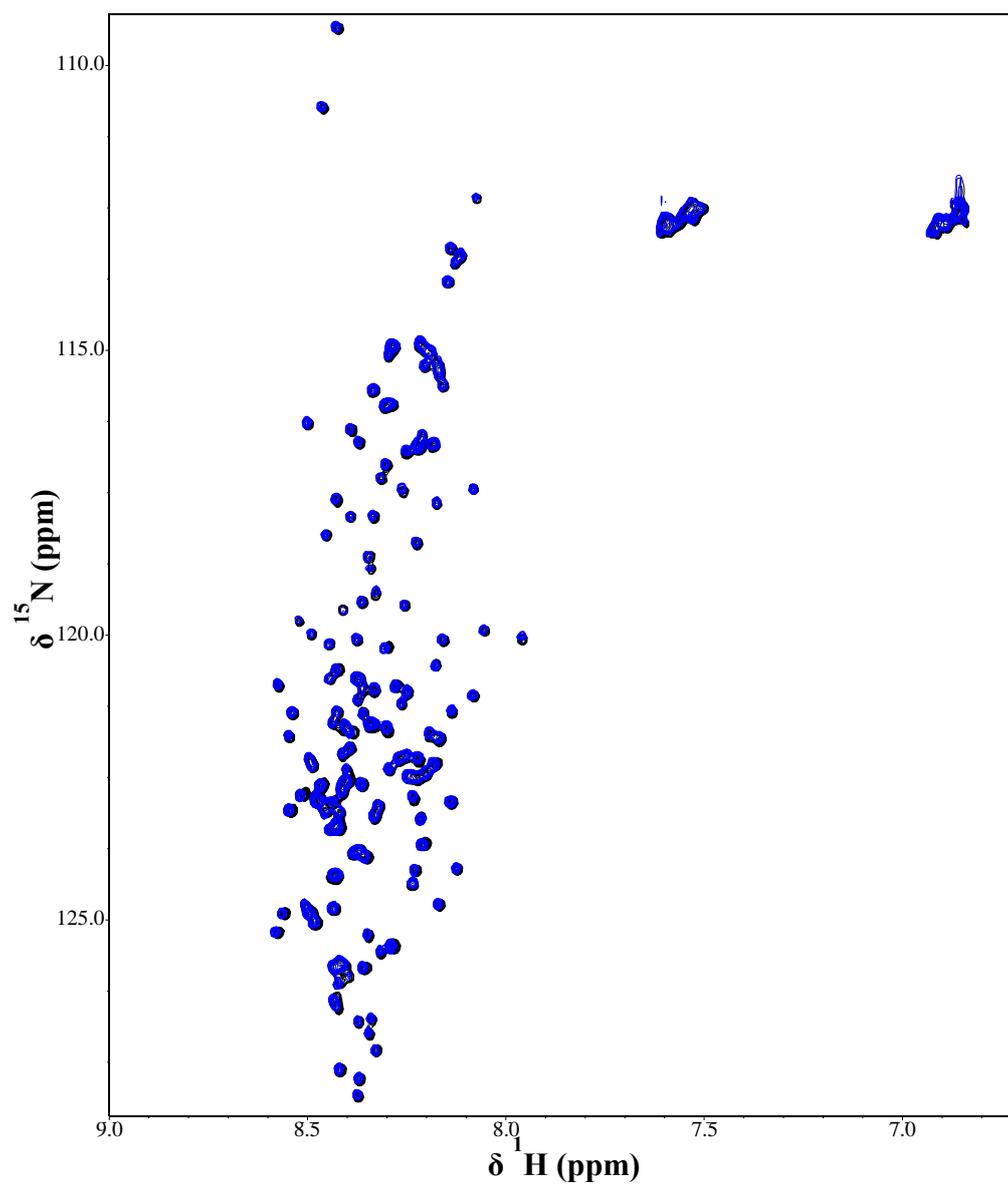


Figure 51 ^1H , ^{15}N HSQC spectrum of N1 in the presence (blue) and absence (black) of DNA and Zn^{2+} . Data was acquired at 298 K on a 700 MHz spectrometer. DNA and ZnAc_2 were both added to achieve molar ratios of 2:1 DNA/ Zn^{2+} :protein.

6.5 Teichoic Acids

S. aureus expresses a number of glycopolymers on the cell surface that play important roles in cell protection and host colonisation (Peschel et al., 1999a, Weidenmaier et al., 2004, Weidenmaier and Peschel, 2008). The most well-characterised are teichoic acids of which *S. aureus* expresses two types; the membrane lipid-linked lipoteichoic acids (LTA) (Grundling and Schneewind, 2007) and the peptidoglycan attached wall teichoic acids (WTA) (Endl et al., 1983, Xia et al., 2010a). Unlike other species of bacteria, non-expression of WTAs does not affect the viability of *S. aureus* under laboratory conditions, although cells lacking WTAs are more sensitive to environmental conditions (Weidenmaier et al., 2004, Pollack and Neuhaus, 1994, Vergara-Irigaray et al., 2008). Contrastingly, LTAs are essential for *S. aureus* viability (Grundling and Schneewind, 2007). WTAs play a pivotal role in host colonisation, with WTA deficient mutants unable to colonise nasal epithelial cells (Weidenmaier et al., 2004). In addition, *S. aureus* lacking WTAs exhibited reduced adhesion to endothelial cells and attenuated virulence in rabbit models of infective endocarditis (Weidenmaier et al., 2005). As such teichoic acids represent a major virulence factor in *S. aureus* infections.

WTAs are covalently attached to peptidoglycan by a phosphodiester linkage to N-acetylmuramic acid sugars (Neuhaus and Baddiley, 2003, Harrington and Baddiley, 1985) (Figure 52). The linkage unit consists of a N-acetyl glucosamine (GlcNAc) disaccharide and up to 3 glycerol-phosphate repeats. A poly-ribitolphosphate main chain follows the linkage unit comprising between 20 and 40 repeats (Endl et al., 1983, Neuhaus and Baddiley, 2003, Fischer, 1994). The main polymeric chain of LTAs is composed of approximately 23 glycerol-phosphate repeat units (Grundling and Schneewind, 2007, Fischer, 1994) (Figure 52). Preceding the main chain, a glucose disaccharide is attached to fatty acid groups via an ester linkage. In turn, the fatty acids anchor LTAs to the cell membrane. The ribitol- and glycerol-phosphate repeats, of WTA and LTA, respectively, are supplemented with D-alanine or GlcNAc. The incorporation of D-alanine and GlcNAc varies depending on the processes teichoic acids are involved in. For example, alanyl residues play an important role in resistance to host antimicrobial peptides (Peschel et al., 1999a, Collins et al., 2002). Differential sugar modifications play a role in eliciting specific

host antibody responses and in resistance to bacteriophage activity (Xia et al., 2010b).

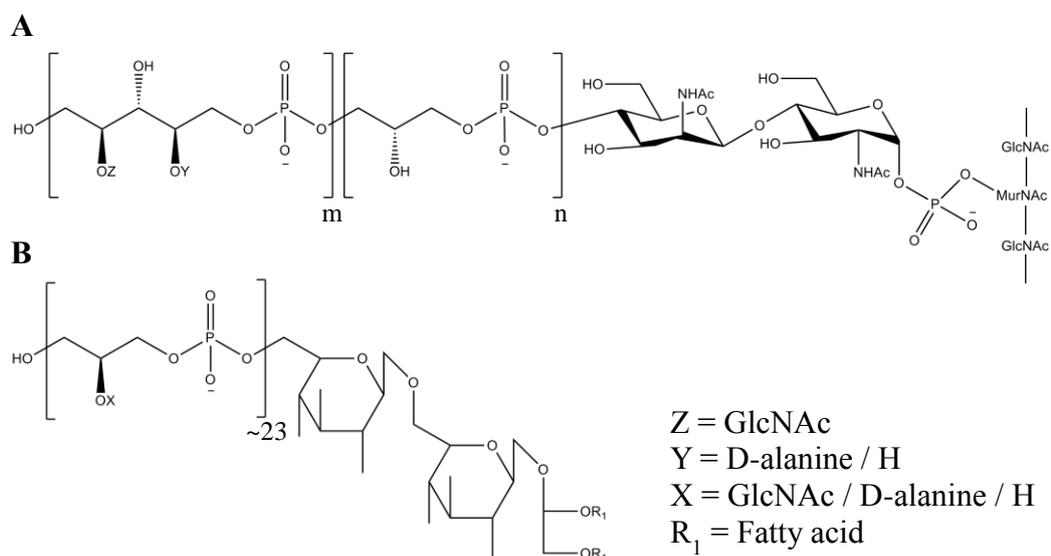


Figure 52 Chemical structures of teichoic acids from *S. aureus* The main chain of WTA (**A**) comprises ribitolphosphate repeats enriched with either D-alanine or GlcNAc. A phosphodiester linkage to acetylmuramic acid achieves covalent attachment to the cell wall. The main chain of LTA (**B**) is composed of D-alanine or GlcNAc enriched glycerol-phosphate repeats. Fatty acid groups anchor LTAs to the cell membrane.

An extracellular variant of teichoic acid (eTA) has also been identified as a structural component of staphylococcal biofilm matrices in both PIA- and protein-mediated biofilms (Kogan et al., 2006). Initially, teichoic acid was found to be a constituent of the *S. epidermidis* biofilm extracellular matrix (Sadovskaya et al., 2005). A more recent study found that teichoic acids were also structural components in biofilms formed by *S. aureus* (Vinogradov et al., 2006). eTAs in the biofilms of both species had the same structure as WTAs, and it has been suggested that WTA is released from the cell surface to become part of the biofilm matrix, although the molecular basis of WTA release is unknown. To date the role of eTAs in protein-dependent biofilms is not understood. Although Zn²⁺-mediated protein dimerisation is generally thought to promote cell accumulation, the concentration of Zn²⁺ required to induce FnBPA dimerisation is thought to be unachievable *in vivo* (Section 6.3.1). Therefore, an interaction between FnBPA and teichoic acids may promote cellular adhesion.

6.5.1 Purification of *S. aureus* WTAs

WTAs were purified from the laboratory *S. aureus* strain 8325-4 kindly supplied by Dr. James Moir, University of York. The purification was carried out largely as described by Peschel *et al.*, with minor modifications made to ensure the WTAs were expressed and purified in conditions similar to those in the biofilm (Peschel *et al.*, 1999a) (section 2.3.8). Principally this entailed tight control of the pH as the alanyl group is labile under basic conditions, with 50% lost per hour at pHs greater than 6.5 (Xia *et al.*, 2010b). FnBP-mediated biofilms form under mildly acidic conditions and it is likely that a significant number of ribitolphosphate repeats will include alanyl groups. *S. aureus* was cultured in LB broth supplemented with 1% glucose, and the final pH found to be pH 5.1. Initially cells were harvested and resuspended in 50mM sodium acetate, 150 mM NaCl, pH 4.5, to be lysed by sonication. However, the cells showed considerable resistance to sonication and lysostaphin was used in subsequent purifications. Although the optimum pH for lysostaphin activity is approximately pH 7.5, it is still active at lower pHs (Trayer and Buckley, 1970). Therefore, to preserve the D-alanine content of the WTAs, incubation with lysostaphin at pH 5.8 for 30 min was followed by adjustment to pH 4.6 prior to lysate clarification. WTAs were released from the cell walls by trichloroacetic acid (TCA) treatment and peptidoglycan removed by centrifugation. The crude WTA extract was extensively dialysed against mildly acidic de-ionised water and lyophilised. Approximately 250 mg of material was isolated from a 1 L culture. The presence of WTA was confirmed by native polyacrylamide gel electrophoresis (native-PAGE) (Figure 53). WTAs were visualised following native-PAGE by staining with alcian blue, a dye that binds acidic polysaccharides (Wolters *et al.*, 1990).

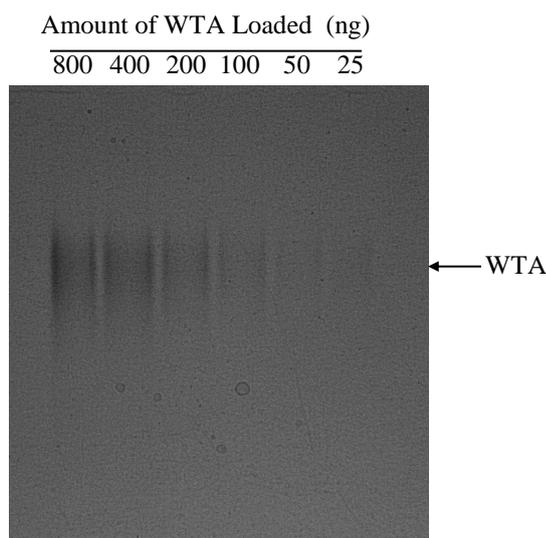


Figure 53 Native-PAGE of crude WTA extract The gel was run at 100 V for 2 h at 4 °C. Varying amounts of WTA were loaded as indicated. WTAs were visualised by Alcian blue staining. The smeared bands are indicative of multiple species with varying main chain polymer lengths.

6.5.2 Preliminary Study of N1 and WTA Interaction

The N2 and N3 subdomains have well characterised roles in ligand binding (Stemberk et al., manuscript in preparation) (Keane et al., 2007b, Wann et al., 2000), whereas N1 has no attributed function. It has been suggested that N1's primary role may be to mediate biofilm formation. Therefore, WTA binding to N1 was investigated. ^1H , ^{15}N HSQC spectra of uniformly ^{15}N labelled N1 were acquired in the absence and presence of WTAs (Figure 54). A significant number of peaks in the spectrum are affected suggesting N1 interacts with WTA in a residue specific manner. This is the first suggestion of a WTA-FnBPA interaction and it may have a significant role in FnBP-mediated biofilm formation. However, to confirm the validity of the interaction, further purification and characterisation of the WTAs was required.

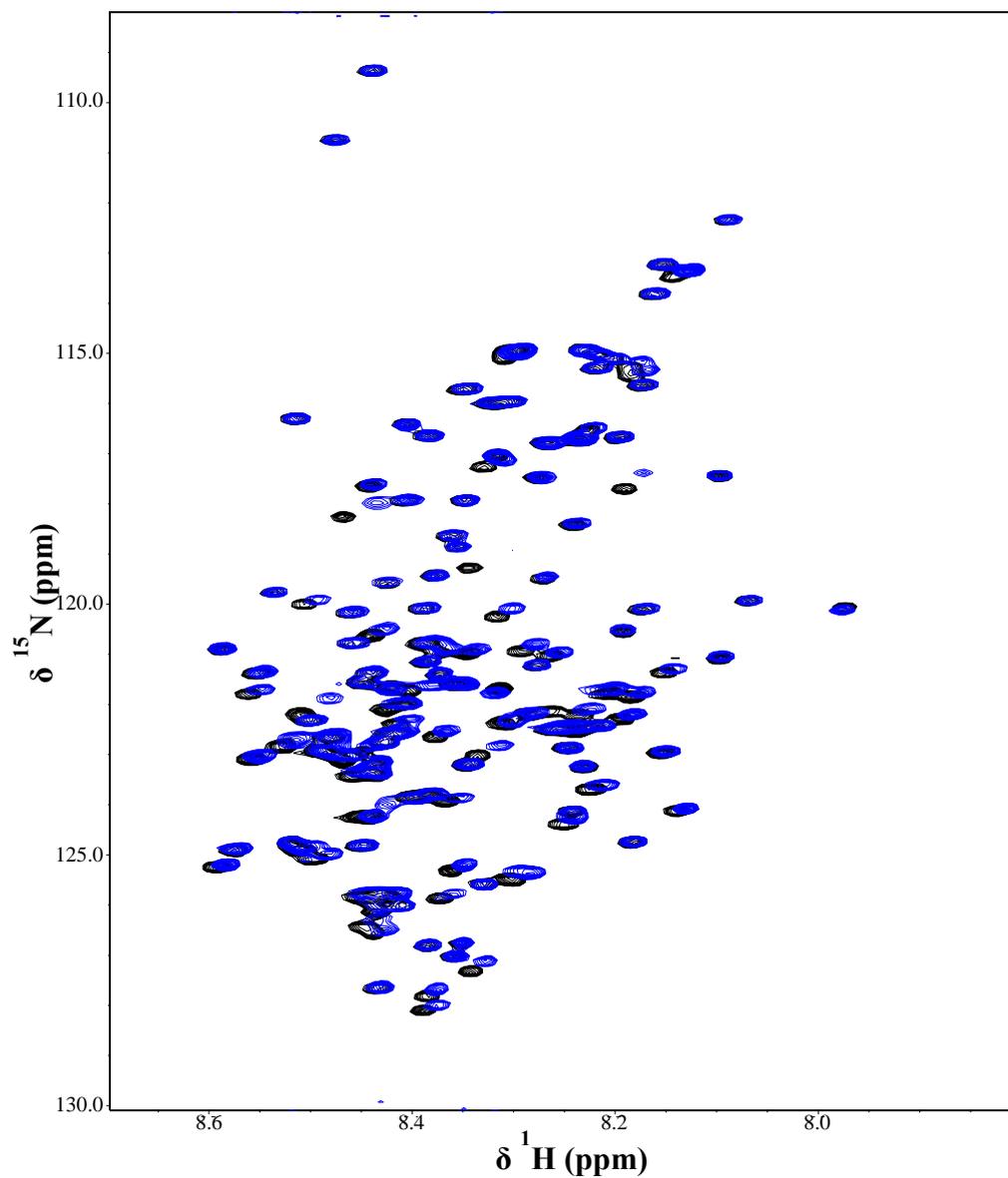


Figure 54 ^1H , ^{15}N -HSQC of ^{15}N N1 in the absence (black) and presence (red) of WTAs Spectra were acquired in 20 mM HEPES, 150 mM NaCl, pH 5.5 at 298 K on a 700 MHz spectrometer. 3.6 mg of WTAs were added to the N1 sample to achieve a 1:1 (w/w) ratio.

6.5.3 Further WTA Purification and Characterisation

The crude WTA extract contains a number of species with varying molecular weights. Although WTAs with varying main chain lengths are likely to produce the multiple species observed by native-PAGE, proteinaceous and DNA contaminants may also be present. To further purify the WTAs, size exclusion chromatography was carried out (Figure 55). Lyophilised material was resuspended in 20 mM HEPES, 150 mM NaCl, pH 5.0 to a final volume of 1 ml and loaded onto a Superdex 75 size exclusion column. During column chromatography purification, elution profiles of teichoic acids are typically monitored by analysing phosphorous content (Leopold and Fischer, 1992). However, this facility was not available and elution profiles were rather monitored by UV absorbance. TAs have a maximum UV absorbance at 205 nm (Eugster and Loessner, 2011) and there is sufficient absorbance at 220 nm to monitor elution profiles. Two species eluted from the column; a main species eluting between 15 and 20 ml and a second species exhibiting a slightly longer elution time. An alcian blue-silver staining procedure enables improved visualisation of purified WTAs and resolution of individual species with varying main chain lengths (Wolters et al., 1990). Native-PAGE analysis with alcian blue-silver staining revealed the main peak is attributable to WTAs (Figure 55), although the resolution of species with different main chain lengths was poor. There was also UV absorbance at 280 nm coinciding with the elution of WTAs that could result from proteinaceous contaminants. The absorbance at 220 nm may also result from DNA contaminants. However, staining with coomassie blue and SYBRsafe detected no protein or DNA contamination, respectively (data not shown).

