This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Shaw, John M

Title:
Biochemical and Structural Characterisation of a Thermostable Cfr-Like Enzyme from Sphaerobacter thermophilus

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.
Biochemical and Structural Characterisation of a Thermostable Cfr-Like Enzyme from *Sphaerobacter thermophilus*

John M. Shaw

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of MSc(Res) in the Faculty of Life Sciences.

Submission Date: Sept 2018

Word count: 24669
Abstract

Cfr is a radical SAM (S-adenosyl-L-methionine) methyltransferase capable of conferring resistance to phenicols, lincomicosides, oxazolidinones, pleuromutilins and streptogramin A (termed the PhLOPSA phenotype) via the C8 methylation of 23s rRNA residue A2503. Radical SAM enzymes utilise a [4Fe-4S] cluster to coordinate the reductive cleavage of SAM producing a 5’Deoxyadenosyl radical, which is used to catalyse reactions at inherently unreactive carbon centres. To date, attempts to crystallise Cfr have been unsuccessful and all structural data originate from RlmN, a Cfr homolog with no resistance phenotype capable of methylating position C2 of the same residue.

To address the void of structural knowledge regarding Cfr, four Cfr-like enzymes were selected for investigation into their biochemical characteristics and potential for crystallisation: the Cfr(B) and Cfr(C) alleles originally identified in resistant pathogens, and the PbCfr (Paenibacillus spp.) and SbCfr (Sphaerobacter thermophilus) homologues from thermophilic organisms. Work presented here shows expression of SbCfr and Cfr in recombinant E. coli BL21* and anaerobic purification by immobilised metal ion chromatography, yielding 10.25 and 22.4 mg.L⁻¹ E. coli culture, respectively. UV-vis spectra of SbCfr display a 420nm peak consistent with the presence of intact [4Fe-4S] clusters before and after reconstitution, while anaerobic activity assays show s-adenosylhomocysteine (SAH) and 5’ deoxyadenosine (DOA) formation by both reconstituted and unreconstituted SbCfr. Activity assays and circular dichroism spectroscopy indicate SbCfr activity up to 55°C and stability up to 65°C. However, SbCfr did not confer the PhLOPSA phenotype on recombinant E. coli, while efforts to use site-directed mutants to trap covalent RNA complexes, indicative of RlmN-like activity, were similarly unsuccessful.

Although crystallisation trials of Cfr were not successful, SbCfr was successfully crystallised and the structure determined to 1.4 Å resolution, revealing a monomeric enzyme containing a partial TIM β-barrel, hosting a [4Fe-4S] cluster coordinated by the canonical radical SAM enzyme CX₃CX₂C motif, and resembling previous structures of E. coli RlmN. The structure
contains an SAH ligand bound to the 4Fe-4S cluster and a methyl group attached to Cys334, which adopts a conformation resembling that observed in RNA-bound RlmN but distinct from that in complexes with SAM. This represents a state not previously described in RlmN structures. This work shows SbCfr to be a stable radical SAM enzyme with likely methyltransferase activity but that is not associated with antimicrobial resistance and suggests that conformational changes previously thought to occur on RNA binding occur earlier in the reaction cycle of radical SAM methyltransferases.
Acknowledgements

I would like to thank the researchers of D57 and D60 for their support and advice throughout this project.

Specifically, I would like to thank Dr Jacqueline Findlay for always keeping the coffee hot.

I would like to thank Dr Philip Hinchcliffe for his tireless efforts, patience and huge contribution to this project. Without whom all I would have are crystal plates of precipitate.

I would also like to thank Dr Jim Spencer, whose trust, advice, expertise and kindness have made this project possible.
Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ............................................................. DATE:..............................
Contents

Abstract ............................................................................................................................................. 3
Chapter one: Introduction ................................................................................................................. 1
  1.1: The emerging threat of antimicrobial resistance (AMR) ....................................................... 1
  1.2: Current clinical impact of AMR ............................................................................................. 2
    1.2.1: Enterococci spp .............................................................................................................. 2
    1.2.2: Staphylococcus aureus ................................................................................................... 5
    1.2.3: Klebsiella pneumoniae ................................................................................................... 9
    1.2.4: Acinetobacter baumannii ................................................................................................ 10
    1.2.5: Pseudomonas aeruginosa ............................................................................................... 11
    1.2.6: Enterobacter spp ......................................................................................................... 12
    1.2.7: Escherichia coli ............................................................................................................ 14
  1.3 Mechanisms of antimicrobial resistance .................................................................................... 15
    1.3.1: Enzymatic inactivation/modification of antimicrobial ..................................................... 15
    1.3.2: Target site modification .................................................................................................. 17
    1.3.3: Efflux of antimicrobials .................................................................................................. 18
    1.3.4: Porin alteration ............................................................................................................... 19
    1.3.5: Overproduction of target ................................................................................................ 20
    1.3.6: Target bypass ................................................................................................................ 20
  1.4 The ribosome as an antimicrobial target site ............................................................................ 21
    1.4.1: Structure and function of prokaryotic ribosome ............................................................. 21
    1.4.2: Sites integral to antimicrobial activity ............................................................................. 24
  1.5 cfr, rlmn and the Radical SAM Methyltransferase superfamily ............................................ 27
    1.5.1: Cfr and RlmN .................................................................................................................. 27
    1.5.2: Radical SAM superfamily ............................................................................................. 27
  1.6 Dissemination and Clinical Impact of cfr ............................................................................... 33
    1.6.1: Cfr in a clinical environment .......................................................................................... 33
    1.6.2: Cfr homologues ............................................................................................................. 34
    1.6.3: Antimicrobial genes associated with cfr ......................................................................... 35
  1.7 Aims of project ........................................................................................................................ 36
Chapter two: Methods and Materials .............................................................................................. 38
  2.1: Selection of candidate genes ................................................................................................... 38
  2.2: Vector construction ............................................................................................................... 38
  2.3: PCR amplification .................................................................................................................. 39
  2.4: Colony PCR .......................................................................................................................... 41
  2.5: Cell lines (incl competent cell generation) ............................................................................. 41
2.6: Recombinant protein expression strains ......................................................... 41
2.7: Minimum inhibitory concentration strains .................................................... 42
2.8: Agarose gel electrophoresis ......................................................................... 42
2.9: Transformation ............................................................................................. 42
2.10: Plasmid purification .................................................................................... 42
2.11: Protein expression trials ............................................................................ 43
2.12: Aerobic NiNTA purification ....................................................................... 43
2.13: Anaerobic protein purification .................................................................. 44
2.14: SDS-page .................................................................................................. 46
2.15: Reconstitution ............................................................................................ 46
2.16: Activity Assay ........................................................................................... 46
2.17: HPLC protocol .......................................................................................... 47
2.19: Anaerobic Crystal trials ............................................................................ 47
2.20: Crystal structure resolution ....................................................................... 48
2.21: Disk testing methods ................................................................................ 48

Chapter three: Recombinant protein production and biochemical characterisation ... 49
3.1: Introduction .................................................................................................. 49
3.2: Vector construction ...................................................................................... 49
3.3: Vector expression trials .............................................................................. 53
3.4: Aerobic protein purification ........................................................................ 59
3.5: Anaerobic protein purification and reconstitution ........................................ 63
   3.5.1: Purification of Cfr .................................................................................. 63
   3.5.2: Purification of SbCfr ............................................................................. 66
   3.5.3: Reconstitution of purified Cfr/SbCfr ...................................................... 68
3.6: S-Adenosylhomocysteine (SAH) and 5’ Deoxyadenosine (DOA) production assays .......................................................... 73
3.7: Structural stability assessment by circular dichroism as a function of temperature .................................................................................. 82
3.8: rRNA:Protein adduct formation .................................................................. 83
3.9: Effect of SbCfr, Cfr and Cfr(B) expression on antimicrobial susceptibility of recombinant E coli .......................................................... 89
3.10: Discussion .................................................................................................. 92

Chapter four: Crystallisation and Structure Determination for SbCfr .................. 99
4.1: Crystal screening .......................................................................................... 99
4.2: Crystallisation screening of Cfr ................................................................... 99
4.3: Crystallisation screening for SbCfr .............................................................. 101
4.4: Diffraction data collection .......................................................................... 105
4.5: Crystal structure solving ............................................................................. 109
Chapter one: Introduction

1.1: The emerging threat of antimicrobial resistance (AMR)

Antimicrobials are key therapeutic agents used in the event of infectious disease. In 2010 the United Kingdom saw 8% of all hospital admissions attributed to infectious diseases caused by bacterial, viral, fungal and parasitic pathogens. In recent years, many developed countries have seen a decline in mortality rate due to infectious disease as a result of improved sanitation. In spite of this, 7% of all deaths in the United Kingdom in 2010 were attributed to infectious disease and when accounting for costs shouldered by the National Health Service (NHS), the labour market and the individual, infectious disease costs the UK around £30b annually [1].

The emergence of resistance to antimicrobials is a natural phenomenon. Directly after the identification of Penicillin in 1940, resistance to its activity was observed, even before its wide dissemination [2]. Many of the antimicrobials in use today were isolated from organisms that produce and release these compounds to promote survival in their respective environments [3, 4]. This highlights the possibility that all antimicrobial therapies derived from these sources are vulnerable to pre-existing resistance mechanisms somewhere in the environment.

It has been estimated that by the year 2050 the global impact of antimicrobial resistance could result in an additional 10 million lives lost annually, with a cost of £80.4 trillion lost in GDP to the world’s economies [5]. Considering the high prevalence of infectious disease and our reliance on antimicrobials as a key therapy the emergence and dissemination of resistance mechanisms proves a disquieting prospect for public health.
1.2: Current clinical impact of AMR

The focus of bacterial infectious disease in the clinical setting is largely based around a subset of specific pathogens, termed the E.S.K.A.P.E.E group: Enterococci spp, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter Spp. and Escherichia coli. These species make up the bulk of nosocomial infections worldwide and understanding of their virulence and pathogenicity is paramount.

All of the E.S.K.A.P.E.E pathogens have been documented to possess some form of antimicrobial resistance determinant; with the frequency of resistant isolate discovery increasing over time [6]. Currently, infections by Gram-negative bacteria are the most clinically problematic. The additional outer membrane and unique efflux systems not present in their Gram-positive counterparts offer intrinsic resistance to a range of antimicrobials [7]. In addition to this, the rise of the Gram-positive Methicillin Resistant Staphylococcus aureus (MRSA) in the late twentieth century forced research focus to Gram-positives. This shift in focus underpins why there are currently more successful treatment options for Gram-positive infections, when compared to Gram-negative. While the arsenal of therapeutic agents with activity against Gram-positive organisms may be more robust it is by no means secure. Resistance to these antimicrobials is continually emerging.

1.2.1: Enterococci spp

The two prominent clinical strains within this group are Enterococcus faecalis and Enterococcus faecium. These two Gram-positive organisms are part of the normal human gastro-intestinal flora and act as opportunistic pathogens in immunocompromised patients [8]. Enterococci spp. have been documented to be the causative agent of endocarditis, urinary tract infections and bacteraemia. In 2016 Public Health England (PHE) reported the rate of bacteraemia, the presence of bacteria in the blood, caused by enterococci at 12.6 per 100,000 population, in contrast with 10.6 per 100,000 population in 2010, as shown in figure 1 [9]. During this period, of all bacteraemias involving Enterococci spp. 42% of cases were attributed
to *E. faecalis* and 33% to *E. Faecium*, with *E. faecalis* totalling 1.9% of all bacteraemia reported.

![Graph showing Enterococcus bacteraemia per 100,000 population in England, Wales and Northern Ireland (2009-2016).](image)

**Figure 1** - *Enterococcus* bacteraemia per 100,000 population in England, Wales and Northern Ireland. 2009 - 2016. Figure adapted from PHE laboratory surveillance 2017 [9].

In 2004 in the United States, a surveillance study by the *Surveillance and Control of Pathogens of Epidemiological Importance* (SCOPE) found that out of 24,179 cases of nosocomial bloodstream infections from 49 US hospitals, *Enterococci spp.* accounted for ~9% of all documented bacteraemias [10]. In both Europe and the United States, the majority of infections caused by *Enterococci* are due to *E. faecalis*, though infections resulting from *E. faecium* are more worrisome because of its greater resistance profile [11].

The standard treatment for an enterococcal infection is the dual administration of two antimicrobials; one that targets the cell wall (Vancomycin) and an aminoglycoside (Gentamicin). This dual agent approach is integral due to the failure of either agent to elicit bactericidal activity alone [12]. All enterococci are intrinsically resistant to cephalosporins and commonly harbour resistance to aminoglycosides via a number of target modifying enzymes, most notably the acetyltransferase/phosphotransferase AAC(6')-APH(2") which confers...
resistance to all members of this class, with the exception of Streptomycin. Enterococci are subject to a high frequency of horizontal gene transfer: the transfer of genetic material from one organism to another. This underpins their adaptability in the face of antimicrobial challenge. Resistance to penicillins in these clinical isolates is common by virtue of a low affinity penicillin binding protein (PBP5). Penicillin binding proteins are a ubiquitous enzyme integral in cell wall synthesis, β-lactam antibiotics can bind to these and inhibit their function, however PBP5 has reduced affinity and can therefore evade interaction with the antibiotic and proceed unhindered [13]. Plasmid mediated β-lactamase genes (e.g \textit{blaZ}) have also been detected in enterococci, these genes have been documented to exist alongside aminoglycoside resistance genes and have even resulted in integration to the host chromosome [14, 15]. In the UK during 2006 the prevalence of Gentamicin resistant enterococci isolated from bloodstream infections reached 30.9% in \textit{E. faecium} and 47.9% in \textit{E. faecalis}. In some European countries this figure has exceeded 50% [16].

In 1988 the first plasmid-mediated Vancomycin resistance was detected, threatening the contemporary conventional treatment [17]. Vancomycin resistance is commonly conferred via acquisition of one (or more) of the six \textit{van} resistance genes: \textit{vanA, B, D-F} (\textit{vanC} is intrinsic to \textit{Enterococcus gallinarum} and \textit{Enterococcus flavescens} [18]). Since the discovery of these genes, vancomycin resistance has become a much greater obstacle. In 2006, in the United Kingdom, Public Health England (PHE) investigated vancomycin-resistant enterococci (VRE) rates and determined that of all enterococci tested 32.1% of \textit{E. faecium} and 2.8% of \textit{E. faecalis} isolates displayed vancomycin resistance. Similar rates have been seen across Cyprus, Germany, Greece, Italy and Portugal [16]. More recently, PHE have reported that in England and Northern Ireland in 2016 the overall vancomycin resistance rate of \textit{Enterococci spp.} currently resides at 15%, following annual increases in the preceding four years. The proportion of \textit{E. faecium} taken from blood samples resistant to ampicillin/amoxicillin was 90.2%, vancomycin 22.2%, teicoplanin 23.4% and linezolid 1.1% [9].
The United States suffers much higher rates of VRE than any European country. The 2004 LEADER surveillance program documented vancomycin non-susceptibility in 72.4% of *E. faecium* and 9.6% of *E. faecalis* isolates [19]. In some intensive care units (ICU) in the US, VRE make up ~25% of all enterococci detected [20]. Therapies utilised to combat aminoglycoside, penicillin and vancomycin resistant enterococci include oxazolidinone antimicrobials, such as linezolid, streptogramins, such as a dalfopristin-quinopristin combination (*for E. faecium*), and daptomycin, a lipopeptide [21-23]. However, resistance to these antimicrobials is already emerging.

### 1.2.2: *Staphylococcus aureus*

The Gram-positive methicillin-resistant *Staphylococcus aureus* (MRSA) has been a major focus of both governments and news outlets for decades. *S. aureus* accounts for a large number of skin and soft tissue infections, but also necrotising pneumonia [24], necrotising fasciitis [25] and bacteraemia [26]. In 1960 methicillin was first used as a treatment for *S. aureus* infections and reports of resistance immediately emerged [27].

Cases of bacteraemia caused by MRSA in the United Kingdom rose sharply in the 1990s, with *S. aureus* being responsible for 40-45% of all cases of bacteraemia (~7000-9000 cases pa) in 2005. Following the government-driven implementation of preventative strategies this figure dropped to 36% by 2007. The dramatic rise seen in the 90s was largely caused by two epidemic strains: EMRSA-15 (ST22-MRSA-IV) and EMRSA-16 (ST36-MRSA-II) [28, 29]. In other parts of Europe MRSA rates commonly exceed 25%, reaching as high as 50% in Greece, Cyprus, Portugal and Malta. Conversely, rates in Scandinavian countries run as low as <2% due to effective “search and destroy” policies, as seen in Figure 2. Rates in the United States are higher than all of Europe, with 2004 data from the LEADER surveillance program (50 laboratories from 33 states) finding a prevalence of 54.2% [19], which rose to 58.1% in 2007 [30].
Figure 2 - MRSA isolates from bacteraemias reported in 2007 by countries covered by the EARSS surveillance scheme. Figure adapted from Woodford et al, 2009 [12].

MRSA prevention strategies have been largely successful in reducing the annual number of cases of MRSA bacteraemia. Between 2009 and 2016 the rate of MRSA bacteraemia dropped by 61%, where the rate of Methicillin susceptible S. aureus (MSSA) has risen by 14.2%, as seen in Figure 3. Currently in the UK, the overall S. aureus bacteraemia rate is 18.2 per 100,000 population; with MRSA and MSSA comprising 1.3 and 14.7 cases per 100,000 population respectively [31].

Methicillin resistance originates from a supplementary penicillin binding protein called PBP2a, which has a much lower affinity for β-lactam antibiotics, resulting in a loss of bactericidal activity. The acquisition of this gene results in the resistance to most β-lactams, although novel cephalosporins and carbapenems such as ceftobiprole and ceftaroline still retain high affinities and therefore antimicrobial activity [32, 33].

PBP2a is encoded by the mecA gene, which is located within the larger Staphylococcal cassette chromosome mec (SCCmec). The mecA gene makes up only a small proportion of the overall DNA element, with the rest occupied by mec regulatory sequences, recombinase
genes, insertion sequences and transposons. SCCmec are categorised into ~8 types and subtypes [34-36] with some even having been documented to include entire additional antimicrobial resistance determinants [37-40].

Figure 3 - Rates of bacteraemia per 100,000 population (England, Wales and Northern Ireland) caused by MRSA, MSSA and S. aureus where susceptibility to methicillin was not reported 2009 to 2016. Figure adapted from PHE laboratory surveillance report [31].

The advent of multilocus sequence typing allowed for the accurate typing of MRSA strains. This approach determined that multiple unique sequence types (ST) of MRSA possessed an identical SCCmec, proving the transmissibility of these genetic elements [41]. It has been hypothesised that a close homolog of mecA identified in Staphylococcus sciuri [42], which confers methicillin resistance when transferred to S. aureus [43], may potentially be the original point of acquisition for mecA in S. aureus [12].
While nosocomial MRSA infections are hugely problematic, in recent years an increase in community associated MRSA (CA-MRSA) has been reported in patients with little or no prior exposure to healthcare environments [44, 45]. CA-MRSA is much more common in the US and fairly rare in Europe and Asia. In contrast to epidemic hospital strains which can host multiple resistance determinants against β-lactams, fluoroquinolones and a range of other antimicrobials including macrolide-lincosamide-streptogramin\(_B\) agents (MLS\(_B\) phenotype), tetracyclines and aminoglycosides, CA-MRSA usually only possess β-lactam resistance conferred by a recent acquisition of SCC\(_{mec}\) elements. CA-MRSA are much more likely to express Panton-Valentine leucocidin (PVL), a virulence factor encoded by a prophage integrated into the \(S.\) \(aureus\) chromosome [46]. PVL is an excreted β-pore-forming toxin that perforates host cells, leading to lysis and the release of nutrients utilised by the pathogen. CA-MRSA is usually associated with skin and tissue infections, though PVL+ strains have been documented to progress to necrotising/haemorrhagic pneumonia in young, otherwise healthy, patients; involving high levels of mortality [47, 48].

Traditionally MRSA and CA-MRSA bacteraemias have been treated with vancomycin. However, in recent years strains with documented both reduced susceptibility and complete resistance have emerged in Japan: Vancomycin/Glycopeptide intermediate \(Staphylococcus\) \(aureus\) (VISA/GISA) is the name given to strains that display reduced susceptibility \((8\) mg.L\(^{-1}\)) when compared to breakpoints recommended by the Clinical and Laboratory Standards institute (CLSI) [49]. Reduced susceptibility is achieved by the expression of thicker than usual cell walls, in addition to a greater concentration of glutamine nonamidated muropeptides, leading to a greater number of exposed D-ala-D-ala termini, resulting in the sequestration of vancomycin, resulting in overall population survival [49-54].

True vancomycin resistance in MRSA (vancomycin MIC >128 mg.L\(^{-1}\)) was detected in Michigan, US, in 2002 alongside a vancomycin resistant \(E.\) \(faecalis\) isolate [55]. Since then, additional vancomycin resistant MRSA strains (VRSA) have been documented. While this resistance is due to the acquisition of the \(vanA\) determinant, it has been reported that resistance levels
between these strains varies as a result of the instability of the gene leading to high frequency loss and much slower induction time [56].

1.2.3: *Klebsiella pneumoniae*

This Gram-negative bacillus is a member of the *Enterobacteriaceae* family. *K. pneumoniae* is a commensal species, frequently found in the flora of the mouth, intestines and skin. It is an opportunistic pathogen associated with severe hospital-acquired infections including UTI, septicaemia, pneumonia and soft tissue infections in compromised individuals.

In 2017 PHE documented a rate of *Klebsiella spp.* bacteraemias of 16.7 per 100,000 population, an increase of 8% from the previous year [57] and a 40% increase from 2009, as shown in figure 4.

![Figure 4 - Klebsiella spp. bacteraemia rate per 100,000 population (England, Wales and Northern Ireland). Figure adapted from PHE laboratory surveillance report 2017 [57].](image_url)
Klebsiella spp. are often susceptible to cephalosporins, carbapenems and aminoglycosides; however, in recent years non-susceptibility rates have been increasing. In the UK K. pneumoniae resistance to netilmicin, an aminoglycoside, rose from 13.11% in 2016 to 28.07% in 2017. Resistance to piperacillin/tazobactam, a penicillin/β-lactamase inhibitor combination, has remained stable during this time however, at 13.86%. In 2016 13% of all K. pneumoniae isolates tested by PHE were resistant to ciprofloxacin and third generation cephalosporins [57]. This increase in cephalosporin resistance is due to the acquisition of β-lactamase genes such as blatem, blasvh and blactx-m.

Carbapenems are the last line antibiotic in the case of broadly resistant Gram-negative infections [58]. At present carbapenem resistance in K. pneumoniae remains contextually low, with resistance to ertapenem and meropenem at 1.2% and 0.72% respectively. However, the emergence of the plasmid-borne blakpc (Klebsiella pneumoniae carbapenemase), encoding a serine β-lactamase, is increasing the frequency of resistance and has also been found harboured in a number of other Enterobacteriaceae [59].

1.2.4: Acinetobacter baumannii

This Gram-negative coccobacillus is a rapidly emerging nosocomial threat, especially in intensive care units. Acinetobacter baumannii are associated with bacteraemia, meningitis, pneumonia, UTIs and skin and soft tissue infections [60]. While most strains of Acinetobacter are soil-based organisms, Acinetobacter baumannii is predominantly found in the hospital environment [61].

PHE report that between 2011 and 2017 the overall incidence of Acinetobacter spp. bacteraemia rose from 1.2 to 1.6 cases per 100,000 population, as shown in figure 5. Of all characterised Acinetobacter spp. isolates 31% were identified as Acinetobacter iwoffii and 20% as Acinetobacter baumannii [62]. A. baumannii has been documented to possess a multidrug resistance (MDR) phenotype. The combination of loss of outer membrane proteins (OMPs), increased efflux, topoisomerase mutations and the acquisition of
β-lactamases and aminoglycoside modifying enzymes have made this species highly problematic [63].

Figure 5 - *Acinetobacter spp.* bacteraemia rate per 100,000 population (England, Wales and Northern Ireland): 2010-2017. Figure adapted from PHE laboratory surveillance report [62].

1.2.5: *Pseudomonas aeruginosa*

This Gram-negative species is an environmental organism that can transiently colonise human gut flora. It is an opportunistic nosocomial pathogen that is associated with bacteraemia, endocarditis, exacerbation of cystic fibrosis and more commonly pneumonia, especially in intensive care unit patients. Mechanically ventilated patients are at especially high risk, with the mortality rate reaching as high as 30% in some ICUs [64]. PHE reports that in 2017 the annual rate of *Pseudomonas spp.* bacteraemia was 8.1 per 100,000 population; an increase on previous years as seen in figure 6 [65].
*P. aeruginosa* are intrinsically resistant to a large number of antimicrobials by virtue of synergistic OMP alteration, upregulated efflux capabilities and the chromosomal *ampC* β-lactamase gene that confers resistance to cephalosporin antimicrobials [66]. The treatment of these resistant isolates usually requires carbapenems, however in recent years the acquisition of class B metallo-β-lactamases such as VIM and IMP have seriously compromised treatment options, however colistin still works in most cases [67].

![Graph showing bacteraemia rate per 100,000 population (England Wales and Northern Ireland) 2009 to 2017](image)

**Figure 6** - *Pseudomonas spp.*, *Stenotrophomonas spp.* and related genera bacteraemia rate per 100,000 population (England Wales and Northern Ireland): 2009 to 2017. Figure adapted from PHE laboratory surveillance report 2018 [65].

