

**Investigation of the Iron(III)-
Siderophore Binding Properties of
Three Bacterial Periplasmic Binding
Proteins**

Ellis Wilde

PhD

University of York

Chemistry

August 2017

Abstract

Bacteria possess complex machinery for uptake of essential iron. Studying iron-uptake provides insight into requirements for bacterial growth, and for development of applications, including novel antibiotics and diagnostic tools.

The presented project investigates the binding of periplasmic binding proteins CeuE, FepB and VctP with a range of iron(III)-siderophore and siderophore-mimic compounds. CeuE was shown to bind iron(III)-n-LICAM (n= 6, 8) (N,N'-hexane-1,6-diylbis (2,3-dihydroxybenzamide and N,N'-octane-1,8-diylbis (2,3-dihydroxybenzamide)) tetradentate siderophore mimics, with Λ -configuration in both crystal and solution phase, with binding constants of 33 ± 8 and 58 ± 8 nM respectively. Comparing these results to those for iron(III)-n-LICAM (n= 4, 5) (N,N'-butane-1,4-diylbis (2,3-dihydroxybenzamide) and N,N'-pentane-1,5-diylbis (2,3-dihydroxybenzamide)) revealed that the highest affinity was found for a five-atom linker. Mutagenesis of His 227 and Tyr 288 that coordinate the iron(III)-centre, proved that Tyr 288 is essential for iron(III)-n-LICAM binding to CeuE. Binding affinity is slightly reduced for all iron(III)-n-LICAM (n= 4, 5, 6, 8) ligands with mutation to His 227. CeuE-H227L-iron(III)-5-LICAM crystal structure determination revealed that iron(III)-5-LICAM bound in the Λ -configuration with one aqua-ligand.

Salmochelin mimic siderophores Sal-n-LICAM (n= 4, 5) were synthesised and iron(III)-binding established *via* Job plot. For both compounds, equilibria of 1:1 and 3:2 ligand:metal ratios were observed, with 3:2 predominating over time. CeuE bound both iron(III)-complexes weakly in the Λ -configuration. Iron(III)-Sal-5-LICAM bound with higher affinity (511 ± 76 nM) than iron(III)-Sal-4-LICAM (15.6 ± 2.3 μ M).

FepB was overexpressed, purified, and its siderophore-binding profile compared with that of CeuE and VctP. It was shown that FepB bound iron(III)-enterobactin with nanomolar affinity, and had micromolar affinity for tetradentate catecholate complexes. CeuE bound tetradentate catecholate complexes with nanomolar affinity, and iron(III)-enterobactin with micromolar affinity. VctP bound tetradentate catecholate complexes with picomolar affinity, iron(III)-enterobactin with mid-nanomolar affinity, and iron(III)-Sal-n-LICAM (n= 4, 5) with low-nanomolar affinity. All three proteins bound iron(III)-MECAM (1,3,5-N,N',N"-tris-(2,3-dihydroxybenzoyl)-triaminomethylbenzene) with low nanomolar affinity.

Table of Contents

Abstract	2
Table of Contents	3
Table of Figures, Schemes and Tables	8
Accompanying Material	29
Acknowledgements	30
Declaration	31
1 Introduction	33
1.1 Overview	33
1.2 Iron Co-ordination	33
1.3 Iron Homeostasis in Bacteria and Mammalian Host	34
1.4 Siderophores	35
Discovery of Siderophores	36
Enterobactin and Catecholate Siderophores	38
Tetradentate Catecholate Siderophores	42
Properties of Siderophores	44
1.5 Bacterial Iron Uptake	45
Iron Uptake in Gram-Positive Bacteria	46
Iron Uptake in Gram-Negative Bacteria	46
Bacterial Iron Regulation via Fur and DxtR	49
Iron Uptake in Escherichia coli	50
Iron Uptake in Campylobacter jejuni	55
Iron Uptake in Vibrio cholerae	59
1.6 Iron Regulation in the Mammalian Host and Stealth Siderophores	64
1.7 Siderophore Mimics	67
1.8 Applications of Siderophores	72
Iron Chelation Therapy	72
Bacterial Detection and Molecular Imaging	73
Antibiotics by Iron Starvation	73
Trojan Horse Antibiotics	74
1.9 Project History	75
[CeuE-iron(III)-MECAM] ₂	76
CeuE-iron(III)-4-LICAM	77
A Linker Length of Five Atoms	79
1.10 Project Aims	81

2	Investigation of the Binding of Tetradentate Siderophore Mimics to CeuE.....	82
2.1	Introduction	82
2.2	CeuE.....	85
2.3	Crystal Structure of CeuE-iron(III)-6-LICAM	85
2.4	Crystal Structure of CeuE-iron(III)-8-LICAM	88
2.5	Determination of Dissociation Constants.....	92
2.6	Circular Dichroism.....	98
2.7	The Optimum Linker Length of n= 5.....	100
2.8	Summary and Conclusion	104
2.9	Future Work	105
3	The study of CeuE Variant Proteins for Investigation of Structure-Function Relationships in the CeuE Binding Cleft	107
3.1	Introduction	107
3.2	CeuE Variant Proteins.....	114
3.3	Crystal Structures of Apo CeuE Variant Proteins.....	115
	Structure Solution for CeuE-H227L	117
	Structure Solution for CeuE-H227A.....	118
	Structure solution for CeuE-Y288F	119
	Structure Solution for CeuE-H227L/Y288F and CeuE-H227A/Y288F	122
	Apo-Variant-CeuE Crystal Structures Summary	124
3.4	Dissociation Constants with Iron(III)-n-LICAM	125
3.5	Crystal Structure of CeuE-H227L-iron(III)-5-LICAM.....	136
3.6	Circular Dichroism.....	141
3.7	Summary and Conclusion	144
3.8	Future Work	146
4	Synthesis and Study of Salmochelin S1 Siderophore Mimics and their Interactions with CeuE.....	147
4.1	Introduction	147
	Computational Modelling of the Binding of CeuE with Iron(III)-Salmochelin S1	148
	Methods for Synthesis of Aryl- β -C-glycosides.....	150
4.2	Design and Synthesis of Salmochelin S1 Mimic Compounds.....	155
	Salmochelin S1.....	155
	Design of a Sulfonamide Linked Salmochelin Mimic	157
	Design of a Carbon-Carbon Bonded Salmochelin Mimic	158
4.3	Iron Binding	167
4.4	Circular Dichroism.....	173

4.5	Determination of Dissociation Constants.....	175
	Ceue-iron(III)-Sal-4-LICAM	175
	Ceue-iron(III)-Sal-5-LICAM	179
4.6	Summary and Conclusion	184
4.7	Future Work	185
5	Comparison of Three Enterobactin-Uptake Periplasmic Binding Proteins	186
5.1	Introduction	186
5.2	Hexadentate Iron(III)-Siderophore Binding of Ceue	195
5.3	FepB Expression and Purification.....	199
5.4	Iron(III)-Siderophore Binding of FepB.....	206
5.5	Iron(III)-Siderophore Binding by VctP.....	212
5.6	Comparison of FepB, Ceue and VctP iron(III)-siderophore binding	221
	FepB: Hexadentate vs Tetradentate	221
	Ceue: Tetradentate vs Hexadentate	222
	VctP: Tetradentate and Hexadentate	222
	Features Favouring Tetradentate Siderophore Binding	222
	Scavenger Versus Exporter Species	223
	MECAM as a Mimic for Enterobactin.....	226
	Salmochelin S1 mimics.....	228
5.7	Summary and Conclusion	230
5.8	Future Work	231
6	Experimental	234
6.1	General Remarks	234
	Materials.....	234
	Instrumentation	234
6.2	Synthesis	235
	2,3-Bis(benzyloxy)benzoic acid (Compound B).....	235
	2,3-Bis(benzyloxy)benzoic acid N-hydroxysuccinimide ester (Compound R) ..	236
	1-Amino, 4-(2,3-dibenzyloxybenzamide) butane (Compound C)	237
	1-Amino, 5-(2,3-dibenzyloxybenzamide) pentane (Compound S).....	238
	Methyl-5-iodo-3-methoxysalicylate (Compound E).....	239
	Methyl-5-iodo-3-hydroxysalicylate (Compound F).....	240
	Methyl-5-iodo-3,4-benzyloxysalicylate (Compound G).....	241
	5-Iodo-3,4-benzyloxysalicylate (Compound H)	242
	5-Iodo-bis(2,3benzyloxy)-4-LICAM (Compound J)	243
	Methoxy-5-aceto- β -D-glucosyl-3, 4-benzyloxysalicylate (Compound M)	244

	Methoxy-5-benzyloxy- β -D-glucosyl-3,4-benzyloxysalicylate (Compound N)...	246
	Benzyloxy-5-benzyloxy- β -D-glucosyl-3,4-benzyloxysalicylate (Compound N)	247
	5-Benzyloxy- β -D-glucosyl-3, 4-benzyloxysalicylate (Compound P)	248
	5-Benzyloxy- β -D-glucosyl-bis(3, 4-benzyloxy)-4-LICAM (Compound Q)	250
	5- β -D-glucosyl-4-LICAM.....	251
	5-Benzyloxy- β -D-glucosyl-bis(3, 4-benzyloxy)-5-LICAM (Compound T)	252
	5- β -D-glucosyl-5-LICAM.....	254
6.3	Sequence Alignments	255
6.4	Protein Production and Purification	255
	FepB Protein Sequence	255
	Primer Design and Cloning	255
	Test Expression	256
	Overexpression.....	257
	Purification	257
6.5	Preparation of CeuE variant proteins	260
6.6	Protein Crystallography	261
	General Protein Crystallisation Procedure	261
	Preparation of Iron(III)-Ligand Complexes	261
	Co-crystallisation of Iron(III)-Ligand Complexes with CeuE and Mutant CeuE H227L	261
	Crystallisation of CeuE Native Mutants	262
6.7	Structure Solution.....	262
	Spacegroups and Structure Solution for CeuE Co-crystal Structures	263
	Spacegroups and Structure Solution for CeuE Variant Proteins.....	263
	Crystallographic Statistics.....	263
6.8	Fluorescence Quenching Titration	270
	Experimental Parameters	270
	Stock Solutions.....	270
	Protein Concentration Determination for Fluorescence Titrations	270
	Standard Binding Method	270
	Weak Binding Method	271
	VctP Tight Binding Method.....	271
	CeuE Fluorescence Control Experiments	271
	VctP Fluorescence Control Experiments	272
	Ligand Absorbance Control Experiments	272
	Averages and Error Calculation	278

6.9	Circular Dichroism.....	278
	Protein Folding Experiments.....	278
	Ligand Binding Experiments	278
	CeuE Iron(III)-5-β-D-glucosyl-4-LICAM and Iron(III)-5-β-D-glucosyl-5-LICAM Circular Dichroism Titrations	279
6.10	Job Plot.....	279
6.11	Analytical HPLC	280
6.12	ESI-MS for Natural Salmochelins	280
	Appendix 1: Fluorescence Quenching Titration Data and Controls	281
	Abbreviations	288
	Iron(III)-Siderophore Complex Nomenclature.....	288
	References	293

Table of Figures, Schemes and Tables

Figure 1: A tris-bidentate octahedral metal complex in a Λ - and Δ -configuration. ¹¹	34
Figure 2: Oxygen containing functional groups common to siderophores. Oxygen atoms that must be deprotonated for iron(III) binding are shown in red. ²⁸	36
Figure 3: Structures of the earliest siderophores discovered, and an example of a fluorescent pyoverdine. ^{28, 39, 40, 42, 48}	37
Figure 4: Protonation resulting in a switch between catecholates and salicylates binding mode. Top: Representation of the iron(III)-bound catecholates and salicylates modes of enterobactin. Bottom: Schematic of the catecholates and salicylates binding modes for one bidentate catecholamide unit. ⁶⁸	39
Figure 5: Triscatecholate siderophores with varied amino acid backbones. ^{28, 73}	40
Figure 6: Amphienterobactin, a range of structurally similar linear hexadentate catecholates siderophores and mixed type siderophore fuscachelin A. ^{80-82, 84, 85, 88}	41
Figure 7: Tetradentate catecholates siderophores from a range of bacterial species with variable linker regions highlighted in blue. ^{91, 93, 96-98, 107}	43
Figure 8: Tridentate and bidentate catecholates siderophores. ²⁸	44
Figure 9: Linear and cyclic hydroxamate siderophores. ^{39, 110}	45
Figure 10: General mechanism for iron(III)-uptake in Gram-positive bacteria. ²²	46
Figure 11: General iron(III)-uptake mechanisms for Gram-negative bacteria. ²⁰	47
Figure 12: Stereo view of siderophore outer membrane receptor FpvA bound to iron(III)-pyoverdine in <i>Pseudomonas aeruginosa</i> (PDB ID: 2IAH). β -barrel is in dark blue with the cork domain in light blue. A: Side view with iron(III)-pyoverdine shown in cylinders coloured by atom type. B: Transmembrane view with iron(III)-pyoverdine shown in spheres coloured by atom type. Figure produced using CCP4mg.	48
Figure 13: Stereo view of periplasmic binding protein FhuD (yellow ribbons) from <i>E. coli</i> bound to siderophore iron(III)-coprogen (cylinders coloured by atom type) (PDB ID: 1ESZ). ¹²² Figure produced using CCP4mg.....	49
Figure 14: Stereo view of the crystal structure of the Fur regulator protein from <i>C. jejuni</i> . (Zn^{2+} ions shown as grey spheres) (PDB ID: 4ETS). ¹²⁹ Figure produced using CCP4mg.	50
Figure 15: Siderophores synthesised and secreted by pathogenic strains of <i>E. coli</i> . ¹³⁸	51
Figure 16: Proposed functions of proteins involved in catecholates siderophore synthesis, secretion and uptake in pathogenic <i>E. coli</i> . Ent is an abbreviation for Enterobactin, S4, S2, S1 and SX are abbreviations for each salmochelin.	52
Figure 17: Stereo view of the crystal structure of FepA, the enterobactin outer-membrane receptor in <i>E. coli</i> . ¹⁴² A: Side view B: Transmembrane view. (PDB ID: 1FEP) Figure produced using CCP4mg.	53

Figure 18: Stereo view of FepB, the enterobactin periplasmic binding protein in complex with iron (III)-enterobactin (PDB ID: 3TLK). ¹⁴⁶ Figure produced using CCP4mg.	54
Figure 19: Proposed functions of proteins involved in catecholate siderophore uptake in <i>C. jejuni</i> . ^{115, 173, 180}	57
Figure 20: Stereo view of CeuE (yellow ribbons) (PDB ID: 5ADW) has bilobate shape with a shallow binding cleft where iron(III)-bisDHBS is bound (cylinders coloured by atom type). ^{186 90} Figure produced using CCP4mg.	58
Figure 21: Known functions of proteins involved in catecholate siderophore synthesis and uptake in <i>V. cholerae</i>	61
Figure 22: Stereo view of catecholate siderophore periplasmic binding proteins VctP (A) (PDB ID: 3TEF) and ViuP bound to iron(III)-vibriobactin (B) (PDB ID:3R5T). Protein shown in yellow ribbons with vibriobactin in cylinders coloured by atom type. ^{208, 209} Figure produced using CCP4mg.	63
Figure 23: Stereo view of crystal structure of human siderocalin in complex with iron(III)-enterobactin (cylinders coloured by atom type) (PDB ID: 3CMP) Siderocalin is shown in surface view, with blue representing positively charged regions, red representing negatively charged regions and white representing neutral or hydrophobic regions. ^{219, 220} Figure produced using CCP4mg.	65
Figure 24: The structures of a selection of catecholate siderophore mimics. ^{248, 249, 253, 255, 257, 261, 262}	70
Figure 25: Tetradentate and bidentate catecholate siderophore mimics. ^{94, 266, 267}	71
Figure 26: Siderophore mimics involved in iron(III)-chelation therapy. ²⁷⁹⁻²⁸³	73
Figure 27: Examples of siderophore-antibiotic Trojan horse conjugates. ^{296, 301}	75
Figure 28: A: Crystal structure of CeuE-iron(III)-MECAM in the 2:2:2 dimer (PDB ID: 2CHU). ¹⁸⁷ CeuE is shown in sea green ribbons, and iron(III)-MECAM in ball and stick coloured by atom type. B: Stereo view of the 2:2 Fe: MECAM arrangement across two CeuE binding pockets shown in ball and stick coloured by atom type. The electron density for the maximum likelihood map was contoured at the 2 σ level. Figure produced using CCP4mg.	76
Figure 29: Crystal structure of CeuE-iron(III)-4-LICAM. ¹⁸⁶ A: CeuE depicted in lawn green ribbons, iron(III)-4-LICAM shown as ball and stick coloured by atom type with coordinating His 227 and Tyr 288 shown as cylinders coloured by atom type. B: Stereo view of iron(III)-4-LICAM shown as ball and stick coloured by atom type with coordinating His 227 and Tyr 288 shown as cylinders coloured by atom type. iron(III) bond distances are labelled in Å. The electron density for the maximum likelihood map was contoured at the 2 σ level. Figure produced using CCP4mg.	78

Figure 30: Top: CeuE depicted in gold green ribbons, iron(III)-bisDHBS shown as ball and stick coloured by atom type with coordinating His 227 and Tyr 288 shown as cylinders coloured by atom type. Bottom: Stereo view of iron(III)-bisDHBS shown as ball and stick. The electron density for the maximum likelihood map contoured at the 1.5σ level.⁹⁰ Figure produced using CCP4mg.80

Figure 31: Stereo view of the crystal structure overlay of CeuE-iron(III)-5-LICAM (cylinders coloured by atom type with carbon in light crimson) with CeuE-iron(III)-bisDHBS (cylinders coloured by atom type with carbon in light gold) superimposed by SSM algorithm based on C α positions. The iron(III)-binding ligands display a very similar binding mode in the CeuE binding pocket. Figure produced using CCP4mg.80

Figure 32: n-LICAM siderophore mimic compounds with extended linker lengths from 4 to 8 carbon atoms.84

Figure 33: Crystal structure of CeuE-iron(III)-6-LICAM (PDB ID: 5A5V). CeuE is depicted as dark purple ribbons. Iron(III)-6-LICAM is depicted as ball and stick coloured by atom type. Iron(III)-coordinating His 227 and Tyr 288 are depicted as cylinders coloured by atom type. Figure produced using CCP4mg.85

Figure 34: Crystal structure of CeuE-iron(III)-6-LICAM in stereo view. A: iron(III)-6-LICAM as ball and stick. The electron density for the maximum likelihood difference map was calculated before introduction of the iron(III)-ligand into the model to avoid phase bias and contoured at the 2σ level (green chicken wire). B: iron(III)-6-LICAM ball and stick, coordinating His 227 and Tyr 288 in cylinders. Iron(III) bond distances are labelled in Å. The electron-density for the maximum likelihood map was contoured at the 2σ level in blue chicken wire. C: iron(III)-6-LICAM in the CeuE binding pocket, with all significant contacts Coordinating residues (cylinders) are labelled. Figure produced using CCP4mg.87

Figure 35: Crystal structure of CeuE-iron(III)-8-LICAM (PDB ID: 5AD1). CeuE is depicted as orange ribbons. Iron(III)-6-LICAM is depicted as ball and stick coloured by atom type. Iron(III)-coordinating His 227 and Tyr 288 are depicted as cylinders coloured by atom type. Figure produced using CCP4mg.88

Figure 36: Crystal structure of CeuE-iron(III)-8-LICAM in stereo view. A: Iron(III)-8-LICAM shown as ball and stick. The electron density for the maximum likelihood difference map was calculated before introduction of the iron(III)-ligand into the model to avoid phase bias and contoured at the 2σ level in green chicken wire. B: Iron(III)-8-LICAM as ball and stick, with coordinating His 227 and Tyr 288 as cylinders. Bond distances are labelled in Å. The electron-density for the maximum likelihood map was

contoured at 2σ level. C: iron(III)-8-LICAM in the CeuE binding pocket, with all significant contacts labelled. Figure produced using CCP4mg.	90
Figure 37: Crystal structure of CeuE-iron(III)-6-LICAM showing that the only tryptophan residue, Trp 287 (circled in orange), is located adjacent to the iron(III)-siderophore binding pocket in CeuE. Figure produced using CCP4mg.	92
Figure 38: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-6-LICAM and their associated non-linear regression fitting data from Dynafit. ³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-6-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	94
Figure 39: Fluorescence quenching titration curves as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-8-LICAM and their associated non-linear regression fitting data from Dynafit. ³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-8-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	95
Figure 40: Overlay of selected binding curves for fluorescence quenching titrations of CeuE titrated with of iron(III)-n-LICAM ($n = 4, 5, 6, 8$) and their associated non-linear regression fitting data calculated using Dynafit ($n = 4, 5$ are taken from the literature). ^{186, 326}	97
Figure 41: Λ and Δ configurations depicted for a general octahedral metal complex, and the Λ -configuration as is present in a schematic of CeuE-iron(III)-n-LICAM and the crystal structure of CeuE-iron(III)-6-LICAM. ^{11, 14} Figure produced using CCP4mg.....	98
Figure 42: Circular dichroism spectra for CeuE plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra are run for 50 μM CeuE in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μM iron(III)-n-LICAM ($n = 5, 6, 8$) ligands as well as the ligands at 50 μM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without CeuE. The spectra show the induction of Λ -configuration upon introduction of CeuE to each iron(III)-n-LICAM ($n = 5, 6, 8$).	99
Figure 43: Stereo view overlay of CeuE-iron(III)-n-LICAM crystal structures showing atom positions of the iron(III)-n-LICAM ligands in the CeuE binding pocket and the iron(III) binding residues His 227 and Tyr 288 for each structure. CeuE-iron(III)-4-LICAM (lawn green), CeuE-iron(III)-5-LICAM (light crimson), CeuE-iron(III)-6-LICAM (dark purple) and CeuE-iron(III)-8-LICAM (orange). CeuE-iron(III)-4-LICAM has catecholate ring (left) displaced out of the optimum plane due to linker strain. Figure produced using CCP4mg.	101

Figure 44: Planes constructed for the six carbon atoms of each catecholate aromatic ring displayed in red (left) and yellow (right). Atom coordinates from CeuE-iron(III)-n-LICAM (n = 4, 5, 6, 8) crystal structures for each ligand are displayed in cylinders coloured by atom type. Interplanar angles are calculated between catecholate aromatic rings for each structure. Images created in Mercury. ³³¹	102
Figure 45: Schematic diagram of CeuE as a potential scaffold for an artificial metalloenzyme. CeuE is shown in green ribbons. The iron(III) siderophore complex is shown in the binding cleft in purple, with a linked inorganic catalyst labelled ‘cat’ in orange. Figure produced using CCP4mg.	106
Figure 46: Structure of CeuE-iron(III)-8-LICAM in stereo view (PDB: 5AD1) showing residues involved in iron(III)-n-LICAM binding Arg118, Arg205, Arg 249, Lys 121, His 227 and Tyr 288. Residues involved in binding are shown in cylinders coloured by atom type. Iron(III) is shown as a grey sphere. 8-LICAM is shown in circles coloured by atom type. Figure produced using CCP4mg.	107
Figure 47: Overlay of crystal structures of CeuE (ice blue) (PDB ID: 3ZKW), YclQ (sea green) (PDB ID: 3GFV), PiuA (light crimson) (PBD ID: 4JCC), VctP (lilac) (PDB ID: 3TEF) and an unnamed PBP (dark grey) (PDB ID: 4MX8) by the SSM algorithm. ³¹⁷ Protein backbones are shown in ribbons. Figure produced using CCP4mg.	108
Figure 48: Sequence alignment of amino acid sequences for proteins superposed in Figure 47. Red blocks are totally conserved, red text indicates close matches or residues of similar properties. Black text shows non-matching residues. Conserved arginine residues are shown with a blue triangle, conserved histidine with a pink star, conserved tyrosine with a green star.	109
Figure 49: Overlay of crystal structures of CeuE (ice blue) (PDB ID: 3ZKW), YclQ (sea green) (PDB ID: 3GFV), PiuA (light crimson) (PBD ID: 4JCC) and VctP (lilac) (PDB ID: 3TEF) by the SSM algorithm in stereo view. ³¹⁷ Protein backbones are shown in ribbons. Conserved Tyr and His residues that are expected to be involved in iron(III)-coordination for these proteins are shown in cylinders. Figure produced using CCP4mg.	110
Figure 50: Overlay of crystal structures of CeuE (ice blue) (PDB ID: 3ZKW) with unnamed PBP (grey) (PDB ID: 4MX8) in stereo view. Selected regions of protein backbone are shown as ribbons with the conserved His and Tyr shown in cylinders. The His of the unnamed PBP is folded away from the binding cleft. Figure produced using CCP4mg.	111
Figure 51: Sequence alignment for proteins similar to CeuE from a range of bacterial species. Residues 235-345 are shown, with conserved His and Tyr shown as residue 260 and 320 as a pink star and green star, respectively. The structures of these proteins are	

currently unsolved. Red blocks are totally conserved across all proteins, red text indicates close matches or residues of similar properties. Black text shows non-matching residues.

.....112

Figure 52: Overlay of the A chain of all mutant structures, showing high similarity in overall fold. H227L is shown in lemon ribbons, H227A in light brown, Y288F in light green with Zinc(II) ions shown as grey spheres, H227L/Y288F in light blue and H227A/Y288F in lilac. Figure produced using CCP4mg.116

Figure 53: Stereo view of overlay of the A chain of all mutant structures with that of wild-type CeuE bound to iron(III)-5-LICAM, showing residues 227 and 288 in cylinders. There is little change in the arrangement of the 227-loop region for any of the mutants, but when iron(III)-5-LICAM is bound, the loop region is shifted towards the iron(III) centre. H227L is shown in lemon ribbons, H227A in light brown, Y288F in light green with zinc(II) ions shown as grey spheres, H227L/Y288F in light blue and H227A/Y288F in lilac. CeuE-iron(III)-5-LICAM is shown in light crimson, with the ligand shown in translucent cylinders. Figure produced using CCP4mg.116

Figure 54: Stereo view of iron(III)-siderophore binding residues Tyr 228, Arg 118, Arg 205 and Arg 249 in the binding cleft of chain A of CeuE-H227L. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.117

Figure 55: Stereo view of loop region in chain A of CeuE-H227L from residues 222 to 230. Residues 222-226, 288-230 shown in lemon cylinders. Leu 227 is shown in cylinders coloured by atom type. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. There is poor electron density for the Leu 227 sidechain. Figure produced using CCP4mg.118

Figure 56: Stereo view of the CeuE-H227A binding cleft (chain B). Protein backbone shown in pale brown ribbons with key sidechains of Arg 118, Arg 205, Arg 249 and Tyr 288 shown as pale brown cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.119

Figure 57: Stereo view of loop region of CeuE-H227A residues 222-226 and 228-230 shown in pale brown cylinders. Ala 227 shown in cylinders coloured by atom type. The electron density for the maximum likelihood weighted map weighted was contoured at the 1.5σ level. Figure produced using CCP4mg.119

Figure 58: Stereo view of zinc(II) chelated by His 210 and Glu 220 in CeuE-Y288F. Protein backbone is depicted in light green ribbons. His 210 and Glu 220 sidechains are displayed as cylinders coloured by atom type. Zinc(II) is shown as a grey sphere with protein contacts as black dashed lines with bond distances annotated in Å. The electron

density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.	120
Figure 59: Stereo view of zinc(II) chelated by His 227 and Glu 183 from a neighbouring protein monomer in CeuE-Y288F. Protein backbones are depicted in light green ribbons. His 227 and Glu 183 sidechains are displayed as cylinders coloured by atom type. Zinc(II) is shown as a grey sphere with protein contacts as black dashed lines with bond distances annotated in Å. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.	121
Figure 60: There is clear electron density for the Y288F mutation in the single mutant variant protein. Protein backbone is shown in light green ribbons in stereo view. Phe 288 is shown in light green cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.	121
Figure 61: Modelling of residue 288 in COOT. ³⁶⁵ The model is shown in dark blue sticks. Weighted electron density map shown in light blue chicken wire, set at 1.5σ . The weighted difference map is shown in red and green chicken wire for negative and positive electron density respectively, set at 1.5σ . A: Tyrosine modelled, with observed negative electron difference density around the tyrosinated oxygen atom. B: Phenylalanine modelled.	122
Figure 62: Stereo view of crystal structure of CeuE-H227L/Y288F showing residues Leu 227 and Phe 288 in the binding cleft region of chain B. Protein backbone is shown as ice blue ribbons with residues Leu 227 and Phe 288 shown as ice blue cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.	123
Figure 63: Stereo view of crystal structure of CeuE-H227A/Y288F showing residues Ala 227 and Phe 288 in the binding cleft region of chain B. Protein backbone is shown as lilac ribbons with residues Ala 227 and Phe 288 shown as lilac cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.	123
Figure 64: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227L/Y288F titrated with iron(III)-5-LICAM. CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-6-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl The titration was performed in triplicate, and a binding curve calculated using Dynafit. ³²⁶ The data were plotted alongside the normalised fluorescence quenching observed when CeuE-H227L/Y288F is diluted with the same amount of 40 mM Tris-HCl pH 7.5, 150 mM NaCl buffer.	126

Figure 65: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227L/Y288F titrated with iron(III)-5-LICAM. CeuE-H227L/Y288F at a concentration of $12\ \mu\text{M}$ in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl was titrated with aliquots of $60\ \mu\text{M}$ iron(III)-6-LICAM in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl. A binding curve was calculated using Dynafit.³²⁶ The quenching was not sufficient for determination of a dissociation constant.128

Figure 66: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-Y288F titrated with iron(III)-5-LICAM. CeuE-Y288F at a concentration of $12\ \mu\text{M}$ in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl was titrated with aliquots of $60\ \mu\text{M}$ iron(III)-6-LICAM in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl. A binding curve was calculated using Dynafit.³²⁶ The quenching was not sufficient for determination of a dissociation constant.129

Figure 67: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227L-iron(III)-5-LICAM and their associated non-linear regression fitting data from Dynafit.³²⁶ CeuE-H227L at a concentration of $240\ \text{nM}$ in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl was titrated with aliquots of $12\ \mu\text{M}$ iron(III)-5-LICAM in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl.130

Figure 68: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227A-iron(III)-5-LICAM and their associated non-linear regression fitting data from Dynafit.³²⁶ CeuE-H227A at a concentration of $240\ \text{nM}$ in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl was titrated with aliquots of $12\ \mu\text{M}$ iron(III)-5-LICAM in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl.131

Figure 69: Selected binding curves for the fluorescence quenching titration of iron(III)-n-LICAM ($n = 4, 5, 6, 8$) with CeuE-H227L, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227L-iron(III)-n-LICAM ($n = 4, 5, 6, 8$) and their associated non-linear regression fitting data from Dynafit. CeuE-H227L at a concentration of $240\ \text{nM}$ in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl was titrated with aliquots of $12\ \mu\text{M}$ iron(III)-n-LICAM in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl.133

Figure 70: Selected binding curves for the fluorescence quenching titration of iron(III)-n-LICAM ($n = 4, 5, 6, 8$) with CeuE-H227A, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227A-iron(III)-n-LICAM ($n = 4, 5, 6, 8$) and their associated non-linear regression fitting data from Dynafit. CeuE-H227A at a concentration of $240\ \text{nM}$ in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl was titrated with aliquots of $12\ \mu\text{M}$ iron(III)-n-LICAM in $40\ \text{mM}$ Tris-HCl pH 7.5, $150\ \text{mM}$ NaCl.134

Figure 71: CeuE-H227L co-crystallised with iron(III)-5-LICAM. Three protein monomers are present in the asymmetric unit, with an iron(III)-5-LICAM in each binding cleft. Chain

A is shown in sea green ribbons, chain B in lawn green ribbons, and chain C in navy ribbons. iron(III)-5-LICAM is shown in spheres coloured by atom type. Figure produced using CCP4mg.	136
Figure 72: Crystal structure of CeuE-H227L-iron(III)-5-LICAM showing just chain A. iron(III)-5-LICAM is shown as ball and stick coloured by atom type. The aqua (or hydroxo) ligand is shown as a red sphere. Iron(III) contacts are shown as black lines. Figure produced using CCP4mg.	137
Figure 73: Co-crystal structure of CeuE-H227L-iron(III)-5-LICAM, showing the binding cleft of Chain A (A) and Chain B (B) in stereo view. An aqua (or hydroxo) ligand the octahedral coordination of iron(III). Tyr 288 is shown in cylinders coloured by atom type. iron(III)-5-LICAM is shown in ball and stick coloured by atom type. The aqua (or hydroxo) ligand is shown as a red sphere. Iron(III) bond distances in Å are labelled. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.	138
Figure 74: Stereo view of the crystal structure of CeuE-H227L-iron(III)-5-LICAM. CeuE-H227L-iron(III)-5-LICAM is shown in sea green ribbons, with key residues coloured by atom type. The iron(III)-coordinating aqua/hydroxo ligand is shown as a red sphere. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Electron density for the flexible loop region is poor for the L227 sidechain. Figure produced using CCP4mg.	139
Figure 75: Stereo view of the overlay of crystal structure of CeuE-H227L-iron(III)-5-LICAM with that of CeuE-iron(III)-5-LICAM and CeuE-H227L via secondary structure matching. ³¹⁷ CeuE-H227L-iron(III)-5-LICAM is shown in sea green cylinders, with key residues coloured by atom type. CeuE-iron(III)-5-LICAM is shown in light crimson cylinders, with key residues coloured by atom type. The water/hydroxide molecule in the CeuE-H227L-iron(III)-5-LICAM structure is shown as a pale blue sphere. CeuE-H227L without a ligand is shown as lemon cylinders. The flexible loop upon which residue 227 is located is folded away from the iron(III) centre when the H227L mutation is present. Figure produced using CCP4mg.	139
Figure 76: Circular Dichroism spectra for all five CeuE variant proteins upon the addition of iron(III)-5-LICAM. Ellipticity in mdeg was plotted vs wavelength from 300 to 700 nm. The spectra were run for 50 μM CeuE-variant in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with 50 μM iron(III)-5-LICAM.	141
Figure 77: Circular Dichroism spectra for CeuE-H227L with each iron(III)-n-LICAM (n= 4, 5, 6, 8) complex. Ellipticity in mdeg was plotted vs wavelength from 300 to 700 nm. The spectra were run for 50 μM CeuE-H227L in 0.11M Tris-HCl pH 7.5, 150 mM NaCl	

buffer with 50 μ M iron(III)-n-LICAM (n = 5, 6, 8) ligands. The spectra show the induction of Λ -configuration.....	143
Figure 78: Stereo view of the overlay of crystal structures of apo wild-type CeuE (PDB ID: 3ZKW) with CeuE-iron(III)-5-LICAM (PDB ID: 5A5D) showing the conformational shift of the His 227-loop region for iron(III) binding. Figure produced using CCP4mg.....	145
Figure 79: CeuE (PDB ID: 3ZKW) shown with surface view, with three identified binding cleft regions. Cavity 1 is the known iron(III)-siderophore binding cleft. Cavity 2 is a secondary binding cleft that may be capable of accommodating a glucose unit. Cavity 3 is a smaller pocket that is unlikely to accommodate a glucose unit. Figure provided by P. S. Bond.....	148
Figure 80: Molecular modelling of chair conformations of the C5 appended glucose in the CeuE secondary binding pocket. Each conformation arose from different simulation parameters. All four conformations provided possible glucose orientations in the binding cleft without unfavourable clashes. Figure provided by P. S. Bond.....	149
Figure 81: Potential hydrogen bonds between glucose and water molecules when modelled in the secondary binding cleft of CeuE. Figure provided by P. S. Bond.....	150
Figure 82: Necessary C-C bond formation between the C5 position of a catecholamide and a glucose unit for the generation of an aryl- β -C-glycoside salmochelin mimic compound. The new C-C bond is shown in red.....	151
Figure 83: HPLC trace for Salmochelin S1. The sample in water containing 0.1% formic acid was injected with a 6-40% gradient of acetonitrile in water with 0.1% formic acid over 25 minutes.....	155
Figure 84: UV spectra of 0.5 mM iron(III)-5-LICAM at pH 7.0 40 mM Tris-HCl 150 mM NaCl overlaid with UV spectra of 0.5 mM iron(III)-S1 sample over a pH range of 7.0-9.1 in 40 mM Tris-HCl 150 mM NaCl.....	156
Figure 85: Simple modelling of the sulfonamide salmochelin mimic in the secondary binding pocket of CeuE. Left: in axial conformation. Right: in equatorial conformation.....	157
Figure 86: Selected UV-Visible spectra of Sal-4-LICAM: iron(III) in the ratios as shown, with a final concentration of 400 nM. Spectra were recorded after 1 hour of equilibration. Vertical lines are shown at the λ_{\max} wavelengths.....	168
Figure 87: Photograph of sample vials of Sal-4-LICAM: iron(III) with highest ligand concentration to lowest ligand concentration from left to right, and highest iron(III) concentration to lowest iron(III) from right to left.....	168
Figure 88: Job plot for Sal-4-LICAM: iron(III), with the x-axis showing the ratio of ligand to iron(III) in %. Absorbance data are plotted at 552 nm (purple triangles) and 492 nm (red circles) across the ratio range, showing a peak at a 50:50 Sal-4-LICAM: iron(III) ratio for	

absorbance at 552 nm, and a peak at 60:40 Sal-4-LICAM: iron(III) ratio for absorbance at 492 nm.....	169
Figure 89: Job plot for Sal-4-LICAM: iron(III) after 7 days, with the x-axis showing the ratio of ligand relative to the percentage of iron(III). Absorbance data are plotted at 552 nm (purple triangles) and 492 nm (red circles) across the ratio range, showing peaks at 60:40 Sal-4-LICAM: iron(III) ratio for absorbance at 552 nm and 492 nm.....	170
Figure 90: Selected UV-Visible spectra of Sal-5-LICAM: iron(III) in the ratios as shown, with a final concentration of 400 nM. Spectra were recorded after 1 hour of equilibration. Vertical lines are shown at the λ_{\max} wavelengths.....	171
Figure 91: Photograph of sample vials of Sal-5-LICAM: iron(III) with highest ligand concentration to lowest ligand concentration from left to right, and highest iron(III) concentration to lowest iron(III) from right to left.....	171
Figure 92: Job plot for Sal-5-LICAM: iron(III), with the x-axis showing the ratio of ligand as a percentage of total Sal-5-LICAM: iron(III). Absorbance data are plotted at 555 nm and 495 nm, showing a peak at a 50:50 Sal-4-LICAM: iron(III) ratio for absorbance at 555 nm, and a peak at 60:40 Sal-4-LICAM: iron(III) ratio at 495 nm.	172
Figure 93: Job plot for Sal-5-LICAM: iron(III) after 7 days, with the x-axis showing the ratio of ligand as a percentage of total Sal-5-LICAM: iron(III). Absorbance data are plotted at 555 nm and 495 nm, showing a peak at a 50:50 Sal-4-LICAM: iron(III) ratio for absorbance at 555 nm, and a peak at 60:40 Sal-4-LICAM: iron(III) ratio at 495 nm.	172
Figure 94: Circular dichroism spectra for 50 mM iron(III)-Sal-4-LICAM and iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer.....	174
Figure 95: Circular dichroism spectra of 50 mM CeuE, CeuE-iron(III)-Sal-4-LICAM and CeuE, iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer. The spectra showed a small Λ -configurational preference upon introduction of CeuE to iron(III)-Sal-4-LICAM, which increases upon introduction of CeuE to iron(III)-Sal-5-LICAM.....	175
Figure 96: Fluorescence quenching titration curves for binding of iron(III)-Sal-4-LICAM to 0.24 μ M CeuE 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	176
Figure 97: Fluorescence quenching titration curve for binding of iron(III)-Sal-4-LICAM to 12 μ M CeuE in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	176
Figure 98: Circular dichroism spectra for the titration of 50 μ M iron(III)-Sal-4-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl with aliquots of CeuE in 0.11M Tris-HCl pH 7.5, 150 mM NaCl.....	177
Figure 99: Binding curves for three independent circular dichroism titrations of 50 μ M iron(III)-Sal-4-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl titrated with CeuE.	

Ellipticity at 330 nm was plotted against CeuE concentration. Non-linear regression fitting data was generated with Dynafit.	178
Figure 100: Binding curve for 0.24 μ M CeuE in 40 mM Tris-HCl pH 7.5, 150 mM NaCl. titrated with 12 μ M iron(III)-Sal-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	179
Figure 101: Circular dichroism spectra for the titration of 50 μ M iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl with aliquots of CeuE.	180
Figure 102: Binding curves for three independent circular dichroism titrations of 50 μ M iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl titrated with CeuE. Ellipticity at 330 nm was plotted against CeuE concentration. non-linear regression fitting data was generated with Dynafit.	181
Figure 103: Binding curves for three independent fluorescence quenching titrations of 2 μ M CeuE in 40 mM Tris-HCl pH 7.5, 150 mM NaCl with 100 μ M iron(III)-Sal-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl. Non-linear regression fitting data was generated with Dynafit.	182
Figure 104: Left: Enterobactin Right: iron(III)-enterobactin displayed as a 3-dimensional representation of the octahedral coordination.	186
Figure 105: A: Co-crystal structure of FepB (coral ribbons) with iron(III)-enterobactin in the binding cleft (cylinders coloured by atom type). B: stereo view of binding cleft of FepB (coral cylinders coloured by atom type, carbon in coral), showing all amino acids involved in binding of the enterobactin scaffold (cylinders coloured by atom type) (structure from PDB ID: 3TLK). ¹⁴⁶ Figure produced using CCP4mg.	188
Figure 106: Stereo view of the binding Cleft of FepB (cylinders coloured by atom type, with carbons in coral) with iron(III)-enterobactin bound (cylinders coloured by atom type). Top: Thr 73, Asn 77, Arg 78 and Thr 200 provide hydrogen bonds with the backbone and amide carbonyl oxygens. Middle: Gly 126 (backbone amide nitrogen) Arg 242 and Arg 301 provide hydrogen bonds to catechol oxygens. Bottom: Trp 209 and Phe 300 provide hydrophobic interactions (PDB ID: 3TLK). ¹⁴⁶ Figure produced using CCP4mg.	189
Figure 107: Sequence alignment of FepB, CeuE and VctP using MUSCLE. ⁴²⁶ Image produced with ESPrpt. ⁴²⁷ Red blocks are totally conserved, red text indicates close matches or residues of similar properties.	191
Figure 108: A: Overlay of Chain A of the structures of CeuE in lemon ribbons (PDB ID: 3ZKW), FepB in coral ribbons (PDB ID: 3TLK) and VctP in lilac ribbons (PDB ID: 3TEF) via the SSM algorithm. B: Stereo view of the binding cleft region of CeuE and VctP: the conserved His and Tyr in CeuE and VctP are absent in FepB. Figure produced using CCP4mg.	192

Figure 109: Hexadentate and tetradentate siderophores and mimic siderophores used to probe periplasmic binding protein accommodating ability. Protons shown in red are deprotonated upon iron(III) binding.	193
Figure 110: HPLC trace of purchased enterobactin. The sample was analysed with a gradient of 10-80% acetonitrile in water over 35 minutes with 0.1% TFA throughout and a flow rate of 1 mL/min. The major peak at a retention time of 21 minutes correlates well with literature retention times for enterobactin. ^{86, 87}	195
Figure 111: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-enterobactin and their associated non-linear regression from Dynafit. ³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-enterobactin in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	196
Figure 112: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-enterobactin and the model from Dynafit. ³²⁶ CeuE at a concentration of 2.4 μM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 120 μM iron(III)-enterobactin in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	197
Figure 113: Fluorescence quenching titration plots of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-MECAM and the model from Dynafit. ³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-MECAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl. .	198
Figure 114: Protein sequence for FepB from <i>E. coli</i> K strain MC1061 from Uniprot (code C6EKA3, residues 1-26 were predicted to be disordered (shown in a black box) and were removed to generate the construct 27-318. Image produced with ESPript. ⁴²⁷	199
Figure 115: SDS PAGE of test overexpression of FepB in BL21 cells. The bands of overexpressed protein are shown in a light blue box. FepB was present in both total and soluble fractions for cells cultured at both 37 and 16 °C.	200
Figure 116: SDS PAGE of overexpression of FepB in 500 mL cultures of BL21 cells. The band of overexpressed protein is shown in a light blue box. FepB was present in both total and soluble fractions for all four cultures.	200
Figure 117: UV-visible absorbance trace at 280 nm (black) for elution of FepB from nickel column over an imidazole gradient (blue) Elution volume from 114 mL to 170 mL was collected for the absorbance band observed at 280 nm.....	201
Figure 118: SDS PAGE analysis confirming the presence of FepB in fractions collected after nickel column chromatography. 4 mL fractions shown in the light blue box were collected and loaded in order of total elution volume. Lane 1: 32-36 mL, lane 2: 52-56 mL,	

lane 3: 114-118 mL, lane 4: 118-122 mL, lane 5: 122-126 mL, lane 6: 130-134 mL, lane 7: 142-146 mL, lane 8: 150-154 mL. All fractions were shown to contain FepB.	202
Figure 119: UV-visible absorbance trace at 280 nm (black) for elution of crude FepB from second nickel column over an imidazole gradient (blue). Fractions from 2 mL to 68 mL were collected for the absorbance band observed at 280 nm.	202
Figure 120: SDS PAGE of fractions after the second nickel column. Lane 1: tagged FepB, lane 2: crude untagged FepB. Untagged protein was confirmed in collected fractions of elution volume from 14 mL to 66 mL, lane 3: 14-18 mL, lane 4: 26-30 mL, lane 5: 62-66 mL. There was no FepB present in the fractions after the imidazole gradient as shown in lanes 6: 108-112 mL, 7: 116-120 mL. FepB fractions are highlighted in the light blue box.	203
Figure 121: Chromatogram showing absorbance at 280 nm (black) for fractions collected of size exclusion chromatography of FepB. Fractions from elution volume 62 mL to 74 mL were collected, giving the major absorbance peak.	204
Figure 122: SDS PAGE analysis of the collected fractions of FepB after size exclusion chromatography. Lane 1: crude FepB before size exclusion, pure FepB was confirmed in collected fractions by elution volume from 62 mL to 74 mL, lane 2: 46-50 mL, lane 3: 54-58 mL, lane 4: 58-62 mL, lane 5: 62-66 mL, lane 6: 66-70 mL, lane 7: 70-74 mL, lane 8: 70-74 (twice concentrated). Collected fractions are highlighted in the light blue box.	204
Figure 123: Circular dichroism spectrum of 0.4 mg mL ⁻¹ FepB in 5 mM Tris-HCl pH 7.5 to confirm correct folding of the protein.	205
Figure 124: Fluorescence quenching titration curves for FepB iron(III)-enterobactin Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for FepB-iron(III)-enterobactin and the model from Dynafit. ³²⁶ FepB at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-enterobactin ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	206
Figure 125: Fluorescence quenching titration curves for FepB-iron(III)-MECAM and the model from Dynafit. ³²⁶ FepB at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-MECAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	207
Figure 126: Circular dichroism spectra for FepB plotted as ellipticity in mdeg vs wavelength from 300 to 650 nm. The spectra were run for 50 μM FepB in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with 50 μM iron(III)-enterobactin or iron(III)-MECAM as well as iron(III)-enterobactin or iron(III)-MECAM at 50 μM in 0.11M Tris-HCl pH 7.5,	

150 mM NaCl buffer without FepB. The spectra show the induction of the Λ -configuration upon introduction of FepB to each iron(III)-ligand.	208
Figure 127: Circular dichroism spectra for FepB plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra are run for 50 μ M FepB in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μ M iron(III)-n-LICAM (n = 4, 5) or iron(III)-bisDHBS ligands as well as the ligands at 50 μ M in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without FepB. The spectra show the induction of Λ -configuration upon introduction of FepB to each iron(III)-ligand.	209
Figure 128: Circular Dichroism spectra for FepB plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm, using 50 μ M FepB in 0.11 M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μ M iron(III)-Sal-n-LICAM (n = 4, 5) ligands as well as the ligands at 50 μ M in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without FepB. The spectra show the induction of the Λ -configuration upon introduction of FepB to each iron(III)-Sal-n-LICAM (n = 4, 5).....	210
Figure 129: HPLC chromatograms showing multiple species present in the purchased salmochelin S4 sample. The sample in water containing 0.1% formic acid was injected with a 6-40% gradient of acetonitrile in water with 0.1% formic acid over 25 minutes. ...	211
Figure 130: Plot of normalised fluorescence emission vs ligand concentration in μ M for VctP-iron(III)-enterobactin and the model from Dynafit. ³²⁶ VctP at a concentration of 2.4 μ M in 40 mM Tris-HCl pH 7.5, 150 mM NaCl titrated with aliquots of 120 μ M iron(III)-enterobactin ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	213
Figure 131: Circular dichroism spectra for VctP plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm for 50 μ M VctP in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with 50 μ M iron(III)-enterobactin as well as iron(III)-enterobactin at 50 μ M in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without VctP. The spectra show the induction of the Λ -configuration upon introduction of VctP iron(III)-enterobactin, with a transition from a Δ -configured ligand complex to a Λ -configured protein-ligand complex.	214
Figure 132: Fluorescence quenching titration curve for VctP-iron(III)-MECAM and the model from Dynafit. ³²⁶ VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M iron(III)-MECAM ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	215
Figure 133: Fluorescence quenching titration curve for VctP-iron(III)-MECAM and the model from Dynafit. ³²⁶ 48 nM VctP in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with 2.4 μ M iron(III)-MECAM ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	216

Figure 134: Circular Dichroism spectra for VctP plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra were run for 50 μ M VctP in 0.11 M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μ M iron(III)-n-LICAM (n = 4, 5) and iron(III)-bisDHBS ligands.	216
Figure 135: Fluorescence quenching titration curve for VctP-iron(III)-Sal-4-LICAM and the model from Dynafit. ³²⁶ VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M iron(III)-Sal-4-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	217
Figure 136: Fluorescence quenching titration curve for VctP-iron(III)-Sal-5-LICAM and the model from Dynafit. ³²⁶ VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M iron(III)-Sal-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	218
Figure 137: Circular Dichroism spectra for VctP plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra were run for 50 μ M VctP in 0.11 M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μ M iron(III)-Sal-n-LICAM (n = 4, 5) ligands. The spectra show the induction of Λ -configuration upon introduction of VctP to each iron(III)-Sal-n-LICAM (n = 4, 5).	219
Figure 138: VctP crystals soaked with iron(III)-Sal-5-LICAM.	220
Figure 139: A: Overlay of VctP (lilac ribbons) with ViuP (dark purple), the vibriobactin binding PBP from <i>V. cholerae</i> . B: Stereo view of vibriobactin (cylinders coloured by atom type) in the binding cleft, and residues involved in binding (cylinders coloured by atom type with carbon in dark purple). C: Stereo view of neighbouring His 239 and Tyr 146 (cylinders coloured by atom type with carbon in dark purple) that may be involved in binding of lower denticity siderophore complexes (PDB ID: 3R5T). ²⁰⁸ Figure produced using CCP4mg.	225
Figure 140: Stereo view of the overlay of CeuE-iron(III)-MECAM crystal structure (PDB ID: 2CHU) with coordinates for iron(III)-enterobactin from FepB-iron(III)-enterobactin crystal structure (PDB ID: 3TLK). Iron(III)-enterobactin is shown in cylinders coloured by atom type with carbon in yellow. Iron(III)-MECAM is shown in cylinders coloured by atom type with carbon in white. A: The catecholamide units of each ligand superposed atom by atom. B: CeuE is depicted in surface view, coloured by electrostatic potential. ¹⁴⁶ . ¹⁸⁷ Figure produced using CCP4mg.	226
Figure 141: UV/Vis spectra of iron(III)-5-LICAM over a range of concentrations in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	273
Figure 142: Absorbance of iron(III)-5-LICAM at 280 nm plotted against concentration. $y = 0.0029x + 0.0049$. $R^2 = 0.9992$	273

Figure 143: UV/visible spectra for iron(III)-6-LICAM in a 1:1 ratio at varying concentrations.	274
Figure 144: UV/visible spectra of iron(III)-8-LICAM in a 1:1 ratio at varying concentrations.	275
Figure 145: Plot of iron(III)-6-LICAM absorbance at 280 nm vs concentration in μM , showing good adherence to the Beer Lambert law. Equation of the line of best fit: $y = 0.0027x - 0.0016$ $R^2 = 0.9991$	275
Figure 146: Linear relationship between iron(III)-8-LICAM absorbance at 280 nm vs concentration in μM . Equation of the line of best fit: $y = 0.0024x + 0.0029$ $R^2 = 0.9996$	276
Figure 147: UV/visible spectra of iron(III)enterobactin in a 1:1 ratio at varying concentrations.	277
Figure 148: Absorbance at 280 nm of iron(III)-enterobactin vs concentration. $y = 0.005x + 0.008$ $R^2 = 0.9999$	277
Figure 149: Equations used for calculation of weighted average and errors for K_d values. a_i = raw data value, σ_i = raw error, N = number of data values used ⁴⁵³	278
Figure 150: Fluorescence spectra of CeuE recorded over 75 minutes.	281
Figure 151: Fluorescence emission of CeuE monitored over time, showing adequate stability of the protein over 75 minutes of exposure to light of the excitation wavelength of 280 nm.	281
Figure 152: Decrease in fluorescence observed when CeuE is titrated with 40 mM Tris-HCl pH 7.5, 150 mM NaCl, shown as both uncorrected normalised data and a linear dilution correction across the titration.	282
Figure 153: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for FepB-Fe(III)-bisDHBS, FepB-Fe(III)-n-LICAM ($n = 4, 5$) and FepB-Fe(III)Sal-n-LICAM ($n = 4, 5$) and their associated non-linear regression fitting data from Dynafit. 240 nM FepB in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM Fe(III)-ligand ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	282
Figure 154: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for VctP titrated with buffer and Fe(III)-NTA. VctP was at concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl and was titrated with aliquots of 12 μM Fe(III)-NTA in 40 mM Tris-HCl pH 7.5, 150 mM NaCl, and the equivalent volume of 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	283
Figure 155: Fluorescence quenching titration curves for VctP-Fe(III)-enterobactin, as a plot of normalised emission vs ligand concentration in μM for VctP-Fe(III)-enterobactin and their associated non-linear regression fitting data from Dynafit. 240 nM VctP in 40	

mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M Fe(III)-enterobactin ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.....	283
Figure 156: Fluorescence quenching titration curve for VctP-Fe(III)-bisDHBS Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μ M for VctP-Fe(III)-bisDHBS and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M Fe(III)-bisDHBS ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	284
Figure 157: Fluorescence quenching titration curve for VctP-Fe(III)-bisDHBS Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μ M for VctP-Fe(III)-bisDHBS and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 48 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 2.4 μ M Fe(III)-bisDHBS ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	284
Figure 158: Fluorescence quenching titration curve for VctP-Fe(III)-5-LICAM Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μ M for VctP-Fe(III)-5-LICAM and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M Fe(III)-5-LICAM ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	285
Figure 159: Fluorescence quenching titration curve for VctP-Fe(III)-5-LICAM Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μ M for VctP-Fe(III)-5-LICAM and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 48 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 2.4 μ M Fe(III)-5-LICAM ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	285
Figure 160: Fluorescence quenching titration curve for VctP-Fe(III)-4-LICAM Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μ M for VctP-Fe(III)-4-LICAM and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M Fe(III)-4-LICAM ⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.	286
Figure 161: Dynafit script used for a general fluorescence titration. V is edited with an estimate of an approximate dissociation constant. W is edited for protein concentration used, X= 1/W, Y= Z/W where Z is the final normalised emission for the titration.....	287

Scheme 1: The Fenton reaction showing the generation of damaging reactive oxygen species. ²⁰	34
Scheme 2: Chemical structures of the amino acids used in mutations of His 227 and Tyr 288 in CeuE variants. Atoms that coordinate to iron(III) after deprotonation are shown in red.....	114
Scheme 3: General strategy for production of aryl-C-glycosides via generation of an electrophilic centre at the anomeric carbon.	151
Scheme 4: Intramolecular Friedel-Crafts reaction, resulting in pure α -anomeric product. ³⁸⁶	152
Scheme 5: A Grignard reagent method for selective generation of a β -anomeric product. ³⁸⁸	152
Scheme 6: The use of the Heck reaction to generate a β -anomeric product. ³⁹³	153
Scheme 7: Suzuki coupling reaction to generate a β -anomeric product. ^{383, 394}	153
Scheme 8: Negishi coupling reaction utilised for the synthesis of β -C5-glucosyl catechol units found in natural salmochelins. ³⁹⁶	153
Scheme 9: Synthesis of benzyl protected aminochelin (compound C).....	158
Scheme 10: Synthetic route for the production of the benzyl protected C5 iodinated 4-LICAM species (Compound J).	159
Scheme 11: Proposed Negishi coupling reaction of Compound J to produce Compound K and subsequent deprotection reactions to yield Sal-4-LICAM.	160
Scheme 12: Synthetic route for the production of Negishi coupled product, compound P.	162
Scheme 13: Amide coupling of the glucosylated catechol unit to benzyl-protected aminochelin, and the subsequent global benzyl deprotection via hydrogenolysis.....	163
Scheme 14: Scheme for the synthesis of a 5-carbon linker analogue of benzyl protected aminochelin (compound S).	165
Scheme 15: Amide coupling of the glucosylated catechol unit to benzyl protected 5-carbon analogue of aminochelin, and the subsequent global benzyl deprotection via hydrogenolysis.	166
Scheme 16: Enterobactin has all catechol units organised on one face of the trilactone backbone before iron(III)-binding. Potential hydrogen bonds are shown in blue, and those involved in preorganisation are shown in red.	228
Scheme 17: MECAM does not have a preorganised structure before iron(III)-binding (hydrogen bonds shown in blue) so may be able to adopt a tetradentate binding mode more easily in the CeuE or VctP binding clefts, by folding one catechol unit away from the iron(III) centre, allowing coordination of His and Tyr.	228

Table 1: Properties of a representative set of cyclic and linear siderophores.	45
Table 2: pFe(III) values for a selection of siderophore mimics.	69
Table 3: Dissociation constants for CeuE titrated with iron(III)-6-LICAM for three independent fluorescence titrations, and their associated error. The average was then calculated via a weighted average method.	94
Table 4: Dissociation constants for CeuE titrated with iron(III)-8-LICAM for three independent fluorescence titrations, and their associated error. The average was then calculated via a weighted average method.	95
Table 5: Dissociation constants for CeuE with iron(III)-n-LICAM (n = 4, 5, 6, 8) each calculated as a weighted average from three independent fluorescence titrations. Errors for each are calculated as an uncertainty in average. ^{310, 322}	96
Table 6: R.m.s.d in Å over number of C α positions for superpositions calculated using the SSM algorithm of CeuE and CeuE-iron(III)-n-LICAM crystal structures showing the similarity in atom positions for all structures. ³¹⁷	100
Table 7: Proteins identified with high structural similarity to CeuE via secondary structure element (SSE) matching by PDBeFold. All contain the conserved histidine and tyrosine residues.	110
Table 8: Comparison of CeuE variant crystal structures to that of wild-type CeuE. R.m.s.d is given for superposition with the wild-type structure via the SSM algorithm and the number of C α positions matched is given in parentheses.	115
Table 9: Calculated K _d values from fluorescence quenching data for CeuE-H227L/Y288F titrated with iron(III)-5-LICAM.	127
Table 10: Dissociation constants for triplicate runs of CeuE-H227L titrated with iron(III)-5-LICAM. The average was calculated using a weighted average method.	132
Table 11: Dissociation constants for triplicate runs of CeuE-H227A titrated with iron(III)-5-LICAM. The average was calculated using a weighted average method.	132
Table 12: Dissociation constants for wild-type CeuE, and variants CeuE-H227L and CeuE-H227A with iron(III)-n-LICAM (n= 4, 5, 6, 8) siderophore mimic compounds. CeuE-iron(III)-4-LICAM and CeuE-iron(III)-5-LICAM dissociation constants were taken from the literature. ³¹⁰	135
Table 13: Dissociation constants for the binding of CeuE to iron(III)-Sal-4-LICAM calculated using Dynafit from circular dichroism titrations.	179
Table 14: Dissociation constants for circular dichroism titration of iron(III)-Sal-5-LICAM with CeuE.	181
Table 15: Dissociation constants for 2 μ M CeuE titrated with iron(III)-Sal-5-LICAM.	183

Table 16: K_d values calculated for CeuE-iron(III)-enterobactin via fluorescence quenching titration and associated errors, giving a weighted average as detailed in Chapter 6.....	197
Table 17: K_d values for CeuE-iron(III)-MECAM from fluorescence quenching titration, with a weighted average calculated as detailed in Chapter 6.....	198
Table 18: Primers used for amplification of the FepB gene.	199
Table 19: K_d values calculated for FepB-iron(III)-enterobactin and FepB-iron(III)-MECAM via fluorescence quenching titration and associated errors, with a weighted average, as in Chapter 6.	207
Table 20: K_d values calculated for VctP-iron(III)-enterobactin via fluorescence quenching titration and associated errors, with a weighted average, as in Chapter 6.	213
Table 21: K_d values calculated for VctP-iron(III)-Sal-4-LICAM and VctP-iron(III)-Sal-5-LICAM via fluorescence quenching titration and associated errors, with a weighted average, as in Chapter 6.	218
Table 22: Dissociation constants for periplasmic binding proteins with iron(III)-siderophores and iron(III)-siderophore mimics. Dissociation constants taken from the literatures are referenced.	221
Table 23: Disordered regions that were not modelled in protein crystal structures of CeuE variant proteins.....	263
Table 24: Crystallographic statistics for crystal structures of CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM.....	264
Table 25: Refinement statistics for crystal structures of CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM.....	265
Table 26: Crystallographic statistics for crystal structures of CeuE-H227L, CeuE- H227L-iron(III)-5-LICAM and CeuE-H227A.	266
Table 27: Refinement statistics for crystal structures of CeuE-H227L, CeuE-H227L-iron(III)-5-LICAM and CeuE-H227A.	267
Table 28: Crystallographic statistics for crystal structures of CeuE-Y288F, CeuE-H227L/Y288F and CeuE- H227A/Y288F.	268
Table 29: Refinement statistics for crystal structures of CeuE-Y288F, CeuE-H227L/Y288F and CeuE- H227A/Y288F.	269

Accompanying Material

A compact disk containing an electronic copy of this thesis, data for synthesis and characterisation, fluorescence quenching and circular dichroism experiments, the Dynafit script, PDB, MTZ and CIF files associated to the X-ray diffraction data is attached at the back of this thesis.

Acknowledgements

With thanks to the EPSRC and Prof. Keith S. Wilson for funding, and to the Royal Society of Chemistry Dalton Division for a generous conference travel grant. Thanks to the University of York for providing excellent research facilities and support.

There are many people who must be thanked for their contributions and support from the beginning of my PhD to final submission. Firstly, my supervisors Prof. Keith S. Wilson and Prof. Anne. K. Duhme-Klair who have always been willing to invest their time and effort into whatever I have asked of them, always providing helpful advice and suggestions.

Thank you to Elena for all of her hard work in teaching me, providing me with purified proteins, and her tireless efforts in crystal screening. Thanks to Johan and Sam for all their hard work in facilitating the crystallography, and to Johan, Olga and Eleanor for their advice in structure solution. Thank you to Justin for his advice on all matters biological. Thank you to Heather, Karl, Graeme and Amanda for providing their help and services in compound characterisation. Thanks to Ross for being a very hardworking summer student. Thank you to Adam, Claire, and all the friendly faces that have welcomed me in YSBL.

This process would have been much more difficult without the friendship, support and advice of everyone in E014, with special mention to Ben, Chris, Danielle, Lewis, Lucy, Luisa, Martin and Natalie. A huge thank you to Tom, who was a massive aid in my synthetic chemistry, and a massive supporter of my baking. A gigantic thank you to Dan, who pretty much holds the lab together, and has been a source of endless help, support and patience across all aspects of my work.

Thank you to my wonderful family, friends and parents for providing your support and encouragement throughout the entire process. I have an amazing support network, and I hope to now spend more quality time with you all. A special mention must be made to my two amazing grandmothers, who taught me the importance of patience, a positive attitude, and good cake.

Finally, thank you to Michael, who deserves a medal for his tireless support, reassurance and patience. I could not have done any of this without him.

Declaration

I declare that I am the sole author of the work presented in this thesis and to the best of my knowledge, the work is original. This work was undertaken by myself at The University of York under the supervision of Prof. Anne-Kathrin Duhme-Klair and Dr. Keith Wilson between October 2013 and August 2017. The work presented within this thesis has not been submitted for any other qualification at this or any other university. Any additional publications are declared below and all sources are acknowledged as references. The following work was carried out by or with the assistance of other researchers:

All mass spectrometry experiments were performed by Mr. K. Heaton, Ms. H. Robinson and Ms. R. Cercola. Elemental analyses were performed by Dr. G. McAllister. HPLC experiments were carried out with the assistance of Ms. A. Dixon. bisDHBS was synthesised and characterised by Dr. D. J. Raines. n-LICAM (n= 4, 5, 6, 8) and MECAM compounds were synthesised and characterised by Dr. D. J. Raines and Prof. A. K. Duhme-Klair. Synthetic strategy for the production of salmochelin mimic compounds Sal-n-LICAM (n= 4, 5) was aided by Dr. T. J. Sanderson. Synthesis of 1-Amino, 5-(2,3-dibenzyloxybenzamide) pentane (Compound S) was carried out by project student Mr. R. P. Thomas. Computational modelling was conducted, and related figures produced by Mr. P. S. Bond. CeuE was produced, purified and characterised by Dr. E. V. Blagova, Dr D. Raines and Mr. A. Hughes. CeuE variant proteins and VctP were produced, purified and characterised by Dr. E. V. Blagova. FepB was produced, purified and characterised with the assistance of Dr. E. V. Blagova. Apo CeuE variant proteins were crystallised and cryoprotected by Dr. E. V. Blagova. Fluorescence quenching titrations for mutant CeuE-H227A were carried out by project student Mr. R. P. Thomas. Data collection at The Diamond Light Source was conducted by Dr. J. P. Turkenburg and Mr. S. Hart.

Citation for publication of results detailed in chapters 2 and 3:

E. J. Wilde, A. Hughes, E. V. Blagova, O. V. Moroz, R. P. Thomas, J. P. Turkenburg, D. J. Raines, A.-K. Duhme-Klair and K. S. Wilson, *Sci. Rep.*, 2017, **7**, 45941.

Citation for an authored book chapter:

Raines D. J., Sanderson T. J., Wilde E. J. and Duhme-Klair A.-K. (2014) Siderophores. In: Reedijk, J. (Ed.) Elsevier Reference Module in Chemistry, Molecular Sciences and Chemical Engineering. Waltham, MA: Elsevier. 23-Jan-15 doi: 10.1016/B978-0-12-409547-2.11040-6.

Authored PBD depositions:

5A5D **DOI:** 10.2210/pdb5a5d/pdb

5A5V **DOI:** 10.2210/pdb5a5v/pdb

5AD1 **DOI:** 10.2210/pdb5ad1/pdb

5LWH **DOI:** 10.2210/pdb5lwh/pdb

5LWQ **DOI:** 10.2210/pdb5lwq/pdb

5MBQ **DOI:** 10.2210/pdb5mbq/pdb

5MBT **DOI:** 10.2210/pdb5mbt/pdb

5MBU **DOI:** 10.2210/pdb5mbu/pdb

5TCY **DOI:** 10.2210/pdb5tcy/pdb

Ellis Wilde

1 Introduction

1.1 Overview

This thesis is concerned with the study of periplasmic binding proteins involved in bacterial iron uptake, the iron(III)-siderophore complexes that are bound by such proteins, and the characterisation of the resulting protein-iron(III)-siderophore complexes. This chapter provides a summary of the area of research that surrounds this project, and puts the succeeding work into context.