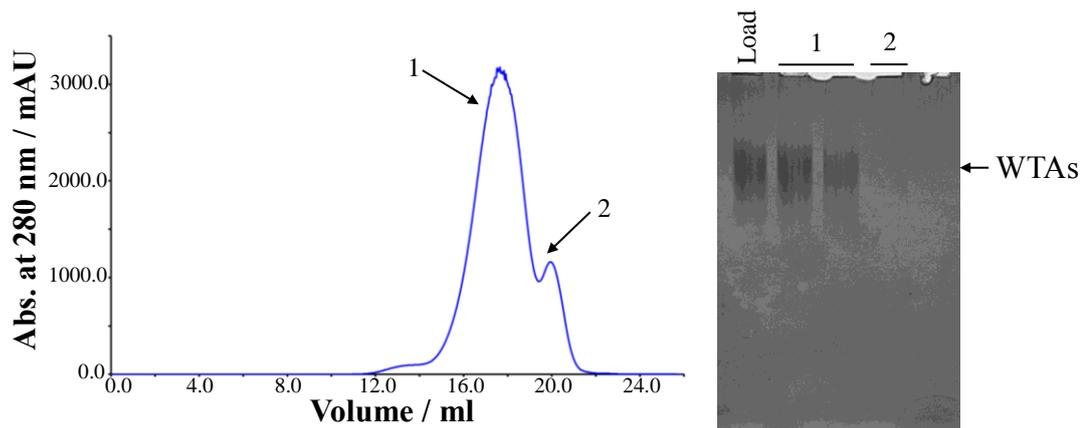


Figure 55 Size exclusion chromatography purification of crude WTA extracts. The chromatogram for the elution of WTAs shows the UV absorbance at 220 nm over 1 column volume (approximately 24 ml). The main species contains purified WTAs as can be seen by alcian blue-silver staining native-PAGE analysis.

6.5.3.1 NMR Characterisation of Purified WTA

Despite the detection of a species with the characteristics of WTAs by native-PAGE, it cannot be unambiguously identified as teichoic acids. Alcian blue does not exclusively interact with teichoic acids and is able to bind multiple acidic polysaccharides. Therefore, NMR spectroscopy was employed to unequivocally identify the purified species. A 1D ^1H spectrum with carbon decoupling and a proton-detected ^1H , ^{13}C HMQC spectrum of the purified WTAs were acquired (Figure 56). By comparison to published spectra (Vinogradov et al., 2006), the purified species was confirmed as WTA.

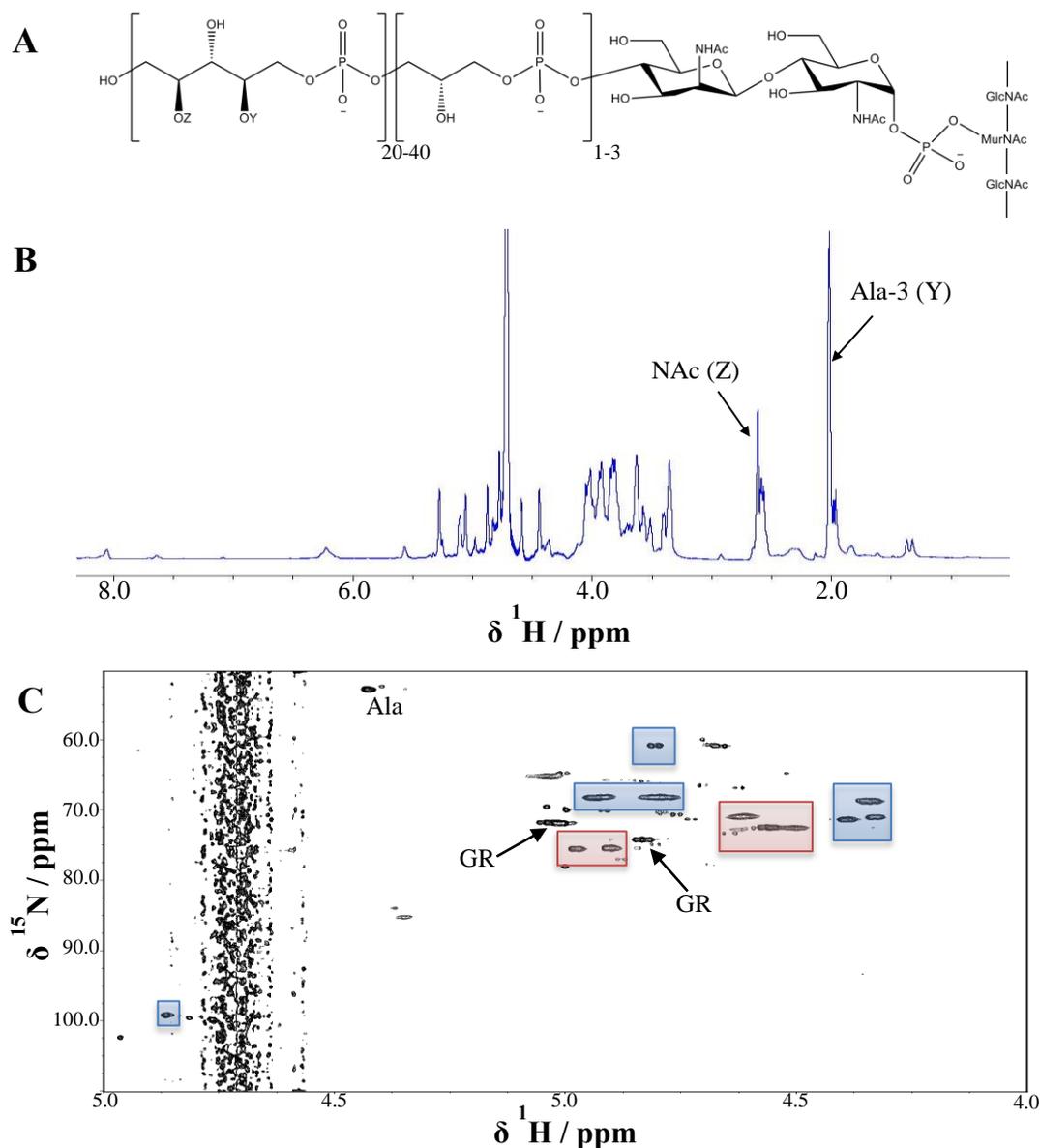


Figure 56 NMR spectra of purified WTAs . The chemical structure (A), 1D ^1H (B) and ^1H , ^{13}C HMQC (C) spectra of purified WTA from *S. aureus*. Certain features are obscured by the residual water signal. Both spectra are similar to published spectra of WTA purified from *S. aureus* MN8m (Vinogradov et al., 2006). The methyl groups of alanine and GlcNAc are clearly visible in the ^1H spectrum. Resonances in the HMQC attributable to GlcNAc are highlighted in blue and alanine $\text{H}^\alpha/\text{C}^\alpha$ correlation indicated. All other resonances are attributed to the ribitolphosphate repeat units and glycerolphosphate linkage units.

6.5.3.2 SEC-MALLS Analysis of Purified WTAs

To fully characterise the purified WTAs, the molecular weight, and hence the number of ribitolphosphate repeat units was calculated. However, as can be seen in the native-PAGE analysis of the WTAs, a number of species with different

molecular weights are present. This is due to the varying length of the polyribitolphosphate main chain, with between 20 and 40 repeat units comprising this region (Endl et al., 1983). Therefore, an average mass of the purified material was calculated using SEC-MALLS (Figure 57). To calculate molecular masses from light scattering data, a dn/dc (specific refractive index increment) value for the compound of interest is required. However, a specific dn/dc value for WTAs is not available and therefore estimates were used. Polysaccharides with similar structures to WTA have dn/dc values ranging from 0.12 to 0.13 (Joyce et al., 2003, Iwasaki et al., 2004, Iwasaki and Akiyoshi, 2006). The calculated mass of WTAs using these two extreme dn/dc values results in average masses of 17480 and 16130 kDa. The linker region of WTAs remains constant, with the main chain the only variable region (Figure 52). Therefore, by subtracting the mass of the linker unit from the mass measured by SEC-MALLS the main chain length can be calculated. Using the aforementioned molecular mass estimates, the main chain of the purified WTAs constitutes an average of either 29 or 38 poly-ribitolphosphate repeats.

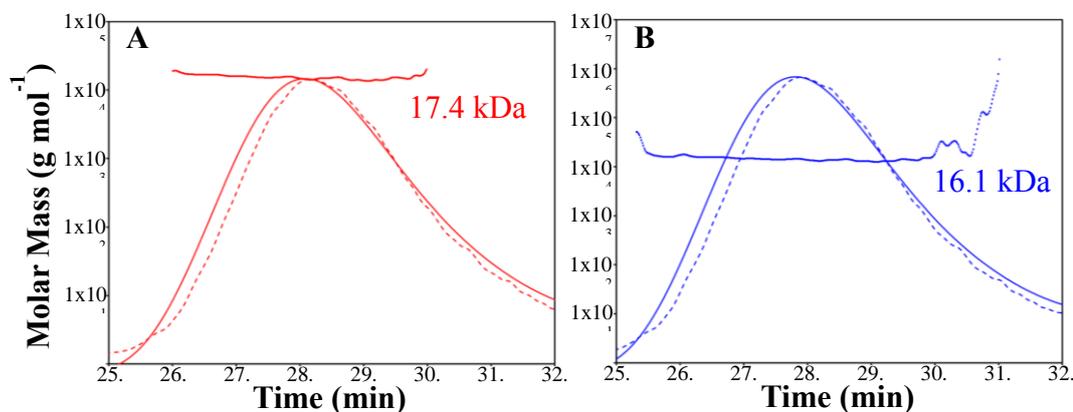


Figure 57 SEC-MALLS analysis of purified WTAs. WTAs were dissolved in 20 mM HEPES, 150 mM NaCl, pH 5.5. Using dn/dc values of 0.12 (A) and 0.13 (B) molecular masses of 17.4 and 16.1 kDa were calculated.

6.5.4 N1 Interacts with WTA in a Residue Specific Manner

The results described previously (section 6.5.2) suggested N1 interacts with WTAs. This is a novel interaction and may be crucial to understanding the molecular basis of *S. aureus* FnBP-mediated biofilm formation. Therefore, a series of 1H , ^{15}N HSQC spectra of uniformly ^{15}N labelled N1 with increasing concentrations of WTA were

recorded (Figure 58). The spectra show a clear, residue specific and concentration dependent interaction between N1 and WTAs. Resonances affected on WTA binding are predominantly in intermediate exchange on the NMR timescale. At low WTA concentrations, small changes in chemical shift and peak linewidths are observed and as a result the movement of individual peaks can be easily tracked. However, as the concentration of WTA increases, chemical shift perturbations become more pronounced and the peaks become less intense. At a 4:1 molar ratio of WTA to N1, a significant number of resonances are no longer observable in the spectrum (Figure 59). As can be seen from the chemical shift perturbation map (Figure 60), the C-terminal region of N1 is most affected by WTA binding. The SSP scores assigned to this region in the previously described chemical shift analysis (section 4.7.1) indicate a propensity for β -strand conformations. Interestingly, all chemical perturbations are unidirectional with decreasing ^1H and ^{15}N chemical shifts. This suggests that on WTA binding, backbone amide groups become de-shielded and that all residues involved in the interaction are affected in a similar way. Typically, this behaviour is observed following pH or temperature changes, factors that affect residues in a non-specific fashion (Baxter and Williamson, 1997, Cho et al., 2012, Kjaergaard et al., 2011). However, as not all residues are affected, the chemical shift changes are likely to be the result of a specific interaction.

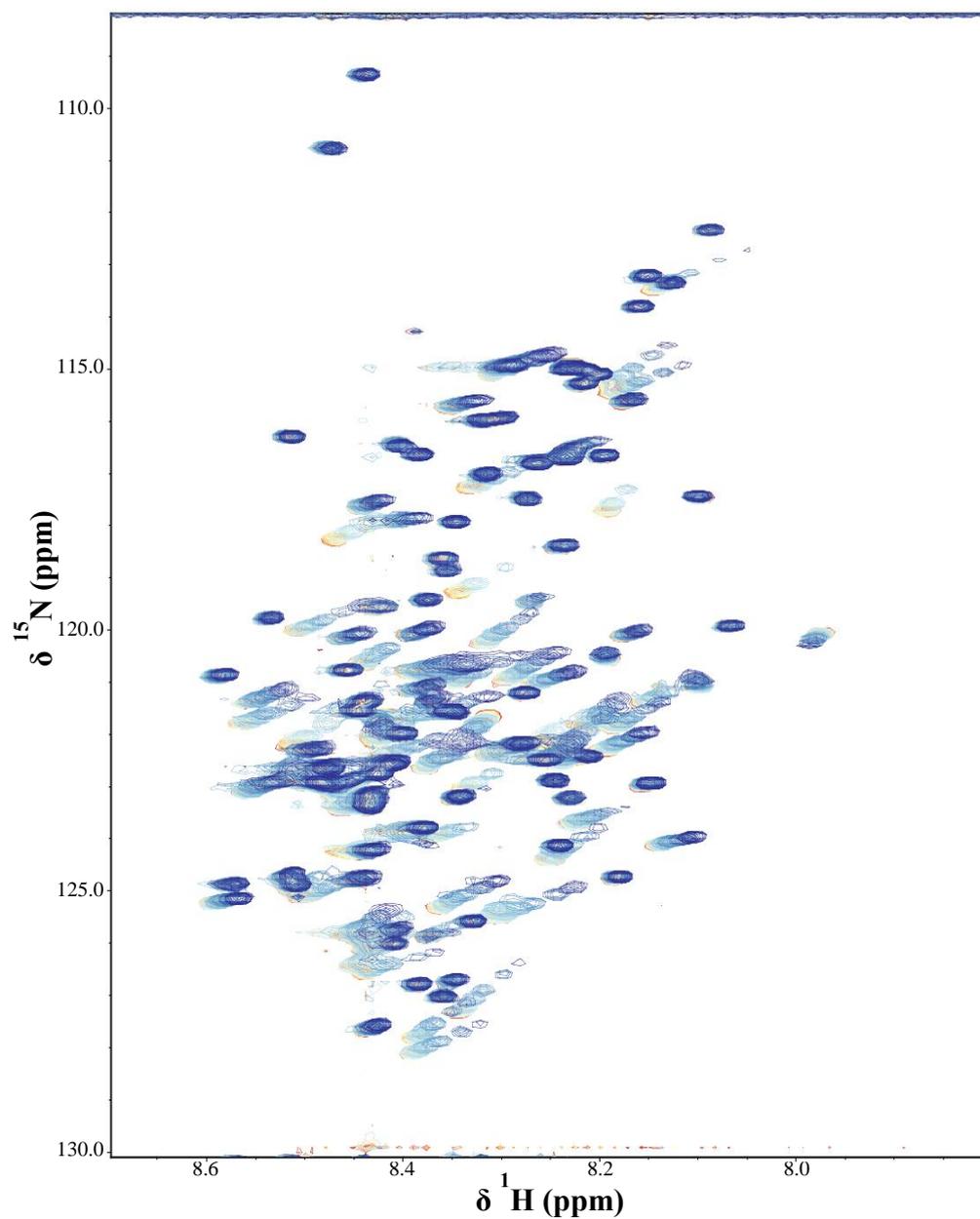


Figure 58 ^1H , ^{15}N HSQC spectra of N1 with increasing concentration of WTAs. Data was acquired at 298 K on a 700 MHz spectrometer. Uniformly ^{15}N labelled N1 dissolved in 20 mM HEPES, 150 mM NaCl, pH 5.5, with freeze dried aliquots of WTA added to give protein:WTA molar ratios of 0.05, 0.10, 0.15, 0.20, 0.40, 0.60, 0.80, 1.00, 1.50, 2.00, 3.00 and 4.00. Spectra are coloured from light to dark as the WTA concentration increases.

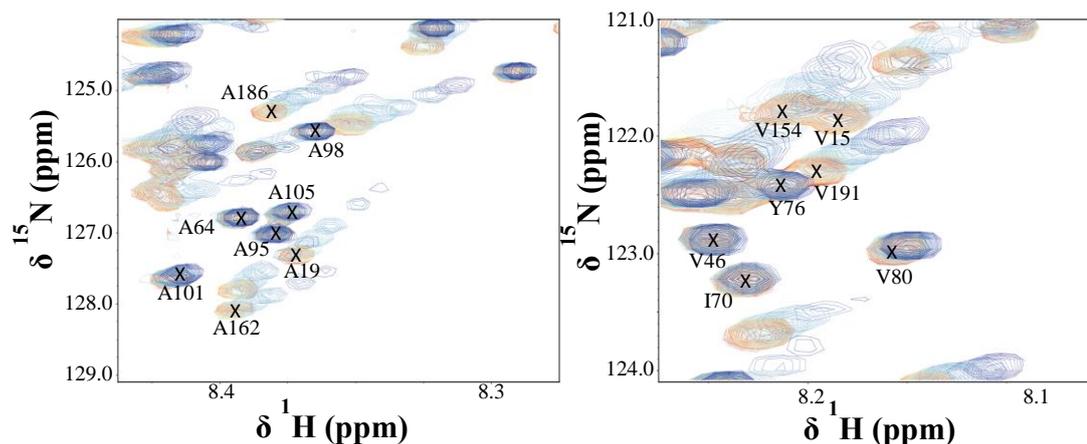


Figure 59 ^1H , ^{15}N -HSQC spectra of N1 with increasing concentration of WTAs. It is clear that the C-terminal residues of N1 are involved in WTA binding, whereas the N-terminal residues are largely unaffected, as the two example regions show. Residues involved in binding WTAs are in intermediate exchange on the NMR timescale.

The effect of Zn^{2+} on the interaction between WTAs and N1 was investigated by adding ZnAc_2 to a final concentration of $300\ \mu\text{M}$ resulting in a 1:1 protein to Zn^{2+} molar ratio. A white precipitate formed on the addition of Zn^{2+} and a significant signal reduction in the ^1H , ^{15}N HSQC of the supernatant suggests that N1 had precipitated. Zn^{2+} does not affect the solubility of N1 or WTAs individually (data not shown), therefore it is reasonable to conclude that Zn^{2+} enhances the N1:WTA interaction, inducing the precipitation of the complex or the formation of larger, insoluble aggregates. The precipitate was solubilised by EDTA, confirming the involvement of metal ions. These results provide evidence for a novel interaction that could describe the molecular basis of protein-mediated biofilm formation. Zn^{2+} is an essential extracellular matrix component of FnBP-mediated biofilms and the enhancement of the N1:WTA interaction could explain the role of Zn^{2+} . Therefore, a new model for FnBP-mediated biofilm formation is proposed in Figure 61.

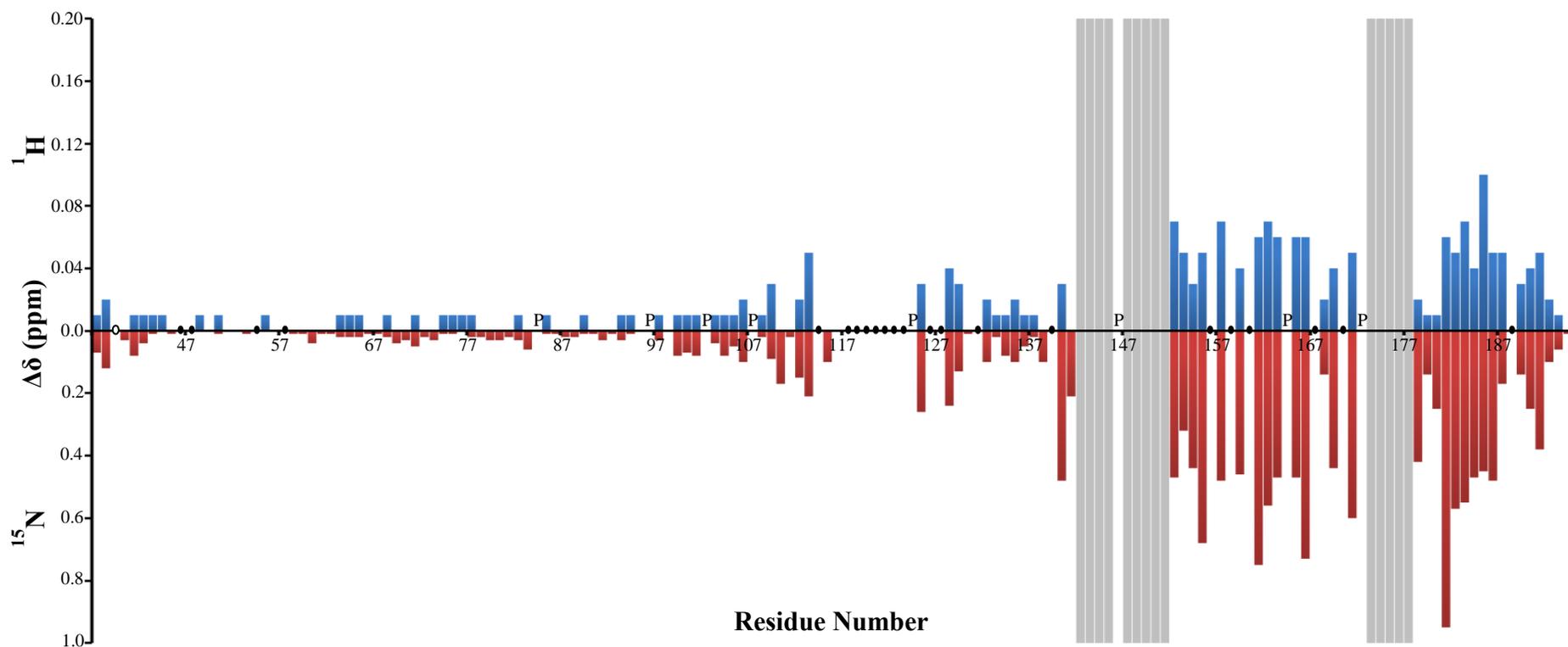


Figure 60 N1 chemical shift perturbation map following the addition of WTAs. Chemical shift perturbations in the ^1H , ^{15}N HSQC of ^{15}N N1 on the addition of a 4x molar excess of WTAs. Blue and red bars indicate ^1H and ^{15}N chemical shift changes, respectively. Peaks that disappear in the WTA-bound spectrum are indicated by a grey bar of arbitrary size. Proline residues are indicated by P; unassigned residues in the apo N1 spectrum by O; unassigned residues in the WTA-bound spectrum by ●.