### 1.2.6: Enterobacter spp

This Gram-negative species contains many commensal strains and are rarely primary pathogens. Ventilated and catheterised patients are particularly vulnerable. Two of the major opportunistic pathogens within this group are *Enterobacter aerogenes* and *Enterobacter cloacae* and are associated with bacteraemia, endocarditis, septic arthritis, osteomyelitis and skin and soft tissue infections [68]. In 2016 PHE reports that the rate of bacteraemias caused...
by *Enterobacter* spp. in England, Wales and Northern Ireland was 3.9 per 100,000, as shown in figure 7 [69].

The broad resistance profile of *Enterobacter* spp. makes their eradication problematic. The presence of the chromosomally encoded AmpC acts as a cephalosporinase which can destroy third generation cephalosporins. In 2016 27% of all *Enterobacter* spp. blood isolates tested by PHE in England and Northern Ireland were resistant to both cefotaxime and ceftazidime; 6% were resistant to the fluoroquinolone ciprofloxacin; 5% were resistant to gentamicin, 8% to tobramycin and 2% to amikacin; 20% were resistant to piperacillin/tazobactam; 2% were resistant to the carbapenem ertapenem, and 1% were resistant to meropenem [69].

![Figure 7 - Enterobacter spp., Serratia spp. and Citrobacter spp. bacteraemia per 100,000 population (England Wales and Northern Ireland): 2009 to 2016. Figure adapted from PHE laboratory surveillance report 2017 [69].](image)

The treatment of cephalosporin resistant *Enterobacter* spp. infections requires the use of carbapenems. However, the emergence of carbapenemase producing *Enterobacteriaceae*
(CPE) are a serious threat. *Enterobacter* spp. have been documented to possess KPC, NDM-1 and IMI-1 [70, 71].

**1.2.7: Escherichia coli**

This versatile Gram-negative member of the *Enterobacteriaceae* is a commensal which is a common inhabitant of human gut flora. It is associated with a number of nosocomial infections such as bacteraemia, meningitis, pneumonia, intra-abdominal infections, UTI, osteomyelitis, endophthalmitis, skin and soft tissue infections [72]. In the UK in 2018 the rate of *E. coli* bacteraemia was 74.3 per 100,000 people, an increase on the 60.4 per 100,000 people seen in 2012 (Fig 8) [73].

![Figure 8 - Trends in the rate of *E. coli* bacteraemia 2012-13 to 2017-18. Figure adapted from PHE laboratory surveillance report 2018 [73].](image)

Pan-resistant *E. coli* strains are regularly being documented and pose a significant risk to public health. The proportion of isolates tested by PHE in 2017 showed that 4.21% were
resistant to a combination of ciprofloxacin, piperacillin/tazobactam and 3rd generation cephalosporins; while 3.66% were resistant to a combination of piperacillin/tazobactam, co-amoxiclav and ciprofloxacin [74].

*E. coli* have been reported to possess β-lactamases such as AmpC, TEM-1, SHV-1, CTX-M and a range of OXAs; carbapenemases such as IMI-1, NDM-1 and KPC; fluoroquinolone resistance in the form of reduced OmpF outer membrane porin production, an increase in antimicrobial efflux, DNA gyrase/topoisomerase IV protection by Qnr and quinolone acetylation by the aminoglycoside -acetyltransferase AAC(6)-lb-cr [72]. In the event of carbapenem resistance in *Enterobacteriaceae* such as *E. coli* the last line polymyxin colistin is used [75]. However, in 2015 a plasmid mediated colistin resistance mechanism (*mcr-1*) was detected in a swine isolate of *E. coli*, effectively threatening the last line antimicrobial for this dangerous pathogen [76].

**1.3 Mechanisms of antimicrobial resistance**

There exist six major mechanisms that facilitate antimicrobial resistance. I. Enzymatic inactivation of the antimicrobial: the destruction of the antimicrobial molecule before it reaches its target II; Target site modification: the alteration of the binding site for the antimicrobial; III. The efflux of antimicrobials: pumping out antimicrobials to lower intracellular concentration; IV. Porin alteration: the reduction of membrane porins to inhibit uptake of extracellular antimicrobials; V. Overproduction of antimicrobial target: upregulation of production of a specific target to overcome the intracellular concentration of antimicrobial; VI. Metabolic target bypass: the presence of an auxiliary enzyme to perform the functions of the one inhibited by the antimicrobial.

**1.3.1: Enzymatic inactivation/modification of antimicrobial**

One of the most widely studied antimicrobial mechanisms is that of enzymatic inactivation/destruction of destruction of the molecule itself. Major examples of antimicrobial modification include acetylation (aminoglycosides, chloramphenicol, streptogramins), phosphorylation (aminoglycosides, chloramphenicol) and adenylation (aminoglycosides,
lincosamides [77, 78]). The modification of these molecules via the addition of these groups results in a decrease of activity, as a result of diminished binding caused by steric hindrance.

Aminoglycoside modifying enzymes (AMEs), a major group within this category, are capable of the covalent modification of the hydroxyl or amino groups of the aminoglycoside molecule. This modification results in diminished interaction with its target site, the 30s ribosomal subunit.

The chloramphenicol acetyltransferases (CATs) are another classic example of antimicrobial modification. Acetylation of this molecule results in the diminished interaction with its target site, the peptidyl transferase region of the 50s ribosomal subunit. Multiple cat genes have been discovered in both Gram positive and negative organisms.

Full inactivation of the antimicrobial is a strategy used in β-lactam resistance. The production of β-lactamases, which destroy β-lactams, results in protection of the organism. β-lactams contain the 4-membered β-lactam ring, which structurally mimics the D-ala-D-ala termini of uncrosslinked peptidoglycan side chains, as shown in figure 9. The introduction of β-lactams to the organism results in the binding of these molecules to penicillin binding proteins (PBPs), which upon catalysis, renders them inert. Due to the critical nature of these enzymes in peptidoglycan synthesis, this has a bactericidal outcome. The presence of β-lactamases destroys the β-lactams before they are able to bind to PBPs, effectively protecting the bacterium.
β-lactam resistance is one of the major health concerns of the 21st century. Many Gram-negative organisms are resistant to multiple groups of antimicrobials, which in a clinical setting forces the use of carbapenems, one of the most potent β-lactams. However, resistance to these antimicrobials is now disseminating rapidly. While some β-lactamases are encoded chromosomally, such as AmpC, a large number are encoded on transmissible genetic elements. To date, the penicillinases TEM-1 and SHV-1, the extended spectrum β-lactamases CTX-M, TEM-3, OXA-11 and the Carbapenemases KPC, OXA-23, OXA-48, IMP, VIM and NDM have all been found located on mobile genetic elements.

1.3.2: Target site modification

The modification of the target site, rather than the antimicrobial, is another key resistance strategy. This can occur in two ways: I. the protection of the target site by an external agent; II. The structural alteration of the target site, reducing binding of the antimicrobial.

Tetracycline resistance determinants TetM and TetO are classical examples of target site protection [79]. In the event of the binding of tetracycline to the bacterial 30s ribosomal subunit, TetM and TetO interact with the ribosome and dislodge the bound tetracycline.

Figure 9 - Structural comparison of the β-lactam penicillin and D-Ala-D-Ala termini of peptidoglycan side chain before cross link.
A second well characterised protection agent is Qnr, a plasmid mediated quinolone resistance determinant. Qnr is a member of the penta-peptide repeat protein family and acts as a DNA homologue, competing for the quinolone binding site of DNA gyrase and DNA topoisomerase IV.

The structural alteration of the antimicrobial binding site can occur in one of two ways: I. the enzymatic modification of the target site and II. the mutation of the target site. Two major examples of resistance by enzymatic target modification are that of the methylation of ribosomal drug binding sites by Erm and Cfr. The erythromycin ribosomal methylation proteins (encoded by \textit{erm}) can mono- or dimethylate an adenine residue in position A2508 of domain V of the 50s ribosomal RNA. This position interferes with the shared binding footprint of 3 major antimicrobials: macrolides, lincosamides and streptogramins (termed the MLS$_B$ phenotype) [80]. Cfr is a radical SAM methyltransferase that is capable of conferring resistance to five classes of antimicrobial that share a binding site adjacent to adenine residue A2503 in domain V of 23s rRNA. Termed the PhLOPS$_a$ phenotype, Cfr confers resistance to phenicols, lincosamides, oxazolidinones, pleuromutilins and streptogramin$_A$ by the methylation of the C8 position of A2503.

One classical example of resistance by mutation is that of rifampicin. Rifampicin is a rifamycin antimicrobial that acts by inhibition of the DNA-dependent RNA polymerase. The rifampicin binding pocket is on the $\beta$ subunit of the RNA polymerase (encoded by the \textit{rpoB} gene) and a single amino acid mutation can result in significantly diminished binding of this agent. A second example of resistance by mutation is that of resistance to the oxazolidinone Linezolid. Linezolid binds the peptidyltransferase region of the 50s ribosomal subunit and mutations in region V of the structural 23s rRNA, and/or mutation in the ribosomal proteins L3 and L4 (described later) can lead to diminished binding.

\textbf{1.3.3: Efflux of antimicrobials}

In order for an antimicrobial to be effective it has to reach critical intracellular concentrations. The presence of efflux pumps can actively extrude antimicrobials from the cytosol. There exist
five major families of efflux pump: I. the major facilitator superfamily (MFS); II. the small multidrug resistance family (SMR); III. the resistance-nodulation-cell-division family (RND); IV. the ATP-binding cassette family (ABC) and V. the multidrug toxic compound extrusion family (MATE) [81].

A classic example of efflux mediated antimicrobial resistance is that of the AcrAB-TolC system found in *Enterobacteriaceae*. AcrAB-TolC is a tripartite complex of the RND family that is widely disseminated in Gram-negative bacteria. AcrAB-TolC acts as a proton antiporter that can extrude a range of molecules, including tetracyclines, chloramphenicol, fusidic acid, novobiocin and fluoroquinolones [82].

1.3.4: Porin alteration

Porins are proteins that span the cell membrane and facilitate transport of small molecules. They are integral to nutrient acquisition but are also a major conduit through which antimicrobials can pass, gaining entry to the microbe. These are especially important to hydrophilic molecules such as β-lactams and tetracyclines which utilise these water-filled diffusion channels to enter the bacterial cell [83].

A shift in porin type, downregulation of porin production and impairment of porin function are the three main methods by which porin alteration can result in a decrease in intracellular antimicrobial concentration. A classic example of porin downregulation is OprD found in *P. aeruginosa*. A report by Quinn *et al* displayed that while initially susceptible, during therapy the downregulation of porins alone gave resistance to imipenem [84]. A classic example of porin shift can be seen in *K. pneumoniae*: Isolates have been detected to have shifted from production of OmpK35 to OmpK36 after challenge by a range of β-lactams, which resulted in a 4-8-fold decrease in susceptibility. It was shown that OmpK36 possesses a smaller porin size, which cannot accommodate certain β-lactams [85, 86].
1.3.5: Overproduction of target

In some circumstances overwhelming the antimicrobial concentration with an excess of target can be an effective way to overcome its activity. An example of this can be seen in resistance to trimethoprim/sulfamethoxazole (TMP-SMX). Most bacteria are unable to import folate from environmental sources and therefore must produce their own, this is the target of TMP-SMX. In order to create purines and some important amino acids folate must be available. TMP-SMX inhibit two enzymes in this integral pathway: dihydropteroic acid synthase (DHPS) and dihydrofolate reductase (DHFR). A mutation in the promoter region encoding these two genes can lead to overproduction which can therefore overcome the TMP-SMX concentration.

1.3.6: Target bypass

By producing new targets which can perform the role of the original but have no/less interaction with the antimicrobial is a powerful mechanism for resistance. Two major examples in this category are the replacement PBPs used by MRSA to protect against penicillin, and the vancomycin resistance conferred by the acquisition of vanA.

As previously discussed β-lactams target the penicillin binding proteins that are integral in peptidoglycan synthesis. These native PBPs perform the transpeptidation and transglycosylation of emergent peptidoglycan molecules. The acquisition of an SCCmec element which encodes the auxiliary PBP2a provides the bacterium with an additional, β-lactam insensitive, PBP able to perform the transpeptidase function of the native enzymes, and consequently confers resistance to penicillins, cephalosporins (up to the most recent generation) and carbapenems. However, the absence of a transglycosylase site within PBP2a highlights the requirement for this function to be performed by another PBP.

Similar to β-lactams, glycopeptides enact bactericidal activity by the inhibition of cell wall synthesis. Glycopeptides do not interact with PBPs however, instead they bind to the exposed D-Ala-D-Ala termini of uncrosslinked peptidoglycan side chains, blocking their crosslinking. The acquisition of van gene clusters can confer resistance to glycopeptides such as
vancomycin and teicoplanin. van clusters encode an array of genes that can remodel the synthesis of peptidoglycan by both changing its terminal D-ala for a D-lactate and destroying existing D-Ala-D-Ala, removing the binding site for glycopeptides.

1.4 The ribosome as an antimicrobial target site

1.4.1: Structure and function of prokaryotic ribosome

The ribosome is an essential ribonucleoprotein responsible for protein synthesis, which due to its critical function is the target of many antimicrobials. Bacterial ribosomes are comprised of three ribosomal RNAs (rRNAs: 5S, 16S and 23S) and ~52 proteins, with the majority of the functional regions being dominated by rRNAs, as shown in figure 10.

**Figure 10 – The *Staphylococcus aureus* ribosome.** The prokaryotic 70S ribosome is comprised of two main subunits: the 30S and the 50S. The 50S unit consists of 23S rRNA (red), 5S rRNA (green) and a number of structural proteins (blue). The 30S unit consists of 16S rRNA (orange) and structural proteins (blue). A. The ribosome in profile orientation. B. The ribosome in side orientation, displaying the mRNA hosting aperture. Figure adapted from Khusainov *et al.*, 2016 [89] Image created by author on Pymol.
Protein synthesis can be broken down into four steps: Initiation, elongation, termination and recycling. Initiation requires the 30S and 50S subunits to associate to form the 70S ribosome, as well as the correct positioning of the messenger RNA (mRNA) against the initiator tRNA in the ribosomal P site. Elongation requires an aminoacylated tRNA to be delivered by Elongation factor Tu (Ef-Tu) to the A site, upon where peptide bond formation occurs between the two tRNA-attached amino acids in the P and A sites. To accommodate further tRNAs translocation of the amino acids residing in the P and A sites occurs, moving them to the P and E sites. This is facilitated by EF-G. As this process repeats, the nascent peptide chain leaves the ribosome through an exit channel in the 50S subunit until the ribosome encounters a stop codon on the mRNA. Stop codons are recognised by termination release factors RF1 and RF2 which hydrolyse the peptidyl-tRNA bond, releasing the peptide chain into the cytosol. Once this process has occurred, the 50S and 30S subunits dissociate and are recycled into further synthesis reactions [90]. This process is highlighted in figure 11.
Figure 11 – Ribosomal protein synthesis pathway and antimicrobial interactions.

Initiation starts with the formation of a 30S and 50S to form a 70S ribosome with the mRNA start codon and the correct tRNA aligned. This process is inhibited by the antibiotics edeine (Ede), kasugamycin (Ksg), pactamycin (Pct) and thermorubin (Thb) on the 30S subunit, and by the orthosomycins avilamycin (Avn) and evernimicin (Evn), as well as thiostrepton (Ths) on the 50S subunit. Next an aminoacylated tRNA is delivered to the to the A site of the ribosome by EF-Tu bound to GDP. This step can be inhibited by streptomycin (Stp), tetracyclines (Tet) and glycylcyclines (tigecycline (Tig)). Peptide bond formation is catalysed between the two tRNA bound amino acids in the P and A regions. This can be inhibited by blasticidin S (BlS), chloramphenicol (Cam), lincosamides (clindamycin (Clm)), oxazolidinones (linezolid (Lnz)), pleuromutilins (Plu), puromycin (Pmn), streptogramin A (Sₐ) and sparsomycin (Spr). Translocation of the tRNAs bound in the P and A region is catalysed by EF-G bound to GDP, moving them to the E and P sites. This step is inhibited by tuberactinomycins capreomycin (Cap) and viomycin (Vio), the aminoglycosides hygromycin B (HygB), neomycin (Neo) and
paromomycin (Par), as well as fusidic acid (Fus), spectinomycin (Spt). Elongation of the nascent peptide chain can be inhibited by the macrolides (erythromycin (Ery)), streptogramin B (Sb) and ketolides (telithromycin (Tel)). The final step is termination, which can be inhibited with certain inhibitors such as Blasticidin S (BLS) and translocation inhibitors such as Fusidic acid. Figure reproduced with permission [90].

1.4.2: Sites integral to antimicrobial activity

The ribosome is considered to be the most important target site for antimicrobials. More antimicrobials target the ribosome than any other region [90]. Figure 11 displays the range of antimicrobial targets throughout ribosomal protein synthesis. Many of the targets of these antimicrobials centre around similar positions. Antimicrobials that target the 30S subunit mostly target sites that interact with the mRNA and tRNA, as shown in figure 12.A. Aminoglycosides have a range of activities when binding to varied sites within the 30S subunit: Kasugamycin can inhibit translation initiation by the prevention of a stable interaction between the initiator tRNA and the start codon at the P-site; Streptomycin can interfere with the delivery of tRNAs the A site (as can tetracycline); neomycin, spectinomycin and pactamycin can inhibit the translocation of the mRNA-tRNA complex through the ribosome.

Antimicrobials that target the 50S subunit often have binding sites that cluster around the site where peptide bonds are formed: the peptidyltransferase (PTC) region, as shown in figure 12.B. PTC bound antimicrobial effect their activity by preventing correct positioning of the tRNA during peptide bond formation. Chloramphenicol, lincosamides and oxazolidinones occlude the A-site tRNA; blasticidin S occludes the P-site tRNA; and pleuromutilins and streptogramin a occlude both. Macrolides and streptogramin b occupy the exit tunnel adjacent to the PTC and do no inhibit peptide bond formation, instead they inhibit the exit of nascent peptides leading to abortion of the process.
A. 30S subunit antimicrobial binding sites

B. 50S subunit antimicrobial binding sites
Figure 12 - 30S, 50S and PTC antimicrobial binding sites. A. Overview of 30S binding sites clustered around the mRNA binding channel, including tetracycline (Tet), spectinomycin (Spt), kasugamycin (two known binding sites: Ksg1 and Ksg2), pactamycin (Pct), edeine (Ede), hygromycin B (HygB), neomycin (Neo), streptomycin (Str), thermorubin (Thb) and tuberactinomycins (Tub). B. Overview of the binding sites of neomycin (neo), evernimicin (Evn) and thiostrepton (Ths) in relation to the 50S ribosome. The A-site tRNA (green), P-site tRNA (blue), E-site tRNA (orange), H43/H44, H69, peptidyl-transferase centre (PTC) and the L1 and L11 stalks are highlighted for reference. C. Enlargement of the peptidyltransferase region antimicrobial binding sites: blasticidin S (Two known binding sites Bls1 and Bls2), sparsomycin (Spr), lincomycin (Lin), linezolid (Lnz), macrolides (Mac), puromycin (Pmn), pleuromutilins (Plu), chloramphenicol (Cam), streptogramins A and B (Sₐ and Sₐ) in relation to the A-site and P-site tRNAs. Figure reproduced with permission [90].
1.5 *cfr*, *rlmN* and the Radical SAM Methyltransferase superfamily

1.5.1: Cfr and RlmN

The chloramphenicol/florfenicol resistance gene (*cfr*) was first discovered in a bovine strain of *Staphylococcus sciuri* in Germany, 2000 [91]. This gene encodes a radical S-adenosyl methionine (SAM) methyltransferase (RSMT) which confers resistance to eight chemically distinct classes of antimicrobial via the methylation of 23S rRNA residue adenine 2503 (A2503). This residue, once methylated at the C8 position, overlaps with the binding site of Phenicols (Chloramphenicol and Florfenicol), Lincosamides (Clindamycin), Oxazolidinones (Linezolid), Pleuromutilins (Tiamulin and Valnemulin) and streptogramin A’s (Pristinamycin II, Virginiamycin M and Dalfopristin) [92]. This resistance profile has been termed the PhLOPS\textsubscript{A} phenotype. Recently, Cfr has been shown to confer resistance to three additional classes of antimicrobial including large macrolides [93], Hygromycin B and A201A [94].

*rlmN* encodes the RlmN housekeeping methyltransferase which shares 32% sequence homology with Cfr [95]. RlmN functions to methylate the C2 position of A2503 and tRNA residue A37 (m\textsuperscript{2}A) [96]. Methylation of the A2503 C2 position confers no resistance phenotype, unlike the methylation of C8 catalysed by Cfr, as shown in figure 13. The function of A37 tRNA methylation remains poorly understood, however methylation of 23S rRNA residue A2503 results in an increase of mRNA stop codon recognition by the ribosome and an improvement of translational proofreading [97]. It has been shown that while RlmN is capable of C2 methylation, Cfr is capable of both C2 and C8 methylation [96, 98]. This bifunctional mechanism is unusual in the context of radical SAM enzymes, which usually only target one site.

1.5.2: Radical SAM superfamily

Both Cfr and RlmN belong to a methyltransferase subgroup of the radical SAM superfamily of enzymes. These enzymes utilise SAM molecules as a substrate for the generation of radicals to drive catalysis. The radical SAM superfamily is characterised by the canonical CX\textsubscript{3}CX\textsubscript{2}C motif found in the active site (where C denotes Cysteine and X denotes no specific amino
acid), as displayed in figure 18. This tri-cysteine motif ligates a [4Fe-4S] cluster via interactions between the Fe$^{2+}$ and the sulphur of the cysteine, as shown in figure 15. The clusters fourth unbound Fe$^{2+}$ (termed the “unique iron”) is the component that interacts with the SAM substrate.

![Figure 13 – Methylation states of 23S rRNA residue A2503 resulting from RlmN and Cfr catalysis. A. Unmethylated A2503. B. C2 methylated A2503 catalysed by RlmN. C. C2 and C8 methylated A2503 catalysed by Cfr.](image)

Enzymes within the radical SAM superfamily utilise this radical to catalyse a wide range of reactions such as thymine dimer repair, biosynthesis of vitamins, coenzymes and antibiotics; and sulphur insertion by biotin and lipoyl synthases [99, 100]. SAM dependent enzymes are responsible for a wide range of reactions, however radical SAM enzymes specifically reflect an intriguing branch of enzymology due to their ability to catalyse reactions at inherently unreactive positions [101, 102].
1.5.3: Radical SAM methyltransferases (RSMTs)

RSMTs are unique in the context of the radical SAM superfamily. All radical SAM enzymes cleave S-adenosyl methionine to methionine and a 5’ Deoxyadenosine radical (DOA), the latter product being used to catalyse reactions. However, the RSMT reaction requires two molecules of SAM to complete a methylation reaction: the first as a methyl donor, producing a S-adenosyl homocysteine by-product and the second for radical formation.

The first SAM molecule is bound via an interaction between the unique iron and the amino acid moiety within the SAM substrate [103] at which point an Sn2 displacement reaction adds the SAM methyl group on to a conserved cysteine (Cys355 in RlmN and Cys338 in Cfr) in the proximity of the [4Fe-4S] cluster [104]. This initial auto-methylation step has been proven essential to the subsequent DOA radical formation [103]. The second SAM molecule is bound in the same position and initiates the reduction of [4Fe-4S] cluster from a 2+ state to a 1+ state via a reducing agent due to the proximity of the sulfonium ion within the SAM and the [4Fe-4S] cluster itself. The reduced cluster then cleaves this second SAM molecule at the adjacent 5’ C-S bond, producing methionine and the DOA radical.
Figure 14 – Methylation of Cfr Cys338. The CH$_3$ group is relocated from the sulphur of SAM to the sulphur of Cys338 via an SN$_2$ displacement reaction, yielding S-adenosylhomocysteine.

Figure 15 – Formation of 5’ methionine and DOA radical from SAM. As the second SAM molecule binds the unique Fe$^{2+}$ of the [4Fe-4S] cluster, it is reductively cleaved to Methionine and a 5’-Dexoyadenosine radical (5’dA•) which is used to catalyse methyl transfer to unreactive sites.
Current understanding of the methylation mechanism largely originates from studies performed on RlmN. The proposed C2 RlmN methylation model by Grove et al [105] states that the first molecule of SAM (SAM1) binds to the unique Fe$^{2+}$ of the [4Fe-4S] cluster, undergoing a standard $S_N2$ displacement reaction to abstract the CH$_3$ group from the sulphur of SAM1 to the sulphur of Cys355, yielding S-adenosylhomocysteine. A second SAM molecule (SAM2) is bound at the same position and reductively cleaved to methionine and a 5’Deoxyadenosine radical (5’dA•). In the presence of its RNA substrate this radical is used to abstract a hydrogen from the CH$_3$ group hosted on Cys355, resulting in a methylene radical. This radical attacks the adenine ring of rRNA residue A2503 at position C2, resulting in the transient relocation from C2 to another base in its proximity (potentially Glu105). This step of the reaction results in a covalent interaction between RlmN and its RNA substrate. A second active site cysteine (Cys118) reductively catalyses the cleavage of the S-CH$_2$ bond adjoining RlmN to the RNA substrate, resulting in RNA release from the complex and the formation of a S-S disulphide bond between the two cysteines. The original C2 hydrogen, currently located on a proximal base is relocated to the CH$_2$ group, resulting in the fully methylated product, as shown in figure 16 [104, 105].
Figure 16 – Methylation of rRNA residue A2503 position C2 by RlmN. A methyl group is donated to Cys355 by SAM1, yielding SAH. SAM2 is cleaved by RlmN to methionine and a 5'-Deoxyadenosine radical (dAdo•). This radical abstracts a hydrogen from the Cys355 bound methyl group creating a CH₂• methylene radical. This radical attacks the C2 position resulting in a S-CH₂ covalent interaction between RlmN and its RNA substrate. The C2 hydrogen is simultaneously relocated to a proximal base. A second cysteine (Cys118) reductively catalyses the cleavage of the S-CH₂ bond, releasing the RNA from the complex and forming an S-S bond with Cys355. The hydrogen originating from position C2 is simultaneously relocated to the CH₂ group at position C2, resulting in the fully methylated product [104, 105]. Figure reproduced with permission: https://pubs.acs.org/doi/10.1021/ar200202c.