1.2 Iron Co-ordination

Iron, the second most abundant metal in the Earth's crust, possesses an electronic configuration with unfilled d-orbitals, enabling access to multiple oxidation states for a range of redox activities.^{1, 2} As such, iron is employed in a broad spectrum of biological systems as both an electron donor and electron acceptor and participates in a range of redox chemistries for functions in catalysis and electron transfer.^{2, 3} Iron can exist in oxidation states between -2 and +6, however the most common are +2 (d^6) and +3 (d^5), for which both high and low spin states can be accessed.^{2, 4}

Iron(III) oxides are the most common iron source in the biosphere, and the most common coordination geometry for iron(III) compounds is octahedral with six coordinating atoms.⁵ Iron(III) is a small and highly-charged cation, with an ionic radius of 0.65 Å and is a hard Lewis acid.⁶ This leads to a preference for ligands that have donor atoms with hard Lewis base character, such as charge-dense cyanide ions, or negatively-charged oxygen donors.^{2, 5} In aqueous solution above pH 1, iron(III) undergoes hydrolytic polymerisation, causing precipitation of insoluble hydroxo and oxo species unless coordinated by suitably stabilising ligands.^{5, 7, 8}

The stereochemistry of metal complexes is important for receptor recognition in biology.⁹ ¹⁰ Multidenticity of ligands around the metal centre can result in chirality, and for iron(III) produces either Λ - or Δ -configured octahedral complexes (Figure 1).^{11, 12} The Λ -configuration is often referred to as an octahedral 'left-handed propeller' of three bidentate ligands around a metal centre, and the Δ -configuration the 'right-handed propeller'. These configurations are possible for hexadentate ligands, or combinations of multidentate ligands.^{11, 13} Complexes may be kinetically inert, such as tris-catecholate chromium(III) complexes, where the Λ - and Δ -configured isomers do not exchange.^{13, 14} Alternatively, complexes may be in equilibrium between the Λ - and Δ -configuration, as is the case for

iron(III), for which it is possible for chiral ligands to influence the preference for one stereoisomer over the other.¹⁵⁻¹⁷

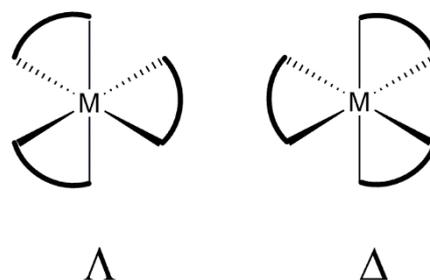


Figure 1: A tris-bidentate octahedral metal complex in a Λ - and Δ -configuration.¹¹

1.3 Iron Homeostasis in Bacteria and Mammalian Host

Iron is the most widely employed transition metal in biology, likely due to its versatility as a redox active centre, and maintenance of strict iron concentrations within living cells is essential.¹ Iron(II) complexes can be used in one-electron reductions, and one-electron oxidation reactions.¹⁸ If the concentration of free iron is too high, this can lead to the production of damaging reactive oxygen species, such as hydroxyl radicals, *via* Fenton chemistry (Scheme 1).¹⁹⁻²¹



Scheme 1: The Fenton reaction showing the generation of damaging reactive oxygen species.²⁰

In a eukaryotic host, the acquisition of excess iron allows the proliferation of unwanted microbial infections.^{1, 21} However, if iron concentrations are too low then key metabolic processes will not function correctly. Iron is incorporated in both eukaryotic and prokaryotic enzymes in the form of organic co-factors such as haem, as inorganic cofactors or iron-sulfur clusters.¹⁹ Although oxidation states of iron(II) and iron(III) are most common in biology, iron(IV) or even iron(V) can be achieved in haem containing proteins.¹⁸ The reduction potentials of biological iron complexes can vary greatly depending on the ligand environment, from +310 mV in oxyhaemoglobin, to -500 mV in transferrin.^{3, 18} The versatile range of enzymes that result from incorporation of iron are required in key cellular processes including respiration, amino acid synthesis, DNA replication and electron transport.²²

Animals obtain the iron they need through diet; however, bacteria must acquire iron from whatever sources are present in their environment. Iron(II) is soluble, but is not abundant in aerobic environments, and so the majority of iron that bacteria encounter is in the form of

iron(III), which is largely insoluble in aqueous media, with a solubility product generally estimated at around 10^{-18} - 10^{-17} mol dm⁻³,^{3, 23} but more recently estimated by more complex models as 1.4×10^{-9} mol dm⁻³ at pH 7,²⁴ due to the formation of soluble ferric hydroxide species.^{20, 25, 26} This means that free iron is not immediately accessible for uptake by living cells. In a mammalian host, iron is locked up in proteins such as haemoglobin, transferrin, and ferritin, and this iron cannot be easily accessed by bacteria.²⁰ Requiring concentrations between 10^{-8} and 10^{-6} mol dm⁻³ for growth,³ iron is almost always restricted, and complex strategies must be employed for acquiring enough iron for survival.^{19, 25}

Bacterial metal-acquisition systems are generally classed into importers of elemental ions, acquisition *via* host iron-sources and extracellular capture *via* secretion of chelating compounds that are named siderophores.¹ Host iron sources are commonly in the form of a haem-containing protein, as in mammals, the majority of iron is found in haemoglobin.^{20, 27} Haem can either be directly taken up through the outer membrane of a bacterial cell *via* a specific receptor, or first sequestered by secreted proteins named haemophores.^{19, 20, 25}

1.4 Siderophores

Siderophores are low molecular weight (<1500 Da) compounds that have a high affinity for iron(III),³ and are used by bacteria, fungi and some plants for iron(III) uptake when under iron deficiency.^{25, 28} They can be endogenous - synthesised and secreted for the specific purpose of iron(III) sequestration or exogenous - as an existing biomolecule that happens to coordinate iron(III). Siderophores must be selective for iron(III) over other metals, so these ligands contain 'hard' oxygen donor atoms, to satisfy the binding requirements of small highly charged iron(III). Although aluminium(III) is abundant in the Earth's crust, siderophores are able to select for iron(III), due to the differences in ionic radius (0.54 Å for aluminium(III) and 0.65 Å for iron(III)).^{28, 29} Siderophores contain one or a mixture of iron-coordinating functional groups, and the most common of these are catecholates, hydroxamates, carboxylates, α -hydroxycarboxylates and phenolates (Figure 2).^{28, 30} The most effective siderophores are hexadentate, that arrange all six iron-coordinating atoms into an entropically favourable octahedral structure, forming a high-spin iron(III) complex.^{2, 28,}

30-32

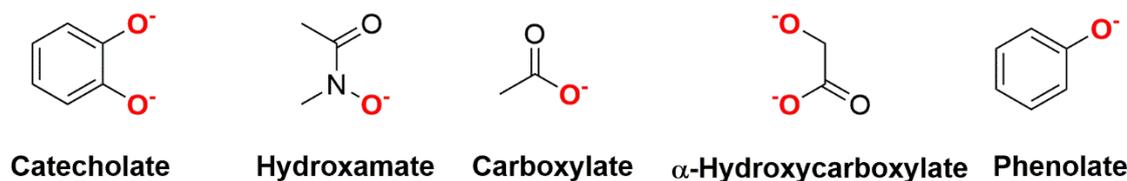


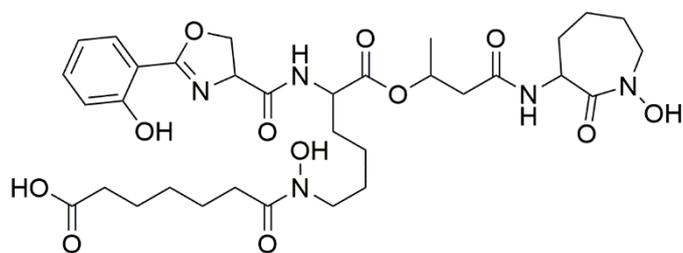
Figure 2: Oxygen containing functional groups common to siderophores. Oxygen atoms that must be deprotonated for iron(III) binding are shown in red.²⁸

As iron(III)-binding units must first be deprotonated, proton competition in aqueous media must be considered when estimating the iron(III)-binding affinity of a siderophore. It is therefore necessary to report true affinities as pFe(III) at a known pH (commonly 7.4), rather than as a simple formation constant (K_f).²⁸ The pFe(III) value is the $-\log[\text{Fe(III)}]_{\text{free}}$ in an aqueous solution where $[\text{Fe}]_{\text{total}}=1 \text{ mM}$ and $[\text{siderophore}]_{\text{total}}=10 \text{ mM}$, and therefore a large pFe(III) value suggests high iron(III) affinity.^{2, 6, 28, 30}

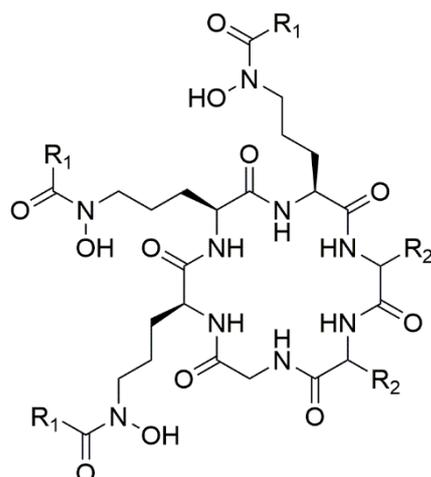
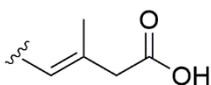
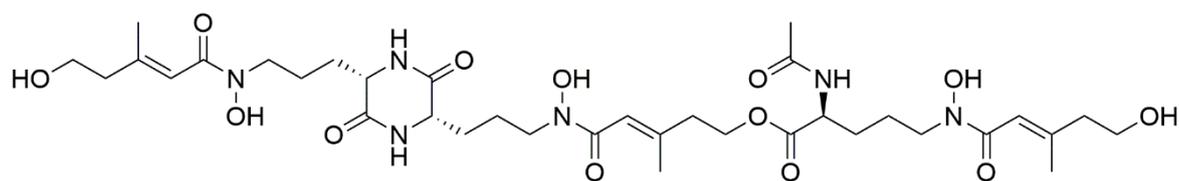
Discovery of Siderophores

Siderophores were discovered as microbial growth factors in the 1950s and the first isolated were ferric complexes of mycobactins, coprogen and ferrichrome (Figure 3).^{2, 28, 33-38} Mycobactins are mixed-type hydrophobic siderophores produced by mycobacteria that are located within the cell envelope, thought to require additional extracellular siderophores to effectively supply the cell with iron(III).²⁵ Coprogen is a fungal hydroxamate siderophore, and has been shown to be taken up by *Escherichia coli*.^{38, 39} Ferrichrome is also a fungal hydroxamate siderophore, but has been shown to support growth of *Pseudomonas aeruginosa*.^{40, 41} The first report to show that siderophore production is iron-regulated involved the observation that ferrichrome A (Figure 3) was produced by iron deficient *Ustilago sphaerogena*,³⁶ followed by confirmation of the structures of both ferrichrome and ferrichrome A in 1961.^{40, 42}

Siderophore identification assays were developed in the 1980s for the rapid identification of natural iron-binding compounds.⁴³ Today, over 500 natural siderophores have been isolated and characterised.²⁸ In marine environments there is a prevalence for amphiphilic siderophores including the marinobactins, aquachelins and amphibactins.^{44, 45} Siderophore units usually contains hydroxamate and/or α -hydroxycarboxylate functional groups, which are appended to an aliphatic chain, that can vary in length depending on the required partition coefficient. It is thought that such siderophores do not diffuse away from the cell envelope, and are in fact embedded in the outer-membrane- and prove a particularly useful feature for species colonising mobile fluid environments.^{2, 44-47}



Hydrophilic Mycobactin T

R₁ = CH₃R₂ = H FerrichromeR₁ =  R₂ = CH₂OH Ferrichrome A

Coprogen

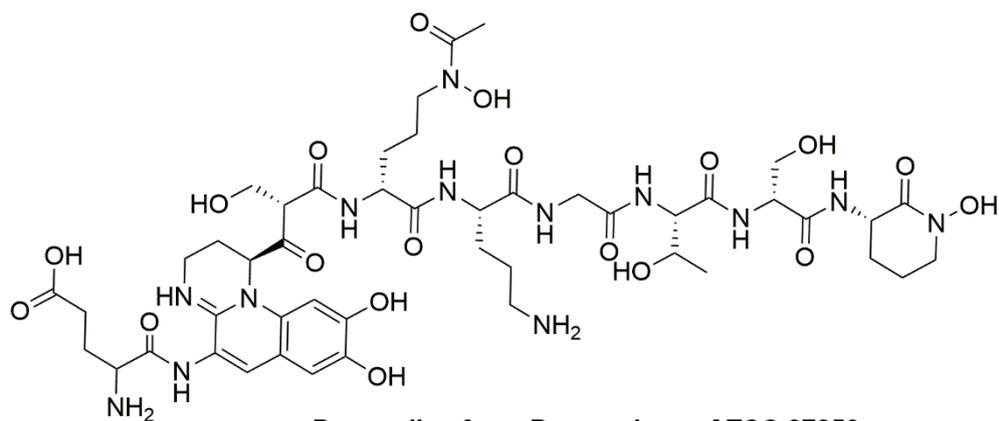
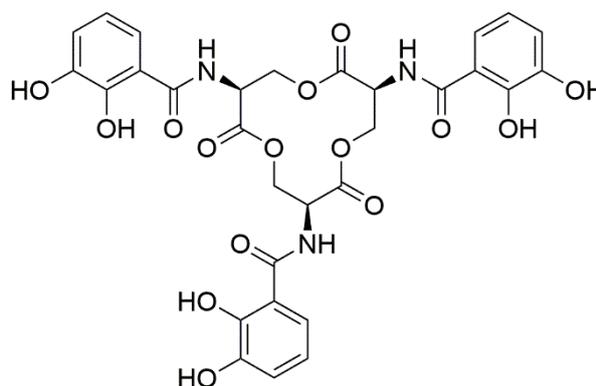
Pyoverdine from *P. aeruginosa* ATCC 27853

Figure 3: Structures of the earliest siderophores discovered, and an example of a fluorescent pyoverdine.^{28, 39, 40, 42, 48}

Notable are the pyoverdine class of siderophores (Figure 3), for which there are over 70 structural analogues.^{28, 49} Pyoverdines are mixed-type fluorescent siderophores produced by *Pseudomonas spp.*, containing two hydroxamate groups, and one catecholate group attached to a fluorophore.⁴⁸⁻⁵² *Pseudomonas aeruginosa* produces a pyoverdine that is known to be important for virulence and biofilm formation, which is a particular threat of lung infections in cystic fibrosis patients.⁵⁰ These siderophores are very effective at iron(III) sequestration, as they act by delivering iron to the periplasm, where the iron is released by reduction to iron(II), and can be recycled without the need for degradation of the siderophore.^{53, 54} Reducing iron(III) to iron(II) decreases thermodynamic stability of the siderophore complex, by decreasing charge density on the metal centre, decreasing the Lewis acidity of the metal, hence decreasing the affinity for the negatively charged donor oxygen atoms and increasing the capacity for ligand exchange.^{55, 56}

Enterobactin and Catecholate Siderophores

Enterobactin, the first triscatecholate siderophore to be identified, was isolated and characterised in 1970, from cultures of both *E. coli* and *Salmonella typhimurium*.⁵⁷⁻⁶²



Enterobactin (2,3-DHB-^LSer)₃

Enterobactin is one of the highest affinity iron(III) binding compounds known at physiological pH with a pFe(III) of 35.5.²⁹ The siderophore has a number of properties that enable such effective iron(III) binding, including a cyclic tris(lactone) backbone, iron binding catecholate units, and the way in which the catecholamides form favourable hydrogen bonds.^{37, 63} The triserine lactone backbone was shown to be the optimum size for spacing and arrangement of the catecholamide groups for iron(III) binding.⁶⁴ The rigidity that is conferred from pre-organisation of the ring provides an entropic contribution to the stability of the iron(III) complex of enterobactin.^{37, 64} Indeed when the ring is hydrolysed, liberating the more hydrophilic trisDHBS, complex-stability is lost.^{6, 65} The backbone also constrains enterobactin to influence the selective formation of the Δ -configured iron(III) complex.⁶⁴

The optimum donor-metal-donor angle is created between iron(III) and the two catechol oxygens of each bidentate catecholamide unit, as observed in the bonding of iron(III)-enterobactin by single-crystal polarized absorption and magnetic circular dichroism spectroscopy.^{37, 66, 67} Fe L-edge X-ray absorption spectroscopy showed that the high stability constants of catecholate complexes are due to a large contribution of σ and π -donation from ligand to metal.⁶⁷ The primary amide adjacent to the catecholate unit is of vital importance to stability of the iron(III)-complex.³⁷ The hydrogen bond that forms between the amide and the ortho-phenolate allows for a lower pK_a , and reduces proton competition, enhancing iron(III)-complex stability. Similar ligands that lack this hydrogen-bonded arrangement only form stable iron(III) complexes at basic pH.³⁷ This amide confers stability of the complex under acidic conditions, and allows for protonation of the meta-phenolate without loss of iron(III). This is achieved *via* adoption of the less stable salicylate binding mode, whereby rotation of the amide allows for direct coordination of the amide carbonyl oxygen with the metal centre (Figure 4).^{68, 69} The salicylate binding mode has been extensively studied, and this less stable form is thought to be important for intracellular iron-release.⁶⁸⁻⁷²

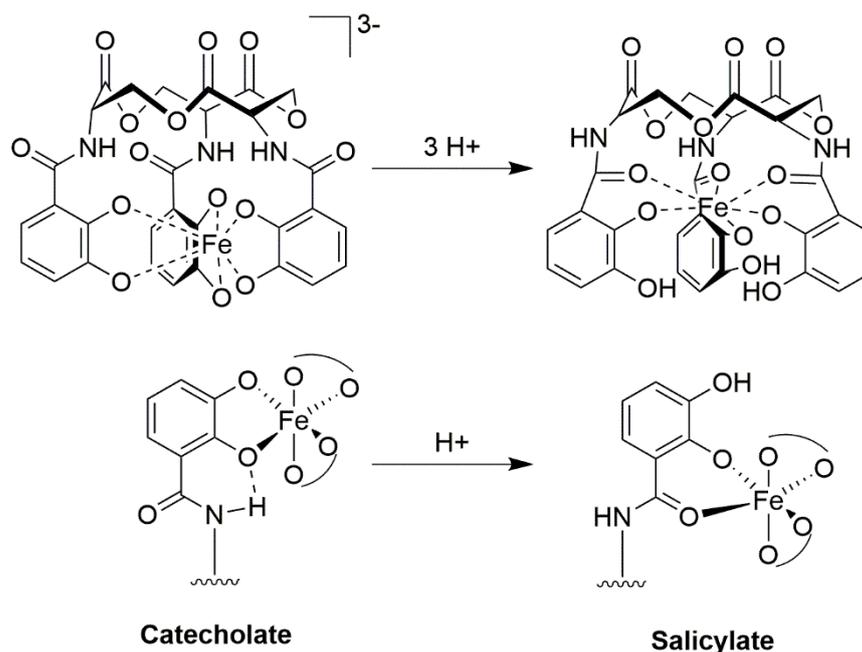


Figure 4: Protonation resulting in a switch between catecholate and salicylate binding mode. Top: Representation of the iron(III)-bound catecholate and salicylate modes of enterobactin. Bottom: Schematic of the catecholate and salicylate binding modes for one bidentate catecholamide unit.⁶⁸

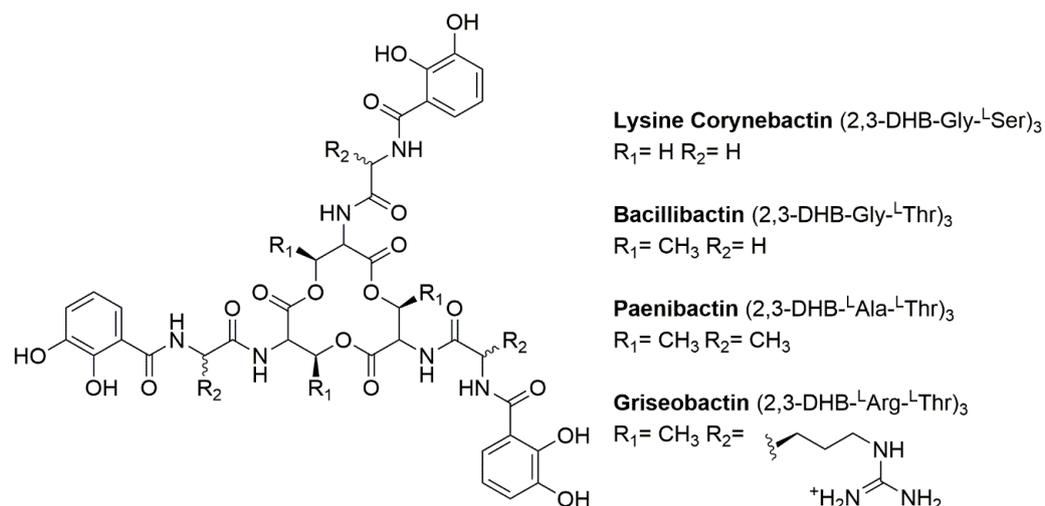


Figure 5: Triscatecholate siderophores with varied amino acid backbones.^{28, 73}

Since the discovery of enterobactin, many catecholate siderophores have been characterised from a range of bacterial species. These include hexadentate, tetradentate and bidentate systems, and a number of these are close variations of enterobactin, incorporating a variety of amino acids in the cyclic trilactone backbone structure (Figure 5).⁷³ Producing structurally similar but slightly different siderophores allows for high-affinity iron(III)-binding but limits theft by other species. This concept is well documented for the range of siderophores synthesised and employed by *Streptomyces* species.⁷⁴ Examples of catecholate siderophore variations are bacillibactin (also commonly named corynebactin) produced by *Bacillus subtilis* and lysine-corynebactin that cannot be used by the *E. coli* for iron delivery,¹⁰ but still have a high iron(III)-affinity cyclic triscatecholate structure.^{75, 76} This may be related to the fact that enterobactin produces a predominantly Δ -configured iron(III) complex, while bacillibactin and lysine-corynebactin, with glycine spacers, both produce a Λ -configured major complex, causing incompatibilities with uptake receptors or the esterases required for iron(III) release.^{17, 37, 77} In addition, the introduction of threonine instead of serine, as is the case in paenibactin produced by *Paenibacillus elgii* and griseobactin produced by *Streptomyces spp.*, increases stability of the backbone against acid or enzymatic hydrolysis.^{10, 78, 79} Other modifications include the insertion of an acylated serine into the trilactone backbone, creating a tetralactone with a fatty acid unit attached in addition to the three catecholamides, as is the case in amphienterobactin siderophores, biosynthesised by *Vibrio campbellii* (Figure 6).^{73, 80, 81} This modification confers hydrophobic character to the siderophore, altering the partition coefficient, whilst still maintaining high selectivity for iron(III). Linear hexadentate catecholate siderophores also exist, including agrobactin, agrobactin A, vibriobactin, fluvibactin, protochelin and trisDHBS (Figure 6).^{28, 82-87} In addition, a number of mixed type hexadentate siderophores are known that contain two

catecholate units, such as the one hydroxamate siderophore fuscachelin A (Figure 6) produced by *Thermobifida fusca*,^{81, 88} or the one phenolate siderophore parabactin (Figure 6) produced in *Paracoccus denitrificans*.^{28, 37, 89}

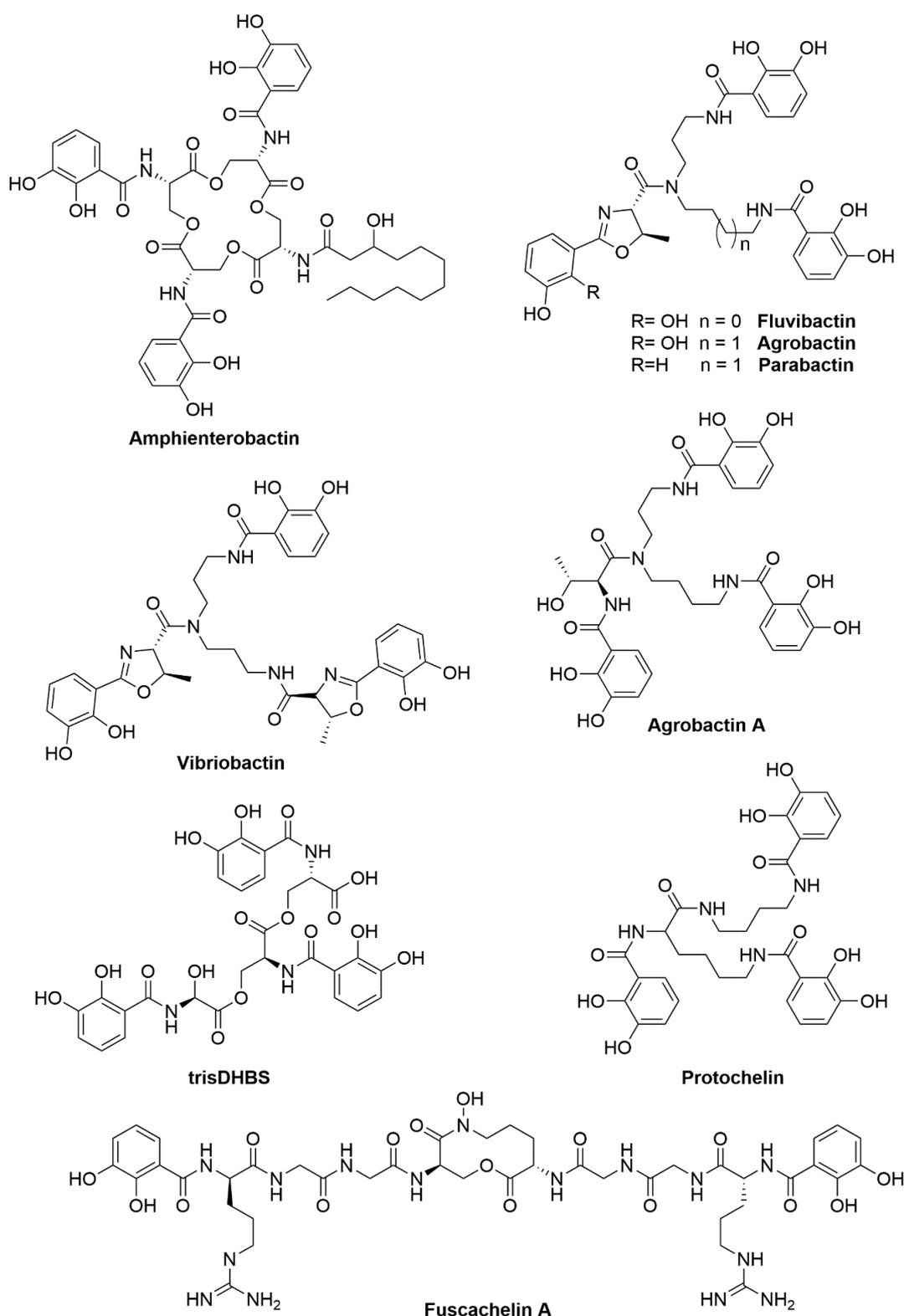


Figure 6: Amphienterobactin, a range of structurally similar linear hexadentate catecholate siderophores and mixed type siderophore fuscachelin A.^{80-82, 84, 85, 88}

Tetradentate Catecholate Siderophores

Siderophores were once thought to form exclusively hexadentate octahedral iron(III) complexes.³ However a number of lower denticity iron(III)-siderophores are known to be of biological relevance.⁹⁰⁻⁹³ The lower the denticity, the weaker the iron(III)-complex stability (as demonstrated later in section 1.4),²⁸ and a range of stabilities may prove useful in natural iron-uptake and release mechanisms. There are a diverse range of tetradentate catecholate siderophores known in nature produced by a wide number of bacterial species.²⁸ Although they are unified in their structures by containing two catecholate iron(III)-binding units, the backbones often vary in functionality and length. Most contain a 5-atom linker, including azotochelin, myxochelin, bisDHBS and salmochelin S1, the enterobactin and salmochelin S4 hydrolysis products respectively and fimsbactins D and E (Figure 7).⁹⁴⁻⁹⁷ The linker can be extended to 6 atoms as in serratiochelin, 8 atoms in pistillarin, or even up to 16 atoms as in amonabactin P 750.^{92, 93, 98} This variation demonstrates that the desirable iron(III)-binding properties are achieved in all structures, but by incorporating an unusual linker design, each species is able to evade the outer-membrane receptors of other species, and thereby limit competition.^{73, 74, 99} Hexadentate siderophores generally form 1:1 octahedral iron(III) complexes, however tetradentate siderophores have two vacant coordination sites when bound in a 1:1 ratio. This phenomenon means that other siderophore: iron(III) ratios are often favourable, with the self-assembly of a 3:2 complex common in aqueous solution.^{28, 90, 100, 101} In addition, several natural tridentate and bidentate catecholate siderophores are known to promote bacterial growth and include chrysobactin, aminochelin, DHBS, itoic acid and acinetobactin (Figure 8).^{28, 102-106}

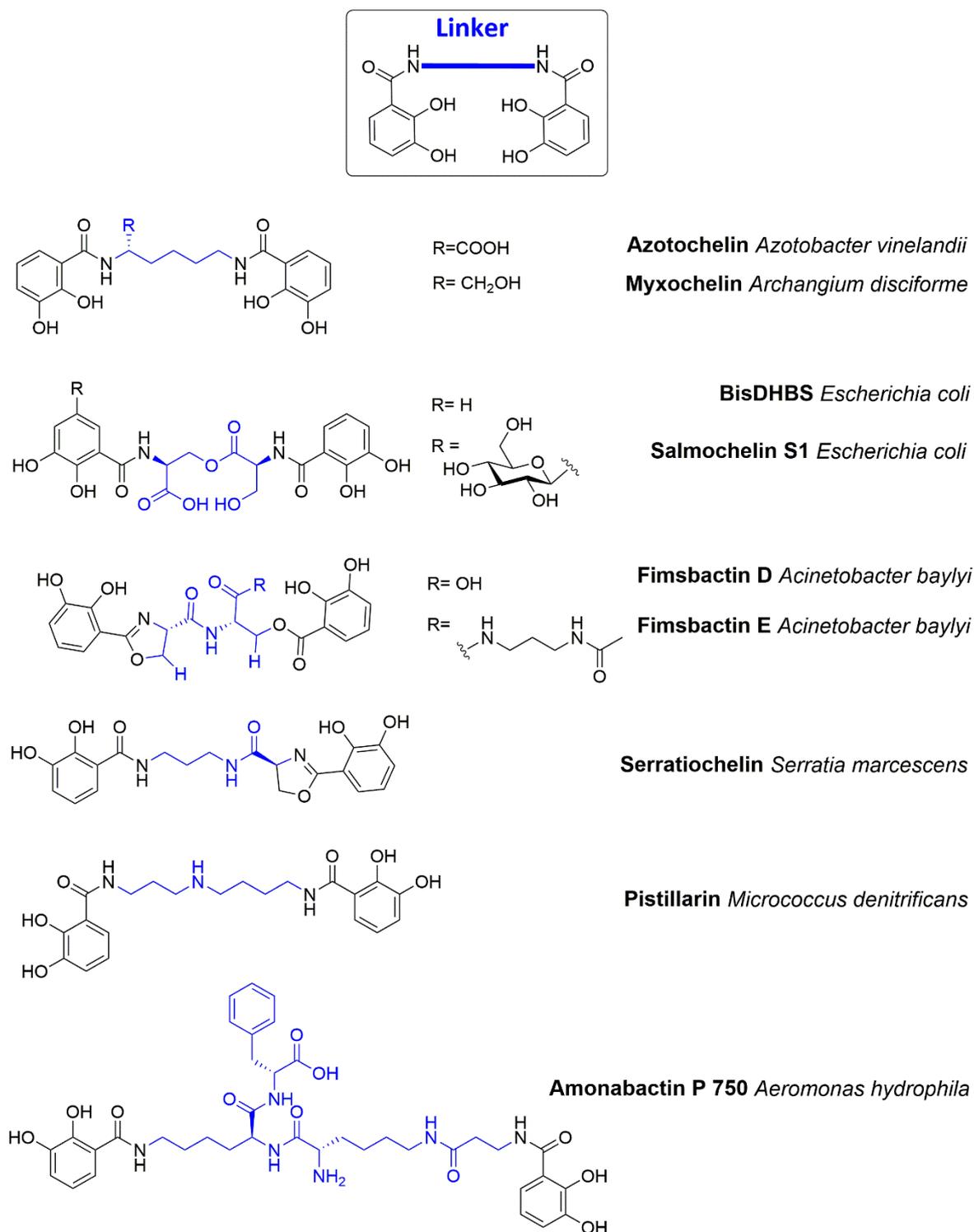


Figure 7: Tetradentate catecholate siderophores from a range of bacterial species with variable linker regions highlighted in blue.^{91, 93, 96-98, 107}

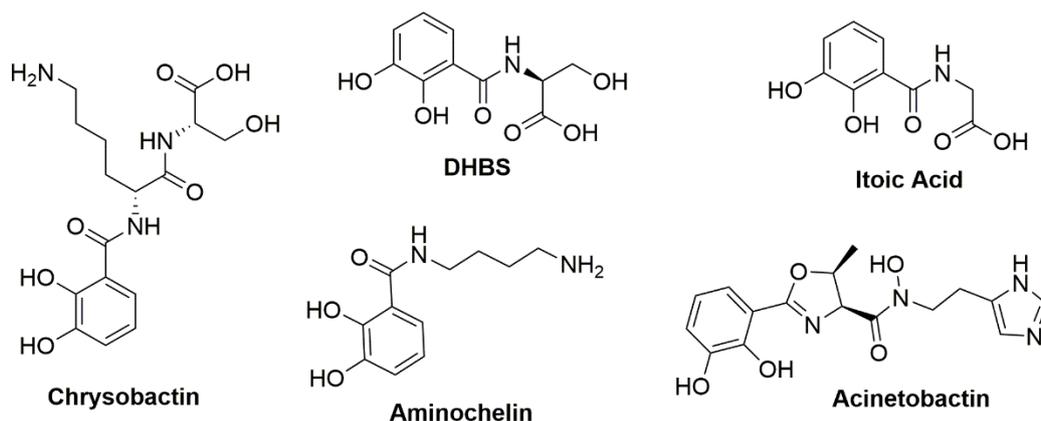


Figure 8: Tridentate and bidentate catechol siderophores.²⁸

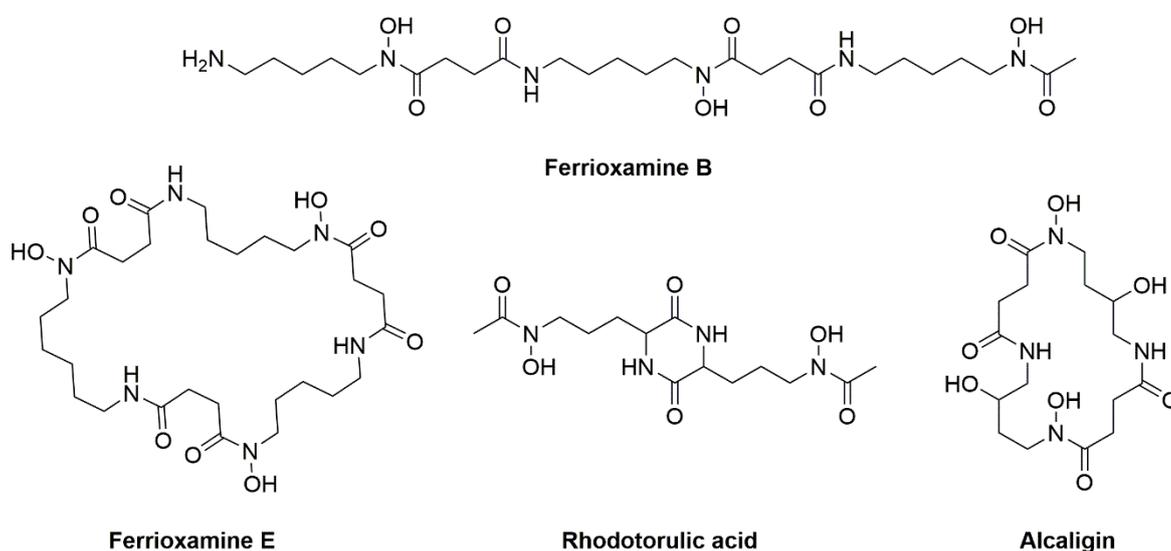
Properties of Siderophores

With very high formation constants for iron(III), siderophores can strip iron from host proteins such as ferritin, transferrin and lactoferrin as well as allowing chelation of any available insoluble ferric hydroxide.²⁵ Catechol siderophores have a pK_a of between 6.5 and 8 for deprotonation of the first proton, and around 11.5 for the second proton.³⁰ Carboxylates have low pK_a values around 4, and hydroxamates around 8.5. This means that carboxylate siderophores are best employed in acidic environments, while catechol siderophores are better suited to neutral pH.^{30, 62}

Cyclic siderophores tend to have high iron(III) binding affinity, due to their rigid pre-organised structure that is entropically favourable for iron(III) chelation as discussed for enterobactin.²⁶ This is demonstrated by ferrioxamine E having higher affinity for iron(III) than linear ferrioxamine B, enterobactin a higher affinity for iron(III) than linear mimic TRENAM and alcaligin a higher affinity for iron(III) than linear rhodotorulic acid as demonstrated by their $pFe(III)$ values (Figure 9) (Table 1).^{26, 63} Furthermore, hexadentate siderophores tend to have higher iron(III) affinity than lower denticity siderophores, and this is reflected in lower $pFe(III)$ values and less negative redox potentials for the resulting iron(III) complexes.¹⁰⁸ This is demonstrated by the lower $pFe(III)$ values for both tetradentate hydroxamate siderophores alcaligin and rhodotorulic acid when compared to hexadentate hydroxamates Ferrioxamine E and B (Figure 9), as well as for the tetradentate catechol siderophore amonabactin T789 when compared to enterobactin (Table 1).¹⁰⁹ This phenomenon allows tetradentate siderophore complexes to be more labile, with a greater rate of iron(III) exchange than hexadentate analogues, which may be useful for rapid iron(III) release *in vivo*.²⁶

Table 1: Properties of a representative set of cyclic and linear siderophores.

Siderophore	Structure	Type	Denticity	pFe(III)
Ferrioxamine E	Cyclic	Hydroxamate	6	27.7 ¹¹⁰
Ferrioxamine B	Linear	Hydroxamate	6	26.6 ¹¹⁰
Enterobactin	Cyclic	Catecholate	6	35.5 ²⁹
TRENCAM	Linear	Catecholate	6	27.8 ⁶³
Alcaligin	Cyclic	Hydroxamate	4	23.0 ^{110, 111}
Rhodotorulic Acid	Linear	Hydroxamate	4	21.8 ¹¹⁰
Amonabactin T789	Linear	Catecholate	4	25.8 ¹⁰⁹

Figure 9: Linear and cyclic hydroxamate siderophores.^{39, 110}

Redox potentials can also be used as a measure of iron(III) complex stability of siderophores.¹⁸ The lower the redox potential, the more energy required to release iron from the complex *via* reduction. Enterobactin has a very low redox potential of -750 mV ($E_{1/2}$ vs NHE), meaning the complex is very stable, and a strong reducing agent is required for iron release. Ferrioxamine B has a higher redox potential of -468 mV ($E_{1/2}$ vs NHE) and so forms a less stable iron(III) complex than enterobactin, and iron release by reduction does not require such a strong reducing agent.^{18, 26, 110} A correlation between redox potential and pFe(III) value has been shown for a number of siderophore and siderophore mimics.¹⁰⁸

1.5 Bacterial Iron Uptake

Competitive versus cooperative behaviour in bacterial siderophore production and uptake is a topic of debate.¹¹²⁻¹¹⁴ It is commonly suggested that bacteria attempt to outcompete each other for available iron sources using a complex variety of siderophores, and it is overlooked

that the siderophores may contribute to the overall growth of a symbiotic colony of mixed species. If bacterial species are able to take up siderophores that they do not produce themselves, then these species benefit from co-colonisation of a host.¹¹⁵ Siderophore biosynthesis comes at a metabolic cost, but the more closely related species are to each other, the more likely they are to be able to share siderophores, so it is possible that similar bacterial species work in a cooperative manner rather than a competitive one.¹¹³ However shared siderophores are downregulated, and more complex specialised siderophores are upregulated in the presence of significantly different opposing species,^{112, 116} and the diversity of siderophores structures is thought to arise due to the advantages of limiting piracy between species.⁷³

Iron Uptake in Gram-Positive Bacteria

Gram-positive bacteria with a thick cell wall and only one cell membrane, have specific binding proteins, often termed ‘substrate binding proteins’ that are anchored in the cell membrane.²⁵ These proteins bind iron sources, and deliver them to ATP-binding cassette transporter systems, which act to deliver the iron complexes into the cytoplasm. A simplified diagram of Gram-positive bacterial iron(III) uptake is depicted in Figure 10.^{19, 20, 117} These systems exist for siderophore complexes as well as haem and transferrin.²⁰

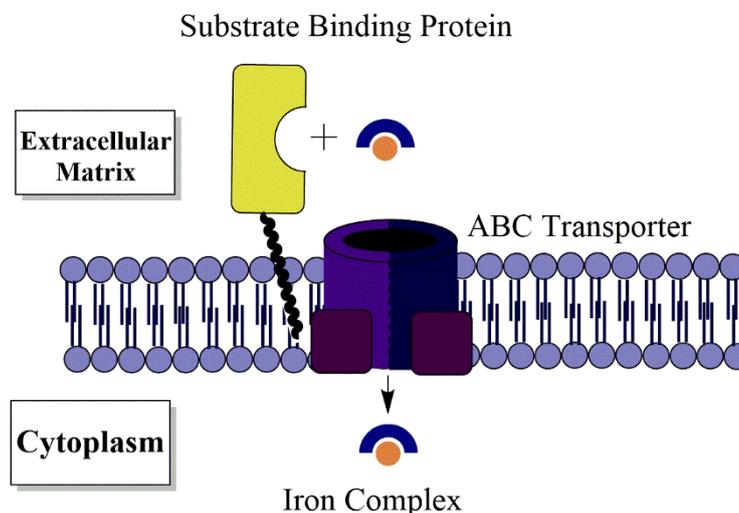


Figure 10: General mechanism for iron(III)-uptake in Gram-positive bacteria.²²

Iron Uptake in Gram-Negative Bacteria

Gram-negative bacteria acquire iron from a number of sources. Iron(II), a soluble ion in anaerobic media can be taken up into the periplasm through porins.¹¹⁸ Active transport by a specific uptake channel then allows the passage of iron(II) into the cytoplasm. The most widespread pathway for iron(II) uptake is the Feo system, and is important for virulence of many species including *E. coli*, *C. jejuni* and *V. cholerae*.¹¹⁸ In aerobic environments,

iron(III) predominates, and must be obtained from soluble sources such as haem and siderophore complexes. A general overview for siderophore and haem derived iron uptake in Gram-negative bacteria is illustrated in Figure 11.

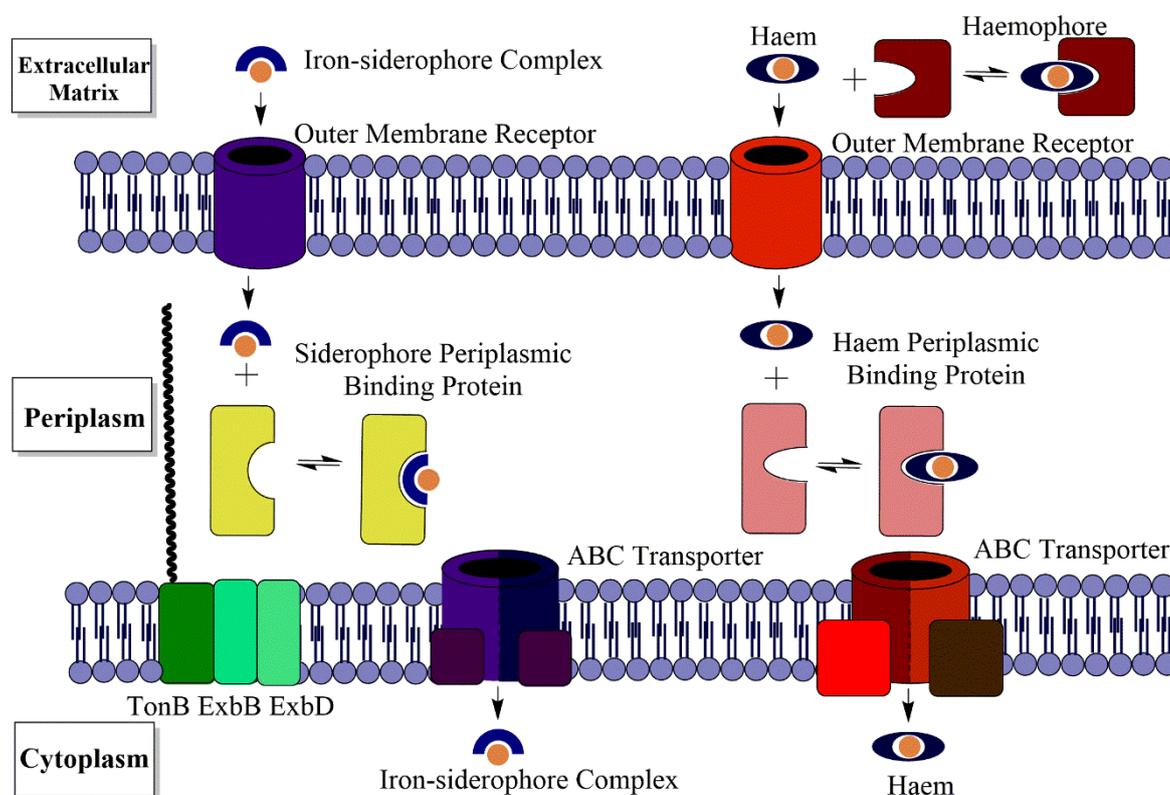


Figure 11: General iron(III)-uptake mechanisms for Gram-negative bacteria.²⁰

Iron(III) complexes are generally too large to be taken up through porins, which have a molecular weight cut-off at around 600 Da, so must be taken up by a specific outer membrane receptor (Figure 11).¹¹⁹ Outer membrane receptors commonly recognise and transport a range of related iron sources,^{19, 20} however Gram-negative bacteria are unable to employ direct active transport of substrates through outer membrane receptors, as there is no transmembrane ATPase for generation of the necessary energy.²⁰ As such, outer membrane receptors are coupled to the inner membrane *via* the inner-membrane anchored proteins TonB, ExbB and ExbD, providing energy generated by the proton motive force of the cytoplasmic membrane to actively transport substrates into the cytoplasm.^{20, 117, 120} Outer membrane receptors, known to uptake catecholate, hydroxamate and carboxylate siderophores, as well as a wide range of other substrates, generally have a 22 stranded β -barrel structure, with a cork domain at the centre of the barrel. The cork domain is involved in the regulation of substrate uptake, acting as a gate between the extracellular matrix and the periplasm.²⁰ The crystal structure of FpvA, the siderophore receptor from *Pseudomonas aeruginosa*, bound to iron(III)-pyoverdine is a good example of this structure (Figure 12).¹²¹

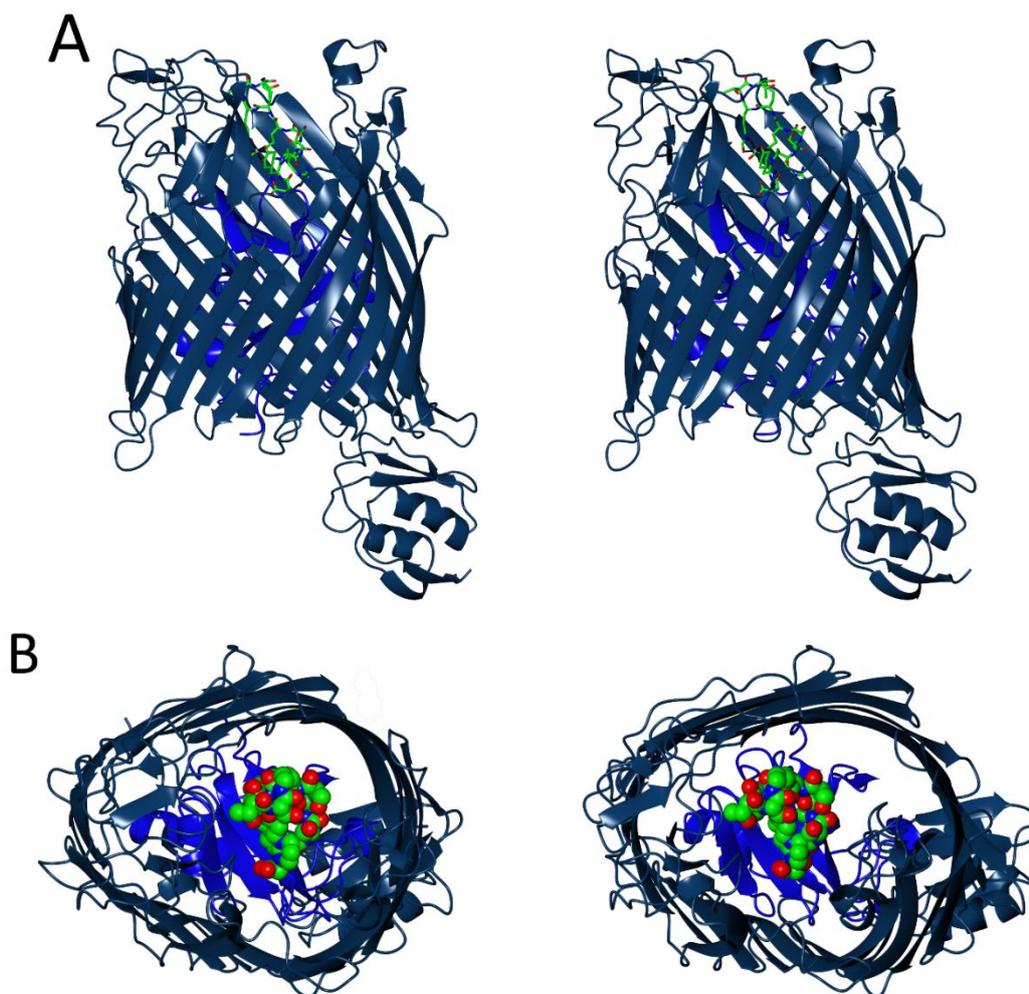


Figure 12: Stereo view of siderophore outer membrane receptor FpvA bound to iron(III)-pyoverdine in *Pseudomonas aeruginosa* (PDB ID: 2IAH). β -barrel is in dark blue with the cork domain in light blue. A: Side view with iron(III)-pyoverdine shown in cylinders coloured by atom type. B: Transmembrane view with iron(III)-pyoverdine shown in spheres coloured by atom type. Figure produced using CCP4mg.

Once the iron(III)-complex is transported into the periplasm, a periplasmic binding protein (with structural similarities to the Gram-positive ‘substrate binding protein’) acts to carry the iron source to the inner membrane ATP-binding cassette transporter, where the complex is delivered into the cytoplasm (Figure 11).²⁰ Siderophore periplasmic binding proteins generally have a bilobate structure, with a shallow binding cleft between the two main domains where the iron(III)-siderophore complex is bound.²⁰ This is demonstrated in the crystal structure of FhuD, a periplasmic binding protein from *E. coli*, bound to iron(III)-coprogen (Figure 13). The iron(III)-coprogen is bound in the region between the two lobes of the protein.¹²² There are currently no known crystal structures for a siderophore inner-membrane ABC transporter.

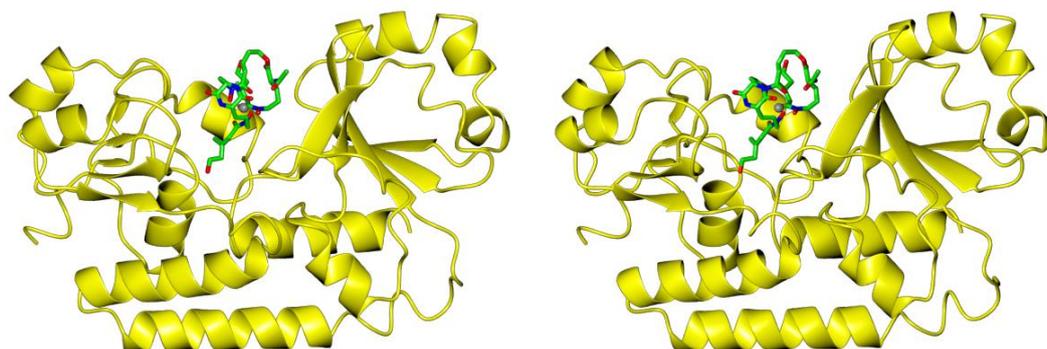


Figure 13: Stereo view of periplasmic binding protein FhuD (yellow ribbons) from *E. coli* bound to siderophore iron(III)-coprogen (cylinders coloured by atom type) (PDB ID: 1ESZ).¹²² Figure produced using CCP4mg.

For haem uptake, haem must first be sequestered from a mammalian host, and this can be done directly *via* binding of the bacterial outer membrane receptor to a host haemoprotein,¹²³⁻¹²⁵ or *via* the synthesis of additional haemophore proteins that are produced and secreted into the extracellular matrix, which then act to scavenge haem and deliver it to the outer membrane receptor^{20, 25, 123, 126}

Once iron(III)-complexes are transported into the cytoplasm, the iron(III) must be released from the complex for storage or use, which can be achieved by degradation of the organic ligand, and/or reduction of the iron(III) to iron(II). Although it was once thought that siderophore complexes would be too stable for iron(III) release by biological reductants,¹⁸ reductase enzymes have been identified for the reduction of iron(III)-siderophore complexes in a number of Gram-negative species, using NADH or NADPH as the reductant.²⁵ These reductases may be specific to iron(III)-siderophore reduction, but general flavin reductases are also reported to reduce iron(III)-siderophore complexes.³²

Bacterial Iron Regulation *via* Fur and DxtR

As discussed in section 1.3 it is important for bacteria to maintain a strict iron concentration within their cells to ensure necessary cell function without generation of damaging reactive oxygen species. In most Gram-negative bacterial species, and a large proportion of Gram-positive species with low genetic guanine-cytosine content, iron levels are controlled by the ferric uptake regulator (Fur).¹²⁷ When genetic guanine-cytosine content is high in Gram-positive species, iron levels are often regulated in a similar manner by the diphtheria toxin regulator (DtxR).³⁰ In the Fur system, the 17 kDa ferric uptake regulator protein (Figure 14) acts as a sensor for iron, and represses the expression of genes that code for the proteins required in siderophore synthesis and those involved in iron-uptake.^{21, 128}

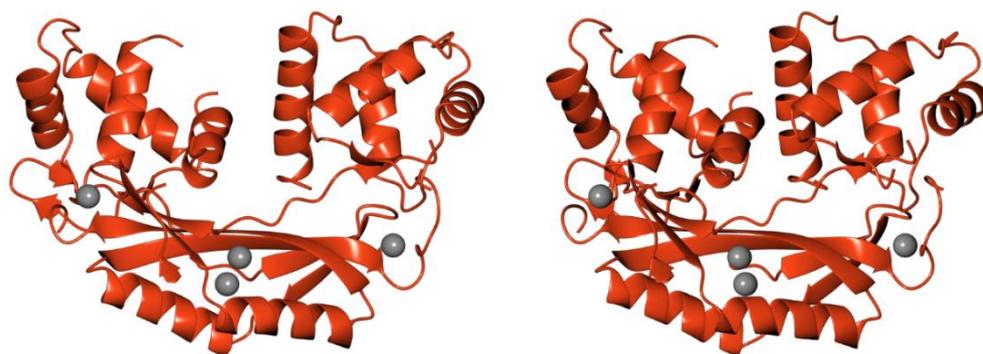


Figure 14: Stereo view of the crystal structure of the Fur regulator protein from *C. jejuni*. (Zn²⁺ ions shown as grey spheres) (PDB ID: 4ETS).¹²⁹ Figure produced using CCP4mg.

This means that when iron concentrations are high, Fur is iron(II)-bound, and represses the upregulation of iron uptake machinery, limiting further iron-uptake into the cell. When cellular iron concentrations decrease, iron is released from the Fur protein, and the genes for expression of iron-uptake machinery are no longer repressed. Iron concentrations can then be replenished as the iron-uptake capacity of the cell is increased.

In many bacteria, including *E. coli*,¹²⁷ *C. jejuni*^{130, 131} and *V. cholerae*,¹³² there is a *fur* gene and a 'fur box' region, with the consensus sequence GATAATGATAATCATTATC located in the promoter sequence of a gene or operon that is iron-regulated. When the Fur protein is iron-loaded, a conformational change is induced, allowing the binding of two dimers of the protein to the Fur Box sequence, preventing access to RNA polymerase.¹³³ Fur has also been shown to act as a transcriptional activator for the genes responsible for essential iron-containing enzymes, and is involved in the regulation of the bacterial ferritin iron-storage proteins. This proves that Fur acts as a global iron regulator protein, prioritising only vital iron consumption when under deficiency.¹³³ In addition to Fur, efflux pumps are known to exist for the rapid export of iron(III) outside of the cell, to prevent cell death by iron(III) overload and generation of reactive oxygen species. In *S. typhimurium*, an iron citrate efflux transporter was found to export iron(III)-citrate when under stress conditions. These systems are however not regulated by Fur, but a two component regulator thought to be linked to the disruption of normal iron concentrations.²¹

Iron Uptake in *Escherichia coli*

E. coli, the most widely studied bacterial species, is Gram-negative and exists in the gut flora of a wide range of animal species, including mammals. *E. coli* is largely aerobic when in the mammalian gut, where the species is able to protect the host from pathogenic colonisation, but can also survive in water and other secondary environments. Pathogenic strains of *E.*

coli, including Shiga Toxin-producing *E. coli* (STEC), are a significant threat to human health, with more than 2.8 million cases of intra and extra-intestinal *E. coli* infections per year.¹³⁴⁻¹³⁶ Iron uptake is of vital importance to successful host colonisation and *E. coli* employs a range of strategies for iron acquisition, including the use of haem as well as a variety of siderophore complexes.^{20, 126, 137} The widely-studied *E. coli* K-12 strain synthesises and secretes enterobactin as its only siderophore. Pathogenic strains of *E. coli* have been found to synthesise salmochelin S4, aerobactin, yersiniabactin as well as enterobactin, and it is likely that these siderophores are involved in virulence (Figure 15).¹³⁸ The ability to produce and utilise a range of structurally-diverse siderophores allows for iron uptake over a broader pH range, and the use of a range of starting materials for siderophore synthesis means growth of the species is possible with a broader range of nutrients, increasing the chances of survival.

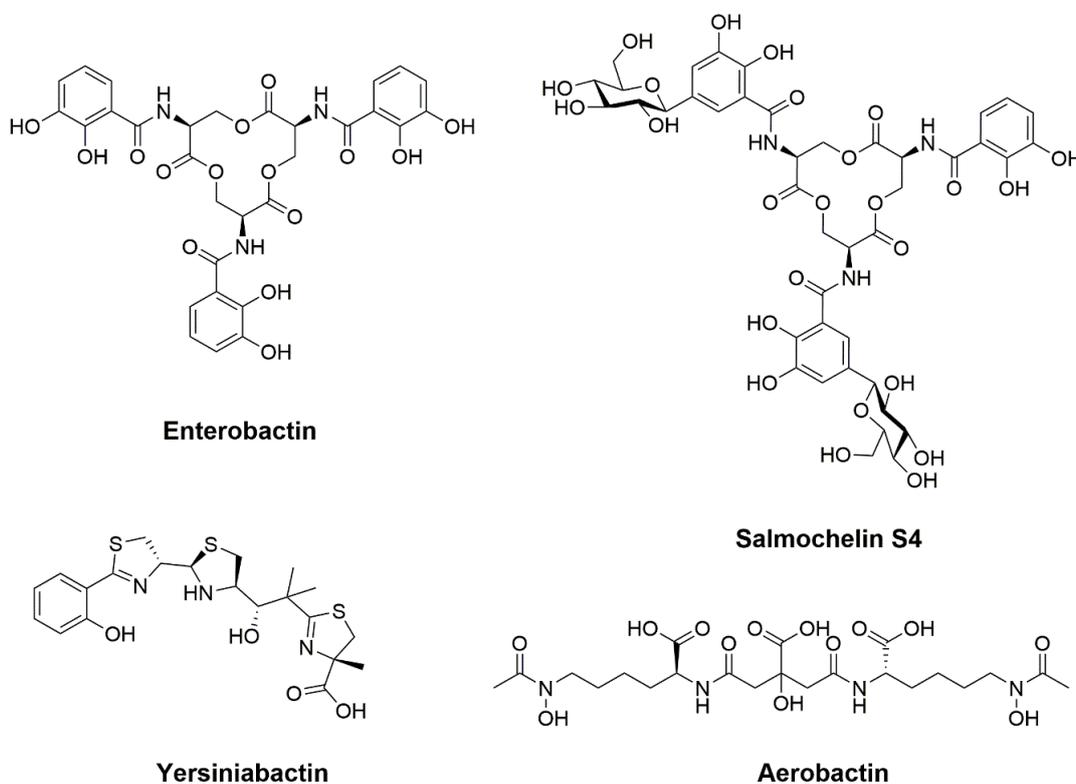


Figure 15: Siderophores synthesised and secreted by pathogenic strains of *E. coli*.¹³⁸

The catecholate siderophores enterobactin and salmochelin S4 are of most interest for this thesis, and their role in *E. coli* will be discussed further. As salmochelin S4 is the bis-C5-glucosylated analogue of enterobactin, the biosynthesis of both siderophores is identical, with a final glucosylation transforming enterobactin into salmochelin S4. Enterobactin biosynthesis is favoured at a lower pH of 5.6, while salmochelin S4 production predominates at pH 7 and above.¹³⁸ The *E. coli* enterobactin biosynthesis and transport genes are held in a

22-kb cluster as follows: *entD-fepA-fes-entF-fepE-fepC-fepG-fepD-fepB-entC-entE-entB-entA-ybdA*. Within the cluster there are 6 operons with three promoter regions and a fur-box.¹³⁹ The genes termed ‘*ent*’ code for six enzymes (EntA-F) that catalyse the production of enterobactin. Catecholate siderophore uptake in *E. coli* is summarised in Figure 16. The ‘*fep*’ genes in the gene cluster code for the proteins FepABCDEG, which are dedicated to the import of ferric enterobactin. FepA is the TonB coupled outer membrane receptor (Figure 17).^{140, 141}

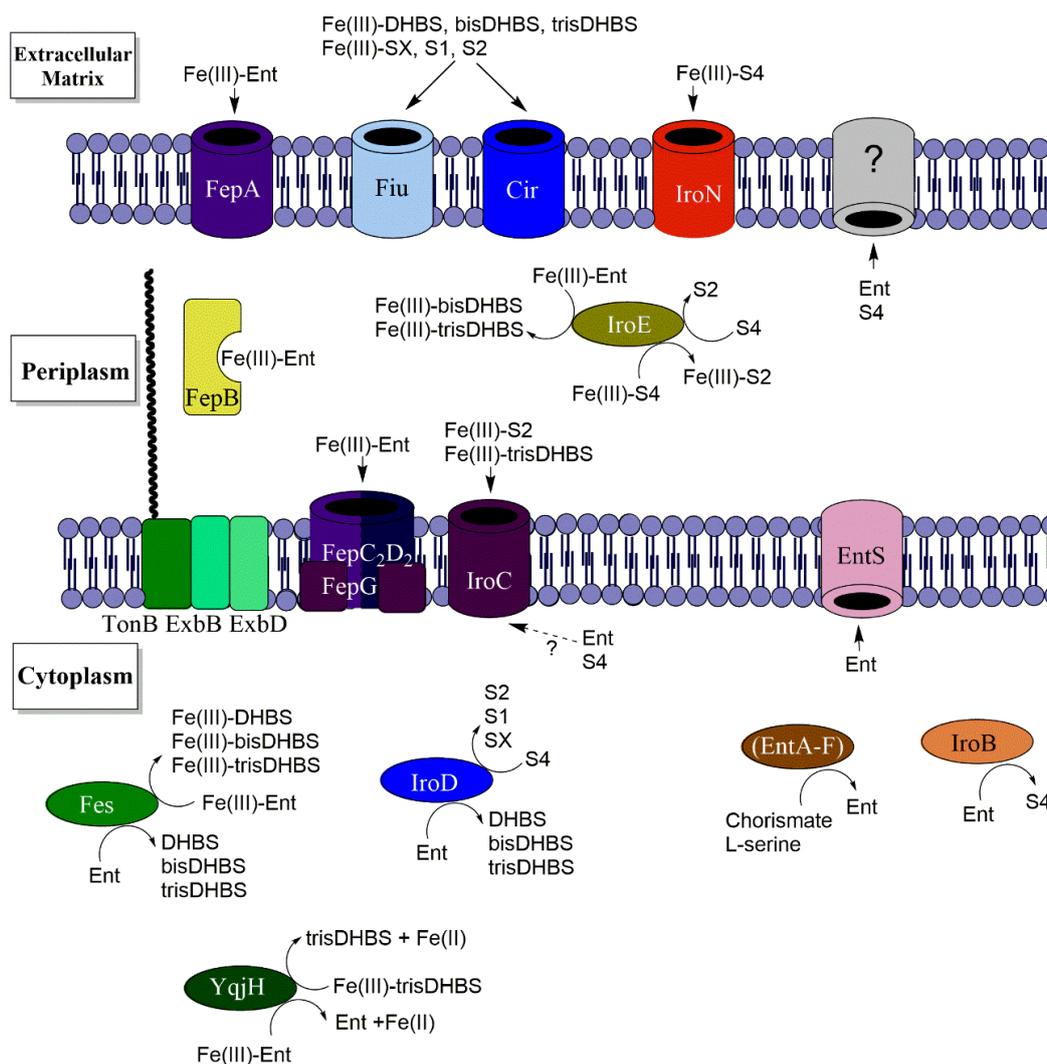


Figure 16: Proposed functions of proteins involved in catecholate siderophore synthesis, secretion and uptake in pathogenic *E. coli*. Ent is an abbreviation for Enterobactin, S4, S2, S1 and SX are abbreviations for each salmochelin.

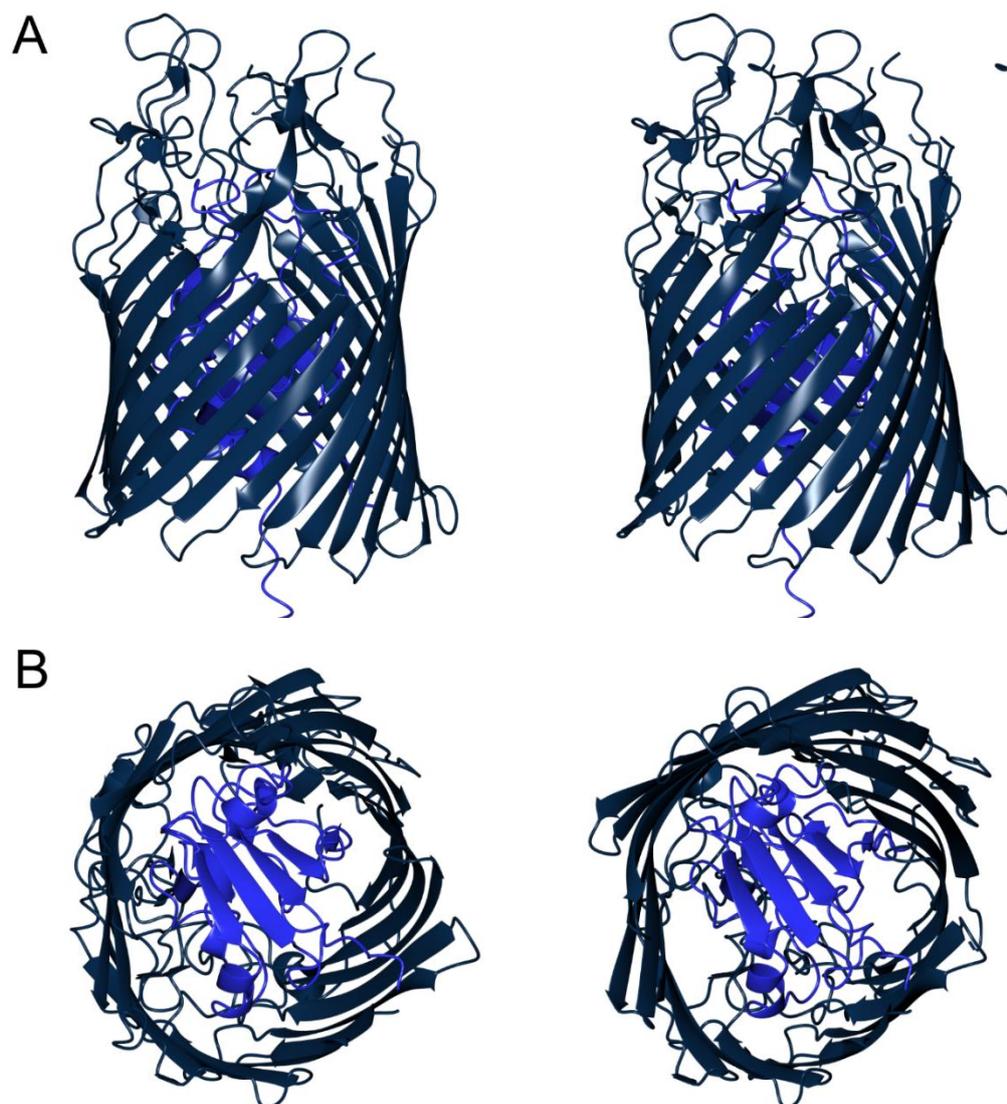


Figure 17: Stereo view of the crystal structure of FepA, the enterobactin outer-membrane receptor in *E. coli*.¹⁴² A: Side view B: Transmembrane view. (PDB ID: 1FEP) Figure produced using CCP4mg.

FepB (Figure 18) is the periplasmic binding protein that delivers iron(III)-enterobactin to the ABC transporter in the cytoplasmic membrane, made up of FepC₂D₂.^{20, 143-146} The function of FepG is less well documented, but is also thought to play a role in uptake at the cytoplasmic membrane.^{3, 20, 147-149} There are no published structures for FepC₂D₂ or FepG. The '*fepE*' gene codes for the protein FepE that is bound to the cytoplasmic membrane, but is in fact an unrelated polysaccharide co-polymerase, involved in the synthesis of polysaccharide chains for cell surface antigens.^{150, 151} The '*fes*' gene codes for the Fes protein, an esterase located in the cytoplasm that causes the hydrolysis of the enterobactin trilactone backbone. The '*ybdA*' gene codes for the EntS protein, which acts as a cytoplasmic enterobactin export protein, and has similarities to efflux pumps of the major facilitator superfamily (Figure 16).^{152, 153}

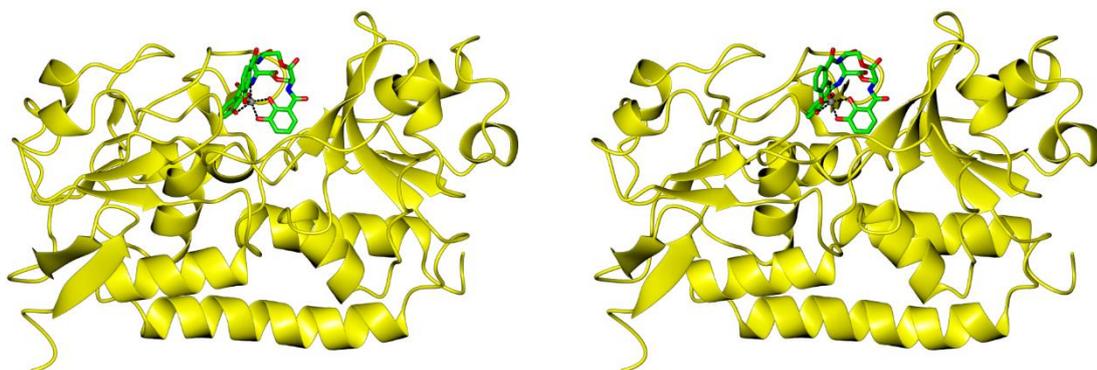
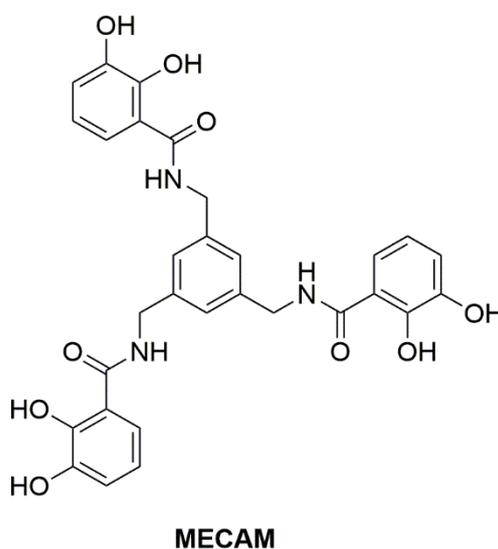


Figure 18: Stereo view of FepB, the enterobactin periplasmic binding protein in complex with iron (III)-enterobactin (PDB ID: 3TLK).¹⁴⁶ Figure produced using CCP4mg.