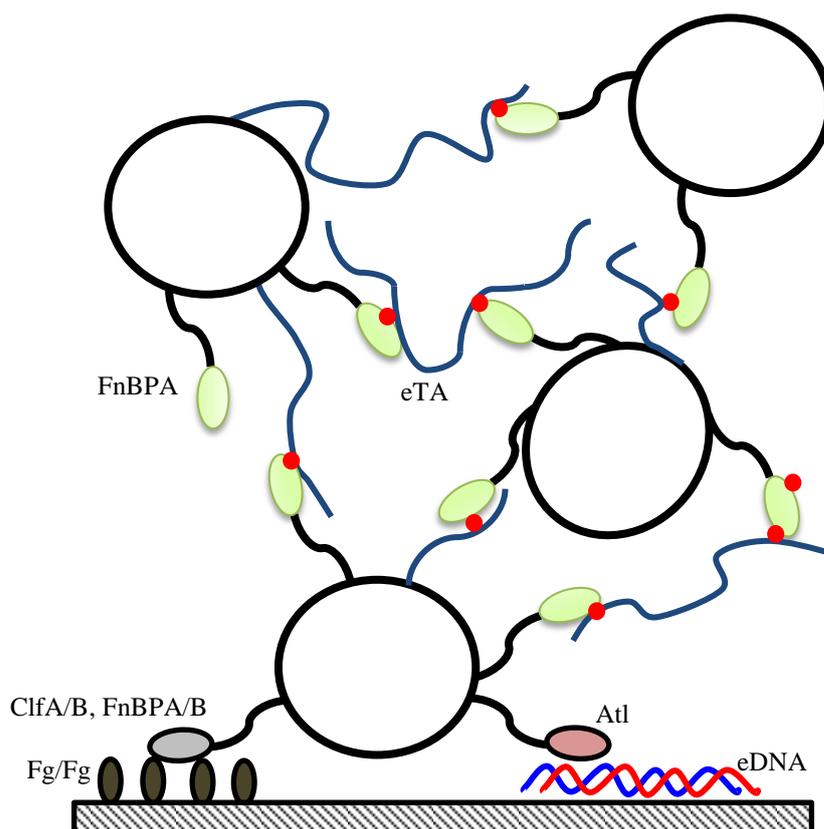


Figure 61 FnBPA/eTA-mediated model of *S. aureus* biofilm formation. Primary attachment to a surface conditioned with host ECM components is facilitated by *S. aureus* surface adhesins such as ClfA. Attachment to hydrophobic or hydrophilic surfaces is mediated by eDNA and Atl. Cell accumulation is mediated by the interaction between FnBPA and eTAs, or FnBPA and WTAs still attached to the cell. An elongated interface is formed via the coordination of Zn^{2+} ions, represented by red circles.

6.6 Summary

Staphylococci are important pathogens due to their interactions with humans. *S. epidermidis* and *S. aureus* are commensals that persistently colonise the skin and anterior nares of significant proportions of the population (Noble et al., 1967, Ziebuhr, 2001). Carriers of *S. aureus* and *S. epidermidis* are more susceptible to infection by their carried strain, particularly when host immune defences are compromised (Wenzel and Perl, 1995). Pathogenic staphylococci have a tendency to form biofilms, particularly following the implantation of prosthetic devices such as catheters (Mack et al., 2006). Growth in a biofilm affords bacteria an increased resistance to antimicrobials, and as a result infections are more difficult to treat (Cha

et al., 2010, Vuong et al., 2004). When forming biofilms, bacteria produce an extracellular matrix comprising various polymers that mediate surface adherence and cell accumulation (Flemming and Wingender, 2010, Izano et al., 2008a, Kaplan et al., 2011). Biofilms are characterised as either *ica*-dependent or *ica*-independent with the matrix composition varying accordingly (Izano et al., 2008a). *ica*-dependent biofilms utilise PIA to mediate cell-to-cell adhesion, whereas in *ica*-independent biofilms proteins facilitate cell accumulation. The current model for protein-mediated cell accumulation is known as the ‘zinc-zipper’ mechanism (Conrady et al., 2008, Geoghegan et al., 2010). Initially proposed for the *S. epidermidis* surface protein Aap, and later the homologous *S. aureus* protein SasG, Zn²⁺-induced dimerisation of proteins on neighbouring cell is suggested to drive cell accumulation.

More recently an FnBP-dependent *S. aureus* biofilm was identified that is triggered by mild acid stress (O'Neill et al., 2009, O'Neill et al., 2008). Expression of either FnBPA or FnBPB was sufficient to restore biofilm in a double *fnbAfnbB* mutant. The region of FnBPA/B responsible for promoting cell adhesion was found to be the A domain (O'Neill et al., 2008). The addition of the Zn²⁺ chelator DTPA to FnBP-mediated biofilms resulted in their dispersal and biofilm was only re-established following the addition of ZnCl₂, suggesting Zn²⁺ is a crucial component of the biofilm matrix (Geoghegan, unpublished). However, to date the molecular basis of FnBP-mediated biofilm formation is poorly understood and the specific roles played by FnBPA and Zn²⁺ are not known.

Initially, the ability of FnBPA to mediate cell accumulation via the ‘zinc-zipper’ mechanism (Conrady et al., 2008) described for Aap and SasG was investigated. SEC-MALLS analysis revealed fA was able to dimerise in a pH and Zn²⁺ dependent manner (Figure 46). Smaller fragments of the A domain were monomeric under all conditions tested suggesting that the formation of an elongated interface is necessary if FnBPA utilises a ‘zinc-zipper’ mechanism. However, a 500-fold molar excess of Zn²⁺ was required to induce protein dimerisation. Free zinc concentration *in vivo* is thought to be in the femtomolar range (Outten and O'Halloran, 2001), and despite being supplemented by cytosolic Zn²⁺ released by cell lysis during the early stages of biofilm formation, the ratio of Zn²⁺ to protein in the biofilm is thought to be

insufficient to induce protein dimerisation. Thus, other components of the biofilm matrix are likely to be involved in cell-to-cell adhesion.

eDNA and teichoic acids are consistently present in the extracellular matrix of *S. aureus* biofilms (Mann et al., 2009, Sadovskaya et al., 2005). eDNA is released during the early stages of biofilm formation following Atl mediated cell lysis, and is known to promote attachment to hydrophilic and hydrophobic surfaces (Houston et al., 2011). In addition to its role in primary attachment, eDNA is a major structural component of *S. aureus* biofilms (Mann et al., 2009, Izano et al., 2008a). Therefore, the ability of FnBPA to interact with DNA was investigated using EMSAs (Figure 49) and NMR spectroscopy (Figure 50). Both techniques indicate that there is no interaction between N1 and DNA and that the addition of Zn^{2+} has no effect (Figure 51).

S. aureus expresses two types of teichoic acid; the membrane linked LTAs and peptidoglycan attached WTAs (Grundling and Schneewind, 2007, Swoboda et al., 2010). Vinogradov *et al.* identified extracellular teichoic acids (eTAs) as a major component the extracellular matrix of a biofilm formed by *S. aureus* MN8m (Vinogradov et al., 2006). The role of eTAs in protein-mediated biofilms is not understood and to date no structural function has been attributed to eTAs. In the absence of other interactions that might stabilise FnBP-mediated biofilms, it is possible FnBPA interacts with eTAs to promote cell-to-cell adhesion. eTAs are structurally identical to WTAs, and it has been suggested that WTAs are released from the cell surface to become part of the biofilm matrix (Vinogradov et al., 2006). Therefore, WTAs from *S. aureus* were isolated and characterised by NMR spectroscopy (Figure 56) and SEC-MALLS (Figure 57). The calculated molar mass of purified WTAs suggested an average of 29 or 38 repeat units constituted the poly-ribitolphosphate main chain.

NMR spectroscopy revealed a novel interaction between N1 and WTAs that may facilitate cell accumulation in FnBP-mediated biofilms (Figure 58). Following the incremental addition of purified WTAs to uniformly ^{15}N labelled N1, sequence specific chemical shift perturbations were observed (Figure 60). WTA binding is restricted to the C-terminal region of N1, with a significant proportion of residues

affected. Chemical shift perturbations following the addition of WTAs are unidirectional suggesting that the residues involved in WTA binding are affected in the same way. Commonly such behaviour is associated with non-specific effects resulting from changes in temperature, pH and ionic strength (Baxter and Williamson, 1997). However, as not all residues are affected and constant temperature, pH and ionic strength was maintained throughout the experiment, the chemical shift changes appear to be as a result of specific interactions.

Although further study is required to determine the role the rest of FnBPA's A domain plays in FnBP-mediated biofilm formation, this is the first evidence that a teichoic acid-protein interaction could promote cellular adhesion in *S. aureus* biofilms. It also represents a novel ligand binding function of the N1 subdomain.

7 Discussion

S. aureus is a major human pathogen associated with a number of life-threatening conditions (Foster, 2004). A significant proportion of healthy individuals are persistently colonised by *S. aureus* and are susceptible to infection by their carried strain if the immune system is compromised (von Eiff et al., 2001). As such, understanding the molecular mechanisms of colonisation is important in combating *S. aureus* infections. *S. aureus* is particularly adept at forming infections following the implantation of indwelling medical devices, commonly resulting in lengthier treatment times for patients, a costly burden for health services (de Lissovoy et al., 2009). A further complication attributed to such infections is the tendency for *S. aureus* to form biofilms (Mack et al., 2006). Growth in biofilms affords bacteria increased resistance to antimicrobials and the host immune system and treatment can be particularly challenging requiring protracted treatments with antibiotics (Hoyle et al., 1992). The *S. aureus* surface protein FnBPA is associated with more invasive *S. aureus* clinical isolates and is a major virulence factor (Peacock et al., 2002). FnBPA can mediate adherence to host endothelial cells via fibronectin binding (Edwards et al., 2010), and recently an FnBP-dependent *S. aureus* biofilm has been identified (O'Neill et al., 2008). The main objectives of the work described here were to determine the structure and function of the previously uncharacterised N1 subdomain, and to elucidate the role played by FnBPA in FnBP-mediated biofilm formation.

7.1 Structural Characterisation of N1

The A domain of FnBPA comprises three subdomains, N1, N2 and N3. N2 and N3 are known interact with fibrinogen, with the C-terminus of the fibrinogen γ -chain binding in the mostly hydrophobic cleft between the two subdomains (Stemberk *et al.*, manuscript in preparation). As a result, both subdomains are well characterised structurally, adopting immunoglobulin-type conformations. Contrastingly, N1 is poorly characterised, despite being the most highly conserved region of the A domain (Loughman et al., 2008).

Initial sequence analysis of N1 revealed characteristics associated with intrinsically disordered proteins (IDPs). N1 contains a relatively high proportion of disorder promoting residues (Glu, Asp, Lys, Pro, Ser and Gln), and correspondingly few residues that would form a stable hydrophobic core. Expectedly therefore, a number of secondary structure prediction tools predicted N1 lacks stable secondary structure. Each tool implements a different algorithm to analyse the protein sequence and predict secondary structure. A consensus between each method, although not conclusive, strongly suggests N1 is an IDP. By contrast, similar analyses of the folded N2N3 subdomains revealed opposing characteristics, containing a higher percentage of order promoting residues (24%) compared to N1(17%), particularly Trp (1 compared to 0), Phe (12 compared to 0) Tyr residues (21 compared to 1). Correspondingly, and as expected, a significant proportion of β -strand was predicted from the sequence by Jpred (Figure 62).

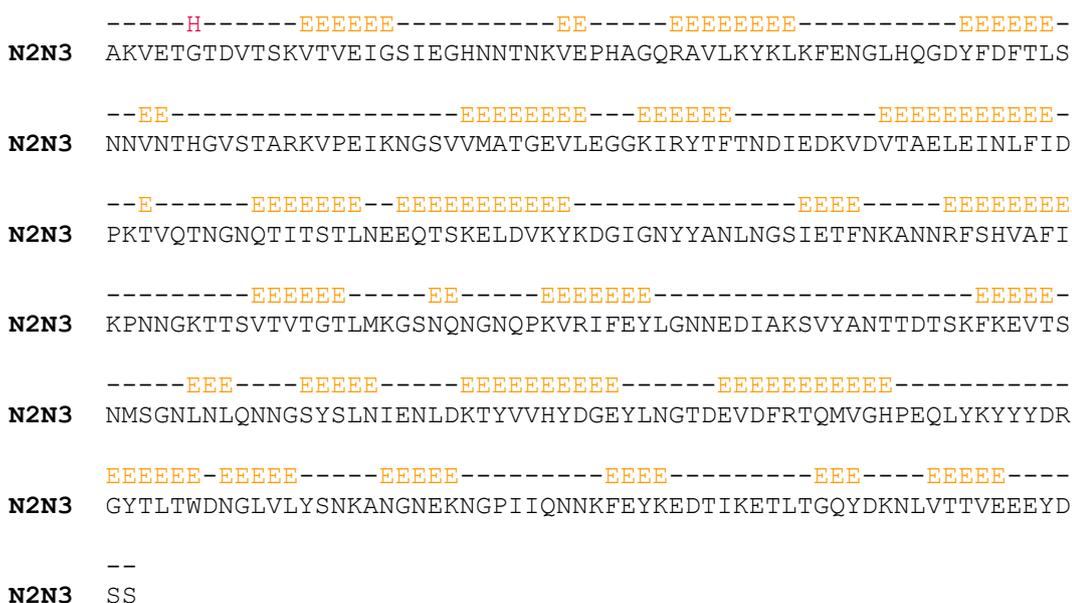


Figure 62 Secondary structure prediction of N2N3. Jpred prediction of secondary structure elements in N2N3 from FnBPA (residues 189 – 515). Helical conformations are indicated by a red ‘H’, strand by a yellow ‘E’ and loops, linker regions and disordered segments by ‘-’.

Preliminary NMR analyses seemed to corroborate the *in silico* predictions that N1 is an IDP. A one-dimensional ^1H spectrum of N1 revealed significant signal overlap in both the amide and aliphatic regions of the spectrum. The ^1H , ^{15}N HSQC of N1 exhibited similarly narrow chemical shift dispersion in the amide region of the

spectrum with the peaks also having narrow linewidths, thereby supporting the hypothesis N1 is an IDP.

The observed disorder in N1 might be the result of incorrectly defined domain boundaries and a subsequent C-terminal truncation. Incorrectly defined domain boundaries can have a significant effect on protein function, as evidenced in the revision of domain boundaries in the Fn binding region of FnBPA (Meenan et al., 2007, Schwarz-Linek et al., 2003). The absence of the N2 and N3 subdomains may also act to destabilise structural elements in N1, a characteristic observed in the B-repeat region of *S. aureus* surface protein SasG (Gruszka et al., 2012). However, NMR spectra of N1N2 and fA contain features indicative of disordered and ordered components. By superimposing the spectrum of N1 onto these spectra, it appears the disordered features result from N1 with the ordered components resulting from the independently folded N2 and N3 subdomains. A number of residues attributed to N1 in the N1N2 spectrum have slightly different chemical shifts compared to those in the isolated N1 spectrum. Principally these are at the C-terminus of N1 with resonances assigned to Gly-194, Glu-192 and Val-180 subject to the most substantial chemical shift changes. For all residues except Gly-194, there is a peak in the N1N2 spectrum in close proximity to the position of the affected residue in the N1 spectrum, in many cases overlapping. Therefore, it is reasonable to surmise the peaks correspond to the same amino acid. The Gly-194 resonance in the N1 spectrum is no longer visible, though the appearance of a peak in a region of the spectrum most commonly populated by glycine resonances is likely attributable to Gly-194. The observed chemical shift changes are not likely to indicate the formation of stable secondary structure elements. The minor chemical shift changes, coupled with the peaks attributed to N1 exhibiting the narrow linewidths associated with IDPs, suggest N1 remains disordered in recombinant N1N2. Rather, the chemical shift changes are likely to result from small changes in the local environment caused by 'nearest neighbour' effects, with nearby amino acids in a polypeptide known to influence chemical shifts in IDPs (Wang and Jardetzky, 2002, Kjaergaard and Poulsen, 2011). Alternatively the conformational freedom of N1 may be slightly reduced when N2 is present.

The TROSY-HSQC spectrum of fA displays similar features to that of N1N2, with features attributable to disordered and ordered components clearly visible. However, due to the size of fA and the different dynamics of the subdomains the spectrum is poor and the influence the N3 subdomain has on specific residues in N1 is difficult to discern. In order to unambiguously confirm N1 is disordered in the context of the intact A domain, a better quality NMR spectrum is required. The TROSY effect is optimal at higher magnetic fields (900 MHz and above) (Pervushin et al., 1997), and acquisition of a spectrum at this frequency may yield better results. In addition, acquiring spectra implementing cross-relaxation enhanced polarisation transfer (CRINEPT) steps rather than the standard insensitive nuclei enhanced by polarisation transfer (INEPT) steps for magnetisation transfer between ^1H and ^{15}N nuclei can produce improved spectra for large proteins (Riek et al., 1999, Riek et al., 2000). Nonetheless, the minimal impact the N2 subdomain has on the structure of N1, and similar features being observed in the spectrum of fA, imply that the domain boundaries between N1 and N2 are correct, and the recombinantly expressed N1 is representative of the subdomain in intact FnBPA.

Ambiguity in the NMR spectra exists as helical proteins give rise to similarly small amide chemical shift dispersions. Therefore, circular dichroism (CD) spectra of N1, N1N2 and fA were acquired. The CD spectra, and subsequent deconvolution, showed no helical characteristics suggesting the conclusion reached through NMR analysis that N1 is an IDP is likely correct.

7.1.1 Resonance Assignment of N1

IDPs often contain transient secondary structure elements or contain regions with a propensity for secondary structure (Dyson and Wright, 2005). Such regions can be identified by chemical shift analysis following completion of the resonance assignment and often provide insight into the regions of proteins likely to be involved in ligand interactions (Marsh and Forman-Kay, 2010, Norris et al., 2011). Assigning large, unstructured proteins is challenging and there are relatively few examples of completed assignments for proteins similar to N1 in size (Libich and Harauz, 2008, Mayer et al., 2012, Szalaine Agoston et al., 2011). In most cases the assignments were crucial to understanding the structure and function of the protein in

question. Therefore, the sequence specific resonance assignment of N1 was carried out. Uniformly ^{13}C , ^{15}N N1 was expressed, but substantial resonance overlap was observed in all acquired spectra. As a result, ambiguities that arose when attempting to sequentially assign N1 could not be solved and the assignment could not be completed. A common approach to improving spectral resolution is to substitute non-labile protons for deuterons, thereby decreasing linewidths (Gardner and Kay, 1998). Traditionally this approach is applied to large, but folded, proteins due to the broad linewidths associated with such species, resulting from their fast relaxation rates. IDPs do not suffer such unfavourable spin relaxation properties, but due to the signal overlap often observed reduced linewidths and improved resolution can aid the assignment procedure. Resolution was dramatically improved by deuterating N1 and nearly all expected peaks were identified in the ^1H , ^{15}N HSQC spectrum.