The C8 methylation mechanism catalysed by Cfr remains poorly understood. All attempts to crystallise Cfr have resulted in failure and all structural data formulating any potential mechanism originate from RlmN. One potential difference that could explain the dual C2-C8 specificity is the presence of an additional cysteine (Cys 110) in the active site. This extra cysteine is present in all of the Cfr and Cfr-like enzymes that can confer the PhLOPS₄ phenotype and could potentially be significant in C8 methylation.
1.6 Dissemination and Clinical Impact of cfr

1.6.1: Cfr in a clinical environment

After first being documented in a bovine strain of *Staphylococcus sciuri* in 1997 (reported first in 2000) [91] it took another five years to be detected in a clinical isolate. In 2005 a patient in Medellin, Colombia was shown to possess a Methicillin-resistant *Staphylococcus aureus* isolate designated CM05 [106]. Since then, *cfr* has been discovered in countries worldwide including Belgium, Brazil, Canada, China, Colombia, Germany, India, Italy, Ireland, Thailand, Spain, The United States and The United Kingdom [91, 106-131], as shown in Figure 17.

![World map displaying locations of cfr discovery to date](image)

**Figure 17 – World map displaying locations of cfr discovery to date.** cfr has been located in Brazil, Canada, China, Colombia, Germany, India, Italy, Ireland, Thailand, Spain, The United States and The United Kingdom.

To date, *cfr* has been predominantly found in Gram-positive isolates, most notably *Staphylococcus spp* [109, 112, 132-136]. The acquisition of *cfr* by clinical strains of *S. aureus* and *S. epidermidis* with resistance to linezolid have resulted in at least four nosocomial outbreaks in the past decade [108, 110, 119, 137]. Recently, *cfr* has been detected in a known pandemic sequence type of MRSA (ST22-MRSA-IV) in Ireland [138]. As of yet these
cfr-positive clinical isolates have not been associated with outbreaks of resistant S. aureus infections.

Though the majority of reports of cfr discovery focus around Staphylococcus spp. a number of other cfr-positive Gram-positive isolates have also been documented. In China at least three reports of cfr-harbouring porcine Bacillus spp. isolates from food producing swine farms have emerged [114, 125, 129]; Macrococcus caseolyticus and Jeotgalicoccus pinnipedialis isolates have also been discovered with cfr [126]. Concerningly, a number of vancomycin resistant Enterococci have been detected to harbour cfr [113, 116, 123], with the first reported cfr-mediated linezolid-resistant VRE outbreak occurring in the United Kingdom in 2017 [118].

To date at least two Gram-negative genera have also been discovered to possess cfr: Escherichia coli and Proteus spp. In both cases the isolates have been of veterinary origin and harbour a portfolio of other resistance determinants. In one circumstance an E. coli isolate containing cfr, tetM, blaTEM-1, floR and QnrS1 was reported [122]; additional reports highlight E. coli containing cfr alongside blaCTX-M [128]. cfr has been identified in porcine isolates of Proteus spp. twice in the last decade, initially found in a P. vulgaris isolate in 2011 and secondly in a P. mirabilis isolate which also harbourd blaCTX-M, fosA3 and aac(6')-Ib-cr [120].

1.6.2: Cfr homologues

Since the initial discovery of cfr, a second clinical Cfr-like enzyme has been discovered. Designated Cfr(B), this new homologue was discovered within the same year in a clinical strains of Enterococcus faecium [139, 140] and Peptoclostridium difficile (formerly Clostridium difficile) [141]. While many of the cfr isoforms discovered have possessed minor deviations from the parent sequence (<1%), Cfr(B) shares only 74.9% sequence identity with the Staphylococcus cfr.

Sequence database comparisons using the Basic Local Alignment Search Tool (BLAST [142]) have yielded a number of Cfr-like homologs from non-pathogenic Bacillales (Bacillus
amyloliquefaciens, Bacillus clausii, Brevibacillus brevis) and Paenibacillus sp. Y412MC10 which when transformed in to E. coli can confer the same PhLOPS$_A$ resistance phenotype conferred by Cfr [143, 144]. This highlights the possibility that cfr-like genes are a natural defence mechanism found in extant environmental bacterial populations.

1.6.3: Antimicrobial genes associated with cfr

While cfr alone is capable of conferring the PhLOPS$_A$ phenotype, it is commonly found in strains that also harbour a range of other resistance mechanisms.

OptrA is a membrane efflux pump of the ABC-F subfamily of ABC pumps capable of lowering intracellular concentrations of chloramphenicol, florfenicol, linezolid and tedizolid. OptrA was originally found on an Enterococcus faecalis strain of both human and animal origin [145], but has since been discovered alongside cfr in porcine strains of Staphylococcus sciuri [146] and in two clinical cases in Italy where the cfr gene was not expressed [113].

FexA is a florfenicol exporter found in both Gram-positive and -negative organisms. Originally discovered in a veterinary isolate of Staphylococcus lentus it has since been discovered at least twice in other animal isolates alongside cfr [114, 147].

FloR is a florfenicol exporter that can cause problems in the veterinary field, originally discovered in an animal strain of Salmonella typhimurium [148]. Due to its good tissue penetration florfenicol is commonly used to treat bacteraemia in food-producing animal neonates. FloR has been detected alongside Cfr in porcine isolates of E. coli [122].

Ribosomal mutation is a common mechanism for resistance against antimicrobials that target the peptidyl transferase centre. The antimicrobial binding pocket of the PTC contains 23s rRNA conserved residues G2061, A2451, C2452, A2503, U2504, G2505, U2506, and U2585 which interact directly with linezolid [149]. Strains that are selected for linezolid resistance often display mutations at residues G2061, C2452, A2503, U2504, and G2505 which are directly adjacent to the linezolid molecule. The most common residue mutation detected is that of G2576U which has been isolated in clinical linezolid-resistant strains of Staphylococcus.
and *Enterococcus* [150, 151]. Mutations in the L3, L4 and L22 protein regions have also been linked with linezolid resistance. L3 is positioned within the 50S subunit but has a loop that extends into the PTC [152]. Mutations in L4 and L22 have been linked to reduced susceptibility to PTC acting antimicrobials [153]. Mutations in these positions have been discovered in clinical isolates that also contain *cfr* [154].

1.7 Aims of project

In order to investigate the relationship between sequence, structure and function in the Cfr/RlmN family, in particular with regard to the basis for Cfr-mediated C8 methylation, three Cfr homologs were selected alongside the *S. sciuri* cfr: *cfr(B)*, a *cfr-like* gene from *Paenibacillus* (PbCfr) whose sequence resides between Cfr and RlmN. Of particular interest, SbCfr, the Cfr/RlmN homologue from the thermophilic Gram-positive organism *Sphaerobacter thermophilus* (SbCfr), contains a fifth cysteine in the active site as shown in Figure 18, but in a different location to Cfr Cys118. A full alignment of all Cfr and Cfr-like targets can be seen in appendix 1.

\[
\begin{align*}
\text{Cfr} & \quad \text{CISSQCGCNFGCKFC} \\
\text{RlmN} & \quad \text{CVSSQVGCALECKFC} \\
\text{Cfr (B)} & \quad \text{CISSQCGCNFGCKFC} \\
\text{SbCfr} & \quad \text{CVSCQVGCAVGCSCF} \\
\text{PbCfr} & \quad \text{CISSQCGCGFGCTFC}
\end{align*}
\]

*Figure 18 – Active site sequence alignment between candidate enzymes Cfr, RlmN, Cfr(B), SbCfr and PbCfr.* Cfr, Cfr(B), SbCfr and PbCfr all possess the active site fifth cysteine that appears synonymous with the PhLOPS\(_A\) phenotype. Notably the fifth cysteine in SbCfr is located two positions forward of those seen in the other sequences.

This project aims to evaluate Cfr and selected homologues for biochemical, and particularly structural, characterisation. These genes will be expressed recombinantly and the proteins purified to homogeneity and subjected to investigation by structural, biochemical and
biophysical methods. In order to confirm their status as functional radical SAM methyltransferases these purified enzymes will be subjected to activity assays that track the generation of SAH during auto-methylation and the formation of 5’Deoxyadenosine from the second step of the reaction. Proteins identified as stable and active, and able to be produced in sufficient amounts, will be subjected to crystallography trials in an effort to resolve structures of Cfr-like enzymes that will shed light on the possible mechanism of C8 methylation.
Chapter two: Methods and Materials

2.1: Selection of candidate genes

Previous work performed by undergraduate student Jake Finan had identified two thermophilic cfr homologs originating in Paenibacillus spp. (Accession number: CP001793) and Sphaerobacter thermophilus (Accession number: AJ420142). Wildtype Staphylococcus sciuri cfr was used as a positive control throughout this project.

2.2: Vector construction

pCDF-Duet1 (Streptomycin\textsuperscript{R}) was the chosen vector for protein overexpression and subsequent purification by IMAC. Containing two multiple cloning sites under the control of Isopropyl β-D-1-thiogalactopyranoside (IPTG), this vector was chosen to allow both flexibility should the inclusion of a second target gene be required and a CFD origin of replication which would not conflict with other co-transformed vectors. Vectors were linearised with HindIII-HF\textsuperscript{®} and Ncol-HF\textsuperscript{®} restriction enzymes in CutSmart\textsuperscript{®} buffer at 37\degree C for 15 minutes and gel purified (see agarose gel section).

pBAD A (Ampicillin\textsuperscript{R}) is the vector chosen for MIC assessments. Using a titratable arabinose-controlled promoter (P\textsubscript{BAD}), which can be further repressed in a low glucose environment, the expression of target genes can be tightly controlled. pBAD utilises a pBR322 origin of replication.

pACYCT2 (Chloramphenicol\textsuperscript{R}) is the vector utilised for rRNA:protein complex capture. Previous investigation into RNA-Protein co-expression strategies by Ponchon \textit{et al} have shown that attempts to co-express both RNA and protein genes each under the control of a T7 promoter can result poor yield. pACYCT2 was chosen because it possesses a tac promoter, which when used in conjunction with a second plasmid utilising a T7 promoter can restore normal expression levels of both RNA and protein [155].
2.3: PCR amplification

Polymerase chain reaction (PCR) was used to generate inserts with flanking regions homologous to target vector linearised ends, for use with the In-Fusion® system (Takara Bio USA). All reactions were carried out in 0.2mL reaction tubes (Eppendorf) in an Alpha Cycler 2 (PCRMax), using Phusion® polymerase (NEB) according to the manufacturer’s instruction. Table 1 documents the primers used to generate plasmid inserts. Cycling parameters: 95°C denaturation for 5 minutes, the 35 cycles of 95°C denaturation for 10 seconds, X°C primer binding (dependent on polymerase) for 10 seconds and 72°C primer extension for 10 seconds, followed by 72°C for 5 minutes final extension.
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Primers</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCDF-SbCfr</td>
<td>F: 5’ - AGG AGA TAT ACC ATG GCA CAT CAC CAT CAC C – 3’</td>
<td>sbcfr inserted in to pCDF-Duet1 MCS site 1 between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td></td>
<td>R: 5’ - AT GCG GCC GCA AGC TTA GGC ACC CGC AAC G – 3’</td>
<td></td>
</tr>
<tr>
<td>pCDF-PbCfr</td>
<td>F:</td>
<td>pbcfr inserted in to pCDF-Duet1 MCS site 1 between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td></td>
<td>R:</td>
<td></td>
</tr>
<tr>
<td>pCDF-Cfr(B)</td>
<td>F: 5’ AGG AGA TAT ACC ATG GCA CAT CAC CAC C 3’</td>
<td>cfr(B) inserted in to pCDF-Duet1 MCS site 1 between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td></td>
<td>R: 5’ ATG CGG CCG CAA GCT TTA CTG GCT GTT CTG TTT CT 3’</td>
<td></td>
</tr>
<tr>
<td>pCDF-Cfr(C)</td>
<td>F: 5’ AGG AGA TAT ACC ATG GCG CAT CAC CAT C 3’</td>
<td>cfr(C) inserted in to pCDF-Duet1 MCS site 1 between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td></td>
<td>R: 5’ ATG CGG CCG CAA GCT TTA TGA GGA ATT GGT CGA AC 3’</td>
<td></td>
</tr>
<tr>
<td>pBAD-SbCfr</td>
<td>F: 5’ GAG GAA TTA ACC ATG GCA CAT CAC CAT CAC C 3’</td>
<td>sbcfr inserted in to pBAD-A between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td></td>
<td>R: 5’ CAA AAC AGC CAA GCT TTA GGC ACC CGC AAC 3’</td>
<td></td>
</tr>
<tr>
<td>pBAD-Cfr(B)</td>
<td>F: 5’ GAG GAA TTA ACC ATG GCA CAT CAC CAC C 3’</td>
<td>cfr(B) inserted in to pBAD-A between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td></td>
<td>R: 5’ CAA AAC AGC CAA GCT TTA CTG GCT GTT CTG TTT C 3’</td>
<td></td>
</tr>
<tr>
<td>pBAD-Cfr(C)</td>
<td>F: 5’ GAG GAA TTA ACC ATG GCG CAT CAC CAT C 3’</td>
<td>Cfr(C) inserted in to pBAD-A between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td></td>
<td>R: 5’ CAA AAC AGC CAA GCT TTA TGA GGA ATT GGT CGA AC 3’</td>
<td></td>
</tr>
<tr>
<td>pMC115</td>
<td>-</td>
<td>cfr inserted in to pCDF-Duet1 MCS site 1 between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td>pMC116</td>
<td>-</td>
<td>cfr inserted in to pBAD-A between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td>pMC118</td>
<td>-</td>
<td>cfr with its natural promoter inserted in to pBAD-A between restriction sites Ncol and HindIII</td>
</tr>
<tr>
<td>pDB1282</td>
<td>-</td>
<td>iscS-iscU-iscA-hscB-hscA-fdx portion of the Isc operon from Azotobacter vinelandii encoding iron sulphur cluster assembly proteins hosted on a pARA13 backbone (Gift of Dr. Dennis Dean, Virginia Tech University)</td>
</tr>
<tr>
<td>pACYCT2-Sb23S</td>
<td>F: 5’ AAG GAG ATA TAC ATA GAA TTC GCC CGG ATA G 3’</td>
<td>S. thermophilus rRNA residues G2522 – C2701 hosted within a tRNA scaffold inserted in to pACYCT2 inserted between Ndel and Xhol (Addgene item ID: 45799)</td>
</tr>
<tr>
<td></td>
<td>R: 5’ CTT TAC CAG ACT CGA AAG CTT AAA AAA AAT CCT TAG CTT T 3’</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 – Table of vectors and primers used during plasmid generation.
2.4: Colony PCR

Confirmation of successful transformation was performed by colony PCR. Colonies of the transformed cell lines were picked and resuspended in 100µL Elgastat water. The resuspended cells were heated on a heat block (Grant QBT2) at 94°C for 3 minutes. This solution was centrifuged at 10,000 x g for one minute and 2µL of supernatant used as template for PCR using the MyTaq mix (Bioline).

2.5: Cell lines (incl competent cell generation)

Chemically competent cells were generated by incubation of a 1:50 inoculation of overnight culture in 50mL LB broth (containing antibiotic if appropriate) until ~0.4 OD600 nm, followed by incubation on ice for 30 minutes. Cells were centrifuged at 4000 rpm for 10 minutes, the supernatant removed, and the pellet resuspended in 30 mLS Transformation Buffer I (TFBI): 30 mM KOAc, 100 mM (RbCl), 10 mM CaCl₂.2H₂O, 50 mM MnCl₂, 3 mM Hexamine cobalt chloride, 15% Glycerol (v/v), pH 5.8. After 15 mins incubation on ice, cells were again pelleted at 4000 rpm and the supernatant removed. Pellets were then gently resuspended in 2 mL Transformation Buffer II (TFBII): 10 mM MOPS, 10 mM RbCl, 75 mM CaCl₂.2H₂O, 15% Glycerol (v/v), pH 6.5. This resuspension was left to incubate on ice for 15 mins, followed by division into aliquots of 100 µL, which were flash frozen in liquid nitrogen and stored at -80°C.

2.6: Recombinant protein expression strains

BL21* E. coli possesses a rne131 gene mutation which results in diminished mRNA degradation and therefore increased translation and protein yield. SOLUBL21 E. coli has been shown to increase the solubility of insoluble proteins during production. Arctic Express E. coli possess cold-adapted Cpn10 and Cpn60 chaperonins from the psychrophilic organism Oleispira antarctica which facilitate protein folding at low temperatures.
2.7: Minimum inhibitory concentration strains

**BW-RlmNΔ E. coli**, a Keio collection strain [156] which possesses a *rlmn* deletion and a kanamycin^R^ gene insertion at that position; **BW-J396-RlmNΔ E. coli**, a second *rlmn* mutant strain, with a silenced RlmN gene and insertion of a kanamycin^R^ gene at a different position; and **BW25113 E. coli**, the parent strain for all BW *E. coli* used in MIC investigations. All MIC strains were a gift from Dr Martin R. Challand.

2.8: Agarose gel electrophoresis

DNA products were assessed by agarose gel electrophoresis. 1% agarose (Bioline) gels were used alongside 1xTAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) with 0.5 µg.mL ethidium bromide. 10 µL samples were loaded onto the gel with Buffer Blue (Bioline) and 5µL DNA 1kb Hyperladder (Bioline). Gels were run at 100V for ~45 minutes.

2.9: Transformation

All cells were transformed using a heat-shock method. 75-100 µL competent cells were incubated on ice for 30 mins with 1-100 ng of DNA. Cells were then incubated at 42ºC in a heat block (Grant QBT2) for 50 seconds exactly, then returned to ice for 2 minutes incubation. 250 µL SOC medium (pre-warmed at 37ºC) was added to each aliquot of cells, which were then placed in a rotating incubator (New Brunswick Scientific Nova 44) at 180 rpm for 1-2 hours at 37ºC. Cells were recovered and pelleted at 8000 x g (Labnet Spectrofuge 16M) for one minute and all supernatant but for 100 µL removed. The remaining 100 µL was used to resuspend the pellet, followed by plating on LB agar with appropriate selection antibiotics and grown overnight at 37ºC.

2.10: Plasmid purification

Colonies were picked from agar plate to generate cultures of 10mL LB broth containing appropriate antibiotic. These cultures were incubated in a rotating incubator (New Brunswick Scientific Nova 44) 180 rpm at 37ºC overnight. These cultures were recovered and purified using the Thermo scientific GeneJet plasmid mini prep according to the manufacturers’ instruction.
2.11: Protein expression trials

Protein expression *E. coli* strains were transformed with a pCDF Duet-1 vector containing the gene of interest and pDB1282, which possesses the iron-sulphur generating *iscS-iscU-iscA-hscB-hscA-fdx* portion of the Isc operon from *Azotobacter vinelandii* and has been demonstrated to improve expression and recovery of [4Fe-4S] cluster-containing proteins in *E. coli* [157]. These strains were inoculated in triplicate into 10 mL LB broth containing 100µg.mL ampicillin and 50µg.mL streptomycin and grown at 37°C, 180rpm, until ~0.6 OD600

nm, at which point pDB1282 was induced in 2/3 cultures via the addition of 20 mM Arabinose and the temperature reduced to 18°C. After 30 minutes the pCDF Duet-1 vector in 1 of the 2 pDB1282 induced cultures was induced by the addition of 1 mM IPTG and grown overnight. After 18 hours 2 mL of each culture was centrifuged at 8000 x g (Labnet Spectrofuge 16M) and the supernatant removed. The cell pellet was resuspended in 100 µL elgastat H2O and assessed via sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE).

2.12: Aerobic NiNTA purification

10 mL LB broth containing 100 µg.mL ampicillin and 50 µg.mL streptomycin was inoculated with a stab of a protein expression *E. coli* strain containing both pCDF Duet-1 (with gene of interest) and pDB1282 and grown at 37°C, 180rpm, until ~0.6 OD600 nm. At this point pDB1282 was induced by the addition of 20 mM Arabinose and the temperature dropped to 18°C. After 30 mins, pCDF Duet-1 was induced by the addition of 1 mM IPTG and left overnight.

The following morning the 10 mL culture was centrifuged at 4000 x g (ALC PK121R) and the pellet resuspended in BugBuster Primary Amine Free (Novagen) according to the manufacturers protocol, with 10 mM Imidazole and left to rotate at room temperature for 30 mins to disrupt the cell. At this point the cells were centrifuged at 4000 x g and the supernatant removed from the pellet.

Single use, 0.5 mL NiNTA centrifuge columns (BioVision) were equilibrated with 0.5 mL BugBuster containing 10 mM imidazole. The soluble fraction was loaded on to the NiNTA column and rotated at 4°C for 1 hour to bind the protein to the nickel beads. The NiNTA
columns were then centrifuged at 4000 x g for 1 minute and the flow-through collected. The column was washed with 1 mL wash buffer: 50 mM Tris, 0.25 M NaCl, 20 mM imidazole, 20% glycerol (v/v), pH 8.4 and the flow-through collected; followed by 1 mL elution buffer: 50 mM Tris, 0.25 M NaCl, 250 mM imidazole, 20% glycerol (v/v), pH 8.4 and the flow-through collected. These samples were assessed via SDS-PAGE.

2.13: Anaerobic protein purification

Large scale protein purification was performed in a Coy Labs type B vinyl anaerobic chamber in a 0PPM O₂ environment. The following four pH 8.4 buffers were made and left to equilibrate anaerobically for at least 48 hours before purification: Binding buffer – 50 mM Bis-Tris propane, 0.5 M KCl, 20% glycerol (v/v), 10 mM imidazole. Buffer A (Wash buffer) – 50 mM Bis-Tris propane, 0.25 M NaCl, 20% glycerol (v/v), 20 mM imidazole. Buffer B (Elution buffer) – 50 mM Bis-Tris propane, 0.25 M NaCl, 20% glycerol (v/v), 500 mM imidazole. Buffer C (Desalting buffer) – 50 mM Bis-Tris propane, 0.25 M NaCl, 10% glycerol (v/v), 5 mM dithiothreitol (DTT) added fresh on the day of use.

10 mL LB broth containing 100 µg.mL ampicillin and 50 µg.mL streptomycin was inoculated with a stab of a protein expression E. coli strain containing both pCDF Duet-1 (with gene of interest) and pDB1282 and grown at 37°C, 180 rpm overnight.

The following morning 8x500mL 2xYT broth with 100 µg.mL ampicillin and 50 µg.mL streptomycin in 2L conical flasks were inoculated with 1% overnight culture and grown at 37°C, 180rpm, until ~0.6 OD600 nm. At this point pDB1282 was induced by the addition of 20 mM Arabinose and the temperature dropped to 18°C. After 30 mins, pCDF Duet-1 was induced by the addition of 1 mM ITPG and left overnight.

The following day the overnight cultures were recovered and centrifuged in 1L flasks at 6500 x g for 10 minutes (Beckman coulter Avanti J-26 XP centrifuge, Beckman JLA-8.1000 rotor). Pellets were transferred to a 50 mL Falcon tube and moved into the glovebox.
50 mL binding buffer was supplemented with 0.5 mg.mL Lysozyme (Sigma), 1 μL.mL DNase I (NEB - 2 units per mL total solution) and an EDTA free protease inhibitor cocktail tablet and rotated anaerobically until all reagents had dissolved fully and left to rotate at 100 rpm on a spinning block for 60 mins. This solution was used to resuspend the pellet up to 50 mL total volume. Once resuspended the cells were suspended in a cold-water bath and sonicated with a 0.5 second cycle, at 60% amplitude until lysed and homogenous (UP50H sonicator, Hielscher).

Protein purifications were carried out using an AKTAPrime chromatography workstation (GE Healthcare) installed inside the anaerobic chamber. A 5 mL HisTrap NiNTA column was anaerobically equilibrated with 5 column volumes of binding buffer, following this a 50 mL Superdex 75 size exclusion column (GE Healthcare) was equilibrated with 5 column volumes of buffer C.

The sonicated mixture was placed in an airtight Oakridge polypropylene centrifuge bottles (Thermo) and sealed with parafilm, removed from the glovebox and centrifuged for 30 minutes at 15,000 x g (Avanti J26XP)). The Oakridge polypropylene centrifuge tube was inserted back into the glovebox and the ~50 mL supernatant containing soluble protein was removed and loaded onto a 5 mL HisTrap HF NiNTA column (GE healthcare).