Salmochelin S4 is synthesised *via* the transfer of a glucose functional group onto the C5 position of two of the catecholamide units of enterobactin *via* the *C*-glucosyltransferase IroB.¹⁵⁴ IroB is derived from the *iroNEDCB* gene cluster. The cluster also contains the genes for a further four proteins involved in siderophore uptake. The function of IroC is disputed as it is commonly reported as a cytoplasmic salmochelin exporter.^{153, 155, 156} A contrasting study suggests that IroC may be the cytoplasmic iron(III)-salmochelin S2 transporter, which is also reported to uptake iron(III)-enterobactin linear hydrolysis product iron(III)-trisDHBS.¹⁵⁷ The confusion is likely to arise from the same naming system being used for similar proteins from *S. typhimurium*.¹⁵⁸ IroC may have a different function in this species when compared to the IroC in *E. coli*, and further studies are required to establish whether this protein really possesses a dual function as both an exporter and importer.^{155, 156} IroD is a cytoplasmic esterase known to break down both salmochelin S4 and enterobactin to their linear hydrolysis products. IroE is a periplasmic esterase that linearises both salmochelin S4 and iron(III)-salmochelin S4 to salmochelin S2. IroE has also been shown capable of the slow linearisation of both enterobactin and iron(III)-enterobactin to trisDHBS and bisDHBS.¹⁵⁹ Finally, IroN is the outer membrane iron(III)-salmochelin S4 importer.^{154, 157} Interestingly there is no known periplasmic binding protein associated with the Iro salmochelin uptake system, meaning there are no known interactions of a periplasmic binding protein with the salmochelins, or indeed the enterobactin hydrolysis products.

Fiu and Cir are outer membrane receptors employed in the uptake of linear hydrolysis products of enterobactin,¹⁵² and it is known that trisDHBS, bisDHBS and DHBS can transport iron into *E. coli*.^{95, 160} The FepBDGC system is proposed to be involved in the transport of the linear enterobactin hydrolysis products, but this requires more rigorous study.^{160, 161} It is possible that IroC may be the transporter for linear enterobactin hydrolysis

products.¹⁵⁷ Once inside the cytoplasm, iron is released from enterobactin *via* hydrolysis by the esterases Fes and IroD to trisDHBS, then bisDHBS and finally to the bidentate DHBS. These hydrolysis events are much more efficient than those of periplasmic IroE.¹⁵⁹ Iron(II) is then released from enterobactin and the lower denticity hydrolysis products by reduction *via* the NADPH dependent reductase YqjH.¹⁶¹⁻¹⁶³ YqjH is *fur* regulated, but is also linked to nickel homeostasis.¹⁶⁴ The reductase reduces iron(III) bound to enterobactin to iron(II), but works most efficiently on iron(III)-trisDHBS, after prior hydrolysis of the cyclic trilactone backbone by Fes.¹⁶³ YqjH can reduce iron(III)-citrate, but is not active on hydroxamate siderophores.¹⁶³ The growth of *E. coli* with iron(III)-MECAM is directly dependent on the action of YqjH.¹⁶⁵



Iron Uptake in *Campylobacter jejuni*

Campylobacter jejuni is a Gram-negative species responsible for gastroenteritis in humans, causing more cases of infection than the sum of both *Salmonella spp.* and *E. coli* and is commonly transmitted through under-cooked poultry.¹⁶⁶⁻¹⁶⁹ It is thought to be particularly good at co-operative survival with other species,¹⁷⁰ and requires a high degree of metabolic flexibility to enable survival in a number of environments upon migration from avian host, to pathogenic colonisation of the human intestinal tract.^{168, 171}

C. jejuni possesses an array of iron(III) uptake machinery to ensure the necessary iron(III) concentration inside the cell for survival and growth,¹³¹ and is able to acquire iron(III) from its environment *via* a number of known sources, including catecholate siderophore complexes, rhodotorulic acid complexes, haem, transferrin and lactoferrin.^{168,172, 173} *C. jejuni* is a scavenger species, able to survive without the ability to produce its own siderophores.^{115,}

¹⁷⁴ Although some strains are thought to produce siderophores in small amounts, the siderophore type has never been classified.^{175, 176}

The full genome of *C. jejuni* has been sequenced, but the complete iron transport system is not yet fully characterised, with several proteins of currently unknown function likely to be involved in iron transport.^{115, 177} Six gene clusters were identified as iron regulated, indicating that these genes may code for proteins involved in iron homeostasis.¹¹⁵ *C. jejuni* possesses typical Gram-negative transport machinery which is regulated by *fur* in a similar manner as discussed for *E. coli*.¹³¹ *Fur* regulation has shown to be involved in prevention of oxidative stress and for allowing the species to survive in acidic environments.¹³⁰ Iron(II)-uptake is known in *C. jejuni* by the Feo system, with FeoB, the outer membrane receptor essential for iron(II) uptake.^{118, 168} The periplasmic binding protein cFbpA is known to bind iron directly, and is selective for iron(II).¹⁷⁸ ChuABCD are *fur* regulated and are involved in the uptake of hemin, and haem from haemoglobin.^{115, 173} The outer-membrane receptor is ChuA, ChuB and ChuC are the permease and ABC transporter components for uptake of haem into the cytoplasm, and ChuD acts as the periplasmic binding protein.¹⁷² Similar systems exist for uptake of iron(III) *via* ferrichrome, rhodotorulic acid and transferrins.¹⁷²

All known strains of *C. jejuni* can employ exogenous enterobactin for iron uptake.^{115, 175} For catecholate enterobactin-derived siderophore complexes, the uptake machinery consists of the outer membrane transporters CfrA and CfrB (structures currently unknown),¹⁷⁹ with three TonB systems which provide active transport of complexes into the periplasm. CeuBCDE (Campylobacter enterobactin uptake) proteins transport enterobactin-derived complexes into the cytoplasm (Figure 19).¹⁷²

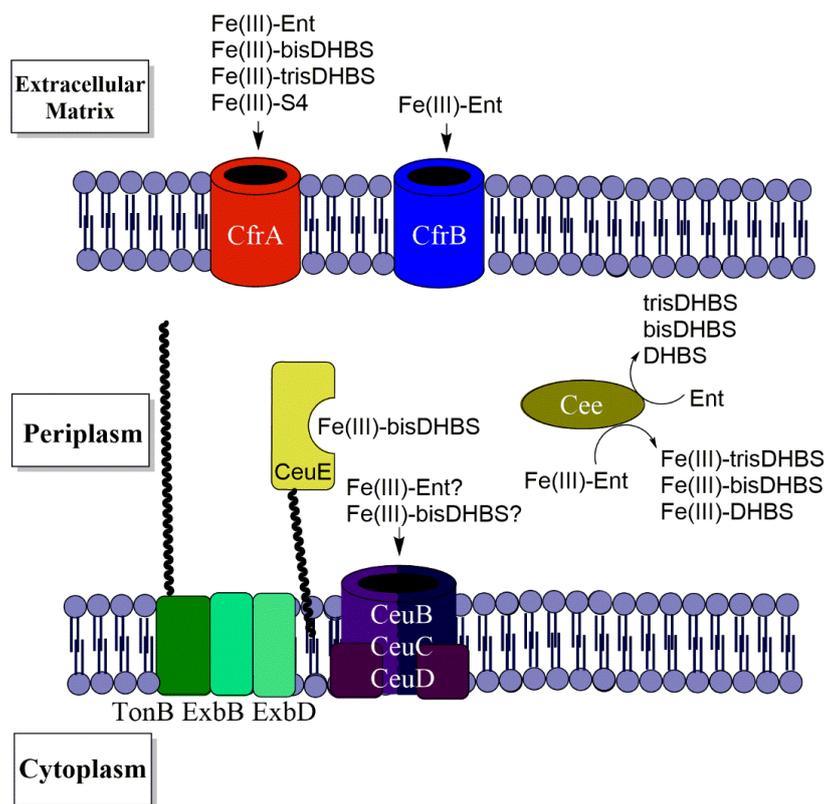


Figure 19: Proposed functions of proteins involved in catechol siderophore uptake in *C. jejuni*.^{115, 173, 180}

CeuBCD share high homology with the iron(III)-enterobactin ABC transporter in *C. coli*.¹¹⁵ CeuB and CeuC are therefore likely to be cytoplasmic membrane permeases, and CeuD, the ATPase that aids the necessary active transport of complexes into the cytoplasm.^{181 172} Although CeuBCD are known to be involved in cytoplasmic uptake of enterobactin derived iron(III)-complexes,¹⁷² it has not been established whether these are complexes of the full cyclic trimer or the hydrolysis products, and there is currently no published crystal structure. CfrA and TonB3 were shown to be required for use of enterobactin, bisDHBS, trisDHBS and salmochelin S4 as iron sources, indicating that CfrA is likely to be involved in the uptake of multiple catechol substrates.¹⁸² This is currently the only study showing that an iron(III)-salmochelin siderophore supports growth of *C. jejuni*.¹⁸² CfrB, a second iron(III)-enterobactin outer membrane receptor, is present in fewer *C. jejuni* strains than CfrA, and more commonly only one of the two outer-membrane receptors is present in each strain, unlike in *C. coli* where both are present.¹⁸³ Strains that contained only CfrB could not always use iron(III)-enterobactin as an iron source.¹⁸³

CeuE (Figure 20), is a periplasmic lipoprotein, characterised as having siderophore-binding properties in 1995.¹⁷¹ Crystal structures revealed it has a typical overall fold of a type III (Cluster A-II) periplasmic binding protein.^{184, 185} The fold consists of two α -helix/ β -sheet

containing lobes linked *via* an α -helix.¹⁸⁶ There is a shallow iron(III)-siderophore binding cleft located between the two lobe regions (Figure 20).^{186, 187} The protein is naturally located in the periplasm, and acts to deliver catecholate iron(III)-siderophore complexes from the outer membrane receptor CfrA to the inner membrane transporter CeuBCD, for uptake of the complexes into the cytoplasm.³⁰ CeuE has been proven not to be essential for iron acquisition from iron(III)-enterobactin, but without CeuE, growth is significantly impaired.^{115, 159} It has been reported that CeuE along with the cytoplasmic transporter made up of CeuBCD are involved in the transport of iron(III)-enterobactin into the cytoplasm.¹⁶⁸ However more recent studies have shown that iron(III)-bisDHBS is a more optimised substrate for CeuE than iron(III)-enterobactin, suggesting that the linear hydrolysis products may be the main source of delivery of cytoplasmic iron(III).⁹⁰ It is likely that confusion arises over the use of impure enterobactin sources, or the lack of consideration of the potential for *in situ* enterobactin hydrolysis.

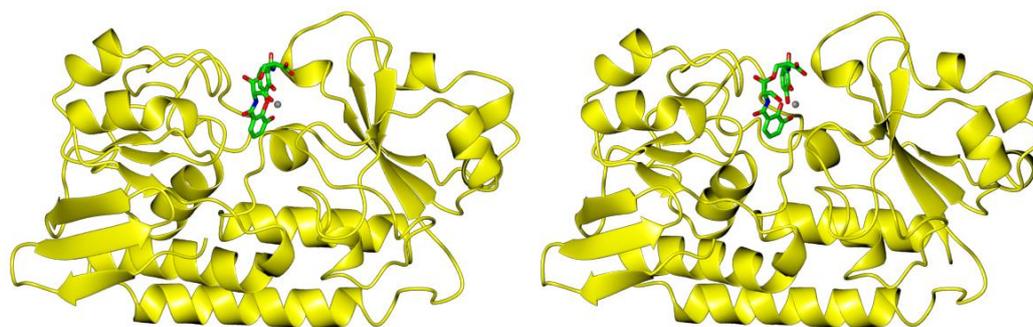


Figure 20: Stereo view of CeuE (yellow ribbons) (PDB 1D: 5ADW) has bilobate shape with a shallow binding cleft where iron(III)-bisDHBS is bound (cylinders coloured by atom type).^{186 90} Figure produced using CCP4mg.

It is known that *C. jejuni* is able to use enterobactin as a sole iron source without a functional CeuBCDE ABC transporter system if Cee is present.¹⁸² Cee is a periplasmic esterase in *C. jejuni* known to degrade iron(III)-enterobactin to iron(III)-trisDHBS, iron(III)-bisDHBS and iron(III)-DHBS, as well as the apo siderophore to trisDHBS, bisDHBS and DHBS.¹⁵⁹ Cee was shown to be essential for iron(III)-enterobactin utilisation with CfrB, however Cee was not required when CfrA was present. As CeuE is optimised for binding of iron(III)-bisDHBS, this may indicate that CfrB can only transport cyclic iron(III)-enterobactin into the periplasm, whereas CfrA is known to transport the linear hydrolysis products as well as cyclic enterobactin.¹⁸² If this is the case, a system with CfrB but no Cee would cause cyclic enterobactin to accumulate in the periplasm, without an esterase capable of hydrolysis of the backbone for uptake of the hydrolysis products into the cytoplasm by CeuBCDE. As Cee

appears to be the only trilactone esterase capable of hydrolysing enterobactin in *C. jejuni*, this supports the theory that the major function of the CeuBCDE system is to transport the linear hydrolysis products into the cytoplasm.¹⁵⁹ The current scheme for catecholate siderophore iron uptake in *C. jejuni* is shown in Figure 19.

The iron(III)-linear hydrolysis products are likely to have higher redox potentials than iron(III)-enterobactin, meaning they require less powerful reducing agents for release of iron(II) from the complex.^{30, 56, 95} It is thus more likely that iron(II) would be released from the linear hydrolysis products with greater efficiency by a cytoplasmic reductase than from cyclic enterobactin. No cytoplasmic iron(III)-siderophore reductase has been characterised, however riboflavin production is *fur* regulated and has been linked to iron(III) to iron(II) reduction capacity, suggesting that a potential reductase is likely to contain a riboflavin derived cofactor such as FAD.¹⁸⁸ This suggests that a reductase with a wide range of functions may act on the linear hydrolysis products of enterobactin.

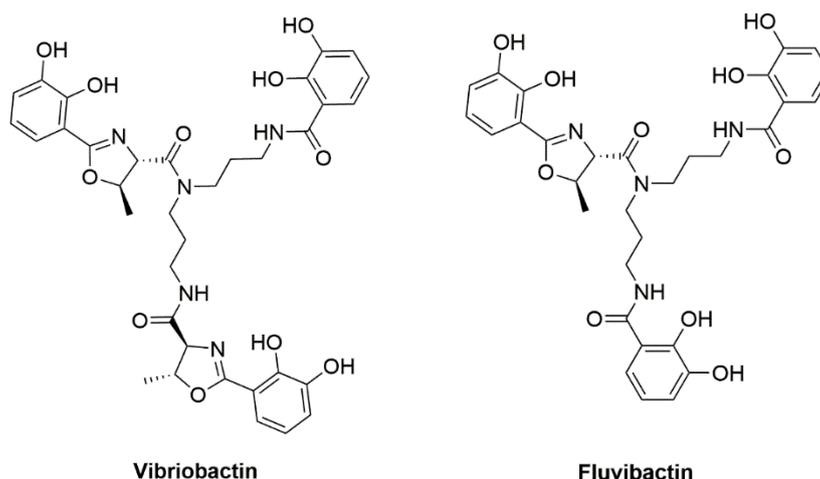
Remaining unanswered questions in the catecholate siderophore uptake of *C. jejuni* include the study of the CeuBCD complex to establish which siderophore complexes are taken up into the cytoplasm, as the fact CeuE binds tightly to iron(III)-bisDHBS and not iron(III)-enterobactin raises the question as to whether CeuBCD is also optimised for the iron(III)-bound hydrolysis products as suggested by studies with Cee and the lack of an iron(III)-enterobactin reductase.¹⁵⁹ Direct protein-siderophore binding studies are required in conjunction with growth studies to confirm this theory. Additionally, since CfrA is able to take up iron(III)-salmochelin S4, and supports growth of the species, it is of interest to study the role of this siderophore in *C. jejuni*, and establish the pathway these siderophores take for release of iron(III) into the cytoplasm.¹⁸²

Iron Uptake in *Vibrio cholerae*

Vibrio cholerae is the Gram-negative pathogen that causes cholera in humans, and is estimated to cause at least 120,000 annual deaths worldwide, with the greatest prevalence in developing countries in Asia and Africa.^{189, 190} Infections usually arise from the consumption of contaminated water, where the species is often found associated with zooplankton.¹⁹¹ Once the species has colonised the latter portion of the small intestine,¹⁹² the production of cholera toxin causes severe diarrhoea, meaning rapid rehydration therapy is required alongside antibiotic treatment.¹⁹⁰ *V. cholerae* is able to form biofilms, as well as being able to switch to a viable but non-culturable state, enhancing the ability of the species to survive in otherwise inhospitable environments.¹⁹¹

Iron regulation has been linked to the virulence of *V. cholerae*,¹⁹⁰⁻¹⁹³ and like *E. coli* and *C. jejuni*, *V. cholerae* possesses a *fur* system, with a conserved Fur box sequence that controls iron homeostasis, and can acquire iron from a number of sources including haem and a range of siderophore complexes.¹³² *V. cholerae* has a *fur* regulated Feo system for the uptake of iron(II), comprising the FeoA, FeoB and FeoC transporter proteins.^{132, 194} Under iron limited conditions *V. cholerae* produces haemolysin to lyse host cells and then uptake haem compounds liberated from the cellular matrix.³ The haem uptake system comprises of HutA, the outer membrane receptor, HutB the periplasmic binding protein, HutCD, the inner membrane ABC transporter, and HutZ, a protein thought to be involved in haem shuttling or storage, more recently suggested to be a haem degradation enzyme.^{132, 195-197}

V. cholerae employs two catechol siderophores, enterobactin (and enterobactin derivatives) and vibriobactin. *V. cholerae* does not produce enterobactin but can poach enterobactin and linear derivatives from the surrounding environment.



Vibriobactin is however synthesised by the species, in a very similar manner to the synthesis of enterobactin by *E. coli*. Once synthesised, vibriobactin is known to be exported from the cytoplasm by the *fur* regulated resistance-nodulation-cell division transporter VexGH, however there is currently no known outer membrane exporter for the siderophore.¹⁹⁸

There is crossover between the uptake machinery for iron(III)-vibriobactin and iron(III)-enterobactin in *V. cholerae* so the two must be described together. The study of the use of iron(III)-enterobactin in *V. cholerae* has suffered from the lack of consideration of the presence of potential hydrolysis products such as trisDHBS, bisDHBS and DHBS. Growth studies involving the use of enterobactin or even crude *E. coli* extracts have often assumed that enterobactin was the only siderophore species present without additional characterisation.^{192, 199-201} The use of enterobactin hydrolysis products rather than the full

cyclic trilactone for iron(III) uptake by *V. cholerae* was suggested in 1984, when growth of the species on enterobactin was delayed.²⁰² The most recent studies have considered linear enterobactin derivatives, and the possible presence of these derivatives can be used to explain previously conflicting results.¹⁶¹ It is likely that enterobactin derivative utilisation is important when in aquatic environments where *E. coli* is abundant, and less so in the host intestine as *E. coli* predominantly resides in the colon.¹⁶¹ *Fur* regulated systems for the uptake of iron(III)-siderophores enterobactin and vibriobactin include VctAPDGC (Vibrio catechol transport APDGC) and ViuAPDGC (Vibrio iron uptake APDGC) (Figure 21).²⁰⁰

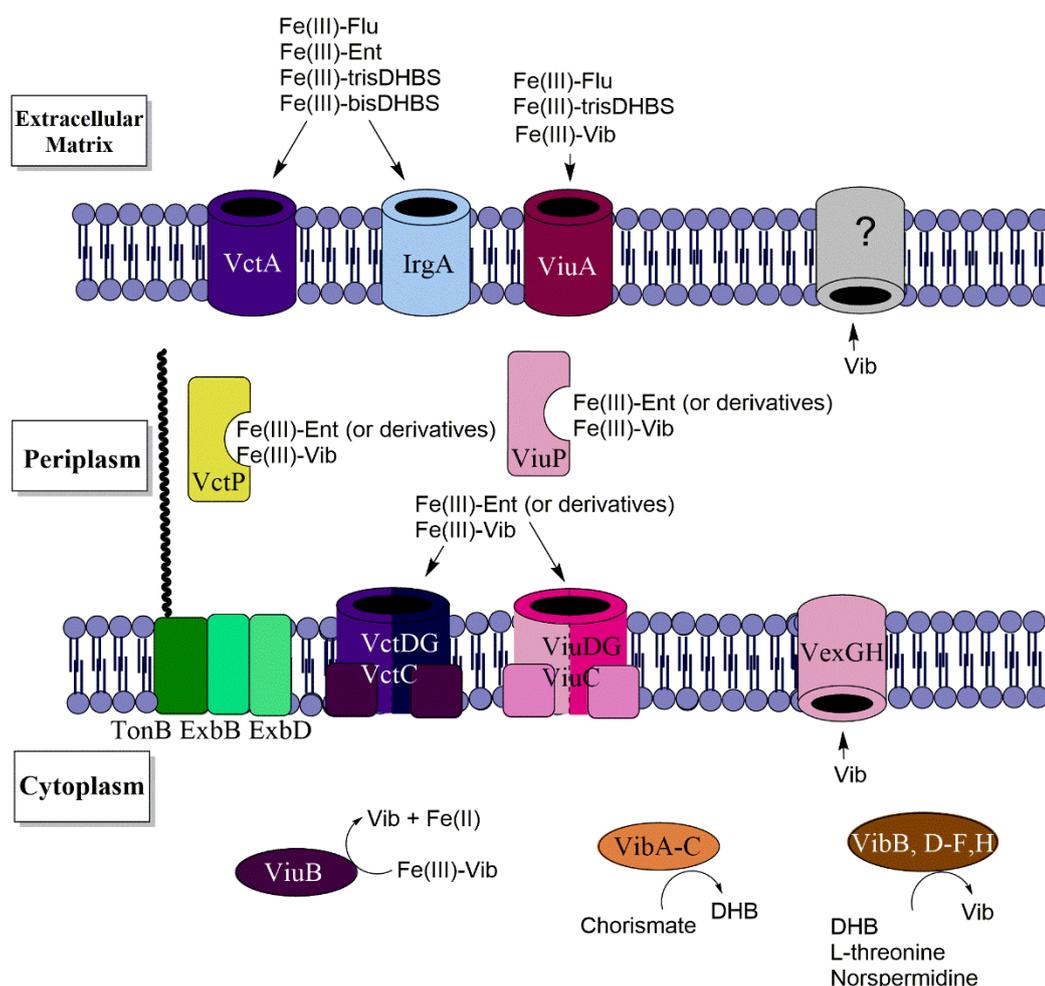


Figure 21: Known functions of proteins involved in catechol siderophore synthesis and uptake in *V. cholerae*.

V. cholerae produces two TonB systems that provide the energy for transport of iron(III)-siderophore complexes across the outer membrane for the two outer membrane receptors VctA and IrgA (iron regulated gene A) for linear enterobactin derivatives,²⁰³ and the receptor for vibriobactin ViuA.^{192, 204, 205} It was first thought that VctA and IrgA were receptors for iron(III)-enterobactin, but it has since been shown that these receptors are optimised for uptake of hydrolysis products iron(III)-trisDHBS and iron(III)-bisDHBS.¹⁶¹ In addition

ViuA may also be capable of uptake of iron(III)-trisDHBS. All three receptors can also be used in the uptake of iron(III)-fluoribactin, a siderophore similar to vibriobactin that could be poached from *Vibrio fluvialis*.^{161, 204} It remains unclear whether cyclic enterobactin can be used to support growth of *V. cholerae* at all, but it is now accepted that iron(III)-trisDHBS and iron(III)-bisDHBS provide the most efficient iron(III) delivery.^{161, 202, 206}

Once inside the periplasm, the periplasmic binding proteins and ABC transporters for uptake of iron(III)-vibriobactin and linear derivatives of iron(III)-enterobactin into the cytoplasm are less specific than outer membrane receptors. Both VctPDGC and ViuPDGC have been shown to transport iron(III)-vibriobactin and iron(III)-enterobactin derivatives as well as the hexadentate catecholate siderophore mimic MECAM into the cytoplasm (Figure 21).^{161, 192, 200, 207} VctPDGC transplanted into *E. coli* deficient in FepB was said to be able transport iron(III)-enterobactin into the cytoplasm, but with the FepC₂D₂-FepG ABC transporter intact, it is difficult to establish whether this was the case, or whether the FepC₂D₂-FepG was simply operating without its periplasmic binding protein.¹⁶¹ As a result, it is unclear which enterobactin-derived species are transported by VctPDGC, but it would be logical to suggest that VctPDGC may be involved in uptake of iron(III)-trisDHBS and iron(III)-bisDHBS if these species are more readily transported into the periplasm than iron(III)-enterobactin.¹⁶¹ Both VctP and ViuP have been crystallised and their structures solved: VctP in apo form, and ViuP with vibriobactin bound (Figure 22).^{208, 209}

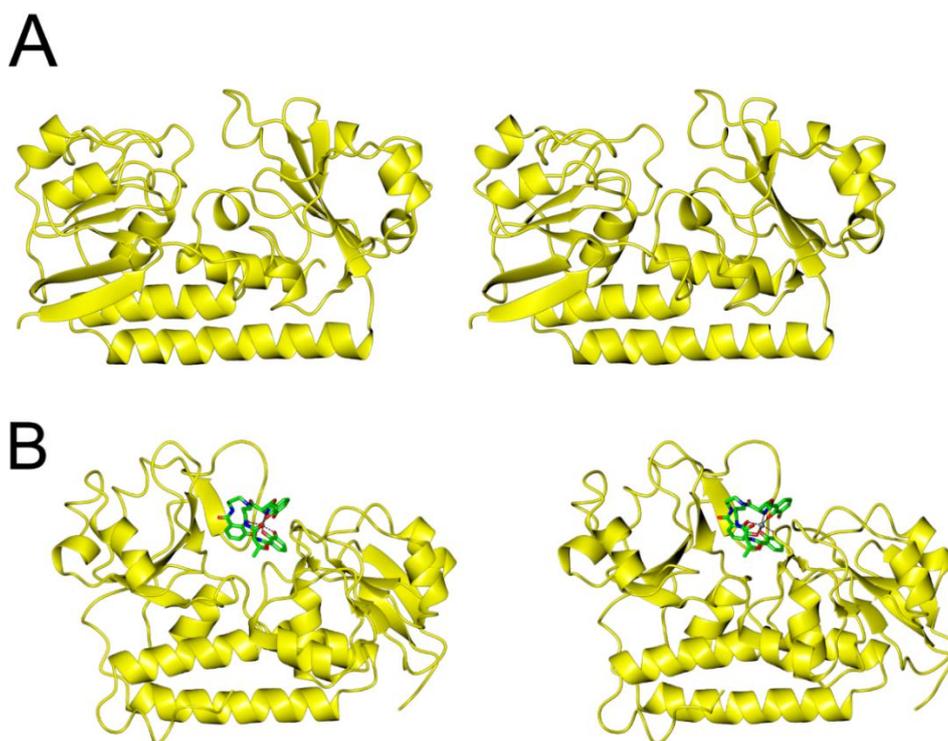


Figure 22: Stereo view of catechol siderophore periplasmic binding proteins VctP (A) (PDB ID: 3TEF) and ViuP bound to iron(III)-vibriobactin (B) (PDB ID:3R5T). Protein shown in yellow ribbons with vibriobactin in cylinders coloured by atom type.^{208, 209} Figure produced using CCP4mg.

ViuB, with homology to the *E. coli* cytoplasmic reductase YqjH acts as a reductase for iron release from vibriobactin in the cytoplasm of *V. cholerae*. ViuB is needed for iron release from vibriobactin but not for enterobactin, enterobactin derivatives or MECAM.¹⁹² *V. cholerae* can use MECAM and linear enterobactin derivatives for growth, which suggests there is likely to be another uncharacterised reductase present that allows iron release from these siderophores and siderophore mimic.¹⁶¹ There is no known cytoplasmic enterobactin esterase in *V. cholerae* that has homology with the cytoplasmic esterase Fes in *E. coli*.²⁰⁴

Siderophore-free uptake of iron(III) has been reported for *V. cholerae*. The studies suggest that VctP may be able to bind iron(III) without the presence of a catechol siderophore. However, the specific interaction was not studied directly, but relied on growth studies, so it is not clear whether the system was truly siderophore-free, or whether an alternative uncharacterised iron-uptake pathway was present.²⁰¹ Direct iron(III)-binding studies with VctP are required to establish whether VctP can indeed bind siderophore-free iron(III). FbpA is reported as the periplasmic binding protein for FbpBC, a ferric iron uptake ABC transporter in the cytoplasmic membrane that is able to uptake iron(III) directly without the

aid of siderophores.^{199-201, 204} In addition, there may be some remaining undiscovered iron uptake proteins in *V. cholerae*.²⁰⁰

Areas still to be addressed in the catecholate siderophores use in iron uptake in *V. cholerae* include the examination of the exact role of VctPDGC and ViuPDGC in the uptake of enterobactin and linear derivatives. To confirm the action of these systems in greater detail, direct protein-siderophore interactions must be studied. A new area of study would involve the potential role of salmochelin siderophores in *V. cholerae* iron uptake. It would be interesting to study whether the Vct or Viu systems can use this siderophore class for iron uptake and growth.

1.6 Iron Regulation in the Mammalian Host and Stealth Siderophores

As bacterial species *E. coli*, *C. jejuni* and *V. cholerae* can colonise the mammalian gut and act as pathogens, it is important to understand the iron availability for these bacteria in this host environment. Macrophages are vital for systemic and cellular iron homeostasis, and limit the iron available to pathogens- known as nutritional immunity.^{1, 210, 211} *E. coli* is known to produce siderophores in the mammalian gut.^{32, 212} Iron regulatory proteins (IRP) produced by host macrophages are essential for preventing colonisation of the gut with such pathogenic species.²¹¹ The proteins cause interference with bacterial iron uptake machinery including the iron(II) uptake systems, and the siderophore production machinery, as well as inducing the production of siderocalins, and limiting the amount of available ferritin: the mammalian iron storage protein.²¹¹

Lipophilic siderophores such as enterobactin are sequestered by immunoproteins such as serum albumin,^{154, 213} however a specific binding protein exists to specifically sequester bacterial siderophores. Siderocalins, also known as Scn or Lipocalin 2, are host immunoproteins that cannot bind free iron(III), but bind iron(III)-bound siderophores, limiting the availability for bacterial iron uptake (Figure 23).²¹⁴⁻²¹⁸ The structures of siderocalins are thought to be optimised for binding the siderophores of bacterial species of particular pathogenic threat.^{37, 218} The human siderocalin protein is a β -barrel with a positively charged binding pocket that promotes the binding of negatively charged iron(III)-siderophore complexes, and a tyrosine that provides direct coordination to the iron(III) centre.²¹⁹⁻²²¹ Siderocalin can bind a number of iron(III)-bound siderophores including enterobactin, bacillibactin, vibriobactin, fluvibactin, carboxymycobactins and pyochelin, as well as a number of siderophore mimic complexes, bidentate catechols, and actinide complexes.²¹⁹⁻²²⁴ Enterobactin binding is predominantly electrostatic, (Figure 23), and has a very high affinity, with a subnanomolar dissociation constant of 0.41 ± 0.11 nM.^{217, 219} Due

to the specificity of the siderocalin binding pocket, some functionalised enterobactin derivatives, and siderophores with vast structural differences cannot be captured by siderocalin.^{215, 218, 219}

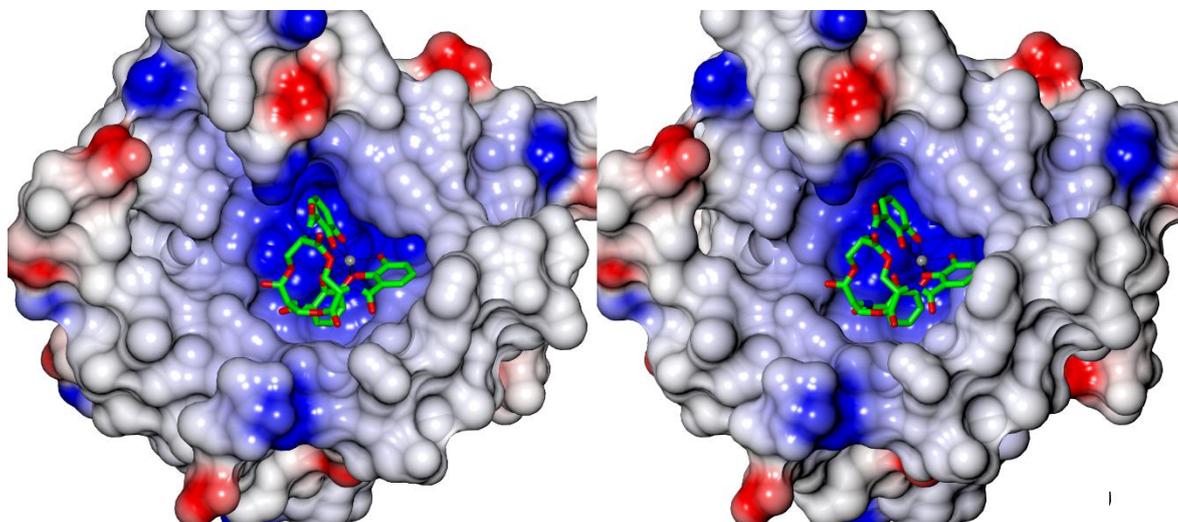


Figure 23: Stereo view of crystal structure of human siderocalin in complex with iron(III)-enterobactin (cylinders coloured by atom type) (PDB ID: 3CMP) Siderocalin is shown in surface view, with blue representing positively charged regions, red representing negatively charged regions and white representing neutral or hydrophobic regions.^{219, 220} Figure produced using CCP4mg.

Bacteria that employ a range of structurally diverse siderophores can survive over a wide pH range, and adapt to a range of environments. For example, while the majority of *E. coli* strains produce enterobactin, as described in section 1.5, the additional production of aerobactin, yersiniabactin and salmochelins is more prevalent in pathogenic strains.^{138, 155} Stealth siderophores are so-called because they are able to evade immunoproteins such as siderocalin.^{215, 218 214, 225} These siderophores are thought to be used by bacteria to more effectively acquire iron from their mammalian host.

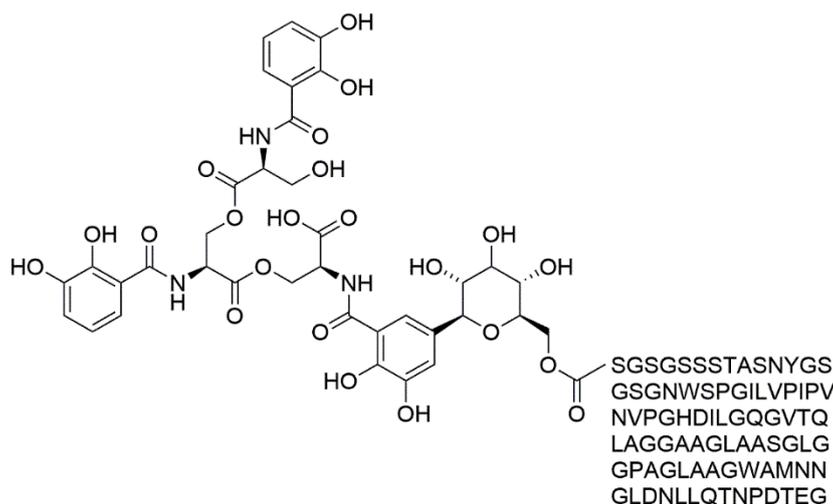
Aerobactin (Figure 15) does not bind to siderocalin due to lack of aromatic groups capable of cation- π interactions, causing structural incompatibility of the siderophore with the siderocalin binding cleft. Yersiniabactin (Figure 15) also possesses structural incompatibilities comprising one aromatic ring and three thiazolidine rings so is not bound by siderocalin.²²⁶⁻²²⁸ Although these siderophores bind iron(III) less effectively than enterobactin, their production provides a selective advantage for survival of *E. coli* under iron deficiency.^{218, 222, 229, 230}

Petrobactin is a mixed-type stealth siderophore with two catechol groups, and one α -hydroxycarboxylate, produced by *Bacillus anthracis*.^{215, 218, 230} The siderophore evades siderocalin by the use of 3,4-catecholates that are sterically incompatible with the siderocalin binding cleft, rather than the 2,3-arrangement that is present in enterobactin and bacillibactin.^{230, 231}



Salmochelin siderophores are C5-glucosylated analogues of enterobactin and the enterobactin hydrolysis products produced under neutral to alkaline conditions by pathogenic strains of *E. coli* (Figure 15).^{138, 157, 232, 233} They incorporate the enterobactin backbone with extremely high iron(III) affinity, but have increased hydrophilicity when compared to enterobactin, and the appended glucose moieties confer steric bulk.^{154, 234} These two properties allow the siderophores to avoid sequestration by immunoproteins.^{235, 236} In addition, salmochelin siderophores were proven to be more effective at iron(III) acquisition in the presence of lipids, when compared to enterobactin, due to the decreased membrane partition coefficient of the hydrophilic siderophore.²³⁴

Salmochelins are derivatised into a number of microcins.⁷³ Siderophore-microcins are short peptides (5-10 kDa) linked to siderophores *via* the C-terminus, used by bacteria to target and kill iron-starved competitors - commonly by forming cytoplasmic membrane pores.²³⁷⁻²⁴⁰ Microcin immunity proteins are expressed in conjunction with microcins to protect the exporter species from its own microcin. For example, microcins secreted under iron limited conditions by *E. coli* are effective against *Salmonella typhimurium*.²³⁸ MccE492 is a microcin produced in *Klebsiella pneumoniae* known to be linked to the linear trimer salmochelin MGE.^{73, 241} This microcin targets *E. coli* *via* the FepA, Fiu, and Cir outer-membrane receptors.^{237, 239, 242-244} Microcins not only evade the host immunoproteins as stealth siderophores, but have a Trojan-horse type antimicrobial effect on competitors to increase the availability of iron.²⁴⁵



Microcin E492

In addition, natural inhibitors for siderophore production are known as a tool for a bacterial species to limit the iron(III)-uptake of neighbouring species. Baulamycin A is produced by *Streptomyces tempisqueus* to block siderophore synthesis in *Staphylococcus aureus* and *Bacillus anthracis*.²⁴⁶ This strategy is particularly effective when siderophore piracy between species is low.

1.7 Siderophore Mimics

Natural siderophores often require complex syntheses, due to their structures containing specific stereochemistries. In addition, they often contain hydrolytically labile linker regions, that are required in biological iron-uptake for hydrolysis of the siderophore for iron(III) release.^{76, 247} This instability can cause challenges for *in vitro* studies, so producing a similar but more stable mimic compound allows for more versatile investigations than may be possible with the natural moiety. In addition, mimic designs can be simplified, allowing individual siderophore features to be investigated. Many mimics have been well studied and their iron(III)-binding properties investigated, but they are also useful for study of biological function, or for development of novel applications.^{2, 108} To investigate how siderophores have such high affinity and specificity for iron(III), denticity and conformational rigidity have been explored with the use of mimics. Computational modelling tools have been employed to establish how siderophore structure influences iron(III)-binding affinity.⁶³

MECAM (Figure 24) is a well-established mimic for enterobactin, and is a triscatecholate with an aromatic backbone that confers rigidity and structural stability over a wide pH range.^{248, 249} Although MECAM binds iron(III) with lower affinity than enterobactin, (pFe(III) values listed in Table 2), this mimic was instrumental to the determination of the

salicylate binding mode in enterobactin.^{29, 68, 248, 250} Sulfonation of MECAM in the C5 position increased the aqueous solubility and resistance to oxidation. In addition, C5-trisulfonated-MECAM (MECAMS) had a slightly higher pFe(III) than MECAM (Table 2).²⁴⁸ Functionalisation of the MECAM backbone with methyl and ethyl groups proved that backbone rigidity is important for ligand pre-organisation, and enhances iron(III) binding (Table 2).²⁵¹ 3,3,4-CYCAM (Figure 24) was a poorer iron(III) chelating ligand than MECAM due to the conformational demands of the flexible ring backbone, that does not have the hydrogen bonding capacity of the triserine backbone in enterobactin.^{2, 249, 252}

TRENCAM (Figure 24) has also been used extensively as an enterobactin mimic. The tertiary amine backbone provides a more flexible structure that results in a lower pFe(III) (Table 2).²⁵³ Alongside a bidentate 2,3-dihydroxyterephthalamide siderophore mimic (*S*)-(PhMe)₂TAM (Figure 24),²⁵⁴ studies proved the importance of the amide nitrogen, and resulting hydrogen bonding, for the stability of enterobactin, and helped to confirm the Δ -configured preference of the iron(III)-enterobactin complex.^{16, 65} Capping the triscatecholate siderophore to create a macrobicyclic structure was explored as a strategy to impart further rigidity on the iron(III) coordination sphere, with the aim to increase the iron(III) binding affinity and kinetic stability by pre-organisation of the ligand. Several analogues, including bicapped TRENCAM (Figure 24) were synthesised, for which a pFe(III) of 30.7 was achieved.²⁵³ The iron binding affinity did not exceed that of enterobactin, likely due to the fact that the rigid structure imparted trigonal prismatic coordination geometry around the iron(III) centre. Trigonal prismatic geometry is usually of higher energy, and therefore disfavoured when compared to octahedral geometry due to ligand repulsion. The large pFe(III) was therefore more likely due to the ease of deprotonation of the catechol units of this ligand rather than from the benefits of a preorganised structure.^{253, 255} Such bicapped structures have been shown to be tuneable for selective encapsulation of different transition metals, by variation of the linker length resulting in different sized cavities.²⁵⁵ The importance of backbone pre-organisation for high affinity iron binding was demonstrated by the design and synthesis of a biscatecholate-hydroxamate siderophore mimic, with a flexible lysine-glutamic acid backbone (H₆L) (Figure 24). This mimic had a low pFe(III) of 18.3, which did not significantly improve with change in pH, this pFe(III) was even lower than linear trishydroxamate ferrioxamine B (Table 2).²⁵⁶ Interestingly, a mimic with higher iron(III) affinity than enterobactin at neutral pH has never been achieved, however at low pH, the mimic hopobactin (Figure 24), with three 3-hydroxy-1-methyl-2(1*H*)-pyridinonate iron-binding units, has been shown to have a higher pFe(III). This is because iron(III)-hopobactin does not protonate at low pH, and therefore is not destabilised by a salicylate-

type binding mode as is the case for enterobactin.²⁵⁷ The use of carbohydrates in siderophore mimic backbones increase the hydrophilicity of iron(III) complexes.^{111, 258, 259}

Table 2: pFe(III) values for a selection of siderophore mimics.

Siderophore Mimic	pFe(III)
Enterobactin	35.5 ²⁹
MECAM	29.1 ²⁴⁸
MECAMS	29.4 ²⁴⁸
MMECAM	31.2 ²⁵¹
EMECAM	32.6 ²⁵¹
3,3,4-CYCAM	23.0 ²⁶⁰
TRENCAM	27.8 ²⁵³
bicappedTRENCAM	30.7 ²⁵³
H ₆ L	18.3 ²⁵⁶
Hopobactin	27.4 ¹⁰⁸
TRICAMS	25.1 ²⁶⁰
Ferrioxamine B	26.6 ¹¹⁰

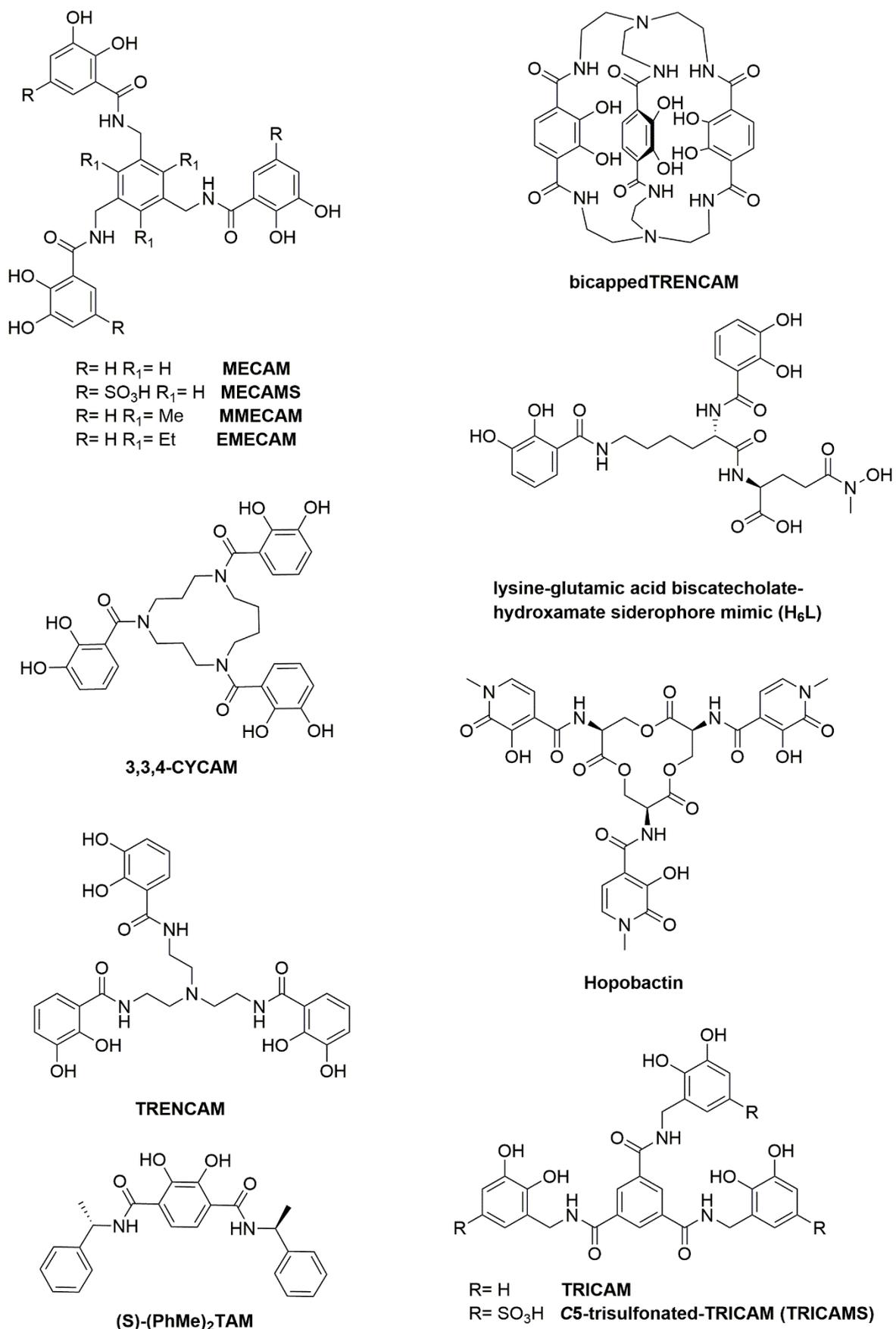


Figure 24: The structures of a selection of catechol siderophore mimics.^{248, 249, 253, 255, 257,}

261, 262

Growth studies showed that several catecholate siderophore mimics could support growth of *E. coli* under iron-limited conditions. MECAM and TRICAM (Figure 25) were shown to support growth, but their sulfonated derivatives MECAMS and TRICAMS (Figure 25) could not. Interestingly, the linear sulfonated triscatecholate, 3,4-LICAMS supported growth. This gave an insight into the specificity of the iron(III)-siderophore uptake machinery in *E. coli*.¹⁶⁵ Mimics of ferrichrome have proven useful in the study of hydrogen bonding in iron(III) chelation, and the specificity of hydroxamate outer membrane receptors.²⁶³ Enterobactin receptor recognition of the *E. coli* outer-receptor FepA was probed using siderophore mimics bound to rhodium(III). These kinetically inert complexes allowed the isolation of the individual properties of key functional units of enterobactin, the iron-binding unit, the amide linkage and the triserine backbone. It was successfully shown that the iron-binding catechol unit and carbonyl were essential for receptor recognition, but the compounds could still be taken up if there was variation at the amide nitrogen and backbone.²⁶⁴ These studies of the properties of triscatecholate siderophores have since been used to inspire high affinity benzene-o-dithiolato donor ligands for titanium(IV) and molybdenum(IV) complexes.²⁶⁵

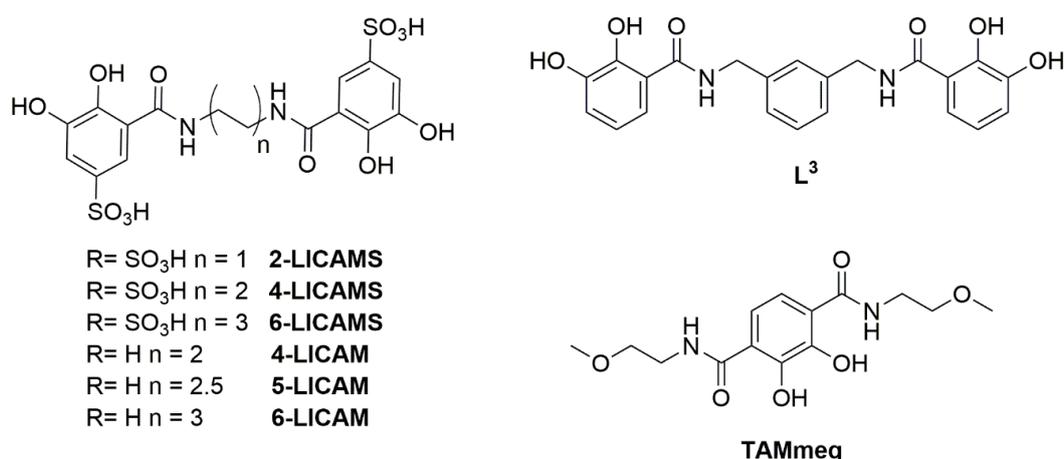


Figure 25: Tetradentate and bidentate catecholate siderophore mimics.^{94, 266, 267}

As well as hexadentate siderophore mimics, catecholate siderophore mimics have been synthesised to mimic tetradentate siderophores such as bisDHBS and azotochelin. n-LICAM and n-LICAMS mimics include an aliphatic backbone linking two catecholamide units, and are more hydrolytically stable than the ester-based backbone of bisDHBS.²⁶⁶ L³ (Figure 25) is a more rigid mimic, with a xylene spacer. L³ and 5-LICAM were synthesised to probe the molybdenum binding ability of tetradentate siderophores.^{94, 268} 2-,4- and 6-LICAMS were synthesised for applications in actinide sequestration.²⁶⁶ In addition, simple bidentate terephthalamide-based ligands have been shown to have high iron(III) binding affinity.^{254,}

²⁶⁷ The highest affinity of which is TAMmeg (Figure 25), and when prepared in a 3:1 iron complexes has a pFe(III) of 24.7.²⁶⁷

1.8 Applications of Siderophores

Iron Chelation Therapy

Patients that are diagnosed with blood disorders such as sickle cell anaemia and thalassemia commonly require multiple blood transfusions. As iron excretion mechanisms are limited within the body, iron overload is a common side effect.²⁶⁷ Excess iron is toxic to mammalian cells, causing problems for the heart, liver and endocrine system, as well as promoting unwanted bacterial infections.^{269, 270} Siderophores and siderophore mimic compounds are therefore in clinical use to enable the excretion of excess iron.^{256, 271} The first approved medical iron chelator was ferrioxamine B, which had to be administered *via* subcutaneous infusion. More recently, the orally active chelator deferasirox (Figure 26) has become a preferred choice, due to higher patient compliance.^{256, 272} Iron overload is still a high cause of death in transfusion-dependent patients, with over 70% of deaths in thalassemia patients due to heart problems, so further work is required to improve the properties of medical iron(III) chelators.^{269, 270, 273, 274} Enterobactin analogues 3,4-LICAMS (Figure 26) and MECAM (Figure 24) have been explored for their ability to remove iron(III) from transferrin. MECAM was shown to outperform enterobactin over a 30 minutes period and 3,4-LICAMS outperforms ferrioxamine B (Figure 26) at high ligand: transferrin ratios.²⁷⁵ It was suggested that enterobactin forms complexes that are too kinetically labile for this application, and thermodynamics as well as kinetics must be considered for design of an efficient therapeutic agent.²⁷⁵ With this in mind, the 2,3-hydroxypyridinone unit, found not to suffer from the proton competition that is characteristic of catechol ligands, may have the necessary thermodynamic stability for design of new therapeutic iron chelators.²⁶⁷

More recently, orally administered iron chelating polymers with a hexadentate 3-hydroxypyridin-4-one based structure have been developed, that tackle iron overload by reducing dietary iron uptake.^{276, 277} Such polymers have the advantage that the chelating agents are not absorbed through the intestine, so carry fewer side effects, and do not promote undesirable bacterial infections.²⁷⁸ Chelation therapy using siderophore mimics can also be extended to the sequestration of heavy metals. Uranium was shown to be successfully

cleared from mammalian kidneys by tetradentate catecholate 5-LICAMS (Figure 25) and 5-LIO(Me-3,2-HOPO) (Figure 26).²⁷⁹⁻²⁸²

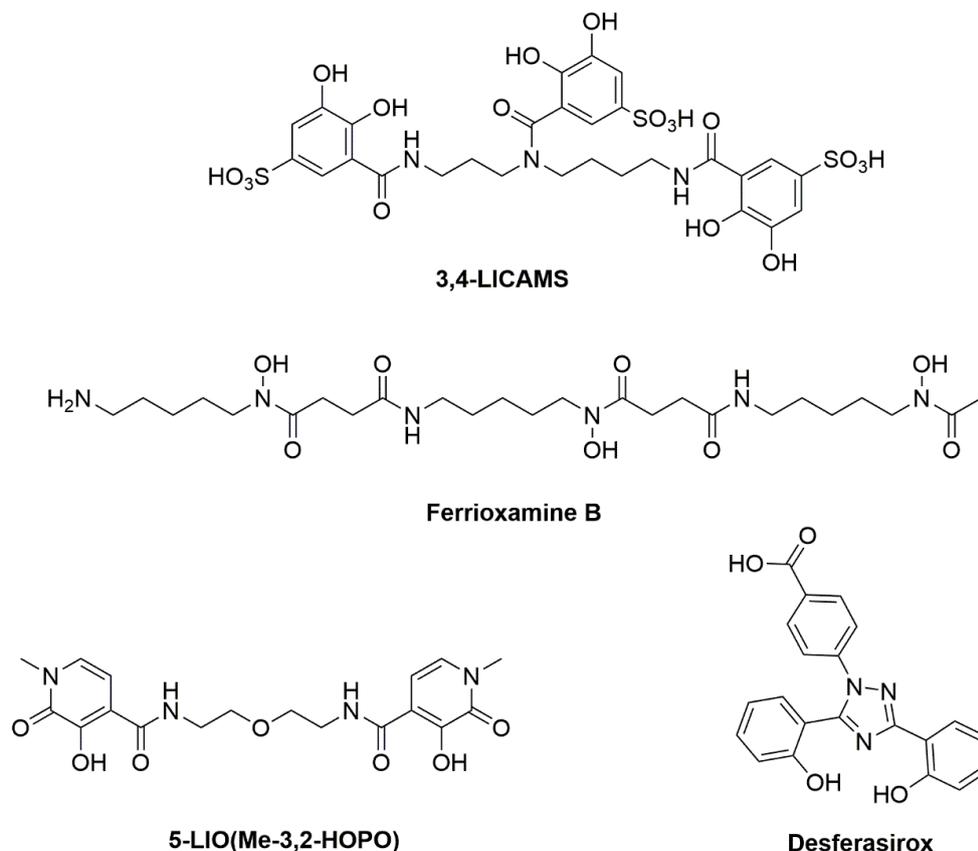


Figure 26: Siderophore mimics involved in iron(III)-chelation therapy.²⁷⁹⁻²⁸³

Bacterial Detection and Molecular Imaging

Incorporation of radioisotopes into siderophore complexes has potential application in positron emission tomography imaging, and as imaging agents with more broad functions.²⁸⁴⁻²⁸⁶ Initial studies have involved the use of gallium-68 and zirconium-89 complexes of ferrioxamine B and E (Figure 9).²⁸⁴ In addition, attachment of a fluorescent moiety to a siderophore can be useful in applications of tracking iron(III), siderophore uptake, or identification of particular bacterial species.^{239, 287, 288} This could prove useful for better diagnosis and more selective treatment of bacterial infections, resulting in more rapid patient recovery.²⁸⁵ Fluorescent analogues of ferrichrome (Figure 3) were successfully tracked in uptake studies of *Pseudomonas* species, and fluorescent catecholate derivatives have been used to track and image *P. aeruginosa* infection in mice.^{289, 290}

Antibiotics by Iron Starvation

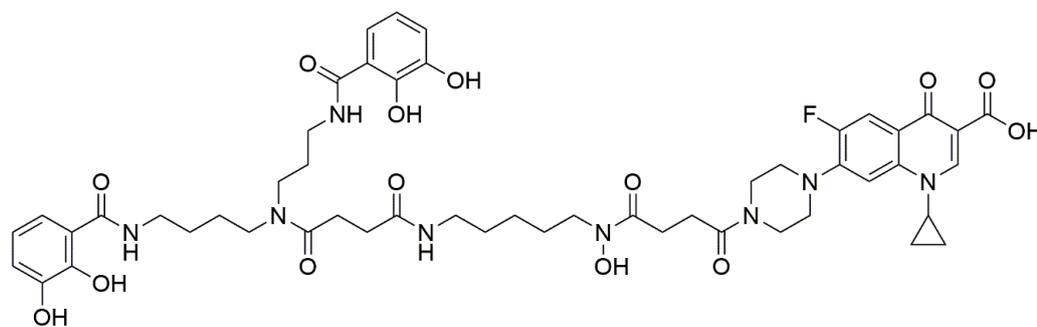
It has been shown that high affinity iron chelators that are not recognised by bacterial outer membrane receptors can be used effectively in bacterial iron starvation.²⁷⁷ This has been

proven using 3-hydroxypyridin-4-ones based siderophore mimics for disruption of staphyloferrin-driven iron(III)-uptake in methicillin resistant *Staphylococcus aureus*. As these systems can be incorporated into polymers, the most promising applications are in external wound infections.^{277, 291}

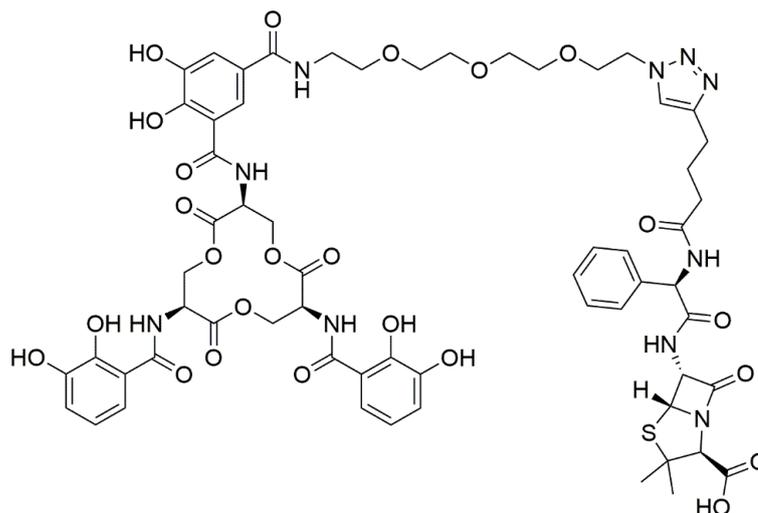
Trojan Horse Antibiotics

A strategy in the development of novel antibiotics is to exploit the absolute requirement for iron of bacteria. By attaching antibiotics to siderophores, the approach relies on the bacterial outer-membrane receptors recognising the iron(III)-loaded siderophore component when under iron-limiting conditions, and importing the whole conjugate into the cell, where the antibiotic can then exert its toxic effects.²⁹² This antibiotic smuggling method may be useful in combating problems of antimicrobial resistance that arises from adaptations in membrane permeability, and drug efflux.²⁹³ Sideromycins are natural antimicrobial conjugated siderophores that inspire this application.^{246, 293}

The focus of this work has involved the incorporation of β -lactam or fluoroquinolone antibiotics onto a siderophore scaffold (Figure 27).²⁹⁴⁻²⁹⁸ Studies have often suffered from reduced activity of the antimicrobial due to the attachment of the siderophore component, and further efforts are required to produce an optimised conjugate.^{299, 300} Progress into increasing potency has been made where the linker region is bio-labile, and the antibiotic can be released from the conjugate.^{294, 301} In addition, hydrophilic salmochelin siderophores that have a reduced mammalian host immune response have been incorporated, for enhanced targeting of specific species, providing better control of antibiotic resistance.²⁹⁷



Biscatecholate-mono-hydroxamate-ciprofloxacin conjugate



Enterobactin-ampicillin conjugate

Figure 27: Examples of siderophore-antibiotic Trojan horse conjugates.^{296, 301}

For antibiotics that are active on Gram-positive but not Gram-negative bacteria, the Trojan horse strategy may improve access of current antibiotics into a wider range of bacterial species. For example, oxazolidinones were proven to be much more potent in *Pseudomonas aeruginosa* when conjugated to a catecholate siderophore mimic.³⁰² Complexation of a DNA intercalator with an iron(III)-hydroxamate siderophore mimic enhanced selectivity of the drug towards bacterial cells over mammalian cells.³⁰³ The Trojan horse strategy is of particular interest for the development of antifungal treatments which commonly encounter selectivity issues due to the fact that mammalian and fungal cells are both eukaryotic. Targeting siderophore uptake machinery which is not present in mammalian cells may enable the reduction of adverse side effects of antifungal drugs.²⁸⁶

1.9 Project History

As described in section 1.5, CeuE is a periplasmic lipoprotein involved in catecholate siderophore uptake in *C. jejuni*. This protein was the focus of previous work in this research group.

[Ceue-iron(III)-MECAM]₂

The first crystal structure of Ceue was obtained from a co-crystal with iron(III)-MECAM.¹⁸⁷ The crystal contained 2:2:2 dimeric complexes, with two MECAM ligands bridging between two Ceue binding clefts, with two iron(III) centres, shared by both MECAM ligands (PDB ID: 2CHU) (Figure 28).¹⁸⁷ This was a surprising ligand arrangement, as it is much more common for similar periplasmic binding proteins to adopt 1:1:1 complexes when bound to iron(III)-siderophore complexes. Known 1:1:1 complexes include FhuD from *V. cholerae* bound to desferrioxamine B (PDB ID: 5GGX), FepB from *E. coli* bound to iron(III)-enterobactin (PDB ID: 3TLK), FeuA from *B. subtilis* bound to iron(III)-enterobactin (PDB ID: 2XUZ) and even the iron(III)-MECAM complex of FeuA from *B. subtilis* (PDB ID: 2XV1).^{122, 146, 304, 305} It was reasoned that the formation of the unusual dimeric complex of [Ceue-iron(III)-MECAM]₂ was favourable due to the stabilisation provided by hydrophobic interactions of the mesitylene backbones.

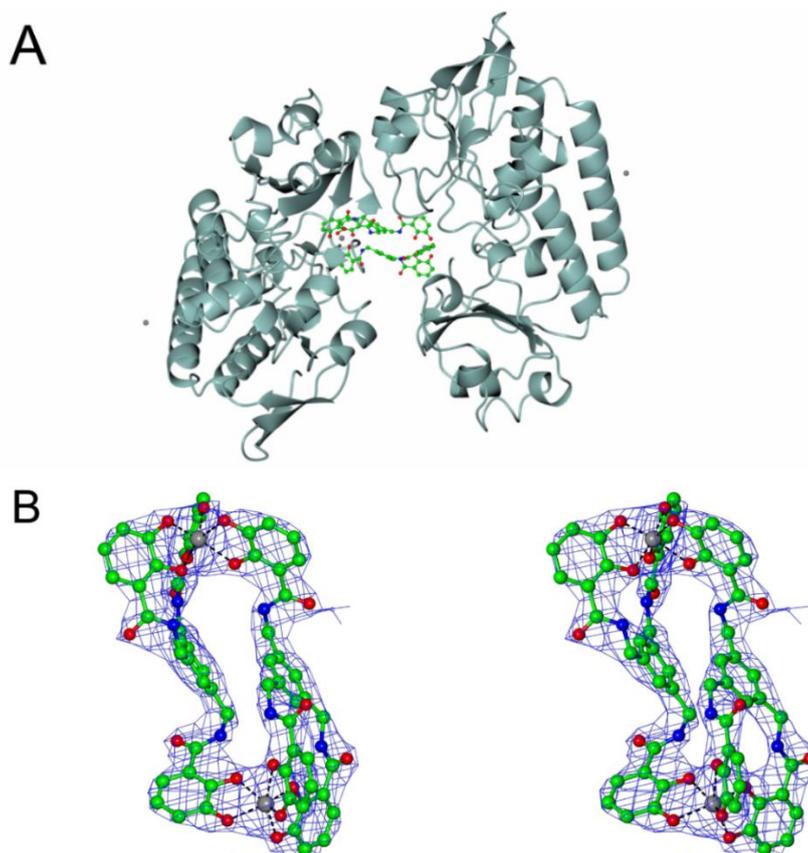


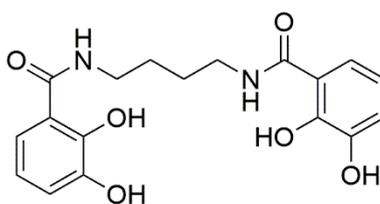
Figure 28: A: Crystal structure of Ceue-iron(III)-MECAM in the 2:2:2 dimer (PDB ID: 2CHU).¹⁸⁷ Ceue is shown in sea green ribbons, and iron(III)-MECAM in ball and stick coloured by atom type. B: Stereo view of the 2:2 Fe: MECAM arrangement across two Ceue binding pockets shown in ball and stick coloured by atom type. The electron density

for the maximum likelihood map was contoured at the 2σ level. Figure produced using CCP4mg.

The complex was arranged in the Λ -configuration in both binding pockets, and this configuration was also present in CeuE-iron(III)-MECAM complexes in solution, as confirmed by circular dichroism.^{11, 13, 187} It was not established whether the 2:2:2 dimer was present in solution phase or indeed in a biologically relevant environment. The CeuE concentration in the periplasm, of somewhere between 0.1 and 1 mM, is a lot lower than in crystal screening and so it is possible that the 2:2:2 dimer is favourable at high concentrations but not at lower biological concentrations.³⁰⁶ It may be that favourable hydrophobic interactions between MECAM backbones causes protein dimerisation at biologically-relevant concentration, rendering the binding clefts inaccessible for interaction with the necessary inner membrane transporter CeuBCD.^{115, 172, 187} Although these findings may render MECAM a poor enterobactin mimic for the CeuE system, it has been postulated that these properties may prove advantageous for future applications such as in the development of antimicrobials.²⁰

CeuE-iron(III)-4-LICAM

CeuE binds the tetradentate siderophore mimic iron(III)-4-LICAM in a 1:1:1 binding ratio and this ligand served as a mimic of the enterobactin hydrolysis product bisDHBS.^{86, 95, 186}



4-LICAM

This was the first reported crystal structure for a periplasmic binding protein crystallised with an iron(III)-bound tetradentate ligand (PDB 1D: 5A1J) (Figure 29).¹⁸⁶ Studies with bidentate and tetradentate siderophores previously showed that lower denticity siderophores are able to support growth of a number of bacterial species.^{307, 308} There are however few protein crystal structures of bacterial uptake machinery in complex with low denticity siderophores.³⁰⁹ This study established the importance of the investigation of tetradentate siderophores, and their direct role in iron(III) uptake *via* interaction with ferric siderophore transport proteins.

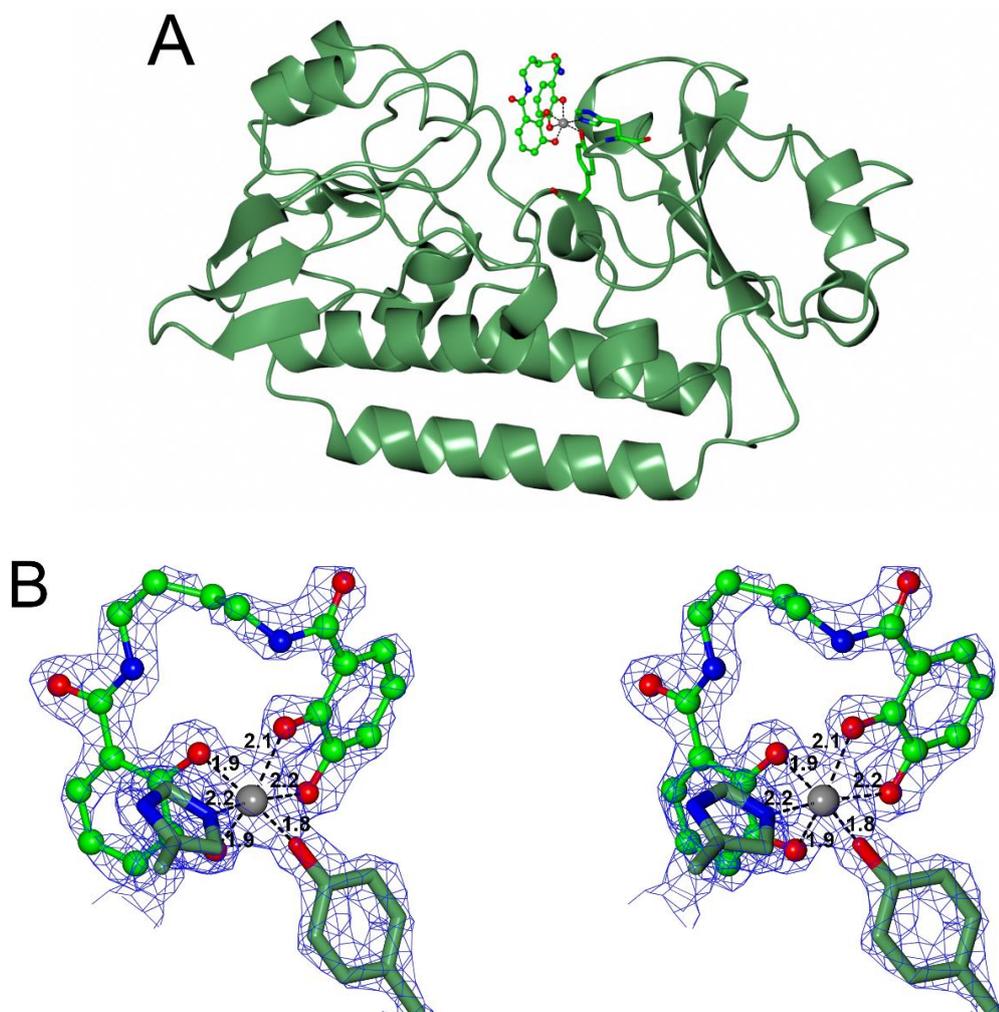


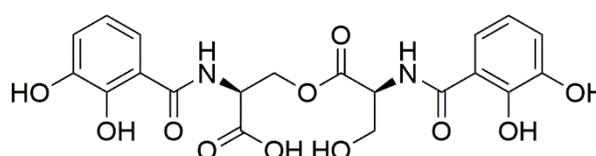
Figure 29: Crystal structure of CeuE-iron(III)-4-LICAM.¹⁸⁶ A: CeuE depicted in lawn green ribbons, iron(III)-4-LICAM shown as ball and stick coloured by atom type with coordinating His 227 and Tyr 288 shown as cylinders coloured by atom type. B: Stereo view of iron(III)-4-LICAM shown as ball and stick coloured by atom type with coordinating His 227 and Tyr 288 shown as cylinders coloured by atom type. iron(III) bond distances are labelled in Å. The electron density for the maximum likelihood map was contoured at the 2σ level. Figure produced using CCP4mg.