Despite the improved resolution achieved through deuteration, the size and disorder of N1 meant that certain regions of the spectra were still congested. Final ambiguities arising from overlapping signals were solved by selectively unlabelling specific amino acids in N1 (Rasia et al., 2012). The addition of non-isotopically enriched metabolic precursors to enriched expression medium resulted in labelled N1 with leucine, valine, proline or isoleucine selectively unlabelled. HNCOC spectra of the unlabelled proteins lack peaks attributed to residues preceded by the unlabelled amino acids. Regions with overlapping peaks therefore became less congested and individual peaks could be resolved. Further, identification of specific residue types provided reference points to the sequence and the interval between such points enabled unambiguous sequential assignment of N1. In the absence of reliable HSQC-TOCSY data, this proved invaluable for unambiguously assigning specific residues and not solely relying on interruptions in triple-resonance experiments connectivity's due to proline residues. Of the 150 peaks observed in the ^1H , ^{15}N HSQC of N1, 149 were assigned.

A potential drawback of selectively unlabelling proteins is a non-specific isotope scrambling effect resulting from the metabolism of the media supplement. Such pitfalls are avoided when supplementing growth media with precursors in the amino acid biosynthetic pathway, as they are not subject to the scrambling effects observed when supplementing with intact amino acids (Rasia et al., 2012). However, L-proline

was also used as a growth supplement to produce ${}^{\text{u}}\text{Pro N1}$, and isotopic scrambling of L-proline can reduce signal intensity throughout the NMR spectra. Glutamate and glutamine residues are particularly affected and subject to up to a 30% decrease in signal intensity. Although the exact peak widths were not measured, no significant drop in signal intensity was observed and the assignment procedure was not hindered.

Notwithstanding the success of implementing a selective unlabelling strategy to complete the resonance assignment of N1 and the use of robust and relatively simple experimental techniques, a number of other methods could have been used with similar success. As mentioned previously, NMR spectra of IDPs often suffer from extensive signal overlap. Typical approaches to protein resonance assignment, such as those implemented in this work, rely on signal detection via the amide proton. A weakness of this approach when studying IDPs is significant signal degeneracy in the amide region of the spectrum. In addition, under certain conditions, particularly alkaline pH's, amide protons rapidly exchange with water, broadening resonances and resulting in decreased sensitivity and resolution (Bai et al., 1993). ${}^{13}\text{C}'$ spins offer superior chemical shift dispersions than other species, and detection through this nucleus can improve spectral resolution. A robust procedure based on magnetisation transfers from side chain ${}^{13}\text{C}$ nuclei to the detected ${}^{13}\text{C}'$ spins has been successfully implemented to assign a number of proteins (Bermel et al., 2008). However, a pitfall of this approach when studying IDPs is reliance on poorly dispersed ${}^{13}\text{C}^{\alpha}/{}^{13}\text{C}^{\beta}$ spins.

More recently, an approach developed by Mäntylähti *et al.* implements H^{α} -detected experiments to complete backbone resonance assignments (Mäntylähti et al., 2010, Mäntylähti et al., 2011). Non-labile protons, such as H^{α} 's, are not subject to the exchange effects detrimental to H^{N} -detection based experiments. Therefore, resolution is dramatically improved in H^{α} -detection based experiments. A further advantage of H^{α} -detection methods is the unidirectional magnetisation transfer pathways they employ, reducing signal degeneracy by suppressing unnecessary transfer routes (Permi and Hellman, 2012). Consequently, it is possible to achieve near complete assignment of disordered proteins where conventional methods fall-down. Another strategy to completing complicated resonance assignments is to

increase the dimensionality of the datasets acquired, thereby reducing the likelihood of signal overlap in higher dimensions. Coupled with recent advances in sparse free induction decay (FID) sampling, collecting 4D or 5D spectra represents an attractive alternative to traditional approaches when assigning IDPs as demonstrated by Wen *et al.* (Wen *et al.*, 2011).

Secondary chemical shift analysis revealed N1 is almost entirely disordered. By employing the chemical shift index method proposed by Wishart *et al.* (Wishart *et al.*, 1991), it was found that N1 contains no regions of stable secondary structure. A limitation of calculating CSI values lies in the inability to detect transient secondary structure elements common in IDPs. Rather than adopt stable conformations, IDPs contain regions with a propensity to form secondary structure, with such elements often indicative of residues involved in ligand binding. A method proposed by Marsh *et al.* assigns a secondary structure propensity (SSP) score according to each residue representing the likelihood it adopts secondary structure. SSP analysis of N1 revealed elements with propensity helical or strand conformations in the C-terminal region of the protein (Marsh *et al.*, 2006). The N-terminus exhibits no propensity for secondary structure and is entirely disordered.

7.1.2 N-terminal Disorder in *S. aureus* Surface Proteins

There are relatively few examples of proteins similar in size to N1 completely lacking stable secondary structure and it is difficult to infer the role such a high level of disorder may play in the function of N1. A number of other *S. aureus* surface proteins contain N1 subdomains within A domains analogous to FnBPA (Figure 63). ClfA/B, FnBPB and SdrC are known to bind fibrinogen via a similar mechanism to FnBPA, also employing N2N3 subdomains that adopt immunoglobulin-type folds (Deivanayagam *et al.*, 2002, Ganesh *et al.*, 2011a, Burke *et al.*, 2011). However, no function has been attributed to their N1 subdomains. Given the structural similarity of the N2N3 subdomains, it was suggested the N1 subdomains in these proteins may also be IDPs, particularly since the initial identification of N1 was through susceptibility to protease activity. Simple analysis of the N1 subdomains reveals the low sequence complexity typically associated with IDPs and a lack of obvious

sequence similarity across these proteins (Figure 64). However, other than sharing the overall characteristics of IDPs, there is little similarity between sequences.

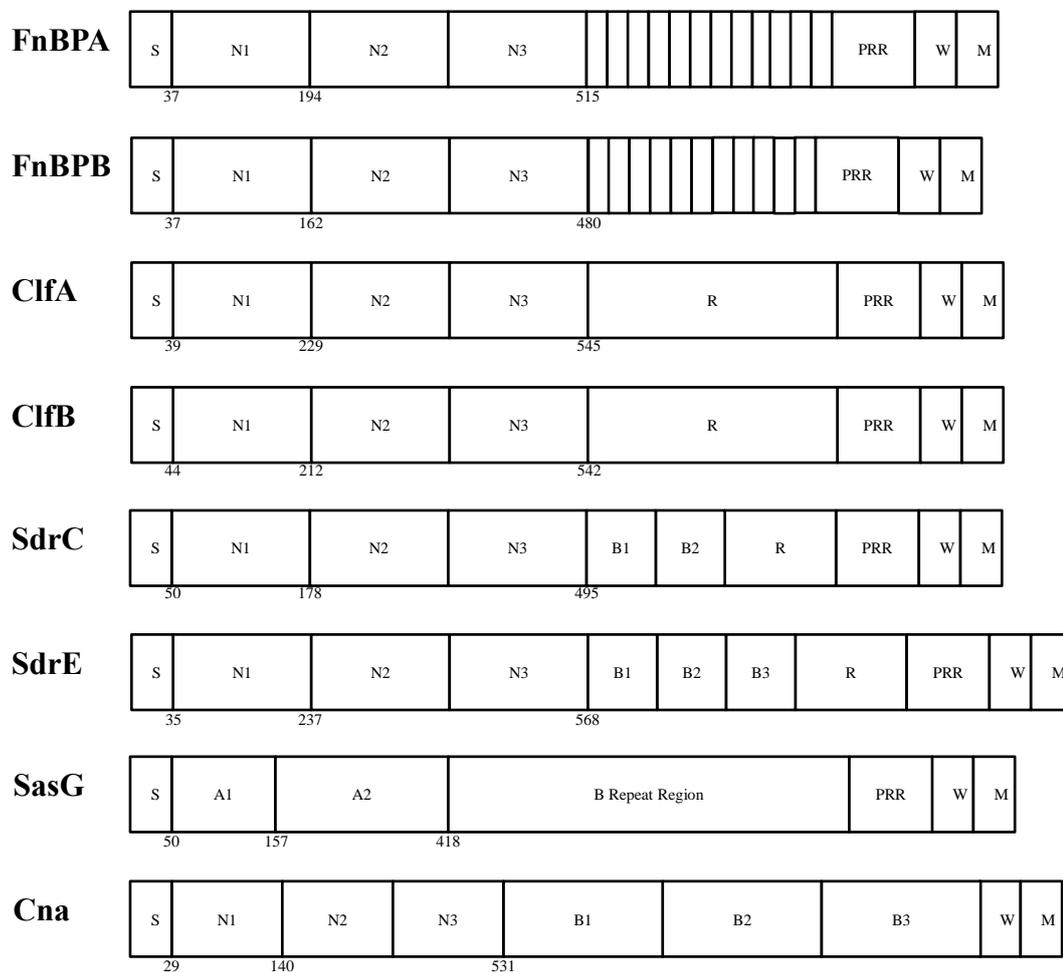


Figure 63 The domain organisation of FnbPA and other cell-wall anchored *S. aureus* surface proteins. The repeat regions (R) and regions of unknown function (B) are labelled accordingly. All A domains are organised and labelled similarly to FnbPA except SasG, which contains a predicted lectin type domain (A2) and a domain analogous to N1 (A1).

FnBPA	ASEQKTTTVEENGNSATD--NKT--SE---TQTTA-----TNVNHIE-ETQSYNATVTEQ	83
FnBFPB	ASEQNNTTVEESGSSATE--SKA--SE---TQTTT-----NNVNTID-ETQSYSATSTEQ	83
ClfA	ASENSVTQSDSASNE-SKSNDSSSVSAAPKTDNTVSDTKTSSNTNNGETSV-AQN----	92
ClfB	ASEQSNDDTTQSSKNNASADSEKNNMIETPQLNTTA-----NDTSDISANTNSANVDSTTK	98
SdrC	-----AAEHTNGELNQ-SKNETTA	67
SdrE	AAENTSTEN-AKQDDATTS DNKEVVSETENNSTT----ENNSTNPIKKETNTDSQPEAKK	106
Cna	-----	
SasG	-----AA----ENNINPTTLKDNVQSKEVKIE	73
FnBPA	PSN--A-----TQVTTEEAPKA-----VQA--PQTAQ---PANIE TVKEEVVKEE--	121
FnBFPB	PSQ--S-----TQVTTEEAPKT-----VQA--PKVET---SR-----	108
ClfA	PAQQETTQSSSTNATTEETPVTGEATTTTNTQANTPATTQSS-NTNAEELVNQTSNET--	149
ClfB	PMSTQT-----SNTTTEEPAST-----NET--PQPTAIKNQATAAKMQDQTVPQE--	141
SdrC	PSENKTTKKVD-----SRQLKDNT---QTATADQPKVTMSD-SAT-----	103
SdrE	ESTSSSTQKQQNNV-----TATTETK-PQNI EKENVKVPSTDKTA	144
Cna	-----AARDIS-STNVTDLT VSPSKIEDG	51
SasG	E-----VTNKDTA-PQGVEAKSEVTSNKDT-	97
FnBPA	----AKPQVKET-T----QSQDN-----SGDQRQ-VDLTP-KKATQNQVA-----ETQ	168
FnBFPB	-----VDLPS-EKVADKETT-----GTQ	125
ClfA	-----TFNDTNTVSSVNSPQNSTNAENVSTTQDTS TEA----TPSNNESAPQSTDASN	198
ClfB	----ANSQVDNK-T----TNDANSIATNSELKNSQT-LDLPQ-SSP-----QT	188
SdrC	-----V--KETSSNMQSPQNATANQSTTKTSNVTNDKS-STTYSNETD-KSNLTQA	151
SdrE	TEDT-SVILEEKKA-----PNNT-NNDVTT---KPSTSEPSTSEIQT KPTTPQ--ESTN	191
Cna	GKTTVKMTFDDKN-----GKIQ-----NGDM	72
SasG	-----IEH---EPSVK---AEDISKKEDTPK--EVAD	121
FnBPA	VEVAQPRTASESK-----PRVTRSADVAEAEKEA-SNA---KVETG-----	194
FnBFPB	VDIAQPSNVSEIK-----PRMKRSTDVTA VAEK-EVV---EETKAT-----	162
ClfA	KDVVN-QAVNTSA-----PRAFSLA AVAADAPA-AG-----	229
ClfB	ISNAQGTSKPSVR-----TRAVRSLAVAEP--V-VNA---ADAKGTNVNDKVT	220
SdrC	KDVSTTPKTTTIK-----PRTLNRMAVNTVAA-----	178
SdrE	IENSQPQPTPSKVDNQVTDATNPKEPVNVSKEE-----LKNNPEKLKELV-	236
Cna	IKVAWPTSGTVKIEG----YSKTVPLTVKGEQVGQAVITPDGATITFNDKVEKLSDVSG	127
SasG	VAEVQPKSSVT--HN----AETPK---VRK-----	142
FnBPA	-----	
FnBFPB	-----	
ClfA	-----	
ClfB	ASNFKLEKT----	229
SdrC	-----	
SdrE	-----	
Cna	FAEFEVQGRNLTQ	140
SasG	-----	

Figure 64 Sequence alignment of N1 subdomains. The multiple sequence alignment was performed using Clustal Omega (Wagner et al., 1983).

$^1\text{D } ^1\text{H}$ NMR spectra of the N1 subdomains from FnBPA, ClfA and the analogous domain of SasG (termed A1) suggest all three proteins are intrinsically unstructured (Figure 65). This may imply the N1 domains fulfil a common role in cell-wall anchored surface proteins. It has been suggested that the N1 subdomain of FnBPA may protect the N2N3 subdomains from action by host proteases preserving ligand binding activity (Burke et al., 2010). As other cell-wall anchored proteins also contain regions with ligand binding activity, an ever-present domain limiting protease activity would be beneficial. An advantage of disorder is an increased

ligand capture radius that could be useful in the early stages of host colonisation, particularly as the N1 subdomains are likely to be most distal from the cell surface. Such a role has been suggested for SasG, with the A1 domain essential for this function (Corrigan et al., 2007). However, as the N1 domain is unnecessary for ClfB-mediated adhesion to squamous tissue a divergence of roles is also possible (O'Brien et al., 2002b).

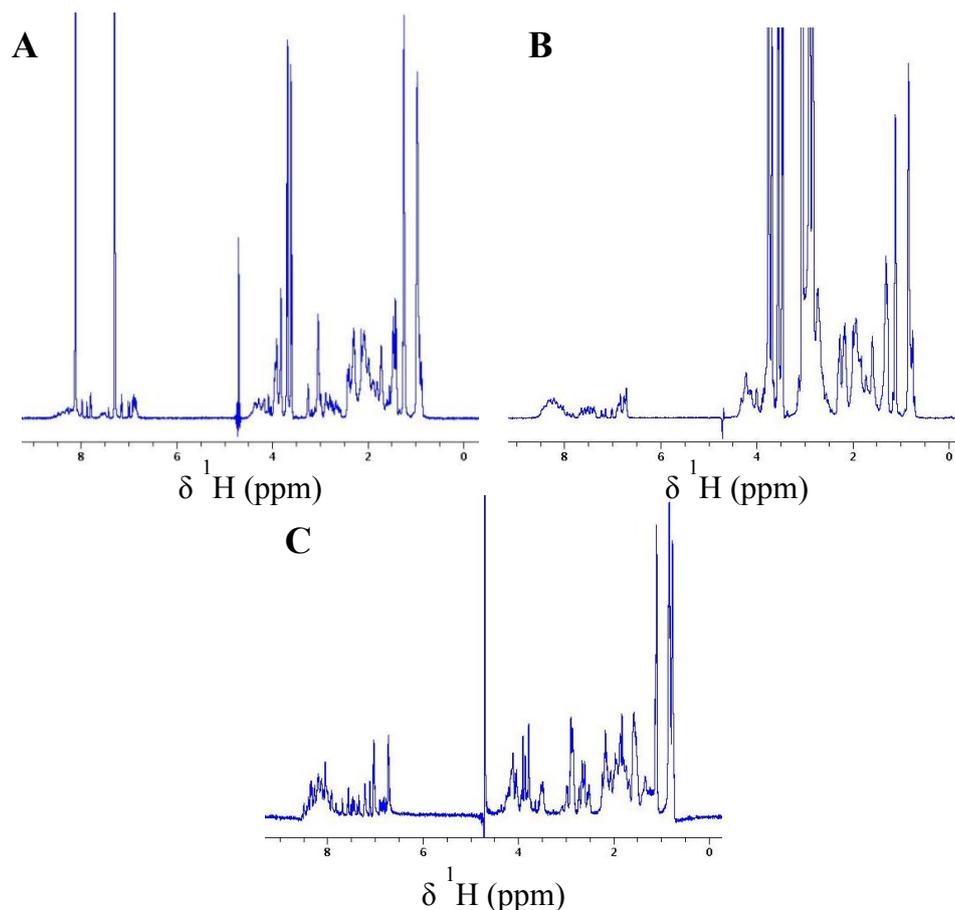


Figure 65 1D ¹H NMR spectra of disordered N-terminal subdomains. NMR spectra of N1 from FnBPA (A) and ClfB (B) and A1 from SasG (C). Data was acquired at 298 K and pH 6.0. Despite the presence of contaminants, all spectra contain features suggestive of disorder.

7.2 A Domain Interactions with Host Factors

Interactions between *S. aureus* surface proteins and host extracellular matrix proteins are of paramount importance in the early stages of a number of diseases (Arrecubieta et al., 2006, Clement et al., 2005). Indeed, the recognition of two ligands, fibronectin and fibrinogen, is thought to underpin FnBPA's importance as a *S. aureus* virulence

factor. As the N1 subdomain has no attributed function, its ligand binding capacity was investigated with a series of blood-plasma pull-down experiments. Glutathione S-transferase (GST) and hexa-histidine tagged constructs of N1, N1N2 and fA were expressed, purified and used as 'bait' proteins and incubated with plasma. Following the incubation, little intact bait protein was successfully recovered. However, when the A domain interacted with plasma components, fA with fibrinogen or N1N2 non-specifically with serum albumin, a small amount of bait protein was recovered. Therefore, it seems the A domain is subject to proteolysis in plasma, probably within the N1 subdomain, and that the cleavage site is protected following ligand binding. In addition to the inability to recover the A domain proteins a significant amount of non-specific binding to the magnetic beads used to anchor the bait proteins was observed. To circumvent the problem of non-specific interactions, a number of proteins encompassing the A domain and the first FnBR were covalently attached to sepharose beads to generate affinity columns over which plasma could be passed. Fibrinogen and fibronectin were both pulled down by constructs containing their respective binding sites. However, N1 did not isolate any plasma components.

A critical step in *S. aureus* infections is adherence to host tissues, enabling subsequent immune evasion (Edwards et al., 2011). Several surface protein adhesins have essential roles in a number of colonisation processes, interacting directly with host cell receptors or via mediating molecules (Sinha et al., 2000b). FnBPA plays such a role in host cell internalisation through the fibronectin-mediated contacts with integrins on the surface of endothelial cells (Edwards et al., 2010). This is a well-characterised process involving the FnBRs. Other *S. aureus* surface proteins with structural similarities to FnBPA, ClfB and SasG, are also able to promote adhesion to host tissues (O'Brien et al., 2002b, Roche et al., 2003). ClfB binds to human type 1 cytokeratin 10, whilst the specific receptor for SasG mediated adherence is unknown. The N2N3 subdomains and disordered A1 subdomain of ClfB and SasG, respectively, are necessary for host cell adhesion. Thus, it was postulated the A domain of FnBPA could fulfil a similar role. The ability of the A domain to promote adherence to endothelial cells was investigated by cell pull-down experiments. Human umbilical vein endothelial cells (HUVECs) were used as bait in analyte solutions containing N1, N2N3 or fA. Western blot analysis revealed fA bound to the HUVECs and is the first evidence of an interaction involving the N1 subdomain.