Once the protein was bound to the 5 mL NiNTA column buffer A was passed through the column at 1 mL.min⁻¹ until the 280 nm UV trace equilibrated. The bound protein was eluted with a buffer A/B gradient according to the following protocol: 20 mM imidazole for two column volumes, rising to 50 mM over one column volume. The imidazole concentration was increased to 500 mM over 10 column volumes and 3 mL fractions collected and monitored via chromatogram (Primeview). The dark brown fractions were pooled and loaded on to a 50 mL Superdex 75 column and eluted with buffer C. 3 mL fractions were collected, pooled and stored at -80°C for downstream use.
2.14: SDS-page

Precast SDS-PAGE gels (Biorad) were used to assess purity of protein at all stages during purification. Samples were incubated with 2x SDS loading dye (100 mM Tris pH 6.8, 20% glycerol (v/v), 0.04% bromophenol blue, 4% SDS) at 94°C for 4 minutes then loaded alongside a protein standard (Unstained Protein Standard, Broad Range (10-200 kDa), NEB). Samples were run at 175V for ~60 minutes and visualised on a Syngene G:Box and analysed on Genesys software.

2.15: Reconstitution

Cfr and SbCfr samples were reconstituted according to the method described by Challand et al.[103]. Protein (2.5 – 5 mg.mL) was defrosted in the glovebox, suspended in a cold ice bath. 5 mM DTT was added and the mixture stirred very gently for 15 mins. 4 molar equivalents of iron (III) chloride were added from a freshly prepared 40 mM stock made with deoxygenated H₂O and the mixture stirred for 15 minutes. 4 molar equivalents Na₂S.9H₂O was added from an anaerobic 40 mM stock and the solution stirred for a further hour. Upon completion the reconstituted sample was then sealed in an airtight Oakridge polypropylene centrifuge bottles (Thermo), sealed with parafilm and taken from the glovebox to be centrifuged at 15,000 x g (Avanti J26 XP) for 5 mins then replaced in the glovebox. 200µL aliquots were flash frozen and stored at -80°C for downstream use.

2.16: Activity Assay

Protein aliquots were taken from -80°C storage, moved into a -20°C Eppendorf freezer block and defrosted anaerobically. Protein samples are treated with 1 mM SAM Iodide (Sigma) to ensure methylation and left to incubate at room temperature for 30 minutes, with an aliquot taken for SAH production assessment. Sample is then buffer exchanged into assay buffer: 50 mM Bis-Tris propane pH 8.4, 10% glycerol (v/v), 0.5 M KCl.

12.5 µM protein was incubated with 1mM SAM equilibrated to 37°C for 5 minutes before the addition of 1mM sodium dithionite to initiate the assay. Reactions were terminated by the addition of 1% perchloric acid and centrifuged at 16,000 x g for one hour to pellet precipitated
protein. The supernatant was removed and frozen at -20°C hours to further denature any remaining protein.

2.17: HPLC protocol

High performance liquid chromatography (HPLC) was carried out on a Shimadzu workstation (LC-20AD pump, DGU-20A55r degassing unit, SPD-M202 diode array detector, SIL-20AHT autosampler, CTO-10AS column oven). Assay samples were defrosted and centrifuged again at 16,000 x g for 30 mins and the supernatant moved in to 300 µL labelled crimp-top HPLC vials (Chromacol - Thermo). 8 µL of each sample was injected on to a Kinetex C18 reverse phase column (Phenomenex) running at 40°C and utilising an initial 0.2 M NH₄OAc, 2.5% MeCN mobile phase. Over a 13-minute time course the MeCN concentration varies as follows: 0 - 2 mins at 2.5% MeCN, 2 - 4.5 mins 12.5% MeCN, 4.5 – 5.5 mins 50% MeCN, 5.5 – 6 mins 50% MeCN, 6 – 6.1 mins 80% MeCN, 6.2 – 7.1 mins 80% MeCN, 7.2 – 13 mins 2.5% MeCN with samples assessed at 254 nm using LabSolutions software.

2.18: Circular dichroism

A JASCO J-815 CD polarimeter was used to measure circular dichroism spectra. 200uL of 0.1 mg.mL⁻¹ SbCfr was loaded in to a 1 mm pathlength quartz cuvette (Sigma) and sealed anaerobically. The sample was heated to 5°C, 25°C, 45°C, 65°C, 75°C and 95°C with UV-Vis wavelength scans taken at each temperature between 260 nm – 210 nm in 1 nm increments.

2.19: Anaerobic Crystal trials

50 µL Preformulated crystal screen reagent kits were loaded in to Swissci 96 well, 2 drop crystallisation plates (Hampton Research) and left to equilibrate anaerobically for 30 mins. Protein samples were loaded in to a 6 PPM O₂ anaerobic glovebox (Belle technologies) and using an OryxNano crystallisation robot (Douglas instruments) dispensed in to the two wells of the plates in a ratio of 1:1 protein:condition and 2:3 protein:condition respectively. Plates were sealed and left to crystallise in the glovebox. Crystals were looped with nylon loops.
(Hampton Research), flash frozen in liquid N\textsubscript{2} and seated in a pre-equilibrated Uni-puck for transport to the synchrotron.

2.20: Crystal structure resolution

The $P_{2_1}2_12$ Iron edge dataset from the xia2 pipeline, after integration in XDS and scaling and merging by XSCALE [158] was used in Autosol [159] (Phenix [160]) to determine phase information. Experimental phases, an electron density map and an initial model were calculated, with two molecules in the asymmetric unit and 8 Fe atoms identified corresponding to the expected [4Fe-4S] content. The initial model built by AutoBuild [161] in Phenix was subsequently refined against the higher resolution (1.42 Å) isomorphous native dataset (0.9686 Å; indexed using XDS and combined and scaled using XSCALE in the xia2 pipeline as above). The model was completed by iterative rounds of manual model building in COOT [162] and refinement in Phenix. Structure validation was assisted by Molprobity [163] and Phenix.

2.21: Disk testing methods

Disc susceptibility testing was performed according to CLSI methodology and interpreted using CLSI performance standards.
Chapter three: Recombinant protein production and biochemical characterisation

3.1: Introduction

The emergence of cfr and its PhLOPS\textsubscript{A} resistance phenotype in the clinical setting has complicated treatment options and limited the repertoire of effective anti-Gram-positive agents available. The mechanism of methylation employed by RlmN and Cfr to methylate 23s rRNA residue A2503 remain poorly characterised. As a starting point for investigations of this mechanism cfr and a number of cfr-like genes were selected for investigation, with the aim of identifying the most tractable system(s) for biophysical and structural study: cfr\textsubscript{(B)}, cfr\textsubscript{(C)}, and homologues from two thermophilic organisms - *Paenibacillus* spp. cfr (pbcfr) and *Sphaerobacter thermophilus* (sbcfr).

3.2: Vector construction

In order to evaluate the suitability of Cfr targets for investigation, two types of construct were generated. For protein expression trials synthetic genes were amplified by PCR and inserted by recombination into the protein expression pCDF Duet-1 vector as described in Methods. In addition to this, a second series of constructs were created using the arabinose inducible pBAD-A vector and transformed into a range of antimicrobial susceptible *E. coli* strains in order to investigate resistance phenotype. This project utilised 4 pre-existing constructs: pMC115 (cfr hosted in pCDF Duet-1) and pMC116 (cfr hosted in pBAD-A) were both generous gifts from Dr Martin Challand. pCDF-SbCfr and pCDF-PbCfr (sbcfr and pbcfr cloned into pCDF-Duet-1) were both created by the author prior to the start of the current project. For the remaining targets, cfr\textsubscript{(B)} and cfr\textsubscript{(C)}, synthetic, codon optimised genes were purchased from Eurofins genomics.

To generate the constructs required, the relevant synthetic genes were used as templates for PCR utilising primers which added a 15bp extension to either end of the gene which had a sequence homologous to the linearised ends of pCDF Duet-1 and pBAD A. The resulting PCR products were then utilised in In-Fusion\textsuperscript{®} reactions, which utilise homologous recombination to combine the PCR product and the linearised vector. These new constructs were then
transformed in to Stellar™ Competent *E. coli* and streaked on agar containing appropriate antibiotic. The pBAD-Cfr(B) transformation yielded only 3 colonies, whereas pCDF-Cfr(B) and pBAD-SbCfr yielded >30 colonies on each plate. For pBAD-Cfr(C) and pCDF-Cfr(C) transformations yielded >30 colonies per plate. To confirm the successful creation of these constructs colony PCR was used to detect the presence of intact *cfr* genes after transformation. All 3 pBAD-Cfr(B) colonies were used as templates for colony PCR (Figure 19, lanes 1-3), of which 2 displayed a band at the position expected of a product with a size of 1188bp. This was followed by the selection of 4 colonies from each of the pCDF-Cfr(B) and pBAD-SbCfr plates for use as colony PCR template (Figure 19 lanes 4-7 and 8-11 respectively). Of the 4 pCDF-Cfr(B) colonies selected 3 displayed a band of the expected size (1188bp). Similarly, 3 of 4 of the pBAD-SbCfr colonies displayed a band of the expected size of the *sbcfr* insert (1128bp). Of the 5 pBAD-Cfr(C) colonies selected 4 yielded a band at the expected position of 1188bp (Figure 20); whilst of the 5 pCDF-Cfr(C) colonies selected 4 also displayed a band of the same size (Figure 21). In each transformation, at least one colony failed to produce a band which implies the presence of “empty” vector which has likely carried over intact through the linearisation process.
Figure 19 – Colony PCR detection of pBAD-Cfr(B), pCDF-Cfr(B) and pBAD-SbCfr constructs created via In-Fusion® after transformation into Stellar™ Competent E. coli. Lanes 1-3: PCR reactions of 3 colonies from a pBAD-Cfr(C) transformation. Lanes 1 and 2 display a ~1.1 kilobase band at around the expected size of the PCR product (1188bp), lane 3 displays no band indicating empty vector. Lanes 4-7: PCR reactions from 4 colonies of a pCDF-Cfr(C) transformation. Lanes 4, 6 and 7 display a ~1 kilobase band at around the expected size of the PCR product (1188bp), while lane 5 displays a faint band at that site. Lanes 8-11: PCR reactions of 4 colonies from a pBAD-SbCfr transformation. Lanes 8-10 display a ~1 kilobase band at around the expected size of the PCR product (1128bp), lane 11 hosts no band at this position, indicating the presence of empty vector.
Figure 20 – Colony PCR detection of pBAD-Cfr(C) created via In-Fusion® after transformation into Stellar™ Competent E. coli. Lane 1: Empty lane. Lanes 2-6: Lanes 2 and 4-6 display a ~1.1 kilobase band at around the expected size of the PCR product (1188bp). Lane 3 contains no band, indicating the presence of empty vector.

Figure 21 - Colony PCR detection of pCDF-Cfr(C) created via In-Fusion® after transformation into Stellar™ Competent E. coli. Lanes 1-5: PCR reactions of 5 colonies from a pCDF-Cfr(C) transformation. Lane 1 displays no band of the expected size of 1188bp but does display a ~700bp band. Lanes 2-5 display a ~1100bp at the expected size of 1188bp as well as a smaller ~700bp band.
3.3: Vector expression trials

Protein expression trials used the pCDF constructs transformed into a range of protein expression *E. coli* strains alongside pDB1282 which contains the iron-sulphur generating *iscS-iscU-iscA-hscB-hscA-fdx* portion of the Isc operon from *Azotobacter vinelandii*. The inclusion of pDB1282 is commonly included during expression of radical SAM enzymes to improve the yield of holoproteins [164, 165]. *E. coli* strains investigated were BL21* which carries deletions in mRNA degradation pathways to increase transcription by prolonging transcript lifetime, and SOLUBL21 which has shown high production of low-solubility products.

In order to investigate recombinant Cfr expression from pCDF vectors three cultures were created from each transformation utilising different permutations of induction: 1. No induction, 2. pDB1282 induction without pCDF induction and 3. Induction of both pDB1282 and pCDF Duet-1. Following overnight growth of these cultures SDS-PAGE assessment of their whole-cell content was performed. (In the following section X:X denotes pDB1282 and pCDF-Cfr either Induced: I or uninduced: U).

Prior investigation into pCDF-SbCfr has shown that expression in BL21* is optimal and provides high yields. Figure 22 displays the whole cell content of three states of induction for pCDF-SbCfr in BL21* with no induction (U:U), induction of pDB1282 without pCDF-SbCfr (I:U) and induction of pDB1282 and pCDF-SbCfr (I:I). Lane U:U displays no significant band at the expected MW of SbCfr (40.63 kDa) when compared to the average intensity of other bands in that lane. Lane I:U, in which pDB1282 is induced, displays a significant band at ~45 kDa which corresponds to the expected size of the IscS cysteine desulphurase product from pDB1282 (45.089 kDa). The final I:I lane displays the culture where both pDB1282 and pCDF-SbCfr are induced. In addition to the IscS ~45 kDa band an additional band of significant intensity can be seen at ~40 kDa, corresponding to the expected size of SbCfr.
Prior investigation has also proven pMC115 (*cfr* hosted in pCDF Duet-1) produces high yields in BL21*. Similarly to pCDF-SbCfr in BL21*, figure 23 displays the whole cell content of the three states of induction for pMC115 in BL21*. In the U:U lane no significant band can be seen. In lane I:U a band of high intensity can be seen at ~45 kDa, which corresponds to the expected position of IscS (45.089 kDa). In lane I:I the IscS ~45 kDa band can be seen in addition to a ~40 kDa band which is the expected size of Cfr (39.861 kDa).

**Figure 22 – Expression profile of pCDF-SbCfr and pDB1282 in BL21* in three states of induction.** U:U – Neither plasmid induced. No significant band can be seen. I:U – pDB1282 induced, pCDF-SbCfr uninduced. A band of strong intensity can be seen at around ~45 kDa, the expected size of IscS from pDB1282 (45.089 kDa). I:I – Both pDB1282 and pCDF-SbCfr induced. In addition to the ~45 kDa IscS band a second band of stronger intensity can be seen at ~40 kDa, which is the expected size of SbCfr (40.63 kDa).
Figure 23 – Expression profile of pMC115 and pDB1282 in BL21* in three states of induction. U:U – Neither plasmid induced. No significant band can be seen. I:U – pDB1282 induced, pMC115 uninduced. A band of strong intensity can be seen at around ~45 kDa, the expected size of IscS from pDB1282 (45.089 kDa). I:I – Both pDB1282 and pMC115 induced. In addition to the ~45 kDa IscS band a second band of stronger intensity can be seen at ~40 kDa, which is the expected size of Cfr (39.861 kDa).

Following the successful expression of pCDF-SbCfr from pMC115, pCDF-PbCfr expression was tested in both BL21* and SOLUBL21 E. coli. Figure 24 displays the whole cell content of the three states of induction for pCDF-PbCfr in BL21*. Lane U:U, where no induction was performed, displays no band of significant intensity as expected from previous experiments. Lane I:U, where only pDB1282 was induced, displays a prominent band at ~45 kDa which corresponds to the expected size of IscS (45.089 kDa). Lane I:I, where both plasmids were induced, displays an identical profile to that of lane I:U indicating that expression of pCDF-PbCfr has not been induced. Figure 25 displays the whole cell content of the three states of induction for pCDF-PbCfr in SOLUBL21. In contrast with the successful results seen in the
previous two expression trials for SbCfr and Cfr, no lanes show an increase in band intensity at any point regardless of the induction state. Thus, induction of expression has not occurred for either PbCfr, or IscS encoded upon pDB1282.

Figure 24 – Expression profile of pCDF-PbCfr and pDB1282 in BL21* in three states of induction. U:U – Neither plasmid induced. No significant bands can be seen. I:U – pDB1282 induced and pCDF-PbCfr uninduced. A band of strong intensity can be seen at around ~45 kDa, the expected size of IscS from pDB1282 (45.089 kDa). I:I – Both plasmids induced. A band of strong intensity can be seen at around ~45 kDa, the expected size of IscS from pDB1282 (45.089 kDa) but no band at ~40kDa is present, indicating the lack of expression of PbCfr.
Figure 25 – Expression profile of pCDF-PbCfr and pDB1282 in SOLUBL21 in three states of induction. U:U – Neither plasmid induced. No significant bands can be seen. I:U – pDB1282 induced and pCDF-PbCfr uninduced. No significant bands can be seen. I:I – Both plasmids induced. No significant bands can be seen.

pCDF-Cfr(B) and pCDF-Cfr(C) were also transformed into BL21* and their expression profile examined. Figure 26 displays the whole cell content of the three states of induction for pCDF-Cfr(B) in BL21*. Lane U:U, where no induction was performed, displays no band of significant intensity as expected. Lane I:U, where only pDB1282 was induced, displays a prominent band at ~45 kDa which corresponds to the size of IscS (45.089 kDa). Lane I:I, where both plasmids were induced, displays a faint additional band at ~40 kDa at the expected size of Cfr(B) (40.718 kDa).

Figure 27 displays the whole cell content of the three states of induction for pCDF-Cfr(C) in BL21*. Lane U:U, where no induction was performed, displays no band of significant intensity as expected. Lane I:U, where only pDB1282 was induced, displays a prominent band at ~45 kDa which corresponds to the size of IscS (45.089 kDa). Lane I:I, where both plasmids were
induced, hosts a similar profile to that of the previous lane, indicating that Cfr(C) expression has been unsuccessful.

Figure 26 – Expression profile of pCDF-Cfr(B) and pDB1282 in BL21* in three states of induction. U:U – Neither plasmid induced. No significant band can be seen. I:U – pDB1282 induced, pCDF-Cfr(B) uninduced. A band of strong intensity can be seen at around ~45 kDa, the expected size of IscS from pDB1282 (45.089 kDa) I:I – Both pDB1282 and pCDF-Cfr(B) induced. In addition to the ~45 kDa IscS band a second band of stronger intensity can be seen at ~40 kDa, which is the expected size of Cfr(B) (40.718 kDa).
Figure 27 – Expression profile of pCDF-Cfr(C) and pDB1282 in BL21* in three states of induction. U:U – Neither plasmid induced. No significant band can be seen. I:U – pDB1282 induced, pCDF-Cfr(C) uninduced. A band of strong intensity can be seen at around ~45 kDa, the expected size of IscS from pDB1282 (45.089 kDa) I:I – Both pDB1282 and pCDF-Cfr(C) induced. A band of strong intensity can be seen at around ~45 kDa, the expected size of IscS from pDB1282 (45.089 kDa) but no band at ~40 kDa is present, indicating the lack of expression of Cfr(C).

3.4: Aerobic protein purification

Assessment of the whole-cell content of all constructs in BL21* has shown that pCDF-SbCfr, pCDF-Cfr(B) and pMC115 are all potentially capable of expressing their respective proteins. The majority of publications investigating radical SAM enzymes state that in order to produce intact, functional protein all experimentation must be performed in an anaerobic environment. However, to determine if these products are amenable to capture via immobilised metal affinity chromatography an aerobic purification step was performed prior to any anaerobic work.

Cultures were created wherein both plasmids induced and left to grow overnight. On the following day these cultures were centrifuged, and the pellets lysed with BugBuster reagent to
release all soluble protein. These soluble fractions were then passed over nickel-nitrilotriacetic acid (NiNTA) media hosted in desktop centrifuge size NiNTA columns to bind any His-tagged protein. The NiNTA media was then washed in a 20mM imidazole buffer. Finally, any bound protein was eluted from the column with a 250mM imidazole buffer. Samples from each of these steps were collected and assessed via SDS-PAGE to determine their content.

Initially a pDB1282 only strain was tested to ascertain if any product from this construct was capable of binding. Figure 28 displays samples of each step from the NiNTA purification of this pDB1282 only strain. The soluble fraction lane displays a band at ~45 kDa – the expected size of IscS (45.089 kDa). The Load lane, taken after being passed over the NiNTA media displays a similar profile. The Wash lane displays a faint band at ~45 kDa. The Elution lane displays a strong band at ~45 kDa. This is significant because it displays the capacity of this ~45 kDa protein to bind to NiNTA media and potentially obscure any assessment downstream.

**Figure 28 – Aerobic NiNTA purification of pDB1282 in BL21*.** Soluble fraction: Displays a high intensity band at ~45 kDa, the expected size of IscS. Load: Displays a band at 45~ kDa of slightly lower intensity than the previous lane. Wash: Displays a faint band at ~45 kDa. Elution: Displays a band of high intensity at ~45 kDa, indicating IscS is potential contaminant during NiNTA purification.
Figure 29 displays the aerobic NiNTA purification of SbCfr expressed from pCDF-SbCfr in BL21* cotransformed with pDB1282. The soluble fraction lane displays a band at ~45 kDa – the expected size of IscS (45.089 kDa) as well as a band at ~40 kDa, the expected size of SbCfr. The Load lane, taken after being passed over the NiNTA media displays a similar profile. The Wash lane displays a faint band at ~45 kDa. The Elution lane displays two bands, one at ~45 kDa and one at ~40 kDa, indicating the potential NiNTA purification of SbCfr.

Figure 30 displays the aerobic NiNTA purification of pMC115 and pDB1282 in Bl21*. The soluble fraction lane displays a band at ~45 kDa – the expected size of IscS (45.089 kDa) as well as a band at ~40 kDa, the expected size of Cfr (39.861 kDa). The Load lane, taken after being passed over the NiNTA media displays a similar profile. The Wash lane displays multiple bands, with one higher intensity band at 45~ kDa. The Elution lane displays two significant bands: one at 45~ kDa and one at ~40 kDa, indicating the potential NiNTA purification of Cfr.

Figure 31 displays the aerobic NiNTA purification of pCDF-Cfr(B) and pDB1282 in Bl21*. The soluble fraction lane displays a band of high intensity at ~45 kDa. The Load lane displays a similar profile. The Wash lane displays a faint band at ~45 kDa. The Elution lane displays a single band at ~45 kDa which indicates no expression of Cfr(B).
**Figure 29** – Aerobic NiNTA purification of pCDF-SbCfr and pDB1282 in BL21*. Soluble fraction: Displays a high intensity band at ~45 kDa, the expected size of IscS, a second band at ~40 kDa, the expected size of SbCfr, and a third band of moderate intensity at ~38 kDa. **Load**: Displays a similar profile to that of the previous lane. **Wash**: Displays a faint band at ~45 kDa. **Elution**: Displays a band of moderate intensity at ~45 kDa, indicating IscS, and a second band at ~40 kDa, indicating potential SbCfr purification.

**Figure 30** – Aerobic NiNTA purification of pMC115 and pDB1282 in BL21*. Soluble fraction: Displays a high intensity band at ~45 kDa, the expected size of IscS. **Load**: Displays a similar profile to that of the previous lane. **Wash**: Displays a faint band at ~45 kDa. **Elution**: Displays a band of moderate intensity at ~45 kDa, indicating IscS, and a second band at ~40 kDa, the expected size of Cfr (39.861 kDa), potentially indicating successful Cfr purification.
Figure 31 – Aerobic NiNTA purification of pCDF-Cfr(B) and pDB1282 in BL21*. Soluble fraction: Displays a high intensity band at ~45 kDa, the expected size of IscS, a second band at ~38 kDa, and a third band of moderate intensity at ~34 kDa Load: Displays a similar profile to that of the previous lane. Wash: Displays a faint band at ~45 kDa, a faint band at ~38 kDa and a faint band at ~34 kDa. Elution: Displays a single band of moderate intensity at ~45 kDa, but no band at ~40 kDa indicating failed Cfr(B) purification.

3.5: Anaerobic protein purification and reconstitution

Having shown expression during aerobic NiNTA purification, SbCfr (expressed in E. coli BL21* from pCDF-SbCfr), and Cfr (expressed in E. coli BL21* from pMC115) were entered into scaled-up anaerobic purification trials.

3.5.1: Purification of Cfr

To purify Cfr, a total of 4 litres of E. coli BL21* pMC115/pDB1282 was grown overnight and the pellet lysed via sonication in a Coy anaerobic chamber at 0ppm O₂. Once complete the solution was centrifuged and the soluble fraction loaded on to a 5mL His-Trap column (GE) via AKTA FPLC, resulting in a dark black band at the top of the media.

The bound protein was gradient eluted from the column, assessed via chromatogram (A280) and collected in 3mL fractions, as shown in figure 32. Lanes 1-13 display samples taken from
the 3mL fractions collected. All lanes have a high intensity band residing ~40 kDa, the expected size of Cfr (39.861 kDa). Additional bands are also visible, indicating some contaminant has also bound to the column. Davies et al have described several native E. coli proteins that are commonly co-purified during IMAC [166]. Interestingly, only a very faint ~45 kDa band can be seen in lanes 4-11, indicating that IscS may have eluted at a different imidazole concentration. 6 additional bands can be seen in each lane of varying intensity, corresponding in size with native E. coli proteins known to co-elute: An intense band appears at ~75 kDa in each lane, which could potentially be 74.2 kDa YfbG (SwissProt access code P77398). A second faint band can be seen at ~29 kDa, the origin of which is unknown. A third more intense band can be seen at ~25 kDa, around the expected size of YadF, a 25 kDa carbonic dehydratase (SwissProt access code P36857). A fourth, less intense band can be seen at ~20 kDa, around the expected size of SlyD, a 20.8 kDa Peptidoylproline cis–trans isomerase (SwissProt access code P30856). A fifth band can be seen at ~16.5 kDa, the expected size of Fur, a 16.7 kDa ferric uptake regulator (SwissProt access code P0A9A9). A final sixth band can be seen at ~14 kDa, the origin of which is unknown. It is not unusual to see these native E. coli proteins co-elute, especially when the NiNTA media has not been saturated with His-tagged protein of interest.