A key finding was the discovery of His 227 and Tyr 288 residues from CeuE that donated N and O chelating atoms respectively to directly coordinate the iron(III) centre, completing the hexadentate octahedral coordination (Figure 29). The resulting complex retained Λ -configuration, as seen for [CeuE-iron(III)-MECAM]₂ and circular dichroism was employed to prove the Λ -configuration was also dominant in solution phase.^{186, 187} Fluorescence quenching titration studies proved that iron(III)-4-LICAM bound tightly to CeuE, with a dissociation constant of 21 ± 6 nM.¹⁸⁶ It was suggested that CeuE is somewhat promiscuous, and may be capable of binding a range of siderophores, and as well as hexadentate

siderophores, the protein could employ His 227 and Tyr 288 to aid full iron(III) coordination for lower denticity siderophores.¹⁸⁶

A Linker Length of Five Atoms

A natural tetradentate siderophore, bisDHBS, was synthesised and studied with CeuE.⁹⁰ This study was fundamental for proving the biological relevance of tetradentate siderophores in *C. jejuni*, and for testing the hypothesis that n-LICAM compounds make good mimics for tetradentate catecholate siderophores. Interestingly, iron(III)-bisDHBS bound tightly to CeuE with a low dissociation constant of 10 ± 4 nM. Circular dichroism proved that the Λ -configuration was adopted upon introduction of CeuE to iron(III)-bisDHBS in solution.⁹⁰



BisDHBS

Crystal structure solutions showed that the CeuE-iron(III)-bisDHBS complex (PDB ID: 5ADW) does indeed adopt a similar ligand binding arrangement to that of CeuE-iron(III)-4-LICAM, with His 227 and Tyr 288 completing the octahedral iron(III) coordination. Interestingly, CeuE exhibits high enantioselectivity, with the asymmetric chiral serine-derived linker between the catecholate binding units always incorporated into the binding pocket in the same orientation. This is demonstrated by the good electron density observed for the free alcohol and acid in fixed orientation in the crystal structure (Figure 30).⁹⁰ Next, CeuE was studied with iron(III)-5-LICAM, a siderophore mimic of iron(III)-bisDHBS with the same number of atoms in the linker region between the two catecholate iron(III) binding units. Iron(III)-5-LICAM binds to CeuE in a very similar manner to that of iron(III)-bisDHBS with r.m.s.d of 0.63 over 286 C α positions upon superposition of the structures with the SSM algorithm (Figure 31).³¹⁰ Iron(III)-5-LICAM binds to CeuE with a very high affinity, <10 nM *via* fluorescence quenching titration.³¹⁰ As such, n-LICAM compounds were chosen as good mimics for tetradentate siderophores such as bisDHBS, and these studies provided a good foundation for further investigation of the CeuE binding pocket *via* manipulation of structurally similar ligands.

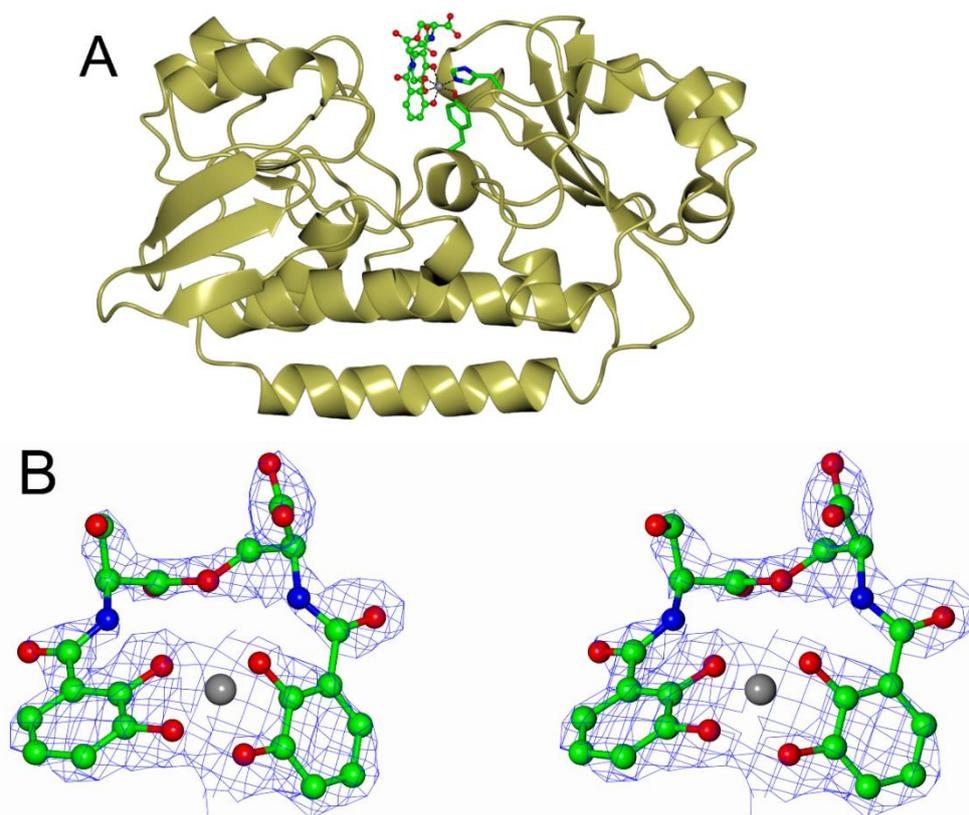


Figure 30: Top: CeuE depicted in gold green ribbons, iron(III)-bisDHBS shown as ball and stick coloured by atom type with coordinating His 227 and Tyr 288 shown as cylinders coloured by atom type. Bottom: Stereo view of iron(III)-bisDHBS shown as ball and stick. The electron density for the maximum likelihood map contoured at the 1.5σ level.⁹⁰ Figure produced using CCP4mg.

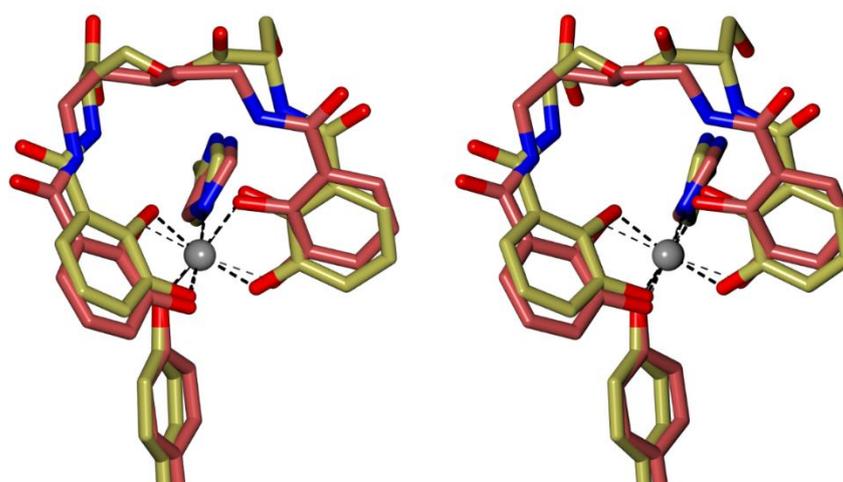


Figure 31: Stereo view of the crystal structure overlay of CeuE-iron(III)-5-LICAM (cylinders coloured by atom type with carbon in light crimson) with CeuE-iron(III)-bisDHBS (cylinders coloured by atom type with carbon in light gold) superimposed by SSM algorithm based on C α positions. The iron(III)-binding ligands display a very similar binding mode in the CeuE binding pocket. Figure produced using CCP4mg.

1.10 Project Aims

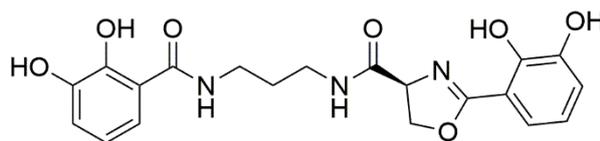
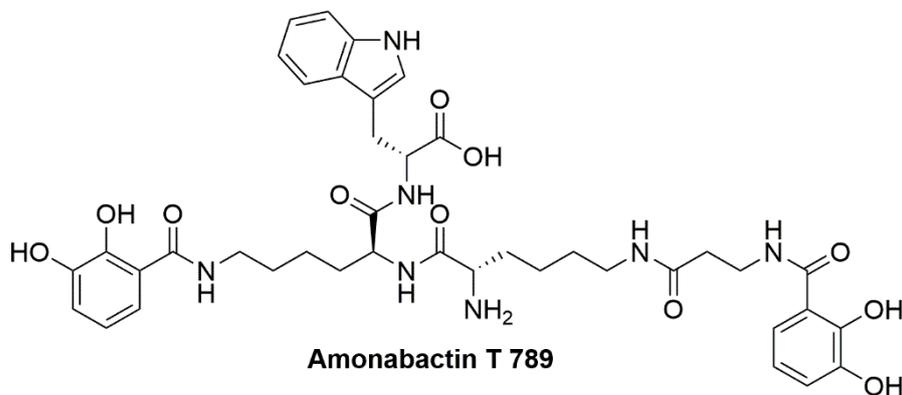
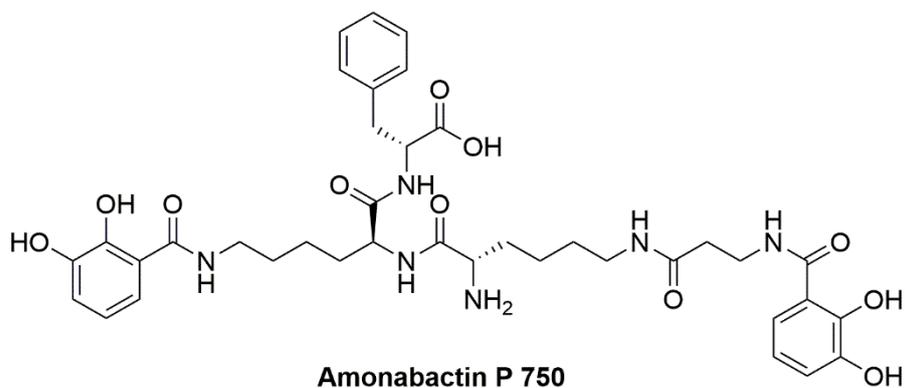
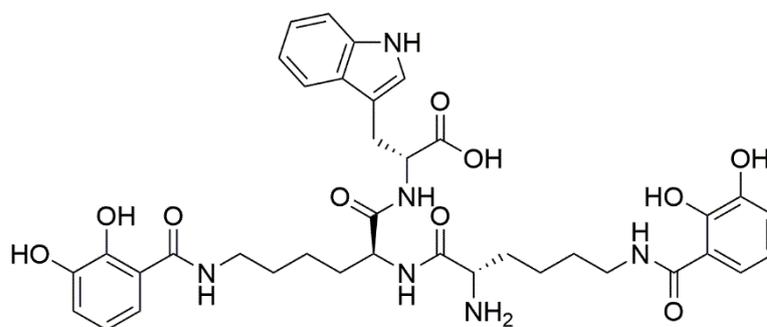
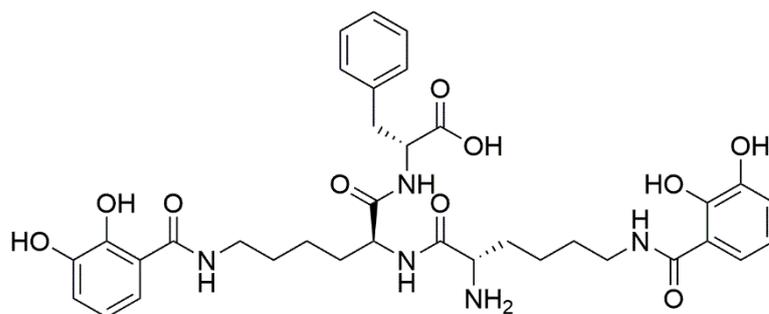
The overall aim was to investigate the interactions between selected bacterial periplasmic binding proteins and iron(III)-siderophore and siderophore mimic complexes. The project was designed to increase knowledge of bacterial iron uptake strategies, which could itself inform the development of novel antimicrobials or artificial metalloenzymes. The investigations are divided into four key areas of study. The first covers how siderophore linker length affects binding of tetradentate siderophores to CeuE. Linker lengths of six and eight atoms were employed in iron(III)-n-LICAM siderophore mimics and studied *via* fluorescence quenching titration, circular dichroism and protein crystallography to probe the protein-ligand interactions and binding affinities (Chapter 2). The second area of study uses CeuE variant proteins to investigate key residues His 227 and Tyr 288, and their importance in iron(III)-siderophore binding *via* the same above techniques (Chapter 3). The third area covers the design and synthesis and iron(III)-binding of tetradentate salmochelin siderophore mimics and their interaction with CeuE (Chapter 4). The fourth area of study aims to tie all three previous areas together, with comparison of three periplasmic binding proteins CeuE, FepB and VctP, and investigation of their binding preferences for hexadentate or tetradentate iron(III)-siderophore complexes, with additional insight into whether these proteins may also employ salmochelin-type siderophores (Chapter 5).

2 Investigation of the Binding of Tetradentate Siderophore Mimics to CeuE

2.1 Introduction

CeuE, the periplasmic catecholate-siderophore binding protein from *C. jejuni*, described in detail in Chapter 1, has been reported to bind the tetradentate siderophore, iron(III)-bis(2,3-dihydroxybenzoyl-L-Ser: a hydrolysis product of enterobactin, and the tetradentate siderophore mimics iron(III)-4-LICAM and iron(III)-5-LICAM with high affinity.^{90, 186, 311}

It is known that *C. jejuni* does not produce its own siderophores, but is able to scavenge iron(III)-bound siderophores, produced by other microorganisms, from its environment.^{115, 131, 168, 172, 175, 312} It is therefore of interest to investigate whether *C. jejuni* would be able to scavenge tetradentate siderophores of different linker lengths if they were present in the environment, and potentially utilise them *via* the CeuE transport system to uptake iron(III).^{95, 172} Tetradentate siderophores of interest include serratiochelin and the amonabactins, with extended linker lengths, produced by the enteric pathogens *Serratia marcescens* and *Aeromonas hydrophila*, respectively, that occupy a similar biological niche to *C. jejuni*.^{93, 99, 313, 314}

**Serratiochelin****Amonabactin T 789****Amonabactin P 750****Amonabactin T 732****Amonabactin P 693**

The aim of the present work is to explore a family of iron(III)-siderophore mimic complexes that can be bound in the CeuE binding cleft, and to establish the binding modes and binding affinities. To further knowledge of the tolerance of CeuE to a range of tetradentate siderophores, siderophore mimics with longer linker lengths were investigated, using *n*-LICAM siderophore mimics where $n = 6$ and $n = 8$ (Figure 32).⁹⁴ *n*-LICAM $n = 6, 8$ ligands were synthesised and provided by Anne K. Duhme-Klair. Such mimics, are useful tools, since the natural siderophores are often unstable in aqueous media or are very challenging to synthesise in large enough quantities.^{77, 253, 256, 315, 316} Mimics are generally simplified in their chemical structures when compared to natural siderophores, allowing for uncomplicated model studies of natural systems.

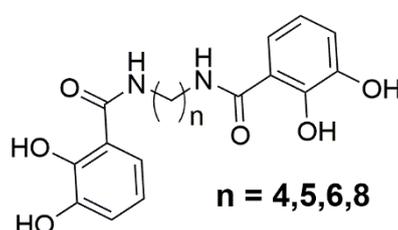


Figure 32: *n*-LICAM siderophore mimic compounds with extended linker lengths from 4 to 8 carbon atoms.

Protein crystallography was employed to determine the structure of siderophore mimic complexes, while fluorescence quenching titrations were used to quantify binding affinities, and circular dichroism to probe binding arrangement in solution. All three techniques were used to assess trends in binding affinity across a range of linker lengths, and to establish the optimum ligand linker length for binding iron(III) catecholate ligands in the CeuE cleft.

2.2 CeuE

Wild-type CeuE was expressed from an overexpression system in BL21 *E. coli* and purified, and characterised in several batches by Daniel J. Raines, Adam Hughes and Elena Blagova via the standard purification procedure detailed for FepB in Chapter 6. The correct construct, containing a C3 cleavable N-terminal hexahistidine tag, was confirmed by DNA sequencing. Purification involved nickel column chromatography, followed by cleavage of the hexahistidine tag, further nickel column chromatography and gel filtration chromatography, with validation via SDS PAGE. The pure protein was characterised with electrospray ionisation mass spectrometry. Circular dichroism spectroscopy was used to confirm that all proteins were folded correctly.^{90, 186, 310}

2.3 Crystal Structure of CeuE-iron(III)-6-LICAM

Protein crystallography revealed how CeuE binds to iron(III)-6-LICAM. This work was analogous to previous co-crystallisation studies carried out for CeuE-iron(III)-4-LICAM.¹⁸⁶ Crystals were obtained *via* co-crystallisation of iron(III)-6-LICAM with CeuE, the procedure for which is detailed in Chapter 6. The crystal was in space group P2₁2₁2₁, containing a single protein monomer in the asymmetric unit. The electron density was well defined for the full CeuE protein backbone, and the overall fold was very similar to previously reported structures.^{90, 186} CeuE adopts a bilobate structure, with the iron(III)-ligand binding cleft situated between two protein domains.

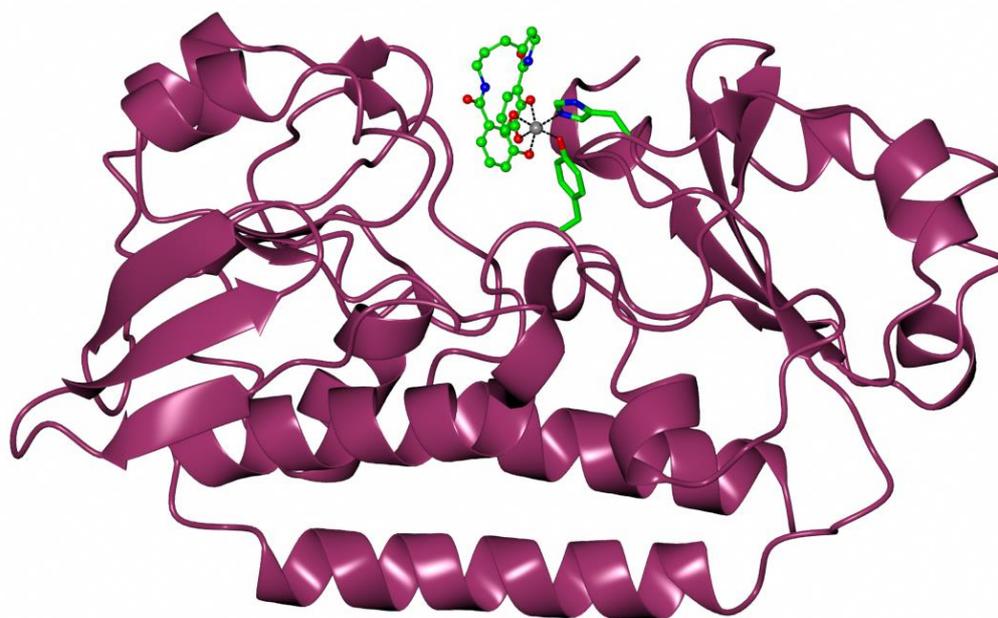


Figure 33: Crystal structure of CeuE-iron(III)-6-LICAM (PDB ID: 5A5V). CeuE is depicted as dark purple ribbons. Iron(III)-6-LICAM is depicted as ball and stick coloured by atom

type. Iron(III)-coordinating His 227 and Tyr 288 are depicted as cylinders coloured by atom type. Figure produced using CCP4mg.

The CeuE-iron(III)-6-LICAM (PDB ID: 5A5V) structure has a root mean square deviation (r.m.s.d.) of 0.48 Å (287 C α positions) from that of CeuE-iron(III)-4-LICAM (PDB ID: 5A1J) upon superposition of the structures calculated using the SSM (secondary structure matching) algorithm,³¹⁷ indicating a very high similarity of the protein fold. There is also a close similarity to the apo-CeuE structure (PDB ID: 3ZKW) with r.m.s.d of 0.94 Å (283 C α positions), indicating that the protein does not significantly rearrange upon ligand binding. This is a common feature of type III periplasmic binding proteins. Bilobate periplasmic binding proteins are known to have a hinge region between the two domains. Type III periplasmic binding proteins have the longest α -helical hinge region, which imparts structural rigidity.^{184, 185}

The positioning of the iron(III)-6-LICAM ligand in the CeuE binding pocket is very similar to that of CeuE-iron(III)-4-LICAM and CeuE-iron(III)-5-LICAM (PDB ID: 5A5D), and there is good electron density for all ligand atoms, with B-values close to the average for the whole structure. This indicates that the ligand had a fixed orientation in the crystal. The iron(III)-centre is coordinated in an octahedral arrangement. The 6-LICAM ligand provides four oxygen donor atoms from the deprotonated catecholate groups. To complete the hexadentate iron(III) coordination, CeuE provides two donor atoms, an oxygen donor from a deprotonated tyrosine residue, Tyr 288, and a nitrogen donor atom from a histidine, His 227 (Figure 34). These are the same donor atoms observed in the CeuE-iron(III)-n-LICAM (n =4, 5) and CeuE-iron(III)-bisDHBS structures. The complex is further stabilised by three arginine residues, Arg 118, Arg 205 and Arg 249, that form hydrogen bonds with the catechol oxygen atoms, and a lysine residue Lys 121 that forms a hydrogen bond with a catecholamide carbonyl oxygen atom (Figure 34). The only other direct hydrogen bonds to the linker region are with water molecules. This suggests that there are few unfavourable interactions between CeuE and the linker region of 6-LICAM, which should allow the protein to accommodate linkers of longer length than the natural linker of five atoms: possessed by bisDHBS.⁹⁰

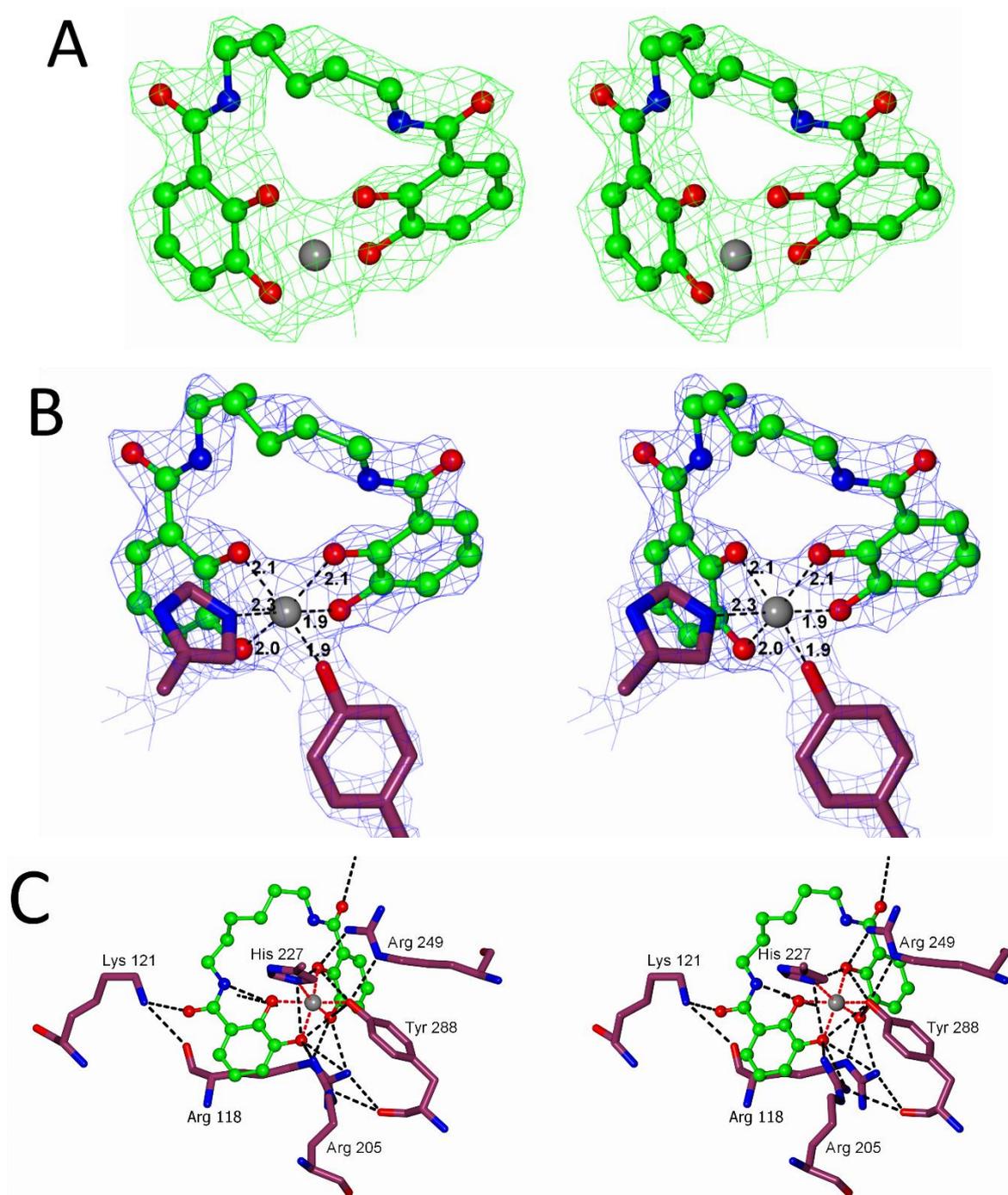


Figure 34: Crystal structure of CeuE-iron(III)-6-LICAM in stereo view. A: iron(III)-6-LICAM as ball and stick. The electron density for the maximum likelihood difference map was calculated before introduction of the iron(III)-ligand into the model to avoid phase bias and contoured at the 2σ level (green chicken wire). B: iron(III)-6-LICAM ball and stick, coordinating His 227 and Tyr 288 in cylinders. Iron(III) bond distances are labelled in Å. The electron-density for the maximum likelihood map was contoured at the 2σ level in blue chicken wire. C: iron(III)-6-LICAM in the CeuE binding pocket, with all significant contacts. Coordinating residues (cylinders) are labelled. Figure produced using CCP4mg.

2.4 Crystal Structure of CeuE-iron(III)-8-LICAM

Crystals of CeuE-iron(III)-8-LICAM were grown by co-crystallisation methods, for which the procedure is detailed in Chapter 6. The high resolution for the structure of 1.32 Å allowed for anisotropic refinement of the model (Figure 36). CeuE bound iron(III)-8-LICAM in a similar manner to iron(III)-6-LICAM. As there are few protein contacts with the n-LICAM backbone, it appears that the longer backbone in 8-LICAM does not cause any unfavourable clashes, but is able to point away from the protein out of the binding cleft into a more solvent exposed region (Figure 35). Interestingly, the iron(III)-8-LICAM ligand is well ordered in the CeuE binding cleft, with clear electron density for fixed atom positions along the whole backbone (Figure 36).

Although all CeuE-iron(III)-n-LICAM (n= 4, 5, 6, 8) crystal structures are in the same space group of P2₁2₁2₁, the unit cell for CeuE-iron(III)-8-LICAM has a quite different packing arrangement. CeuE-iron(III)-8-LICAM has unit cell parameters of a= 42.98 Å, b= 55.98 Å c= 140.08 Å, while for iron(III)-6-LICAM, the cell parameters are a= 61.37 Å, b= 66.08 Å c= 68.96 Å- close to those of iron(III)-n-LICAM (n= 4, 5).¹⁸⁶ This suggests that the solvent exposed longer backbone of iron(III)-8-LICAM results in different crystal contacts. This change in packing arrangement is likely to be the reason for the higher resolution achieved for this structure.

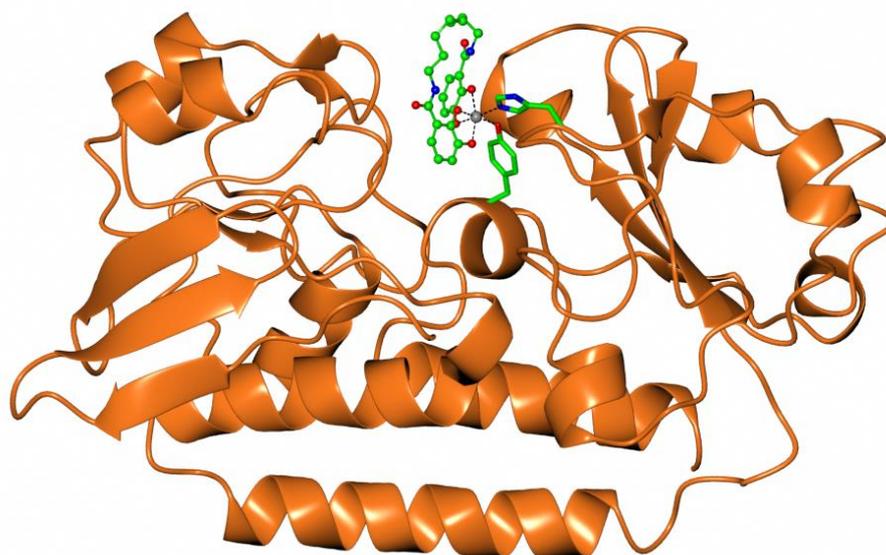


Figure 35: Crystal structure of CeuE-iron(III)-8-LICAM (PDB ID: 5AD1). CeuE is depicted as orange ribbons. Iron(III)-6-LICAM is depicted as ball and stick coloured by atom type. Iron(III)-coordinating His 227 and Tyr 288 are depicted as cylinders coloured by atom type. Figure produced using CCP4mg.

His 227 and Tyr 288 act to complete the hexadentate octahedral coordination of the iron(III) centre as seen for previous structures.^{90, 186} As for the CeuE-iron(III)-6-LICAM structure, the longest iron(III) bond is 2.3 Å, for that of the iron(III)-nitrogen from His 227. This is likely due to the iron(III) being a small, highly-charged cation, with hard Lewis acidity.³¹⁸ The negatively charged oxygen atoms from the 8-LICAM ligand and the deprotonated Tyr 288 have hard Lewis base character and make a good match, forming strong iron(III)-O bonds of between 2.0 and 2.2 Å. As the histidine nitrogen is a weaker Lewis base than the donor oxygens, it is a weaker electron donor, and forms a longer bond to the iron(III) of 2.3 Å (Figure 36). In addition to the iron(III) binding residues, Arg 118, Lys 121, Arg 205 and Arg 249 act to stabilise the ligand binding *via* hydrogen bonding (Figure 36).

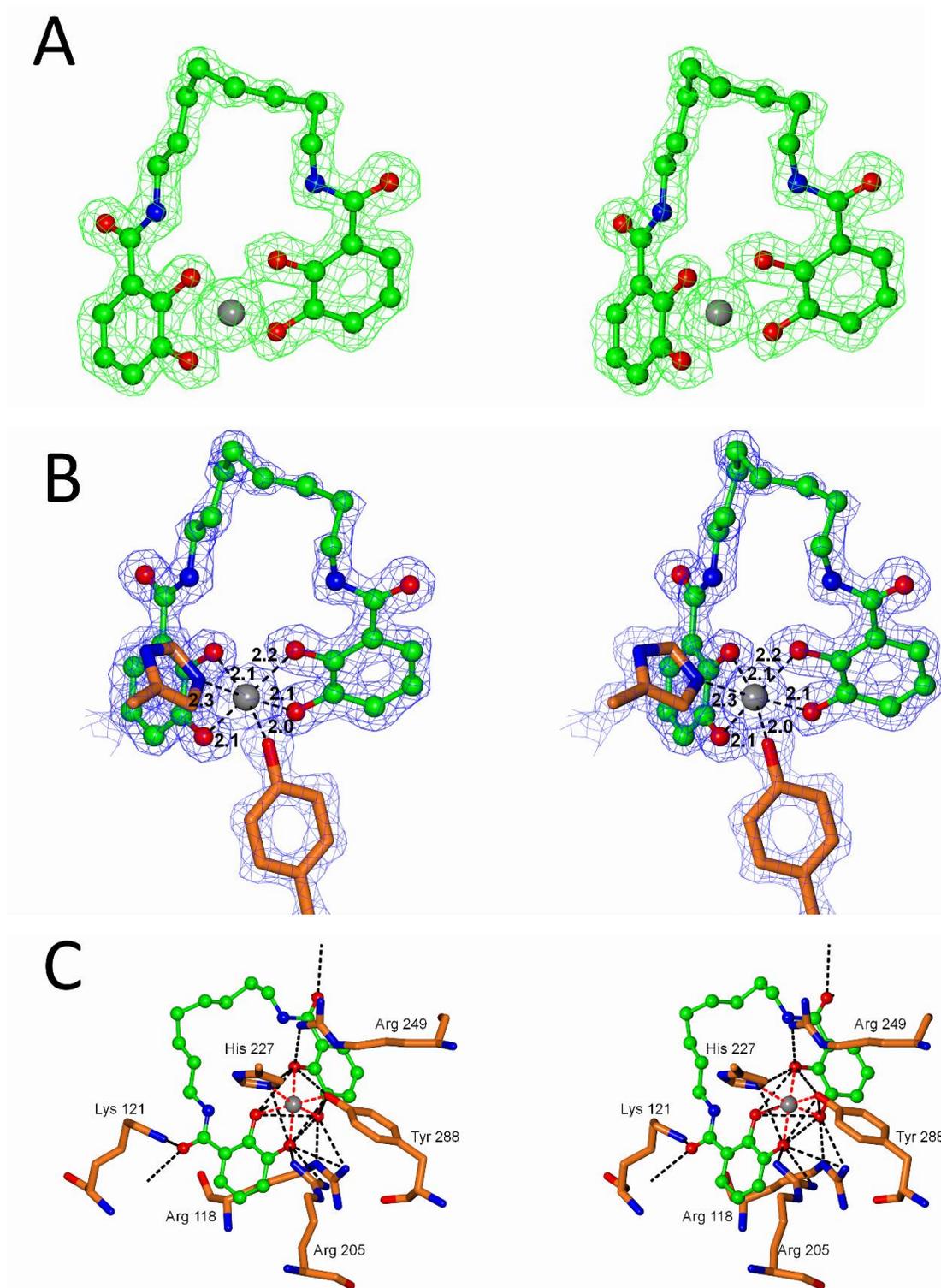


Figure 36: Crystal structure of CeuE-iron(III)-8-LICAM in stereo view. A: Iron(III)-8-LICAM shown as ball and stick. The electron density for the maximum likelihood difference map was calculated before introduction of the iron(III)-ligand into the model to avoid phase bias and contoured at the 2σ level in green chicken wire. B: Iron(III)-8-LICAM as ball and stick, with coordinating His 227 and Tyr 288 as cylinders. Bond distances are labelled in Å. The electron-density for the maximum likelihood map was contoured at 2σ level. C:

iron(III)-8-LICAM in the CeuE binding pocket, with all significant contacts labelled. Figure produced using CCP4mg.

2.5 Determination of Dissociation Constants

Dissociation constants were determined for CeuE with iron(III)-6-LICAM and iron(III)-8-LICAM *via* fluorescence quenching titration. This was possible due to the close-proximity of the only tryptophan residue (Trp 287) to the iron(III)-siderophore binding pocket (Figure 37). When excited at 280 nm, a broad emission peak is observed between around 310 nm and 410 nm.³¹⁹

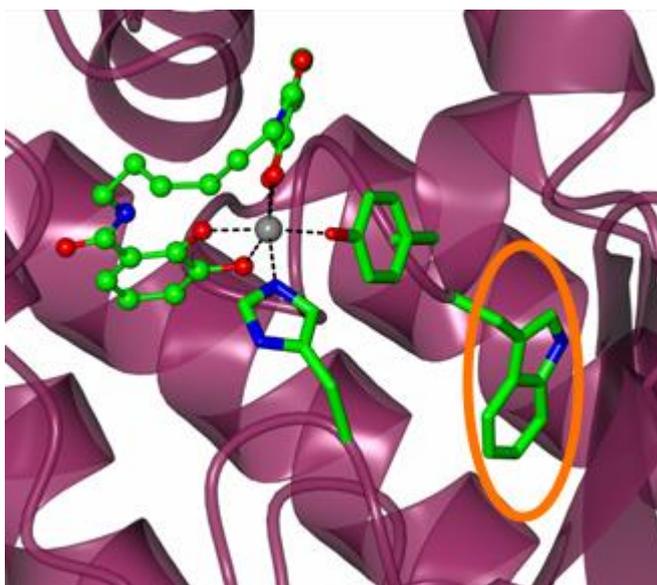


Figure 37: Crystal structure of CeuE-iron(III)-6-LICAM showing that the only tryptophan residue, Trp 287 (circled in orange), is located adjacent to the iron(III)-siderophore binding pocket in CeuE. Figure produced using CCP4mg.

Addition of iron(III)-bound ligands into the CeuE binding pocket causes quenching of tryptophan fluorescence. Subsequently, when CeuE is saturated with iron(III)-bound ligand, the addition of further iron(III)-bound ligand does not cause any further fluorescence quenching. The fluorescence emission observed was therefore recorded over a series of additions of ligand, and the area of the emission band for each ligand concentration calculated and plotted against concentration to obtain a binding curve. This is a commonly used technique for determination of protein-ligand binding affinities,^{219, 223, 231, 320, 321} and was previously used for the determination of dissociation constants of CeuE with iron(III)-4-LICAM and iron(III)-5-LICAM.^{186, 310, 322} It was important to establish that the observed fluorescence quenching was indeed from the binding of the iron(III)-complex to CeuE. Control experiments were recorded and are described in detail in Chapter 6.

It was proven that CeuE was adequately stable, with negligible change in intrinsic fluorescence of the protein, over 75 minutes of repeat exposure to the UV radiation set at the

excitation wavelength of 280 nm. Upon dilution of CeuE by 10%, and application of a linear dilution correction, decrease in fluorescence was observed to be within an acceptable range. This ensured that any fluorescence quenching observed was a direct consequence of the addition of iron(III)-n-LICAM (n= 6, 8) and not due to dilution or oxygen exposure during mixing. It was decided that for a calculated dissociation constant to be significant, the final normalised emission must be below 0.4 to ensure that protein fluorescence was adequately quenched. Ligand absorbance limits were set to ensure no significant contributions from primary inner filter effects.³²³⁻³²⁵ The upper permissible concentrations for iron(III)-6-LICAM and iron(III)-8-LICAM were 37.3 μM and 41.1 μM , respectively.

Fluorescence quenching titrations for CeuE with iron(III)-6-LICAM and iron(III)-8-LICAM were performed according to a published method and so could be directly compared to those already published for iron(III)-4-LICAM and iron(III)-5-LICAM.^{90, 310} The titrations were recorded in triplicate, then the fluorescence signal observed was buffer subtracted and the area of each peak was calculated between 310 and 410 nm. The data were normalised and K_d values and associated error calculated using non-linear regression analysis with Dynafit fitting software for a 1:1 binding model.³²⁶ A weighted average and uncertainty in average calculation were used to calculate the final K_d and error from the three independent titrations.

The curves for each titration of CeuE with iron(III)-6-LICAM are displayed in Figure 38 and the K_d values calculated from each curve recorded in Table 3. The curves for each titration of CeuE with iron(III)-8-LICAM are displayed in Figure 39 and the K_d values calculated from each curve recorded in Table 4.

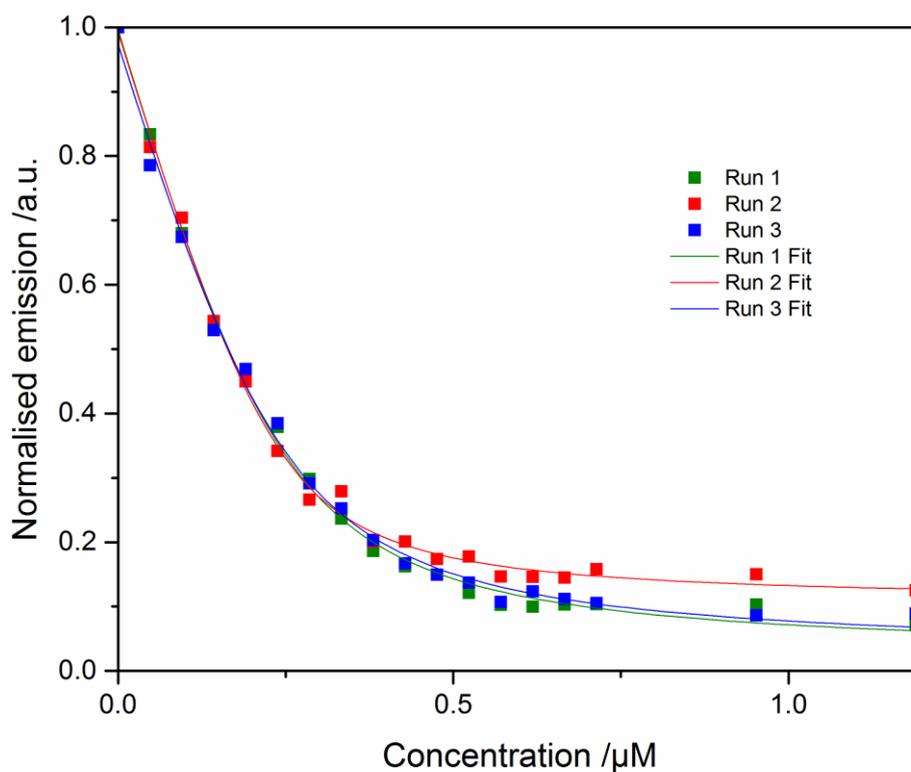


Figure 38: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-6-LICAM and their associated non-linear regression fitting data from Dynafit.³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-6-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Table 3: Dissociation constants for CeuE titrated with iron(III)-6-LICAM for three independent fluorescence titrations, and their associated error. The average was then calculated *via* a weighted average method.

CeuE-iron(III)-6-LICAM	K_d /nM	Error /nM
Run 1	41.1	4.6
Run 2	23.9	3.4
Run 3	44.1	5.8
Average	33	8

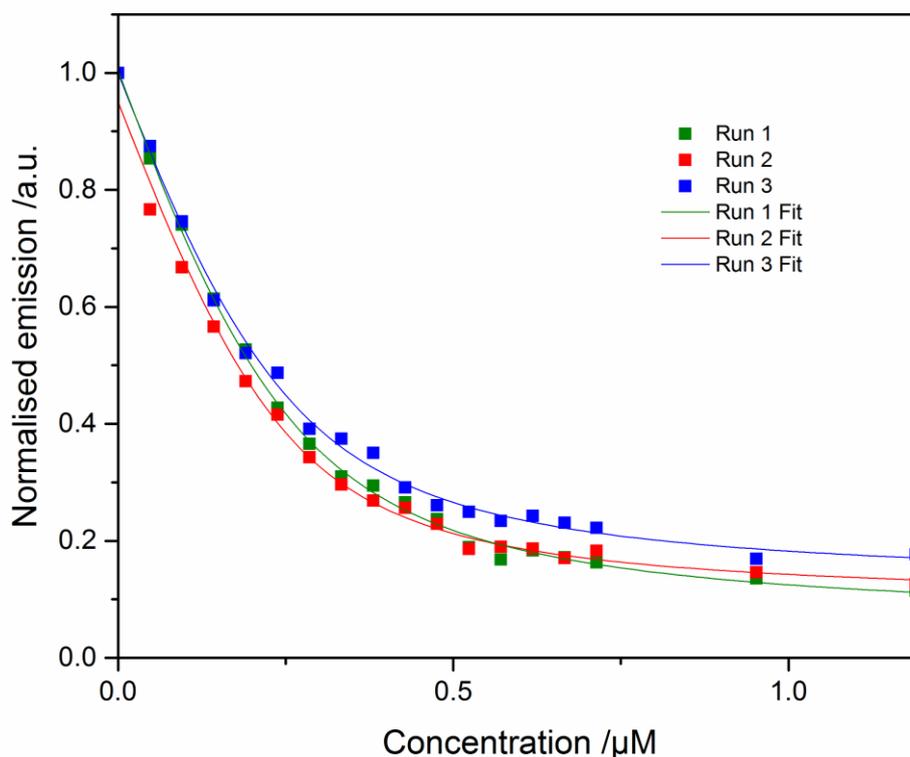


Figure 39: Fluorescence quenching titration curves as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-8-LICAM and their associated non-linear regression fitting data from Dynafit.³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-8-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Table 4: Dissociation constants for CeuE titrated with iron(III)-8-LICAM for three independent fluorescence titrations, and their associated error. The average was then calculated *via* a weighted average method.

CeuE-iron(III)-8-LICAM	K_d/nM	Error /nM
Run 1	62.8	5.0
Run 2	47.2	7.5
Run 3	59.7	6.7
Average	58	8

The dissociation constants could then be compared to the literature values for iron(III)-4-LICAM and iron(III)-5-LICAM (Table 5). All titrations were carried out with 240 nM CeuE in 40 mM Tris-HCl pH 7.5, 150 mM NaCl with aliquots of 12 μM iron(III)-n-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.^{186, 310}

Table 5: Dissociation constants for CeuE with iron(III)-n-LICAM (n = 4, 5, 6, 8) each calculated as a weighted average from three independent fluorescence titrations. Errors for each are calculated as an uncertainty in average.^{310, 322}

Iron(III)-n-LICAM	CeuE K_d/nM	Error /nM
4	21	6
5	< 10	-
6	33	8
8	58	8

It is clear that iron(III)-5-LICAM binds most tightly to CeuE, with a K_d of <10 nM. However, the dissociation constant could not be accurately determined from the recorded data, with the binding being too tight to provide an adequate curve for satisfactory fitting.³¹⁰

Iron(III)-4-LICAM and iron(III)-6-LICAM bind less strongly than iron(III)-5-LICAM, with comparable affinity to each other, given the associated error values, with a K_d of 21 ± 6 nM and 33 ± 8 nM respectively.^{310, 322} Iron(III)-8-LICAM has the weakest affinity of the four, with a K_d of 58 ± 8 nM. The binding curves display the differences in binding affinity, with the steepest curve arising from CeuE-iron(III)-5-LICAM and the shallowest from CeuE-iron(III)-8-LICAM (Figure 40).

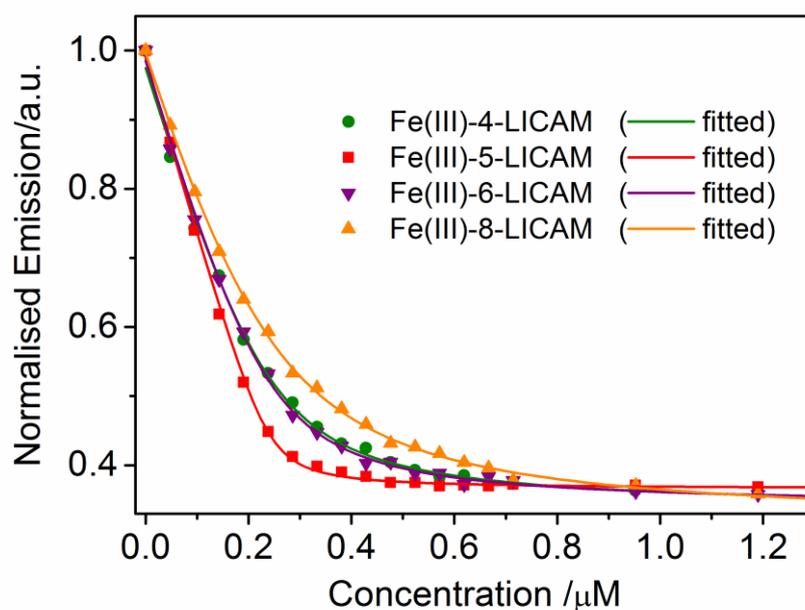


Figure 40: Overlay of selected binding curves for fluorescence quenching titrations of CeuE titrated with of iron(III)-n-LICAM ($n = 4, 5, 6, 8$) and their associated non-linear regression fitting data calculated using Dynafit ($n = 4, 5$ are taken from the literature).^{186, 326}

The binding affinities of CeuE to the iron(III)-n-LICAM ($n = 4, 5, 6, 8$) siderophore mimics are comparable to the binding affinity of iron(III)-bisDHBS reported as 10.1 ± 3.8 nM.⁹⁰ All dissociation constants lie within the nanomolar range, with iron(III)-bisDHBS and iron(III)-5-LICAM, both with a 5 atom linker, binding to CeuE with the highest affinities.

Isothermal titration calorimetry was trialled as an alternative method to attempt to validate the fluorescence quenching titration method for dissociation constant calculation but unfortunately proved unsuitable.^{327, 328} The additional equilibria between iron(III) and n-LICAM ligands in 1:1 and 2:3 ratio, and the necessary addition of iron(III) complexed with NTA to inhibit $\text{Fe}(\text{OH})_3$ precipitation represented a complicated array of equilibria.^{90, 186} As such it was not possible to isolate the specific CeuE to iron(III)-n-LICAM binding event from the thermal effects of all other binding processes.³²⁷ Microscale thermophoresis also gave inconclusive results, likely due to a similar complications of multiple equilibria.³²⁹

The fluorescence quenching titration method overcomes these complications by having the CeuE-based tryptophan as the sole reporter. As such the iron(III) binding equilibria are fluorescence silent, and the CeuE to iron(III)-n-LICAM binding event is the only observable event *via* this method.

2.6 Circular Dichroism

The crystal structures of CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM show that the siderophore analogues are bound in the Λ -configuration, as previously reported for CeuE-iron(III)-4-LICAM, CeuE-iron(III)-bisDHBS and CeuE-iron(III)-MECAM (Figure 41).^{186,90,11, 187}

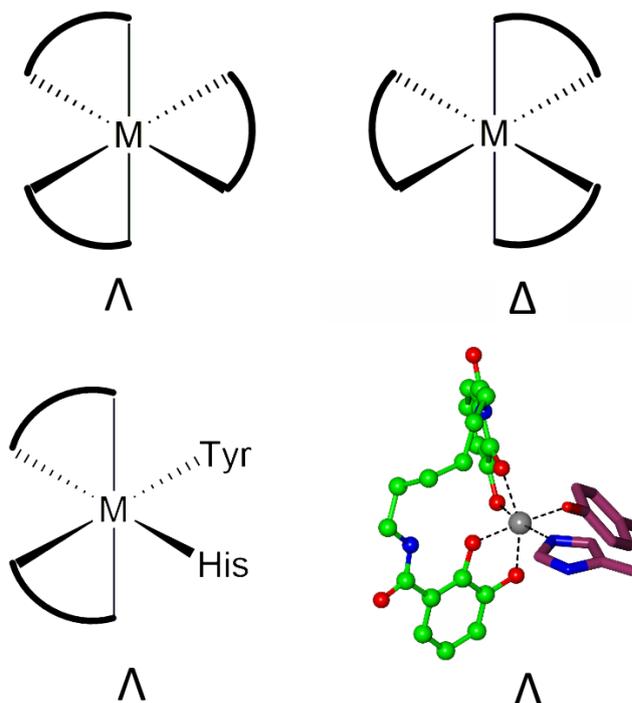


Figure 41: Λ and Δ configurations depicted for a general octahedral metal complex, and the Λ -configuration as is present in a schematic of CeuE-iron(III)-n-LICAM and the crystal structure of CeuE-iron(III)-6-LICAM.^{11, 14} Figure produced using CCP4mg.

Circular dichroism spectroscopy was used to confirm whether the Λ -configuration is also present in solution for CeuE-iron(III)-n-LICAM (n= 5, 6, 8). This was possible as the ligand to metal charge transfer band (LMCT) can be monitored in the visible region from 300 to 700 nm, and the observed signal compared to similar but enantiomerically pure kinetically inert Λ or Δ -configured complexes.^{13, 95, 330}

If a lack of signal is observed in the circular dichroism spectrum from 300 to 700 nm, this indicates no preference for Λ or Δ -configured complexes. When CeuE is analysed with no iron(III) bound ligand, there is no significant observed signal, and the same is true when the iron(III)-n-LICAM (n= 5, 6, 8) is analysed with no protein (Figure 42). This confirms that the protein does not absorb over this wavelength range without iron(III)-n-LICAM (n= 5, 6, 8), and the iron(III)-n-LICAM (n= 6,8) itself has no significant configurational preference,

and is a racemic mixture of both Λ and Δ -configured complexes. Upon introduction of CeuE to iron(III)-*n*-LICAM ($n = 5, 6, 8$), a substantial preference for the Λ -configuration is observed (Figure 42). This confirms that the protein selectively binds the Λ -configuration and drives the equilibrium of free Λ and Δ -configured complexes to a Λ -configured majority.

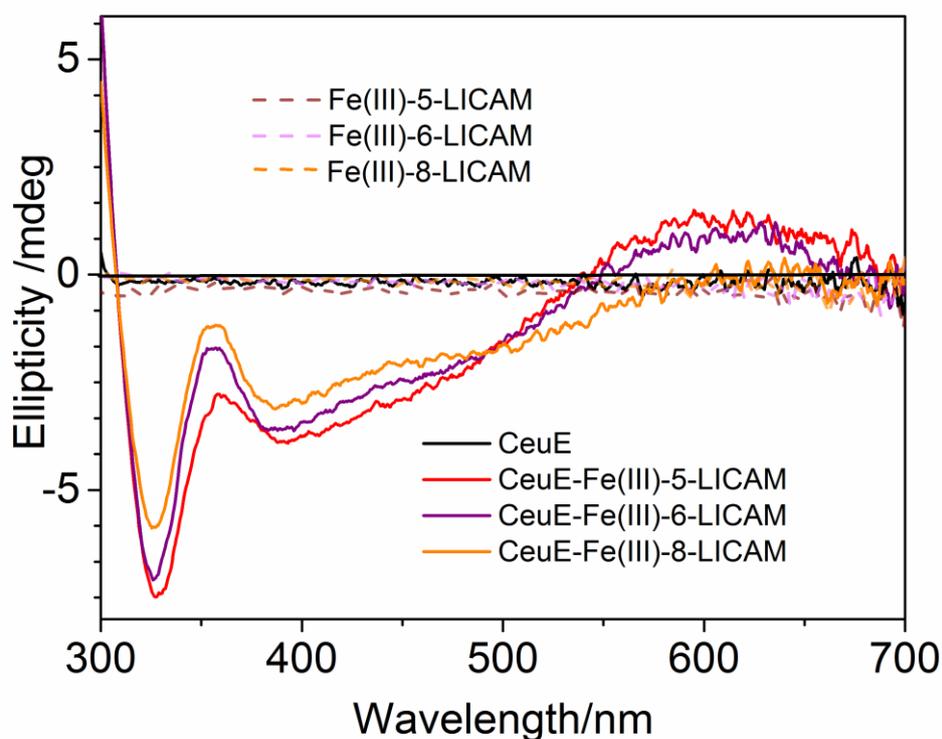


Figure 42: Circular dichroism spectra for CeuE plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra are run for 50 μM CeuE in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μM iron(III)-*n*-LICAM ($n = 5, 6, 8$) ligands as well as the ligands at 50 μM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without CeuE. The spectra show the induction of Λ -configuration upon introduction of CeuE to each iron(III)-*n*-LICAM ($n = 5, 6, 8$).

All CeuE-iron(III)-*n*-LICAM ($n = 5, 6, 8$) circular dichroism spectra were run at the same concentration of 50 μM so spectra could be directly compared to each other. The amplitude of the signal is largest for CeuE-iron(III)-5-LICAM and decreases as the number of spacer atoms increases (Figure 42). This suggests that the binding observed by fluorescence titration can be correlated to the amount of induced Λ -configured complex. The tighter the binding, the more Λ -CeuE-iron(III)-*n*-LICAM ($n = 5, 6, 8$) protein-ligand complex is present in solution, relative to free ligand and protein, and therefore a larger Λ signal is observed.

2.7 The Optimum Linker Length of n= 5

Overall, the CeuE-iron(III)-n-LICAM (n= 4, 5, 6, 8) crystal structures are all very similar, with only subtle differences in overall protein fold and ligand-binding arrangement. The global similarities between structures were quantified by comparing the r.m.s.d in C α positions upon superposition of the structures (Table 6).³¹⁷ The most similar structures are CeuE-iron(III)-4-LICAM and CeuE-iron(III)-5-LICAM with r.m.s.d of 0.38 over 288 C α positions, and all structures containing ligands were more similar to one another than to apo-CeuE, confirming there is very limited structural rearrangement of the protein upon ligand binding.

Table 6: R.m.s.d in Å over number of C α positions for superpositions calculated using the SSM algorithm of CeuE and CeuE-iron(III)-n-LICAM crystal structures showing the similarity in atom positions for all structures.³¹⁷

Crystal Structure	Apo- CeuE	4-LICAM	5-LICAM	6-LICAM	8-LICAM
Apo- CeuE (chain A)		0.68 /286	0.62 /286	0.70 /285	0.63 /286
4-LICAM			0.38 /288	0.48 /287	0.45 /288
5-LICAM				0.58 /254	0.43 /289
6-LICAM					0.54 /288
8-LICAM					

To explain the differences in binding affinity for iron(III)-n-LICAM (n= 4, 5, 6, 8) to CeuE, the crystal structures were overlaid and analysed. Although all four ligands bind in a very similar Λ -configured orientation, there are small differences in precise binding mode. When the structures are viewed with His 227 oriented towards the back, the catecholate units viewed on the right-hand side have atom positions that are almost identical for all four structures. However, differences in the structures arise for the left hand catecholate units. Although the catecholate oxygen atoms are all in very similar positions, the catecholate aromatic ring in the CeuE-iron(III)-4-LICAM structure is displaced upwards relative to the other three structures that contain longer linkers (Figure 43).¹⁸⁶

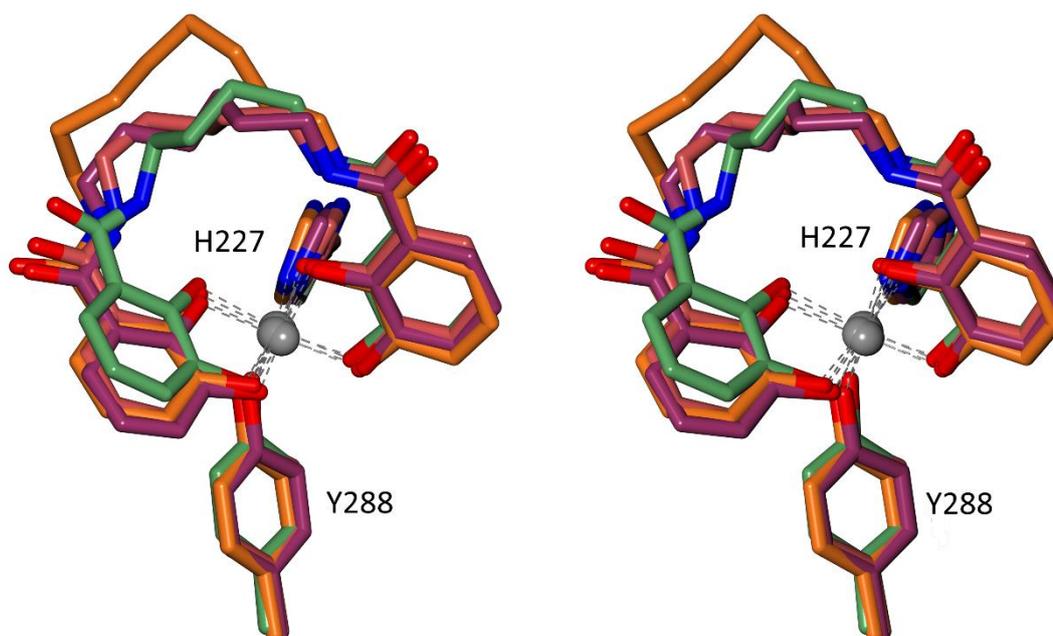


Figure 43: Stereo view overlay of CeuE-iron(III)-n-LICAM crystal structures showing atom positions of the iron(III)-n-LICAM ligands in the CeuE binding pocket and the iron(III) binding residues His 227 and Tyr 288 for each structure. CeuE-iron(III)-4-LICAM (lawn green), CeuE-iron(III)-5-LICAM (light crimson), CeuE-iron(III)-6-LICAM (dark purple) and CeuE-iron(III)-8-LICAM (orange). CeuE-iron(III)-4-LICAM has catecholate ring (left) displaced out of the optimum plane due to linker strain. Figure produced using CCP4mg.

It is likely that this displacement in the catecholate ring seen for iron(III)-4-LICAM is present due to the iron(III)-4-LICAM linker being shorter than is optimum for the ideal binding arrangement. Although the octahedral iron(III) coordination is possible, there is strain in the linker of 4-LICAM. This difference can be quantified by measuring the interplanar angle between the catecholate aromatic rings. Using the atom positions of each ligand in each crystal structure, planes for each ring were generated and the angle between planes calculated using Mercury (Figure 44).³³¹

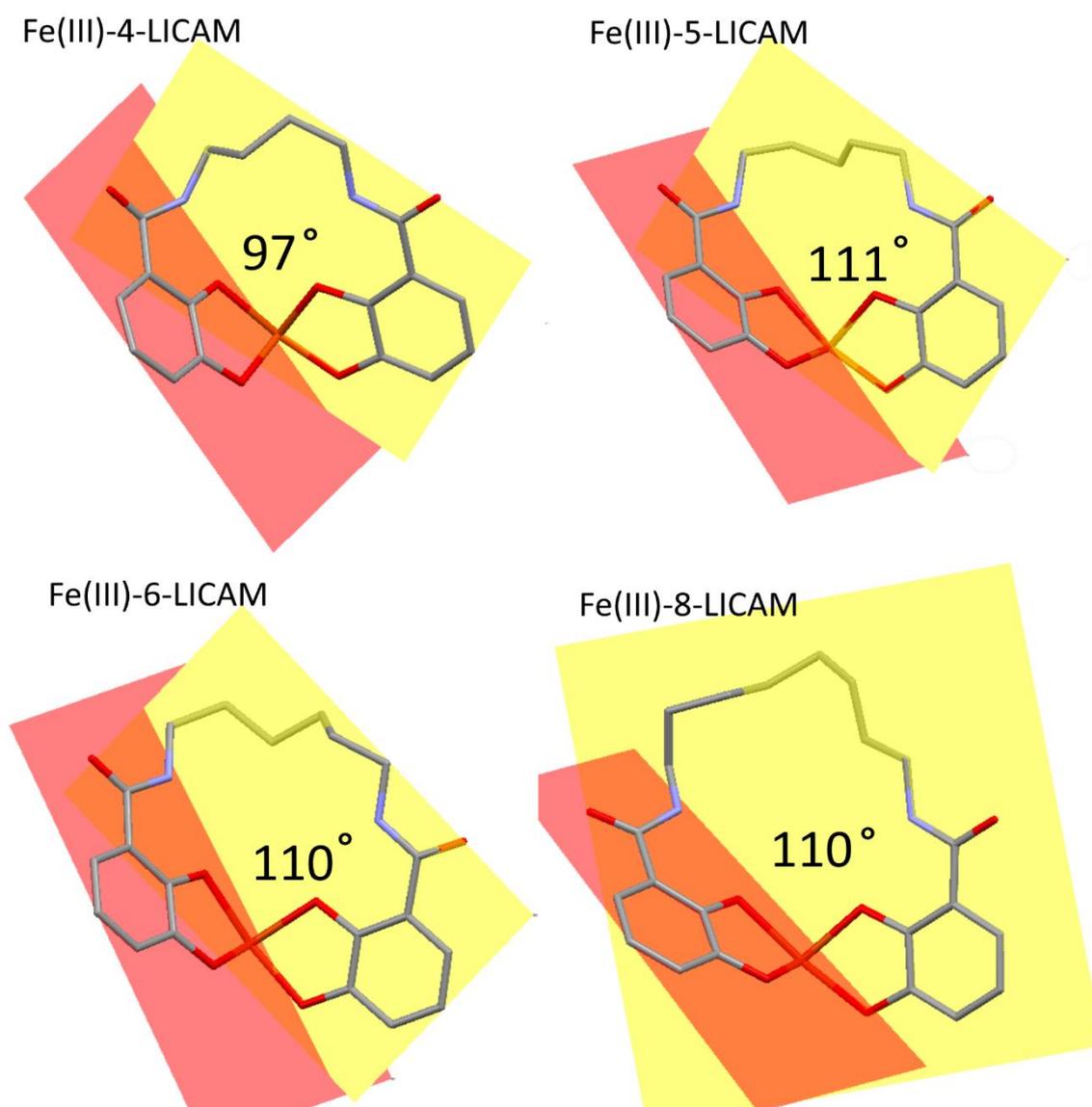


Figure 44: Planes constructed for the six carbon atoms of each catecholate aromatic ring displayed in red (left) and yellow (right). Atom coordinates from CeuE-iron(III)-*n*-LICAM ($n = 4, 5, 6, 8$) crystal structures for each ligand are displayed in cylinders coloured by atom type. Interplanar angles are calculated between catecholate aromatic rings for each structure. Images created in Mercury.³³¹

For CeuE-iron(III)-4-LICAM, the angle is 97° , but for iron(III)-*n*-LICAM ($n = 5, 6, 8$) the angles are 111° , 110° and 110° , respectively. It appears that the ideal angle is around 110° , and a smaller angle is forced when the linker is too short to achieve the optimum. This helps to explain why iron(III)-4-LICAM binds more weakly to CeuE than does iron(III)-5-LICAM. Iron(III)-5-LICAM is better suited for enthalpic and geometric requirements of the iron(III) centre as well as the CeuE binding pocket.

Although iron(III)-n-LICAM (n=6, 8) are able to satisfy the enthalpic requirements, and allow for the optimum catecholate interplanar angle, when the linker region is extended, the binding affinity for CeuE weakens. This is likely due to the entropic cost of fixing a longer linker with an increased number of degrees of freedom. Protein-ligand binding entropy can be estimated using computational calculations.³³²⁻³³⁴ For the purpose of calculations, the overall entropic cost of ligand binding is often separated into conformational and vibrational entropy. The entropic cost of ligand binding is likely due to fewer possible ligand rotamers, and is therefore principally conformational entropy. There will however also be contributions from the changes in vibrational entropy.³³² The calculations are however very complicated when also taking the flexibility and motion of the protein into account.^{332, 334} For the CeuE system, the protein does not have a large conformational change upon ligand binding, and the protein conformation is very similar when each iron(III)-n-LICAM (n=4, 5, 6, 8) are bound. This is shown by the high similarity in atom position upon superposition of the crystal structures (Table 6). It is therefore likely that the majority of the difference in conformational entropy of the system is caused by the ligand rather than the protein, and a longer linker comes with a larger entropic cost.

It seems that the shortest linker of five atoms in iron(III)-5-LICAM that achieves the ideal interplanar angle without strain is the tightest binder due to a balance of enthalpic and entropic optimisation. Unsurprisingly, the known natural ligand, iron(III)-bisDHBS, has a linker region containing five atoms.⁹⁰ The binding affinities of the longer linkers are however within an acceptable biological range: periplasmic chaperone proteins are commonly known to have submicromolar binding affinities for their natural substrate.³³⁵ These findings therefore may indicate that it is possible for CeuE to bind natural iron(III)-bound siderophores of longer linker length. This means there is a good chance that *C. jejuni* is able to acquire iron(III) from a more structurally diverse set of tetradentate catecholate siderophores, other than just the enterobactin hydrolysis product bisDHBS. This could give *C. jejuni* a competitive advantage over a wide number of tetradentate siderophore producing species, or conversely may allow for wider cooperative virulence of *C. jejuni* with a larger range of bacterial species than first thought.¹¹³

2.8 Summary and Conclusion

It was shown that CeuE can bind iron(III)-*n*-LICAM siderophore mimic complexes with linker lengths of $n=6$ and $n=8$, which adds to findings already published in the literature that it can bind a range of linker lengths including $n=4$ and $n=5$.^{186, 310} The CeuE iron(III)-*n*-LICAM complexes adopt the Λ -configuration in both crystal and solution phase, proven by protein crystallography and solution phase circular dichroism. Fluorescence quenching titrations with CeuE showed that the tightest binding complex is iron(III)-5-LICAM with a K_d of <10 nM. Iron(III)-4-LICAM and iron(III)-6-LICAM bind more weakly than iron(III)-5-LICAM with K_d values of 21 ± 6 nM and 33 ± 8 nM, respectively. The weakest binder is iron(III)-8-LICAM with a K_d of 58 ± 8 nM. The interplanar angles obtained from the crystal structures show that CeuE-iron(III)-4-LICAM does not possess the optimum geometry around the iron(III) centre, and this is suggested to be caused by strain from having the shortest linker length. Iron(III)-6-LICAM and iron(III)-8-LICAM are suggested to bind less tightly than iron(III)-5-LICAM due to the entropic cost of fixing a longer linker in the CeuE binding pocket. Iron(III)-5-LICAM with the highest binding affinity to CeuE is the closest mimic of iron(III)-bisDHBS, the natural siderophore of CeuE.⁹⁰

2.9 Future Work

To complete the study, a good starting point would be to test the limits of linker lengths that CeuE is able to bind. To do this, the n-LICAM series should be pushed to extreme linker lengths. To test the shortest linker, it would be necessary to synthesise and repeat studies with 3-LICAM. It is likely that this linker length would either be too short to establish the necessary octahedral binding arrangement in the CeuE binding pocket, or the formation of 2:3 Fe: ligand complexes would be much more stable than the 1:1 complexes required for CeuE binding. As such, it is likely that the binding affinity of iron(III)-3-LICAM to CeuE would be dramatically diminished. It would be interesting to test this hypothesis.