The reason why only fA and not N1 or N2N3 bound to HUVECs could be explained by subtle structural differences between the intact A domain and isolated subdomains. The structural characterisation of N1 revealed that the C-terminal region exhibits a propensity to form secondary structure. The presence of the independently folded N2 subdomain could potentially stabilise these transient conformations, enabling binding to endothelial cell surface receptors. In the absence of N2, the increased entropic cost of stabilising N1 is not overcome by binding, and so the interaction is not permitted. Conducting similar experiments with recombinant N1N2, or N2N3 with an extended N-terminus, could investigate this hypothesis. Another important consideration is the apparent proteolysis N1 undergoes when subjected to plasma, suggesting FnBPA would not survive *S. aureus* bacteraemia intact and potentially inhibiting fA mediated endothelial cell adherence. Therefore, it is important to identify the cleavage site within the A domain and the minimum region required to adhere to endothelial cells before the physiological relevance can be concluded.

Despite the aforementioned caveats, the ability of the A domain to mediate adherence to endothelial cells has potentially far reaching consequences. As an opportunistic pathogen, the capacity of *S. aureus* to colonise different environments is an important facet of its virulence. The ligand binding redundancy between FnBPA and other *S. aureus* surface proteins is well documented (Deivanayagam et al., 2002, Ganesh et al., 2011b, Keane et al., 2007a, Wann et al., 2000). However, interplay between surface proteins able to mediate host colonisation is poorly understood. Expression of multiple proteins that bind to different tissues would be advantageous during the spread of infections. ClfB can mediate nasal colonisation (O'Brien et al., 2002b), however there is no evidence to suggest ClfB can adhere to host vasculature. Therefore, FnBPA, able to mediate adherence through either the A domain or FnBRs, might be employed by *S. aureus* to colonise different tissues and establish an infection. Further investigation is required to confirm the viability of A domain-mediated adherence *in vivo*, although it is the first evidence suggesting a function for the N1 subdomain.

7.3 FnBPA-mediated Biofilms

An important facet of *S. aureus* virulence is its tendency to form biofilms on implanted medical devices. Growth in a biofilm provides bacteria with an increased resistance to the host immune response and antimicrobial agents and as a result infections are more difficult to treat. The molecular mechanism of protein-mediated biofilm formation is poorly understood, particularly for FnBP-mediated biofilms. A study found that the A domain of FnBPA is necessary and sufficient for biofilm formation, though no mechanistic details of its role were investigated. Cell accumulation in other protein-mediated biofilms is thought to result from Zn²⁺-induced dimerisation of protein molecules on neighbouring cells (Conrady et al., 2008, Geoghegan et al., 2010). While the conditions necessary for FnBP-mediated biofilms differ from those of other protein-mediated biofilms, requiring mildly acidic condition resulting from glucose metabolism, Zn²⁺ is essential for biofilm integrity. Therefore, it was suggested FnBPA might also utilise a ‘zinc-zipper’ mechanism to promote cell aggregation.

The full A domain of FnBPA did dimerise under the conditions known to induce biofilm formation; at pH 5.5 and in the presence of Zn²⁺. Smaller fragments of the A domain were always monomeric under the conditions tested suggesting an extended interface between protein molecules is required if FnBPA utilises the ‘zinc-zipper’ mechanism. A caveat to describing FnBP-mediated biofilm formation using the zinc-zipper model is the concentration of Zn²⁺ required to induce dimerisation. A 500-fold molar excess of Zn²⁺ was necessary for dimers to form, far exceeding the protein to zinc ratio thought to be achievable *in vivo* where free zinc concentration is in the low femtomolar range. Therefore, it seems more likely a different interaction facilitates cell accumulation in FnBP-mediated biofilms.

Extracellular DNA (eDNA) and teichoic acids (eTA) are major components of the biofilm matrix, along with proteins and PIA. The relative concentration of each component varies depending on the type of biofilm formed. Although their exact role is not known, both eDNA and eTA are ubiquitous in protein-mediated biofilms with eDNA known to be key to structural integrity in the early stages on biofilm development and able to mediate primary attachment. The coordination of metal ions

plays an important role in a number of DNA/protein interactions (Nowotny et al., 2005, Beese and Steitz, 1991, Schreiter et al., 2006). Typically, the carboxyl groups of glutamate and aspartate residues and phosphate groups of DNA coordinate metal ions to establish the interaction. The N1 subdomain of FnBPA contains 26 aspartate and glutamate residues that could mediate an interaction with DNA by Zn^{2+} coordination. Thus, the ability of FnBPA to interact with DNA, and the potential involvement of Zn^{2+} was investigated. EMSAs and NMR spectroscopy indicate that there is no interaction between N1 and DNA and that the addition of Zn^{2+} has no effect. However, Izano *et al.* found that eDNA fragments greater than 10 kb in length were necessary for biofilm viability (Izano et al., 2008a). The DNA fragments used in this study were approximately 6 kb in length and may not be able to form stable and interactions with N1. The sensitivity of NMR enables weak interactions to be detected and it seems unlikely the even suboptimal interface formation would not be detected.

The role of eTAs in protein-mediated biofilm formation is not yet understood, despite them being ever-present. In the absence of other interactions that might stabilise FnBP-mediated biofilms, it is possible FnBPA and teichoic acids interact to drive cell accumulation. Wall teichoic acids (WTA), found to be structurally identical to eTA (Sadovskaya et al., 2006), were isolated and binding to N1 investigated by NMR spectroscopy. The addition of an increasing concentration of WTAs to ^{15}N labelled N1 resulted in residue specific chemical shift perturbations. These chemical shift changes were unidirectional suggesting residues involved in WTA binding were affected in the same way. Typically such chemical shift changes are associated with non-specific interactions or changes in temperature and pH (Baxter and Williamson, 1997). All spectra were acquired at 298 K with the temperature tightly controlled throughout the titration. Following each addition of WTA, the pH was also measured and corrected as required to pH 5.5. Therefore, it seems unlikely either of these variables would account for the observed chemical shift changes. However, WTA are known scavengers of divalent metal cations, and despite extensive dialysis, an increased ion concentration may have resulted from the addition of WTA to the N1 sample, accounting for the chemical shift perturbations. The observed line broadening of a number of resonances could also be attributed to ionic effects (Lambert et al., 2005).

The similarity of the chemical shift changes observed in the N1 NMR spectra may also reflect how N1 interacts with WTAs. Electrostatic interactions can result in systematic chemical shift perturbations (Live et al., 1984, Wagner et al., 1983). Therefore, the chemical shift changes in the N1 spectrum could result from an extensive electrostatic interface between N1 and WTAs. The polymeric nature of WTAs may result in the same region of the repeat unit interacting with N1 backbone amide groups, inducing the similar chemical shift perturbations observed for affected residues.

The addition of Zn^{2+} resulted the formation of a precipitate containing N1 (data not shown). Zn^{2+} has no effect on the solubility of either N1 or WTAs when in isolation, suggesting Zn^{2+} enhances the interaction between N1 and WTAs resulting in the observed precipitate. Although the Zn^{2+} binding capacity of WTAs is unknown, binding to other divalent metal cations, such as Mg^{2+} and Cd^{2+} , has been well characterised (Halye and Rice, 2010, Wickham et al., 2009). However, Zn^{2+} is an essential component of the FnBP-mediated biofilm matrix and its role may be to stabilise the interface between FnBPA and eTAs or WTAs.

To confirm the validity of the N1:WTA interaction further spectroscopic methods could be employed. NOEs have been employed to distinguish between non-specific interactions and specific structural rearrangements the result from interactions with metal ions (Riek et al., 1999), and a similar approach could be applied to N1. Changes to carbon chemical shifts could also be monitored on the addition of WTA to determine if perturbations attributable to specific or non-specific interactions are observed, with uniform and unidirectional movement indicative of the latter. A different technique less sensitive to ionic effects, such as surface plasmon resonance or mobility shifts assays could also be employed to confirm the legitimacy of the interaction. Implementation of either of these techniques would also be useful in determining the involvement of other regions of the A domain outside the scope of this work.

7.4 Role of Disorder in N1 Function

In many instances identification of secondary structure propensity in IDPs is indicative of regions involved in ligand binding interactions. The SSP analysis of N1 revealed the C-terminal region exhibits a propensity to form elements of secondary structure. Both the identified interactions involving N1 indicate this region is necessary for binding. In each case a stabilisation of these transient conformations may result in the formation of a stable interface, enabling FnBPA to fulfil the required role. A recent study has suggested that a truncated FnBPA construct lacking approximately the first 100 residues of N1 is the minimal region required for stable surface expression and to mediate biofilm formation (Geoghegan *et al.*, submitted manuscript). Coincidentally, a similarly truncated variant of FnBPA is produced by the action of thrombin, a protease present in plasma. These findings suggest that the regions of N1 adopting transient secondary structure, identified through the structural characterisation and thought to be key in the interactions described here, are physiologically relevant and are likely to be crucial for FnBPA's function.

The entropic penalty of stabilising disordered proteins is offset by a high enthalpic contribution (Dyson and Wright, 2001). As a result interactions involving IDPs are highly specific but low affinity. Recognition of multiple ligands is a well-documented facet of IDP function and coupled with the aforementioned thermodynamic characteristics enables IDPs to fulfil a variety of roles (Kim *et al.*, 2008). For example, recognition of multiple kinases enables regulation of a number of aspects of the cell cycle by kinase inhibitor p21^{Cip1} (Kriwacki *et al.*, 1996). Despite this well-studied binding promiscuity, there are no examples of IDPs fulfilling two such distinct roles as those proposed for N1 within the FnBPA A domain; adhesion to endothelial cells and binding to WTAs. The supposition that in both cases ligand interactions stabilise local folds is a phenomenon observed in many systems, from protein:protein interactions (Schwarz-Linek *et al.*, 2003) to protein:DNA interactions (Spolar and Record, 1994). To confirm the feasibility that N1 could contribute to separate processes requires further investigation, particularly in deducing the role played by the rest of the A domain.

References

- AGBACK, P., MALTSEVA, T. V., YAMAKAGE, S. I., NILSON, F. P., FOLDESI, A. & CHATTOPADHYAYA, J. 1994. The differences in the T2 relaxation rates of the protons in the partially-deuteriated and fully protonated sugar residues in a large oligo-DNA ('NMR-window') gives complementary structural information. *Nucleic acids research*, 22, 1404-12.
- ANWAR, H., DASGUPTA, M. K. & COSTERTON, J. W. 1990. Testing the susceptibility of bacteria in biofilms to antibacterial agents. *Antimicrobial agents and chemotherapy*, 34, 2043-6.
- ARRECUBIETA, C., ASAI, T., BAYERN, M., LOUGHMAN, A., FITZGERALD, J. R., SHELTON, C. E., BARON, H. M., DANG, N. C., DENG, M. C., NAKA, Y., FOSTER, T. J. & LOWY, F. D. 2006. The role of *Staphylococcus aureus* adhesins in the pathogenesis of ventricular assist device-related infections. *Journal of Infectious Diseases*, 193, 1109-19.
- ATKIN, K. E., BRETNALL, A. S., HARRIS, G., BINGHAM, R. J., ERAT, M. C., MILLARD, C. J., SCHWARZ-LINEK, U., STAUNTON, D., VAKONAKIS, I., CAMPBELL, I. D. & POTTS, J. R. 2010. The streptococcal binding site in the gelatin-binding domain of fibronectin is consistent with a non-linear arrangement of modules. *J Biol Chem*, 285, 36977-83.
- ATREYA, H. S. & CHARY, K. V. 2001. Selective 'unlabeling' of amino acids in fractionally ¹³C labeled proteins: an approach for stereospecific NMR assignments of CH₃ groups in Val and Leu residues. *Journal of biomolecular NMR*, 19, 267-72.
- BABA, T., TAKEUCHI, F., KURODA, M., YUZAWA, H., AOKI, K., OGUCHI, A., NAGAI, Y., IWAMA, N., ASANO, K., NAIMI, T., KURODA, H., CUI, L., YAMAMOTO, K. & HIRAMATSU, K. 2002. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet*, 359, 1819-27.
- BAI, Y., MILNE, J. S., MAYNE, L. & ENGLANDER, S. W. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins*, 17, 75-86.
- BANDYK, D. F. 2008. Vascular surgical site infection: risk factors and preventive measures. *Semin Vasc Surg*, 21, 119-23.
- BAXTER, N. J. & WILLIAMSON, M. P. 1997. Temperature dependence of ¹H chemical shifts in proteins. *Journal of biomolecular NMR*, 9, 359-69.
- BEENKEN, K. E., DUNMAN, P. M., MCALEESE, F., MACAPAGAL, D., MURPHY, E., PROJAN, S. J., BLEVINS, J. S. & SMELTZER, M. S. 2004. Global gene expression in *Staphylococcus aureus* biofilms. *Journal of Bacteriology*, 186, 4665-84.
- BEESE, L. S. & STEITZ, T. A. 1991. Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *The EMBO journal*, 10, 25-33.
- BERMEL, W., FELLI, I. C., KUMMERLE, R. & PIERATTELLI, R. 2008. (¹³C) direct-detection biomolecular NMR. *Concepts in Magnetic Resonance Part A*, 32A, 183-200.
- BINGHAM, R. J., RUDINO-PINERA, E., MEENAN, N. A., SCHWARZ-LINEK, U., TURKENBURG, J. P., HOOK, M., GARMAN, E. F. & POTTS, J. R. 2008. Crystal structures of fibronectin-binding sites from *Staphylococcus*

- aureus FnBPA in complex with fibronectin domains. *Proc Natl Acad Sci U S A*, 105, 12254-8.
- BORDOLI, L., KIEFER, F. & SCHWEDE, T. 2007. Assessment of disorder predictions in CASP7. *Proteins*, 69 Suppl 8, 129-36.
- BOZZI, M., BIANCHI, M., SCIANDRA, F., PACI, M., GIARDINA, B., BRANCACCIO, A. & CICERO, D. O. 2003. Structural characterization by NMR of the natively unfolded extracellular domain of beta-dystroglycan: toward the identification of the binding epitope for alpha-dystroglycan. *Biochemistry*, 42, 13717-24.
- BRYAN, C. S., KIRKHART, B. & BRENNER, E. R. 1984. Staphylococcal bacteremia: current patterns in nonuniversity hospitals. *Southern medical journal*, 77, 693-6.
- BURKE, F. M., DI POTO, A., SPEZIALE, P. & FOSTER, T. J. 2011. The A domain of fibronectin binding protein B of *Staphylococcus aureus* contains a novel fibronectin binding site. *FEBS J*.
- BURKE, F. M., MCCORMACK, N., RINDI, S., SPEZIALE, P. & FOSTER, T. J. 2010. Fibronectin-binding protein B variation in *Staphylococcus aureus*. *BMC Microbiol*, 10, 160.
- CAVANAGH, J. 1996. *Protein NMR spectroscopy : principles and practice*, San Diego, Academic Press.
- CHA, H. Y., KIM, H. O., JIN, J. S. & LEE, J. C. 2010. Emergence of vancomycin-intermediate *Staphylococcus aureus* from predominant methicillin-resistant *S. aureus* clones in a Korean hospital. *Journal of microbiology*, 48, 533-5.
- CHEN, Y. H., YANG, J. T. & CHAU, K. H. 1974. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. *Biochemistry*, 13, 3350-9.
- CHESNEY, P. J., BERGDOLL, M. S., DAVIS, J. P. & VERGERONT, J. M. 1984. The disease spectrum, epidemiology, and etiology of toxic-shock syndrome. *Annual review of microbiology*, 38, 315-38.
- CHO, M. K., XIANG, S., KIM, H. Y., BECKER, S. & ZWECKSTETTER, M. 2012. Cold-induced changes in the protein ubiquitin. *PLoS One*, 7, e37270.
- CLARKE, S. R., WILTSHIRE, M. D. & FOSTER, S. J. 2004. IsdA of *Staphylococcus aureus* is a broad spectrum, iron-regulated adhesin. *Molecular microbiology*, 51, 1509-19.
- CLARO, T., WIDAA, A., O'SEAGHDHA, M., MIAJLOVIC, H., FOSTER, T. J., O'BRIEN, F. J. & KERRIGAN, S. W. 2011. *Staphylococcus aureus* protein A binds to osteoblasts and triggers signals that weaken bone in osteomyelitis. *PLoS One*, 6, e18748.
- CLEMENT, S., VAUDAUX, P., FRANCOIS, P., SCHRENZEL, J., HUGGLER, E., KAMPF, S., CHAPONNIER, C., LEW, D. & LACROIX, J. S. 2005. Evidence of an intracellular reservoir in the nasal mucosa of patients with recurrent *Staphylococcus aureus* rhinosinusitis. *The Journal of infectious diseases*, 192, 1023-8.
- COHEN, S. N., CHANG, A. C. & HSU, L. 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 69, 2110-4.
- COLE, A. M., TAHK, S., OREN, A., YOSHIOKA, D., KIM, Y. H., PARK, A. & GANZ, T. 2001. Determinants of *Staphylococcus aureus* nasal carriage. *Clinical and diagnostic laboratory immunology*, 8, 1064-9.

- COLE, C., BARBER, J. D. & BARTON, G. J. 2008. The Jpred 3 secondary structure prediction server. *Nucleic acids research*, 36, W197-201.
- COLLINS, L. V., KRISTIAN, S. A., WEIDENMAIER, C., FAIGLE, M., VAN KESSEL, K. P., VAN STRIJP, J. A., GOTZ, F., NEUMEISTER, B. & PESCHEL, A. 2002. Staphylococcus aureus strains lacking D-alanine modifications of teichoic acids are highly susceptible to human neutrophil killing and are virulence attenuated in mice. *The Journal of infectious diseases*, 186, 214-9.
- CONRADY, D. G., BRESCIA, C. C., HORII, K., WEISS, A. A., HASSETT, D. J. & HERR, A. B. 2008. A zinc-dependent adhesion module is responsible for intercellular adhesion in staphylococcal biofilms. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 19456-61.
- CORRIGAN, R. M., MIAJLOVIC, H. & FOSTER, T. J. 2009. Surface proteins that promote adherence of Staphylococcus aureus to human desquamated nasal epithelial cells. *BMC microbiology*, 9, 22.
- CORRIGAN, R. M., RIGBY, D., HANDLEY, P. & FOSTER, T. J. 2007. The role of Staphylococcus aureus surface protein SasG in adherence and biofilm formation. *Microbiology*, 153, 2435-46.
- COSTERTON, J. W., LEWANDOWSKI, Z., CALDWELL, D. E., KORBER, D. R. & LAPPIN-SCOTT, H. M. 1995. Microbial biofilms. *Annual review of microbiology*, 49, 711-45.
- CRAMTON, S. E., GERKE, C., SCHNELL, N. F., NICHOLS, W. W. & GOTZ, F. 1999. The intercellular adhesion (ica) locus is present in Staphylococcus aureus and is required for biofilm formation. *Infection and immunity*, 67, 5427-33.
- CRESPI, H. L. & KATZ, J. J. 1969. High resolution proton magnetic resonance studies of fully deuterated and isotope hybrid proteins. *Nature*, 224, 560-2.
- CSIZMOK, V., FELLI, I. C., TOMPA, P., BANCI, L. & BERTINI, I. 2008. Structural and dynamic characterization of intrinsically disordered human securin by NMR spectroscopy. *Journal of the American Chemical Society*, 130, 16873-9.
- CUCARELLA, C., SOLANO, C., VALLE, J., AMORENA, B., LASA, I. & PENADES, J. R. 2001. Bap, a Staphylococcus aureus surface protein involved in biofilm formation. *Journal of Bacteriology*, 183, 2888-96.
- CUFF, J. A. & BARTON, G. J. 2000. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins*, 40, 502-11.
- DAROUCHE, R. O. 2004. Treatment of infections associated with surgical implants. *The New England journal of medicine*, 350, 1422-9.
- DE LENCASTRE, H., SA FIGUEIREDO, A. M., URBAN, C., RAHAL, J. & TOMASZ, A. 1991. Multiple mechanisms of methicillin resistance and improved methods for detection in clinical isolates of Staphylococcus aureus. *Antimicrobial agents and chemotherapy*, 35, 632-9.
- DE LISSOVOY, G., FRAEMAN, K., HUTCHINS, V., MURPHY, D., SONG, D. & VAUGHN, B. B. 2009. Surgical site infection: incidence and impact on hospital utilization and treatment costs. *American journal of infection control*, 37, 387-97.