![Figure 32 - Collected fractions from anaerobic Cfr 5 mL NiNTA purification](image)

Lanes 1-13 display samples from each 3mL fraction collected from the AKTA FPLC. A very strong ~40
kDa band can be seen at the expected position of Cfr (39.861 kDa). Several other bands are also present, which potentially result from several known native E. coli proteins that commonly co-elute during IMAC.

Prior experimentation has shown that should the sample idle in high imidazole buffer (>25 mM), it immediately begins to precipitate. To circumvent this once these fractions had been purified from the NiNTA media, they were immediately pooled and loaded on to a 50mL Superdex 75 size-exclusion column which was pre-equilibrated with a non-imidazole buffer. This de-salted sample was collected in 3mL fractions. Figure 33 displays fractions taken from the Superdex 75 size-exclusion column (assessed via chromatogram A\textsubscript{280}). Lanes 1-7 display a strong ~40 kDa band at the expected size of Cfr (39.861 kDa) in addition to several fainter bands. Comparison between figures 32 and 33 display that this S75 purification has greatly reduced the level of contaminant within the sample. After assessment by SDS-PAGE the desalted sample was pooled and measured at 1.28 mg.mL, as measured by 280 nm absorbance. Several aliquots were taken, flash frozen and stored at -80°C for downstream use.

![Figure 33](image)

**Figure 33 – Collected fractions from anaerobic Cfr 50 mL Superdex 75 size exclusion purification.** Lanes 1-7 display samples from each 3mL fractions taken from the Superdex 75 elute. A very strong ~40 kDa band can be seen at the expected position of Cfr (39.861 kDa).
Several other bands are also present, which potentially result from several known native *E. coli* proteins that commonly co-elute during IMAC. Compared to the figure 32, it can be seen that this step of the purification has greatly reduced the level of contaminant in the sample. Once pooled this sample was measured at 1.28 mg.mL.

**3.5.2: Purification of SbCfr**

Following successful purification of Cfr, anaerobic purification of recombinant SbCfr was next attempted. A total of 4 litres of *E. coli* BL21* transformed with pCDF-SbCfr/pDB1282 was grown overnight and the pellet lysed via sonication in a Coy anaerobic chamber at 0ppm O₂. Once complete the solution was centrifuged and the soluble fraction loaded on to a 5mL His-Trap column, resulting in a dark black band that encompassed the majority of the media.

The bound protein was gradient eluted from the column, assessed via chromatogram (A₂₈₀) and collected in 3mL fractions, as shown in figure 34. All lanes have a high intensity band residing ~40 kDa, the expected size of SbCfr (40.630 kDa). Additional bands are also visible, indicating some contaminant has also bound to the column, though considerably less than was seen in the Cfr NiNTA purification (figure 32). This could be due to the greater concentration of his-tagged protein in the loaded sample out competing potential contaminant for the NiNTA binding sites.

Fractions were pooled and immediately loaded on to a 50mL Superdex 75 size-exclusion column and buffer exchanged in to a low imidazole buffer. 5mL fractions were collected. Figure 35 displays fractions taken from the Superdex 75 peak (assessed via chromatogram A₂₈₀). Lanes 2-10 display a strong ~40 kDa band and a negligible amount of contaminant at other sizes, indicating successful purification of SbCfr at 3.3 mg.mL. Throughout this purification SbCfr was a highly stable protein and experienced none of the minor precipitation observed with the Cfr sample. This could potentially be attributed to its origin in a thermophilic organism.
Figure 34 – Collected fractions from anaerobic SbCfr 5 mL NiNTA purification. Lanes 1-13 display samples from each 3mL fraction collected from the AKTA FPLC. A very strong ~40 kDa band can be seen at the expected position of SbCfr (40.630 kDa). Additional bands are also visible, indicating some contaminant has also bound to the column, though considerably less than was seen in the Cfr NiNTA purification (figure 32).

Figure 35 – Collected fractions from anaerobic SbCfr 50 mL Superdex 75 size exclusion purification. Lanes 2-10 display samples from each 3mL fractions taken from the Superdex 75 elute. A very strong ~40 kDa band can be seen at the expected position of SbCfr (40.630 kDa). Negligible amount of contaminant at other sizes, indicating successful purification of SbCfr at 1.9 mg.mL.
3.5.3: Reconstitution of purified Cfr/SbCfr

Prior work by Challand et al [157], as well as previous studies by other groups working with different radical SAM enzymes and other Fe-S cluster-containing proteins [167-169], has demonstrated that directly after purification, an Iron-Sulphur reconstitution step is usually required to ensure complete incorporation of fully intact Fe-S cofactors. The protein samples were incubated in a reducing Fe-S solution for 90 minutes, which allows the correct refolding of protein and incorporation of Fe-S clusters. Spectroscopic analysis was used to investigate the Fe-S content of our protein samples before and after reconstitution. Hagen et al [170] have reported that when assessed via UV-Vis spectroscopy, [4Fe-4S] clusters display a broad peak or "shoulder" between 350-420 nm.

Following reconstitution, the samples were concentrated to 5 mL and purified a final time on 150 mL Superdex 75 size exclusion column to purify to homogeneity and assessed via SDS-PAGE. Figure 36 displays the Cfr fractions collected from the 150 mL Superdex 75 size exclusion column. Lanes 2-6 show a single ~40 kDa band, with no contaminant. After being pooled the Cfr concentration was 1.9 mg.mL. Figure 37 displays the SbCfr fractions collected from the 150mL Superdex 75 size exclusion column. Lanes 2-6 show a single ~40 kDa band with no contaminant. After being pooled the SbCfr concentration was 8.96 mg.mL.
Figure 36 – Collected 2 mL fractions from anaerobic Cfr 150 mL Superdex 75 size exclusion purification. Lanes 2-6 display a single, uncontaminated band at ~40 kDa, the expected size of Cfr (39.861 kDa). After being pooled the Cfr concentration was 4.1 mg.mL

Figure 37 - Collected 2 mL fractions from anaerobic SbCfr 150mL Superdex 75 size exclusion purification. Lanes 2-6 display a single, uncontaminated band at ~40 kDa, the expected size of SbCfr (40.630 kDa). After being pooled the SbCfr concentration was 8.96 mg.mL.
Figure 38 displays the UV-Vis wavelength scan (300-700nm) of recombinantly produced wildtype Cfr before and after reconstitution. The unreconstituted Cfr sample (red trace) displays a region between 350-420nm that is characterised by a gradual sloping decline, looking dissimilar to the shoulder feature noted by Hagen et al. This implies the large majority of folded protein in this unreconstituted sample is lacking intact [4Fe-4S] clusters. Conversely, the reconstituted sample (blue trace) conforms closely to the 350-420nm shoulder characterised by Hagen et al. This implies that after reconstitution, a much larger proportion of the protein sample contains intact [4Fe-4S] clusters.

Figure 39 displays the UV-Vis wavelength scan (300-700 nm) of recombinantly produced SbCfr before and after reconstitution. The unreconstituted SbCfr sample (red trace) displays a 350-420 nm region which in stark contrast to that of unreconstituted wildtype Cfr conforms closely to that region as described by Hagen et al. This implies that prior to reconstitution a majority of protein within this sample possesses intact [4Fe-4S] clusters. The reconstituted SbCfr sample displays a similar 350-420 nm region, though at a higher absorbance. This is potentially due to an increase in overall intact [4Fe-4S] incorporation.
Figure 38 - UV-Vis wavelength scan (300-700 nm) of recombinantly produced wildtype Cfr before and after reconstitution. The unreconstituted Cfr sample (red trace) displays a region between 350-420 nm is that characterised by a gradual sloping decline, looking dissimilar to the shoulder feature noted by Hagen et al [170]. This implies the large majority of folded protein in this unreconstituted sample is lacking intact [4Fe-4S] clusters. Conversely, the reconstituted sample (blue trace) conforms closely to the 350-420 nm shoulder characterised by Hagen et al. This implies that after reconstitution, a much larger proportion of the protein sample contains intact [4Fe-4S] clusters.
Figure 39 – UV-Vis wavelength scan (300-700 nm) of recombinantly produced SbCfr before and after reconstitution. The unreconstituted SbCfr sample (red trace) displays a 350-420 nm region which in stark contrast to that of unreconstituted wildtype Cfr conforms closely to that region as described by Hagen et al. [170]. This implies that prior to reconstitution a majority of protein within this sample possesses intact [4Fe-4S] clusters. The reconstituted SbCfr sample displays a similar 350-420 nm region, though at a higher absorbance. This is potentially due to an increase in overall intact [4Fe-4S] incorporation.
3.6: S-Adenosylhomocysteine (SAH) and 5' Deoxyadenosine (DOA) production assays

Prior work performed by Challand et al [103] has shown that in the absence of its RNA substrate, Cfr is capable of interaction with a SAM substrate and in a process termed "uncoupled turnover" have been documented to utilise SAM for auto-methylation and DOA formation. In the event that the enzyme is unmethylated, the first molecule of SAM binds and the CH₃ group situated on the sulphur of SAM is displaced via an SN₂ reaction and hosted on a nearby Cysteine residue (Cys 355 in RlmN, Cys 338 in Cfr and Cys 334 in SbCfr), resulting in the formation of S-adenosylhomocysteine (SAH). In the event that the protein has been methylated, the binding of the "second" SAM results in the rapid oxidation of the [4Fe-4S] cluster and the cleavage of SAM to 5'deoxyadenosine (DOA) and methionine. Thus, by measuring SAH and DOA production, it is possible to demonstrate activity of purified reconstituted preparations notwithstanding the absence of an appropriate RNA substrate.

Accordingly, purified Cfr and SbCfr were assayed for SAH and DOA formation. During the reconstitution step both enzymes were incubated with 1mM SAM iodide in a non-reducing environment. This was intended to result in the auto-methylation of the protein, ensuring maximal activity in DOA formation assays as previous work demonstrated that methylation of Cfr is essential to catalysis of DOA production. Aliquots of these incubated samples were taken after 45 minutes and assessed via High Performance Liquid Chromatography (HPLC) on a C18 reverse phase column (Phenomenex) alongside SAM, SAH and DOA standards.

Figure 40 displays the standards used during this experiment and their respective elution times over a 10-minute chromatography run. SAM (black trace) is highly polar and therefore has a short retention time, eluting at ~1 minute. SAH is retained until 2 minutes, at which point the standard elutes in a tight, concise peak. DOA is retained until ~5.8 mins, at which point it elutes in a clear, concise peak. It is worth noting that SAM iodide (Sigma) is ~80% pure and therefore contains several additional products but can still be used for the purposes of defining retention time.
**Figure 40** – SAM, SAH and DOA standard elution points from a C18 reverse phase HPLC column over a 10-minute time course. SAM (black trace) is highly polar and therefore has a poor retention time, eluting at ~1 minute. SAH is retained until 2 minutes, at which point the standard elutes in a tight, concise peak. DOA is retained until ~5.8 mins, at which point it elutes in a clear, concise band.

Figure 41 (a) displays the elution profile during minutes 2 - 2.6 from 8 µL of a Cfr SAH formation assay injected on to a C18 reverse phase chromatography column (Phenomenex). Figure 41(b) Displays the pmol SAH in these injected samples, as determined by SAH standard. All reactions were performed with negative controls which displayed no elution at the 2 min mark in the absence of protein or SAM (data not shown), indicating a lack of SAH formation in the absence of either of these reagents. This assay was performed on an unknown quantity of protein prior to concentration, as a cursory investigation into methyl transfer activity.
Figure 41 – Representative SAH production by reconstituted and unreconstituted Cfr over 45 mins at 37°C. (a) The elution profile during minutes 2 - 2.6 from 8 µL of a reconstituted (blue) and unreconstituted (red) SbCfr SAH formation assay injected on to a C18 reverse phase chromatography column (Phenomenex). (b) Concentration of SAH (pmol) in these injected samples, as determined by SAH standard. Negative controls produced no ~2-minute peak, and therefore no SAH.

Figure 42 (a) displays the elution profile during minutes 2 - 2.6 from 8 µL of a SbCfr SAH formation assay injected on to a C18 reverse phase chromatography column (Phenomenex). Figure 41(b) Displays the pmol SAH in these injected samples, as determined by SAH standard. All reactions were performed with negative controls which displayed no elution at the 2 min mark in the absence of protein or SAM (data not shown), indicating a lack of SAH formation in the absence of either of these reagents. This assay was performed on an unknown quantity of protein prior to concentration, as a cursory investigation into methyl transfer activity.
Figure 42 – Representative SAH production by reconstituted and unreconstituted SbCfr over 45 mins at 37°C. (a) The elution profile during minutes 2 - 2.6 from 8 µL of a SbCfr SAH formation assay injected on to a C18 reverse phase chromatography column (Phenomenex). (b) concentration of SAH (pmol) in these injected samples, as determined by SAH standard. Negative controls produced no ~2-minute peak, and therefore no SAH.

Following SAM incubation, the samples were buffer exchanged in to an assay buffer with a NAP-5 purification column (GE) which separated the protein from smaller molecules including residual SAM and SAH. Uncoupled DOA generation by Cfr and Sbcfr was examined by incubation of 100 pmol enzyme with 1mM SAM and 1mM dithionite as reducing agent and incubated at 37°C. A sample of the assay was taken every 10 minutes for 60 minutes and precipitated with 1% perchloric acid then assessed via reverse phase HPLC alongside DOA standards.

Figure 43 displays the total DOA production (pmol) by 100 picomole Cfr over 60 minutes at 37°C: 0 min – 12.13 pmol, 10 min – 177.47 pmol, 20 min – 301.55 pmol, 30 min – 408.03 pmol, 40 min – 507.6 pmol, 50 min – 566.59 pmol, 60 min – 586.41 pmol, with a specific activity of 3.13 x 10³µmol.min⁻¹.mg⁻¹.
Figure 44 displays the total DOA production (pmol) by 100 picomole SbCfr over 60 minutes at 37ºC: 0 min – 13.19 pmol, 10 min – 41.28 pmol, 20 min – 69.9 pmol, 30 min – 100.4 pmol, 40 min – 123.25 pmol, 50 min – 145.26 pmol, 60 min – 152.35 pmol, with a specific activity of 6.87 x 10^{-4} µmol.min^{-1}.mg^{-1}.

In order to assess the impact of reconstitution on DOA production, both 12.5 µM reconstituted and unreconstituted Cfr and SbCfr were incubated for 60 minutes at 37ºC with 1mM SAM and 1mM dithionite as reducing agent. Figure 45(a) displays the chromatogram showing the elution during minutes 5 – 7 of the reconstituted (Blue) and unreconstituted (Red) Cfr DOA production assay. Figure 45(b) Displays the total DOA concentration for an 8 µL injection of this assay, ascertained using a range of DOA standards. The reconstituted sample produced 586.41 pmol in 60 minutes, with a specific activity of 3.13 x 10^{-3} µmol.min^{-1}.mg^{-1}. The unreconstituted sample produced 509.92 pmol in 60 minutes.

Figure 46(a) displays the chromatogram showing the elution during minutes 5 – 7 of the reconstituted (Blue) and unreconstituted (Red) SbCfr DOA production assay. Figure 41(b) Displays the total DOA concentration for an 8 µL injection of this assay, ascertained using a range of DOA standards. The reconstituted sample produced 152.35 pmol in 60 mins, with a specific activity of 6.87 x 10^{-4} µmol.min^{-1}.mg^{-1}. The unreconstituted assay produced 108.34 pmol om 60 minutes.
Figure 43 - Total DOA production (pmol) by 100 picomole Cfr over 60 minutes: 0 min – 12.13 pmol, 10 min – 177.47 pmol, 20 min – 301.55 pmol, 30 min – 408.03 pmol, 40 min – 507.6 pmol, 50 min – 566.59 pmol, 60 min – 586.41 pmol, with a specific activity of $3.13 \times 10^{-3} \mu\text{mol.min}^{-1}.\text{mg}^{-1}$.

Figure 44 - Total DOA production (pmol) by 100 picomole SbCfr over 60 minutes: 0 min – 13.19 pmol, 10 min – 41.28 pmol, 20 min – 69.9 pmol, 30 min – 100.4 pmol, 40 min – 123.25 pmol, 50 min – 145.26 pmol, 60 min – 152.35 pmol, with a specific activity of $6.87 \times 10^{-4} \mu\text{mol.min}^{-1}.\text{mg}^{-1}$.
Figure 45 – DOA production by 100 picomole reconstituted and unreconstituted Cfr over 60 mins at 37°C. (a) Elution time point of DOA from reconstituted (blue) and unreconstituted (red) DOA formation assays. (b) Displays the total DOA concentration for an 8 µL injection of this assay, ascertained using a range of DOA standards. The reconstituted sample produced 586.41 pmol in 60 minutes, with a specific activity of $3.13 \times 10^3 \mu \text{mol.min}^{-1}.\text{mg}^{-1}$. The unreconstituted sample produced 509.92 pmol over 60 minutes.
Figure 46 – DOA production by 100 picomole reconstituted and unreconstituted SbCfr over 60 mins at 37°C. (a) Elution time point of DOA from reconstituted (blue) and unreconstituted (red) DOA formation assays. (b) Displays the total DOA concentration for an 8 µL injection of this assay, ascertained using a range of DOA standards. The reconstituted sample produced 152.35 pmol in 60 mins, with a specific activity of $6.87 \times 10^{-4}$ µmol.min$^{-1}$.mg$^{-1}$. The unreconstituted assay produced 108.34 pmol over 60 minutes.
Sphaerobacter thermophilus is a thermophilic organism cultured at 55°C and therefore SbCfr may potentially have greater activity at this temperature. A second assay was performed at 55°C wherein which 12.5 µM Cfr and SbCfr were incubated with 1mM SAM and 1mM dithionite for 1 hour and the results assessed by reverse phase HPLC. Figure 47 displays the relative DOA (pmol) production by 100 picomoles reconstituted/unreconstituted Cfr and SbCfr at 55°C. All four samples performed poorly when compared to their respective 37°C datasets, however for the first time SbCfr produced a higher yield of product than Cfr. Reconstituted Cfr, unreconstituted Cfr, reconstituted SbCfr and unreconstituted SbCfr producing 5.33 pmols, 4.4 pmols, 19.07 pmols and 18.83 pmols respectively.

Figure 47 - Relative DOA (pmol) production by 100 pmol reconstituted/unreconstituted Cfr and SbCfr at 55°C. All four samples performed poorly when compared to their respective 37°C datasets, however for the first time SbCfr produced a higher yield of product than Cfr. Reconstituted Cfr, unreconstituted Cfr, reconstituted SbCfr and unreconstituted SbCfr producing 5.33 pmols, 4.4 pmols, 19.07 pmols and 18.83 pmols respectively.
3.7: Structural stability assessment by circular dichroism as a function of temperature

Circular dichroism (CD) is a powerful tool used to evaluate certain characteristics of protein secondary structure. CD can determine the presence of α-helices and β-sheets within a protein sample at very low concentrations. Here we used CD to measure the α-helical content of SbCfr under temperature challenge. Figure 55 displays the circular dichroism spectra of SbCfr at a range of increasing temperatures: 5°C, 25°C, 45°C, 65°C, 75°C and 95°C. Intact α-helices typically display negative bands at 222 nm and intact β-sheets display positive bands at 195 nm. To prevent precipitation of the protein in this case however the buffer contained 0.25 M NaCl which obscured results below 210 nm, precluding the possibility of β-sheet observation. The SbCfr CD assessment in figure 55 displays significant negative values (-18 mdeg to -16 mdeg) at 222nm in the 5°C, 25°C, 45°C and 65°C data and a significant reduction at this point for samples above 75°C. This implies a structurally intact protein up until ≤ 65°C, which along with the observed stability up to this point is consistent with the designation of SbCfr as a thermostable protein. This procedure was repeated with Cfr, however the inherent fragility of this protein resulted in significant precipitation before the assessment began.
Figure 48 – Circular dichroism spectra of SbCfr at a range of increasing temperatures: 5°C, 25°C, 45°C, 65°C, 75°C and 95°C. The 5°C, 25°C, 45°C and 65°C datasets display significant negative values (-18 mdeg to -16 mdeg) at 222nm. The data taken above 75°C indicate a significant reduction in negative values at 222nm. Taken together, these data imply that SbCfr remains structurally intact protein up until ≤ 65°C.

3.8: rRNA:Protein adduct formation

Work performed by McCusker et al [171] has highlighted the role of RlmN Cysteine 118 in the resolution of the transient Cys355-RNA covalent interaction. The mutation of Cys118 to an alanine residue has resulted in the inability of this mutant to resolve this covalent interaction, trapping the RNA substrate within the protein.

In addition to the cysteine residue located at the same site in SbCfr (Cys 103), there exists an additional cysteine residue (Cys 106). The role (if any) of this additional cysteine was examined by the creation of Cys to Serine mutants to elucidate which of these residues was responsible for this RNA adduct resolution.
Initially three mutants were generated: C103S, C106S and C103/106S (Confirmed by sequencing). The work performed by McCusker et al display via SDS-PAGE that in addition to a band the usual size of RlmN, a secondary much heavier band can also be seen, which reflects the RlmN:RNA adduct. To this end, these three mutant constructs were transformed in to BL21* and expressed under the usual protocol. Figure 49 displays the SDS-PAGE assessment of the whole cell content of these strains. In all three lanes a heavy band can be seen at ~40 kDa, the expected size of SbCfr. However, no heavier bands are apparent.

![Figure 49 – SbCfr C103S, C106S, C103/106S RNA:SbCfr adduct formation whole cell native PAGE assessment. Lanes 1-3 highlight a ~40 kDa band as expected, but no additional heavier bands can be seen.](image-url)
One potential reason for the failure of recombinantly produced SbCfr to bind to native *E. coli* 23S rRNA may lie in the sequence. Work performed by Yan *et al.* [96] has displayed 3 RNA regions within the substrate that are integral to Cfr and RlmN binding, without which methyltransferase activity declines rapidly. Figure 50 displays helices 89-93 which are integral to high level methyltransferase activity. Helix 89 is shown in green, helices 90-92 are shown in yellow and helix 93 is shown in violet. The A2503 site of action of Cfr and RlmN resides between Helix 89 and Helices 90-92.

**Figure 50 – Helices 89-93 are integral to high level methyltransferase activity.** Helix 89 is shown in green, helices 90-92 are shown in yellow and helix 93 is shown in violet. The A2503 site of action of Cfr and RlmN resides between Helix 89 and Helices 90-92. The absence of any of these regions greatly diminishes methyltransferase activity. Adapted with permission from Yan, F., *et al.*, RlmN and Cfr are radical SAM enzymes involved in methylation of ribosomal RNA. J Am Chem Soc, 2010. 132(11): p. 3953-64. Copyright 2010 American Chemical Society
A sequence alignment of this region within *E. coli K12*, *S. aureus* and *S. thermophilus* was performed to determine if any significant differences could be observed. Figure 51 displays the alignment between these regions in *Staphylococcus aureus* (2454-2643), *E. coli* (2447-2625) and *Sphaerobacter thermophilus* (2511-2689) 23s rRNA, colour coded to highlight which sequence resides in which helix. Helix 89 is shown in green, helices 90-92 are shown in yellow and helix 93 is shown in violet.

In order to determine if SbCfr would be more amenable to binding its native 23S rRNA sequence, a new construct was created. Work performed by Ponchon *et al* [172, 173] has highlighted the stability of RNA transcripts when hosted in a tRNA scaffold, as displayed in figure 52. Transcription that results in a tRNA scaffold sequence at both termini are not polyadenylated and therefore benefit from an increased duration in cytoplasm before degradation occurs.
Figure 52 - tRNA scaffold schematic. The tRNA scaffold designed by Poncho et al allows for the hosting of RNA transcripts with increased stability. Transcripts hosted in this structure are protected from polyadenylation, which greatly increases persistence in the cytosol. A series of cleavage sites have been added for ease of sequence introduction. The usual anticodon site has been replaced with a sequence that incorporates a sephadex aptamer region, which allows purification downstream.
Additional work performed by Ponchon et al. [155] has described the difficulty surrounding the co-expression of RNA and protein from two constructs that both utilise a T7 RNA polymerase binding site. Thus, in an attempt to provide recombinant E. coli expressing SbCfr with sufficient copies of Sphaerobacter-specific rRNA sequences that may be capable of forming covalent adducts, it was decided to construct a separate plasmid able to express a target RNA sequence from a T7-independent promoter. To achieve this, the pACYCT2 vector was selected. pACYCT2 expresses recombinant inserts under control of the tac promoter (a synthetic hybrid of the trp and lac promoters) and is anticipated to be capable of achieving near-equivalent levels of expression to those of gene targets under control of the T7 promoter.

A DNA fragment containing the S. thermophilus 23S rRNA sequence (bases 2511 – 2689) hosted within the tRNA scaffold shown in Figure 52 was synthesised by Eurofins genomics and inserted into the pACYCT2 plasmid (Addgene). The resulting plasmid, pACYCT2-Sb23S, was co-transformed into E. coli BL21* along with pDB1282 and a pCDF Duet-1 vector containing one of the three SbCfr mutants: SbCfr C103S, SbCfr C106S and SbCfr103/106S. These strains were grown overnight and induced using the regular protocol.