It would then be of interest to try a series of longer linkers. As the amonabactins can contain a linker length of 16 atoms, 16-LICAM would be an key choice of ligand to establish the upper limit of linker length tolerance of CeuE.¹⁰⁹

CeuE is now known to bind tetradentate catecholate siderophore mimics of increased backbone length, and this poses the question as to whether CeuE is able to bind analogous natural tetradentate siderophores such as serratiochelin and the amonabactins. The above studies could be repeated with these natural siderophores to investigate whether the binding affinities are comparable by fluorescence quenching titration. It would then be of interest, if these siderophores did have a biologically relevant binding affinity, to study whether they were able to bind to CeuE with the observed Λ -configured binding arrangement both in crystal and solution phase. Protein crystallography with these ligands would establish the role of His 227 and Tyr 288 and other neighbouring residues in the binding configuration with these new natural siderophores.

It would then be of interest to determine whether any natural tetradentate siderophores promote growth of *C. jejuni* under iron limited conditions. This study would validate whether any interactions of CeuE with the iron(III) bound siderophores were of biological relevance. Studies could be extended to the use of *C. jejuni* mutants that are unable to produce CeuE, to confirm whether CeuE is essential for the uptake of the iron(III) bound siderophores for effective growth of the species under iron limited conditions.

This study could be extended to the n-LICAM siderophores, to probe whether these siderophore mimics are able to be used in the full CeuBCDE iron(III)-siderophore uptake system for delivery of iron(III) to the cytoplasm.

The high binding affinity of iron(III)-siderophore complexes to CeuE make this system an attractive starting point for the development of artificial metalloenzymes (Figure 45).³³⁶

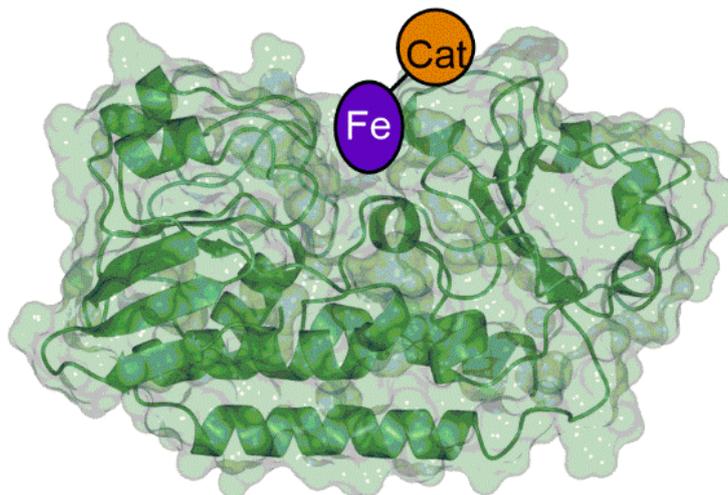


Figure 45: Schematic diagram of CeuE as a potential scaffold for an artificial metalloenzyme. CeuE is shown in green ribbons. The iron(III) siderophore complex is shown in the binding cleft in purple, with a linked inorganic catalyst labelled 'cat' in orange. Figure produced using CCP4mg.

The knowledge that CeuE is able to bind a range of linker lengths allows for a range of catalyst-siderophore anchor designs that do not need to be limited to a 5-atom linker. This knowledge may prove important for the successful attachment of an inorganic catalyst to a siderophore backbone for anchorage in the CeuE binding cleft, or to optimising function of the artificial metalloenzyme, by positioning a catalyst in a specific orientation in the CeuE binding cleft.

3 The study of CeuE Variant Proteins for Investigation of Structure-Function Relationships in the CeuE Binding Cleft

3.1 Introduction

Residues that contribute to iron(III)-siderophore binding include a basic triad of arginine residues, Arg 118, Arg 205 and Arg 249, that form hydrogen bonds with the deprotonated catecholate oxygens, and a lysine, Lys 121, that provides a hydrogen bond to the carbonyl oxygen of the catecholamide backbone as detailed in the literature and in Chapter 2.^{90, 186, 187, 310} Crucially, for tetradentate iron(III)-siderophore binding (iron(III)-n-LICAM and iron(III)-bisDHBS), a histidine, His 227, and tyrosine, Tyr 288 coordinate directly to the iron(III) centre (Figure 46).^{90, 186, 310} To better understand the binding of iron(III)-complexes in the CeuE binding cleft, the structure-function relationships of these key residues must be studied in greater detail.

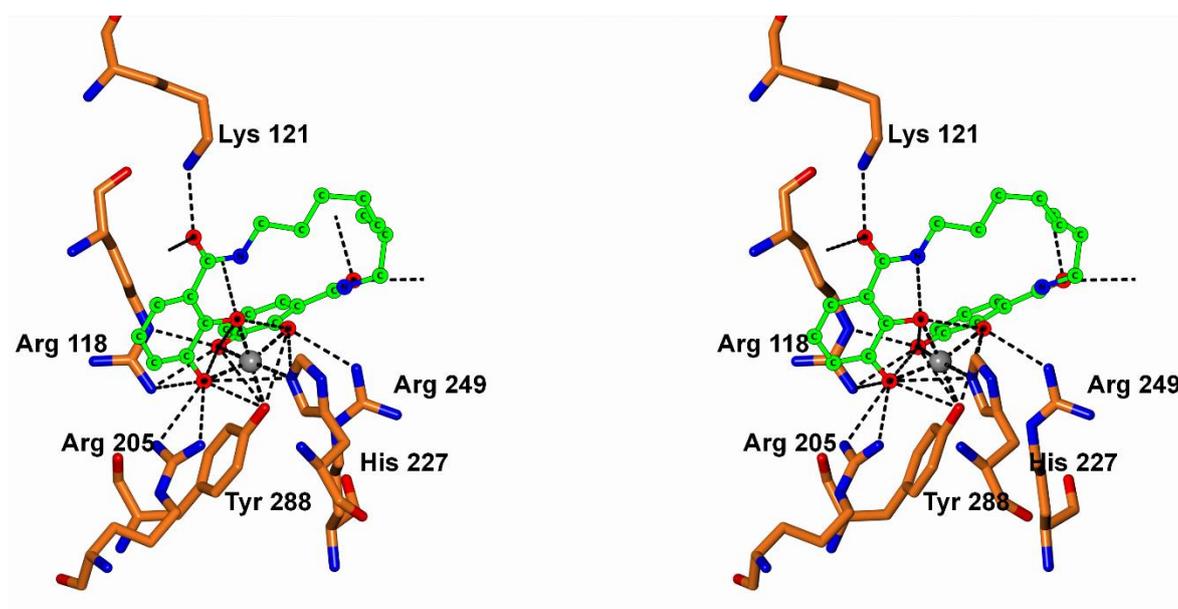


Figure 46: Structure of CeuE-iron(III)-8-LICAM in stereo view (PDB: 5AD1) showing residues involved in iron(III)-n-LICAM binding Arg118, Arg205, Arg 249, Lys 121, His 227 and Tyr 288. Residues involved in binding are shown in cylinders coloured by atom type. Iron(III) is shown as a grey sphere. 8-LICAM is shown in circles coloured by atom type. Figure produced using CCP4mg.

Residues involved in iron(III) binding in CeuE are conserved across a range of iron(III)-siderophore periplasmic binding proteins from other bacterial species.^{90, 186, 337} Previously, the Dali server was used to identify periplasmic binding proteins that are structurally similar to CeuE.^{90, 338} To follow up this study and provide an up-to-date assessment, a structural match search was conducted with PDBeFold.^{317, 339-341} 387 residues from chain A of the apo

CeuE crystal structure (PDB ID: 3ZKW) were used in a search model against structures deposited in the PDB with 70% or greater secondary structural similarity. 21 distinct proteins were found, and the most homologous structures, at above 85% secondary structure similarity, included YclQ (PDB ID: 3GFV) from *Bacillus subtilis*,²³¹ an unnamed PBP (PDB ID: 4MX8) from *Xylanimonas cellulosilytica*, PiuA (PDB ID: 4JCC) from *Streptococcus pneumoniae* and VctP (PDB ID: 3TEF) from *Vibrio cholerae*.²⁰⁹ These four proteins were previously identified, suggesting no crystal structures of very closely matching (above 85% SSE match) proteins to CeuE have been solved in the last two years.⁹⁰ The structures of CeuE and the four identified proteins were superposed using the SSM algorithm (Figure 47).³¹⁷ All structures had the characteristic type III periplasmic binding protein extended α -helix hinge region, and an overall bilobate shape.^{184, 185}

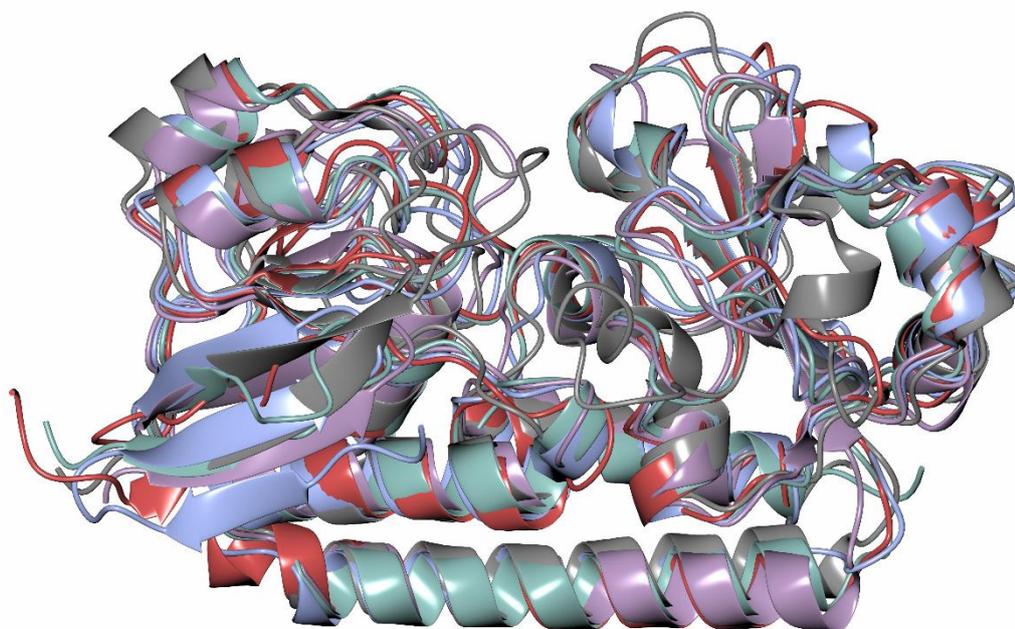


Figure 47: Overlay of crystal structures of CeuE (ice blue) (PDB ID: 3ZKW), YclQ (sea green) (PDB ID: 3GFV), PiuA (light crimson) (PDB ID: 4JCC), VctP (lilac) (PDB ID: 3TEF) and an unnamed PBP (dark grey) (PDB ID: 4MX8) by the SSM algorithm.³¹⁷ Protein backbones are shown in ribbons. Figure produced using CCP4mg.

Sequence alignments revealed that there were a number of conserved residues across these structurally similar proteins, including the His and Tyr involved in iron(III) binding in CeuE, as well as two of the Arg residues involved in hydrogen bonding in CeuE-iron(III)-tetradentate catecholate complexes (Figure 48).

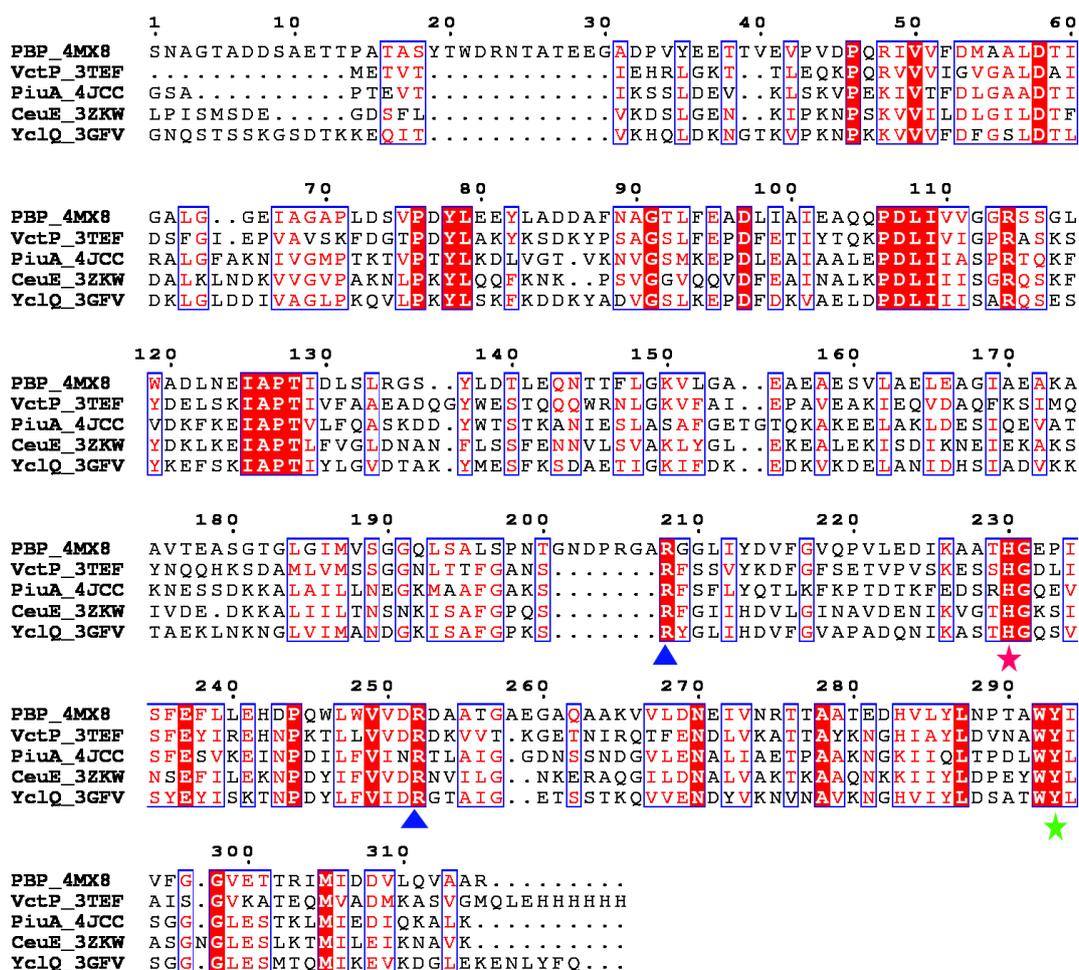


Figure 48: Sequence alignment of amino acid sequences for proteins superposed in Figure 47. Red blocks are totally conserved, red text indicates close matches or residues of similar properties. Black text shows non-matching residues. Conserved arginine residues are shown with a blue triangle, conserved histidine with a pink star, conserved tyrosine with a green star.

Using CCP4mg, the binding cleft regions were superposed with the apo CeuE crystal structure (PDB ID: 3ZKW) *via* the SSM algorithm, to identify whether the potential iron(III) binding histidine and tyrosine residues were in similar positions in all structures.³¹⁷ The four structures, their structural similarity by secondary structure elements in % to CeuE, and the numbering of their conserved histidine and tyrosine residues are detailed in Table 7. The position of the conserved histidine and tyrosine residues in the crystal structures are all very similar to each other for CeuE YclQ, PiuA and VctP (Figure 49). The greater variation in histidine position, relative to the tyrosine position, is likely due to the histidine residue being located on a more mobile loop region of each protein. For the unnamed PBP (PDB ID: 4MX8), the conserved histidine is situated away from the binding cleft (Figure 50).

Table 7: Proteins identified with high structural similarity to CeuE *via* secondary structure element (SSE) matching by PDBeFold. All contain the conserved histidine and tyrosine residues.

PBP	Organism	PBD ID	% SSE match to CeuE	SSM match to CeuE in Å / residues matched	His	Tyr
YclQ	<i>Bacillus subtilis</i>	3GFV	89	1.23 /272	H214	Y275
PiuA	<i>Streptococcus pneumoniae</i>	4JCC	88	1.53 /266	H238	Y300
VctP	<i>Vibrio cholerae</i>	3TEF	100	1.78 /263	H248	Y310
PBP-2	<i>Xylanimonas cellulosilytica</i>	4MX8	85	1.64 /258	H228	Y291

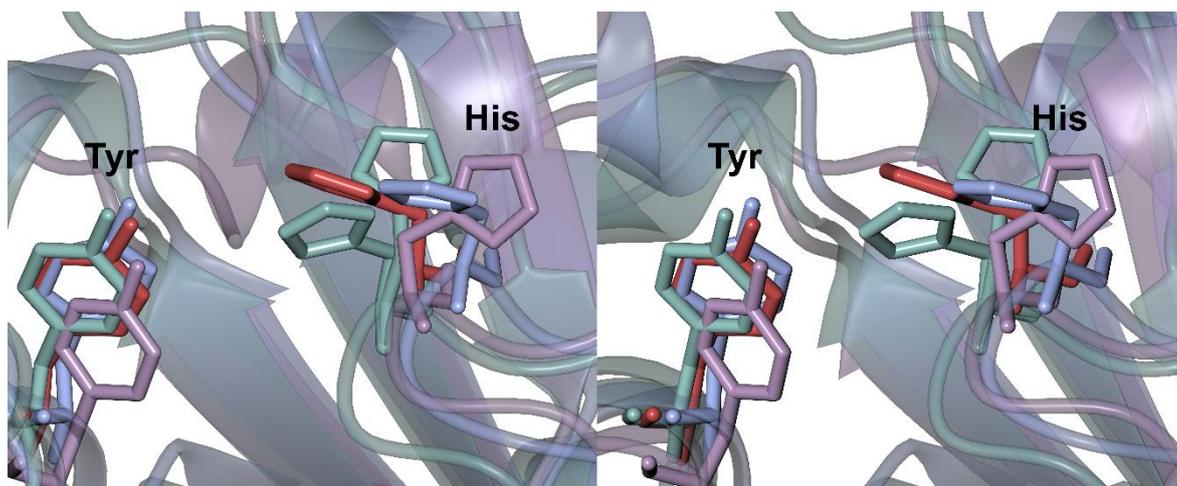


Figure 49: Overlay of crystal structures of CeuE (ice blue) (PDB ID: 3ZKW), YclQ (sea green) (PDB ID: 3GFV), PiuA (light crimson) (PDB ID: 4JCC) and VctP (lilac) (PDB ID: 3TEF) by the SSM algorithm in stereo view.³¹⁷ Protein backbones are shown in ribbons. Conserved Tyr and His residues that are expected to be involved in iron(III)-coordination for these proteins are shown in cylinders. Figure produced using CCP4mg.

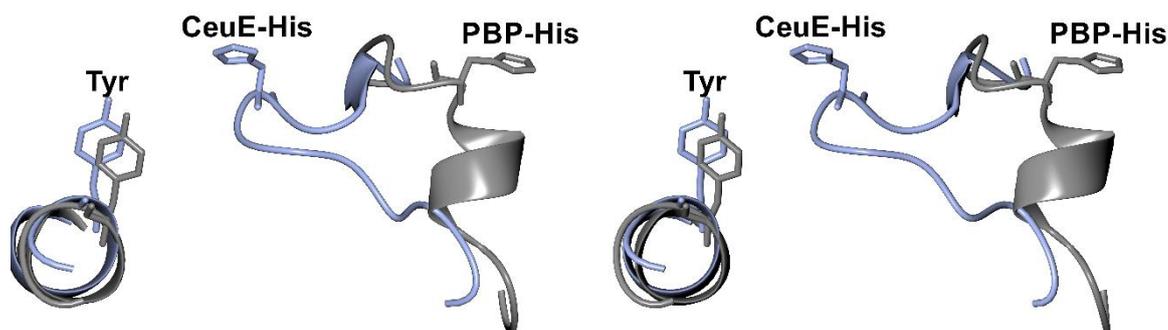


Figure 50: Overlay of crystal structures of CeuE (ice blue) (PDB ID: 3ZKW) with unnamed PBP (grey) (PDB ID: 4MX8) in stereo view. Selected regions of protein backbone are shown as ribbons with the conserved His and Tyr shown in cylinders. The His of the unnamed PBP is folded away from the binding cleft. Figure produced using CCP4mg.

Interestingly, three out of the four identified proteins were from Gram-positive bacteria, with only VctP from a Gram-negative species *V. cholerae*. As Gram-positive species have a single cell membrane, the periplasmic binding protein equivalent, termed the substrate-binding protein, is often anchored in the cytoplasmic membrane, in close proximity to the ABC-transporter to which it delivers iron(III)-siderophore complexes.²³¹ This is also the case for CeuE, as although the protein is located in the periplasmic space, it is a lipoprotein that is tethered to the cytoplasmic membrane. Lipoprotein character was confirmed by a characteristic signal sequence as well as post-translational modification with palmitic acid when expressed in *E. coli*.^{171, 342} It is thought that similar post-translational modification allows anchorage of the protein to the cytoplasmic membrane in *C. jejuni*.¹⁷¹ The structural similarity of CeuE to Gram-positive analogues may be explained by this similarity in function -with the need for membrane anchorage.

Many more putative periplasmic binding proteins, containing the conserved histidine and tyrosine residues, were found by protein sequence alignment, but the 3D structures of these are yet to be solved. Many were from Gram-positive species, and many of these were from the Bacillaceae family. In addition, there were a number of species of medical relevance that are likely to have a tetradentate catecholate siderophore binding protein with conserved His and Tyr residues (Figure 51). Species included *Haemophilus influenzae*, commonly known as bacterial influenza, *Chlamydia trachomatis*, the species responsible for human chlamydia infection, and *Akkermansia muciniphilia*, a species thought linked to obesity and diabetes.³⁴³⁻
³⁴⁶ *Staphylococcus aureus*, *Enterococcus faecium* and *Enterococcus raffinosus* were also identified to express a protein with sequence similarities for CeuE. These are all species that are known to pose a risk of antibiotic resistance (Figure 51).³⁴⁷⁻³⁵⁰

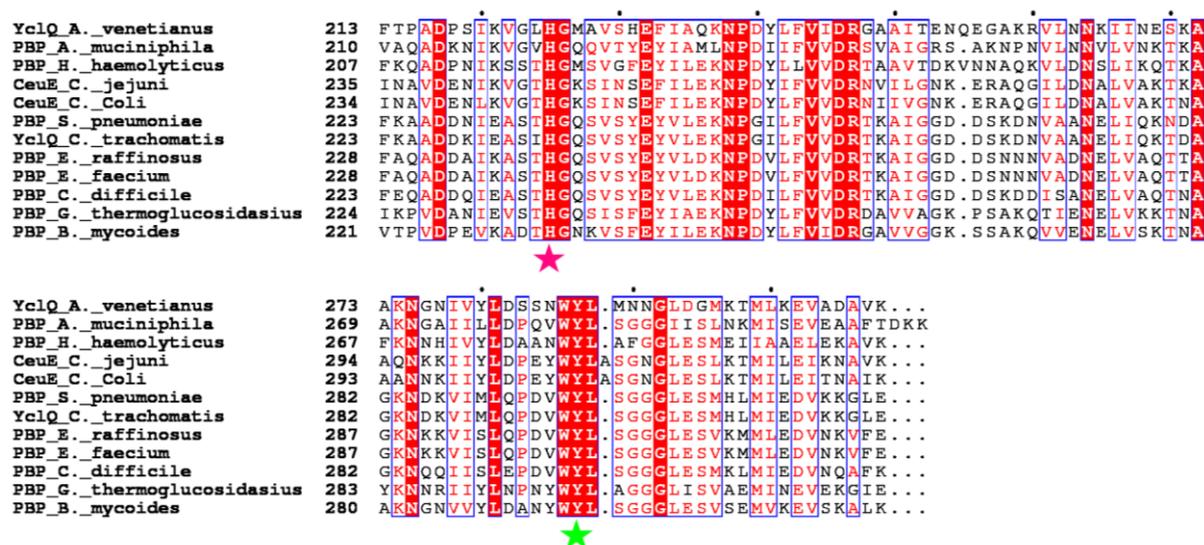


Figure 51: Sequence alignment for proteins similar to CeuE from a range of bacterial species. Residues 235-345 are shown, with reserved His and Tyr shown as residue 260 and 320 as a pink star and green star, respectively. The structures of these proteins are currently unsolved. Red blocks are totally conserved across all proteins, red text indicates close matches or residues of similar properties. Black text shows non-matching residues.

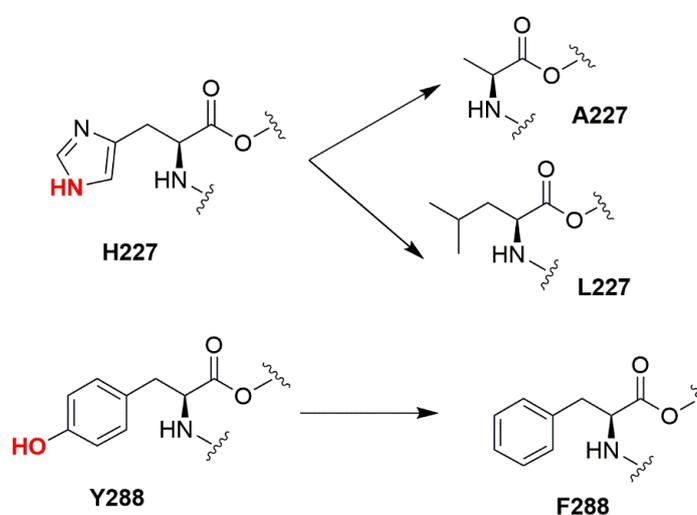
It is of interest to study the role of His 227 and Tyr 288 residues in CeuE, as the results may reveal a general siderophore-binding mode for this identified range of similar putative periplasmic binding proteins from different pathogenic bacterial species. CeuE thus provides a model for understanding this whole subfamily of periplasmic binding proteins. It may then be of medical interest to use these proteins as a tool to gain a better understanding of the iron uptake capacities of the above species, and understanding the iron uptake pathways of pathogenic species may lead to more effective design of new antibiotics.^{295, 299, 301} In some cases, these proteins may even prove suitable in the development of vaccines.³⁵¹

A number of thermophilic species including *Geobacillus thermoglucosidasius* were identified in the sequence alignment. *Geobacillus thermoglucosidasius* is a thermophilic Gram-positive species that is capable of growth up to around 70 °C, with an optimum range of 61-63 °C.³⁵²⁻³⁵⁴ A protein analogous to CeuE from a thermophilic species may prove to be more thermostable than CeuE. Thermostable proteins are desirable candidates for the protein scaffold of potential artificial metalloenzymes, as this would allow for catalysis over a wider temperature range, and therefore have potential for greater catalytic turnover frequencies than those possible at lower temperatures.³⁵⁵⁻³⁵⁷ These species are already known for their applications in thermostable biocatalysis, such as in the production of bioethanol at 60 °C and isobutanol at 50 °C.^{352, 358, 359}

The work detailed in this chapter aims to investigate the individual roles of key residues His 227 and Tyr 288 in the CeuE binding cleft. The objective was to study CeuE variant proteins, deficient in each of these residues, with tetradentate siderophore mimics iron(III)-n-LICAM (n =4, 5, 6, 8). Using the same experimental techniques that were used in the studies in Chapter 2, including protein crystallography, fluorescence titration, and circular dichroism, it was possible to directly compare variant CeuE proteins and wild-type CeuE. Differences in iron(III)-n-LICAM (n =4, 5, 6, 8) binding could then be used to infer the individual structure-function relationships of His 227 and Tyr 288. This information may provide a model for understanding a whole subfamily of periplasmic binding proteins, and could inform the development of new antibiotics or thermostable artificial metalloenzymes.

3.2 CeuE Variant Proteins

CeuE variants were designed without the functional regions of residues His 227 and Tyr 288, replacing them with amino acids with no iron(III)-coordinating atoms in their side chains. Histidine was replaced with alanine and leucine, and tyrosine replaced with phenylalanine (Scheme 2). Alanine, with a CH₃ sidechain was a non-coordinating alternative to histidine. Leucine, with a bulkier aliphatic sidechain was a closer match to histidine in terms of steric bulk, so it was decided the two mutations would be compared to investigate any influence of sterics on iron(III)-n-LICAM binding. Phenylalanine provided similar steric bulk and aromatic rigidity to tyrosine, but without the hydroxy group this residue was rendered unable to participate in iron(III)-coordination.



Scheme 2: Chemical structures of the amino acids used in mutations of His 227 and Tyr 288 in CeuE variants. Atoms that coordinate to iron(III) after deprotonation are shown in red.

Site-directed mutagenesis is a common technique used for determining structure-function relationships of significant residues in proteins. Systematic methods can be employed when functional residues are not already known, and this is particularly helpful when structural information cannot easily be obtained.³⁶⁰⁻³⁶³ Database-derived software SIFT can even be used to predict whether a mutation might affect protein function.³⁶⁴

CeuE variant proteins were expressed, purified, and characterised by Dr E. V. Blagova. The necessary mutations were made by PCR-based site directed mutagenesis. Firstly, three mutant proteins were produced with the mutations: H227L, H227A and Y288F. A second cycle of PCR produced the double mutants H227L/Y288F and H227A/Y288F. All mutations were confirmed by DNA sequencing. The mutations were then confirmed in the expressed proteins by electrospray ionisation mass spectrometry. Circular dichroism spectroscopy was performed to confirm the mutants were folded correctly.⁹⁰

3.3 Crystal Structures of Apo CeuE Variant Proteins

CeuE variants were crystallised in their apo form and cryoprotected by Dr E. V. Blagova. Crystallisation procedures are detailed in Chapter 6. Data for the following structure solution were collected at the Diamond Light Source and structure solution and refinement procedures are detailed in Chapter 6. All variants, with the exception of CeuE-Y288F, crystallised in space group P1, and were isomorphous with the wild-type, with three independent molecules in the asymmetric unit (Table 8). As might be expected, the mutants adopted an essentially identical fold to that of apo wild-type CeuE (Figure 52) and residues 227 and 288 were located in very similar locations in all mutants, with an open binding cleft in the absence of ligands (Figure 53).

Table 8: Comparison of CeuE variant crystal structures to that of wild-type CeuE. R.m.s.d is given for superposition with the wild-type structure *via* the SSM algorithm and the number of C α positions matched is given in parentheses.

Mutation	Spacegroup	Cell parameters (Å, °)	R.m.s.d (Å)	PDB ID
Wild-type	P1	a= 56.95 b= 62.74 c= 67.98 α = 82.19 β = 76.74 γ = 75.96	-	3ZKW ¹⁸⁶
H227L	P1	a= 56.92 b= 62.56 c= 67.79 α = 82.21 β = 76.99 γ = 76.18	0.27 (860)	5LWQ
H227A	P1	a= 56.90 b= 62.61 c=67.79 α = 82.11 β = 77.21 γ = 76.28	0.90 (848)	5MBQ
Y288F	P3 ₂ 21	a= 65.52 b= 65.52 c= 145.66 α = 90.00 β = 90.00 γ =120.00	0.68 (284)	5LWH
H227L/Y288F	P1	a= 56.72 b=62.36 c= 67.71 α = 82.73 β = 77.33 γ = 76.54	0.54 (840)	5MBT
H227A/Y288F	P1	a= 56.90 b= 62.61 c= 67.79 α = 82.23 β = 76.96 γ = 75.98	0.38 (847)	5MBU

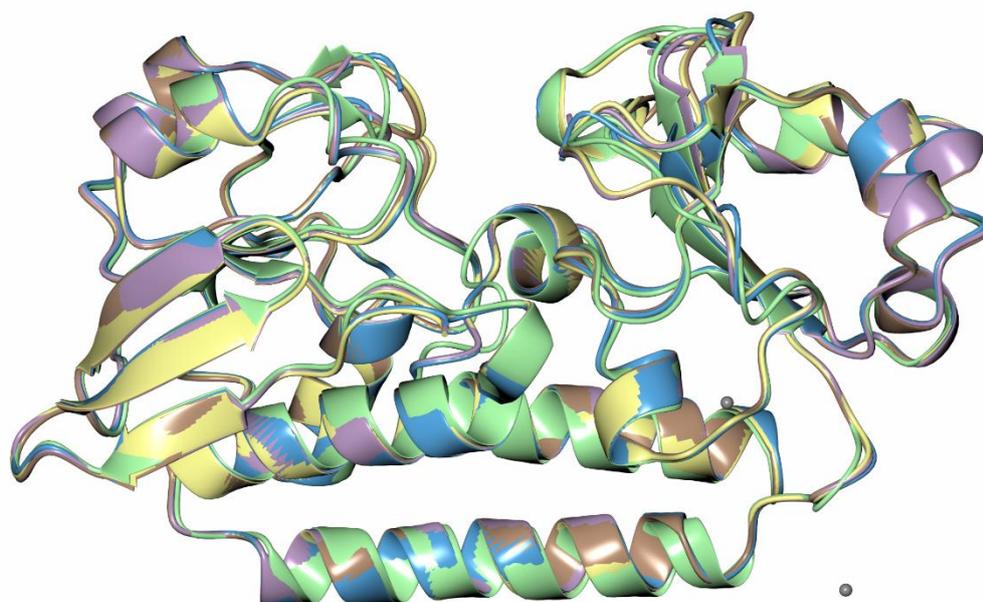


Figure 52: Overlay of the A chain of all mutant structures, showing high similarity in overall fold. H227L is shown in lemon ribbons, H227A in light brown, Y288F in light green with Zinc(II) ions shown as grey spheres, H227L/Y288F in light blue and H227A/Y288F in lilac. Figure produced using CCP4mg.

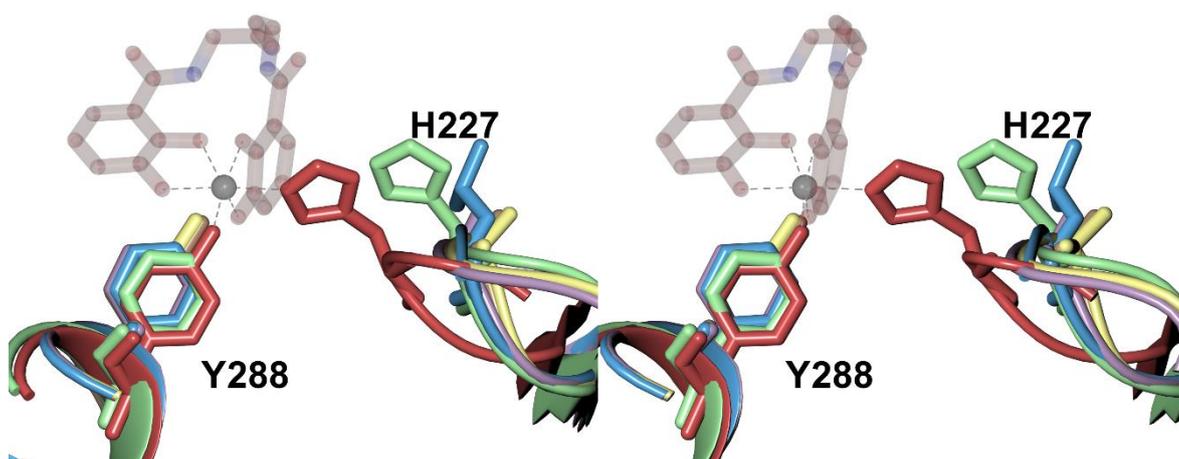


Figure 53: Stereo view of overlay of the A chain of all mutant structures with that of wild-type CeuE bound to iron(III)-5-LICAM, showing residues 227 and 288 in cylinders. There is little change in the arrangement of the 227-loop region for any of the mutants, but when iron(III)-5-LICAM is bound, the loop region is shifted towards the iron(III) centre. H227L is shown in lemon ribbons, H227A in light brown, Y288F in light green with zinc(II) ions shown as grey spheres, H227L/Y288F in light blue and H227A/Y288F in lilac. CeuE-

iron(III)-5-LICAM is shown in light crimson, with the ligand shown in translucent cylinders. Figure produced using CCP4mg.

Structure Solution for CeuE-H227L

The structure of CeuE-H227L was well ordered with an average B value of 25.7 \AA^2 for the protein. The most similar chain to apo wild-type CeuE was chain B with r.m.s.d of 0.14 \AA over 287 $\text{C}\alpha$ positions. The CeuE-H227L binding cleft contained good electron density for iron(III)-chelating residue Tyr 288 and for Arg 205. There was some electron density for the sidechains of Arg 118 and Arg 249 which act as stabilising residues (Figure 54).^{90, 186} The mutated H227L was on a moveable loop, and while there was good electron density for the $\text{C}\alpha$ backbone of the loop region, the electron density for side chains in this region was weak, indicating that they were flexible and disordered. There was a no significant electron density for the side chain of Leu 227 (Figure 55).

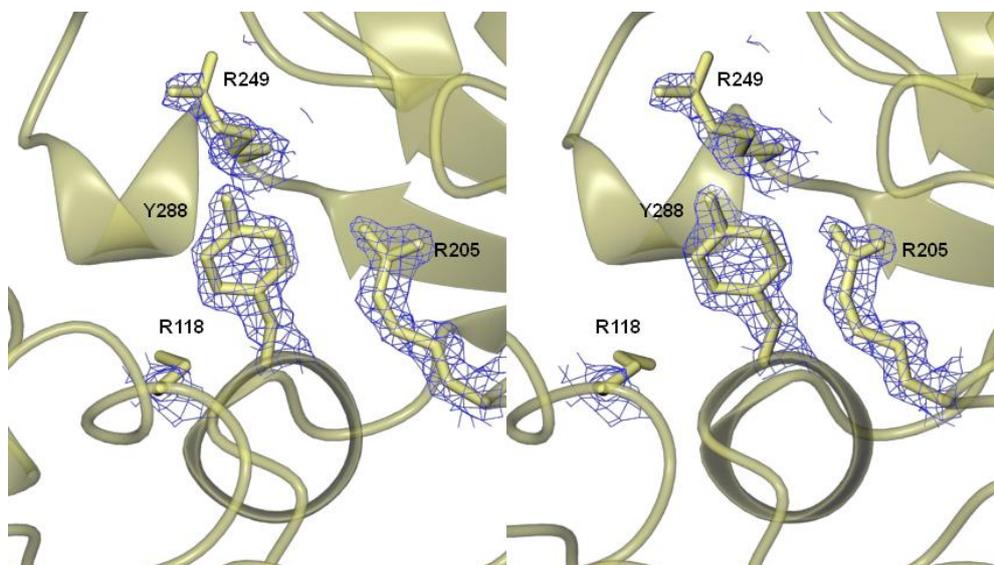


Figure 54: Stereo view of iron(III)-siderophore binding residues Tyr 228, Arg 118, Arg 205 and Arg 249 in the binding cleft of chain A of CeuE-H227L. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.

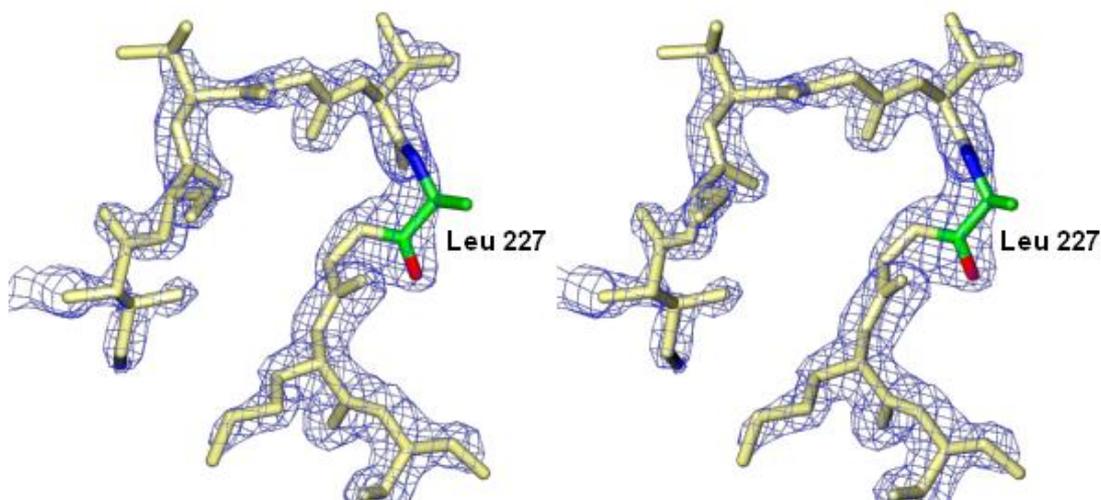


Figure 55: Stereo view of loop region in chain A of CeuE-H227L from residues 222 to 230. Residues 222-226, 288-230 shown in lemon cylinders. Leu 227 is shown in cylinders coloured by atom type. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. There is poor electron density for the Leu 227 sidechain. Figure produced using CCP4mg.

Structure Solution for CeuE-H227A

The structure of CeuE-H227A was solved as for CeuE-H227L, and the average B value for the whole structure was 25.2 \AA^2 . Some flexible loop regions of the protein (detailed in Chapter 6) could not be modelled, as there was not sufficient electron density present in these regions, indicating that these residues were mobile in the crystal. Residues Tyr 288, Arg 205 and Arg 249 were well ordered in the iron(III)-siderophore binding cleft (Figure 56). As in CeuE-H227L, the Ala 227 residue located on the moveable loop region of the CeuE binding cleft possessed good electron density for the $C\alpha$ protein backbone, but the sidechains were disordered (Figure 57).

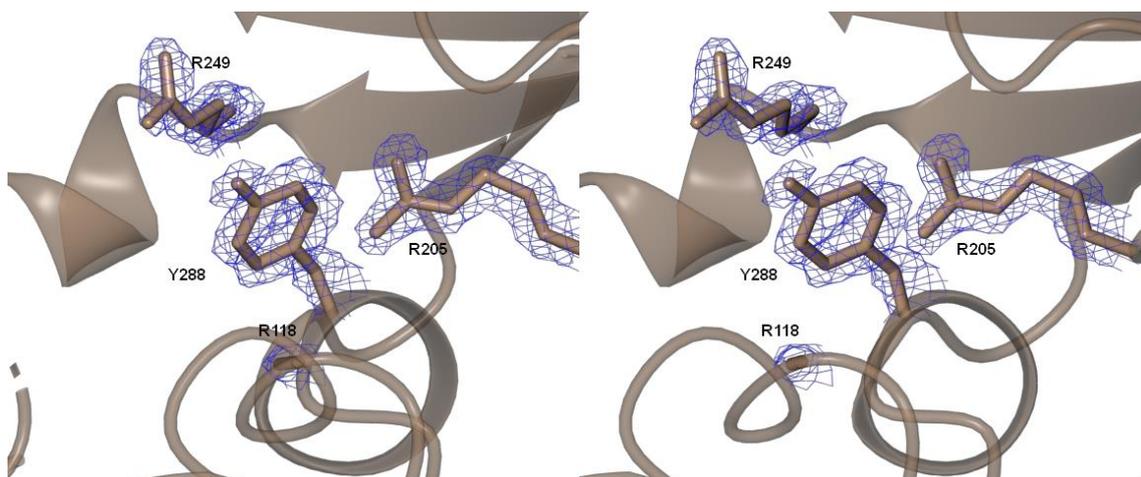


Figure 56: Stereo view of the CeuE-H227A binding cleft (chain B). Protein backbone shown in pale brown ribbons with key sidechains of Arg 118, Arg 205, Arg 249 and Tyr 288 shown as pale brown cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.



Figure 57: Stereo view of loop region of CeuE-H227A residues 222-226 and 228-230 shown in pale brown cylinders. Ala 227 shown in cylinders coloured by atom type. The electron density for the maximum likelihood weighted map weighted was contoured at the 1.5σ level. Figure produced using CCP4mg.

Structure solution for CeuE-Y288F

CeuE-Y288F was in P3₂21, a spacegroup not previously observed for CeuE crystals. The cell parameters gave one particularly long cell length, in the c dimension of 145.65 Å, and two equal cell lengths of 65.52 Å for a and b. There was one protein monomer in the asymmetric unit, and there was clear electron density for all residues in the protein backbone and the average B value for the structure was 24.4 Å. The crystal structure contained two

zinc(II) ions coordinated by residues His 210 and His 227, a result of the crystal being grown in conditions containing ZnCl_2 . An intramolecular zinc(II) complex arose from chelation of the zinc(II) by His 210 and a neighbouring residue Glu 220 (Figure 58).

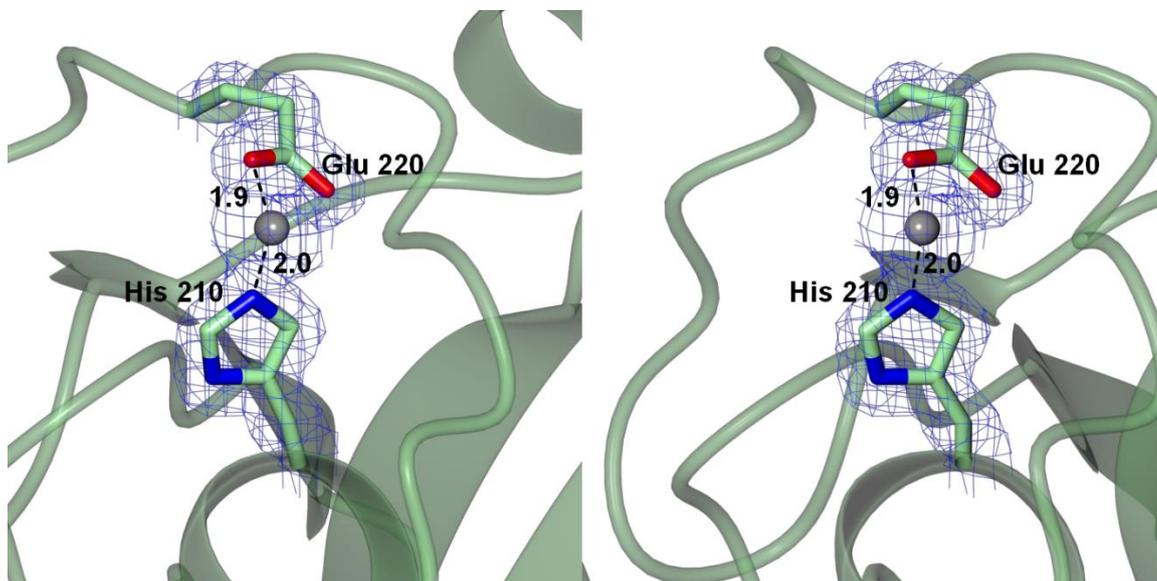


Figure 58: Stereo view of zinc(II) chelated by His 210 and Glu 220 in CeuE-Y288F. Protein backbone is depicted in light green ribbons. His 210 and Glu 220 sidechains are displayed as cylinders coloured by atom type. Zinc(II) is shown as a grey sphere with protein contacts as black dashed lines with bond distances annotated in Å. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.

The second set of zinc(II) interactions were intermolecular, formed from two protein monomers, one providing a nitrogen donor atom from His 227, and the other oxygen donor atoms from Glu 183. Water molecules completed the zinc(II) coordination sphere (Figure 59). This zinc(II) complex between protein monomers is likely an important factor in the crystal packing arrangement of CeuE-Y288F, which is a probable cause of the different spacegroup of $P3_221$ for this variant protein, instead of the more common $P1$. The mutation of the iron(III) binding tyrosine to a phenylalanine was clearly visible in the binding cleft, with a distinct lack of electron density for the tyrosinate oxygen atom (Figure 60).

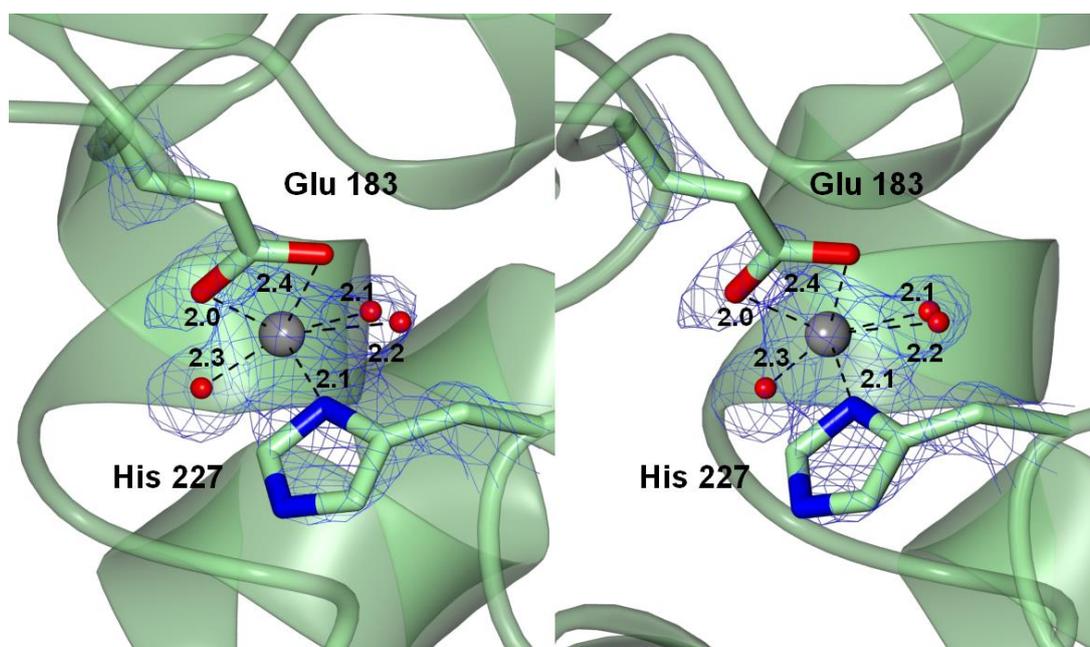


Figure 59: Stereo view of zinc(II) chelated by His 227 and Glu 183 from a neighbouring protein monomer in CeuE-Y288F. Protein backbones are depicted in light green ribbons. His 227 and Glu 183 sidechains are displayed as cylinders coloured by atom type. Zinc(II) is shown as a grey sphere with protein contacts as black dashed lines with bond distances annotated in Å. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.

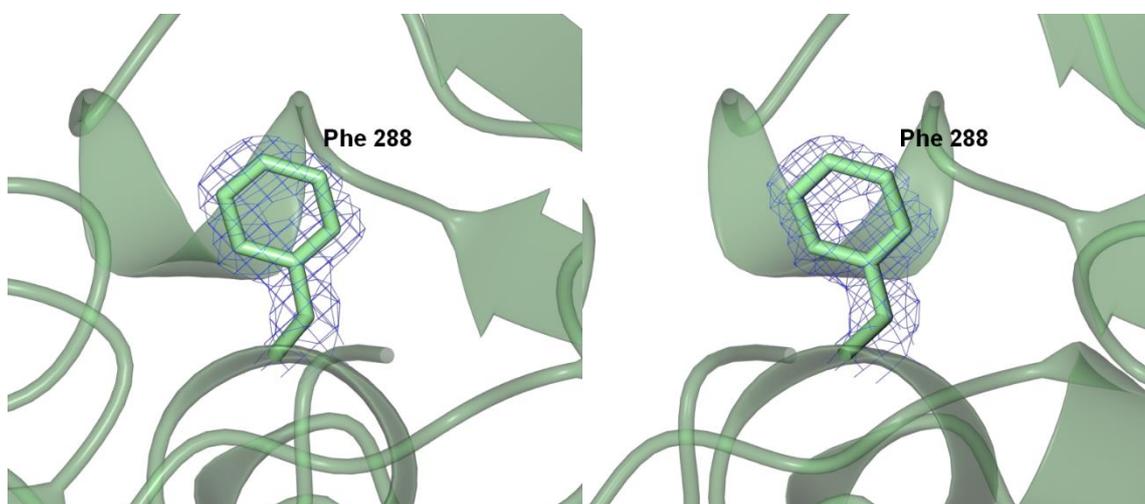


Figure 60: There is clear electron density for the Y288F mutation in the single mutant variant protein. Protein backbone is shown in light green ribbons in stereo view. Phe 288 is shown in light green cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.

This mutation was very clear when modelling in a tyrosine or phenylalanine in COOT.³⁶⁵ When the structure was refined with a tyrosine as residue 288, there was a clear peak in the

negative electron density map over the oxygen atom. When this atom was removed, and the residue was a phenylalanine, the model provided a much better match for the observed electron density (Figure 61).

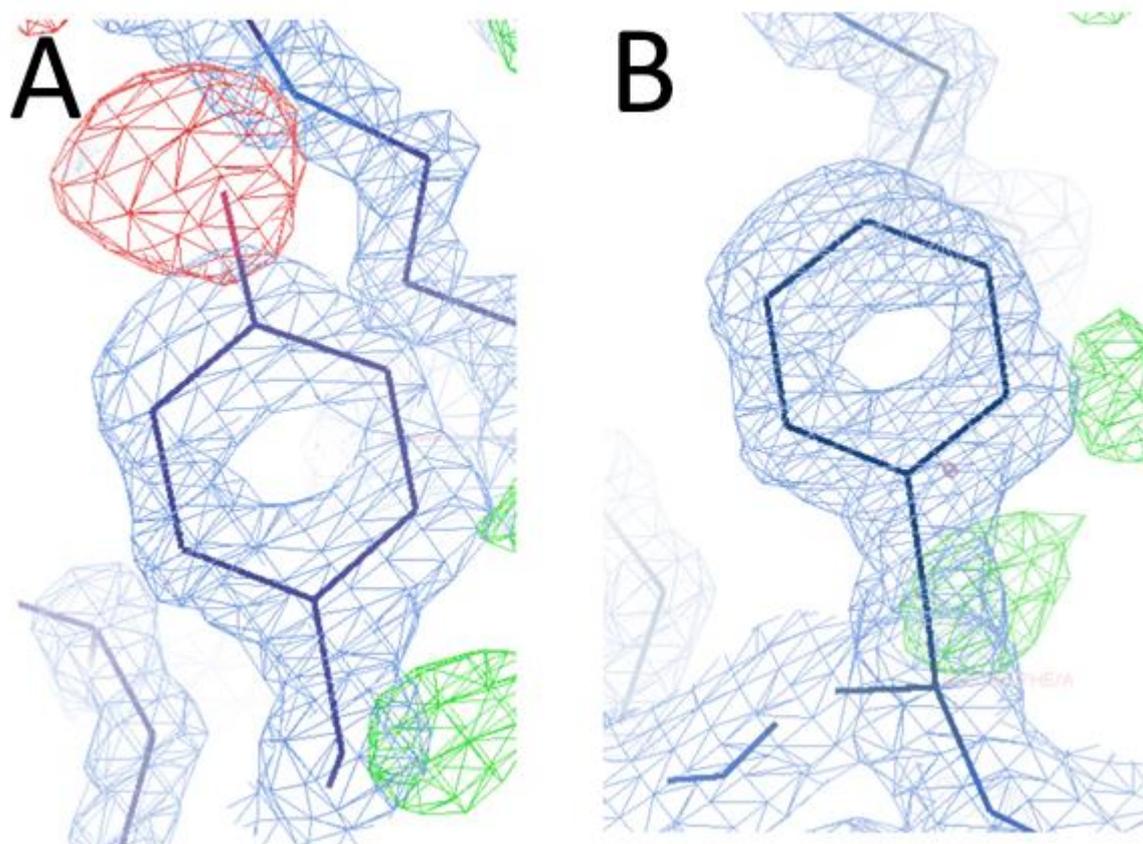


Figure 61: Modelling of residue 288 in COOT.³⁶⁵ The model is shown in dark blue sticks. Weighted electron density map shown in light blue chicken wire, set at 1.5σ . The weighted difference map is shown in red and green chicken wire for negative and positive electron density respectively, set at 1.5σ . A: Tyrosine modelled, with observed negative electron difference density around the tyrosinated oxygen atom. B: Phenylalanine modelled.

Structure Solution for CeuE-H227L/Y288F and CeuE-H227A/Y288F

The double mutant CeuE-H227L/Y288F and CeuE-H227A/Y288F structures were well ordered, with overall B values of 30.0 \AA^2 and 35.9 \AA^2 , respectively. Both variants contained a number of disordered regions that could not be modelled due to a lack of electron density, detailed in Chapter 6. Features present in the binding clefts of both variants were equivalent to those in the single mutant structures, with the H227(L/A) mutation not well resolved, but Y288F clearly present in all cases, with a lack of electron density for the tyrosinate oxygen atom (Figure 62 and Figure 63).

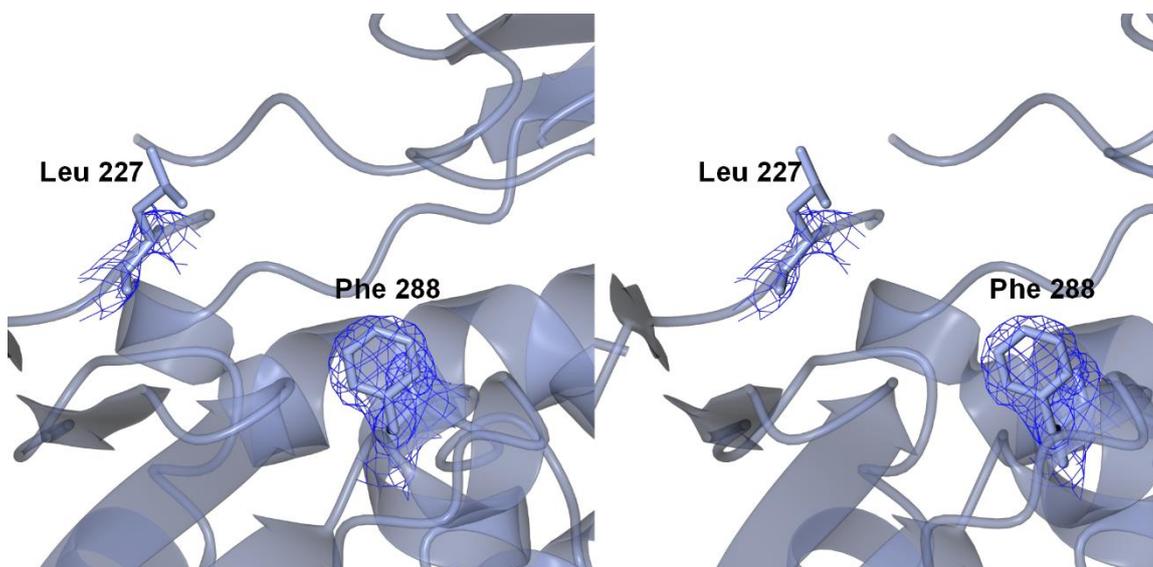


Figure 62: Stereo view of crystal structure of CeuE-H227L/Y288F showing residues Leu 227 and Phe 288 in the binding cleft region of chain B. Protein backbone is shown as ice blue ribbons with residues Leu 227 and Phe 288 shown as ice blue cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.

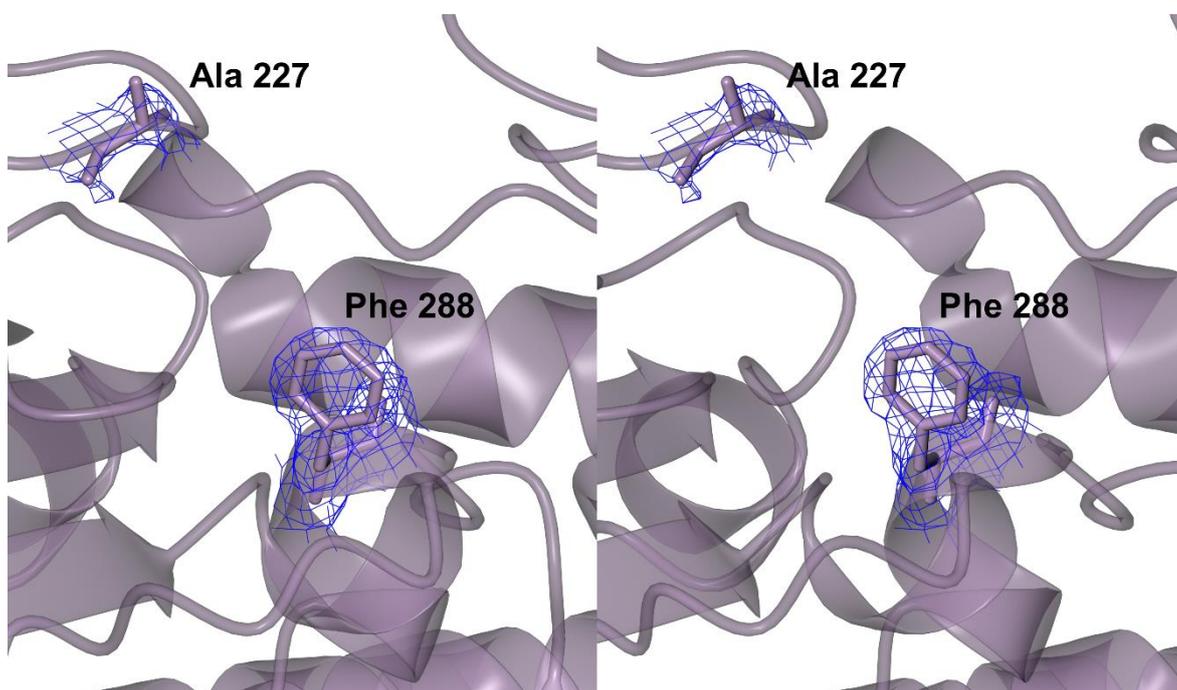


Figure 63: Stereo view of crystal structure of CeuE-H227A/Y288F showing residues Ala 227 and Phe 288 in the binding cleft region of chain B. Protein backbone is shown as lilac ribbons with residues Ala 227 and Phe 288 shown as lilac cylinders. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.

Apo-Variant-CeuE Crystal Structures Summary

The crystal structures of all apo-CeuE variants were used to validate the results from DNA sequencing and ESI-MS that the confirmed mutation was present for variants containing the Tyr 288, and although the H227(L/A) sidechains are not resolved with complete certainty, there is a lack of electron density for a histidine sidechain in this position. The Tyr 288 mutation was well resolved due to its location in a rigid α -helical region, whereas the poorly-resolved His 227 mutations in a flexible loop region, that had less conformational order. In addition, the structures of all apo-CeuE variants confirmed that the mutations left the overall fold of CeuE unchanged, demonstrated by only minor differences in $C\alpha$ positions when the structures are superposed.³¹⁷ This allowed further study *via* fluorescence quenching titration and circular dichroism spectroscopic titrations of binding of ligands that were known to bind to wild-type CeuE.^{186, 310} If there were differences in binding between wild-type CeuE and a variant protein with an iron(III)-siderophore complex, then this could be related to the necessity of the key iron(III) binding residues, rather than the possibility of a misfolding of the protein.

3.4 Dissociation Constants with Iron(III)-n-LICAM

All five variants were studied *via* fluorescence quenching titration with iron (III)-5-LICAM, the siderophore mimic with the highest binding affinity with wild-type CeuE (as discussed in Chapter 2),³¹⁰ to assess whether the mutations to the key iron(III)-binding residues affected the binding affinity. This was possible as all variants retained Trp 287 in a very similar position to that of wild-type CeuE. The fluorescence quenching of Trp 287 in all variants upon addition of iron(III)-siderophore was therefore monitored by the standard fluorescence quenching titration method as detailed in Chapter 6.

The double variants CeuE-H227L/Y288F and CeuE-H227A/Y288F showed very poor iron(III)-5-LICAM binding. This was not surprising, as without an oxygen donor atom from Phe 288, or a nitrogen donor atom from Leu/Ala 227, the hexadentate coordination required by iron(III) could not be fulfilled in the binding cleft of these proteins. The fluorescence quenching data for CeuE-H227L/Y288F titrated with iron(III)-5-LICAM and fitted curves are plotted in Figure 64. The decrease in fluorescence observed when the protein is titrated with an equal volume of 40 mM Tris-HCl pH 7.5, 150 mM NaCl buffer is also displayed to demonstrate the significance of the binding curves. As the normalised fluorescence does not decrease below 0.4 and quenching appears to be only slightly greater than is observed for dilution, the binding affinity was concluded to be very weak.

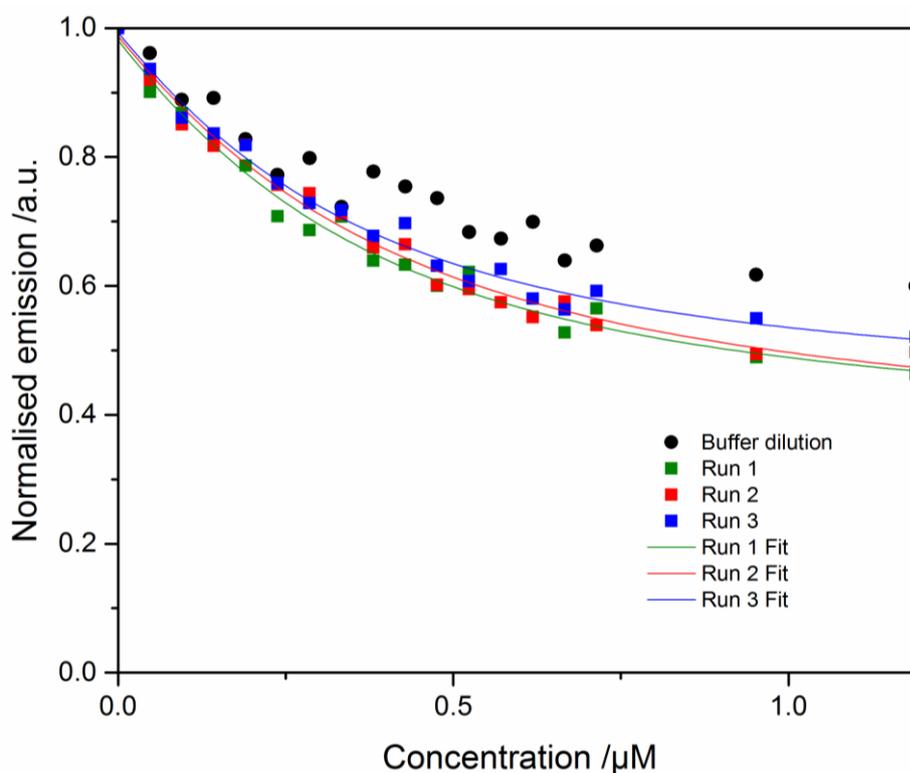


Figure 64: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227L/Y288F titrated with iron(III)-5-LICAM. CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-6-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl. The titration was performed in triplicate, and a binding curve calculated using Dynafit.³²⁶ The data were plotted alongside the normalised fluorescence quenching observed when CeuE-H227L/Y288F is diluted with the same amount of 40 mM Tris-HCl pH 7.5, 150 mM NaCl buffer.

The calculated K_d values and their weighted average and error are displayed in Table 9. Although the errors appear within an acceptable range, the calculated K_d values are not calculated from a true binding curve, as the end-point is not reached, and so these values must be interpreted with caution. It is more likely that the true K_d is a lot higher, probably in excess of 1 μM . To calculate this accurately, the protein and ligand concentrations were increased in an attempt to drive the equilibrium towards complex formation. This method should produce more pronounced fluorescence quenching with a definite end-point, and a better curve for more accurate fitting with Dynafit.³²⁶ In practice this was not possible due to the large concentration of iron(III)-5-LICAM required for full quenching, and the resulting absorbance of the ligand at the excitation wavelength of 280 nm.

Table 9: Calculated K_d values from fluorescence quenching data for CeuE-H227L/Y288F titrated with iron(III)-5-LICAM.

CeuE-H227L/Y288F-iron(III)-5-LICAM	K_d /nM	Error /nM
Run 1	233	49
Run 2	270	58
Run 3	219	45
Average	236	44

It was decided that at 280 nm a ligand absorbance under 0.1 was acceptable, but exceeding 0.1 was likely to be too large an absorbance for estimation of an accurate K_d value. This is due to the primary inner filter effect caused by the ligand absorbing at the excitation wavelength, and ‘artificially’ decreasing the emission intensity due to inadequate excitation of the protein. The maximum iron(III)-5-LICAM concentration was calculated by UV-visible spectroscopic methods detailed in Chapter 6 as 67.3 μM . To obtain a suitable binding curve, it was necessary to introduce at least five equivalents of ligand to protein to ensure saturation of the protein. As such, the maximum protein concentration possible was 13.4 μM . Given the maximum protein concentration of 13.4 μM , a titration with 12 μM for CeuE-H227L/Y288F was trialled, to be sure not to exceed the maximum permissible ligand concentration. This titration gave a shallow titration curve, without a significant amount of fluorescence quenching upon addition of iron(III)-5-LICAM. The data could not be fitted with an adequate curve. The binding affinity could therefore not be accurately quantified by this fluorescence quenching titration method (Figure 65).

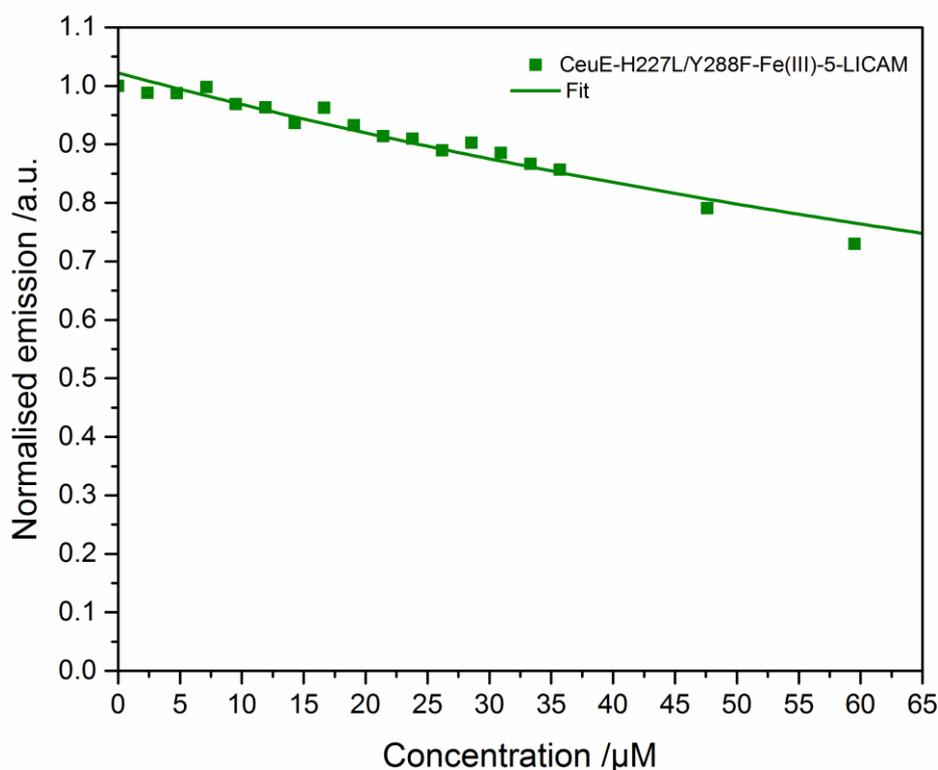


Figure 65: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for Ceue-H227L/Y288F titrated with iron(III)-5-LICAM. Ceue-H227L/Y288F at a concentration of $12 \mu\text{M}$ in $40 \text{ mM Tris-HCl pH } 7.5$, 150 mM NaCl was titrated with aliquots of $60 \mu\text{M}$ iron(III)-6-LICAM in $40 \text{ mM Tris-HCl pH } 7.5$, 150 mM NaCl . A binding curve was calculated using Dynafit.³²⁶ The quenching was not sufficient for determination of a dissociation constant.