- DEDENT, A., BAE, T., MISSIAKAS, D. M. & SCHNEEWIND, O. 2008. Signal peptides direct surface proteins to two distinct envelope locations of *Staphylococcus aureus*. *The EMBO journal*, 27, 2656-68.
- DEIVANAYAGAM, C. C., WANN, E. R., CHEN, W., CARSON, M., RAJASHANKAR, K. R., HOOK, M. & NARAYANA, S. V. 2002. A novel variant of the immunoglobulin fold in surface adhesins of *Staphylococcus aureus*: crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *The EMBO journal*, 21, 6660-72.
- DELAGLIO, F., GRZESIEK, S., VUISTER, G. W., ZHU, G., PFEIFER, J. & BAX, A. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR*, 6, 277-93.
- DELEO, F. R., OTTO, M., KREISWIRTH, B. N. & CHAMBERS, H. F. 2010. Community-associated meticillin-resistant *Staphylococcus aureus*. *Lancet*, 375, 1557-68.
- DIEKEMA, D. J., PFALLER, M. A., SCHMITZ, F. J., SMAYEVSKY, J., BELL, J., JONES, R. N. & BEACH, M. 2001. Survey of infections due to *Staphylococcus* species: frequency of occurrence and antimicrobial susceptibility of isolates collected in the United States, Canada, Latin America, Europe, and the Western Pacific region for the SENTRY Antimicrobial Surveillance Program, 1997-1999. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 32 Suppl 2, S114-32.
- DONLAN, R. M. 2001. Biofilm formation: a clinically relevant microbiological process. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 33, 1387-92.
- DOSZTANYI, Z., CSIZMOK, V., TOMPA, P. & SIMON, I. 2005a. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*, 21, 3433-4.
- DOSZTANYI, Z., CSIZMOK, V., TOMPA, P. & SIMON, I. 2005b. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *Journal of molecular biology*, 347, 827-39.
- DUGUID, I. G., EVANS, E., BROWN, M. R. & GILBERT, P. 1992. Effect of biofilm culture upon the susceptibility of *Staphylococcus epidermidis* to tobramycin. *The Journal of antimicrobial chemotherapy*, 30, 803-10.
- DUNKER, A. K., BROWN, C. J., LAWSON, J. D., IAKOUCHEVA, L. M. & OBRADOVIC, Z. 2002. Intrinsic disorder and protein function. *Biochemistry*, 41, 6573-82.
- DYSON, H. J. & WRIGHT, P. E. 2001. Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Methods in enzymology*, 339, 258-70.
- DYSON, H. J. & WRIGHT, P. E. 2002. Coupling of folding and binding for unstructured proteins. *Current opinion in structural biology*, 12, 54-60.
- DYSON, H. J. & WRIGHT, P. E. 2004. Unfolded proteins and protein folding studied by NMR. *Chemical Reviews*, 104, 3607-3622.
- DYSON, H. J. & WRIGHT, P. E. 2005. Intrinsically unstructured proteins and their functions. *Nature reviews. Molecular cell biology*, 6, 197-208.
- EDWARDS, A. M., BOWDEN, M. G., BROWN, E. L., LAABEI, M. & MASSEY, R. C. 2012. *Staphylococcus aureus* Extracellular Adherence Protein Triggers

- TNFalpha Release, Promoting Attachment to Endothelial Cells via Protein A. *PLoS One*, 7, e43046.
- EDWARDS, A. M., POTTER, U., MEENAN, N. A., POTTS, J. R. & MASSEY, R. C. 2011. Staphylococcus aureus Keratinocyte Invasion Is Dependent upon Multiple High-Affinity Fibronectin-Binding Repeats within FnBPA. *PLoS One*, 6, e18899.
- EDWARDS, A. M., POTTS, J. R., JOSEFSSON, E. & MASSEY, R. C. 2010. Staphylococcus aureus host cell invasion and virulence in sepsis is facilitated by the multiple repeats within FnBPA. *PLoS Pathog*, 6, e1000964.
- ENDL, J., SEIDL, H. P., FIEDLER, F. & SCHLEIFER, K. H. 1983. Chemical composition and structure of cell wall teichoic acids of staphylococci. *Archives of microbiology*, 135, 215-23.
- ENGLANDER, S. W., SOSNICK, T. R., ENGLANDER, J. J. & MAYNE, L. 1996. Mechanisms and uses of hydrogen exchange. *Current opinion in structural biology*, 6, 18-23.
- EUGSTER, M. R. & LOESSNER, M. J. 2011. Rapid analysis of Listeria monocytogenes cell wall teichoic acid carbohydrates by ESI-MS/MS. *PLoS One*, 6, e21500.
- FABRINI, R., DE LUCA, A., STELLA, L., MEI, G., ORIONI, B., CICCONE, S., FEDERICI, G., LO BELLO, M. & RICCI, G. 2009. Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. *Biochemistry*, 48, 10473-82.
- FARLEY, J. E., ROSS, T., KRALL, J., HAYAT, M., CASTON-GAA, A., PERL, T. & CARROLL, K. C. 2012. Prevalence, risk factors, and molecular epidemiology of methicillin-resistant Staphylococcus aureus nasal and axillary colonization among psychiatric patients on admission to an academic medical center. *American journal of infection control*.
- FEY, P. D. & OLSON, M. E. 2010. Current concepts in biofilm formation of Staphylococcus epidermidis. *Future microbiology*, 5, 917-33.
- FISCHER, W. 1994. Lipoteichoic acid and lipids in the membrane of Staphylococcus aureus. *Medical microbiology and immunology*, 183, 61-76.
- FITZGERALD, J. R., LOUGHMAN, A., KEANE, F., BRENNAN, M., KNOBEL, M., HIGGINS, J., VISAI, L., SPEZIALE, P., COX, D. & FOSTER, T. J. 2006. Fibronectin-binding proteins of Staphylococcus aureus mediate activation of human platelets via fibrinogen and fibronectin bridges to integrin GPIIb/IIIa and IgG binding to the FcgammaRIIa receptor. *Mol Microbiol*, 59, 212-30.
- FITZPATRICK, F., HUMPHREYS, H. & O'GARA, J. P. 2005. Evidence for icaADBC-independent biofilm development mechanism in methicillin-resistant Staphylococcus aureus clinical isolates. *Journal of clinical microbiology*, 43, 1973-6.
- FLEMMING, H. C. & WINGENDER, J. 2010. The biofilm matrix. *Nature reviews. Microbiology*, 8, 623-33.
- FOGH, R. H., BOUCHER, W., VRANKEN, W. F., PAJON, A., STEVENS, T. J., BHAT, T. N., WESTBROOK, J., IONIDES, J. M. & LAUE, E. D. 2005. A framework for scientific data modeling and automated software development. *Bioinformatics*, 21, 1678-84.
- FOSTER, T. J. 2004. The Staphylococcus aureus "superbug". *Journal of Clinical Investigation*, 114, 1693-1696.

- FROMAN, G., SWITALSKI, L. M., SPEZIALE, P. & HOOK, M. 1987. Isolation and characterization of a fibronectin receptor from *Staphylococcus aureus*. *The Journal of biological chemistry*, 262, 6564-71.
- GANESH, V. K., BARBU, E. M., DEIVANAYAGAM, C. C., LE, B., ANDERSON, A. S., MATSUKA, Y. V., LIN, S. L., FOSTER, T. J., NARAYANA, S. V. & HOOK, M. 2011a. Structural and biochemical characterization of *Staphylococcus aureus* clumping factor B/ligand interactions. *The Journal of biological chemistry*, 286, 25963-72.
- GANESH, V. K., BARBU, E. M., DEIVANAYAGAM, C. C., LE, B., ANDERSON, A. S., MATSUKA, Y. V., LIN, S. L., FOSTER, T. J., NARAYANA, S. V. & HOOK, M. 2011b. Structural and biochemical characterization of *Staphylococcus aureus* Clumping Factor B:ligand interactions. *J Biol Chem*.
- GARDNER, K. H. & KAY, L. E. 1998. The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annual review of biophysics and biomolecular structure*, 27, 357-406.
- GARNER, E., CANNON, P., ROMERO, P., OBRADOVIC, Z. & DUNKER, A. K. 1998. Predicting Disordered Regions from Amino Acid Sequence: Common Themes Despite Differing Structural Characterization. *Genome informatics. Workshop on Genome Informatics*, 9, 201-213.
- GASTEIGER, E., GATTIKER, A., HOOGLAND, C., IVANYI, I., APPEL, R. D. & BAIROCH, A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*, 31, 3784-8.
- GEOGHEGAN, J. A., CORRIGAN, R. M., GRUSZKA, D. T., SPEZIALE, P., O'GARA, J. P., POTTS, J. R. & FOSTER, T. J. 2010. Role of surface protein SasG in biofilm formation by *Staphylococcus aureus*. *Journal of Bacteriology*, 192, 5663-73.
- GERKE, C., KRAFT, A., SUSSMUTH, R., SCHWEITZER, O. & GOTZ, F. 1998. Characterization of the N-acetylglucosaminyltransferase activity involved in the biosynthesis of the *Staphylococcus epidermidis* polysaccharide intercellular adhesin. *The Journal of biological chemistry*, 273, 18586-93.
- GOTO, N. K., GARDNER, K. H., MUELLER, G. A., WILLIS, R. C. & KAY, L. E. 1999. A robust and cost-effective method for the production of Val, Leu, Ile (δ 1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. *Journal of biomolecular NMR*, 13, 369-74.
- GOTO, N. K. & KAY, L. E. 2000. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Current opinion in structural biology*, 10, 585-92.
- GRUNDLING, A. & SCHNEEWIND, O. 2007. Synthesis of glycerol phosphate lipoteichoic acid in *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 8478-83.
- GRUSZKA, D. T., WOJDYLA, J. A., BINGHAM, R. J., TURKENBURG, J. P., MANFIELD, I. W., STEWARD, A., LEECH, A. P., GEOGHEGAN, J. A., FOSTER, T. J., CLARKE, J. & POTTS, J. R. 2012. Staphylococcal biofilm-forming protein has a contiguous rod-like structure. *Proceedings of the National Academy of Sciences of the United States of America*, 109, E1011-8.
- HALYE, J. L. & RICE, C. V. 2010. Cadmium chelation by bacterial teichoic acid from solid-state nuclear magnetic resonance spectroscopy. *Biomacromolecules*, 11, 333-40.

- HARRINGTON, C. R. & BADDILEY, J. 1985. Biosynthesis of wall teichoic acids in *Staphylococcus aureus* H, *Micrococcus varians* and *Bacillus subtilis* W23. Involvement of lipid intermediates containing the disaccharide N-acetylmannosaminyl N-acetylglucosamine. *European journal of biochemistry / FEBS*, 153, 639-45.
- HEILMANN, C., HUSSAIN, M., PETERS, G. & GOTZ, F. 1997. Evidence for autolysin-mediated primary attachment of *Staphylococcus epidermidis* to a polystyrene surface. *Molecular microbiology*, 24, 1013-24.
- HEILMANN, C., SCHWEITZER, O., GERKE, C., VANITTANAKOM, N., MACK, D. & GOTZ, F. 1996. Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis*. *Molecular microbiology*, 20, 1083-91.
- HERRICK, S., BLANC-BRUDE, O., GRAY, A. & LAURENT, G. 1999. Fibrinogen. *The international journal of biochemistry & cell biology*, 31, 741-6.
- HERSH, A. L., CHAMBERS, H. F., MASELLI, J. H. & GONZALES, R. 2008. National trends in ambulatory visits and antibiotic prescribing for skin and soft-tissue infections. *Archives of internal medicine*, 168, 1585-91.
- HIRAMATSU, K., HANAKI, H., INO, T., YABUTA, K., OGURI, T. & TENOVER, F. C. 1997. Methicillin-resistant *Staphylococcus aureus* clinical strain with reduced vancomycin susceptibility. *The Journal of antimicrobial chemotherapy*, 40, 135-6.
- HORE, P. J. 1995. *Nuclear magnetic resonance*, Oxford ; New York, Oxford University Press.
- HOUSTON, P., ROWE, S. E., POZZI, C., WATERS, E. M. & O'GARA, J. P. 2011. Essential role for the major autolysin in the fibronectin-binding protein-mediated *Staphylococcus aureus* biofilm phenotype. *Infection and immunity*, 79, 1153-65.
- HOYLE, B. D., ALCANTARA, J. & COSTERTON, J. W. 1992. *Pseudomonas aeruginosa* biofilm as a diffusion barrier to piperacillin. *Antimicrobial agents and chemotherapy*, 36, 2054-6.
- HSU, C. Y., LIN, M. H., CHEN, C. C., CHIEN, S. C., CHENG, Y. H., SU, I. N. & SHU, J. C. 2011. Vancomycin promotes the bacterial autolysis, release of extracellular DNA, and biofilm formation in vancomycin-non-susceptible *Staphylococcus aureus*. *FEMS immunology and medical microbiology*, 63, 236-47.
- IKURA, M., KAY, L. E. & BAX, A. 1990. A novel approach for sequential assignment of ¹H, ¹³C, and ¹⁵N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry*, 29, 4659-67.
- ISHIDA, T. & KINOSHITA, K. 2008. Prediction of disordered regions in proteins based on the meta approach. *Bioinformatics*, 24, 1344-8.
- IWASAKI, Y. & AKIYOSHI, K. 2006. Synthesis and characterization of amphiphilic polyphosphates with hydrophilic graft chains and cholesteryl groups as nanocarriers. *Biomacromolecules*, 7, 1433-8.
- IWASAKI, Y., NAKAGAWA, C., OHTOMI, M., ISHIHARA, K. & AKIYOSHI, K. 2004. Novel biodegradable polyphosphate cross-linker for making biocompatible hydrogel. *Biomacromolecules*, 5, 1110-5.
- IZANO, E. A., AMARANTE, M. A., KHER, W. B. & KAPLAN, J. B. 2008a. Differential roles of poly-N-acetylglucosamine surface polysaccharide and

- extracellular DNA in *Staphylococcus aureus* and *Staphylococcus epidermidis* biofilms. *Applied and environmental microbiology*, 74, 470-6.
- IZANO, E. A., SADOVSKAYA, I., WANG, H., VINOGRADOV, E., RAGUNATH, C., RAMASUBBU, N., JABBOURI, S., PERRY, M. B. & KAPLAN, J. B. 2008b. Poly-N-acetylglucosamine mediates biofilm formation and detergent resistance in *Aggregatibacter actinomycetemcomitans*. *Microbial pathogenesis*, 44, 52-60.
- JONES, D. P., CARLSON, J. L., SAMIEC, P. S., STERNBERG, P., JR., MODY, V. C., JR., REED, R. L. & BROWN, L. A. 1998. Glutathione measurement in human plasma. Evaluation of sample collection, storage and derivatization conditions for analysis of dansyl derivatives by HPLC. *Clinica chimica acta; international journal of clinical chemistry*, 275, 175-84.
- JONKERS, D., ELENBAAS, T., TERPORTEN, P., NIEMAN, F. & STOBBERINGH, E. 2003. Prevalence of 90-days postoperative wound infections after cardiac surgery. *Eur J Cardiothorac Surg*, 23, 97-102.
- JONSSON, K., SIGNAS, C., MULLER, H. P. & LINDBERG, M. 1991. Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*. The complete nucleotide sequence and characterization of the second gene. *European journal of biochemistry / FEBS*, 202, 1041-8.
- JOYCE, J. G., ABEYGUNAWARDANA, C., XU, Q., COOK, J. C., HEPLER, R., PRZYSIECKI, C. T., GRIMM, K. M., ROPER, K., IP, C. C., COPE, L., MONTGOMERY, D., CHANG, M., CAMPIE, S., BROWN, M., MCNEELY, T. B., ZORMAN, J., MAIRA-LITRAN, T., PIER, G. B., KELLER, P. M., JANSEN, K. U. & MARK, G. E. 2003. Isolation, structural characterization, and immunological evaluation of a high-molecular-weight exopolysaccharide from *Staphylococcus aureus*. *Carbohydrate research*, 338, 903-22.
- JULANDER, I. 1985. Unfavourable prognostic factors in *Staphylococcus aureus* septicemia and endocarditis. *Scandinavian journal of infectious diseases*, 17, 179-87.
- KAPLAN, J. B. 2011. Antibiotic-induced biofilm formation. *The International journal of artificial organs*, 34, 737-51.
- KAPLAN, J. B., JABBOURI, S. & SADOVSKAYA, I. 2011. Extracellular DNA-dependent biofilm formation by *Staphylococcus epidermidis* RP62A in response to subminimal inhibitory concentrations of antibiotics. *Research in microbiology*, 162, 535-41.
- KEANE, F. M., CLARKE, A. W., FOSTER, T. J. & WEISS, A. S. 2007a. The N-terminal A domain of *Staphylococcus aureus* fibronectin-binding protein A binds to tropoelastin. *Biochemistry*, 46, 7226-32.
- KEANE, F. M., LOUGHMAN, A., VALTULINA, V., BRENNAN, M., SPEZIALE, P. & FOSTER, T. J. 2007b. Fibrinogen and elastin bind to the same region within the A domain of fibronectin binding protein A, an MSCRAMM of *Staphylococcus aureus*. *Mol Microbiol*, 63, 711-23.
- KEELER, J. 2005. *Understanding NMR spectroscopy*, Chichester, England ; Hoboken, NJ, Wiley.
- KIM, P. M., SBONER, A., XIA, Y. & GERSTEIN, M. 2008. The role of disorder in interaction networks: a structural analysis. *Molecular systems biology*, 4, 179.
- KINNIMENT, S. L. & WIMPENNY, J. W. 1992. Measurements of the distribution of adenylate concentrations and adenylate energy charge across *Pseudomonas aeruginosa* biofilms. *Applied and environmental microbiology*, 58, 1629-35.

- KJAERGAARD, M., BRANDER, S. & POULSEN, F. M. 2011. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *Journal of biomolecular NMR*, 49, 139-49.
- KJAERGAARD, M. & POULSEN, F. M. 2011. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of biomolecular NMR*, 50, 157-65.
- KLEVENS, R. M., MORRISON, M. A., NADLE, J., PETIT, S., GERSHMAN, K., RAY, S., HARRISON, L. H., LYNFIELD, R., DUMYATI, G., TOWNES, J. M., CRAIG, A. S., ZELL, E. R., FOSHEIM, G. E., MCDUGAL, L. K., CAREY, R. B. & FRIDKIN, S. K. 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA : the journal of the American Medical Association*, 298, 1763-71.
- KOCH-WESER, J. & SELLERS, E. M. 1976a. Binding of drugs to serum albumin (first of two parts). *The New England journal of medicine*, 294, 311-6.
- KOCH-WESER, J. & SELLERS, E. M. 1976b. Drug therapy. Binding of drugs to serum albumin (second of two parts). *The New England journal of medicine*, 294, 526-31.
- KOGAN, G., SADOVSKAYA, I., CHAIGNON, P., CHOKR, A. & JABBOURI, S. 2006. Biofilms of clinical strains of *Staphylococcus* that do not contain polysaccharide intercellular adhesin. *FEMS microbiology letters*, 255, 11-6.
- KRIWACKI, R. W., HENGST, L., TENNANT, L., REED, S. I. & WRIGHT, P. E. 1996. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 11504-9.
- KROPEC, A., MAIRA-LITRAN, T., JEFFERSON, K. K., GROUT, M., CRAMTON, S. E., GOTZ, F., GOLDMANN, D. A. & PIER, G. B. 2005. Poly-N-acetylglucosamine production in *Staphylococcus aureus* is essential for virulence in murine models of systemic infection. *Infection and immunity*, 73, 6868-76.
- LAMBERT, M. N., HOERTER, J. A., PEREIRA, M. J. & WALTER, N. G. 2005. Solution probing of metal ion binding by helix 27 from *Escherichia coli* 16S rRNA. *RNA*, 11, 1688-700.
- LEE, C. B., SWATEK, K. N. & MCCLURE, B. 2008. Pollen proteins bind to the C-terminal domain of *Nicotiana glauca* arabinogalactan proteins. *The Journal of biological chemistry*, 283, 26965-73.
- LEOPOLD, K. & FISCHER, W. 1992. Heterogeneity of lipoteichoic acid detected by anion exchange chromatography. *Archives of microbiology*, 157, 446-50.
- LEVITT, M. H. 2001. *Spin dynamics : basics of nuclear magnetic resonance*, Chichester ; New York, John Wiley & Sons.
- LIBICH, D. S. & HARAUZ, G. 2008. Backbone dynamics of the 18.5 kDa isoform of myelin basic protein reveals transient alpha-helices and a calmodulin-binding site. *Biophysical journal*, 94, 4847-66.
- LIVE, D. H., DAVIS, D. G., AGOSTA, W. C. & COWBURN, D. 1984. LONG-RANGE HYDROGEN-BOND MEDIATED EFFECTS IN PEPTIDES - N-15 NMR-STUDY OF GRAMICIDIN-S IN WATER AND ORGANIC-SOLVENTS. *Journal of the American Chemical Society*, 106, 1939-1941.
- LOUGHMAN, A., SWEENEY, T., KEANE, F. M., PIETROCOLA, G., SPEZIALE, P. & FOSTER, T. J. 2008. Sequence diversity in the A domain of *Staphylococcus aureus* fibronectin-binding protein A. *BMC Microbiol*, 8, 74.
- LOWY, F. D. 1998. *Staphylococcus aureus* infections. *N Engl J Med*, 339, 520-32.