Figure 53 displays the whole cell assessment via native PAGE. Similar to the SbCfr mutant strains a band can be seen at ~40 kDa, though lane 2 has run poorly. A second band can be seen in lanes 1 and 3 at around 38 kDa. The expected size of the RNA transcript is 179 kDa, so it is unlikely this band represents the RNA:SbCfr adduct, though this could be confirmed via western blot.
Figure 53 – SbCfr C103S, C106S, C103/106S RNA:SbCfr adduct formation in the presence of pACYCT2-Sb23S whole cell native PAGE assessment. A band can be seen at ~40 kDa, though lane 2 has run poorly. A second band can be seen in lanes 1 and 3 at around 38 kDa. The expected size of the RNA transcript is 179 kDa, so it is unlikely this band represents the RNA:SbCfr adduct, though this could be confirmed via western blot.

3.9: Effect of SbCfr, Cfr and Cfr(B) expression on antimicrobial susceptibility of recombinant E coli

As described elsewhere [174] Cfr expression confers upon recombinant E. coli the PHLOPS\textsubscript{A} phenotype associated with resistance to multiple ribosome-acting antibiotics. Thus, in order to complement in vitro studies described above, the effect of expressing Cfr homologues on antimicrobial susceptibility of E. coli was investigated. As the majority of PHLOPS\textsubscript{A} agents have limited anti-Gram-negative activity due to poor penetration, effects upon phenicol (chloramphenicol, florfenicol) antibiotics were selected as markers of activity. Three E. coli
strains were used for susceptibility testing: BW25113 (BW), the parent strain, BW-RlmNΔ (ΔRlmN), a RlmN knock out strain with RlmN replaced by a KanamycinR cassette and BW-J396-RlmNΔ (J396), a RlmN silenced strain, with a KanamycinR cassette inserted elsewhere. As literature reports conflict concerning the involvement of RlmN-catalysed rRNA modifications upon antimicrobial susceptibility, the ΔRlmN strain was used in an effort to identify any possible effects of RlmN-like, rather than Cfr, activity. BW-J396-RlmNΔ was chosen as an additional control, to account for any possible interference with cell survivability arising from replacement of RlmN with the KanamycinR cassette.

Table 2 displays the measured zones of inhibition surrounding each antibiotic disk after overnight growth. Cfr, SbCfr and Cfr(C) hosted in a pBAD vector were induced with 5% arabinose, alongside a strain-only and empty pBAD controls, and challenged with 30µg chloramphenicol (C) and 30µg florfenicol disks. The following day the zone of inhibition was measured (mm) and compared. The control strains were all inhibited by chloramphenicol and florfenicol, with clear zones of inhibition evident with diameters ranging between 22-26mm. Similar values were observed with the empty pBAD control strains, with values of 24-25mm.

In comparison, strains containing pBAD-Cfr displayed considerable reduction in susceptibility resistance. In all strain backgrounds expression of Cfr led to smaller inhibition zones, of 17 – 19 mm diameter for chloramphenicol, and 14 - 15 mm diameter for florfenicol, implying that Cfr activity was protecting the organism from antibiotic action. Conversely, production of SbCfr by all three strains conferred no change in antimicrobial susceptibility compared to controls, implying that SbCfr is not acting to modify C8 of E. coli A2503. This experiment was kindly performed by Dr Jacqueline Findlay.

Although Cfr(C) constructs in the pCDF Duet-1 expression vector were unable to produce detectable recombinant protein (see above), the effect of expression in pBAD on phenicol susceptibility was also tested. However, the observed zones of inhibition for Cfr(C) strains were in all cases consistent with those seen in the negative controls (25-28mm) implying that
Cfr(C) was not being produced in this construct or was inactive in recombinant *E. coli* under these conditions.

<table>
<thead>
<tr>
<th>Strain</th>
<th>CAM mm</th>
<th>FFC mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>BW- 1</td>
<td>23</td>
<td>22</td>
</tr>
<tr>
<td>BW- 2</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>J396 - 1</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>J396 - 2</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>ΔRlmN - 1</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>ΔRlmN - 2</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>BW Empty pBAD - 1</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>BW Empty pBAD - 2</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>J396 Empty pBAD - 1</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>J396 Empty pBAD - 2</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>ΔRlmN Empty pBAD - 1</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>ΔRlmN Empty pBAD - 2</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>BW Cfr - 1</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>BW Cfr - 2</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>J396 Cfr - 1</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>J396 Cfr - 2</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>ΔRlmN Cfr - 1</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>ΔRlmN Cfr - 2</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>BW SbCfr - 1</td>
<td>26</td>
<td>25</td>
</tr>
<tr>
<td>BW SbCfr - 2</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>J396 SbCfr - 1</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>J396SbCfr - 2</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>ΔRlmN SbCfr - 1</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>ΔRlmN SbCfr - 2</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>BW Cfr(C) - 1</td>
<td>26</td>
<td>25</td>
</tr>
<tr>
<td>BW Cfr(C) - 2</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>J396 Cfr(C) - 1</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>J396 Cfr(C) - 2</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>ΔRlmN Cfr(C) - 1</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>ΔRlmN Cfr(C) - 2</td>
<td>28</td>
<td>26</td>
</tr>
</tbody>
</table>

**Table 2 – Cfr, SbCfr and Cfr(C) tested against chloramphenicol and florfenicol, displayed in mm zone of inhibition.** The empty BW, J396 and ΔRlmN control strains were all inhibited by chloramphenicol and florfenicol with mm values ranging between 22-26mm zones of inhibition. Similar values were observed with the pBAD only strains, where mm values ranged between 24-25. pBAD-Cfr strains displayed considerable resistance by comparison where all strains had lower inhibition zones of 17-19mm when challenged with chloramphenicol and 14-15mm when challenged with florfenicol, implying that Cfr activity was protecting the organism. Conversely, the production of SbCfr by all three strains had no
antimicrobial resistance phenotype when compared to the controls and Cfr, implying that the role of SbCfr in *S. thermophilus* may be as a RlmN, rather than a Cfr. The zones of inhibition documented for Cfr(C) strains reflected those seen in the negative controls (25-28mm) implying that Cfr(C) was not being produced in this construct. This experiment was kindly performed by Dr Jacqueline Findlay.

### 3.10: Discussion

While initially Cfr, SbCfr, PbCfr and Cfr(B) were the four targets for this project, the discovery of Cfr(C) was made during the project and was included in an effort to broaden our investigation and increase the chance of characterisation of these products.

After successful creation of SbCfr, Cfr(B) and Cfr(C) expression vectors and subsequent transformation in to expression strains, we were able to display protein production in pCDF-SbCfr, Cfr(B) and Cfr(C). Initially we opted to characterise the expression profile of each plasmid in the expression strain (pCDF Duet-1 and pDB1282) by performing triplicates of 50 mL overnight cultures and inducing them in different permutations: one with no induction, one with only pDB1282 induced and one with both pCDF Duet-1 and pDB1282 induced. Immediately it became apparent that the expression of pDB1282 resulted in a large band at ~45 kDa, the expected size of IscS, a cysteine desulferase. Knowing that the target protein sizes were between 39 kDa and 42 kDa this informed our decision to run all SDS-PAGE gels with a low quantity of sample (4 µL) at 100V to allow for the delineation between this potential IscS band and our product.

A protein band at ~40 kDa was seen clearly in the pDB1282+pCDF Duet-1 bands of the Cfr and SbCfr expression strains, and to a lesser degree in the Cfr(B) strain. The PbCfr and Cfr(C) strains provided no indication that expression had occurred.

Next, we subjected the Cfr, SbCfr and Cfr(B) strains to a cursory aerobic purification using 0.5 mL of NiNTA media in desktop centrifuge spin column to determine if purification was
possible. Initially we performed the purification of a pDB1282-only strain in order to determine if any of the Isc operon products would bind. This showed us that the large band seen at ~45 kDa was a candidate for NiNTA purification and again highlighted the need for vigilance when trying to determine which band belonged to which product. The Cfr and SbCfr expression strains produced a product which upon purification yielded a faint band at ~40 kDa, in addition to the band at ~45 kDa. Cfr(B), however failed to produce a band at this site whatsoever and was therefore removed from the pool of constructs being investigated. Cfr and SbCfr purifications were scaled up and performed anaerobically.

In the early stages of the project, the anaerobic purification protocol devised by Dr Martin Challand was followed. This protocol was similar to what was performed in the latter stages of the project, however the addition of one-hour anaerobic incubation at room temperature and stirring at 100rpm prior to sonication of the cells (a deviation from the original protocol) greatly increased the time the cells were in contact with lysozyme and DNAse I and reduced the time it took to lyse the cells by sonication. In addition, to this the original protocol utilised a column containing 25 mL NiNTA media for IMAC, which resulted in the co-purification of a range of contaminants. While this column was used in the initial stages of this project, it was eventually replaced with a 5 mL His-Trap NiNTA column (GE) which considerably reduced the amount of contaminant binding, while retaining almost identical levels of purification of our protein of interest.

The overnight cultures were scaled up from 50 mL to 4 Litres and the cells pelleted, moved in to the glovebox, resuspended in anaerobic binding buffer and purified as described in the method section. Conveniently, as an Fe-S enzyme, the product of interest exhibited a very dark brown colour and was easy to track throughout all purification steps. On the advice from Dr Challand the purified protein from the 5 mL NiNTA column was immediately pooled (with samples taken from each fraction for assessment by SDS-PAGE) and desalted to avoid precipitation in a high imidazole solution. Figure 32 displays the 13 fractions taken from this initial purification step, and all show a high intensity band at ~40 kDa, the expected size of Cfr,
as well as a number of other bands of contaminant. After being pooled and desalted on a 50 mL Superdex 75 size exclusion column the level of contaminant had decreased, as shown in figure 33. Conversely the purification of SbCfr on a 5 mL NiNTA column occurred with minimal contaminant. Figure 34 shows the 13 fractions taken from the NiNTA purification, and when compared to the Cfr gel displays fewer bands of contamination.

Fe-S reconstitution is a common protocol performed on recombinantly purified Fe-S enzymes. Work performed by Hagen et al displayed that when assessed via UV-Vis spectrophotometry a sample containing intact [4Fe-4S] clusters will display a peak or “shoulder” between 350-420 nm. The UV-Vis spectra [in figure 36] indicates that unreconstituted Cfr (red trace) does not possess this peak. After reconstitution however, the same sample (with a standardised concentration of 1.08 mg.mL displays this shoulder, indicating a much larger proportion of this sample contains protein with intact Fe-S cofactors. Conversely, the UV-Vis scan of unreconstituted SbCfr displays a peak at 420 nm, indicating the presence of intact Fe-S prior to reconstitution. This would indicate that the enzyme is more capable of retaining its structure during the purification process than wildtype Cfr. The reconstituted SbCfr UV-Vis spectra indicate a similar 420 nm peak as seen in the unreconstituted sample, though at a higher absorbance in spite of a standardised protein concentration of 3.1 mg.mL between the two samples. This difference could potentially be explained as the reconstituted sample having a slightly higher percentage of intact Fe-S cofactor. Interestingly, almost all contaminant was removed from these samples during reconstitution. During the protocol, the sample would contain a visible concentration of precipitated protein, which upon completion would be centrifuged and removed, leaving a homogeneous sample of our protein of interest.

produced During reconstitution SAM was added and incubated (alongside negative controls) at room temperature for 45 minutes. After this point a small sample was taken and precipitated to remove any functional Cfr/SbCfr, centrifuged to remove precipitate, and assessed via HPLC
to determine if SAH formation had occurred in the presence of our purified enzyme. When compared against a range of SAH standards injected on to a C18 reverse phase HPLC (RP-HPLC) column (8 µL). While the protein concentration during reconstitution is unknown, compared to the negative controls (protein in reconstitution mix and SAM in reconstitution mix) the generation of SAH only occurred in the presence of the enzyme and SAM substrate. Figure 41(B) displays that the reconstituted Cfr sample produced 613.4 picomoles of SAH in 45 minutes, and unreconstituted Cfr produced 579.9 picomoles of SAH in that same time. Figure 42(B) displays that reconstituted SbCfr produced 522.25 picomoles of SAH, and unreconstituted SbCfr produced 254.67 in that same time. While this SAH production assay provides only a cursory insight in to SAH production by these two enzymes, it does prove that SAH formation, likely by auto-methylation, is indeed possible.

DOA formation was measured against a known concentration of protein. When compared to negative controls and assessed via HPLC is can be seen that both Cfr and SbCfr are capable of multiple turnovers of SAM in to DOA. Figures 43 and 44 display a time course of DOA production by Cfr and SbCfr respectively. Reactions were initiated, and a sample taken every 10 minutes for 60 minutes to determine DOA formation. An 8 µL sample of each timepoint was injected on to a C18 RP-HPLC column and the total concentration of DOA in that sample calculated with the use of DOA standards. Each 8 µL injection contains the product of 100 picomoles of protein. Figure 43 displays the total concentration of DOA in each 8 µL injection of each time point from a Cfr DOA formation assay. It can be seen that over time the DOA concentration increases: 0 min – 12.13 pmol, 10 min – 177.47 pmol, 20 min – 301.55 pmol, 30 min – 408.03 pmol, 40 min – 507.6 pmol, 50 min – 566.59 pmol, 60 min – 586.41 pmol, with a specific activity of $3.13 \times 10^{-3} \mu \text{mol.min}^{-1}.\text{mg}^{-1}$. Figure 44 shows the same trend during a SbCfr DOA formation assay: 0 min – 13.19 pmol, 10 min – 41.28 pmol, 20 min – 69.9 pmol, 30 min – 100.4 pmol, 40 min – 123.25 pmol, 50 min – 145.26 pmol, 60 min – 152.35 pmol, with a specific activity of $6.87 \times 10^{-4} \mu \text{mol.min}^{-1}.\text{mg}^{-1}$. When compared to the negative controls,
in which negligible DOA formation occurred, it can be seen that Cfr and SbCfr are capable of DOA production in the absence of an RNA substrate, with Cfr having the highest activity.

When comparing the total DOA (pmol) formation over one hour between our reconstituted and unreconstituted samples, it can be seen that while unreconstituted protein can still produce DOA, the reconstituted samples have a higher activity when at the same concentration. Figure 45 displays the comparison of total DOA (pmol) in an 8 µL injection produced by reconstituted and unreconstituted Cfr over 60 minutes. 100 picomoles reconstituted Cfr produced 586.41 picomoles of DOA in 60 minutes, whereas 100 picomoles unreconstituted Cfr produced 509.92 (86.9% of what the reconstituted sample produced), implying the reconstituted sample contained a higher percentage of fully intact Fe-S cofactors. SbCfr performed at a lower activity, with 100 picomoles unreconstituted SbCfr producing 152.35 picomoles of DOA in 60 minutes and unreconstituted SbCfr producing 108.34 picomoles (71.1% of what the reconstituted sample produced), again implying that the reconstitute sample may possess a larger population of protein with fully intact [4Fe-4S] clusters.

SbCfr originates from *Sphaerobacter thermophilus*, a thermophilic bacterial species which is cultured at 55ºC. To determine if SbCfr would be more active at this temperature, we repeated the DOA formation assay at this temperature and assessed via HPLC. Figure 47 displays the total DOA concentration in an 8 µL sample of this reactions. Compared to 37ºC both enzymes experience a significant drop in activity. With reconstituted Cfr, unreconstituted Cfr, reconstituted SbCfr and unreconstituted SbCfr producing 5.33 pmols, 4.4 pmols, 19.07 pmols and 18.83 pmols respectively. In addition, the protein aliquots were equilibrated at 55ºC for 5 minutes before the assay was initiated, during which time high levels of precipitation were observed in the Cfr samples and slight precipitation (though to a much lesser degree) were observed in the SbCfr samples. This significant decrease in activity at this temperature highlights that the previous temperature of 37ºC provides better results.

The stability of SbCfr was then assessed via circular dichroism (CD). Figure 48 displays the circular dichroism spectra of SbCfr in the far-UV region at a range of increasing temperatures:
5°C, 25°C, 45°C, 65°C, 75°C and 95°C. Intact α-helices produce intense negative bands at 222nm, while intact β-sheets display positive bands at 195nm. Unfortunately to safeguard the protein 0.25 M NaCl was included in the buffer, which obscures any results below 210 nm. However, the α-helical content was assessed to be intact at 5°C, 25°C, 45°C and 65°C, with structure destabilisation observed at 75°C. When Cfr assessment by CD was attempted, significant precipitation was observed prior to the procedure, in spite of attempts to seal the cuvette from the aerobic environment.

Attempts were made to repeat work performed by McCusker et al where an RNA substrate became covalently bound to RlmN during C2 methylation, by the mutation of Cys 118 to alanine. E. coli One key difference between SbCfr and RlmN is the presence of a fifth cysteine in SbCfr, potentially located in the active site: Cys 106. To determine if this cysteine was of mechanistic importance three cysteine to serine mutants were generated: C103S, C106S and C103/106S. The McCusker experiments assessed crude extracts of recombinant E. coli expressing the RlmN C118A mutant which were assessed by SDS-PAGE and displayed a second heavy band in addition to that of RlmN, one with a higher molecular weight which corresponds to a RlmN:RNA adduct. We transformed these mutants in to BL21* E. coli and expressed the product as per the usual protocol. Upon assessment E. coli by SDS-PAGE we could visualise our SbCfr, but no heavier bands were apparent.

E. coli E. coli Work performed by Yan et al [96] has documented the role three specific regions within the RNA substrate that are integral to RlmN and Cfr binding, and without which activity declines considerably. Figure 50 displays helices 89-93 which are integral to high level methyltransferase activity. When aligned, the E. coli, S. aureus and S. thermophilus RNA sequences at these positions show that while they do align, key residues in integral loops do not match, as shown in figure 51. There exists the possibility that SbCfr is incapable of binding to an RNA substrate with the E. coli sequence.

Work performed by Ponchon et al [144, 152], has highlighted that RNA transcripts which possess the tRNA scaffold are not polyadenylated and therefore persist in the cell for longer.
These same publications also highlight issues surrounding the co-expression of protein and RNA from vectors that utilise the same promoter sequence. To circumvent this, they created a vector (pACYCT2 – Addgene Plasmid #45799) which utilises a tac promoter (a synthetic hybrid of the trp and lac promoters) to host their RNA expression sequence and return normal expression of protein and RNA to the cell.

To examine if SbCfr would be amenable to binding its native RNA sequence, a *S. thermophilus* RNA substrate hosted inside a tRNA scaffold was synthesised, inserted into pACYCT2 (Addgene Plasmid #45799) alongside pDB1282 and pCDF Duet-1 and grown overnight as per the usual protocol. Similarly, however, the crude sample was assessed via SDS-PAGE and no heavier band was apparent, as shown in figure 53.

Antimicrobial disk testing has shown that while Cfr is capable of conferring resistance to phenicols, SbCfr expression appears to yield no antimicrobial resistance (AMR) phenotype. At the outset of this project every *cfr*-like gene that conferred an AMR phenotype possessed an additional cysteine near the CX₃CX₂C active site motif. It was hypothesised that this extra cysteine was somehow involved in C8 methylation, and the presence of an extra cysteine in this motif of SbCfr, alongside its origins in a thermophilic organism, made it an interesting candidate for investigation. During the project however the discovery of *cfr(C)*, which has only 4 cysteines in this region like RlmN, likely disproves this hypothesis by conferring a PhLOPSₐ phenotype.

The circular dichroism data, taken alongside modest DOA formation at 55°C and the possible retention of Fe-S cofactors during initial purification highlight the stability of SbCfr, and contribute evidence toward the notion that this is a thermostable enzyme. In spite of the lack of an AMR phenotype in cells producing SbCfr, information gathered from the characterisation of this enzyme could help elucidate not only the mechanism of Cfr, but the mechanism behind a range of radical SAM enzymes.
Chapter four: Crystallisation and Structure Determination for SbCfr

4.1: Crystal screening

Currently all data regarding the structural basis for RNA methylation by Cfr and RlmN originate from crystal structures of *E. coli* RlmN. While providing valuable insights into the mechanisms of radical SAM enzymes as a whole, and of RNA methyltransferases in particular, the description of the catalytic cycle remains incomplete, and the basis for activity of Cfr at the C8, rather than C2, position of adenosine remains unclear. Obtaining a crystal structure of Cfr would represent a substantial step towards answering some of these questions. Preliminary observations of SbCfr highlight its activity and stability, as assessed by assays of SAM conversion to SAH/DOA and circular dichroism spectropolarimetry, respectively. These data, in conjunction with the presence of a fifth cysteine residue (Cys) also present in Cfr, suggest that SbCfr would be a reasonable candidate for crystallisation trials alongside recombinant Cfr. Accordingly, crystallisation of purified reconstituted Cfr and SbCfr was attempted with the aim of obtaining structural information on one or both of these targets.

4.2: Crystallisation screening of Cfr

The original RlmN structural publications [175] reported crystallisation of the protein at a concentration of 100mg.mL. Attempts to replicate these conditions revealed that SbCfr started to precipitate above 11 mg.mL, and Cfr above 9 mg.mL when subjected to anaerobic pressure concentration. This limitation dictated protein concentration in our crystallisation drops. All crystal plates were made in a second anaerobic glovebox (Belle Technologies) at 6ppm O₂, utilising an OryxNano crystallisation robot (Douglas instruments).

Initially five preformulated reagent crystal screens were chosen to test against Cfr: Morpheus® HT-96, JCSG-plus™ HT-96, MIDAS™, TOP96 (Molecular Dimensions), and Grid Screen™ Salt HT (Hampton Research). These trials enabled determination of Cfr stability in a range of conditions. 200µL Cfr (50mM Bis Tris Propane, 0.25M NaCl, 10% glycerol) at 9 mg.mL was aliquotted, flash frozen, situated in an Eppendorf freezer block and put into the crystallography glovebox alongside plates containing these five crystal screens, with 500µL reagent in each
well. Upon entering these plates were given 30 mins to equilibrate and the OryxNano used to load one well of a 96-well 2-drop MRC crystallisation plate (Swissci) with a 0.5 µL drop in a 1:1 protein:condition ratio (drop 1) and the second in a 3:2 protein:condition ratio (drop 2).

Almost immediately after dispensing, protein in >95% of the wells began to precipitate. Figure 54 displays representative examples of Cfr wells directly after plating. Figure 54(A) displays Drop 2 of well D9 from the Morpheus® HT-96 screen. Figure 54(B) displays drop 1 of well E7 from the HT-96, MIDAS™ screen. Figure 54(C) displays drop 2 of well A6 from the JCSG-plus™ screen. Accordingly, this process was repeated with these 5 initial screens using single drops of a 2:3 protein:condition ratio to determine whether a lower protein concentration would resolve the problem, but after daily assessment for 4 days almost identical results were observed.

Experiments on SbCfr (Section 4.3) suggested that additional reducing agent might contribute to reducing precipitation in crystallisation experiments. To that end a further series of the 5 initial screens containing 9 mg.mL reconstituted Cfr supplemented with 1mM tris(2-carboxyethyl)phosphine (TCEP) as reducing agent were dispensed, only to result in the same widespread precipitation observed previously.
4.3: Crystallisation screening for SbCfr

Following the unsuccessful crystallisation screening of Cfr, we progressed to SbCfr to determine whether its increased stability would make it more amenable to remaining soluble in the dispensed drops. Initially we repeated the same five preformulated crystal screens using the same 1:1 and 3:2 protein:condition mixes in drop 1 and 2 respectively, plating both reconstituted and unreconstituted SbCfr at 11 mg.mL. When compared to Cfr the improvement was immediately noticeable. Approximately ~40% of the drops had precipitated, though a large number had been successfully dispensed and were left to crystallise anaerobically at 19ºC.

Both drops of every condition in each of the 10 plates were manually assessed daily for two weeks after which no crystallisation had occurred. At this point another 4 preformulated crystal screens were selected to test: JCSG++ Eco screen(Jena Bioscience), ProPlex™ HT-96 Eco
Screen, PACT premier™ HT-96 and Structure Screen 1 + 2 HT-96 screen (Molecular Dimensions). Both reconstituted and unreconstituted SbCfr were dispensed into plates containing these screens, with the same 1:1 and 3:2 protein:condition mixes in drop 1 and 2 respectively.

All drops in the 18 active SbCfr plates were manually assessed daily for another two weeks, by which point no crystallisation had occurred and an additional ~40% of the drops had precipitated in the unreconstituted SbCfr plates and ~25% in the plates with reconstituted SbCfr. It was apparent that the reconstituted protein was more stable. It was also hypothesised that the addition of a reducing agent might help stabilise the protein even further.

Accordingly, a further 9 plates were dispensed, using all the preformulated crystal screens previously tested. Reconstituted SbCfr at 11mg.mL, supplemented with 1mM TCEP, was dispensed using the same 1:1 and 3:2 protein:condition mixes in drops 1 and 2, respectively, and allowed to incubate anaerobically at 19°C.