Based on the limitation of the method described above, a variation was attempted for titration of Ceue-Y288F with iron(III)-5-LICAM. A $6 \mu\text{M}$ solution of Ceue-Y288F was titrated with 10 equivalents of iron(III)-5-LICAM, and the binding curve observed was shallow, giving a dissociation constant of $19.4 \mu\text{M} \pm 1.9 \mu\text{M}$ (Figure 66). As there was no well-defined endpoint, this dissociation constant was not accurate. As the limit of the method was reached at this protein concentration, further ligand cannot be added before primary inner filter effects become a significant problem. As such, the dissociation constant was estimated as above $1 \mu\text{M}$. Although His 227 was still present in this variant, and capable of iron(III) chelation, iron(III)-5-LICAM had only very weak binding affinity to this protein.

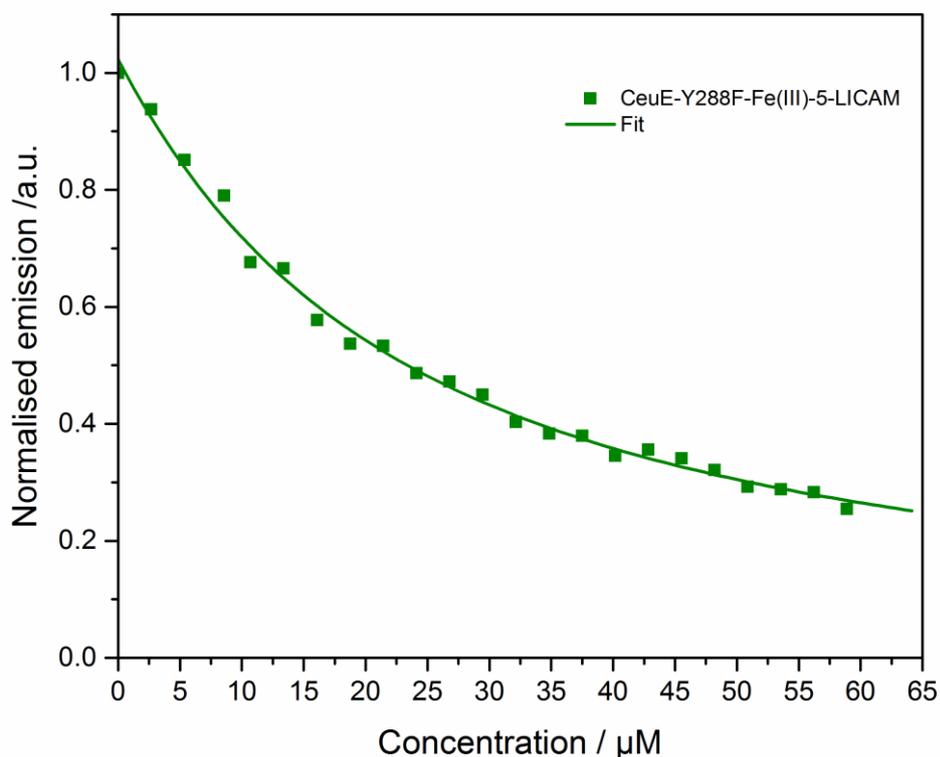


Figure 66: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-Y288F titrated with iron(III)-5-LICAM. CeuE-Y288F at a concentration of $12 \mu\text{M}$ in $40 \text{ mM Tris-HCl pH } 7.5, 150 \text{ mM NaCl}$ was titrated with aliquots of $60 \mu\text{M}$ iron(III)-6-LICAM in $40 \text{ mM Tris-HCl pH } 7.5, 150 \text{ mM NaCl}$. A binding curve was calculated using Dynafit.³²⁶ The quenching was not sufficient for determination of a dissociation constant.

In contrast, iron(III)-5-LICAM did bind to the variants CeuE-H227L and CeuE-H227A. The fluorescence quenching titrations, at the standard protein concentration of $0.24 \mu\text{M}$,⁹⁰ were recorded in triplicate and the binding curves displayed in Figure 67 and Figure 68. All fluorescence quenching titrations for CeuE-H227A were recorded by project student R. P. Thomas.

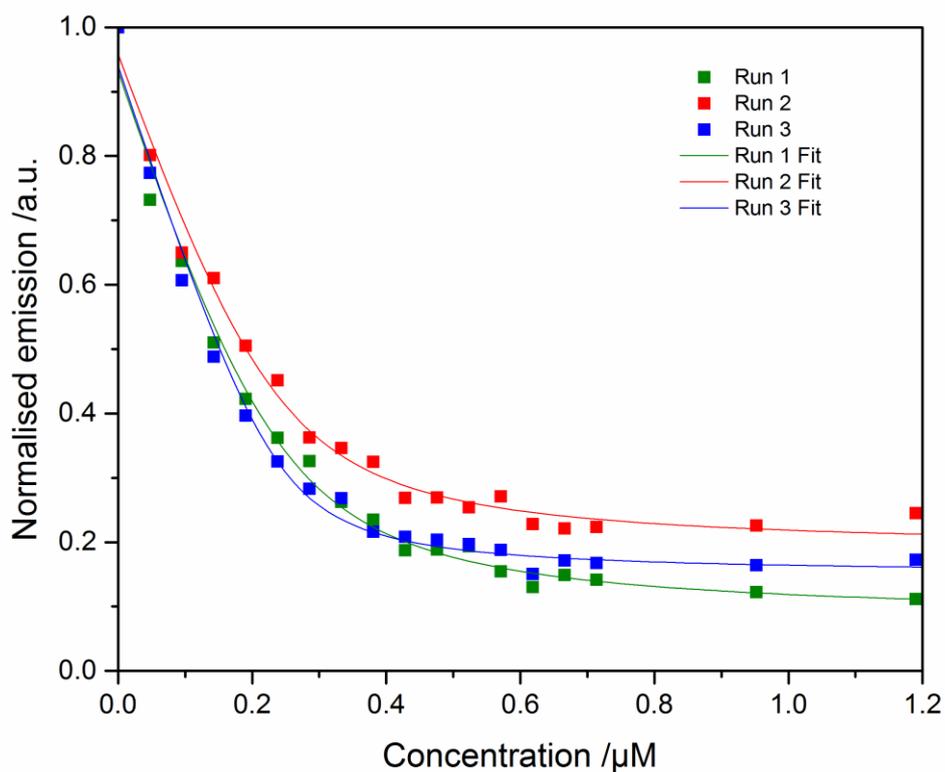


Figure 67: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227L-iron(III)-5-LICAM and their associated non-linear regression fitting data from Dynafit.³²⁶ CeuE-H227L at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

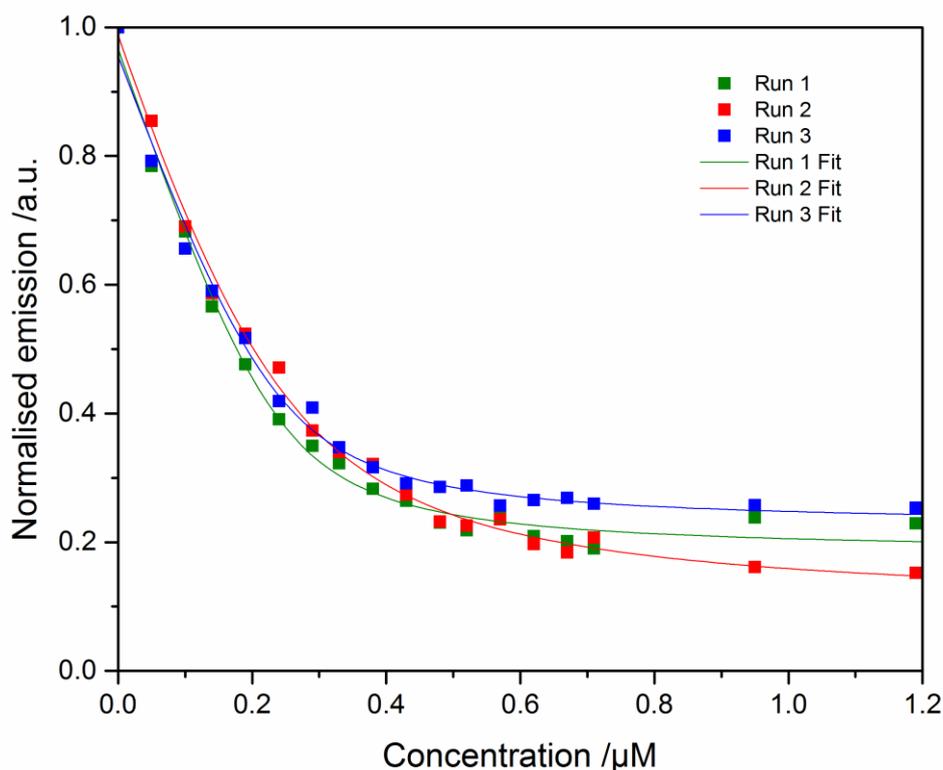


Figure 68: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission *vs* ligand concentration in μM for CeuE-H227A-iron(III)-5-LICAM and their associated non-linear regression fitting data from Dynafit.³²⁶ CeuE-H227A at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

The dissociation constants, calculated *via* non-linear regression of the binding curves in Dynafit, are reported in Table 10 for CeuE-H227L titrated with iron(III)-5-LICAM and in Table 11 for CeuE-H227A titrated with iron(III)-5-LICAM.³²⁶ The dissociation constants were within the nanomolar range previously observed for wild-type CeuE binding to iron(III)-n-LICAM ($n = 4, 5, 6, 8$). CeuE-H227L bound iron(III)-5-LICAM slightly more tightly than CeuE-H227A with a dissociation constant of $22.0 \text{ nM} \pm 10.0 \text{ nM}$ for CeuE-H227L and $35.2 \text{ nM} \pm 13.8 \text{ nM}$ for CeuE-H227A. With the associated errors of the method, it was concluded that the binding affinities for iron(III)-5-LICAM were broadly similar for both CeuE-H227L/A variants. Both variants had weaker binding affinities than that of wild-type CeuE with a K_d of $<10 \text{ nM}$.

Table 10: Dissociation constants for triplicate runs of CeuE-H227L titrated with iron(III)-5-LICAM. The average was calculated using a weighted average method.

CeuE-H227L-iron(III)-5-LICAM	K_d /nM	Error /nM
Run 1	36.6	8.3
Run 2	32.9	7.7
Run 3	14.6	4.3
Average	22	10

Table 11: Dissociation constants for triplicate runs of CeuE-H227A titrated with iron(III)-5-LICAM. The average was calculated using a weighted average method.

CeuE-H227A-iron(III)-5-LICAM	K_d /nM	Error /nM
Run 1	58.4	7.4
Run 2	23.2	5.2
Run 3	36.7	0.8
Average	35	14

These findings indicate that Tyr 288 is essential for CeuE to bind iron(III)-5-LICAM. While His 227 is not essential, it does enhance the binding affinity for CeuE to iron(III)-5-LICAM.

Next, the effect of linker length was explored with the variants to probe whether His 227 had an influence on the binding affinities of shorter or longer linkers compared to the optimum 5-linker. CeuE-H227L and CeuE-H227A were titrated with iron(III)-n-LICAM (n = 4, 5, 6, 8) *via* the standard fluorescence quenching method,⁹⁰ and their associated dissociation constants calculated. The binding curves with associated fitting data for selected runs for each protein and each iron(III)-n-LICAM (n = 4, 5, 6, 8) are shown in Figure 69 for CeuE-H227L and Figure 70 for CeuE-H227A.

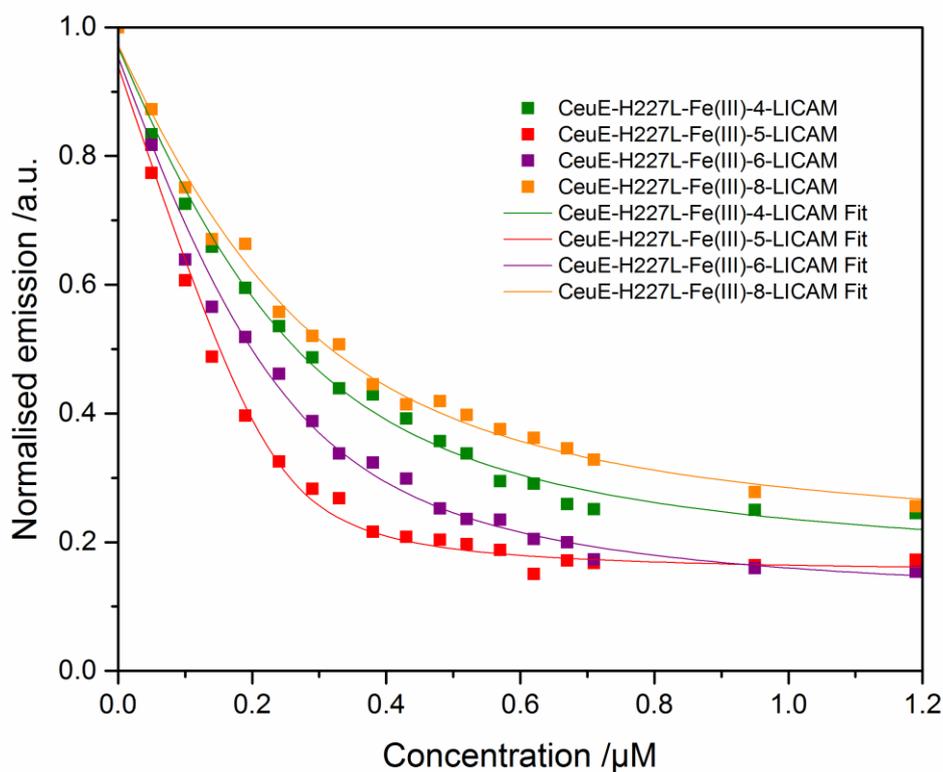


Figure 69: Selected binding curves for the fluorescence quenching titration of iron(III)-n-LICAM ($n = 4, 5, 6, 8$) with CeuE-H227L, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227L-iron(III)-n-LICAM ($n = 4, 5, 6, 8$) and their associated non-linear regression fitting data from Dynafit. CeuE-H227L at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-n-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

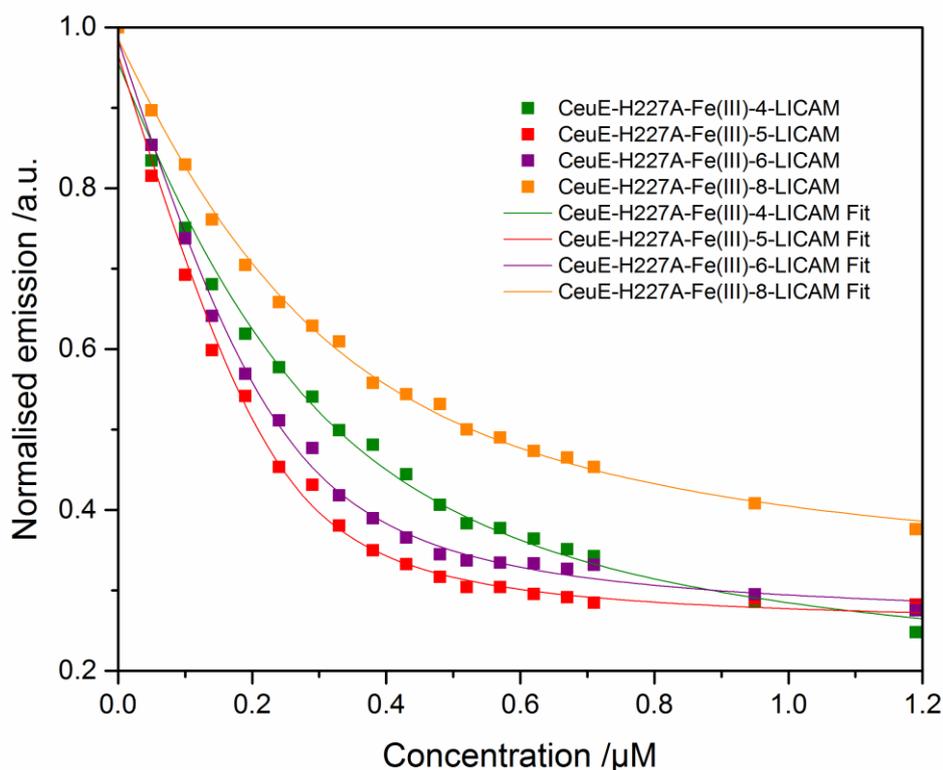


Figure 70: Selected binding curves for the fluorescence quenching titration of iron(III)-n-LICAM ($n = 4, 5, 6, 8$) with CeuE-H227A, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-H227A-iron(III)-n-LICAM ($n = 4, 5, 6, 8$) and their associated non-linear regression fitting data from Dynafit. CeuE-H227A at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-n-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

The binding data from triplicate runs were used to calculate three independent dissociation constants using Dynafit.³²⁶ The average dissociation constant is reported in Table 12. Wild-type CeuE dissociation constants are included for comparison.³¹⁰

Iron(III)-n-LICAM binding was weaker across all linker lengths for both variants compared to wild-type CeuE. It is likely that although His 227 does not appear to be essential to iron(III)-n-LICAM binding, that it does provide a stabilising effect that enhances the binding affinity for all iron(III)-n-LICAM linker lengths.

Table 12: Dissociation constants for wild-type CeuE, and variants CeuE-H227L and CeuE-H227A with iron(III)-n-LICAM (n= 4, 5, 6, 8) siderophore mimic compounds. CeuE-iron(III)-4-LICAM and CeuE-iron(III)-5-LICAM dissociation constants were taken from the literature.³¹⁰

Iron(III)-n-LICAM	CeuE K_d /nM	CeuE-H227L K_d /nM	CeuE-H227A K_d /nM
n= 4	21 ± 6	90 ± 30	131 ± 28
n= 5	< 10	22 ± 10	35 ± 14
n= 6	33 ± 8	65 ± 21	41 ± 7
n= 8	58 ± 8	112 ± 21	125 ± 36

For the iron(III)-n-LICAM (n = 4, 5, 6, 8) ligand series, a similar trend in binding affinities is observed for the CeuE-H227L/A variants as was reported for wild-type CeuE. A linker length of 5 carbon atoms provides the iron(III)-n-LICAM with the highest binding affinity, while longer linkers of 6 and 8 carbon atoms provide a weaker binding affinity, as linker length is increased. iron(III)-4-LICAM binds more weakly than iron(III)-5-LICAM, probably due to linker strain as reported for wild-type CeuE as discussed in Chapter 2.³¹⁰ It appears that this effect of a linker shorter than 5 atoms providing a weaker binding affinity is more pronounced in the CeuE-H227L/A variants than in wild-type CeuE. The dissociation constants for iron(III)-4-LICAM are larger than those of iron(III)-6-LICAM, whereas with wild-type CeuE, iron(III)-4-LICAM was observed to bind more tightly than iron(III)-6-LICAM. This is also observed with shallower binding curves for iron(III)-4-LICAM than the curves for iron(III)-6-LICAM for the titrations with both CeuE-H227L and CeuE-H227A variant proteins.

It may be that the CeuE-H227L/A variants are slightly more tolerant of longer linker lengths, with the flexible loop with H227L/A located away from the binding pocket. It could be that His 227 somehow provides a particularly stabilising effect for the binding of iron(III)-4-LICAM in wild-type CeuE. Given the significant errors in a number of dissociation constants, it was not possible to draw firm conclusions on the reasons for differences in observed binding affinities. It was concluded that the CeuE-H227L and CeuE-H227A variants bind the iron(III)-n-LICAM (n = 4, 5, 6, 8) siderophore mimics with a similar trend in binding affinities to wild-type CeuE, but with overall weaker binding for all ligands.

3.5 Crystal Structure of CeuE-H227L-iron(III)-5-LICAM

Variant CeuE-H227L was co-crystallised with iron(III)-5-LICAM and data were collected at the Diamond Light Source. Details of crystallisation and structure solution are detailed in Chapter 6. The data were indexed in space group P1 with three CeuE-H227L monomers in the asymmetric unit. All three binding clefts contained electron density that allowed iron(III)-5-LICAM to be modelled (Figure 71).

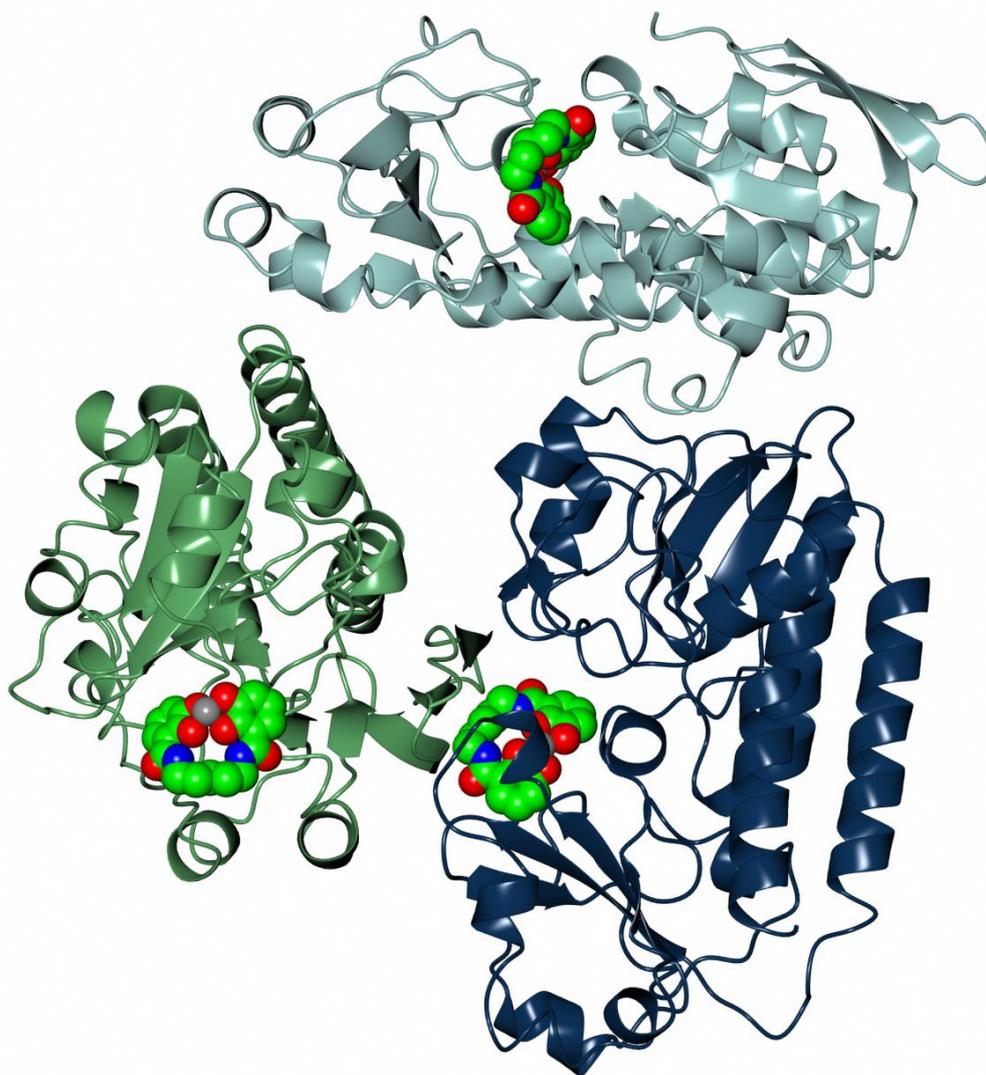


Figure 71: CeuE-H227L co-crystallised with iron(III)-5-LICAM. Three protein monomers are present in the asymmetric unit, with an iron(III)-5-LICAM in each binding cleft. Chain A is shown in sea green ribbons, chain B in lawn green ribbons, and chain C in navy ribbons. iron(III)-5-LICAM is shown in spheres coloured by atom type. Figure produced using CCP4mg.

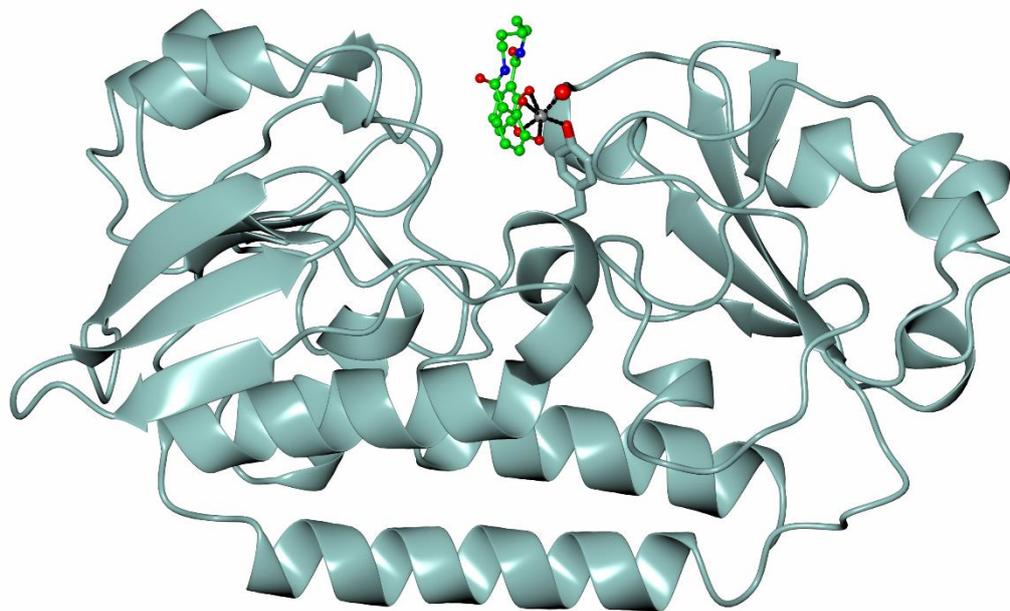


Figure 72: Crystal structure of CeuE-H227L-iron(III)-5-LICAM showing just chain A. iron(III)-5-LICAM is shown as ball and stick coloured by atom type. The aqua (or hydroxo) ligand is shown as a red sphere. Iron(III) contacts are shown as black lines. Figure produced using CCP4mg.

The iron(III) centres in all three chains are coordinated by four oxygen donor atoms from the 5-LICAM ligand scaffold, and by the oxygen donor atom of Tyr 288 (Figure 72). In chains A and B, an aqua (or hydroxo) ligand completes the octahedral coordination (Figure 73).

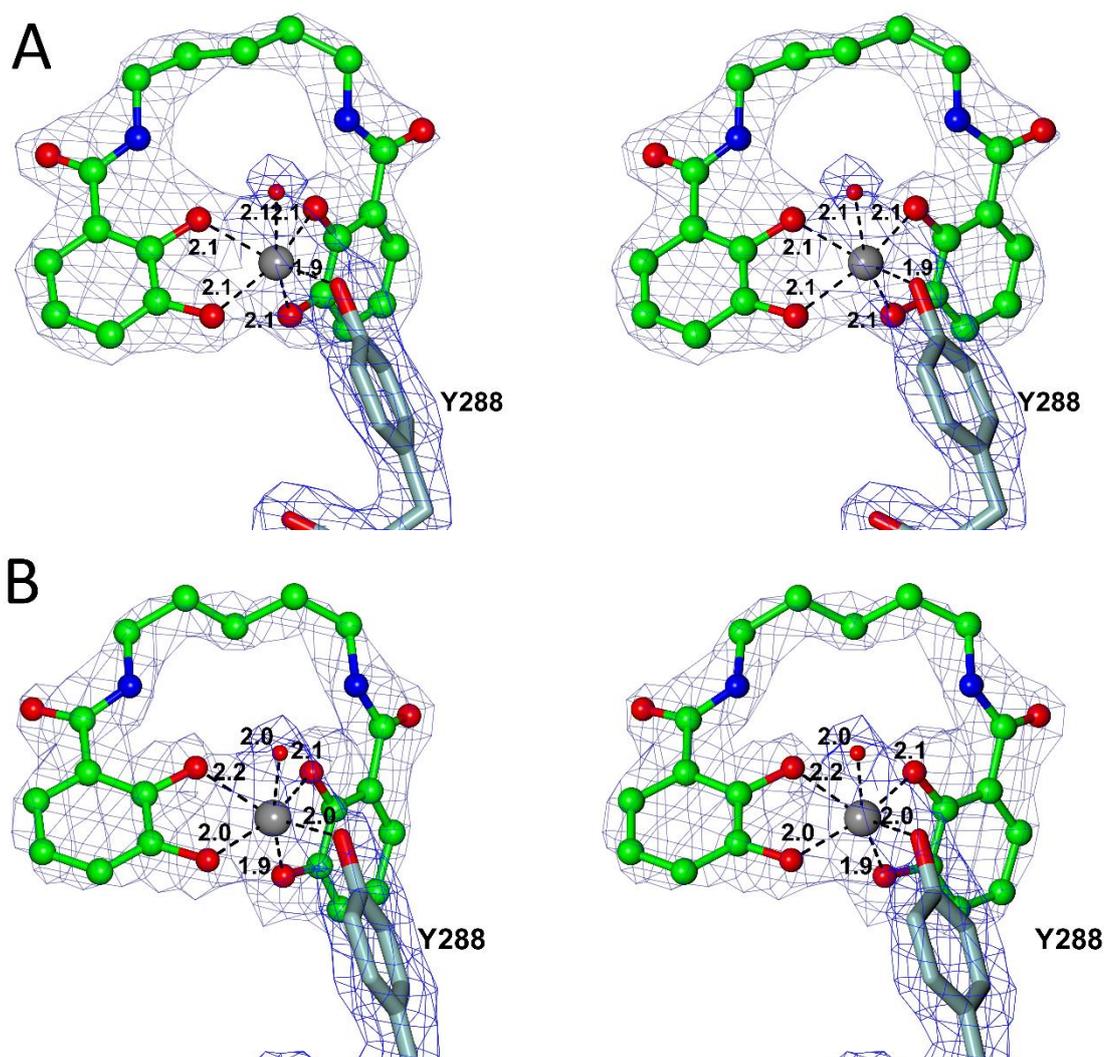


Figure 73: Co-crystal structure of CeuE-H227L-iron(III)-5-LICAM, showing the binding cleft of Chain A (A) and Chain B (B) in stereo view. An aqua (or hydroxo) ligand the octahedral coordination of iron(III). Tyr 288 is shown in cylinders coloured by atom type. iron(III)-5-LICAM is shown in ball and stick coloured by atom type. The aqua (or hydroxo) ligand is shown as a red sphere. Iron(III) bond distances in Å are labelled. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Figure produced using CCP4mg.

The loop region of the protein, where the H227L mutation is contained, is disordered, with poor electron density for the Leu 227 sidechain (Figure 74). The loop is located away from the binding cleft, probably since there are no available donor atoms for iron(III) chelation. Overlaying the structure with that of wild-type CeuE-iron(III)-5-LICAM shows that there are only subtle differences in the binding arrangement of the iron(III)-5-LICAM ligand in the binding cleft, and the Λ -configuration is adopted around the iron(III) centre (Figure 75). The co-crystal structure therefore confirms that iron(III)-5-LICAM can still be bound in the Λ -configuration by CeuE that is deficient in H227.

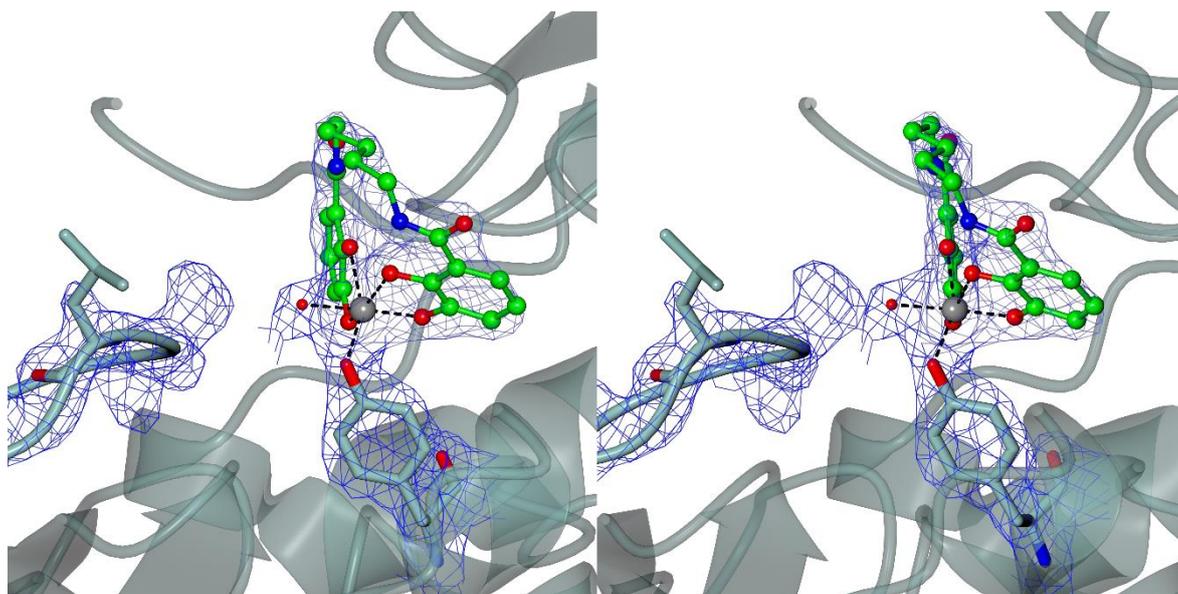


Figure 74: Stereo view of the crystal structure of CeuE-H227L-iron(III)-5-LICAM. CeuE-H227L-iron(III)-5-LICAM is shown in sea green ribbons, with key residues coloured by atom type. The iron(III)-coordinating aqua/hydroxo ligand is shown as a red sphere. The electron density for the maximum likelihood weighted map was contoured at the 1.5σ level. Electron density for the flexible loop region is poor for the L227 sidechain. Figure produced using CCP4mg.

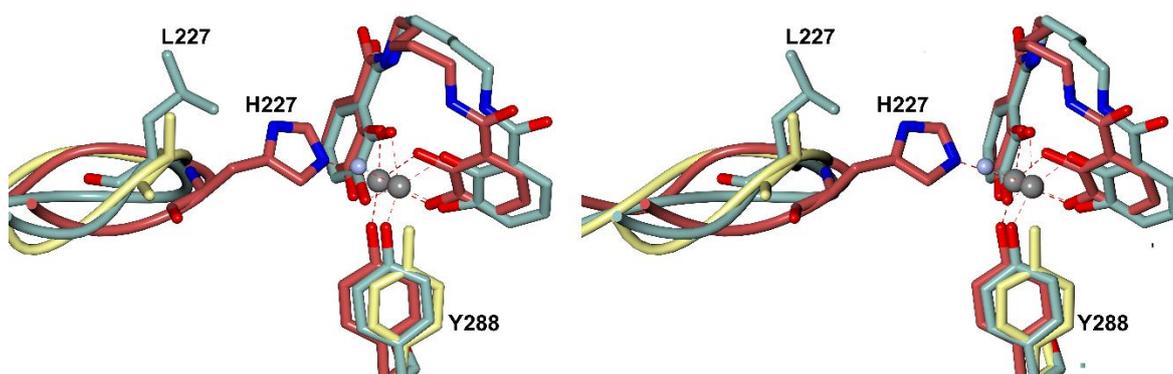


Figure 75: Stereo view of the overlay of crystal structure of CeuE-H227L-iron(III)-5-LICAM with that of CeuE-iron(III)-5-LICAM and CeuE-H227L *via* secondary structure matching.³¹⁷ CeuE-H227L-iron(III)-5-LICAM is shown in sea green cylinders, with key residues coloured by atom type. CeuE-iron(III)-5-LICAM is shown in light crimson cylinders, with key residues coloured by atom type. The water/hydroxide molecule in the CeuE-H227L-iron(III)-5-LICAM structure is shown as a pale blue sphere. CeuE-H227L without a ligand is shown as lemon cylinders. The flexible loop upon which residue 227 is

located is folded away from the iron(III) centre when the H227L mutation is present. Figure produced using CCP4mg.

3.6 Circular Dichroism

Circular dichroism spectroscopy was employed to probe the interaction of the CeuE variant proteins with iron(III)-n-LICAM ($n = 4, 5, 6, 8$) siderophore mimics. Iron(III)-5-LICAM was mixed with each CeuE variant, and the circular dichroism spectra recorded. All five spectra were overlaid in Figure 76.

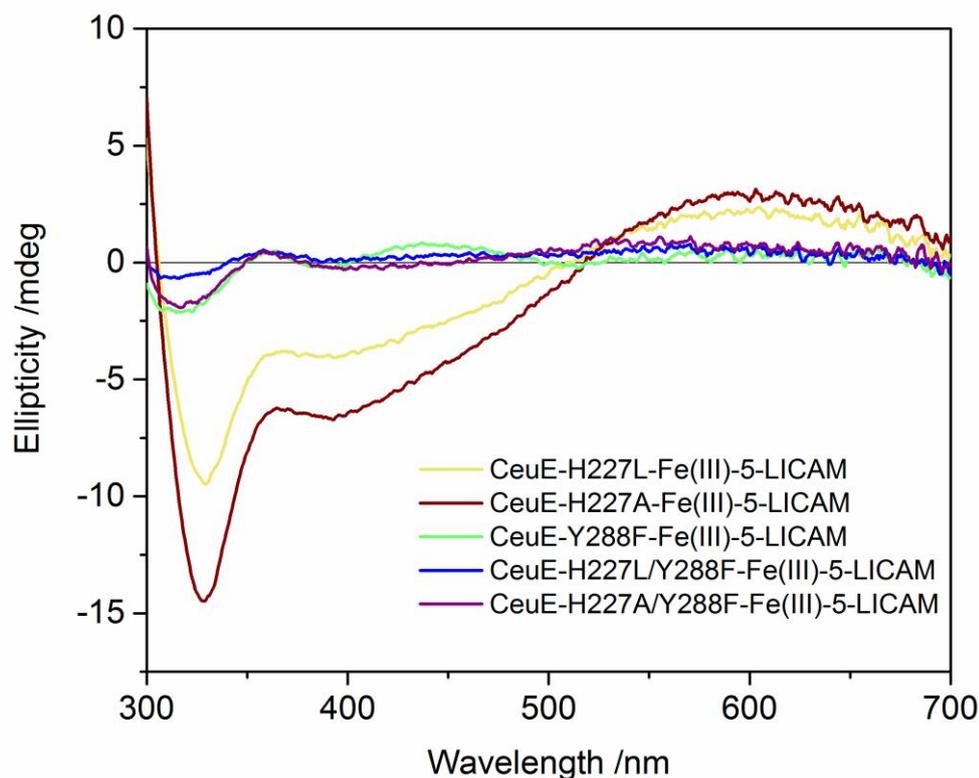


Figure 76: Circular Dichroism spectra for all five CeuE variant proteins upon the addition of iron(III)-5-LICAM. Ellipticity in mdeg was plotted vs wavelength from 300 to 700 nm. The spectra were run for 50 μ M CeuE-variant in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with 50 μ M iron(III)-5-LICAM.

As fluorescence quenching titration suggested very weak binding affinity for iron(III)-5-LICAM to the proteins containing the Y288F mutation (section 3.4), it was of interest to see whether variants CeuE-Y288F, CeuE-H227L/Y288F and CeuE-H227A/Y288F exerted any chiral preference on the Λ/Δ equilibrium of iron(III)-5-LICAM in solution. As the signal remained very close to zero for all three of these variants, it was confirmed that the observed weak binding by fluorescence quenching titration was either very weak, completely non-binding or a non-specific binding event.

The variants containing His 227 mutations, but with an intact Tyr 288, bound iron(III)-5-LICAM with nanomolar affinity as shown by fluorescence quenching titration in section 3.4. The circular dichroism spectra showed a significant Λ signal for both variants when mixed with iron(III)-5-LICAM (Figure 76),^{11, 14} as seen for CeuE-H227L-iron(III)-5-LICAM in the crystal structure. This proved that the Λ -configuration is indeed retained in the solution phase, and that His 227 is not required to induce the Λ -configuration binding mode in CeuE.

Circular dichroism spectra were then recorded with the full iron(III)-n-LICAM (n= 4, 5, 6, 8) series for CeuE-H227L to investigate whether the Λ -configuration was induced for all four linker lengths, and whether the amplitudes of the signal correlated with the observed binding constants from fluorescence quenching studies. All iron(III)-n-LICAM (n= 4, 5, 6, 8) ligands showed a signal for Λ -configured complexes. The strongest signal was observed when iron(III)-5-LICAM was bound, which was not surprising as this ligand had the lowest dissociation constant of the four. Iron(III)-6-LICAM gave the next strongest signal, with iron(III)-4-LICAM and iron(III)-8-LICAM giving the weakest (Figure 77). These spectra correlated well with the dissociation constants recorded for these complexes with CeuE-H227L, and backed up the finding that iron(III)-4-LICAM has a particularly weak interaction when compared to the dissociation constants for iron(III)-4-LICAM with wild-type CeuE.³¹⁰

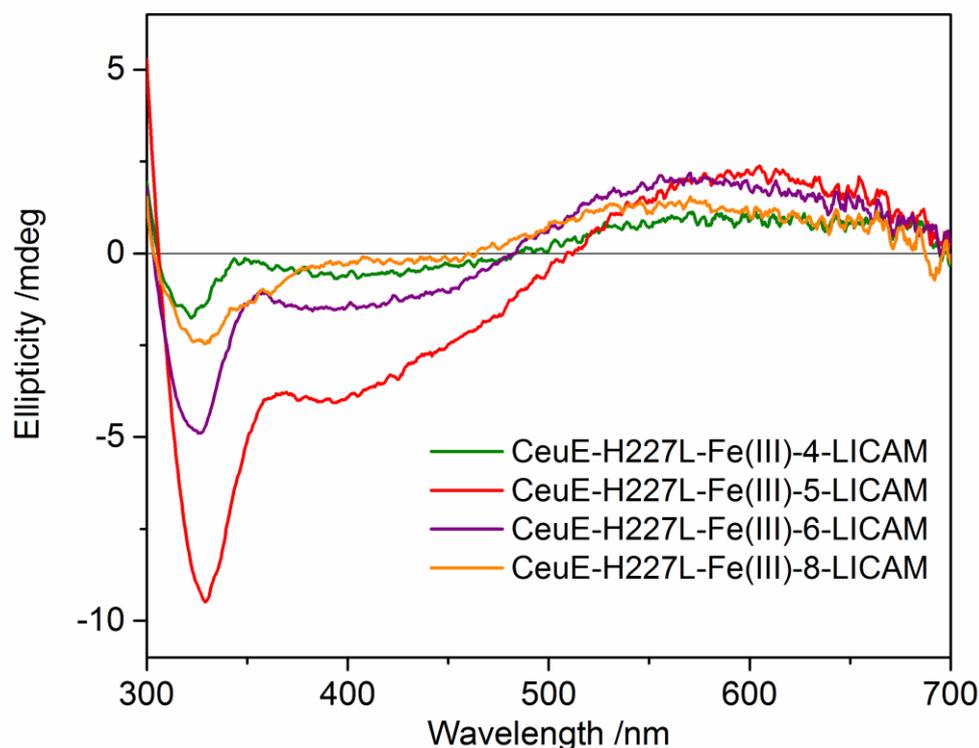


Figure 77: Circular Dichroism spectra for CeuE-H227L with each iron(III)-n-LICAM (n= 4, 5, 6, 8) complex. Ellipticity in mdeg was plotted vs wavelength from 300 to 700 nm. The spectra were run for 50 μ M CeuE-H227L in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with 50 μ M iron(III)-n-LICAM (n = 5, 6, 8) ligands. The spectra show the induction of Λ -configuration.

The magnitudes of the circular dichroism spectra for variant CeuE-H227L were compared with those of wild-type CeuE in Chapter 2. The magnitudes for each protein when in complex with iron(III)-5-LICAM are quite similar, as might be expected for dissociation constants of <10 nM and 20 ± 10 nM for wild-type CeuE and CeuE-H227L respectively, that are within the associated error. For iron(III)-6-LICAM and iron(III)-8-LICAM, the magnitude of the circular dichroism signal is significantly reduced for CeuE-H227L compared to wild-type CeuE, in keeping with the observed binding affinities. Iron(III)-6-LICAM and iron(III)-8-LICAM display increased binding constants of 65 ± 21 nM and 112 ± 21 nM respectively with CeuE-H227L, when compared to wild-type CeuE of 33 ± 8 nM and 58 ± 8 , respectively.

3.7 Summary and Conclusion

The study of CeuE variants containing the Y288F mutation by fluorescence quenching titration showed that they are unable to bind iron(III)-5-LICAM with any measurable binding affinity, with circular dichroism spectra in the presence of iron(III)-5-LICAM showing no significant signal for a Λ -configured complex. These findings indicate that CeuE iron(III)-5-LICAM binding is significantly impaired when Tyr 288 is not available to donate an oxygen atom for iron(III) coordination and indicates that Tyr 288 is fundamental to iron(III)-5-LICAM binding.

Variants containing both the H227L and H227A mutations retained the ability to bind iron(III)-5-LICAM with nanomolar affinity, as was proven by fluorescence quenching titrations. Both variants were able to bind the full iron(III)-n-LICAM (n= 4, 5, 6, 8) series, with larger dissociation constants for all linker lengths when compared to the dissociation constants with wild-type CeuE. This indicates that His 227 is not essential for CeuE to bind iron(III)-n-LICAM (n= 4, 5, 6, 8), but its presence does enhance the binding affinity. As the dissociation constant for iron(III)-4-LICAM was significantly larger for the binding with H227L and H227A variants when compared to the wild-type CeuE, it may be that His 227 provides a particularly stabilising effect for such strained complexes. The crystal structure of CeuE-H227L-iron(III)-5-LICAM backs up the above findings, showing that iron(III)-5-LICAM is able to adopt a very similar binding mode in the binding cleft of CeuE-H227L as that in wild-type CeuE. The vacant coordination site, where His 227 would usually provide a nitrogen donor atom, is filled by a water or hydroxide molecule. The observed Λ -configuration in the CeuE-H227L-iron(III)-5-LICAM co-crystal structure is retained in solution phase, as circular dichroism confirms that CeuE-H227L exerts a preference for the Λ -configuration upon introduction to iron(III)-n-LICAM (n= 4, 5, 6, 8). The crystal structure and circular dichroism spectroscopic results therefore prove that His 227 is not required for the complexes to adopt the Λ -configuration in the CeuE binding cleft.

These findings that Tyr 288 is essential for iron(III)-n-LICAM (n= 4, 5, 6, 8) binding, but His 227 is not, can be rationalised by the positioning of these residues in the tertiary fold of CeuE and chemical properties of these residues. Tyr 288 is located on a rigid α -helical region of CeuE, and so its position and conformational flexibility are limited. His 227 in contrast is located on a flexible loop region, and so from the unbound to the bound state, the protein must undergo a conformational rearrangement in this region to bring the chelating His 227 nitrogen towards the iron(III) centre (Figure 78). This means there is an entropic cost to the

fixing of His 227 into the bound position. As Tyr 288 is more conformationally rigid, it is likely already in the optimum position for iron(III) coordination.

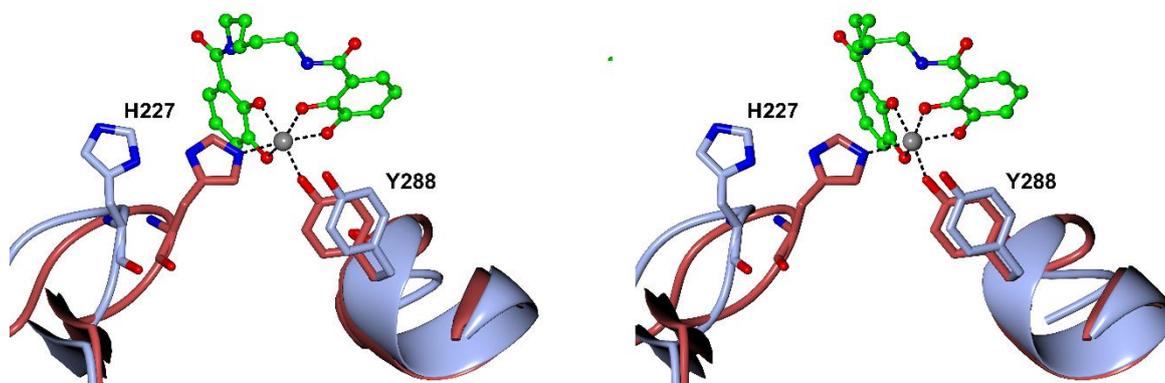


Figure 78: Stereo view of the overlay of crystal structures of apo wild-type CeuE (PDB ID: 3ZKW) with CeuE-iron(III)-5-LICAM (PDB ID: 5A5D) showing the conformational shift of the His 227-loop region for iron(III) binding. Figure produced using CCP4mg.

Additionally, it is likely that the Tyr O-iron(III) bond is stronger than the His N-iron(III) bond, due to nitrogen being less electronegative than oxygen, which means that nitrogen is a more intermediate-type donor, and a softer Lewis base. Oxygen is a hard Lewis base, which is more able to complement the hard Lewis acid character of the small and highly charged iron(III) centre. Combining these factors along with the experimental findings make a clear conclusion that Tyr 288 is much better optimised for iron(III)-n-LICAM coordination in the CeuE binding pocket than His 227.

This study provides a general model of the binding arrangement of iron(III)-tetradentate catecholate siderophores and the key binding residues in a family of periplasmic binding proteins, as were identified by sequence alignments in section 3.1. These results should prove useful for the further study of PBPs from pathogenic species, for the development of siderophore-conjugated antibiotics.^{301, 366, 367}

3.8 Future Work

To extend this study, further mutations could be made to CeuE to explore the wider binding pocket and uncover the contribution of other amino acids that are involved in iron(III)-siderophore binding. For example, the arginine residues Arg 118, Arg 205 and Arg 249 that are known to provide stabilising hydrogen bonds could be replaced with non-hydrogen bonding residues and the iron(III)-siderophore binding to be quantified. This would be of interest for gaining information about the wider family of periplasmic binding proteins identified in section 3.1 and may be useful in understanding iron uptake pathways in a number of pathogenic bacteria. It may then be of interest to extend the study to analogous proteins from such species, to identify whether iron(III)-siderophore binding mode is in fact similar to CeuE, which could lead to identification of new antimicrobial targets.^{299, 301, 303}

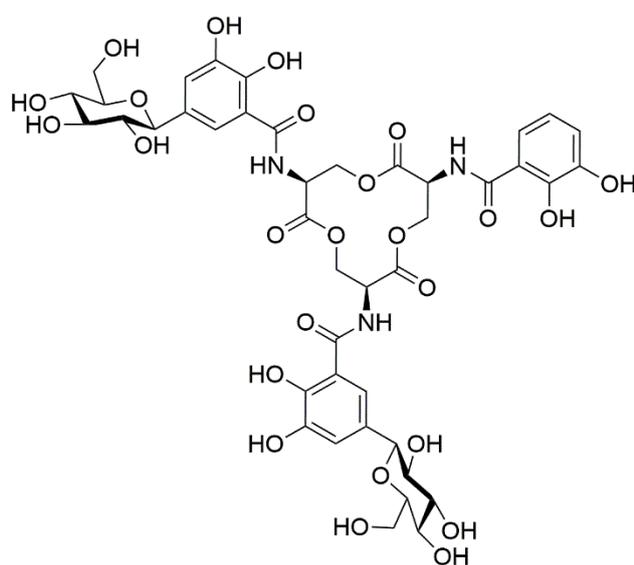
A wider variety of ligands could be explored with CeuE variant proteins. It would be interesting to validate the study by investigating whether the natural substrate iron(III)-bisDHBS gives a similar binding profile with the CeuE-H227A/L and CeuE-Y288F variant proteins as the n-LICAM siderophore mimics do.⁹⁰ This would prove whether the n-LICAM series are adequate mimics for the natural substrate in these studies. As CeuE is known to only bind iron(III)-enterobactin with weak affinity,⁹⁰ it may be of interest to study whether CeuE variant proteins missing the residues needed for optimum tetradentate siderophore binding are able to bind this hexadentate siderophore with a significant difference in affinity.

This work could further inform site-directed mutagenesis of CeuE for applications in biotechnology. For example, if the binding affinity for a siderophore-catalyst conjugate could be increased, then this would be useful for the development of artificial metalloenzymes. If the siderophore component could be tightly anchored into the binding cleft, then this would enhance efficiency and potentially the enantioselectivity of the system.^{336, 355-357} To enhance the binding affinity, one strategy might be to replace His 227 with a residue that could provide an oxygen donor atom, such as tyrosine, aspartic acid or glutamic acid, which could be a better match for iron(III) chelation in terms of Lewis acidity. Another approach could be to provide more hydrogen-bonding residues in the Arg 118, Arg 205, Arg 249 coordination sphere, for increased specificity and fixation of the overall ligand. There is also scope to introduce unnatural amino acids into the protein backbone, to functionalise CeuE for a range of applications.³⁶⁸⁻³⁷⁰ This could be extended further by investigating analogues of CeuE from thermophilic species such as *Geobacillus thermoglucosidasius*, with the aim to develop an artificial metalloenzyme capable of functioning at temperatures in excess of 60 °C.³⁵²

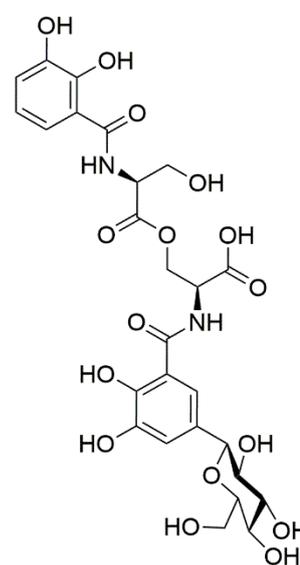
4 Synthesis and Study of Salmochelin S1 Siderophore Mimics and their Interactions with CeuE

4.1 Introduction

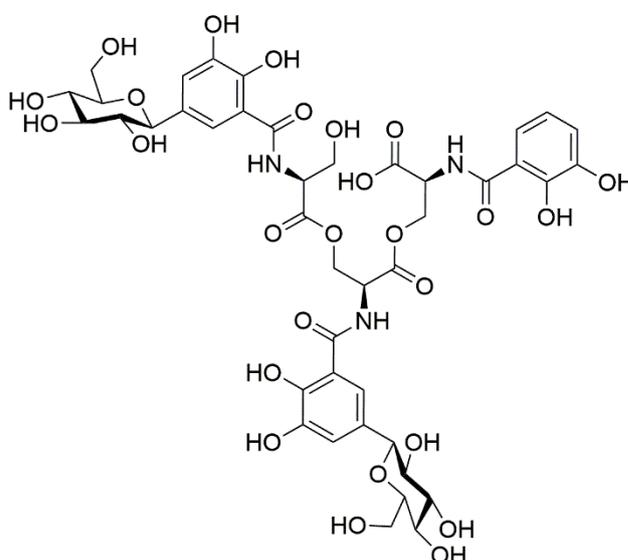
The salmochelin family of siderophores are enzymatically C-glucosylated analogues of enterobactin and the resulting hydrolysis products, produced by bacterial species including *E. coli* and *S. enterica*.^{154, 235, 371} As discussed previously in Chapter 1, salmochelins are thought to exist in nature as stealth siderophores that are able to evade capture by the mammalian immunoprotein siderocalin.^{218, 230, 243, 372} Because enterobactin is captured by siderocalin, the use of salmochelin siderophores allows bacterial species to more effectively acquire iron from their host.^{219, 220, 236, 373}



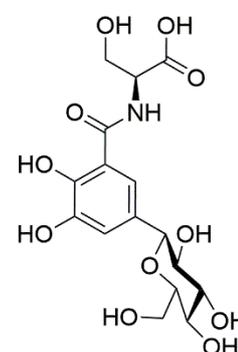
Salmochelin S4



Salmochelin S1



Salmochelin S2



Salmochelin SX

It was postulated that as CeuE readily binds iron(III)-bisDHBS, it may also be able to bind the iron(III) complex of the glucosylated analogue of bisDHBS, salmochelin S1.^{90, 154} It is known that *C. jejuni* is able to survive and grow using salmochelin S4 as an iron source under otherwise iron-limited conditions.¹⁸² However, the stability and hydrolysis of salmochelin S4 was not explored in the growth study, and commercially-sourced salmochelin S4 was used as supplied without additional characterisation.¹⁸² It is therefore likely that some salmochelin S1 was present in the sample, or salmochelin S4 was hydrolysed during the assay to its tetradentate and bidentate hydrolysis products by the periplasmic trilactone esterase Cee.¹⁵⁹ It appeared possible that if CeuE binds iron(III)-salmochelin S1 in the periplasm, iron delivery to the cytoplasm with the CeuBCDE iron(III)-uptake system may be possible *via* this siderophore.^{172, 342}

Computational Modelling of the Binding of CeuE with Iron(III)-Salmochelin

S1

To test the hypothesis that CeuE may be able to bind iron(III)-salmochelin S1, computational modelling of the CeuE binding cleft was conducted by a collaborator, P. S. Bond, details for which are listed in Chapter 6. Interestingly, the modelling indicated that there was a potential secondary binding pocket (Cavity 2, Figure 79) adjacent to the known iron(III)-siderophore binding cleft (Cavity 1, Figure 79).

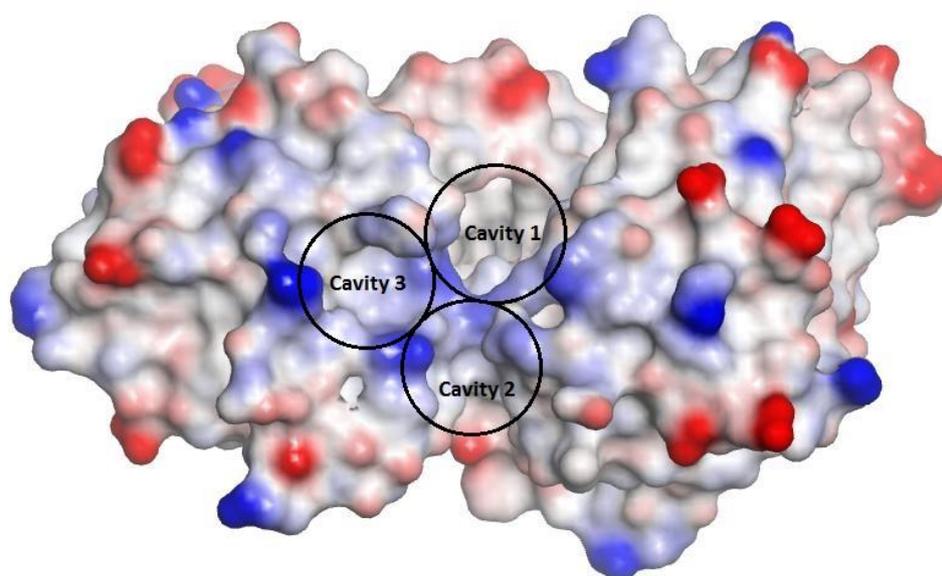


Figure 79: CeuE (PDB ID: 3ZKW) shown with surface view, with three identified binding cleft regions. Cavity 1 is the known iron(III)-siderophore binding cleft. Cavity 2 is a secondary binding cleft that may be capable of accommodating a glucose unit. Cavity 3 is a

smaller pocket that is unlikely to accommodate a glucose unit. Figure provided by P. S. Bond.

Using the coordinates from the CeuE-iron(III)-4-LICAM crystal structure (PDB ID:5A1J), a glucose moiety was attached *via* a carbon-carbon bond at the C5 position of an iron(III)-bound catecholate siderophore that was modelled into the CeuE binding cleft, in the previously known tetradentate siderophore-binding arrangement.¹⁸⁶ The glucose was positioned towards the secondary binding cleft. It was observed upon energy minimisation, *via* a standard dynamics cascade method, that it was possible for this glucose to be accommodated in the secondary binding pocket.^{332, 374} The glucose was able to adopt a number of chair conformations that did not cause unfavourable interactions with the surrounding protein (Figure 80).

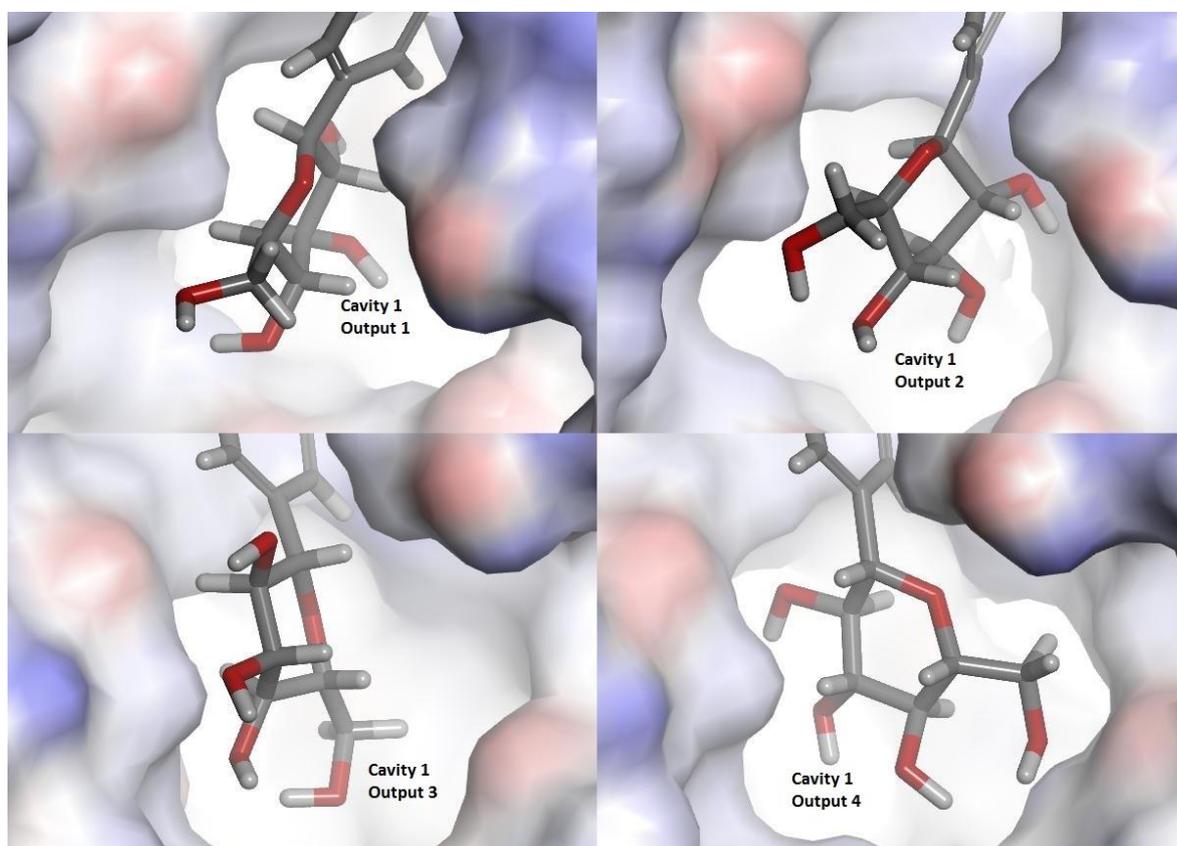


Figure 80: Molecular modelling of chair conformations of the C5 appended glucose in the CeuE secondary binding pocket. Each conformation arose from different simulation parameters. All four conformations provided possible glucose orientations in the binding cleft without unfavourable clashes. Figure provided by P. S. Bond.

The energy upon addition of the glucose into the secondary binding pocket was calculated to decrease from $-4.284 \times 10^4 \pm 25 \text{ kJ mol}^{-1}$ to $-4.324 \times 10^4 \pm 10 \text{ kJ mol}^{-1}$, indicating a lack of intramolecular clashes of the protein. The likelihood of hydrogen bonding with the

glucose unit was shown to be predominantly solvent based, and simulations included a number of potential hydrogen bonds to water molecules (Figure 81).

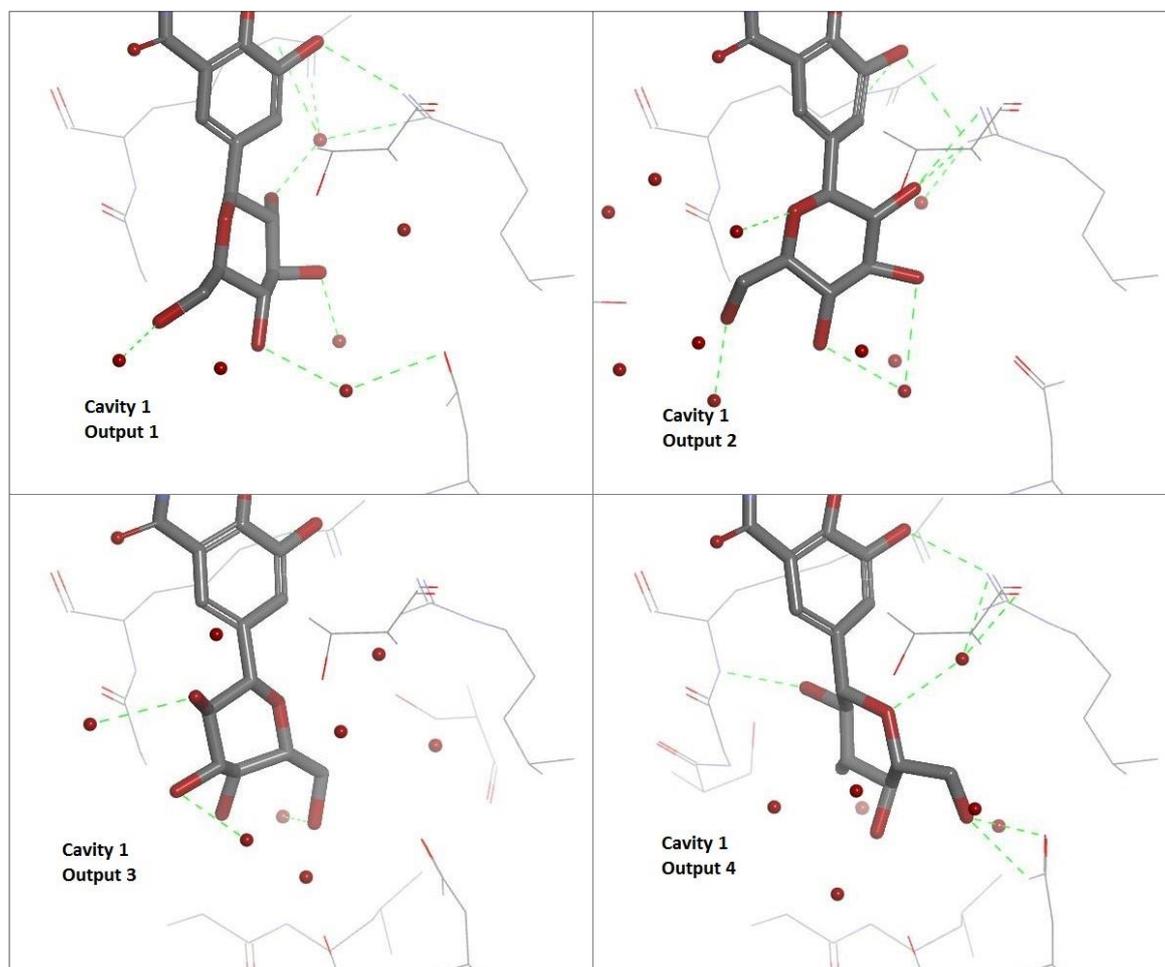


Figure 81: Potential hydrogen bonds between glucose and water molecules when modelled in the secondary binding cleft of CeuE. Figure provided by P. S. Bond.

The results of the simulations indicated that there was a good probability that CeuE would be able to bind iron(III)-salmochelin S1. Although an additional third smaller potential binding cleft was studied (Cavity 3, Figure 79) it was established that it was less likely that salmochelin S4, with two glucose units, would be readily accommodated by all three potential binding pockets.

Methods for Synthesis of Aryl- β -C-glycosides

To establish experimentally whether iron(III)-salmochelin S1 could be bound by CeuE, it was necessary to design a salmochelin S1 mimic. A key step in the synthesis involved the installation of a carbon-carbon bond in the C5 position of a catechol aromatic ring to generate an aryl- β -C-glycoside (Figure 82).

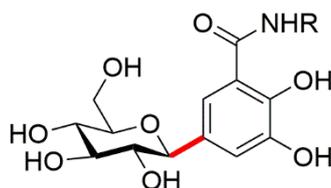
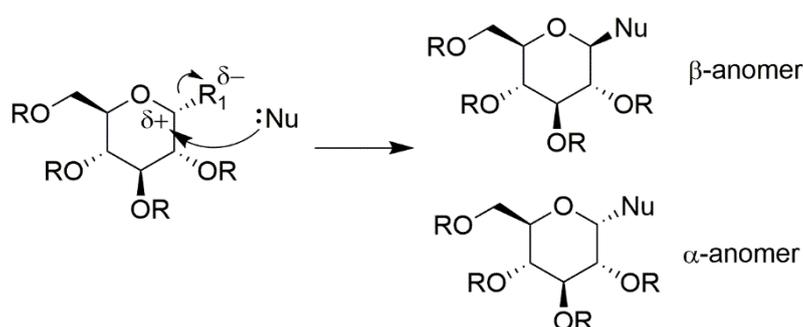


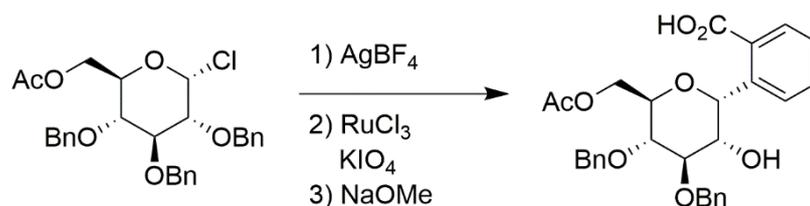
Figure 82: Necessary C-C bond formation between the C5 position of a catecholamide and a glucose unit for the generation of an aryl- β -C-glycoside salmochelin mimic compound. The new C-C bond is shown in red.