- MACK, D., FISCHER, W., KROKOTSCH, A., LEOPOLD, K., HARTMANN, R., EGGE, H. & LAUFS, R. 1996a. The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan: purification and structural analysis. *Journal of Bacteriology*, 178, 175-83.
- MACK, D., HAEDER, M., SIEMSEN, N. & LAUFS, R. 1996b. Association of biofilm production of coagulase-negative staphylococci with expression of a specific polysaccharide intercellular adhesin. *The Journal of infectious diseases*, 174, 881-4.
- MACK, D., ROHDE, H., HARRIS, L. G., DAVIES, A. P., HORSTKOTTE, M. A. & KNOBLOCH, J. K. 2006. Biofilm formation in medical device-related infection. *The International journal of artificial organs*, 29, 343-59.
- MANN, E. E., RICE, K. C., BOLES, B. R., ENDRES, J. L., RANJIT, D., CHANDRAMOHAN, L., TSANG, L. H., SMELTZER, M. S., HORSWILL, A. R. & BAYLES, K. W. 2009. Modulation of eDNA release and degradation affects *Staphylococcus aureus* biofilm maturation. *PLoS One*, 4, e5822.
- MANTYLAHTI, S., AITIO, O., HELLMAN, M. & PERMI, P. 2010. HA-detected experiments for the backbone assignment of intrinsically disordered proteins. *Journal of biomolecular NMR*, 47, 171-81.
- MANTYLAHTI, S., HELLMAN, M. & PERMI, P. 2011. Extension of the HA-detection based approach: (HCA)CON(CA)H and (HCA)NCO(CA)H experiments for the main-chain assignment of intrinsically disordered proteins. *Journal of biomolecular NMR*, 49, 99-109.
- MARKLEY, J. L., PUTTER, I. & JARDETZKY, O. 1968. High-resolution nuclear magnetic resonance spectra of selectively deuterated staphylococcal nuclease. *Science*, 161, 1249-51.
- MARKUS, M. A., DAYIE, K. T., MATSUDAIRA, P. & WAGNER, G. 1994. Effect of deuteration on the amide proton relaxation rates in proteins. Heteronuclear NMR experiments on villin 14T. *Journal of magnetic resonance. Series B*, 105, 192-5.
- MARLEY, J., LU, M. & BRACKEN, C. 2001. A method for efficient isotopic labeling of recombinant proteins. *Journal of biomolecular NMR*, 20, 71-5.
- MARSH, J. A. & FORMAN-KAY, J. D. 2010. Sequence determinants of compaction in intrinsically disordered proteins. *Biophysical journal*, 98, 2383-90.
- MARSH, J. A., SINGH, V. K., JIA, Z. & FORMAN-KAY, J. D. 2006. Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein science : a publication of the Protein Society*, 15, 2795-804.
- MARTI, M., TROTONDA, M. P., TORMO-MAS, M. A., VERGARA-IRIGARAY, M., CHEUNG, A. L., LASA, I. & PENADES, J. R. 2010. Extracellular proteases inhibit protein-dependent biofilm formation in *Staphylococcus aureus*. *Microbes and infection / Institut Pasteur*, 12, 55-64.
- MASSEY, R. C., KANTZANOU, M. N., FOWLER, T., DAY, N. P., SCHOFIELD, K., WANN, E. R., BERENDT, A. R., HOOK, M. & PEACOCK, S. J. 2001. Fibronectin-binding protein A of *Staphylococcus aureus* has multiple, substituting, binding regions that mediate adherence to fibronectin and invasion of endothelial cells. *Cell Microbiol*, 3, 839-51.

- MATSUDA, M., YOSHIDA, N., AOKI, N. & WAKABAYASHI, K. 1978. Distribution of cold-insoluble globulin in plasma and tissues. *Annals of the New York Academy of Sciences*, 312, 74-92.
- MAYER, C., SLATER, L., ERAT, M. C., KONRAT, R. & VAKONAKIS, I. 2012. Structural analysis of the Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. *The Journal of biological chemistry*, 287, 7182-9.
- MCDEVITT, D., FRANCOIS, P., VAUDAUX, P. & FOSTER, T. J. 1994. Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus. *Molecular microbiology*, 11, 237-48.
- MEENAN, N. A., VISAI, L., VALTULINA, V., SCHWARZ-LINEK, U., NORRIS, N. C., GURUSIDDAPPA, S., HOOK, M., SPEZIALE, P. & POTTS, J. R. 2007. The tandem beta-zipper model defines high affinity fibronectin-binding repeats within Staphylococcus aureus FnBPA. *J Biol Chem*, 282, 25893-902.
- MESELSON, M. & YUAN, R. 1968. DNA restriction enzyme from E. coli. *Nature*, 217, 1110-4.
- MIDIC, U. & OBRADOVIC, Z. 2012. Intrinsic disorder in putative protein sequences. *Proteome science*, 10 Suppl 1, S19.
- MIELKE, S. P. & KRISHNAN, V. V. 2004. An evaluation of chemical shift index-based secondary structure determination in proteins: influence of random coil chemical shifts. *Journal of biomolecular NMR*, 30, 143-53.
- MITTAG, T. & FORMAN-KAY, J. D. 2007. Atomic-level characterization of disordered protein ensembles. *Current opinion in structural biology*, 17, 3-14.
- MOHAMED, N., VISAI, L., SPEZIALE, P. & ROSS, J. M. 2000. Quantification of Staphylococcus aureus cell surface adhesins using flow cytometry. *Microb Pathog*, 29, 357-61.
- MOTACKOVA, V., NOVACEK, J., ZAWADZKA-KAZIMIERCZUK, A., KAZIMIERCZUK, K., ZIDEK, L., SANDEROVA, H., KRASNY, L., KOZMINSKI, W. & SKLENAR, V. 2010. Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments. *Journal of biomolecular NMR*, 48, 169-77.
- MULLER-SPATH, S., SORANNO, A., HIRSCHFELD, V., HOFMANN, H., RUEGGER, S., REYMOND, L., NETTELS, D. & SCHULER, B. 2010. From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 14609-14.
- MURGA, R., MILLER, J. M. & DONLAN, R. M. 2001. Biofilm formation by gram-negative bacteria on central venous catheter connectors: effect of conditioning films in a laboratory model. *Journal of clinical microbiology*, 39, 2294-7.
- MYLOTTE, J. M., MCDERMOTT, C. & SPOONER, J. A. 1987. Prospective study of 114 consecutive episodes of Staphylococcus aureus bacteremia. *Rev Infect Dis*, 9, 891-907.
- NAIMI, T. S., LEDELL, K. H., COMO-SABETTI, K., BORCHARDT, S. M., BOXRUD, D. J., ETIENNE, J., JOHNSON, S. K., VANDENESCH, F., FRIDKIN, S., O'BOYLE, C., DANILA, R. N. & LYNFIELD, R. 2003. Comparison of community- and health care-associated methicillin-resistant Staphylococcus aureus infection. *JAMA*, 290, 2976-84.

- NELSON, D. L., COX, M. M. & LEHNINGER, A. L. P. O. B. 2008. *Lehninger principles of biochemistry*, New York ; Basingstoke, W.H. Freeman.
- NEUHAUS, F. C. & BADDILEY, J. 2003. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiology and molecular biology reviews : MMBR*, 67, 686-723.
- NICKEL, J. C., RUSESKA, I., WRIGHT, J. B. & COSTERTON, J. W. 1985. Tobramycin resistance of *Pseudomonas aeruginosa* cells growing as a biofilm on urinary catheter material. *Antimicrobial agents and chemotherapy*, 27, 619-24.
- NOBLE, W. C., VALKENBURG, H. A. & WOLTERS, C. H. 1967. Carriage of *Staphylococcus aureus* in random samples of a normal population. *J Hyg (Lond)*, 65, 567-73.
- NORRIS, N. C., BINGHAM, R. J., HARRIS, G., SPEAKMAN, A., JONES, R. P., LEECH, A., TURKENBURG, J. P. & POTTS, J. R. 2011. Structural and functional analysis of the tandem beta-zipper interaction of a Streptococcal protein with human fibronectin. *The Journal of biological chemistry*, 286, 38311-20.
- NOVICK, R. P. 2000. Sortase: the surface protein anchoring transpeptidase and the LPXTG motif. *Trends Microbiol*, 8, 148-51.
- NOWOTNY, M., GAIDAMAKOV, S. A., CROUCH, R. J. & YANG, W. 2005. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell*, 121, 1005-16.
- O'BRIEN, L., KERRIGAN, S. W., KAW, G., HOGAN, M., PENADES, J., LITT, D., FITZGERALD, D. J., FOSTER, T. J. & COX, D. 2002a. Multiple mechanisms for the activation of human platelet aggregation by *Staphylococcus aureus*: roles for the clumping factors ClfA and ClfB, the serine-aspartate repeat protein SdrE and protein A. *Molecular microbiology*, 44, 1033-44.
- O'BRIEN, L. M., WALSH, E. J., MASSEY, R. C., PEACOCK, S. J. & FOSTER, T. J. 2002b. *Staphylococcus aureus* clumping factor B (ClfB) promotes adherence to human type I cytokeratin 10: implications for nasal colonization. *Cellular microbiology*, 4, 759-70.
- O'GARA, J. P. 2007. *ica* and beyond: biofilm mechanisms and regulation in *Staphylococcus epidermidis* and *Staphylococcus aureus*. *FEMS microbiology letters*, 270, 179-88.
- O'NEILL, A. J. 2010. *Staphylococcus aureus* SH1000 and 8325-4: comparative genome sequences of key laboratory strains in staphylococcal research. *Letters in applied microbiology*, 51, 358-61.
- O'NEILL, E., HUMPHREYS, H. & O'GARA, J. P. 2009. Carriage of both the *fnbA* and *fnbB* genes and growth at 37 degrees C promote FnBP-mediated biofilm development in methicillin-resistant *Staphylococcus aureus* clinical isolates. *Journal of medical microbiology*, 58, 399-402.
- O'NEILL, E., POZZI, C., HOUSTON, P., HUMPHREYS, H., ROBINSON, D. A., LOUGHMAN, A., FOSTER, T. J. & O'GARA, J. P. 2008. A novel *Staphylococcus aureus* biofilm phenotype mediated by the fibronectin-binding proteins, FnBPA and FnBPB. *Journal of Bacteriology*, 190, 3835-50.
- O'NEILL, E., POZZI, C., HOUSTON, P., SMYTH, D., HUMPHREYS, H., ROBINSON, D. A. & O'GARA, J. P. 2007. Association between methicillin susceptibility and biofilm regulation in *Staphylococcus aureus* isolates from device-related infections. *Journal of clinical microbiology*, 45, 1379-88.

- OLDFIELD, C. J., CHENG, Y., CORTESE, M. S., BROWN, C. J., UVERSKY, V. N. & DUNKER, A. K. 2005. Comparing and combining predictors of mostly disordered proteins. *Biochemistry*, 44, 1989-2000.
- OTTO, M. 2004. Virulence factors of the coagulase-negative staphylococci. *Frontiers in bioscience : a journal and virtual library*, 9, 841-63.
- OTTO, M. 2008. Staphylococcal biofilms. *Current topics in microbiology and immunology*, 322, 207-28.
- OTWINOWSKI, Z., SCHEVITZ, R. W., ZHANG, R. G., LAWSON, C. L., JOACHIMIAK, A., MARMORSTEIN, R. Q., LUISI, B. F. & SIGLER, P. B. 1988. Crystal structure of trp repressor/operator complex at atomic resolution. *Nature*, 335, 321-9.
- OUTTEN, C. E. & O'HALLORAN, T. V. 2001. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science*, 292, 2488-92.
- OWEN, W. F., JR., LEW, N. L., LIU, Y., LOWRIE, E. G. & LAZARUS, J. M. 1993. The urea reduction ratio and serum albumin concentration as predictors of mortality in patients undergoing hemodialysis. *The New England journal of medicine*, 329, 1001-6.
- PANCSA, R. & TOMPA, P. 2012. Structural disorder in eukaryotes. *PLoS One*, 7, e34687.
- PARK, K. E., INEROWICZ, H. D., WANG, X., LI, Y., KOSER, S. & CABOT, R. A. 2012. Identification of karyopherin alpha1 and alpha7 interacting proteins in porcine tissue. *PLoS One*, 7, e38990.
- PATTI, J. M., ALLEN, B. L., MCGAVIN, M. J. & HOOK, M. 1994a. MSCRAMM-mediated adherence of microorganisms to host tissues. *Annual review of microbiology*, 48, 585-617.
- PATTI, J. M., BREMELL, T., KRAJEWSKA-PIETRASIK, D., ABDELNOUR, A., TARKOWSKI, A., RYDEN, C. & HOOK, M. 1994b. The *Staphylococcus aureus* collagen adhesin is a virulence determinant in experimental septic arthritis. *Infection and immunity*, 62, 152-61.
- PATZER, S. I. & HANTKE, K. 2000. The zinc-responsive regulator Zur and its control of the *znu* gene cluster encoding the ZnuABC zinc uptake system in *Escherichia coli*. *The Journal of biological chemistry*, 275, 24321-32.
- PEACOCK, S. J., DAY, N. P., THOMAS, M. G., BERENDT, A. R. & FOSTER, T. J. 2000. Clinical isolates of *Staphylococcus aureus* exhibit diversity in *fnb* genes and adhesion to human fibronectin. *J Infect*, 41, 23-31.
- PEACOCK, S. J., MOORE, C. E., JUSTICE, A., KANTZANOU, M., STORY, L., MACKIE, K., O'NEILL, G. & DAY, N. P. 2002. Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus aureus*. *Infection and immunity*, 70, 4987-96.
- PELTON, J. G., TORCHIA, D. A., MEADOW, N. D., WONG, C. Y. & ROSEMAN, S. 1991. ¹H, ¹⁵N, and ¹³C NMR signal assignments of IIIgIc, a signal-transducing protein of *Escherichia coli*, using three-dimensional triple-resonance techniques. *Biochemistry*, 30, 10043-57.
- PERMI, P. & HELLMAN, M. 2012. Alpha proton detection based backbone assignment of intrinsically disordered proteins. *Methods in molecular biology*, 895, 211-26.
- PERVUSHIN, K., RIEK, R., WIDER, G. & WUTHRICH, K. 1997. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large

- biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 12366-71.
- PESCHEL, A., OTTO, M., JACK, R. W., KALBACHER, H., JUNG, G. & GOTZ, F. 1999a. Inactivation of the *dlt* operon in *Staphylococcus aureus* confers sensitivity to defensins, protegrins, and other antimicrobial peptides. *The Journal of biological chemistry*, 274, 8405-10.
- PESCHEL, A., OTTO, M., JACK, R. W., KALBACHER, H., JUNG, G. N. & GOTZ, F. 1999b. Inactivation of the *dlt* Operon in *Staphylococcus aureus* Confers Sensitivity to Defensins, Protegrins, and Other Antimicrobial Peptides. *Journal of Biological Chemistry*, 274, 8405-8410.
- PETERSEN, F. C., TAO, L. & SCHEIE, A. A. 2005. DNA binding-uptake system: a link between cell-to-cell communication and biofilm formation. *Journal of Bacteriology*, 187, 4392-400.
- PETERSON, P. K., VERHOEF, J., SABATH, L. D. & QUIE, P. G. 1977. Effect of protein A on staphylococcal opsonization. *Infection and immunity*, 15, 760-4.
- PIROTH, L., QUE, Y. A., WIDMER, E., PANCHAUD, A., PIU, S., ENTENZA, J. M. & MOREILLON, P. 2008. The fibrinogen- and fibronectin-binding domains of *Staphylococcus aureus* fibronectin-binding protein A synergistically promote endothelial invasion and experimental endocarditis. *Infect Immun*, 76, 3824-31.
- PLOUIN-GAUDON, I., CLEMENT, S., HUGGLER, E., CHAPONNIER, C., FRANCOIS, P., LEW, D., SCHRENZEL, J., VAUDAUX, P. & LACROIX, J. S. 2006. Intracellular residency is frequently associated with recurrent *Staphylococcus aureus* rhinosinusitis. *Rhinology*, 44, 249-54.
- POLLACK, J. H. & NEUHAUS, F. C. 1994. Changes in wall teichoic acid during the rod-sphere transition of *Bacillus subtilis* 168. *Journal of Bacteriology*, 176, 7252-9.
- POWERS, R., GARRETT, D. S., MARCH, C. J., FRIEDEN, E. A., GRONENBORN, A. M. & CLORE, G. M. 1992. 1H, 15N, 13C, and 13CO assignments of human interleukin-4 using three-dimensional double- and triple-resonance heteronuclear magnetic resonance spectroscopy. *Biochemistry*, 31, 4334-46.
- PRILUSKY, J., FELDER, C. E., ZEEV-BEN-MORDEHAI, T., RYDBERG, E. H., MAN, O., BECKMANN, J. S., SILMAN, I. & SUSSMAN, J. L. 2005. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics*, 21, 3435-8.
- QIN, Z., OU, Y., YANG, L., ZHU, Y., TOLKER-NIELSEN, T., MOLIN, S. & QU, D. 2007. Role of autolysin-mediated DNA release in biofilm formation of *Staphylococcus epidermidis*. *Microbiology*, 153, 2083-92.
- QUE, Y. A., FRANCOIS, P., HAEFLIGER, J. A., ENTENZA, J. M., VAUDAUX, P. & MOREILLON, P. 2001. Reassessing the role of *Staphylococcus aureus* clumping factor and fibronectin-binding protein by expression in *Lactococcus lactis*. *Infection and immunity*, 69, 6296-302.
- QUE, Y. A., HAEFLIGER, J. A., PIROTH, L., FRANCOIS, P., WIDMER, E., ENTENZA, J. M., SINHA, B., HERRMANN, M., FRANCIOLI, P., VAUDAUX, P. & MOREILLON, P. 2005. Fibrinogen and fibronectin binding cooperate for valve infection and invasion in *Staphylococcus aureus* experimental endocarditis. *J Exp Med*, 201, 1627-35.
- RANI, S. A., PITTS, B., BEYENAL, H., VELUCHAMY, R. A., LEWANDOWSKI, Z., DAVISON, W. M., BUCKINGHAM-MEYER, K. & STEWART, P. S.