The following day every drop was assessed manually, and the formation of a multitude of small (20 µm x 20 µm) length uncoloured/clear crystals were observed in drop 2 (3:2 ratio of protein:condition) of Morpheus® HT-96 screen condition H7: 0.02M L-Na-Glutamate; 0.02M Alanine (racemic); 0.02M Glycine; 0.02M Lysine HCl (racemic); 0.02M Serine (racemic), 0.1M Sodium HEPES; MOPS, pH 7.5, 20% v/v Glycerol; 10% w/v PEG 4000. Figure 55 displays crystal formation in condition H7 of Morpheus® HT-96 screen. While these crystals were too small to loop, they were dyed using bromophenol blue to confirm they were protein as opposed to salt. Protein crystals contain solvent channels large enough to accommodate the small molecule dye, whereas salt crystals do not.
Figure 55 – Uncoloured crystal formation in well H7 of Morpheus® HT-96 screen. (A) Multitude of small 20 µm x 20 µm crystals in well H7. (B) Well H7 dyed with bromophenol blue to confirm protein crystals.

Prior literature reports Fe-S cluster proteins to produce brown-red crystals [175, 176], which in this case had not occurred. To mitigate this, 9 additional crystal trials utilising the same 9 preformulated reagent kits were dispensed using protein left in its reconstitution mixture (i.e. in the presence of excess FeSO₄) with 1mM TCEP. In addition, 9 crystal trials were dispensed under the same circumstances but containing 1mM SAM iodide, and a further 9 containing SAH.

These 27 plates were manually checked daily for two days. On the second a day an array of brown crystals was obtained in the Morpheus® HT-96 screen containing 1mM SAM. Drops B6 (1:1), H7 (3:2), D11 (1:1), E11 (1:1), E11 (3:2) and F11 (1:1) (6 different conditions) contained crystals. Drop B6 (0.03M Sodium fluoride, 0.03M Sodium bromide, 0.03M sodium iodide, 0.1M Sodium HEPES/MOPS pH 7.5, 20% v/v Ethylene glycol, 10% w/v PEG 8000) contained a single large brown crystal (dimensions approx. 120 µm x 100 µm), as shown in
figure 56 (A). Drop H7 contained 4 crystals, two light brown hexagonal crystals and two dark brown rectangular rod-shaped crystals, as shown in figure 56 (B). Drop D11 (1:1) (0.02M 1,6-Hexanediol; 0.02M 1-Butanol; 0.02M 1,2-Propanediol; 0.02M 2-Propanol; 0.02M 1,4-Butanediol; 0.02M 1,3-Propanediol, 0.1M Tris (base)/BICINE pH 8.5, 20% v/v Glycerol, 10% w/v PEG 4000) contained a range of small brown crystals; drop E11 (1:1) (0.03M Diethylene glycol; 0.03M Triethylene-glycol; 0.03M Tetraethylene glycol; 0.03M Pentaethylene glycol, 0.1M Tris (base)/BICINE pH 8.5, 20% v/v Glycerol, 10% w/v PEG 4000) contained 5 small brown crystals; drop E11 (3:2) contained one small brown rhombus shaped crystal and drop F11 (1:1) (0.02M D-Glucose; 0.02M D-Mannose; 0.02M D-Galactose; 0.02M L-Fucose; 0.02M D-Xylose; 0.02M N-Acetyl-D-Glucosamine, 0.1M Tris (base)/BICINE pH 8.5, 20% v/v Glycerol, 10% w/v PEG 4000) contained 4 small brown crystals.

All crystals from wells B6 and H7 were looped anaerobically and dipped into the well condition to act as cryoprotectant for 1 min, then flash frozen in liquid nitrogen and transferred to a uni-puck (Molecular Dimensions). 1 crystal from well D11, 1 crystal from E11 (1:1) and 2 crystals from F11 were looped in the same fashion.

Figure 56 – Brown crystals obtained from Morpheus® HT-96 screen drops B6 (1:1) and H7 (3:2) with 1mM SAM and 1mM TCEP. (A) Drop B6 contained a single large brown cuboid shape crystal (approx. 120 µm x 100 µm). (B) Drop H7 contained 4 crystals, two light brown hexagonal crystals and two dark brown rectangular rod-shaped crystals.
Figure 57 – Brown crystals obtained from Morpheus® HT-96 screen drops D11 (1:1), E11 (1:1), E11 (3:2) and F11 (1:1) with 1mM SAM and 1mM TCEP. (A) Drop D11 (1:1) contained a range of small brown (<Xum) crystals. (B) Drop E11 (1:1) contained 5 small brown crystals. (C) Drop E11 (3:2) contained one small brown rhombus shaped crystal. (D) Drop F11 (1:1) Contained 4 small brown crystals.

4.4: Diffraction data collection

The puck containing looped crystals was transported to Diamond Light Source and shot on the I24: microfocus MX beamline. Of all crystals to be looped 4 were fully intact: 1 from B6 and 3 from H7. All crystals were initially assessed for diffraction from three images, 45 degrees apart, at a wavelength of 0.9686 Å, exposure time of 0.040 seconds, resolution of 1.80 Å and a beamsize of 30x30 µm.

Figure 58 displays the three crystals from well H7, all of which diffracted to a low resolution (4 Å or worse). Figure 59 displays the single crystal from well B6, which was the largest crystal, which had been looped successfully, and with minimal contamination. Initial diffraction tests suggested that this crystal diffracted X-rays beyond 1.5 Å resolution.
Figure 58 – Three looped crystals from Morpheus® HT-96 well H7 (3:2). All three crystals were loaded on to the I24 beamline and assessed for diffraction from three angles. All data gathered from these crystals suggest they diffracted poorly (4 Å or worse). (A) Hexagonal crystal from H7. (B) Second hexagonal crystal from H7. (C) Rectangular rod-shaped crystal from H7.
Figure 59 – Single brown crystal from Morpheus® HT-96 well B6 (1:1) in three 45-degree angles tested for diffraction. Initial diffraction suggested that this dataset could exceed 1.5 Å resolution.

Following successful preliminary assessment of diffraction from the B6 crystal, a native dataset was collected at a 0.9686 Å wavelength, 0.010 second exposure and 20 µm x 20 µm beamsize (Table 3). An X-ray fluorescence scan was performed on this crystal to locate the Fe edge; a scan around the Fe edge (7080-7150 eV) revealed a clear peak indicating the presence of Fe in the crystal. A further diffraction data set was then collected at a wavelength (1.73002 Å) close to the peak (maximal f") wavelength for the iron edge, in order to collect anomalous data that might assist with phasing the macromolecular structure. Integration (see below) identified crystal symmetry as P orthorhombic, with a likely solvent content of 57 % indicating two molecules in the asymmetric unit. Scaling established resolutions of 1.42 Å for the high-resolution and 1.8 Å for the iron-edge datasets, respectively, as adjudged on the basis of the correlation coefficient CC 1/2. Data processing statistics are presented in Table 3.
Figure 60 – X-ray diffraction pattern of SbCfr B6 crystal. 0.010 second exposure (45° rotation) collected on DLS beamline I24.
4.5: Crystal structure solving

Initially the autosolving software (Autosharp, Autobuild and Crank2) hosted on the Diamond Light Source servers attempted to solve the structure in the P 2\textsubscript{1} 2\textsubscript{1} 2\textsubscript{1} space group, however these data yielded electron density maps that we could not satisfactorily model. Attempts to solve the phase problem using molecular replacement with the extant RlmN structure (PDB: 3RFA) also failed. With guidance from Dr Neil Paterson (Diamond Light Source) we noticed the Xia2 [177] pipeline had integrated and scaled the raw data in two space groups: P 2\textsubscript{1} 2\textsubscript{1} 2\textsubscript{1} and P 2\textsubscript{1} 2\textsubscript{1} 2\textsubscript{1}. We therefore took the P 2\textsubscript{1} 2\textsubscript{1} 2\textsubscript{1} data from the xia2 pipeline, after integration in XDS and scaling and merging by XSCALE [158] and used this dataset in Autosol [159] (Phenix [160]) to determine phase information. Experimental phases, an electron density map and an initial model were calculated, with two molecules in the asymmetric unit and 8 Fe atoms identified corresponding to the expected [4Fe-4S] content. Figures of merit were 0.36 before, and 0.68 after, density modification (DM), as shown in Table 3. The initial model built by AutoBuild [161] in Phenix was subsequently refined against the higher resolution (1.42 Å) isomorphous native dataset (0.9686 Å; indexed using XDS and combined and scaled using XSCALE in the xia2 pipeline as above). The model was completed by iterative rounds of manual model building in COOT [162] and refinement in Phenix. Structure validation was assisted by Molprobity [163] and Phenix. Figure 61 displays the Ramachandran plot describing 98.07% of residues in favoured positions, with 1.93% residing in acceptable positions and no outliers. The final structure contains 340 (chain A) and 339 (chain B) amino acids modelled as two continuous polypeptides.
### Table 3 – Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>Fe-edge</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>p 21 21 2</td>
<td>p 21 21 2</td>
</tr>
<tr>
<td>Molecules/ASU</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>131.8, 53.11, 113.54</td>
<td>131.8, 53.14, 113.57</td>
</tr>
<tr>
<td>a, β, γ (°)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Wavelength(s) (Å)</td>
<td>1.73002</td>
<td>0.96859</td>
</tr>
<tr>
<td>Resolution (Å)*</td>
<td>86.02 – 1.88 (1.91 – 1.88)</td>
<td>57.01 – 1.42 (1.44 – 1.42)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.1028 (0.7304)</td>
<td>0.1318 (2.6908)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.9969 (0.3512)</td>
<td>0.9971 (0.4936)</td>
</tr>
<tr>
<td>l / σ(l)</td>
<td>14.41 (0.96)</td>
<td>9.95 (0.95)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>86.47 (24.92)</td>
<td>99.99 (100.00)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>10.1 (1.83)</td>
<td>13.07 (12.58)</td>
</tr>
<tr>
<td><strong>Phasing Calculations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOM¹</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>FOM after DM²</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td></td>
<td>49.29–1.42</td>
</tr>
<tr>
<td>No. reflections</td>
<td></td>
<td>150914</td>
</tr>
<tr>
<td>Rwork / Rfree</td>
<td></td>
<td>0.166 / 0.187</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td></td>
<td>5360</td>
</tr>
<tr>
<td>Solvent</td>
<td></td>
<td>644</td>
</tr>
<tr>
<td>Fe-S cluster cofactor</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>SAH</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td></td>
<td>27.63</td>
</tr>
<tr>
<td>Solvent</td>
<td></td>
<td>40.42</td>
</tr>
<tr>
<td>Fe-S cluster cofactor</td>
<td></td>
<td>17.76</td>
</tr>
<tr>
<td>SAH</td>
<td></td>
<td>20.12</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td></td>
<td>1.109</td>
</tr>
<tr>
<td>Ramachandran (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outliers</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Favoured</td>
<td></td>
<td>98.07</td>
</tr>
</tbody>
</table>

* *Values in parenthesis are for highest resolution shell

¹ Figure of merit

² Figure of merit after density modification
Figure 61 – Ramachandran plot describing accuracy of amino acid residues modelled in to structure. 98.07% of residues reside in favoured positions, with no outlying residues. All residues are either in favoured positions, with only 8 residues in allowed positions. Residues plotted in β-sheets at the top left, residues plotted within α-helices in the middle-left and residues plotted within left-handed α-helices plotted in the middle-right.
4.6: Crystal Structure of SbCfr

The crystal structure of SbCfr reveals two protein molecules in the asymmetric unit, as shown in Figure 62. The two chains are near identical, with a Cα rmsd of 0.2677 Å over 339 residues. The core of the enzyme is composed of a partial (6-stranded) TIM (triosephosphate isomerase) β-barrel, hosting a [4Fe-4S] cluster coordinated by the canonical radical SAM enzyme CX₃CX₂C motif, with an N-terminal ancillary helical domain (residues 1 – 62) connected to the core by three additional β-strands as shown in Figure 63(A). In total, the structure contains 14 α-helices and 12 β-sheets, as shown in Figure 63(B). The [4Fe–4S] cluster sits above one end of the central β-barrel, with the CX₃CX₂C motif situated on an extended hairpin loop (residues 110 – 124) that closes over the open barrel end. The overall structure resembles that of E. coli RlmN (42% sequence identity - PDB: 3RFA) as shown in Figure 64 with a Cα rmsd of 1.78 Å over 320 residues.

PDB searches indicate that crystal structures are now available for 96 different radical SAM enzymes. Hence our structure was compared to a range of existing radical SAM crystal structures using the DALI search function [178]. Aside from six previously described E. coli RlmN structures (PDB: 3RFA, 3RF9, 5HR6, 5HR7, 4PL1 and 4PL2), more distant structural relationships with a range of other radical SAM structures of enzymes more distantly related to SbCfr (~11-19% sequence identity) were identified, four of which are shown in Figure 65 (A) 2A5H - Lysine-2,3-aminomutase from Clostridium subterminale, which shares 16% identity with SbCfr. (B) 3CB8 - [4Fe-4S]-Pyruvate formate-lyase activating enzyme in complex with SAM which shares 19% identity with SbCfr. (C) 5V1T - Crystal structure of Streptococcus suis SuiB, which shares 11% identity with SbCfr. (D) 5VSL - Crystal structure of viperin with bound [4Fe-4S] cluster and S-adenosylhomocysteine (SAH), which shares 15% identity with SbCfr. Like SbCfr, these structures utilise an incomplete TIM β-barrel to bind a [4Fe-4S] cluster, with access to the cluster possible from the incomplete side of the barrel. These enzymes also all contain the conserved GGE motif (GGD in 2A5H - Lysine-2,3-
aminomutase, a conservative substitution to another negatively charged amino acid) which is situated adjacent to the binding site of SAM.

*SbCfr active site*

Consistent with the location of 4 Fe sites per protein monomer by the AutoSol pipeline, final electron density maps clearly resolve the presence of an intact [4Fe-4S] cluster in the active site. This cluster is coordinated by the conserved cysteine residues of the canonical radical SAM CX₃CX₂C motif (Cys110, Cys114 and Cys117), with the fourth Fe atom (often termed the unbound iron) co-ordinated to a molecule of bound cofactor via the α-amino and α-carboxy moieties (distances 2.3 Å). This molecule of cofactor is clearly defined in electron density maps (Figures 66, 67) and is refined with B-factors below the average values for the protein and approaching those of the Fe atoms (Table 3). The quality and resolution of our final structure allow us to assign this as SAH, rather than SAM. In addition to interactions involving the amino and carboxyl groups (above) SAH is oriented such that the sulphonium ion resides 3.2 Å away, from the unbound iron atom (Figure 68). The distance measurements for SAH:Fe interactions are typical of those for cofactor interactions with a range of structurally characterised radical SAM enzymes [175, 176]. As expected Cys106 is present in the active site, although its positioning makes it unlikely that it plays a catalytic role.

SAH would be produced as a result of the relocation of the methyl group from SAM to Cys334, thus its presence in crystals exposed to excess SAM might indicate high levels of methylation of the methyl acceptor Cys334. Inspection of Fo-Fc electron density maps reveals elongated electron density at the sulphur atom of Cys334, compared to similar maps for residues not subject to modification (Figure 69). Cys334 was thus refined in the methylated form, with the methyl group refined at an occupancy of 1.0 and with a B-factor of 62.61 in the finished model (compared to a B-factor of 51.23 for the adjacent sulphur atom). Notably, the entire 6 amino acid residue loop (aa. 332-337; AACGQL) near to the C-terminus of the protein, and on which methylated Cys334 (SMC) resides, is well defined in electron density maps for both chains of
our structure. In contrast, in RlmN crystal structures determined in the absence of bound RNA, this is loop is poorly defined and/or unmodeled (see below).

Inspection of electron density (Fo-Fc omit map) around the active site region also identified a well-defined water molecule positioned close (2.1 Å proximity) to cysteine 103 (Figure 70). In studies of RlmN Cys103 has been implicated in the resolution of the transient covalent RNA:protein [179] suggesting involvement of a bound water molecule in this location in this process.

**Figure 62 - Asymmetric unit of SbCfr crystal.** The two subunits are near identical, with a SSMSuperpose highlighting a Cα rmsd of 0.2677 Å. Chain A is in green, Chain B is in blue.
Figure 63 – SbCfr folding showing secondary structural elements. (A) Chain terminus orientation. C-terminus is highlighted in red and N-terminus highlighted in blue. (B) \( \alpha \)-helix and \( \beta \)-sheet content. In total the structure contains 14 \( \alpha \)-helices (red) and 12 \( \beta \)-sheets (yellow).
Figure 64 - Structural alignment of SbCfr (green) bound to SAH and RlmN (red) bound to SAM.
Figure 65 – Range of radical SAM enzyme crystal structures. Many other radical SAM enzymes utilise the partial β-barrel structure to facilitate access to the 4Fe-4S groups. (A) 2A5H - Lysine-2,3-aminomutase from Clostridium subterminale. (B) 3CB8 - 4Fe-4S-Pyruvate formate-lyase activating enzyme in complex with SAM (C) 5V1T - Crystal structure of Streptococcus suis SuiB. (D) 5VSL - Crystal structure of viperin with bound [4Fe-4S] cluster and S-adenosylhomocysteine (SAH).
Figure 66 – Representative 2Fo-Fc electron density of the SbCfr active site. 2Fo-Fc electron density of the SbCfr active site, contoured at 1 σ, highlighting SAH and the [4Fe-4S] cluster.
Figure 67 –2Fo-Fc Electron density surrounding SAH ligand. In contrast to SAM, SAH is lacking a methyl (CH₃) group on its sulphur atom. The density surrounding this atom in our structure accounts for the density expected from a sulphur atom, but not for a methylated sulphur atom. 2Fo-Fc density (blue mesh) contoured at 1 σ.
Figure 68 – SbCfr active site. A 4Fe-4S cluster is coordinated by the canonical radical SAM enzyme CX₃CX₂C motif. SAH is bound via the molecules α-amino and α-carboxy moieties to the uncoordinated Fe at 2.3 Å, alongside the sulfonium ion which resides 3.2 Å away. Three additional Cysteines are present: Cys334, a key catalytic residue, Cys103, which is in 2.1 Å to a water molecule.
Figure 69 – Fo-Fc electron density surrounding an unmethylated and methylated cysteine. Fo-Fc electron density suggests the presence of a methylated sulphur on Cys334 when compared to the density seen on an unmethylated cysteine. (A) Cys103, an unmethylated cysteine. (B) Cys334, Fo-Fc data displays elongated density originating from the sulphur atom, suggesting methylation. Electron density (green mesh) contoured at 3 \( \sigma \).

Figure 70 - Fo-Fc omit map (green) displaying the presence of a water molecule in 2.1 Å proximity to cysteine 103. Fo-Fc data (green mesh contoured at 3 \( \sigma \)) suggests the presence of a water molecule in close (2.1 Å) proximity to Cys103, an important residue involved in the resolution of the covalent RNA:protein adduct. 2Fo-Fc density (blue mesh) is contoured at 1 \( \sigma \).
Relationship of SbCfr Structure to Structures of RlmN

As one of the most heavily studied radical SAM enzymes, and the prototypical radical SAM RNA methyltransferase, RlmN provides an obvious point of comparison for the SbCfr structure. As described above, six crystal structures of RlmN are available: the native enzyme (3RF9); its complex with SAM (3RFA); and the C118A mutant (4PL2); SAM-bound (4PL1); E. coli tRNA-bound (5HR6) and in vitro transcribed tRNA-bound (5HR7) forms. Here the SbCfr structure is compared with these structures, focusing in particular on binding of SAM/SAH and on the environment of the methylated Cys334.

Figure 73 displays the distances observed between the SAM ligand and the RlmN [4Fe-4S] cluster in comparison with binding of SAH to SbCfr. In the SAM-bound RlmN structure (3RFA) the distances between the unbound Fe of the [4Fe-4S] cluster and the SAM sulphur atom and α-amino and α-carboxy moieties are 3.2 Å, 2.3 Å and 2.5 Å, respectively. The distance between the SAM sulphur atom and the nearest sulphur within the cluster is 3.8 Å, the distance between this sulphur atom and the nearest carbon of the adenine ring is 4.2 Å, and the distance between the same sulphur and the exocyclic nitrogen 8.1 Å. In the SAH-bound SbCfr structure the distances between the unbound Fe and the SAH sulfonium sulphur, α-amino and α-carboxy moieties are 3.2 Å, 2.3 Å and 2.3 Å respectively. The distance between the sulfonium sulphur and the nearest sulphur atom within the cluster is 3.8 Å; the distance between the nearest sulphur and the nearest adenine ring carbon is also 3.8 Å, and the distance between the nearest Sulphur and the exocyclic nitrogen is 7.8 Å. The last two values are shorter than their equivalents in the RlmN:SAM complex (see above) indicating a reduced distance between the [4Fe-4S] cluster and SAH, when compared to SAM, specifically with respect to the positioning of the adenine ring.
Figure 71 - Distances between SbCfr and RlmN [4Fe-4S] clusters and their respective ligands. (A) SAM bound to RlmN [4Fe-4S] cluster. The distances to the uncoordinated Fe for the Sulphur atom, and the α-amino and α-carboxy moieties are 3.2 Å, 2.3 Å and 2.5 Å respectively. The distance between the SAM sulphur atom and the nearest Sulphur within the cluster is 3.8 Å, the distance between the SAM sulphur atom and the closest C of the adenine ring is 4.2 Å, with the distance between the same sulphur and the adenine exocyclic nitrogen 8.1 Å. (B) SAH bound to SbCfr [4Fe-4S]. The distances to the uncoordinated Fe for the Sulphur atom, and the α-amino and α-carboxy moieties are 3.2 Å, 2.3 Å and 2.3 Å respectively. The distance between the sulphur atom and the nearest Sulphur within the cluster is 3.8 Å, as is the distance between the nearest Sulphur and the closest C of the adenine ring. The distance between the nearest Sulphur and the exocyclic nitrogen is 7.8 Å.
SAM binding elicits conformational changes in RlmN affecting residues Cys118 and Met176 adjacent to the ligand binding site. In the RlmN crystal structure with no bound ligand (3RF9) Cys 118 is positioned facing away from the ligand binding site, 4.4 Å from the Met 176 carbonyl oxygen. Upon binding of SAM (3RFA), Cys 118 changes rotamer, shifting the thiol group 2.8 Å towards the ligand binding site and inducing a flip in the Met 176 – Gly 177 peptide bond resulting in a shift of 2.4 Å of the Met 176 carbonyl oxygen towards the ligand. This conformational change reduces the distance between Cys 118 and the Met 176 carbonyl oxygen from 4.4 Å with no bound ligand, to 3.5 Å after ligand binding. In the present SAH-bound structure of SbCfr, the conformations of the equivalent residues Cys103 and Met157 aligns closely with those of the RlmN SAM-bound structure, with Cys 103 positioned 3.5 Å from the Met 157 carbonyl oxygen. This data highlights potential mechanistic similarities between RlmN and SbCfr.

Two mutant RlmN crystal structures bound to tRNA have been published [CITATION], both with a Cys 118 to alanine mutation. This mutation[171, 176] removes the enzymes capacity to resolve the intermediate covalent interaction between the methylated Cys 334 (SMC) and C2 of the bound adenine residue, resulting in an RNA:Protein adduct. [176]Crystal structure alignments of RlmN bound to SAM and the RlmN C118A mutant bound to RNA, Methionine and DOA display a significant conformational shift in the position of the methylated Cys 334. Figure 74 displays a structural comparison between the methylated Cys 334 loop in the SAM bound RlmN and the tRNA bound C118A RlmN. In the SAM bound structure the methylated Cys 334 sulphur is positioned 5.3 Å away from the SAM sulphur atom. In the tRNA bound C118A RlmN structure the methylated Cys 334 loop is in a different conformation, facing the tRNA binding site, with the distance between the Cys 334 sulphur and SAM sulphur at 7.8 Å. When comparing the two structures the distance between the two positions of Cys 334 is 6.9 Å. Comparing these structures to our SbCfr structure shows that during SAH binding the methylated Cys 103 loop aligns closely with the RNA, methionine and DOA bound RlmN loop.
conformation. Figure 75 displays this alignment, wherein the distance between the SAH sulphur atom and the Cys 103 sulphur atom is 7.9 Å.
Figure 72 – RlmN Cys 118 rotamer shift and Met 176 flip upon SAM binding. RlmN with no ligand bound (blue) hosts a Cys 118 which faces away from the ligand binding site and a Met 176 carboxyl oxygen 4.4 Å away. Upon SAM binding Cys 118 undergoes a change in rotamer and shifts 2.8 Å towards the ligand binding site, at the same time Met 176 flips, with the carboxyl oxygen moving 2.4 Å towards the ligand binding site. During this movement the distance between Cys 118 and the Met 118 carboxyl oxygen diminishes from 4.4 Å to 3.5 Å.