Interest in aryl-C-glycosides in the literature is predominantly focussed on the synthesis of natural products, for study into their biological importance or medicinal properties.³⁷⁵⁻³⁷⁷ There are a number of synthetic methods for the generation of aryl-C-glycosides known, including arylation, alkylation and C-C cross coupling, and the area of chemistry has been thoroughly reviewed.^{295, 377-384} Each strategy generally involves the production of an electrophilic centre at the anomeric carbon of the glucose unit, that can be reacted with the desired nucleophile, such as an appropriately functionalised aryl ring (Scheme 3).



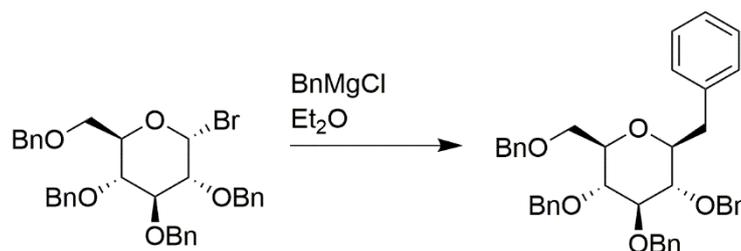
Scheme 3: General strategy for production of aryl-C-glycosides *via* generation of an electrophilic centre at the anomeric carbon.

Early approaches for aryl-C-glycoside formation included a Friedel-Crafts type reaction *via* the generation of acetate protected glucosyl chlorides (Scheme 3 R = Ac, R₁ = Cl) that reacted with aromatic compounds when heated with aluminium chloride.³⁸⁵ However, these initial studies lacked stereoselective control in the production of the β -anomer, which is required to mimic the salmochelins. In an attempt to achieve stereocontrol, the Friedel-Crafts method was adapted with an oxidation, followed by an intramolecular migration of a benzyl group to the anomeric carbon, exclusively producing the α -anomeric product.³⁸⁶ However, this synthesis involved multiple steps, produced the undesired anomer and was not widely applicable to a range of substrates (Scheme 4).³⁸⁶



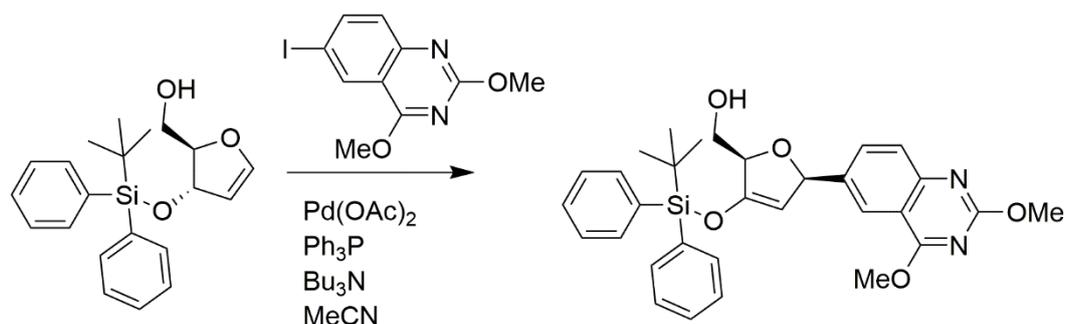
Scheme 4: Intramolecular Friedel-Crafts reaction, resulting in pure α -anomeric product.³⁸⁶

By carefully selecting protecting groups for the hydroxy units (Scheme 3 shown as R) and with the correct choice of leaving group (Scheme 3 shown as R₁) it was possible to influence the production of the β anomeric product *via* reaction with a Grignard reagent.^{387, 388} Full conversion to the β -anomer was achieved with the use of benzyl protecting groups and a bromide leaving group (Scheme 5).³⁸⁸ Numerous studies reported in the literature expand on the use of Grignard reagents for the selective generation of both α and β -aryl-*C*-glycosides with the use of comparable organometallic reagents such as aryllithium species, organoindium halides and organocuprates.^{380, 381, 389-392}

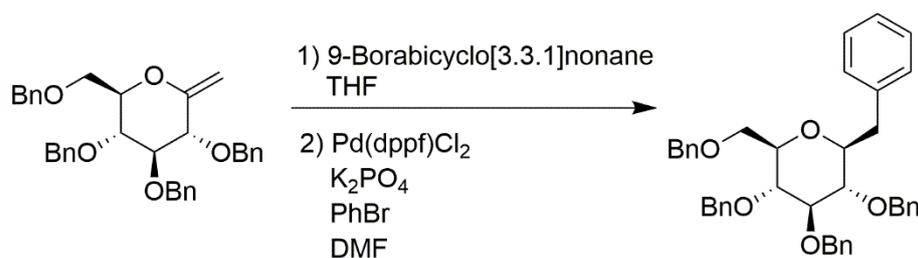


Scheme 5: A Grignard reagent method for selective generation of a β -anomeric product.³⁸⁸

More recent studies into selective generation of aryl- β -*C*-glycosides include the use of *C*-*C* cross coupling reactions.^{378, 380, 383} The Heck reaction was employed for synthesis of β -anomeric products of pyrimidine nucleosides, for applications in DNA base biotechnology (Scheme 6).^{378, 393} The method relied on the use of a suitably bulky protecting group on the 3-hydroxy group of the sugar unit, ensuring exclusive attack on the less sterically hindered face by the organopalladium reagent.³⁷⁸ In addition, Suzuki coupling of palladium-catalysed aryl boronic acids with *exo*-glycals proved successful for the generation of β -*C*-glycosides (Scheme 7).^{383, 394}

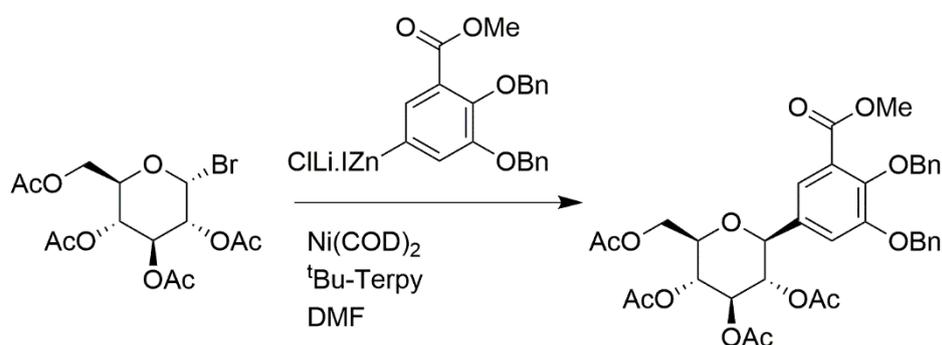


Scheme 6: The use of the Heck reaction to generate a β-anomeric product.³⁹³



Scheme 7: Suzuki coupling reaction to generate a β-anomeric product.^{383, 394}

Key to this project were the studies by Gong and co-workers into the use of Negishi coupling reactions in the synthesis of the salmochelins.³⁹⁵⁻³⁹⁷ The group first published a number of different substrates in Negishi coupling reactions of glucose analogues, with resulting α- and β-selectivity dependent on the reaction conditions.^{395, 396} The most β-selective conditions (10 mol % Ni(COD)₂, 15 mol % ^tBu-Terpy, 150 mol % ArZnI.LiCl in DMF) were then applied to synthesise a key β-C5-glucosyl protected catechol species in a 20:1 β:α ratio (Scheme 8).³⁹⁶ This species was then adapted for coupling to a di-serine, tri-serine and cyclic tri-serine backbone for syntheses of salmochelin S1, salmochelin S2 and salmochelin S4 respectively.³⁹⁷ Logically, this study was used as a starting point for the synthesis of the salmochelin mimics detailed in section 4.2.



Scheme 8: Negishi coupling reaction utilised for the synthesis of β-C5-glucosyl catechol units found in natural salmochelins.³⁹⁶

The research presented in this chapter aimed to identify whether salmochelin S1 or related mimics could be accommodated in the known binding cleft, and the adjacent pocket (Cavity 2, Figure 79) of CeuE. The aim was to further knowledge into the iron(III)-uptake capacity of *C. jejuni*, and give an explanation of the role of stealth siderophores in the growth and survival of the bacterial species. The work focussed on the synthesis of salmochelin S1 mimics and establishing iron-binding properties of these compounds *via* a Job Plot method. The complexes were then studied for CeuE binding *via* fluorescence quenching titration and circular dichroism. The use of these mimics enabled an insight into the potential ability of CeuE to accommodate iron(III)-bound salmochelin S1.

4.2 Design and Synthesis of Salmochelin S1 Mimic Compounds

Salmochelin S1

Salmochelin S1 was purchased from EMC Microcollections and the sample was characterised by ESI-MS and analytical HPLC, as detailed in Chapter 6. Both methods indicated that salmochelin S1 was present. The HPLC trace gave a single peak with strong absorbance at 254 nm, with a retention time of 10 minutes (Figure 83). As the analytical HPLC parameters were set as close to those described in the literature as possible, the retention time when compared with the literature value of around 12 minutes, was deemed within the expected range.²³⁵ ESI-MS gave a molecular ion peak with m/z of 625.1510, corresponding to the $[M-H]^-$ species.

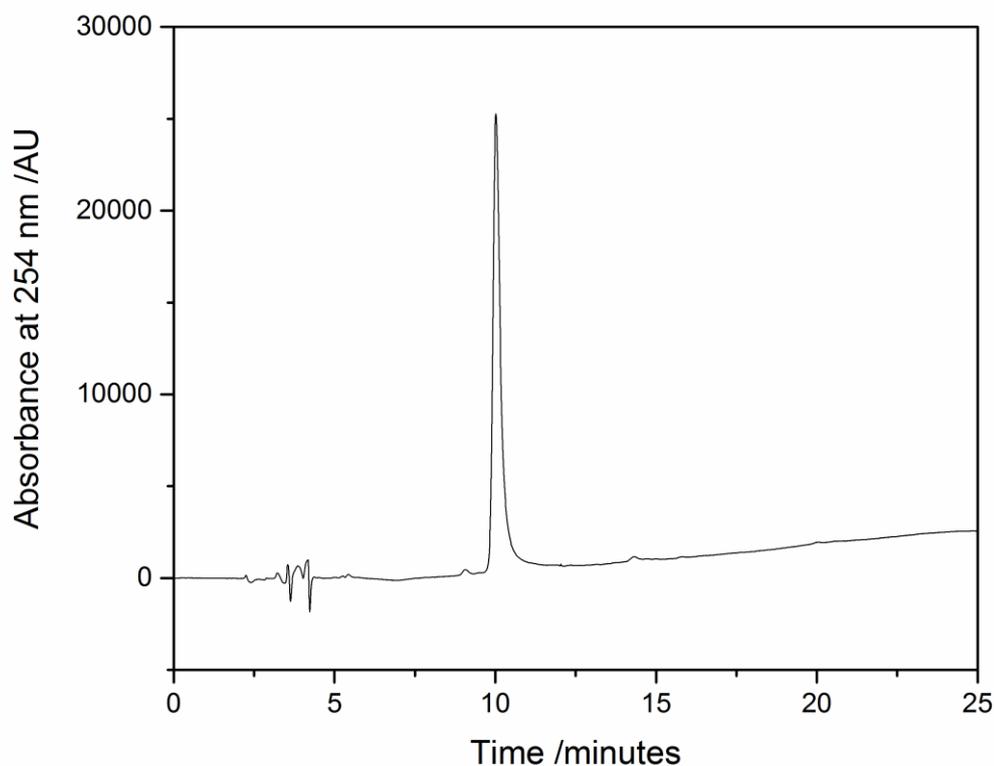


Figure 83: HPLC trace for Salmochelin S1. The sample in water containing 0.1% formic acid was injected with a 6-40% gradient of acetonitrile in water with 0.1% formic acid over 25 minutes.

Upon addition of one equivalent of iron(III) to the salmochelin S1 sample at a concentration of 0.5 mM at pH 7 in 0.40 mM Tris-HCl 150 mM NaCl, no typical red or purple colouration was observed. This was confirmed by UV-visible spectroscopy, with no absorbance for the distinctive metal to ligand charge transfer band typical for iron(III)-catecholate

complexes.^{66, 90, 254, 398, 399} Considering that the pK_a values for the catechol hydroxy groups were around 9 and 13,²⁸ the spectra were run across a range of neutral to basic pH values to promote deprotonation of the catechol units to facilitate iron(III) binding, but no metal to ligand charge transfer band was observed. For reference, a spectrum of the same concentration of iron(III)-5-LICAM was recorded, to show the intensity of absorbance bands that might be expected for iron(III)-Salmochelin S1. For iron(III)-5-LICAM a broad absorbance band was observed between 450 and 650 nm, a very similar absorbance band to those documented in the literature for iron(III)-4-LICAM and iron(III)-bisDHBS.^{90, 322}

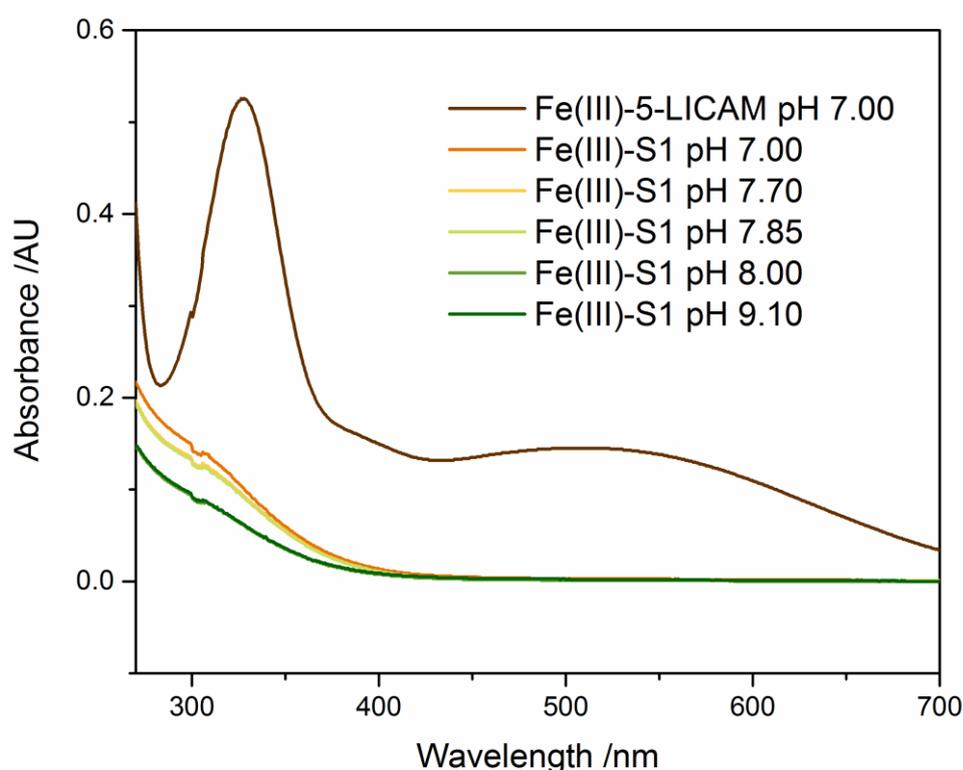
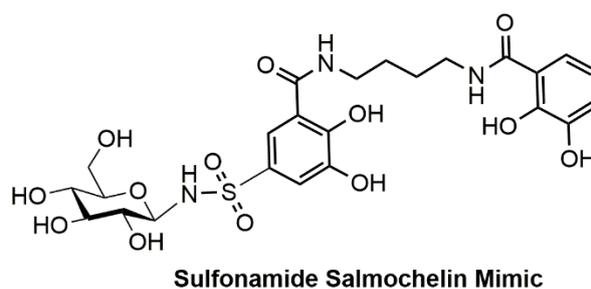


Figure 84: UV spectra of 0.5 mM iron(III)-5-LICAM at pH 7.0 40 mM Tris-HCl 150 mM NaCl overlaid with UV spectra of 0.5 mM iron(III)-S1 sample over a pH range of 7.0-9.1 in 40 mM Tris-HCl 150 mM NaCl.

These findings suggested that the sample either contained only a trace amount of salmochelin S1, detectable by ESI-MS and HPLC, or the sample contained a form of salmochelin S1 that is incapable of iron(III) chelation. As the sample supplied was unable to bind iron(III) in the characteristic catecholate mode, it was deemed unsuitable for further study.

Design of a Sulfonamide Linked Salmochelin Mimic

With no other commercial suppliers of salmochelin S1, it was necessary to design and synthesise a compound that could effectively mimic salmochelin S1, and could therefore be used to explore whether a glucose unit could be accommodated in the secondary binding pocket of CeuE. A logical design based on work detailed in Chapters 2 and 3, involved the use of a LICAM-type backbone, with a link to the anomeric carbon of a glucose moiety from the C5 position of one catechol aromatic ring. The first design (sulfonamide salmochelin mimic) involved a sulfonamide linkage between the aromatic catechol and the glucose unit.



Upon modelling of this mimic into the secondary binding cleft of CeuE, it was clear that the linker region was bulky, and the glucose unit was not in an optimum position in the secondary binding pocket. It appeared that an energetically disfavoured axial conformation would provide a better fit in the binding pocket than a more favourable equatorial conformation. With some doubt as to whether this mimic may allow for a good fit in the secondary binding pocket, as well as issues with initial synthetic reproducibility, this design was not pursued.

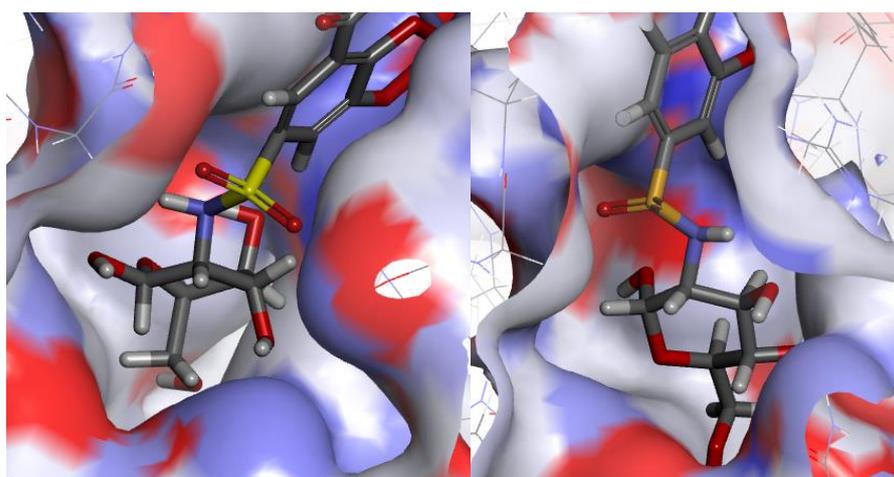
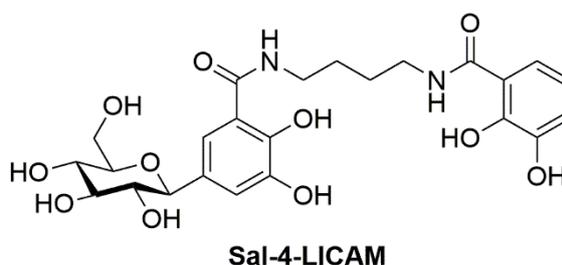


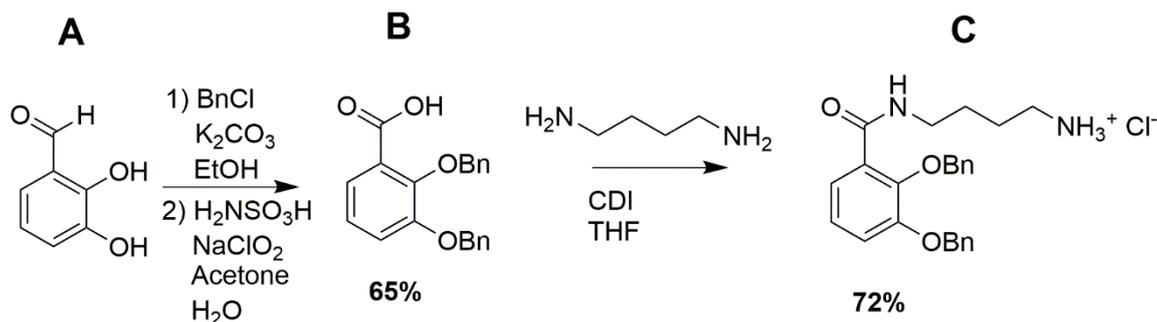
Figure 85: Simple modelling of the sulfonamide salmochelin mimic in the secondary binding pocket of CeuE. Left: in axial conformation. Right: in equatorial conformation.

Design of a Carbon-Carbon Bonded Salmochelin Mimic

The second salmochelin mimic design involved a more synthetically challenging C5 carbon-carbon bond formation between a LICAM catechol unit and a glucose unit (Figure 82). This was already proven as a favourable linkage for optimum glucose arrangement in the secondary binding pocket from initial modelling results (section 4.1) and was a closer match for salmochelin S1.



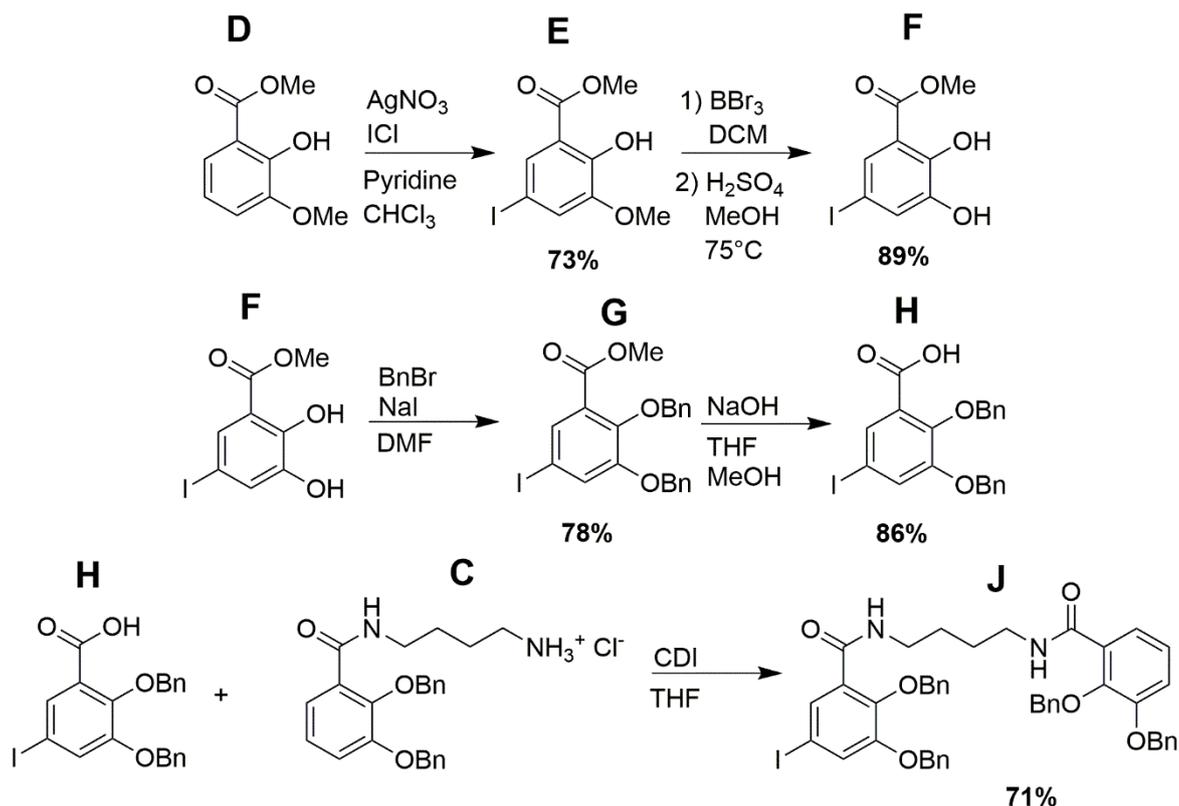
The first C-C bonded target compound (Sal-4-LICAM) was based on a 4-LICAM backbone. This linker length was chosen as the least synthetically challenging, as the synthesis of the intermediate, benzyl protected aminochelin (compound C), was already well documented in the literature.^{85, 106} Compound C was synthesised over two steps in an overall yield of 47% (Scheme 9). Firstly, the catechol oxygen atoms of Compound A were benzyl protected, followed by an oxidation of the aldehyde to a carboxylic acid. Addition of Compound B and CDI to an excess of 1,4-diaminobutane in a dropwise manner yielded Compound C.



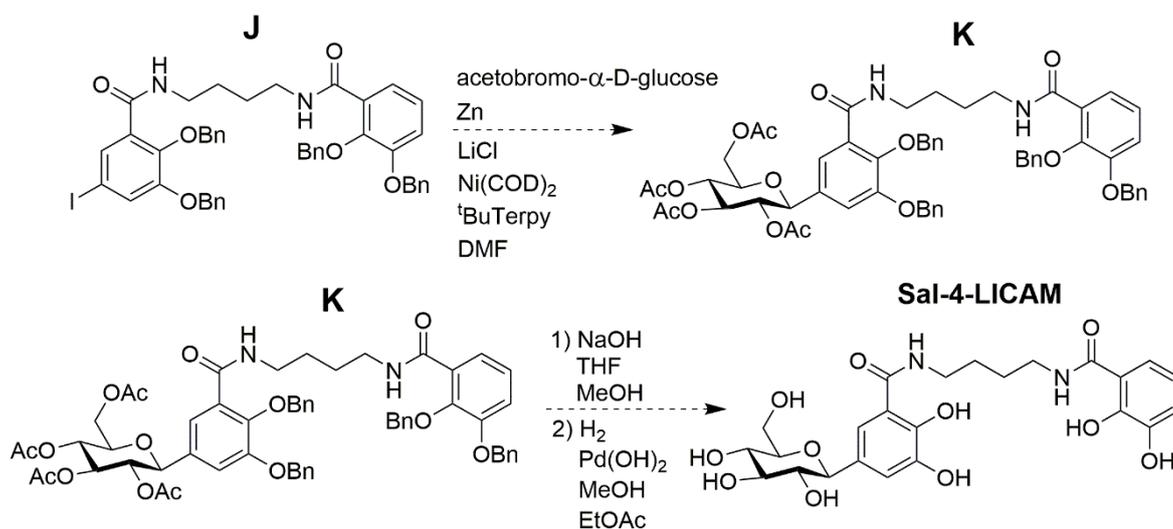
Scheme 9: Synthesis of benzyl protected aminochelin (compound C).

Salmochelin syntheses were documented in the literature, as discussed in section 4.1, the reported procedures for installation of the β -D-glucose unit were adapted for the synthesis of Sal-4-LICAM (Scheme 10, Scheme 11 and Scheme 12).³⁹⁵⁻³⁹⁷ Synthetic design was carried out in collaboration with Dr. T. J. Sanderson, who required intermediate compounds detailed in Scheme 12 for the production of a salmochelin inspired antibiotic conjugate.²⁹⁵ All compounds were synthesised and characterised individually, however synthetic strategy and resulting knowledge was shared.

The full synthesis of Sal-4-LICAM was initially designed with the production of the *C5* iodinated benzyl-protected 4-LICAM (compound **J**) over five steps from commercially available starting material (compound **D**) (Scheme 10). Compound **J** was designed as the starting material for the Negishi coupling to install a *C5* C-C bonded acetate-protected β -D-glucose unit (compound **K**) (Scheme 11).³⁹⁶ It was hoped that base deprotection of the acetate protecting groups, followed by hydrogenolysis to remove the benzyl groups would produce the desired target Sal-4-LICAM over seven synthetic steps.



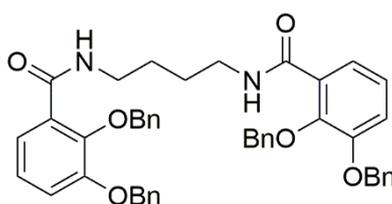
Scheme 10: Synthetic route for the production of the benzyl protected *C5* iodinated 4-LICAM species (Compound **J**).



Scheme 11: Proposed Negishi coupling reaction of Compound **J** to produce Compound **K** and subsequent deprotection reactions to yield Sal-4-LICAM.

Compound **J** was successfully produced in an overall yield of 31% in good purity (Scheme 10), as confirmed by ¹H NMR spectroscopy, with a characteristic doublet at 8.07 ppm for one of the aromatic protons ortho to the iodine, with a ⁴J_{H-H} coupling of 1.8 Hz to the other ortho proton. ¹³C NMR spectroscopy gave characteristic peaks for the two amide carbonyls at 164.94 and 163.42 ppm. The characterisation of compound **J** was also validated with ESI-MS and elemental analysis.

Unfortunately, the Negishi coupling of compound **J** with acetobromo- α -D-glucose proved not to be viable. Rigorous moisture and air free techniques were used in an attempt to obtain compound **K**, but only trace amounts were observed with a peak for the [M+Na]⁺ species with *m/z* of 1073.4122 (calculated 1073.4024, mean error -7.4 ppm) observable by ESI-MS. The major product isolated was benzyl protected 4-LICAM, whereby the iodine in the C5 position had been replaced with a proton.

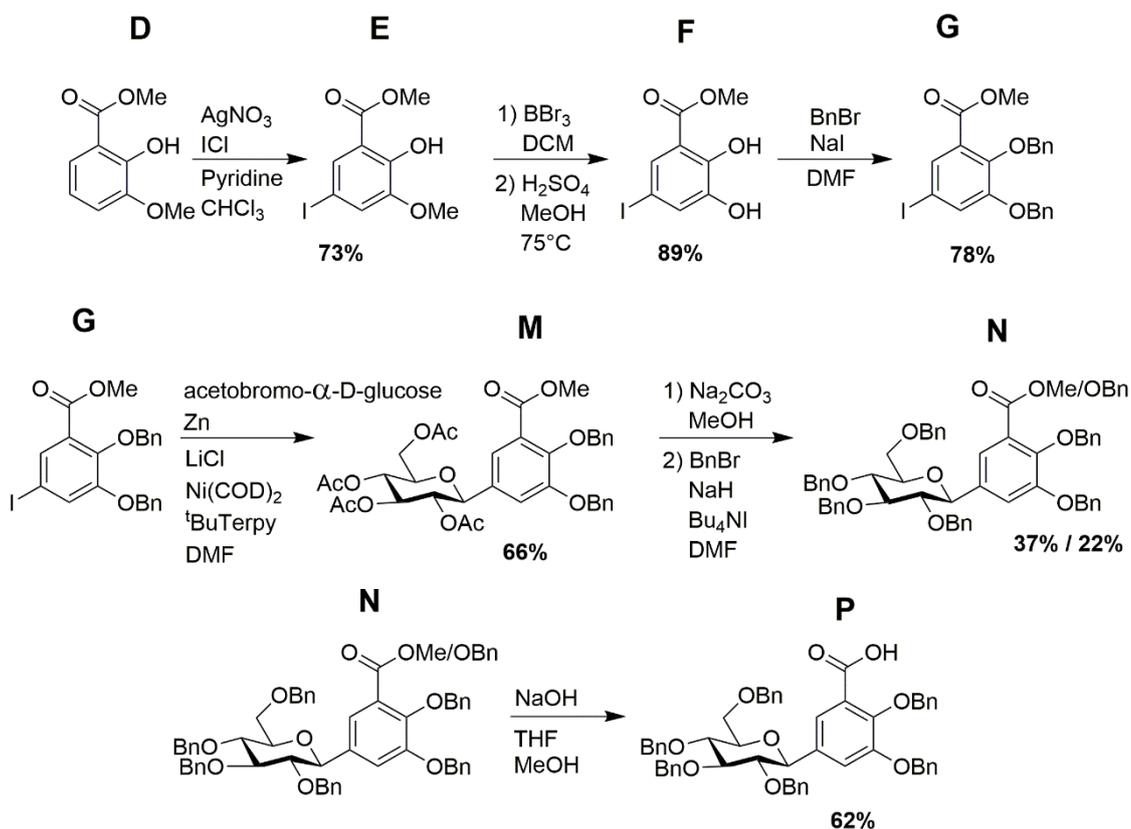


Benzyl Protected 4-LICAM

The presence of this compound was confirmed by ESI-MS with a peak for the [M+Na]⁺ species with *m/z* of 743.3048 by ESI-MS. ¹H NMR showed an absence of four distinct peaks for each set of benzyl CH₂ protons for an unsymmetrical compound at 5.17 ppm, 5.12 ppm, 5.06 ppm and 5.01 ppm. Instead two resonances were observed at 5.28 ppm and 5.18 ppm,

with relative integrations for four protons each, arising from the symmetry that resulted from the exchange of the iodine for a proton. It was not clear why compound **J** was not compatible with the Negishi coupling reaction conditions (Scheme 11). Elemental analysis and ^1H NMR analysis of the starting material ruled out the possibility of contamination of the starting material with water, or other solvent. It was however noted in the literature that Negishi couplings were slower with bulkier substrates.^{396, 400} It was necessary for compound **J** to be converted to an activated arylzinc reagent before the coupling reaction, and it is likely that either this arylzinc compound was not formed, or its decomposition was faster than the desired *C-C* bond formation. It is unclear why this was the case, but literature studies suggested that steric bulk may have been a factor.^{396, 400}

Due to the synthetic problems encountered, the synthesis of Sal-4-LICAM was redesigned with a smaller iodinated substrate for Negishi coupling (compound **G**) (Scheme 12). The Negishi coupling step to yield compound **M** was planned on a monocatechol substrate (compound **G**), with the amide coupling step to produce the 4-LICAM backbone planned at a later stage of the synthesis (Scheme 13), after installation of the glucose-containing unit (compound **M**). This newly designed route required ten synthetic steps, three more than the previously designed route, but had the advantage of a reported successful Negishi coupling of a substrate known in the literature, and therefore only required small alterations to the documented synthetic method.^{396, 397}

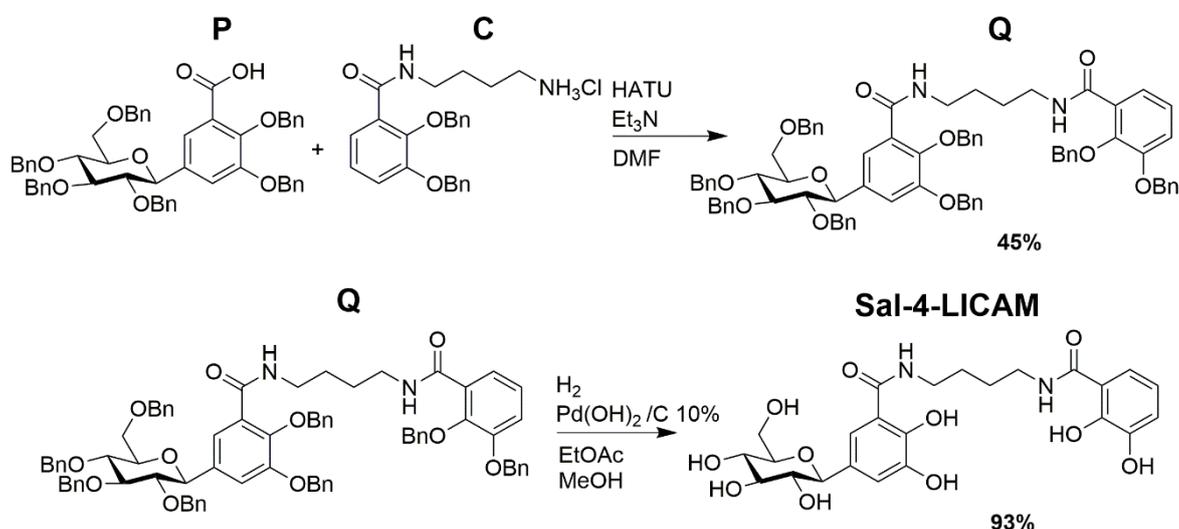


Scheme 12: Synthetic route for the production of Negishi coupled product, compound **P**.

First, a commercially-available mono methyl protected catechol starting material (compound **D**) was iodinated in the *C5* position to yield compound **E** in 73% yield. Methyl protecting groups allowed for the necessary functional group directing effects for selective iodination in the *C5* position.^{295, 401} The methyl group protecting the catechol was next removed to yield compound **F** in 89% yield, followed by benzyl protection of both the catechol hydroxy groups to afford compound **G** in a 78% yield.

The benzyl-protected iodinated catechol product (compound **G**) was then used for a Negishi coupling reaction, whereby a carbon-carbon bond formation allowed for the installation of an acetate-protected β -D-glucose unit in the *C5* position to yield compound **M** in a 66% yield. The glucose unit was selectively installed as the β -anomer, as confirmed by the coupling constant of the anomeric proton observed in the ^1H NMR spectrum. Literature $^3J_{\text{H-H}}$ coupling constants for diaxial protons, as present in the β -configuration, have a large value, typically between 7 and 9 Hz. If the α -configuration was present, the coupling constant would be much smaller, typically between 2 and 4 Hz, for an equatorial-axial proton coupling.^{392, 402} The doublet at 4.39 ppm for the anomeric proton in compound **M** had a coupling constant of $J = 10.1$ Hz, confirming the presence of the β -anomer.³⁹⁷

The acetate protecting groups on the glucose unit were required in the Negishi coupling for selective formation of the β -anomer due to their role in neighbouring group participation.^{295, 396, 403} Acetate protecting groups however proved unsuitable for later synthetic steps, due to their instability under basic conditions.⁴⁰⁴ As such, the acetate protecting groups were removed, and benzyl protecting groups installed in their place to yield both the methyl and benzyl ester compounds **N**. The compounds were individually purified by column chromatography to allow for their characterisation, however as the following step required deprotection to yield the carboxylic acid (compound **P**), the two products were recombined. The percentage yield, in terms of the combined moles of each species was calculated as 59%. A base deprotection yielded the free carboxylic acid in a 62% yield (compound **P**).



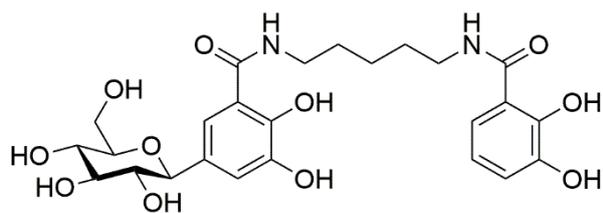
Scheme 13: Amide coupling of the glucosylated catechol unit to benzyl-protected aminochelin, and the subsequent global benzyl deprotection *via* hydrogenolysis.

The free carboxylic acid species (compound **P**) was combined with benzyl protected aminochelin (compound **C**) in an amide coupling reaction, using the coupling agent HATU, to produce the benzyl-protected, mono-glucosylated 4-LICAM, compound **Q**, in 45% yield. After purification, all benzyl protecting groups were removed in one step *via* hydrogenolysis to yield Sal-4-LICAM over 10 synthetic steps in an overall yield of 5%. The specific rotation was recorded as +3.2 at 0.245 g/100 mL in methanol, and the signal for the anomeric proton in the ^1H NMR spectrum was observed as a doublet at 4.02 ppm with a coupling constant of $J = 9.2$ Hz. This proved that the β -anomer was retained in the final product, as required to act as a suitable mimic of the β -configured glucose unit in salmochelin S1.

The fact that the successful Negishi coupling reaction conditions for the conversion of compound **G** to compound **M** (Scheme 12) were almost identical to those used for the

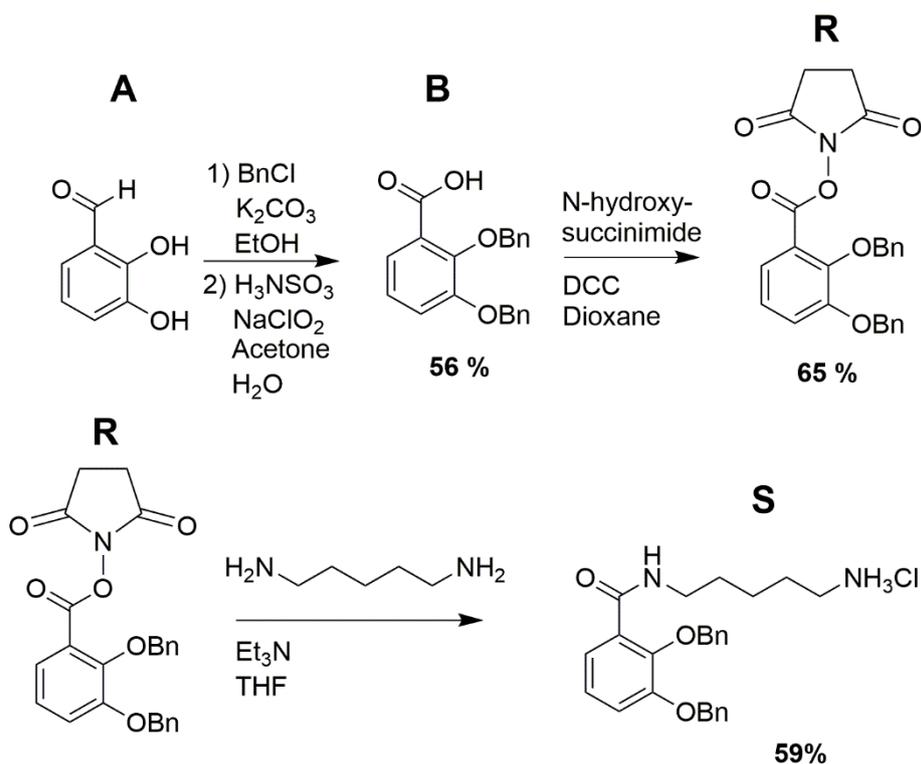
attempted reaction of compound **J** to compound **K** (Scheme 11) further suggested that there was substrate incompatibility of compound **J**, rather than the possibility of a contamination.

Following successful synthesis of Sal-4-LICAM, the 5-LICAM analogue was synthesised. This salmochelin mimic was a more desirable synthetic target, as it contains a 5-atom linker that matches the linker length of salmochelin S1.¹⁵⁴ In addition, previous studies in Chapter 2 showed that iron(III)-5-LICAM was the tetradentate siderophore mimic, in the n-LICAM (n = 4, 5, 6, 8) series, with the highest binding affinity for CeuE.³¹⁰



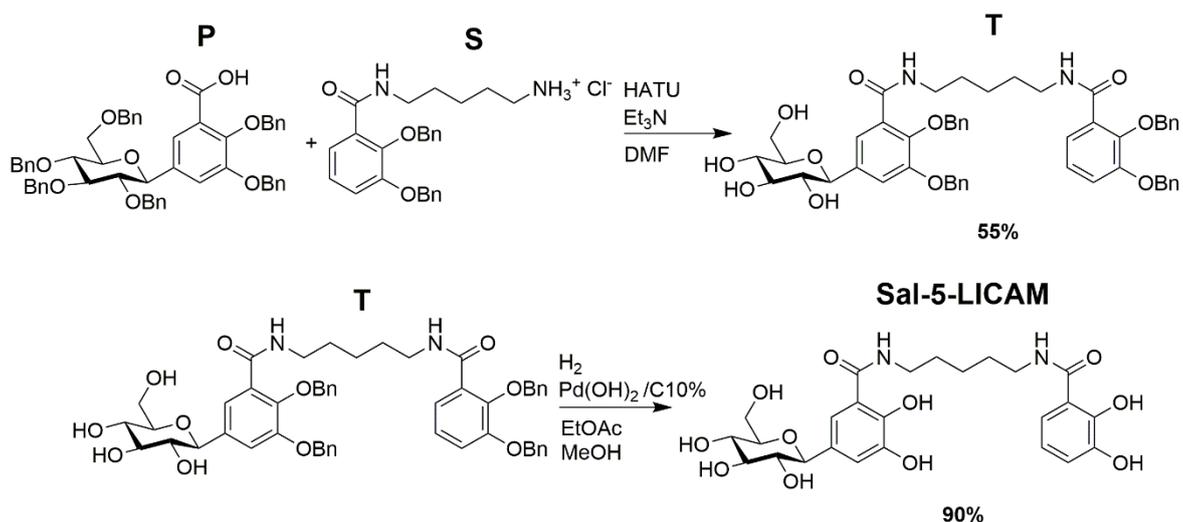
Sal-5-LICAM

Isolation of the benzyl protected aminochelin analogue containing a 5-carbon linker (compound **S**) proved problematic under the same conditions as the synthesis of compound **C** (Scheme 9). This was likely due to the amphiphilic nature of the mono-amine product, and incompatibility with the work up procedure. The synthesis of compound **S** was therefore redesigned (Scheme 14). The method was based on literature procedures for synthesis of 4-LICAM.^{15, 186, 405}



Scheme 14: Scheme for the synthesis of a 5-carbon linker analogue of benzyl protected aminochelin (compound **S**).

This synthesis, carried out by project student R. P. Thomas, although requiring three synthetic steps instead of two, afforded compound **S** in an overall yield of 21%. Compound **S** could then be coupled to the glucose-functionalised catechol product (compound **P**) in a similar manner to that of the Sal-4-LICAM procedure (Scheme 13), yielding a benzyl-protected 5-LICAM salmochelin mimic in a 55% yield (compound **T**) (Scheme 15). Hydrogenolysis was used to remove all benzyl groups to produce the final 5-LICAM salmochelin mimic in a 90% yield. The final product, Sal-5-LICAM was obtained in an overall yield of 6%, over the longest synthetic route.



Scheme 15: Amide coupling of the glucosylated catechol unit to benzyl protected 5-carbon analogue of aminochelin, and the subsequent global benzyl deprotection *via* hydrogenolysis.

Both salmochelin mimics were characterised by ¹H and ¹³C NMR, ESI-MS, IR and elemental analysis, and all intermediates were characterised by ¹H and ¹³C NMR, ESI-MS and IR as a minimum. HMQC, COSY and HMBC were used to aid NMR assignments and supplemented with HMBC where necessary. All data are presented in Chapter 6. Sal-4-LICAM was found to be in a form containing 2.1 equivalents of H₂O that could not be removed with extensive drying methods. Sal-5-LICAM was found to be in a form containing 0.9 equivalents of EtOH and 1.2 equivalents of H₂O that could not be removed with extensive drying methods. The present solvents were observed by ¹H NMR, and CHN elemental analysis data correlated with the calculated equivalents. The solvents were taken into account for further calculations and studies. All further studies were carried out with portions of each salmochelin mimic that had been analysed *via* all characterisation techniques, including elemental analysis.

4.3 Iron Binding

Each salmochelin mimic, Sal-4-LICAM and Sal-5-LICAM, was studied for iron binding stoichiometry using a Job Plot method adapted from a method documented for the iron(III) binding stoichiometry of bisDHBS.⁹⁰ A Job plot, also known as the Continuous Variation Method, was used to determine the stoichiometry of a binding event by studying the physical properties of the system over a range of ratios.⁴⁰⁶ UV-visible spectroscopy is a common technique for determination of speciation of coloured metal complexes,^{90, 281, 322, 406, 407} but techniques such as NMR can also be employed.^{408, 409} The method has previously been used for monocatecholate substrates epinephrine and norepinephrine for the study of their iron binding stoichiometries.⁴¹⁰ The absorbance of solutions of iron(III)-Sal-n-LICAM (n= 4,5) of varying iron(III) : ligand ratios with a total concentration of 400 nM, were recorded. For the iron(III) complexes of Sal-4-LICAM, the λ_{\max} was observed to shift across the range of ratios, from around 492 nm to around 552 nm as iron(III) to ligand ratio was increased (Figure 86). The colour of the range of solutions transitioned from red to purple as the iron(III) to ligand ratio was increased, as would be expected from the shift in λ_{\max} in the UV-visible spectra (Figure 87).

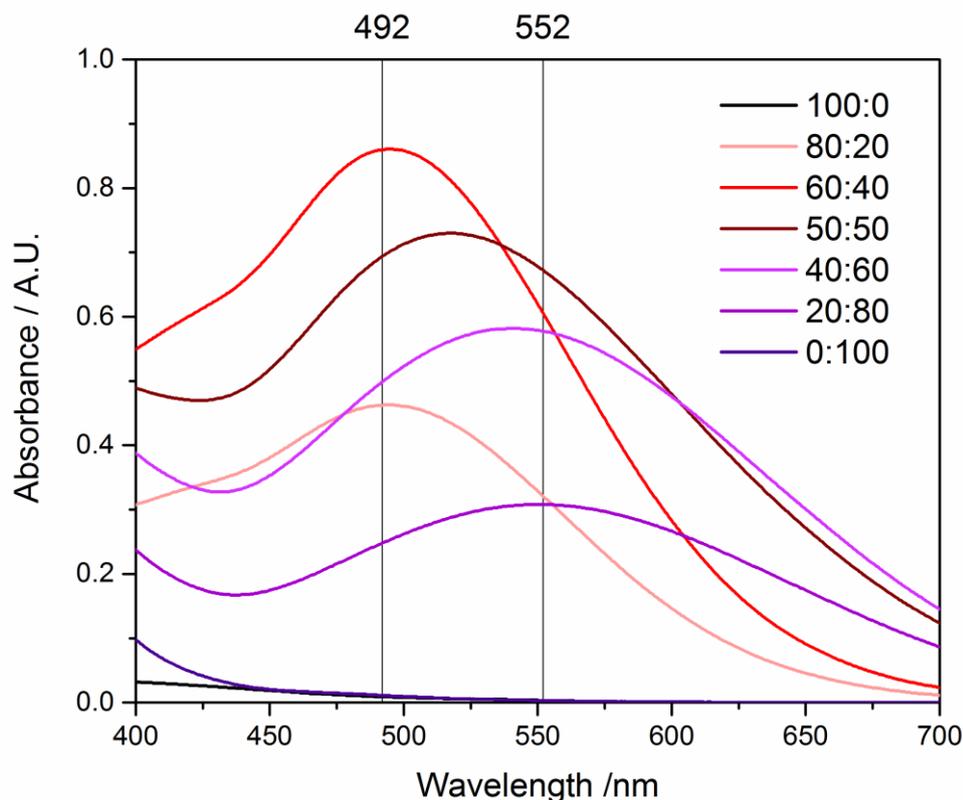


Figure 86: Selected UV-Visible spectra of Sal-4-LICAM: iron(III) in the ratios as shown, with a final concentration of 400 nM. Spectra were recorded after 1 hour of equilibration. Vertical lines are shown at the λ_{\max} wavelengths.

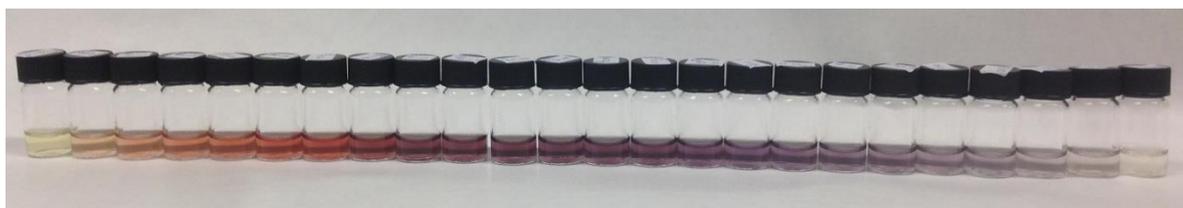


Figure 87: Photograph of sample vials of Sal-4-LICAM: iron(III) with highest ligand concentration to lowest ligand concentration from left to right, and highest iron(III) concentration to lowest iron(III) from right to left.

Plotting the absorbance at 552 nm and 492 nm for each spectrum over the range of Sal-4-LICAM: iron(III) ratios, gave the Job plot shown in Figure 88. The maximum absorbance at 552 nm was at a 50:50 Sal-4-LICAM: iron(III) ratio, and at 492 nm was at a 60:40 Sal-4-LICAM: iron(III) ratio. This indicated an equilibrium between a 1:1 species and a 3:2 species in solution. This is a similar set of equilibria to that observed for bisDHBS : iron(III), where exact speciation was confirmed by NMR, using gallium(III) as a diamagnetic replacement for paramagnetic iron(III).⁹⁰

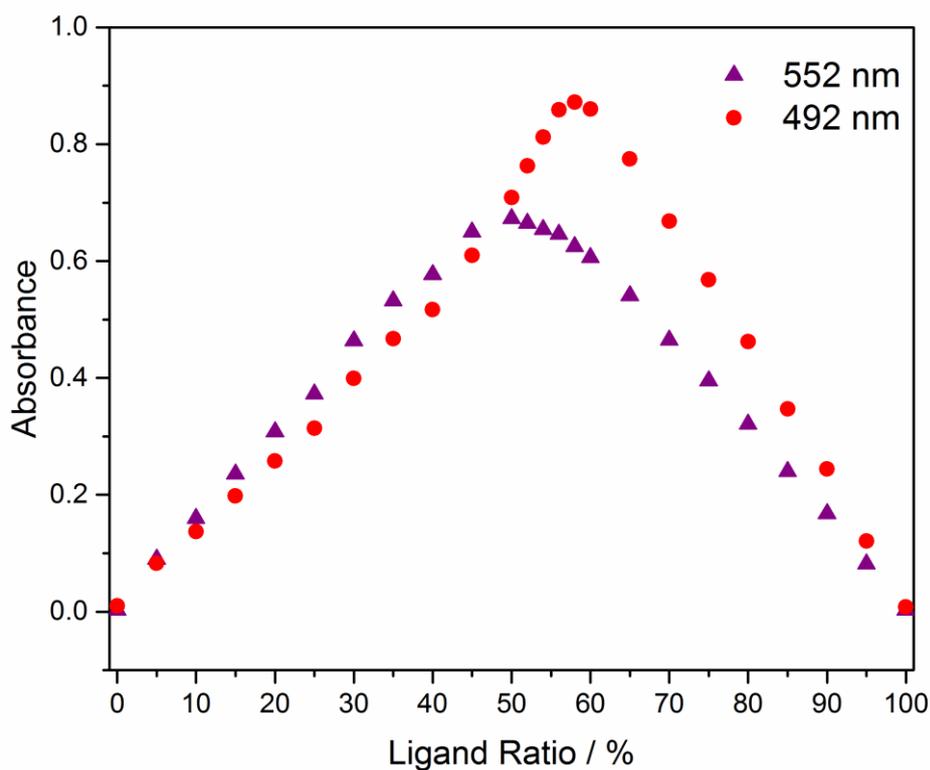


Figure 88: Job plot for Sal-4-LICAM: iron(III), with the x-axis showing the ratio of ligand to iron(III) in %. Absorbance data are plotted at 552 nm (purple triangles) and 492 nm (red circles) across the ratio range, showing a peak at a 50:50 Sal-4-LICAM: iron(III) ratio for absorbance at 552 nm, and a peak at 60:40 Sal-4-LICAM: iron(III) ratio for absorbance at 492 nm.

After 7 days, the spectra were recorded again, to establish whether equilibrium had been reached after 1 hour, or whether a slower equilibrium was present. For both wavelengths, 552 nm and 492 nm, the maximum absorbance was achieved at a 60:40 ratio of Sal-4-LICAM: iron(III) (Figure 89). This indicated that the 3:2 species predominated over time, rather than the previously observed 1:1 species.

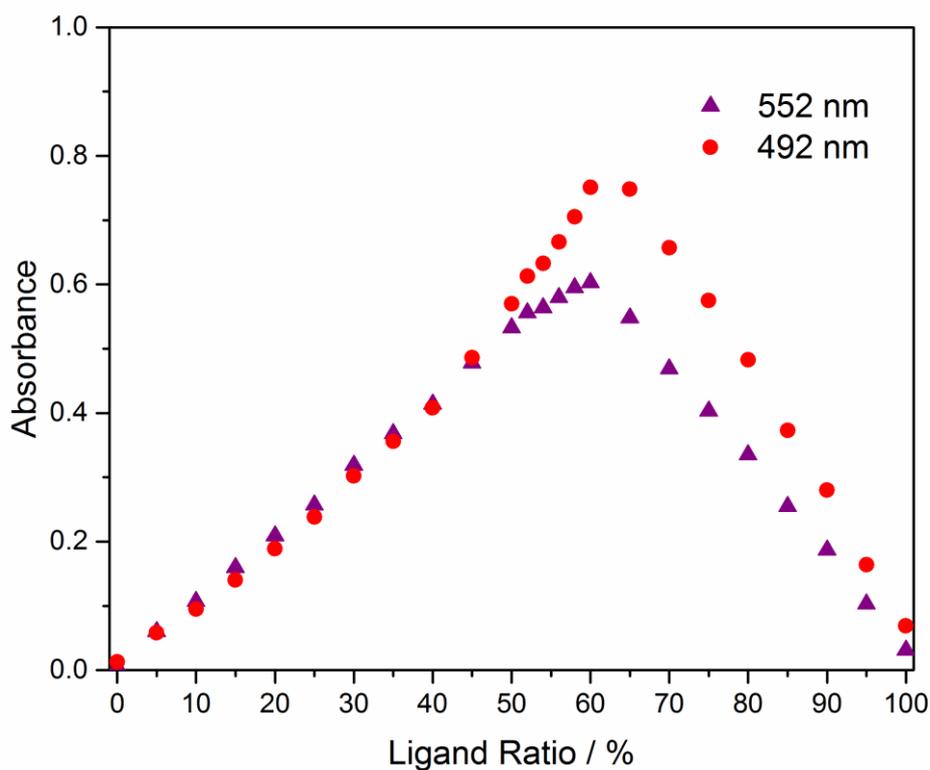


Figure 89: Job plot for Sal-4-LICAM: iron(III) after 7 days, with the x-axis showing the ratio of ligand relative to the percentage of iron(III). Absorbance data are plotted at 552 nm (purple triangles) and 492 nm (red circles) across the ratio range, showing peaks at 60:40 Sal-4-LICAM: iron(III) ratio for absorbance at 552 nm and 492 nm.

The studies were then repeated with Sal-5-LICAM. The λ_{\max} , recorded across the same range of ratios, was shifted from around 495 nm to around 555 nm as iron(III) to ligand ratio was increased (Figure 90). A similar array of coloured solutions was observed across the range of ratios, with a gradient from red to purple as the iron(III) to ligand ratio was increased (Figure 91).

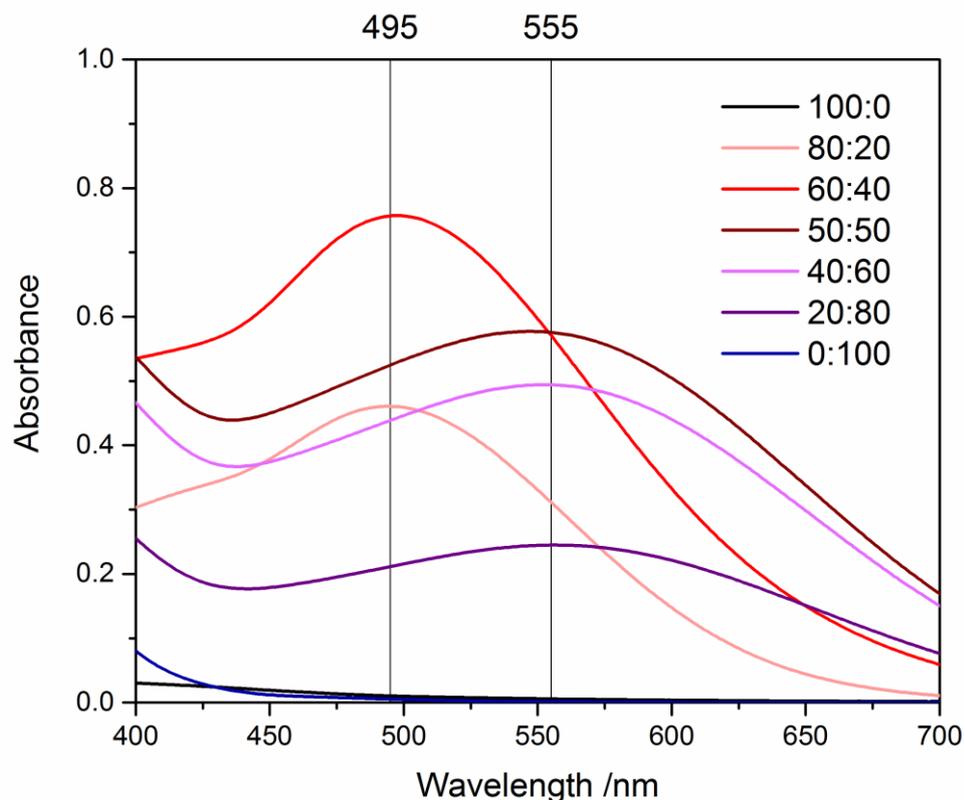


Figure 90: Selected UV-Visible spectra of Sal-5-LICAM: iron(III) in the ratios as shown, with a final concentration of 400 nM. Spectra were recorded after 1 hour of equilibration. Vertical lines are shown at the λ_{\max} wavelengths.



Figure 91: Photograph of sample vials of Sal-5-LICAM: iron(III) with highest ligand concentration to lowest ligand concentration from left to right, and highest iron(III) concentration to lowest iron(III) from right to left.

The Job plot showed a similar result to the iron(III) ligand speciation for Sal-4-LICAM, with a 1:1 and 3:2 species, shown by these ratios resulting in the maximum absorbance at each λ_{\max} wavelength of 555 nm and 495 nm respectively (Figure 92). After 7 days, the maximum absorbance at both λ_{\max} wavelength resulted from the 60:40 ratio of Sal-5-LICAM: iron(III), suggesting that the 3:2 species predominated after longer equilibration time (Figure 93).

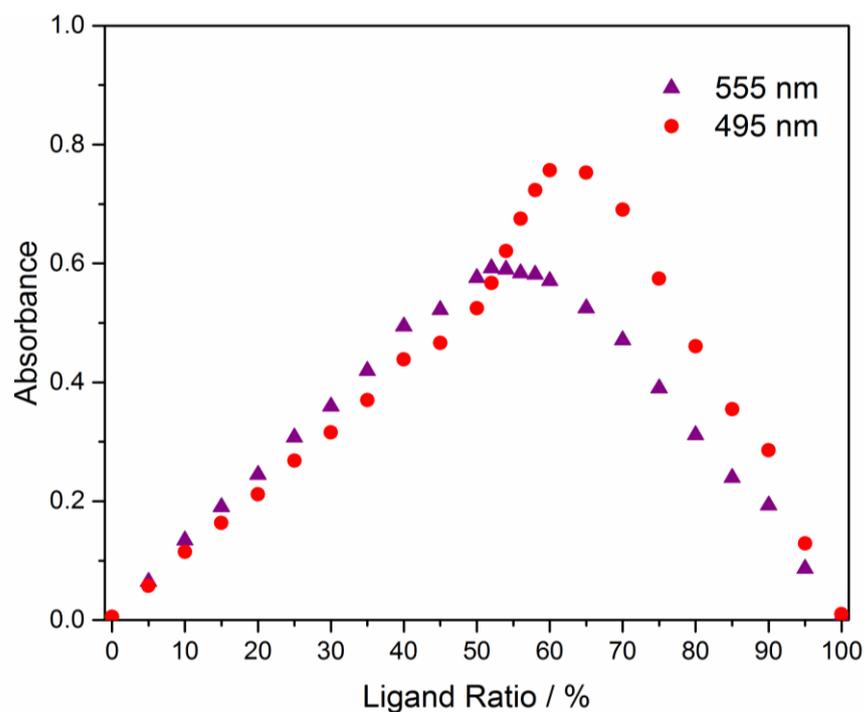


Figure 92: Job plot for Sal-5-LICAM: iron(III), with the x-axis showing the ratio of ligand as a percentage of total Sal-5-LICAM: iron(III). Absorbance data are plotted at 555 nm and 495 nm, showing a peak at a 50:50 Sal-4-LICAM: iron(III) ratio for absorbance at 555 nm, and a peak at 60:40 Sal-4-LICAM: iron(III) ratio at 495 nm.

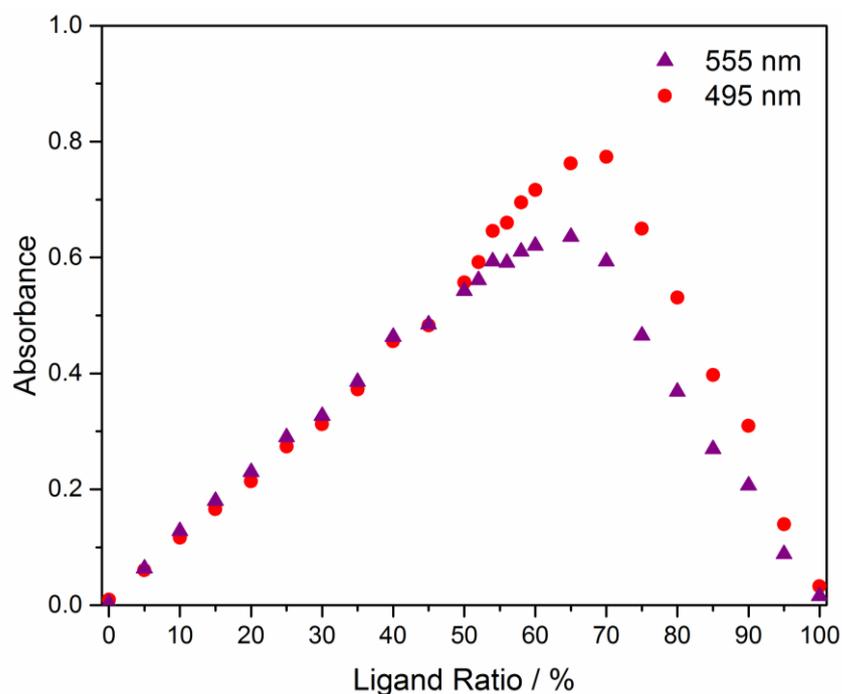


Figure 93: Job plot for Sal-5-LICAM: iron(III) after 7 days, with the x-axis showing the ratio of ligand as a percentage of total Sal-5-LICAM: iron(III). Absorbance data are plotted at 555 nm and 495 nm, showing a peak at a 50:50 Sal-4-LICAM: iron(III) ratio for absorbance at 555 nm, and a peak at 60:40 Sal-4-LICAM: iron(III) ratio at 495 nm.

The similar λ_{\max} values of 492 nm and 552 nm for iron(III): Sal-4-LICAM and 495 nm and 555 nm for iron(III): Sal-5-LICAM, with very similar Job plots, indicate that the difference in linker length between the two salmochelin mimics has very little effect on the iron(III) binding ratios of the species. For both iron(III): Sal-n-LICAM (n = 4, 5), initially there is an equilibrium of 1:1 and 3:2 Ligand: iron(III) species that have maximum absorbances at the corresponding Ligand: iron(III) ratios. This is similar to the equilibria present for iron(III) : bisDHBS and iron(III) : 4-LICAM.^{90, 322} After 7 days of equilibration, it appears that the 3:2 becomes the major species, as the maximum absorbance at both wavelengths peaks with the 3:2 ratio. It is unclear why a 3:2 ratio is favoured over 1:1, as it could be reasoned that the steric bulk of the glucose unit would disfavour close arrangement of three ligands around two iron(III) centres, when compared to the LICAM siderophores. As the ligands are not symmetrical, and have one glucose unit, it is likely that a number of geometric isomers exist in the 3:2 ratio, where the glucose units either align at a single iron(III) centre or alternate between the two centres.

Salmochelin siderophores are known to have a reduced membrane partition coefficient when compared to their enterobactin equivalents.²³⁴ It may be that the hydrophilic nature of the glucose units, and their propensity to hydrogen bond in the water based buffer allows the ligands to form a favourable 3:2 arrangement with solvent exposed sugars somewhat shielding the aligned hydrophobic LICAM backbones. To explore this further it would be necessary to study a wider coordination sphere of the iron(III) centres, and particularly the role of hydrogen bonding of the glucose-containing ligands, to understand all interactions involved in complex assembly.⁴¹¹⁻⁴¹³

4.4 Circular Dichroism

Co-crystallisation was tried for both iron(III)-Sal-n-LICAM (n = 4, 5) with CeuE, but unfortunately no crystals were obtained. It was therefore necessary to employ other methods to establish whether iron(III)-Sal-n-LICAM (n= 4, 5) were able to bind to CeuE.

Sal-n-LICAM (n= 4,5) were studied by circular dichroism to identify whether there was any chiral preference of iron(III) complexes in solution. When complexed with iron(III) in a 1:1 ratio, the circular dichroism signal remained close to the baseline for both Sal-n-LICAM (n= 4,5) complexes in solution, indicating a lack of chiral preference (Figure 94).¹³ It is therefore likely that there was a mixture of both Λ and Δ configured complexes in solution for both ligand complexes.¹¹

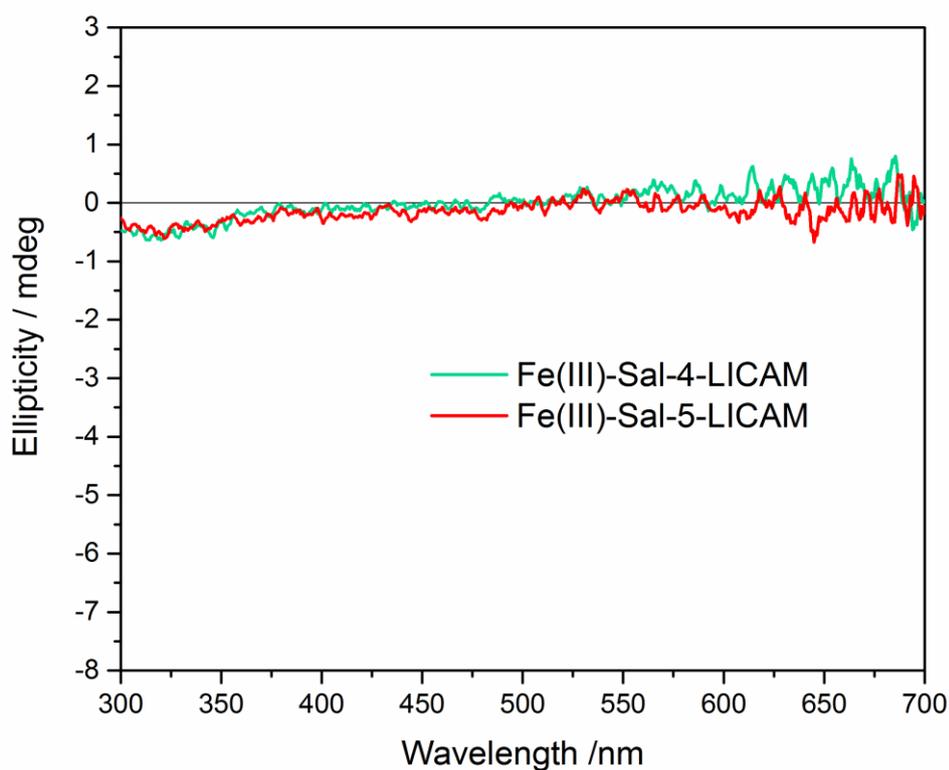


Figure 94: Circular dichroism spectra for 50 mM iron(III)-Sal-4-LICAM and iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer.