2007. Spatial patterns of DNA replication, protein synthesis, and oxygen concentration within bacterial biofilms reveal diverse physiological states. *Journal of Bacteriology*, 189, 4223-33.
- RASIA, R. M., BRUTSCHER, B. & PLEVIN, M. J. 2012. Selective isotopic unlabeled proteins using metabolic precursors: application to NMR assignment of intrinsically disordered proteins. *Chembiochem : a European journal of chemical biology*, 13, 732-9.
- REZAEI-GHALEH, N., BLACKLEDGE, M. & ZWECKSTETTER, M. 2012. Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *Chembiochem : a European journal of chemical biology*, 13, 930-50.
- RICHARDSON, J. M., COLLOMS, S. D., FINNEGAN, D. J. & WALKINSHAW, M. D. 2009. Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. *Cell*, 138, 1096-108.
- RIEK, R., PERVUSHIN, K. & WUTHRICH, K. 2000. TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends in biochemical sciences*, 25, 462-8.
- RIEK, R., WIDER, G., PERVUSHIN, K. & WUTHRICH, K. 1999. Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 4918-23.
- RIPPON, J. E. & VOGELSANG, T. M. 1956. Carriage of pathogenic staphylococci in the upper respiratory tract of children. *Acta pathologica et microbiologica Scandinavica*, 39, 284-96.
- ROCHE, F. M., DOWNER, R., KEANE, F., SPEZIALE, P., PARK, P. W. & FOSTER, T. J. 2004. The N-terminal A domain of fibronectin-binding proteins A and B promotes adhesion of *Staphylococcus aureus* to elastin. *J Biol Chem*, 279, 38433-40.
- ROCHE, F. M., MEEHAN, M. & FOSTER, T. J. 2003. The *Staphylococcus aureus* surface protein SasG and its homologues promote bacterial adherence to human desquamated nasal epithelial cells. *Microbiology*, 149, 2759-67.
- ROHDE, H., BURDELSKI, C., BARTSCHT, K., HUSSAIN, M., BUCK, F., HORSTKOTTE, M. A., KNOBLOCH, J. K., HEILMANN, C., HERRMANN, M. & MACK, D. 2005. Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. *Molecular microbiology*, 55, 1883-95.
- ROHDE, H., FRANKENBERGER, S., ZHRINGER, U. & MACK, D. 2010. Structure, function and contribution of polysaccharide intercellular adhesin (PIA) to *Staphylococcus epidermidis* biofilm formation and pathogenesis of biomaterial-associated infections. *European journal of cell biology*, 89, 103-11.
- ROMERO, P., OBRADOVIC, Z., LI, X., GARNER, E. C., BROWN, C. J. & DUNKER, A. K. 2001. Sequence complexity of disordered protein. *Proteins*, 42, 38-48.
- SADOVSKAYA, I., CHAIGNON, P., KOGAN, G., CHOKR, A., VINOGRADOV, E. & JABBOURI, S. 2006. Carbohydrate-containing components of biofilms produced in vitro by some staphylococcal strains related to orthopaedic prosthesis infections. *FEMS immunology and medical microbiology*, 47, 75-82.

- SADOVSKAYA, I., VINOGRADOV, E., FLAHAUT, S., KOGAN, G. & JABBOURI, S. 2005. Extracellular carbohydrate-containing polymers of a model biofilm-producing strain, *Staphylococcus epidermidis* RP62A. *Infection and immunity*, 73, 3007-17.
- SADOVSKAYA, I., VINOGRADOV, E., LI, J. & JABBOURI, S. 2004. Structural elucidation of the extracellular and cell-wall teichoic acids of *Staphylococcus epidermidis* RP62A, a reference biofilm-positive strain. *Carbohydrate research*, 339, 1467-73.
- SCHMITT, A. P. & MCENTEE, K. 1996. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 5777-82.
- SCHOMMER, N. N., CHRISTNER, M., HENTSCHKE, M., RUCKDESCHEL, K., AEPFELBACHER, M. & ROHDE, H. 2011. *Staphylococcus epidermidis* uses distinct mechanisms of biofilm formation to interfere with phagocytosis and activation of mouse macrophage-like cells 774A.1. *Infection and immunity*, 79, 2267-76.
- SCHREITER, E. R., WANG, S. C., ZAMBLE, D. B. & DRENNAN, C. L. 2006. NikR-operator complex structure and the mechanism of repressor activation by metal ions. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 13676-81.
- SCHROEDER, K., JULARIC, M., HORSBURGH, S. M., HIRSCHHAUSEN, N., NEUMANN, C., BERTLING, A., SCHULTE, A., FOSTER, S., KEHREL, B. E., PETERS, G. & HEILMANN, C. 2009. Molecular characterization of a novel *Staphylococcus aureus* surface protein (SasC) involved in cell aggregation and biofilm accumulation. *PLoS One*, 4, e7567.
- SCHWARZ-LINEK, U., WERNER, J. M., PICKFORD, A. R., GURUSIDDAPPA, S., KIM, J. H., PILKA, E. S., BRIGGS, J. A., GOUGH, T. S., HOOK, M., CAMPBELL, I. D. & POTTS, J. R. 2003. Pathogenic bacteria attach to human fibronectin through a tandem beta-zipper. *Nature*, 423, 177-81.
- SCHWARZINGER, S., KROON, G. J., FOSS, T. R., CHUNG, J., WRIGHT, P. E. & DYSON, H. J. 2001. Sequence-dependent correction of random coil NMR chemical shifts. *Journal of the American Chemical Society*, 123, 2970-8.
- SCHWARZINGER, S., KROON, G. J., FOSS, T. R., WRIGHT, P. E. & DYSON, H. J. 2000. Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *Journal of biomolecular NMR*, 18, 43-8.
- SHANKS, R. M., MEEHL, M. A., BROTHERS, K. M., MARTINEZ, R. M., DONEGAN, N. P., GRABER, M. L., CHEUNG, A. L. & O'TOOLE, G. A. 2008. Genetic evidence for an alternative citrate-dependent biofilm formation pathway in *Staphylococcus aureus* that is dependent on fibronectin binding proteins and the GraRS two-component regulatory system. *Infection and immunity*, 76, 2469-77.
- SHINJI, H., YOSIZAWA, Y., TAJIMA, A., IWASE, T., SUGIMOTO, S., SEKI, K. & MIZUNOE, Y. 2011. Role of Fibronectin-Binding Proteins A and B in In Vitro Cellular Infections and In Vivo Septic Infections by *Staphylococcus aureus*. *Infect Immun*, 79, 2215-23.
- SIGALOV, A. B., AIVAZIAN, D. A., UVERSKY, V. N. & STERN, L. J. 2006. Lipid-binding activity of intrinsically unstructured cytoplasmic domains of

- multichain immune recognition receptor signaling subunits. *Biochemistry*, 45, 15731-9.
- SINHA, B., FRANCOIS, P., QUE, Y. A., HUSSAIN, M., HEILMANN, C., MOREILLON, P., LEW, D., KRAUSE, K. H., PETERS, G. & HERRMANN, M. 2000a. Heterologously expressed *Staphylococcus aureus* fibronectin-binding proteins are sufficient for invasion of host cells. *Infection and immunity*, 68, 6871-8.
- SINHA, B., FRANCOIS, P., QUE, Y. A., HUSSAIN, M., HEILMANN, C., MOREILLON, P., LEW, D., KRAUSE, K. H., PETERS, G. & HERRMANN, M. 2000b. Heterologously expressed *Staphylococcus aureus* fibronectin-binding proteins are sufficient for invasion of host cells. *Infect Immun*, 68, 6871-8.
- SINHA, B., FRANCOIS, P. P., NUSSE, O., FOTI, M., HARTFORD, O. M., VAUDAUX, P., FOSTER, T. J., LEW, D. P., HERRMANN, M. & KRAUSE, K. H. 1999. Fibronectin-binding protein acts as *Staphylococcus aureus* invasin via fibronectin bridging to integrin $\alpha 5\beta 1$. *Cellular microbiology*, 1, 101-17.
- SINHA, B. & HERRMANN, M. 2005. Mechanism and consequences of invasion of endothelial cells by *Staphylococcus aureus*. *Thrombosis and haemostasis*, 94, 266-77.
- SPOLAR, R. S. & RECORD, M. T., JR. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science*, 263, 777-84.
- STROP, P., WIDER, G. & WUTHRICH, K. 1983. Assignment of the ^1H nuclear magnetic resonance spectrum of the proteinase inhibitor IIA from bull seminal plasma by two-dimensional nuclear magnetic resonance at 500 MHz. *Journal of molecular biology*, 166, 641-65.
- SWOBODA, J. G., CAMPBELL, J., MEREDITH, T. C. & WALKER, S. 2010. Wall teichoic acid function, biosynthesis, and inhibition. *Chembiochem : a European journal of chemical biology*, 11, 35-45.
- SZALAINÉ AGOSTON, B., KOVACS, D., TOMPA, P. & PERCZEL, A. 2011. Full backbone assignment and dynamics of the intrinsically disordered dehydrin ERD14. *Biomolecular NMR assignments*, 5, 189-93.
- TANAKA, H., YAMASHITA, T., ASADA, M., MIZUTANI, S., YOSHIKAWA, H. & TOHYAMA, M. 2002. Cytoplasmic p21(Cip1/WAF1) regulates neurite remodeling by inhibiting Rho-kinase activity. *The Journal of cell biology*, 158, 321-9.
- THURLOW, L. R., HANKE, M. L., FRITZ, T., ANGLE, A., ALDRICH, A., WILLIAMS, S. H., ENGBRETSEN, I. L., BAYLES, K. W., HORSWILL, A. R. & KIELIAN, T. 2011. *Staphylococcus aureus* biofilms prevent macrophage phagocytosis and attenuate inflammation in vivo. *Journal of immunology*, 186, 6585-96.
- TOMPA, P. 2002. Intrinsically unstructured proteins. *Trends in biochemical sciences*, 27, 527-33.
- TRAYER, H. R. & BUCKLEY, C. E., 3RD 1970. Molecular properties of lysostaphin, a bacteriolytic agent specific for *Staphylococcus aureus*. *The Journal of biological chemistry*, 245, 4842-6.
- TUGARINOV, V. & KAY, L. E. 2004. Stereospecific NMR assignments of prochiral methyls, rotameric states and dynamics of valine residues in malate synthase G. *Journal of the American Chemical Society*, 126, 9827-36.

- TZAGOLOFF, H. & NOVICK, R. 1977. Geometry of cell division in *Staphylococcus aureus*. *Journal of Bacteriology*, 129, 343-50.
- UHLEN, M., GUSS, B., NILSSON, B., GOTZ, F. & LINDBERG, M. 1984. Expression of the gene encoding protein A in *Staphylococcus aureus* and coagulase-negative staphylococci. *Journal of Bacteriology*, 159, 713-9.
- UVERSKY, V. N. 2003. Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: which way to go? *Cellular and molecular life sciences : CMLS*, 60, 1852-71.
- VALLE, J., LATASA, C., GIL, C., TOLEDO-ARANA, A., SOLANO, C., PENADES, J. R. & LASA, I. 2012. Bap, a Biofilm Matrix Protein of *Staphylococcus aureus* Prevents Cellular Internalization through Binding to GP96 Host Receptor. *PLoS pathogens*, 8, e1002843.
- VAUDAUX, P., PITTET, D., HAEBERLI, A., HUGGLER, E., NYDEGGER, U. E., LEW, D. P. & WALDVOGEL, F. A. 1989. Host factors selectively increase staphylococcal adherence on inserted catheters: a role for fibronectin and fibrinogen or fibrin. *The Journal of infectious diseases*, 160, 865-75.
- VEENSTRA, G. J., CREMERS, F. F., VAN DIJK, H. & FLEER, A. 1996. Ultrastructural organization and regulation of a biomaterial adhesion of *Staphylococcus epidermidis*. *Journal of Bacteriology*, 178, 537-41.
- VENTERS, R. A., FARMER, B. T., 2ND, FIERKE, C. A. & SPICER, L. D. 1996. Characterizing the use of perdeuteration in NMR studies of large proteins: ¹³C, ¹⁵N and ¹H assignments of human carbonic anhydrase II. *Journal of molecular biology*, 264, 1101-16.
- VERGARA-IRIGARAY, M., MAIRA-LITRAN, T., MERINO, N., PIER, G. B., PENADES, J. R. & LASA, I. 2008. Wall teichoic acids are dispensable for anchoring the PNAG exopolysaccharide to the *Staphylococcus aureus* cell surface. *Microbiology*, 154, 865-77.
- VERGARA-IRIGARAY, M., VALLE, J., MERINO, N., LATASA, C., GARCIA, B., RUIZ DE LOS MOZOS, I., SOLANO, C., TOLEDO-ARANA, A., PENADES, J. R. & LASA, I. 2009. Relevant role of fibronectin-binding proteins in *Staphylococcus aureus* biofilm-associated foreign-body infections. *Infection and immunity*, 77, 3978-91.
- VINOGRADOV, E., SADOVSKAYA, I., LI, J. & JABBOURI, S. 2006. Structural elucidation of the extracellular and cell-wall teichoic acids of *Staphylococcus aureus* MN8m, a biofilm forming strain. *Carbohydrate research*, 341, 738-43.
- VON EIFF, C., BECKER, K., MACHKA, K., STAMMER, H. & PETERS, G. 2001. Nasal carriage as a source of *Staphylococcus aureus* bacteremia. Study Group. *The New England journal of medicine*, 344, 11-6.
- VOYICH, J. M., BRAUGHTON, K. R., STURDEVANT, D. E., WHITNEY, A. R., SAID-SALIM, B., PORCELLA, S. F., LONG, R. D., DORWARD, D. W., GARDNER, D. J., KREISWIRTH, B. N., MUSSER, J. M. & DELEO, F. R. 2005. Insights into mechanisms used by *Staphylococcus aureus* to avoid destruction by human neutrophils. *Journal of immunology*, 175, 3907-19.
- VRANKEN, W. F., BOUCHER, W., STEVENS, T. J., FOGH, R. H., PAJON, A., LLINAS, M., ULRICH, E. L., MARKLEY, J. L., IONIDES, J. & LAUE, E. D. 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, 59, 687-96.
- VUONG, C., KOCIANOVA, S., VOYICH, J. M., YAO, Y., FISCHER, E. R., DELEO, F. R. & OTTO, M. 2004. A crucial role for exopolysaccharide

- modification in bacterial biofilm formation, immune evasion, and virulence. *The Journal of biological chemistry*, 279, 54881-6.
- WAGNER, G., PARDI, A. & WUTHRICH, K. 1983. HYDROGEN-BOND LENGTH AND H-1-NMR CHEMICAL-SHIFTS IN PROTEINS. *Journal of the American Chemical Society*, 105, 5948-5949.
- WAGNER, G. & WUTHRICH, K. 1982. Sequential resonance assignments in protein ¹H nuclear magnetic resonance spectra. Basic pancreatic trypsin inhibitor. *Journal of molecular biology*, 155, 347-66.
- WANG, Y. & JARDETZKY, O. 2002. Investigation of the neighboring residue effects on protein chemical shifts. *Journal of the American Chemical Society*, 124, 14075-84.
- WANN, E. R., GURUSIDDAPPA, S. & HOOK, M. 2000. The fibronectin-binding MSCRAMM FnbpA of *Staphylococcus aureus* is a bifunctional protein that also binds to fibrinogen. *J Biol Chem*, 275, 13863-71.
- WARD, J. J., SODHI, J. S., MCGUFFIN, L. J., BUXTON, B. F. & JONES, D. T. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of molecular biology*, 337, 635-45.
- WATANABE, I., ICHIKI, M., SHIRATSUCHI, A. & NAKANISHI, Y. 2007. TLR2-mediated survival of *Staphylococcus aureus* in macrophages: a novel bacterial strategy against host innate immunity. *Journal of immunology*, 178, 4917-25.
- WEIDENMAIER, C., KOKAI-KUN, J. F., KRISTIAN, S. A., CHANTURIYA, T., KALBACHER, H., GROSS, M., NICHOLSON, G., NEUMEISTER, B., MOND, J. J. & PESCHEL, A. 2004. Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections. *Nature medicine*, 10, 243-5.
- WEIDENMAIER, C. & PESCHEL, A. 2008. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature reviews. Microbiology*, 6, 276-87.
- WEIDENMAIER, C., PESCHEL, A., XIONG, Y. Q., KRISTIAN, S. A., DIETZ, K., YEAMAN, M. R. & BAYER, A. S. 2005. Lack of wall teichoic acids in *Staphylococcus aureus* leads to reduced interactions with endothelial cells and to attenuated virulence in a rabbit model of endocarditis. *The Journal of infectious diseases*, 191, 1771-7.
- WEN, J., WU, J. & ZHOU, P. 2011. Sparsely sampled high-resolution 4-D experiments for efficient backbone resonance assignment of disordered proteins. *Journal of magnetic resonance*, 209, 94-100.
- WENZEL, R. P. & PERL, T. M. 1995. The significance of nasal carriage of *Staphylococcus aureus* and the incidence of postoperative wound infection. *The Journal of hospital infection*, 31, 13-24.
- WHITMORE, L. & WALLACE, B. A. 2004. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic acids research*, 32, W668-73.
- WICKHAM, J. R., HALYE, J. L., KASHTANOV, S., KHANDOGIN, J. & RICE, C. V. 2009. Revisiting magnesium chelation by teichoic acid with phosphorus solid-state NMR and theoretical calculations. *The journal of physical chemistry. B*, 113, 2177-83.
- WISHART, D. S., BIGAM, C. G., HOLM, A., HODGES, R. S. & SYKES, B. D. 1995a. ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common

- amino acids. I. Investigations of nearest-neighbor effects. *Journal of biomolecular NMR*, 5, 67-81.
- WISHART, D. S., BIGAM, C. G., YAO, J., ABILDGAARD, F., DYSON, H. J., OLDFIELD, E., MARKLEY, J. L. & SYKES, B. D. 1995b. ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *Journal of biomolecular NMR*, 6, 135-40.
- WISHART, D. S. & NIP, A. M. 1998. Protein chemical shift analysis: a practical guide. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 76, 153-63.
- WISHART, D. S. & SYKES, B. D. 1994. The ¹³C chemical-shift index: a simple method for the identification of protein secondary structure using ¹³C chemical-shift data. *Journal of biomolecular NMR*, 4, 171-80.
- WISHART, D. S., SYKES, B. D. & RICHARDS, F. M. 1991. Simple techniques for the quantification of protein secondary structure by ¹H NMR spectroscopy. *FEBS letters*, 293, 72-80.
- WISHART, D. S., SYKES, B. D. & RICHARDS, F. M. 1992. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, 31, 1647-51.
- WOLTERS, P. J., HILDEBRANDT, K. M., DICKIE, J. P. & ANDERSON, J. S. 1990. Polymer length of teichuronic acid released from cell walls of *Micrococcus luteus*. *Journal of Bacteriology*, 172, 5154-9.
- WU, J. A., KUSUMA, C., MOND, J. J. & KOKAI-KUN, J. F. 2003. Lysostaphin disrupts *Staphylococcus aureus* and *Staphylococcus epidermidis* biofilms on artificial surfaces. *Antimicrobial agents and chemotherapy*, 47, 3407-14.
- WUTHRICH, K., WIDER, G., WAGNER, G. & BRAUN, W. 1982. Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance. *Journal of molecular biology*, 155, 311-9.
- XIA, G., KOHLER, T. & PESCHEL, A. 2010a. The wall teichoic acid and lipoteichoic acid polymers of *Staphylococcus aureus*. *International journal of medical microbiology : IJMM*, 300, 148-54.
- XIA, G., MAIER, L., SANCHEZ-CARBALLO, P., LI, M., OTTO, M., HOLST, O. & PESCHEL, A. 2010b. Glycosylation of wall teichoic acid in *Staphylococcus aureus* by TarM. *The Journal of biological chemistry*, 285, 13405-15.
- XUE, B., DUNKER, A. K. & UVERSKY, V. N. 2012. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *Journal of biomolecular structure & dynamics*, 30, 137-49.
- YAMAZUMI, T., MARSHALL, S. A., WILKE, W. W., DIEKEMA, D. J., PFALLER, M. A. & JONES, R. N. 2001. Comparison of the Vitek Gram-Positive Susceptibility 106 card and the MRSA-screen latex agglutination test for determining oxacillin resistance in clinical bloodstream isolates of *Staphylococcus aureus*. *Journal of clinical microbiology*, 39, 53-6.
- ZENG, Y., HE, Y., YANG, F., MOONEY, S. M., GETZENBERG, R. H., ORBAN, J. & KULKARNI, P. 2011. The cancer/testis antigen prostate-associated gene 4 (PAGE4) is a highly intrinsically disordered protein. *The Journal of biological chemistry*, 286, 13985-94.
- ZHANG, X., PERUGINI, M. A., YAO, S., ADDA, C. G., MURPHY, V. J., LOW, A., ANDERS, R. F. & NORTON, R. S. 2008. Solution conformation,

- backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *Journal of molecular biology*, 379, 105-21.
- ZIEBUHR, W. 2001. Staphylococcus aureus and Staphylococcus epidermidis: emerging pathogens in nosocomial infections. *Contributions to microbiology*, 8, 102-7.
- ZOLL, S., PATZOLD, B., SCHLAG, M., GOTZ, F., KALBACHER, H. & STEHLE, T. 2010. Structural basis of cell wall cleavage by a staphylococcal autolysin. *PLoS pathogens*, 6, e1000807.