Figure 73 – SbCfr overlay of RlmN Cys 118/Met 176 site showing SbCfr residues in ligand binding conformation. SbCfr overlay of RlmN Cys 118/Met 176 site shows SAH bound SbCfr to occupy the same positions as the ligand bound RlmN structure.
Figure 74 – Structural alignment of methylated Cys 355 loop of RlmN bound to SAM (red) and C118A RlmN bound to tRNA A37, methionine and DOA (Cyan). In the SAM bound structure (Red) the methylated Cys 355 sulphur is positioned 5.3 Å away from the SAM sulphur atom. In the tRNA bound C118A RlmN structure the methylated Cys 355 loop is in a different conformation, facing the tRNA binding site, with the distance between the Cys 355 sulphur and SAM sulphur at 7.8 Å. When comparing the two structures the distance between the two positions of Cys 355 is 6.9 Å.
Figure 75 – Structural alignment of methylated Cys 355 loop of RlmN bound to SAM (red) and C118A RlmN bound to tRNA A37, methionine and DOA (cyan) alongside the SAH (green) bound SbCfr methylated Cys 334 loop. When compared the SbCfr Cys 334 loop aligns closely with the tRNA bound C118A RlmN Cys 355 loop. The distance between the Cys 334 sulphur and the sulphur of the SAH ligand is 7.9 Å.
4.7: Discussion

The unsuccessful crystallisation trials utilising Cfr highlight the structural vulnerability, i.e. poor stability and high oxygen sensitivity, of the enzyme. During purification in the COY anaerobic chamber at 0 ppm O\textsubscript{2} the reconstituted samples remained soluble provided that they were stored at low temperature. Upon introduction to the crystallography glovebox, a ~6 ppm O\textsubscript{2} environment, precipitation of these samples began almost immediately after they were unsealed and was evident in almost all of the wells of crystallisation screening experiments that were attempted. ~10 crystallography plates of reconstituted Cfr have since been attempted manually in the 0 ppm O\textsubscript{2} COY anaerobic chamber, but >90% of the wells precipitated within the first hour and no crystals have been detected in the remaining wells. Successful crystallisation of Cfr has so far not been reported. In contrast SbCfr remained stable at room temperature in a 0 ppm O\textsubscript{2} environment, and in the 6 ppm O\textsubscript{2} environment after being unsealed. This, taken alongside the evidence for presence of [4Fe-4S] clusters in the UV-Vis wavelength scan and for secondary structure in reconstituted CD spectra, further evidences of the stability of SbCfr and indicates its suitability for crystallisation trials.

The initial SbCfr crystallisation plates proved unsuccessful, though protein in a large percentage of the wells remained soluble. In the hope that reduction of the [4Fe-4S] cluster would contribute to overall protein stability and therefore crystallisation, it was considered prudent to add a more stable reducing agent that the 5 mM Dithiothreitol (DTT) present in the buffer SbCfr was stored, as DTT remains stable in solution for a short time 12-40 hours. The addition of 1 mM TCEP to the SbCfr aliquot prior to dispensing proved successful, with the emergence of small and uncoloured crystals in one condition (Morpheus® HT-96 reagent mix well H7) in less than 24 hours. As a range of published literature regarding the crystallisation of enzymes containing Fe-S clusters has documented the brown nature of the crystals obtained, it was hypothesised that under the crystallisation conditions the protein was potentially losing its [4Fe-4S] cluster, but still retaining its structure, potentially because of the overall high stability of SbCfr. To overcome this, it was decided that SbCfr would be dispensed
into crystallisation experiments in its reconstitution liquor, supplemented with 1 mM SAM. After dispensing the protein in identical conditions to the previous successful well it took less than 24 hours for the discovery of deep brown crystals in well H7 and B6, as well as several less well-formed crystals in other wells. A single large crystal from condition B6 yielded a native dataset to a resolution of 1.42 Å; following this the same crystal was shot at the Fe-edge wavelength of 1.73002 Å, which provided a 2.25 Å dataset.

Molecular replacement implemented in the Diamond automated pipeline was unsuccessful in solving the SbCfr structure, despite the multiple RlmN structures in the PDB. Similarly, the solving of these datasets using anomalous methods in the P 2₁ 2₁ 2₁ space group initially yielded density that could not be satisfactorily modelled or refined. Ultimately successful structure solution followed correct identification of the P 2₁ 2₁ 2 space group, which was not always achieved in the automated pipeline.

The high resolution and high quality of our electron density help to define aspects of the active site architecture and ligand chemistry. For example, electron density surrounding the sulphur atom on the SAM ligand clearly displays absence of a methyl group, indicating that the bound species is SAH rather than SAM. This was unexpected considering that, in order to methylate the sample, the protein was previously treated with excess SAM and, following this treatment, the entire sample was desalted to remove any unbound material. As shown by the chromatogram for SAM standards (Figure 40) negligible quantities of SAH were present in the initial sample; suggesting that SAH has been generated in situ by activity of SbCfr and remains bound to the enzyme even in the presence of excess SAM. The implication is that SbCfr shows preferential binding of SAH over SAM.

Assignment of the ligand as SAM rather than SAH was one point of consideration in the decision to model Cys334 as methylated, as loss of the methyl group from SAM would be consistent with its presence on the enzyme. This, in conjunction with the elongated electron density at the Cys334 site of methylation, compared to that observed for an unmodified cysteine residue, convinced the author that this position was methylated. Observation of a
methyalted cysteine is consistent with structures for wild-type RlmN obtained in the presence of SAM, where the equivalent Cys355 is modelled as methylated, although in RlmN structures the C-terminal portion of the protein is poorly defined and either associated with higher than average B-factors or entirely absent from the final model.

The conformation of the C-terminal (Cys334) loop, and its relationship to RlmN structures, is of particular interest. When SbCfr was overlaid onto RlmN, the SbCfr methylCys334 loop aligned closely with the C-terminal Cys355 loop in RNA-bound RlmN, and not the SAM-bound conformation. While both the SAM-bound RlmN and SAH-bound SbCfr structures possess methylated cysteines within this loop, they are in clearly different conformations and thus could potentially reflect different states on the catalytic pathway. The SAM-bound RlmN structure is considered to represent binding of the second SAM equivalent, responsible for 5' DOA radical generation, to the enzyme, with the first having methylated Cys355. This would explain the presence of both the methylated Cys and the SAM ligand. Conversely, the SAH-bound SbCfr structure would represent the structure after binding of the first molecule of SAM, methylation of Cys334 and conversion to SAH, and as such could be an earlier stage in the reaction cycle.

This explanation is however complicated by the clear conformational shift that can be seen in the SbCfr C-terminal loop relative to its position in RlmN. During the binding of SAM to RlmN, this loop is positioned in a way which could be conducive to methyl transfer between SAM and Cys355, with the two sites only 3.3 Å apart. Conversely, Cys118Ala RlmN bound to RNA displays a significant shift upon RNA binding (around 6.9 Å movement of Cys355 away from its previous position). It has been proposed that interactions with bound RNA support this conformational change and bring Cys355 into position to react with the adenine methyl acceptor by moving the methyl Cys group closer to the DOA• radical produced by the reductive cleavage of SAM by the [4Fe-4S] cluster. It is this DOA• radical which is thought to abstract a hydrogen from the Cys355-bound methyl group to catalyse formation of the covalent bond to RNA. Supporting this contention, in the absence of RNA, the Cys355 loop is weakly defined or absent in RlmN structures. However, in SbCfr the corresponding region is well defined with
methylCys334 apparently positioned to react with an incoming RNA cosubstrate. This could suggest either that conformational change in the C-terminal loop can take place on methyl transfer to Cys334, and does not require binding of RNA, or that this conformation is a result of the presence of SAH, rather than SAM, with the implication that SAH can “trap” the protein in a conformation that is normally accessed only when RNA is present.

A further mechanistically relevant observation is that our structure clearly provides evidence for the presence of a water molecule, (in both chains and confirmed by Fo-Fc electron density omit maps), in very close proximity (2.1 Å) to SbCfr Cys103. This is significant as the equivalent residue (Cys118 in RlmN, Cys105 in Cfr) has been proposed to act as a general base during resolution of the covalent RNA complex, and C8 methylation catalysed by Cfr has very recently been shown to involve exchange of protons with solvent [179], suggesting involvement of a water molecule in the reaction. However, equivalent exchange has not been observed for (RlmN-catalysed) C2 methylation, and available RlmN crystal structures do not contain an appropriately positioned water. Our structure could thus potentially provide the first evidence of this mechanistically relevant water molecule close to the candidate general base.
Chapter five: General discussion and reflections

While Cfr is a newly emerging antimicrobial resistance threat, the importance of characterisation of Cfr and Cfr-like enzymes cannot be understated. Under certain conditions the acquisition of cfr in a clinical setting, especially in conjunction with other AMR determinants, could prove fatal to human life. Cfr is unusual among antimicrobial resistance determinants in the range of agents against which producer bacteria are protected (the PHLOPS_A antibiotics) and the breadth of its distribution which extends to both Gram-positive and Gram-negative bacteria. However, no Cfr or Cfr-like enzyme has been structurally characterised. At the outset of this project we identified 4 main targets for investigation: Cfr, SbCfr, Cfr(B) and PbCfr. Upon the identification of the novel Cfr-like enzyme Cfr(C), which was reported during this project, it was added to the roster of candidates for characterisation.

Here we have shown that Cfr and SbCfr can be recombinantly produced and purified to homogeneity. The yield of purified reconstituted Cfr was 10.25 mg.L⁻¹ of E. coli culture, which is an increase on the 7.1 mg.L⁻¹ yield seen by Booth et al [157], and the 0.3 mg.L⁻¹ seen by Yan et al [96] in the absence of the isc proteins, strongly highlighting the importance of the inclusion of pDB1282. The yield of purified reconstituted SbCfr was 22.4 mg.L⁻¹, more than twice the yield of Cfr. This higher Cfr yield compared to earlier publications could be attributed to the modifications made to the original protocol. The decision to include a much longer lysozyme incubation period before sonication could have resulted in increased lysis and therefore a lower time subjected to sonication, potentially resulting in a reduced denaturation and an increase in properly folded enzyme. Previous attempts by other researchers in the same laboratory had shown unreconstituted Cfr incapable of converting SAM to DOA above a negligible quantity (M.R. Challand, unpublished), this is the first time that DOA production in a quantity almost comparable with that produced by reconstituted Cfr has been observed.

As Cfr(B), Cfr(C) and PbCfr did not express to any detectable levels in E. coli BL21*, they were only subjected to preliminary investigation due to the successes seen with Cfr and SbCfr, and further investigation is warranted. The pCDF Cfr(B), Cfr(C) and PbCfr constructs could be
re-evaluated and redesigned in an effort to elicit expression. Change of expression vector, relocation of the 6-His tag to the opposing termini and expression attempts in further *E. coli* expression strains are three changes that could be easily implemented.

We were able to prove that Cfr and SbCfr were biochemically active through investigation into SAH production, and demonstration of DOA production of in the presence of reductant and the absence of RNA substrate. The UV-Vis spectra, alongside the DOA formation assays, indicate that reconstitution likely results in the incorporation of [4Fe-4S] clusters in enzyme without intact Fe-S cofactor. The UV-Vis spectrum of SbCfr before reconstitution possesses the 420 nm peak associated with intact [4Fe-4S] incorporation, which is increased in intensity after reconstitution. In contrast the unreconstituted Cfr UV-Vis spectrum lacks this 420 nm peak, which is introduced upon reconstitution. These data, in conjunction with the circular dichroism spectra displaying intact α-helices up to at least 65°C, highlight the stability of SbCfr through purification and under heat challenge.

Attempts to replicate the RNA:protein adduct formation experiments performed by McCusker *et al* on RlmN were unsuccessful. Our initial hypothesis that SbCfr may be incapable of binding to an *E. coli* RNA substrate was tested by the inclusion of the pACYCT2 construct which encoded an RNA product containing the homologous *S. thermophilus* sequence harboured within a tRNA scaffold. In practice this experiment yielded no detectable extra heavier molecular weight bands, indicating the absence of any RNA:protein adduct. However, the expression profile of the RNA substrate has yet to be assessed and optimisation of this protocol remains ongoing. In the future it would be prudent to pursue investigation into the profile of rRNA substrate binding of SbCfr. The *in-vitro* production of a range of lengths of *S. thermophilus* 23S rRNA transcripts that span regions of known importance (helices 89-93) and subsequent incubation with SbCfr could elucidate the impact of an RNA substrate on DOA production and the respective mechanistic roles of Cys103 and the additional Cys 106. This experiment also assumes that the target of SbCfr methylation is the 23S rRNA; although RlmN is demonstrated to methylate tRNA, as well as rRNA. It may be necessary to investigate a
wider range of candidate RNA substrates, extending beyond 23S rRNA, for SbCfr, in order to demonstrate activity. It is also assumed, based on the failure of recombinant *E. coli* expressing SbCfr to demonstrate the PHLOPS\(_A\) phenotype in antimicrobial susceptibility tests, that SbCfr possesses C2, and not C8, methyltransferase activity. As trapping of covalent RNA adducts formed by Cfr has not been demonstrated by this approach, suggesting that this does not work for C8 methyltransferase enzymes, it is possible that SbCfr possesses C8 methyltransferase activity at a site not associated with reduction in antimicrobial sensitivity.

While attempts to crystallise Cfr were unsuccessful, the successful crystallisation and determination of the SbCfr crystal structure justify our initial hypothesis that this was a tractable candidate for investigation. We were able to show that SbCfr is indeed a radical SAM enzyme, hosting a [4Fe-4S] cluster ligated by the canonical CX\(_3\)CX\(_2\)C tri-cysteine motif, located inside a partial TIM \(\beta\)-barrel, much like RlmN. By crystallising SbCfr we were able to contribute evidence to the idea that a thermostable homologue could provide biochemical and crystallographic data about a protein in the event that the product of interest cannot be recombinantly produced or is resistant to crystallisation. This approach has been adopted to yield structures of thermophilic homologues of other radical SAM enzymes [180] and for other high profile but challenging structural biology targets [181].

Comparisons between SAM-bound RlmN (3RFA), RNA-bound RlmN (5HR7) and our SAH-bound SbCfr structure reveal interesting characteristics of these structures related to the conformations of the C-terminal (methylCys) loop and with possible mechanistic implications. The discovery that methylCys334 and the associated loop in the SAH bound SbCfr structure occupy the same orientation as the equivalent residues in the RNA bound RlmN structure, but not that in the RlmN:SAM complex, indicates that movement of this loop may occur after methylation of Cys334 and SAM conversion to SAH, implying that RNA binding does not drive this conformational change. This is consistent with previous observations that Cys methylation is required for reductive cleavage (DOA production) of the second SAM equivalent (SAM2; source of the DOA radical), and the observed modification to the active site could explain how
radical SAM methyltransferases use a single Fe-S cluster to bind two molecules of SAM that undergo two different fates.

Further mechanistic relevance of our structure comes from a recent publication by Bauerle et al [179] highlighting the probable involvement of a tightly bound water molecule in the proximity of the active site monoprotic base, likely Cys 118, in the C8 methylation of A2503 by Cfr. The presence of a water molecule in the proximity of Cys 103, the homologous residue, seen in our SbCfr structure could potentially provide the first crystallographic evidence for this mechanism. As this study argues that such a water molecule may be involved in C8, but not C2, methylation, and equivalent water molecules are not observed in available RlmN crystal structures, this might imply that SbCfr either differs mechanistically from RlmN, and/or acts at a substrate and position other than the adenosine C2 atom. Identification of a suitable RNA substrate for use in in vitro experiments would go some way to resolving such issues, and as well as providing a route to identifying the precise nature of the catalysed methyl transfer reaction could permit solvent proton exchange experiments confirming involvement of the observed water in catalysis.

Although efforts to crystallise Cfr were not successful, successful generation of high-quality SbCfr crystals opens the door to a new range of crystallisation trials experimenting with potential cross-homologue crystal seeding. Work performed by Abuhammad et al [182] has displayed the potential for using crystals from one homologue as nucleation material for another in seeding experiments. In the future crystal plates for experiments crystallising Cfr should be dispensed using SbCfr crystal seed in an effort to determine if this approach might yield any successful results. This could also be applied to any of the other Cfr-like homologues should difficulties with their expression be overcome.

In conclusion, work presented here demonstrates that SbCfr is a member of the radical SAM methyltransferase enzyme family with properties (thermostability, stable [4Fe–4S] cluster) that suit it to structural and biochemical characterisation. Although investigations of Cfr did not yield a crystal structure, a high-resolution structure of SbCfr was obtained and reveals an
unexpected orientation of the methylCys loop, observed during SAH binding, with potentially novel implications for the mechanism of action for radical SAM methyltransferases. Although the crystallisation and structural characterisation of the wildtype Cfr enzyme was unsuccessful, the study has been able to obtain new information regarding the incompletely characterised mechanism of RSMTs, enzymes that are of both great biochemical interest and considerable clinical importance to the growing field of antimicrobial resistance.
Chapter six: References and list of figures


110. Baos, E., et al., Characterization and monitoring of linezolid-resistant clinical isolates of Staphylococcus epidermidis in an intensive care unit 4 years after an outbreak of


Appendix 1: Sequence alignment of Wildtype Cfr, SbCfr, Cfr (b), Cfr (c), S. aureus RlmN and Paenibacillus Cfr.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staph Cfr</td>
<td></td>
<td>MNF-NKTGYGKIQEFLRSNNEP</td>
</tr>
<tr>
<td>SbCfr</td>
<td></td>
<td>MAHHHHHSSGLEVLQFGPMALYDLTIALEELRAADGVP</td>
</tr>
<tr>
<td>Cfr(b)</td>
<td></td>
<td>MQ-QKNKYIRQRQFLKQNKFP</td>
</tr>
<tr>
<td>Cfr(c)</td>
<td></td>
<td>MSKYKKMKQOL1ADMRLP</td>
</tr>
<tr>
<td>S. aureus RlmN</td>
<td></td>
<td>MITAEKKNKLFNFDKQSTSYSLRFDEMQNWLVQEQOO</td>
</tr>
<tr>
<td>Paenibacillus Cfr</td>
<td></td>
<td>MAHHHHHSSGLEVLQFGPMK-YLSKYEKIRKLISNLQP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>41</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staph Cfr</td>
<td></td>
<td>DYRIKQITNAIFKQRISRFDKMLPKLLREDLINTNGGET</td>
</tr>
<tr>
<td>SbCfr</td>
<td></td>
<td>RYRARQTFHAYRQLAVDYDAMTVKFLTRADLTRLPLT</td>
</tr>
<tr>
<td>Cfr(b)</td>
<td></td>
<td>NYRNQITTNAIFPGIRNSFETIVLPSKLRLMLIEFGES</td>
</tr>
<tr>
<td>Cfr(c)</td>
<td></td>
<td>EYRYQLLDAVFLQGIRMFEDKLLPKLRTKLRQKEVGPEG</td>
</tr>
<tr>
<td>S. aureus RlmN</td>
<td></td>
<td>KFRAKQIFEWLYQKVRDSIDEHMNLKDLRQQLKDQNTVT</td>
</tr>
<tr>
<td>Paenibacillus Cfr</td>
<td></td>
<td>NYSQITEAIIFKNKIGNFEAMNNLFPVRNLKIELGNN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>81</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staph Cfr</td>
<td></td>
<td>VLNKLAEQSNSEQVTKLFVESKNERVEVTNMKYGKE</td>
</tr>
<tr>
<td>SbCfr</td>
<td></td>
<td>PLTVREVQTDGTITKTLRFTVDEQEHEITVLMWYEDR-T</td>
</tr>
<tr>
<td>Cfr(b)</td>
<td></td>
<td>ILINIPLKAQQSTQVKVLQISGDEKIKETVNMKYGKE</td>
</tr>
<tr>
<td>Cfr(c)</td>
<td></td>
<td>VVEIKAHHEKMQTDKLVFELSDGNYETVGLFQKEFEN</td>
</tr>
<tr>
<td>S. aureus RlmN</td>
<td></td>
<td>TLTVKQESK-DGTIKFLQDGYTETVLMQRYDYG-N</td>
</tr>
<tr>
<td>Paenibacillus Cfr</td>
<td></td>
<td>VLSITPKMEQKSNQVSKILFAIPGDEIESVRLSYQGWE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>121</th>
<th>160</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staph Cfr</td>
<td></td>
<td>SFCISSQCGCNFGCKFATGDIGMRLKLKNLTVEIDTVQNYF</td>
</tr>
<tr>
<td>SbCfr</td>
<td></td>
<td>TVCVSCQVQGCAVGCSFCATGMMLRNLTAGEMVAQVAAA</td>
</tr>
<tr>
<td>Cfr(b)</td>
<td></td>
<td>SFCISSQCGCNFGCKFCATGDIGMRLKLKNLTDEIDTVQNYF</td>
</tr>
<tr>
<td>Cfr(c)</td>
<td></td>
<td>SFCISSQCGCNFGCKFCATGTLGLRLNLTDVQNYF</td>
</tr>
<tr>
<td>S. aureus RlmN</td>
<td></td>
<td>SVCVTTQVCRGCRCTFCASTGGLGLRNLEAGEIVSQVLT</td>
</tr>
<tr>
<td>Paenibacillus Cfr</td>
<td></td>
<td>SYCISISSQCGCNFGCCTFCATGTLGLKLNRLLTDVQNYF</td>
</tr>
</tbody>
</table>

148
Staph Cfr  H------LLGHQIDSISFMGMGEALANRQ-VFDALDSFTDPN
SbCfr  ARRAEAGRTLTVNMGMGEFPQNYEATOMVRILHREE
Cfr(b)  H------LQGHSIDSISFMGMGEALANVQ-VFDALNVLTDPA
Cfr(c)  M------QOGCSINSISFMGMGEFPANEQ-VFELAHDLTAP
S. aureus RlmN  QKALDTEERVQIVQMIGEPENFYDEMMDLFRIVNDDN
Paenibacillus Cfr  T------LNNHFLDSVSFMGMGEALANFY-VFDALHLLTDPK

Staph Cfr  LFALSPRRLSISTIGIIPSIKKIQYQPVQVNLTSFLHSPY
SbCfr  GMNFAGARRITVSTSLVPLPFIDRLAREPFFQVKLAVSLHAPN
Cfr(b)  LFALSPRRLSISTIGIIPNI1KKLTQNYPQVNLTSFLHSPF
Cfr(c)  LFGLSKARITISTIGVPGIQKLTREYPQVNLAYSILHAPT
S. aureus RlmN  SLDNAGRHTIVSTSGIIIPRIYDFADEDIQINFAVSLHAAK
Paenibacillus Cfr  LFGLGHRRITVSTIGLPLPGVKKLTKFQVQINLSFLHSPF

Staph Cfr  SEERSKLMIPINDRYPIDEVMNILEDEHIRTSEKTVYAYIM
SbCfr  DLLRSSLVPLNKRYPFGEIAACRRYVGETGRRVTFEYVL
Cfr(b)  NEQRSELMPINEFYLSDVMNTDLDEHIRTSEKTVIYAYIM
Cfr(c)  DRLRETLMIPITKTYFLMFLDLQVLDLQHRTQTNKRFLAYIM
S. aureus RlmN  DEVRSLMPINRAYVEKLIEATQYYQEKNNRVTFEYGL
Paenibacillus Cfr  HDQRSELMPINHNFFPLEEVMTVLDEHIQQTKRKVYDIYIL

Staph Cfr  LPGVNDSELHANEVSVLLKRSRKYSGKLYHVNLIYRNTIS
SbCfr  IDGVNDSDNAXELARLLRG------ILCHVNLIPLNPTPA
Cfr(b)  LHGVNDSIHEAKEVNNLKGKRSKNLYHVNIIYRNTPSV
Cfr(c)  IKDVSNDRHARQILTKLLFHKHKKYFLYHDILTPYNTTV
S. aureus RlmN  FGGVNDQLEHARELAIKG------INCHYNFLPVNHEVE
Paenibacillus Cfr  LRGINDSTKHAKAVDDLRLERSWHELHYNLIPYNSDA
<table>
<thead>
<tr>
<th>321</th>
<th>360</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staph Cfr</td>
<td>APEMYGEANEGQVEAFYKVLKSAGIHVTIRSQFGIDIDAA</td>
</tr>
<tr>
<td>SbCfr</td>
<td>A--PFGRPSVERINRFQILRARGIPATVYRSGVDISAA</td>
</tr>
<tr>
<td>Cfr(b)</td>
<td>SRMRFEEANEKCLVNYFKKLKSAGIKVTIRSQFGIDIDAA</td>
</tr>
<tr>
<td>Cfr(c)</td>
<td>TETM--VPSHHTRIKAFCRRIIHNAGISINIRTFQFGSDINAA</td>
</tr>
<tr>
<td>S. aureus RlmN</td>
<td>R--NYVTAKNDIFKEKELKRLGINATIRREQGSGIDAA</td>
</tr>
<tr>
<td>Paenibacillus Cfr</td>
<td>TSQSFVESDQNSINMFLRILKSGIHVTVRTQFGSDINAA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>361</th>
<th>400</th>
</tr>
</thead>
</table>
| Staph Cfr | CGQLYGNYQNSQ--
| SbCfr | CGQLRAEYEAVAGA--
| Cfr(b) | CGQLYGNYQKTNSQ--
| Cfr(c) | CGQLACAYRDDDQCGERTMSARDVKSFGEEVCEYGFYNQG |
| S. aureus RlmN | CGQLRAKERVETR--
| Paenibacillus Cfr | CGQLYGSNGNI-- |

<table>
<thead>
<tr>
<th>401</th>
<th>408</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staph Cfr</td>
<td>------</td>
</tr>
<tr>
<td>SbCfr</td>
<td>------</td>
</tr>
<tr>
<td>Cfr(b)</td>
<td>AMRSNSS</td>
</tr>
<tr>
<td>Cfr(c)</td>
<td>AMRSNSS</td>
</tr>
<tr>
<td>S. aureus RlmN</td>
<td>------</td>
</tr>
<tr>
<td>Paenibacillus Cfr</td>
<td>------</td>
</tr>
</tbody>
</table>