The salmochelin mimics Sal-*n*-LICAM (*n*= 4,5) were next studied for their potential interactions with CeuE by circular dichroism. CeuE induced a slight preference for the Λ -configured complex of iron(III)-Sal-4-LICAM (Figure 95), indicating that there must be a CeuE-iron(III)-Sal-4-LICAM interaction in solution. Interestingly, CeuE induced a more significant preference for the Λ -configured complex of iron(III)-Sal-5-LICAM (Figure 95), with the larger magnitude suggesting higher binding affinity of CeuE for iron(III)-Sal-5-LICAM than for iron(III)-Sal-4-LICAM.

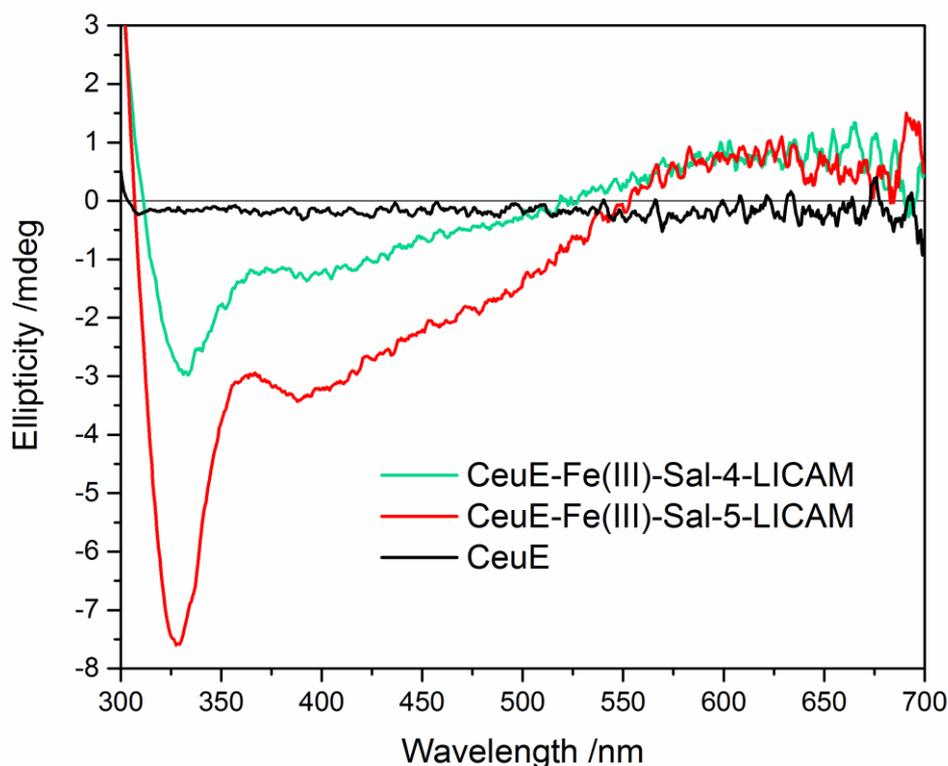


Figure 95: Circular dichroism spectra of 50 mM CeuE, CeuE-iron(III)-Sal-4-LICAM and CeuE, iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer. The spectra showed a small Λ -configurational preference upon introduction of CeuE to iron(III)-Sal-4-LICAM, which increases upon introduction of CeuE to iron(III)-Sal-5-LICAM.

The circular dichroism results suggested that CeuE did indeed interact with the synthesised salmochelin mimics iron(III)-Sal- n -LICAM ($n=4,5$). To understand the interactions further, it was necessary to quantify the binding affinities of iron(III)-Sal- n -LICAM ($n=4,5$) with CeuE.

4.5 Determination of Dissociation Constants

CeuE-iron(III)-Sal-4-LICAM

The standard fluorescence quenching titration method, Chapter 6, was employed to identify whether iron(III)-Sal-4-LICAM was able to bind to CeuE. The curves obtained were shallow, and a defined end-point was not reached, suggesting a weak binding affinity (Figure 96). As such, any calculated dissociation constants were unlikely to be accurate to the true value due to the shallow curves obtained.⁴¹⁴

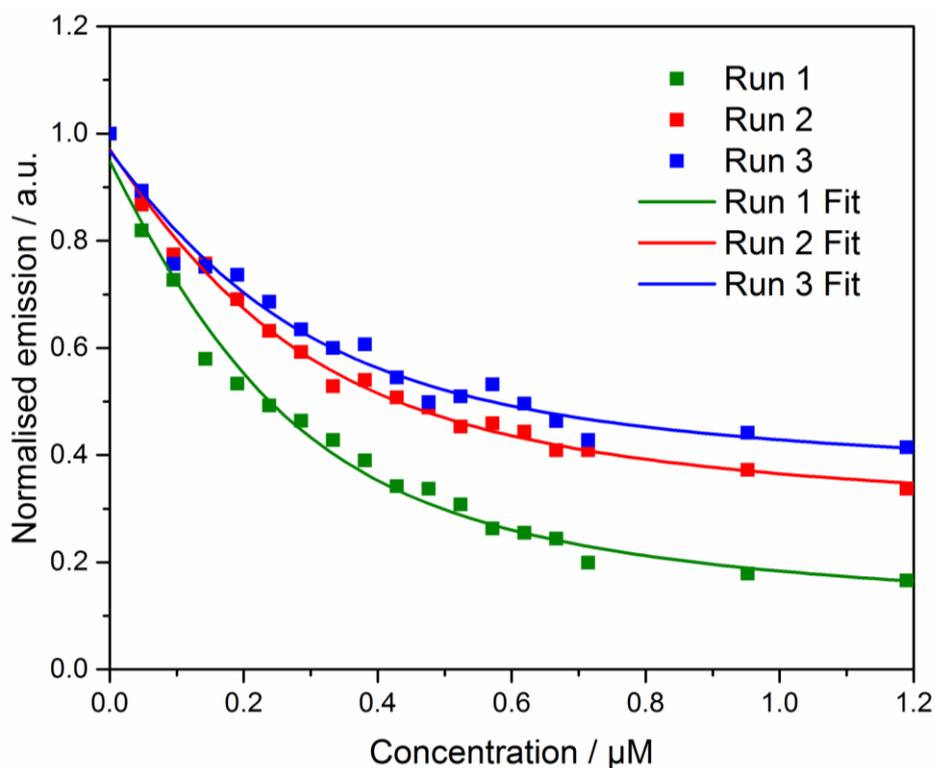


Figure 96: Fluorescence quenching titration curves for binding of iron(III)-Sal-4-LICAM to 0.24 μM CeuE 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

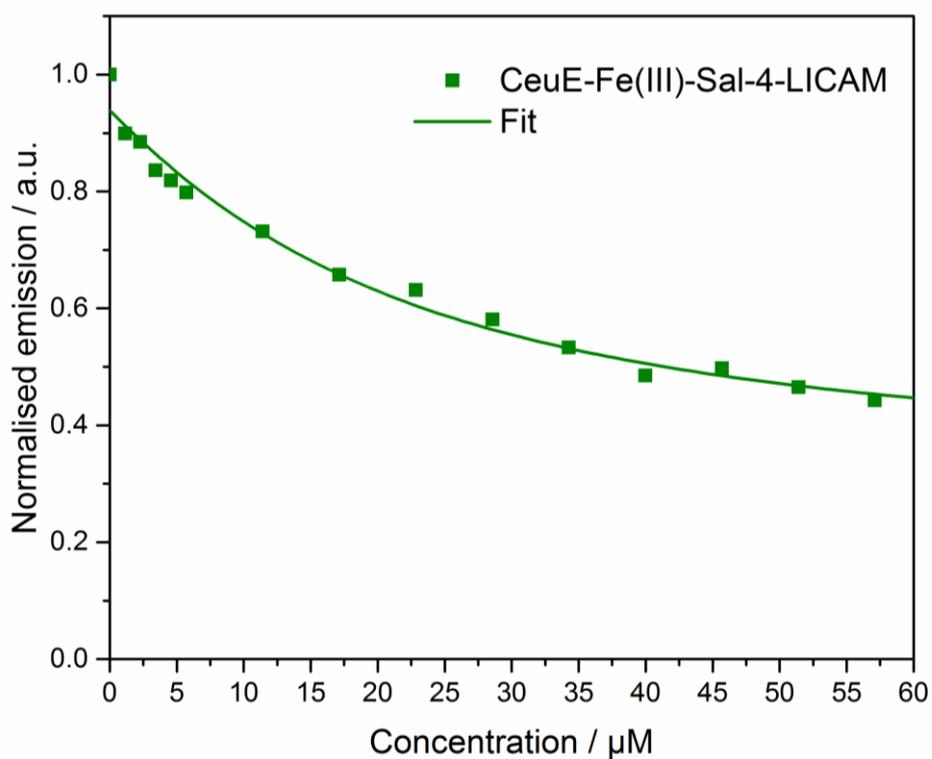


Figure 97: Fluorescence quenching titration curve for binding of iron(III)-Sal-4-LICAM to 12 μM CeuE in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Fluorescence quenching titrations were attempted at an increased CeuE concentration of 12 μM (established as the upper permissible CeuE concentration in Chapter 3) to promote CeuE-iron(III)-Sal-4-LICAM complex formation, but similar a shallow curve was obtained, confirming weak binding. Because the dissociation constant could not be accurately quantified from fluorescence quenching titration, a circular dichroism titration method was developed. Primary inner filter effect caused by ligand absorbance at the excitation wavelength limits the maximum ligand concentration that is possible *via* fluorescence quenching titration (Chapter 3). Circular dichroism can be run at higher concentrations without such problems assuming the ligand does not absorb at the wavelength of interest, as was demonstrated in the featureless circular dichroism spectrum of iron(III)-Sal-4-LICAM in Figure 94. Complex formation is more favourable at higher protein and ligand concentrations, so this method allowed for sufficient complex formation to produce an adequate binding curve.^{414, 415} The increase in signal observed upon additions of aliquots of CeuE to iron(III)-Sal-4-LICAM at 50 μM is shown in Figure 98. 330 nm was chosen as the wavelength to monitor, as the largest change in signal was observed in this region. Ellipticity at 330 nm was plotted against concentration to produce binding curves for triplicate titrations (Figure 99).

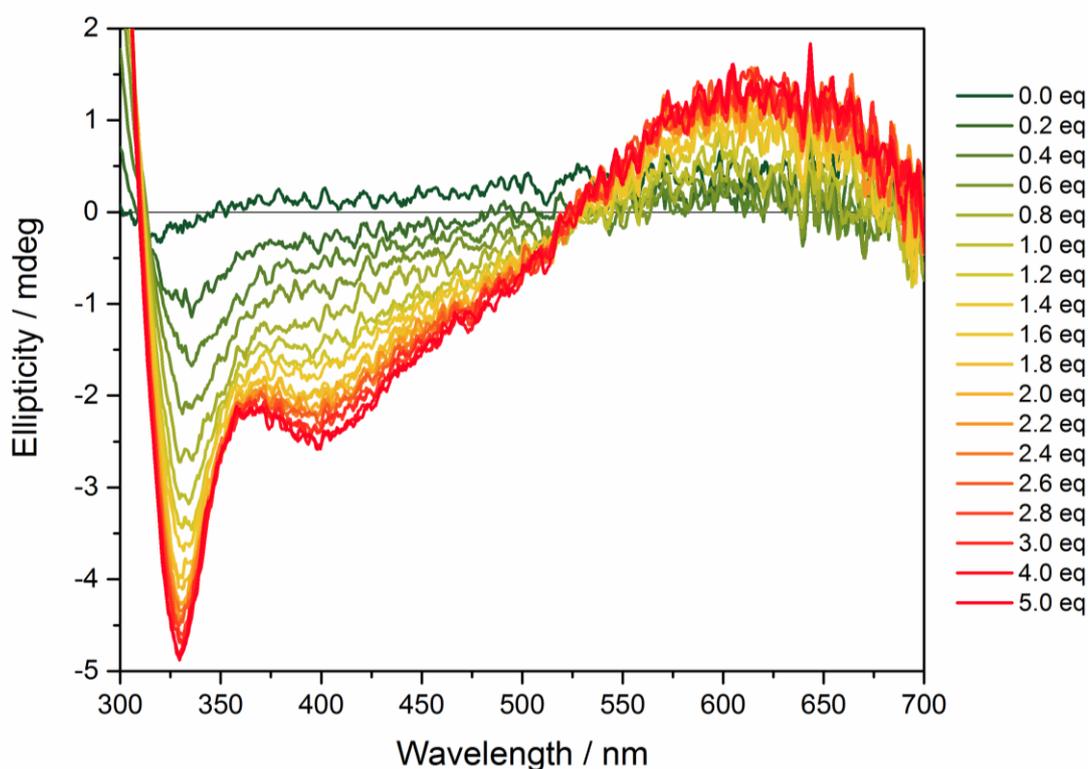


Figure 98: Circular dichroism spectra for the titration of 50 μM iron(III)-Sal-4-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl with aliquots of CeuE in 0.11M Tris-HCl pH 7.5, 150 mM NaCl.

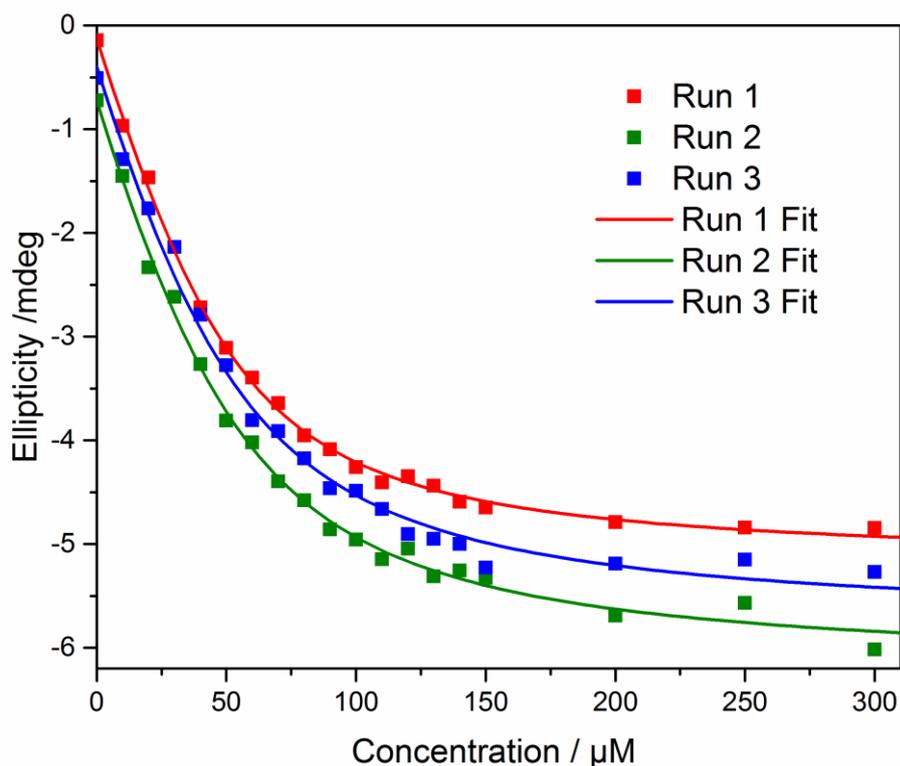


Figure 99: Binding curves for three independent circular dichroism titrations of 50 μM iron(III)-Sal-4-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl titrated with CeuE. Ellipticity at 330 nm was plotted against CeuE concentration. Non-linear regression fitting data was generated with Dynafit.

From the three independent titrations, dissociation constants were calculated with their associated errors and then averaged using a weighted average calculation that is detailed in Chapter 6 (Table 13). The shapes of the obtained curves were more acceptable with a defined end-point, suggesting the concentration range used in the titration was more appropriate for promotion of complex formation than the lower concentration ranges used in the previous fluorescence titrations. This ensured that the data fitting and dissociation constant calculations were more likely to be accurate to the true value.⁴¹⁴ The average K_d was calculated to be $15.6 \mu\text{M} \pm 2.3 \mu\text{M}$, a weak binding affinity compared to the low nanomolar dissociation constants observed for CeuE binding to iron(III)-n-LICAM ($n = 4, 5, 6, 8$) siderophore mimics (Chapter 2). The binding affinity of CeuE to iron(III)-Sal-4-LICAM was around 750 to 800 times weaker than to iron(III)-4-LICAM.

Table 13: Dissociation constants for the binding of CeuE to iron(III)-Sal-4-LICAM calculated using Dynafit from circular dichroism titrations.

CeuE-iron(III)-Sal-4-LICAM	Kd / μ M	Error / μ M
Run 1	14.5	1.0
Run 2	18.9	2.3
Run 3	18.8	2.7
Average	15.6	2.3

CeuE-iron(III)-Sal-5-LICAM

The dissociation constant for CeuE binding to iron(III)-Sal-5-LICAM was established by both fluorescence and circular dichroism titration techniques. First a standard fluorescence titration was used to estimate whether the binding was of high or low affinity. It was observed that CeuE binding to iron(III)-Sal-5-LICAM was likely not in the low nanomolar range seen for iron(III)-n-LICAM (n= 4, 5, 6, 8) siderophore mimics. This was assessed by obtaining a shallow binding curve without a defined end-point: suggesting weak binding (Figure 100).

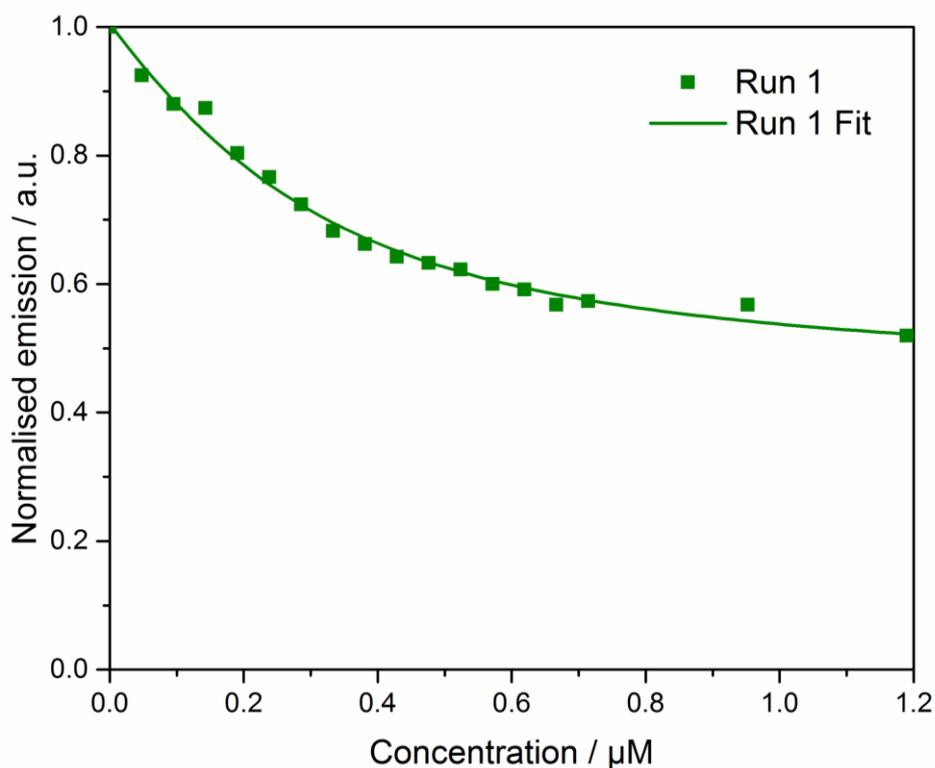


Figure 100: Binding curve for 0.24 μ M CeuE in 40 mM Tris-HCl pH 7.5, 150 mM NaCl. titrated with 12 μ M iron(III)-Sal-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

As binding was weaker than the iron(III)-n-LICAM (n= 4, 5, 6, 8) series in Chapter 2, circular dichroism titrations were carried out in triplicate as for CeuE with iron(III)-Sal-4-LICAM. The signal increased in magnitude with the addition of CeuE, as more iron(III)-Sal-5-LICAM was converted to Λ -configured complex upon CeuE binding (Figure 101).

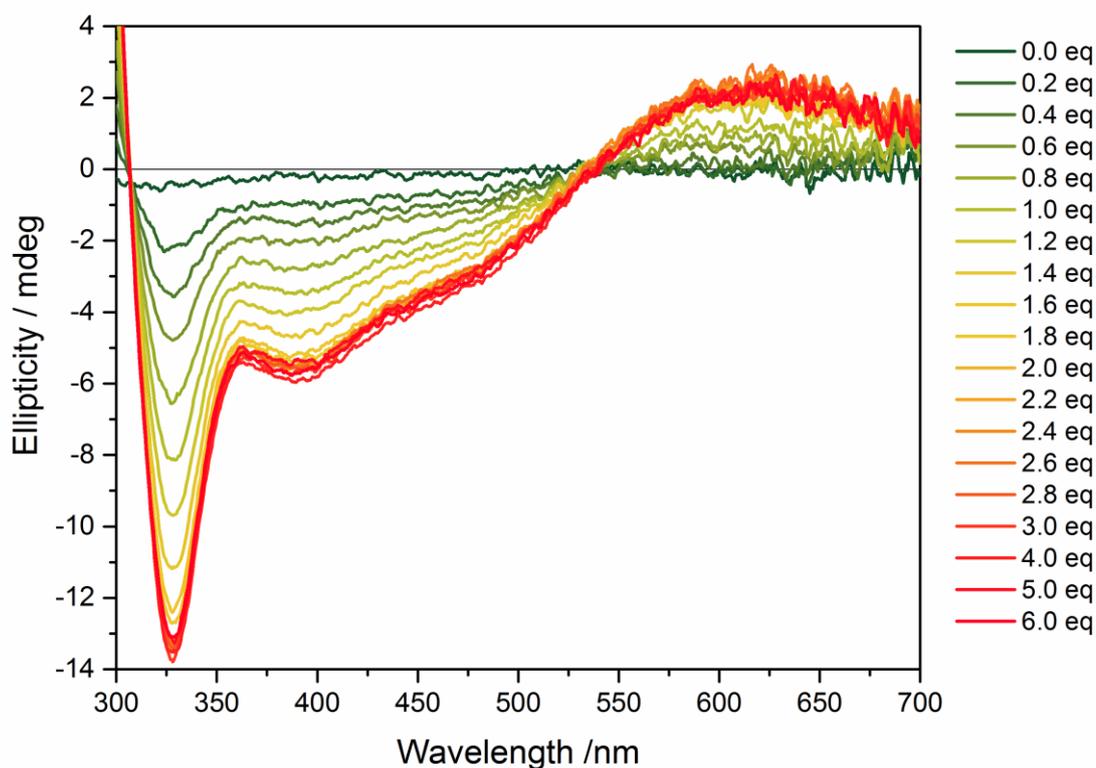


Figure 101: Circular dichroism spectra for the titration of 50 μ M iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl with aliquots of CeuE.

Ellipticity at 330 nm was recorded for each addition of CeuE, and binding curves constructed (Figure 102). The binding curve was steep, without many data points around the curved region, suggesting that binding was in fact tighter than expected. This caused problems upon fitting of the data, resulting in large errors on the K_d values (Table 14).

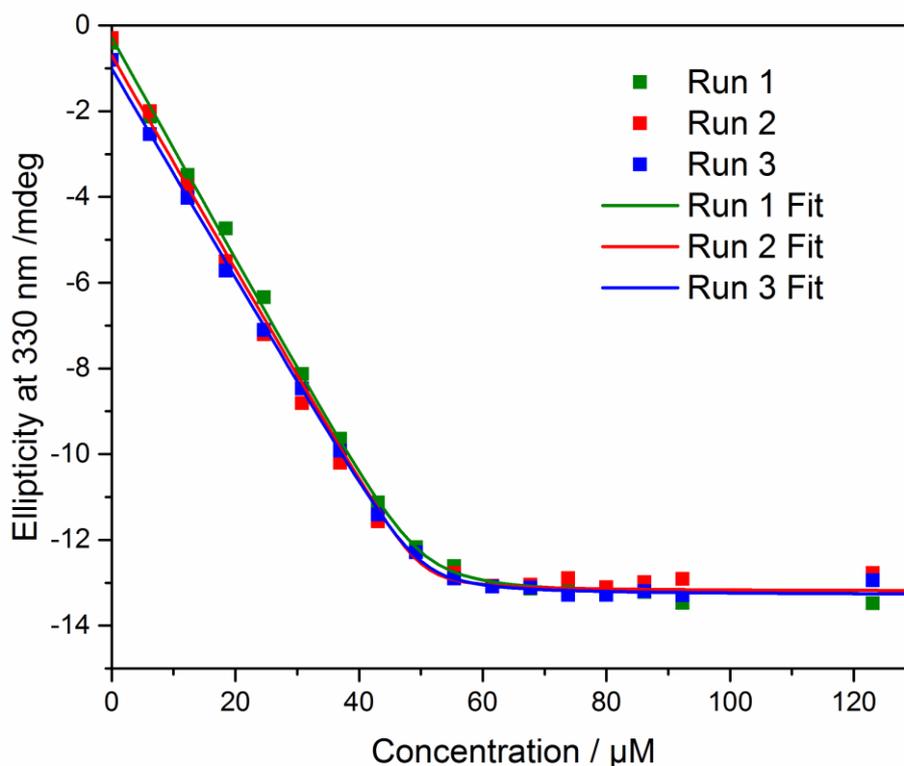


Figure 102: Binding curves for three independent circular dichroism titrations of 50 μM iron(III)-Sal-5-LICAM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl titrated with CeuE. Ellipticity at 330 nm was plotted against CeuE concentration. non-linear regression fitting data was generated with Dynafit.

Table 14: Dissociation constants for circular dichroism titration of iron(III)-Sal-5-LICAM with CeuE.

CeuE-iron(III)-Sal-5-LICAM	Kd /nM	Error /nM
Run 1	340	110
Run 2	140	130
Run 3	225	81
Average	240	127

A fluorescence quenching titration method at lower concentration than that used for the circular dichroism titration was used in an attempt to capture a good binding curve with a defined end-point as well as an acceptable number of data points around the curved region. The initial fluorescence quenching titration was run at a CeuE concentration of 0.24 μM ,

whereas the circular dichroism titration was run at a concentration of 50 μM . The second fluorescence quenching titration method was run at a more intermediate CeuE concentration of 2 μM , and gave a better set of binding curves (Figure 103).

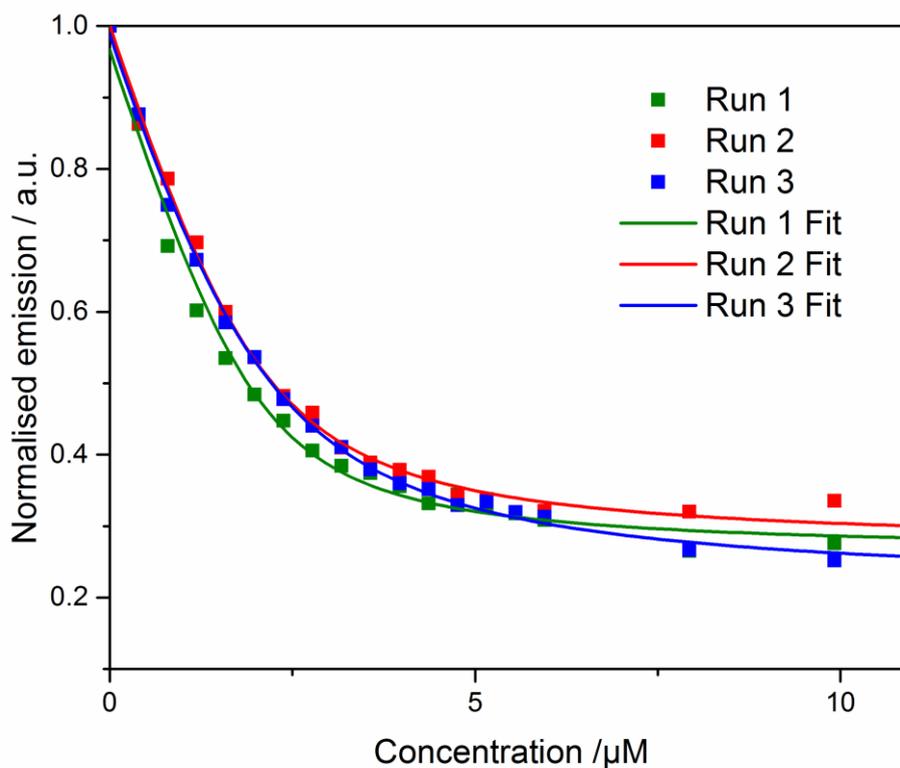


Figure 103: Binding curves for three independent fluorescence quenching titrations of 2 μM CeuE in 40 mM Tris-HCl pH 7.5, 150 mM NaCl with 100 μM iron(III)-Sal-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl. Non-linear regression fitting data was generated with Dynafit.

The dissociation constants calculated from the fitted binding curves are given in Table 15. The average dissociation constant calculated by this fluorescence method is likely to be the better estimate, due to the increased number of data points in the curved region of the binding curve, resulting in better fitting of the data in Dynafit, with lower corresponding error values.

Table 15: Dissociation constants for 2 μM CeuE titrated with iron(III)-Sal-5-LICAM.

CeuE-iron(III)-Sal-5-LICAM	K _d /nM	Error /nM
Run 1	536	59
Run 2	396	56
Run 3	563	42
Average	511	76

The binding constants for the iron(III) bound salmochelin mimics iron(III)-Sal-n-LICAM (n = 4,5) to CeuE were quite different to each other depending on linker length of the LICAM backbone. The CeuE-iron(III)-Sal-4-LICAM had a binding constant of $15.6 \mu\text{M} \pm 2.3 \mu\text{M}$ whereas that for CeuE-iron(III)-Sal-5-LICAM was much smaller, at $511 \text{ nM} \pm 76 \text{ nM}$, showing that iron(III)-Sal-5-LICAM bound around 30 times more strongly. Neither iron(III)-Sal-4-LICAM nor iron(III)-Sal-5-LICAM bound as tightly as the iron(III)-n-LICAM siderophore mimics did. Iron(III)-5-LICAM with a K_d value of $<10 \text{ nM}$ bound at least 50 times more strongly than iron(III)-Sal-5-LICAM. This suggests that the appended glucose unit hinders binding to CeuE, and may not be well accommodated in the secondary binding pocket as predicted by computational modelling in section 4.1. It appears that the shorter linker length of iron(III)-Sal-4-LICAM particularly affects the fit of the glucose into the secondary binding pocket. As iron(III)-5-LICAM is known to bind most tightly to CeuE out of the whole n-LICAM (n= 4, 5, 6, 8) series, it could be that this linker length, closer to that of the natural siderophore, salmochelin S1 is predisposed for a more favourable fit of the glucosyl unit into the secondary binding pocket.

Assuming Sal-5-LICAM is a good mimic for salmochelin S1, these findings suggest that while CeuE may be able to utilise this salmochelin S1 for iron uptake, it is unlikely that it is optimised for the uptake of this stealth siderophore. The affinity of non-functionalised iron(III)-n-LICAM siderophore mimics, and iron(III)-bisDHBS is much greater for CeuE than the salmochelin siderophore mimics, so it is likely that bisDHBS is the major siderophore used in CeuE iron(III) uptake. The fact that co-crystallisations of iron(III)-Sal-n-LICAM (n= 4, 5) with CeuE were unsuccessful is likely due to the poor binding affinities when compared to the affinities of previously studied tetradentate analogues.^{90, 310}

4.6 Summary and Conclusion

Salmochelin mimics Sal-4-LICAM and Sal-5-LICAM were successfully synthesised in overall yields of 5% and 6% respectively over 10 steps for each compound.

The iron(III) binding ratios of each mimic were measured using a Job Plot. After 1 hour, there was an equilibrium of 1:1 and 3:2 ligand: iron(III) complexes in solution, and after 7 days the 3:2 complex predominated for both linker lengths. These are similar to the iron(III) binding stoichiometries of 1:1 and 3:2 found for 4-LICAM and bisDHBS.^{90, 322}

When mixed with CeuE, each iron(III)-Sal-n-LICAM (n = 4, 5) complex was arranged in a majority Λ -configuration- proving that CeuE is able to bind both iron(III)-bound salmochelin S1 mimic compounds. This indicates that CeuE may be able to bind to salmochelin S1. Dissociation constants determined by both fluorescence quenching titration and circular dichroism titration showed that iron(III)-Sal-4-LICAM is bound around 30 times more weakly than iron(III)-Sal-5-LICAM, while the latter is bound at least 50 times more weakly by CeuE than the non-C5-functionalised iron(III)-5-LICAM siderophore mimic, and the natural substrate iron(III)-bisDHBS. This probably indicates that although CeuE is able to bind the salmochelin mimic compounds, the periplasmic binding protein is better suited to the binding of the enterobactin hydrolysis product bisDHBS than salmochelin S1.

4.7 Future Work

Further efforts are needed to obtain crystal structures of the iron(III) bound salmochelin mimics in complex with CeuE. These would give information on the binding arrangement of the salmochelin mimics in the CeuE binding pocket, and provide insight into the potential role of the secondary binding pocket region of CeuE as predicted by computational modelling. Studying potential interactions of the salmochelin mimics with new regions of the binding cleft would help to further knowledge of CeuE function in *C. jejuni* iron-uptake.

Building on the work in Chapter 2, longer linker length salmochelin mimics such as Sal-n-LICAM (n = 6, 8) could be synthesised. The longer salmochelin compounds could then be tested for their iron(III) binding, and subsequent CeuE binding affinities. It would be interesting to see whether linker length influences the ability of the salmochelin mimic to position the glucosyl unit in the CeuE binding cleft. To prove whether the salmochelin mimic compounds iron(III)-Sal-n-LICAM (n = 4, 5) were indeed good mimics of natural Salmochelin S1, it would be necessary to repeat all studies with the natural siderophore that contains a serine backbone. To acquire salmochelin S1 it would be necessary to find a new supplier of the natural siderophore or to synthesise the compound. It may be possible to combine the synthetic knowledge gained in the synthesis detailed in section 4.2 with the syntheses of bisDHBS and the salmochelins documented in the literature.^{90, 396, 397}

Salmochelin S4 mimics could be synthesised based on vibriobactin, or the mimic MECAM. The proposed synthesis of salmochelin S1 could be extended to the synthesis of salmochelin S4. It would be interesting to study whether CeuE may bind hexadentate salmochelins, and compare the affinity with that of CeuE to iron(III)-enterobactin.

With the identification of a family of tetradentate periplasmic binding proteins described in Chapter 3, it would be interesting to study other similar periplasmic binding proteins to establish whether any periplasmic binding proteins are able to bind the salmochelin mimics with higher affinity than CeuE. If periplasmic binding proteins were found that bound the salmochelin mimics with biologically relevant binding constants,¹⁸⁵ then this would provide insight into the iron uptake capacities of range of species that are able to take advantage of tetradentate salmochelin siderophores. As iron uptake often has relevance in bacterial virulence, this knowledge may be of medical relevance for understanding the host colonisation of pathogenic species. This may provide information for the design of antibiotics and vaccines to specifically target such periplasmic binding proteins. Chapter 5 addresses this area of study, and involves the investigation of the interactions of periplasmic binding proteins FepB from *E. coli* and VctP from *V. cholerae* with the salmochelin mimics.

5 Comparison of Three Enterobactin-Uptake Periplasmic Binding Proteins

5.1 Introduction

This chapter compares the type III periplasmic binding proteins involved in iron(III)-enterobactin uptake from three gram-negative bacterial species that are known to cause gastric illness in the mammalian gut: CeuE from *C. jejuni*, FepB from *Escherichia coli*, and VctP from *V. cholerae*. As described in Chapter 1, all three species are known to have strains that cause diarrheal disease by production of enterotoxins.^{134, 135, 166, 167, 189, 416} Iron uptake is commonly known to be related to bacterial virulence,^{1, 30, 132, 193} so the strategies employed by the above pathogenic species may give an insight into their prevalence as infectious agents. As both pathogenic strains of *E. coli* and *V. cholerae* have been noted as species at risk of widespread antibiotic resistance, the resulting knowledge may prove useful for development of new design strategies for novel antimicrobials.^{134, 135, 417} In addition, the study aims to increase knowledge of PBP-iron(III)-siderophore combinations as potential candidates for the development of novel artificial metalloenzymes, with tight-binding siderophore complexes as anchors for attachment of inorganic catalysts.^{336, 356}

CeuE is named from a family of Ceu (*Campylobacter* Enterobactin Uptake) proteins, as it is generally accepted that these are involved in the direct uptake of iron(III)-enterobactin into the cytoplasm (Figure 104).^{115, 168, 172, 418}

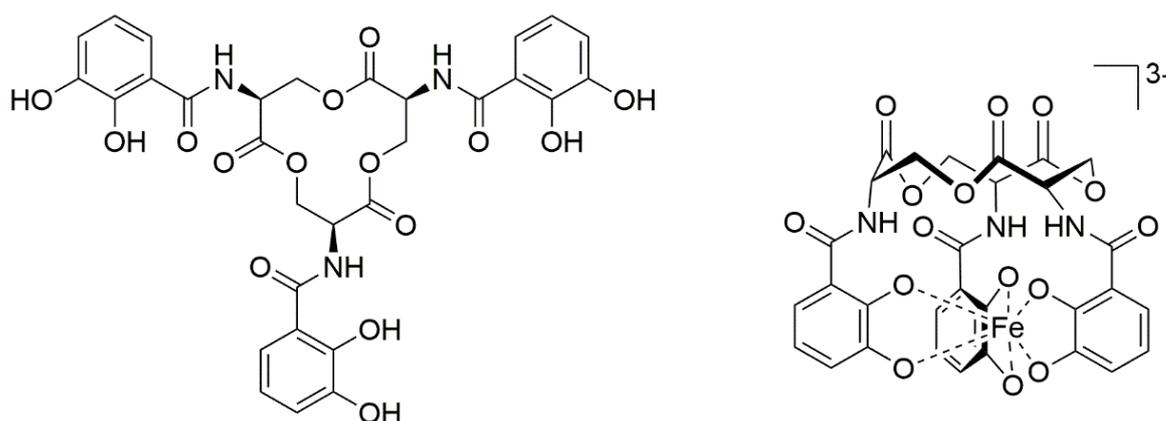


Figure 104: Left: Enterobactin Right: iron(III)-enterobactin displayed as a 3-dimensional representation of the octahedral coordination.

As it has been commonly reported that iron(III)-enterobactin is transported by the Ceu proteins into the cytoplasm before iron(III) release,^{159, 168} the CeuE- iron(III)-enterobactin binding interaction was of interest to study. The circular dichroism spectra for CeuE- iron

(III)-enterobactin were previously reported.⁹⁰ iron(III)-enterobactin had predominant Δ -configuration in aqueous solution, but the equilibrium between Δ - and Λ -configured complexes was shifted to predominant Λ -configuration in the presence of CeuE. This indicated that there must be a CeuE- iron(III)-enterobactin interaction, and the K_d was estimated to be around 400 nM.⁹⁰ Based on the high binding affinity of iron(III)-bisDHBS with CeuE, and the iron(III)-n-LICAM series discussed in Chapter 2, CeuE appears rather to be optimised for binding of the tetradentate enterobactin hydrolysis product instead of iron(III)-enterobactin.^{90, 310} This theory fits with the lack of an esterase capable of hydrolysing enterobactin, or a suitable reductase in the cytoplasm of *C. jejuni*.¹⁵⁹

FepB, the periplasmic binding protein in the FepABCDG (Ferric Enterobactin-binding Proteins) family of iron(III)-enterobactin uptake proteins in *E. coli* has been extensively studied and is known to bind iron(III)-enterobactin with high affinity.¹⁴³⁻¹⁴⁶ Fluorescence quenching titrations gave K_d values of between 30 and 170 nM.^{144, 145} Whole cell binding assays using radioactive ⁵⁵Fe(III) proved that FepB is required for delivery of iron(III)-enterobactin to the inner membrane,¹⁴³ and proteinase K digest assays showed that FepB is stabilised in the presence of iron(III)-enterobactin.¹⁴⁵ The structure of FepB was first solved by NMR, and chemical shift perturbations were observed with the addition of gallium(III)-enterobactin.¹⁴⁵ The crystal structure of FepB in complex with iron(III)-enterobactin was subsequently solved.¹⁴⁶ While the interpretation of the results in that study placed a large emphasis on the trimeric symmetry of FepB in the crystal structure,¹⁴⁶ the trimer does not seem likely to be of biological relevance as there were no direct protein-protein contacts that would indicate that it is able to hold together at biological concentrations. A 1:1 relationship is observed in NMR and binding affinity experiments, and it is widely documented that a 1:1 complex is required for delivery of iron(III)-siderophore complexes to the uptake machinery located in the cytoplasmic membrane,^{20, 145, 305, 419-422} Inspection of the binding pocket of FepB from the co-crystal structure (PDB ID: 3TLK) (Figure 105) reveals a complex array of binding interactions between amino acid residues in the FepB binding cleft and the iron(III)-enterobactin scaffold (Figure 106).¹⁴⁶ These include hydrogen bonding to catechol oxygen atoms from Gly 126 (amide nitrogen) Arg 242 and Arg 301, hydrogen bonding to the backbone carbonyl oxygens from Arg78 and Thr200, and hydrogen bonding to the amide carbonyls from Thr 73 and Asn 77 (Figure 106). In addition, Trp 209 and Phe 300 at the base of the binding cleft are likely to provide favourable hydrophobic interactions with the aromatic rings (Figure 106). The arginine residues, Arg 75, Arg 239 and Arg 298 that act as a basic triad for iron(III)-enterobactin binding, indicate that the FepB binding cleft

shares some similarities to that of CeuE, which contains a similar basic triad as discussed in Chapter 2.^{146, 186, 310}

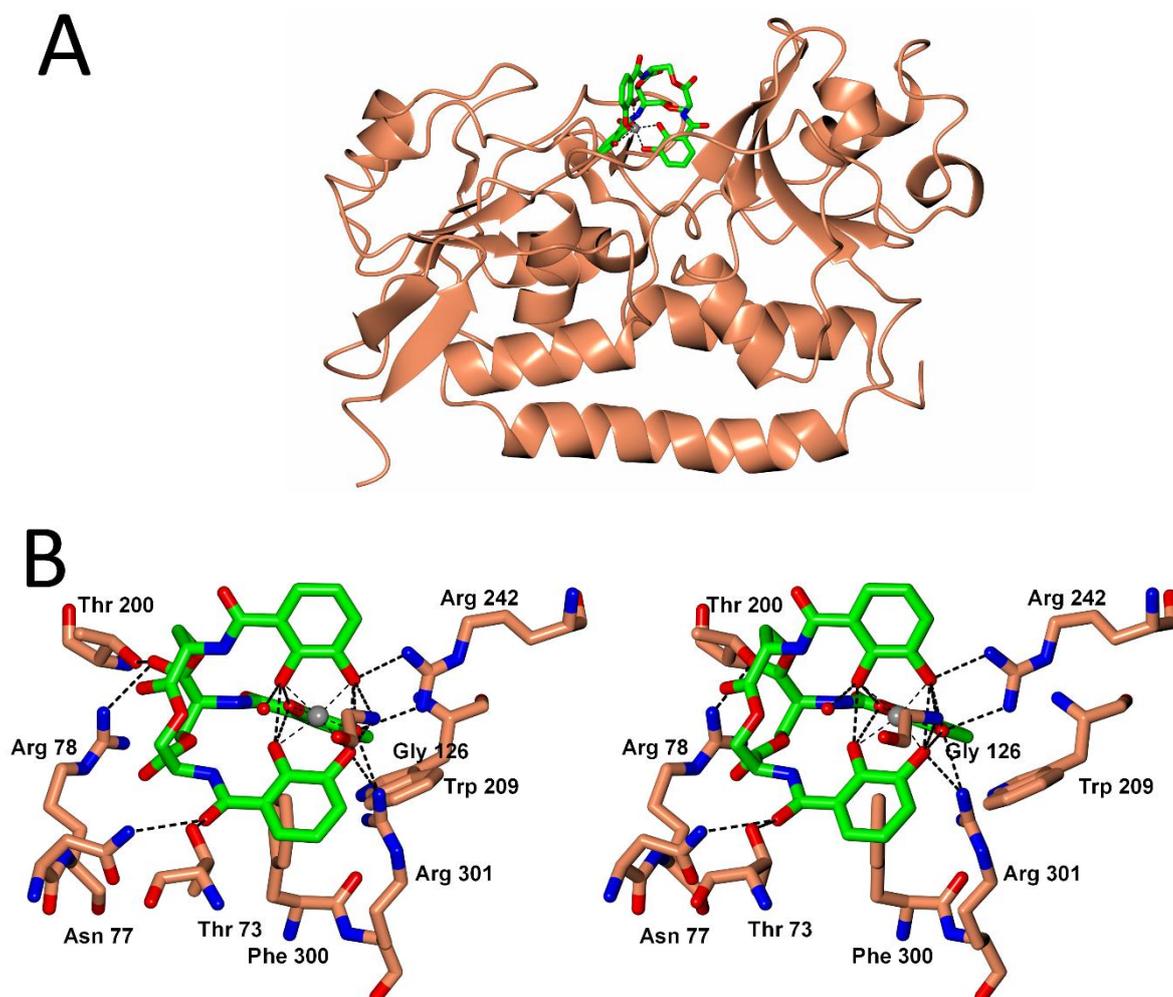


Figure 105: A: Co-crystal structure of FepB (coral ribbons) with iron(III)-enterobactin in the binding cleft (cylinders coloured by atom type). B: stereo view of binding cleft of FepB (coral cylinders coloured by atom type, carbon in coral), showing all amino acids involved in binding of the enterobactin scaffold (cylinders coloured by atom type) (structure from PDB ID: 3TLK).¹⁴⁶ Figure produced using CCP4mg.

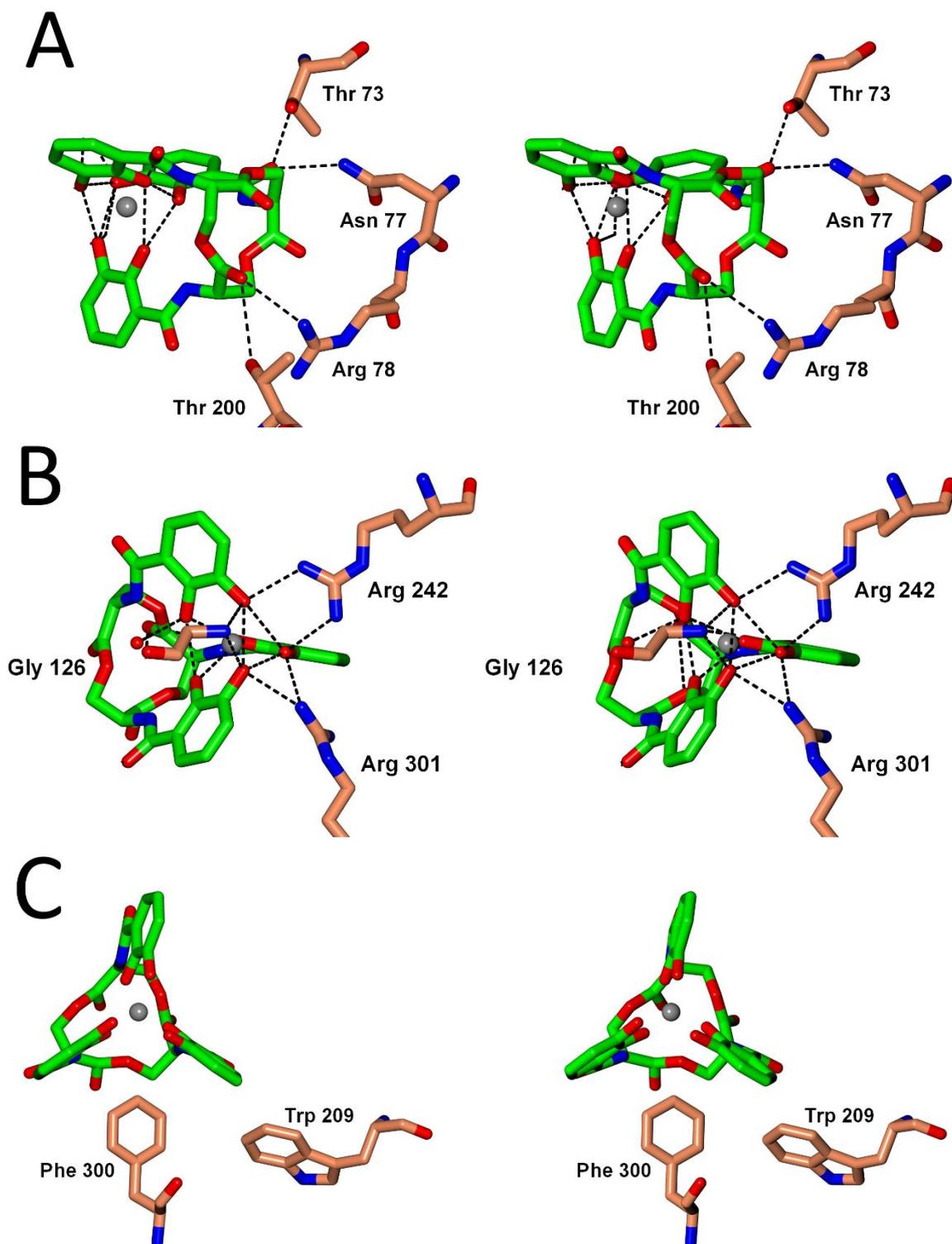


Figure 106: Stereo view of the binding Cleft of FepB (cylinders coloured by atom type, with carbons in coral) with iron(III)-enterobactin bound (cylinders coloured by atom type). Top: Thr 73, Asn 77, Arg 78 and Thr 200 provide hydrogen bonds with the backbone and amide carbonyl oxygens. Middle: Gly 126 (backbone amide nitrogen) Arg 242 and Arg 301 provide hydrogen bonds to catechol oxygens. Bottom: Trp 209 and Phe 300 provide hydrophobic interactions (PDB ID: 3TLK).¹⁴⁶ Figure produced using CCP4mg.

E. coli uses the hexadentate catecholate siderophore mimic MECAM for effective delivery of iron(III) for growth.⁴²³ Growth was hindered in the absence of the *fepB* gene product, indicating that FepB is required for MECAM uptake, and an *E. coli* reductase was shown to reduce iron(III) to iron(II) from iron(III)-MECAM.^{165, 423} Although many previous studies focussed on the iron(III)-enterobactin binding capacity of FepB, the ability of FepB to bind tetradentate catecholate siderophores such as iron(III)-bisDHBS remains largely unexplored. It was therefore of interest to investigate whether FepB is capable of binding both hexadentate (including iron(III)-MECAM) and tetradentate iron(III)-catecholate complexes.

VctP is the periplasmic binding protein from the VctACDGP (*Vibrio cholerae* Transport) family of iron(III)-uptake proteins in *V. cholerae*.^{200, 201} VctP, contains the conserved histidine and tyrosine residues known to be involved in iron(III)-siderophore binding in CeuE,¹⁸⁶ and is known to interact with the ABC transporter VctDGC in the inner membrane of *V. cholerae*, reported for the transport of both iron(III)-enterobactin and iron(III)-vibriobactin as discussed in Chapter 1.¹⁹²

The use of enterobactin hydrolysis products rather than the full cyclic trilactone for iron(III) uptake by *V. cholerae* is a topic of debate, with mixed results from growth studies of the species with enterobactin.²⁰² It has been suggested that there may not be an enterobactin outer membrane receptor, due to the inability of the species to accumulate iron(III)-enterobactin.²⁰⁶ In the studies where enterobactin uptake was reported, crude *E. coli* extract was used as an enterobactin source without purification, and the linear hydrolysis products were not discussed, suggesting possible enterobactin hydrolysis was not taken into account.^{192, 201} More recently, pure enterobactin was proven to be a poor iron(III) delivery source for growth of *V. cholerae* due to a lack of an outer-membrane receptor, as well as the lack of an esterase to degrade the trilactone backbone for iron(III) release once inside the cell.¹⁶¹ In contrast, iron(III)-bisDHBS and iron(III)-trisDHBS supported growth when taken up by both outer-membrane receptors VctA and IrgA- those previously reported for enterobactin uptake.^{161, 192} The structure of apo-VctP (PDB ID: 3TEF) was determined by protein crystallography, with high structural homology to CeuE as discussed in Chapter 3.²⁰⁹ While there are no structures of VctP complexed with an iron(III)-bound ligand in the PDB, computational modelling of iron(III)-enterobactin into the binding cleft region of VctP has been attempted, using the coordinates from the apo crystal structure.^{161, 209} So far, no dissociation constants have been estimated for the binding of any iron(III)-ligands to VctP, so it was of interest to establish whether VctP is able to bind hexadentate and/or tetradentate iron(III)-catecholate siderophores as suggested in previous uptake studies, and measure

binding affinities of each iron(III)-siderophore for the protein. These studies aim to clear up the current confusion over the role of VctP in the use of iron(III)-enterobactin as an iron source for *V. cholerae*.^{161, 200, 201, 424}

Comparing the amino acid sequences of CeuE, FepB and VctP, it is obvious that the sequences are quite different (Figure 107). FepB only has 18.6% identity to CeuE, with 35.1% similarity, and 23.1% identity to VctP with 35.4% similarity, as calculated using EMBOSS Needle.⁴²⁵ CeuE and VctP have slightly higher similarity to each other, and share 25.3% identical and 47.6% similar residues.⁴²⁵

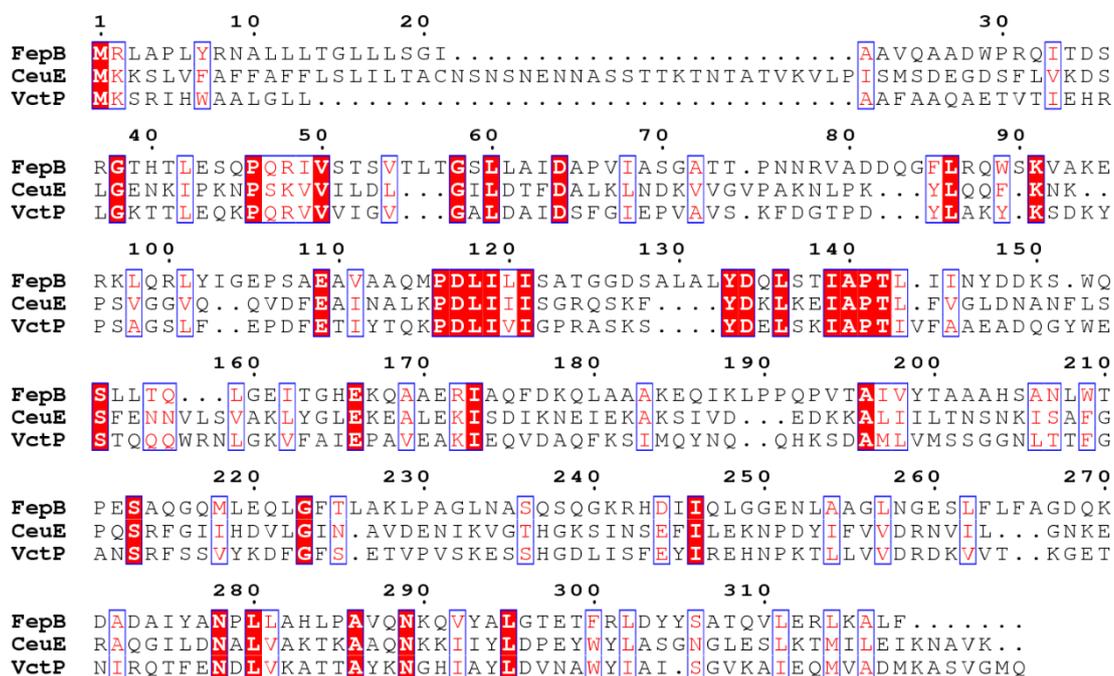


Figure 107: Sequence alignment of FepB, CeuE and VctP using MUSCLE.⁴²⁶ Image produced with ESPript.⁴²⁷ Red blocks are totally conserved, red text indicates close matches or residues of similar properties.

The overlaid crystal structures of the three proteins reveal high structural similarity (Figure 108). CeuE and VctP are most similar, with r.m.s.d of 1.78 Å over 263 C α positions and a 100% match for secondary structure elements.³¹⁷ Both contain conserved histidine and tyrosine residues in the binding cleft, as well as two conserved arginines, as discussed in Chapter 3. FepB has r.m.s.d of 3.29 Å over 238 C α positions with CeuE with 80% match for secondary structure elements, and r.m.s.d of 3.27 Å over 226 C α positions with VctP with 79% match for secondary structure elements. FepB does not contain the conserved histidine or tyrosine in the binding cleft, and instead contains a flexible loop where the tyrosine would be, and an extended loop where the histidine would be (Figure 108). These differences

suggested that FepB may have different preferences in siderophore binding to CeuE and VctP, whereas the latter were expected to share a more similar siderophore binding profile.

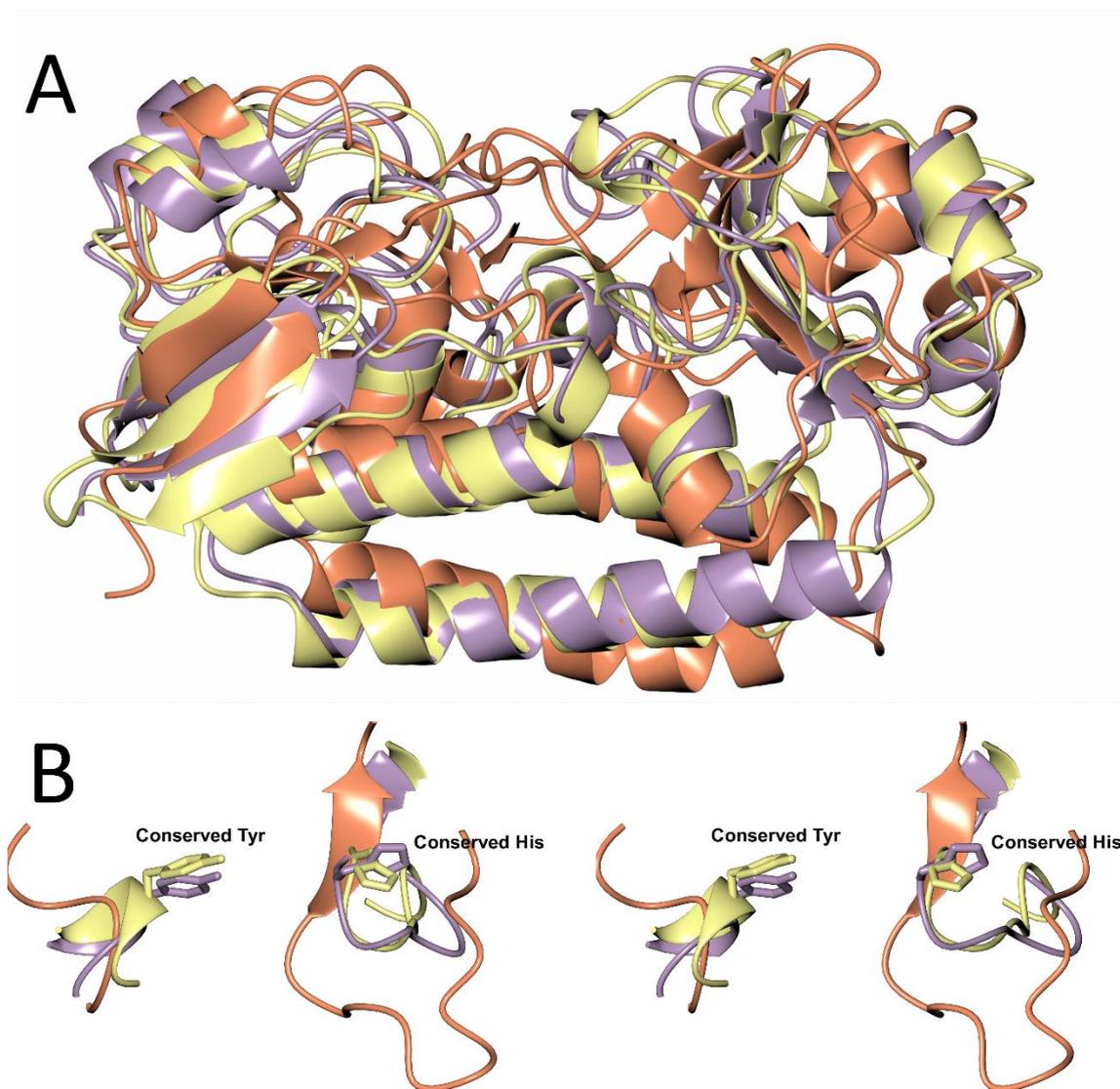


Figure 108: A: Overlay of Chain A of the structures of CeuE in lemon ribbons (PDB ID: 3ZKW), FepB in coral ribbons (PDB ID: 3TLK) and VctP in lilac ribbons (PDB ID: 3TEF) via the SSM algorithm. B: Stereo view of the binding cleft region of CeuE and VctP: the conserved His and Tyr in CeuE and VctP are absent in FepB. Figure produced using CCP4mg.

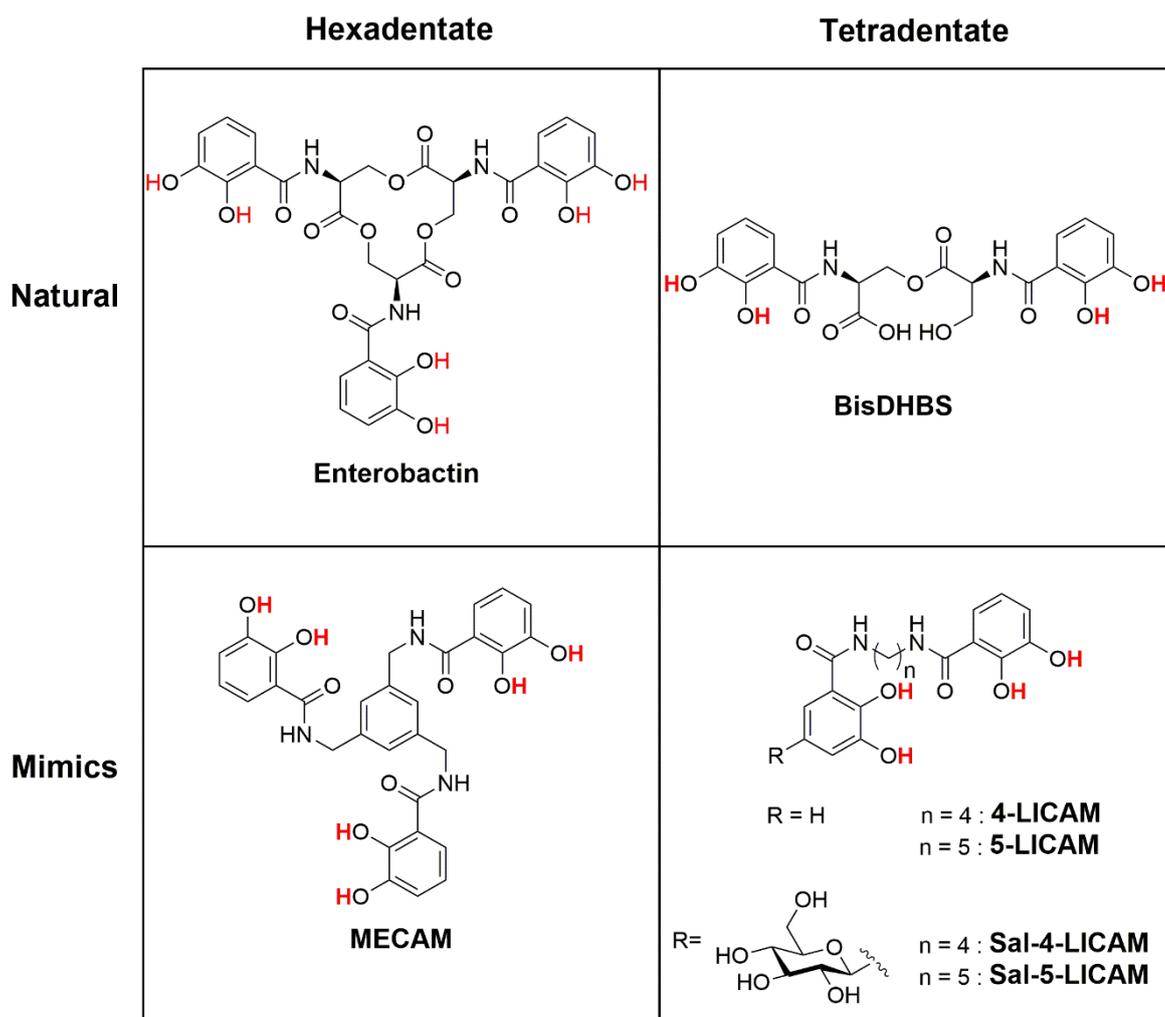


Figure 109: Hexadentate and tetradentate siderophores and mimic siderophores used to probe periplasmic binding protein accommodating ability. Protons shown in red are deprotonated upon iron(III) binding.

To examine the binding abilities of each protein, a range of hexadentate and tetradentate siderophores and mimics were selected (Figure 109), and fluorescence quenching titration were used to quantify the binding affinities. As already described in Chapter 2, this was possible for CeuE due to the intrinsic fluorescence of Trp 276 adjacent to the iron(III)-siderophore binding pocket region, that was quenched upon addition of iron(III)-siderophore. This Trp residue is conserved in VctP, as seen in position 300 in the sequence alignment in Figure 107. This Trp provides a similar fluorescence quenching event upon addition of iron(III)-siderophore to VctP. FepB also contains a Trp adjacent to the Fe(III)-siderophore binding pocket as shown in Figure 106C, providing fluorescence quenching upon addition of iron(III)-siderophore to FepB. As all three periplasmic binding proteins contained tryptophan adjacent to the iron(III)-siderophore region, all were compatible with the standard fluorescence titration method detailed in Chapter 6.

Circular dichroism was then used to assess binding modes, and the results were analysed for similarities and differences between the proteins based on binding cleft features, overall iron-uptake strategies and biological roles of the native organism from which each protein was derived. Crystal structure determination was attempted *via* co-crystallisation of proteins and ligands.

5.2 Hexadentate Iron(III)-Siderophore Binding of CeuE

A commercial sample was confirmed to be enterobactin by ESI-MS (Chapter 6), and HPLC (Figure 110). The maximum concentration of iron(III)-enterobactin that could be used in fluorescence quenching titrations was determined as 18.40 μM , as detailed in Chapter 6.

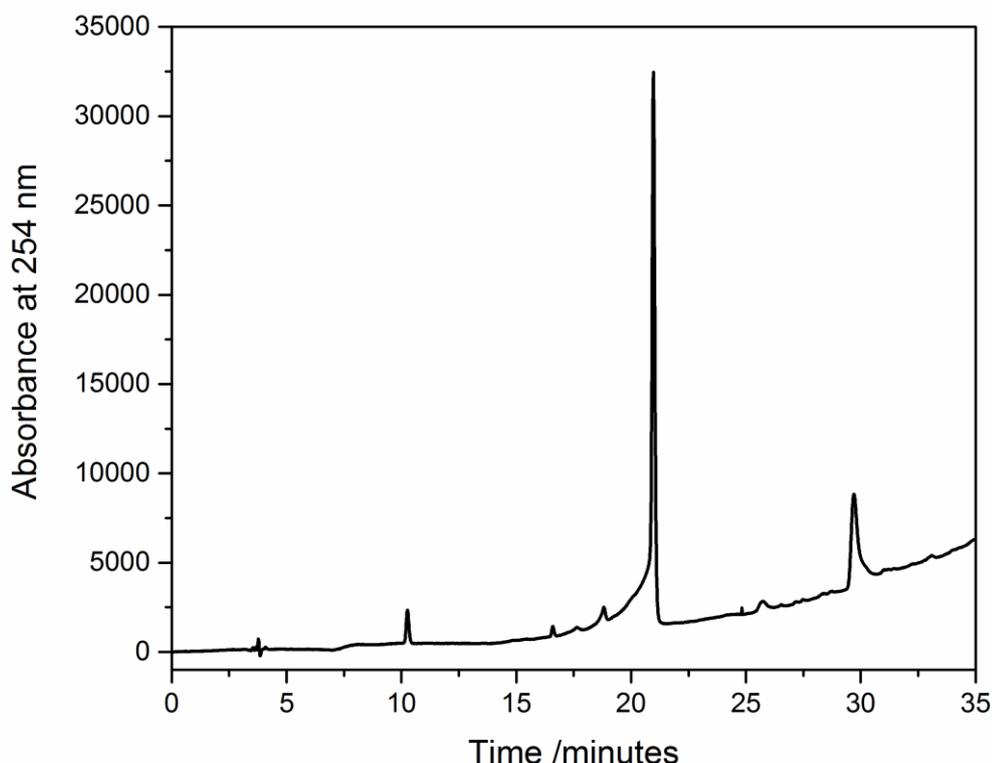


Figure 110: HPLC trace of purchased enterobactin. The sample was analysed with a gradient of 10-80% acetonitrile in water over 35 minutes with 0.1% TFA throughout and a flow rate of 1 mL/min. The major peak at a retention time of 21 minutes correlates well with literature retention times for enterobactin.^{86, 87}

To test the hypothesis that enterobactin may not be the major substrate for CeuE, fluorescence quenching titrations were performed to quantify the binding affinity.⁹⁰ Circular dichroism proved that there must be a CeuE iron(III)-enterobactin interaction,⁹⁰ while the fluorescence quenching titrations in the literature gave dissociation constants of limited accuracy as a true end-point was not achieved.⁹⁰ The studies were therefore optimised and repeated to estimate a more accurate K_d . The first titration at a protein concentration of 240 nM gave a shallow binding curve that did not reach a clear endpoint with five equivalents of iron(III)-enterobactin (Figure 111). The K_d for iron(III)-enterobactin binding to CeuE was

much larger than for the tetradentate analogues iron(III)-n-LICAM ($n = 4, 5, 6, 8$) as discussed in Chapter 2, or iron(III)bisDHBS.⁹⁰

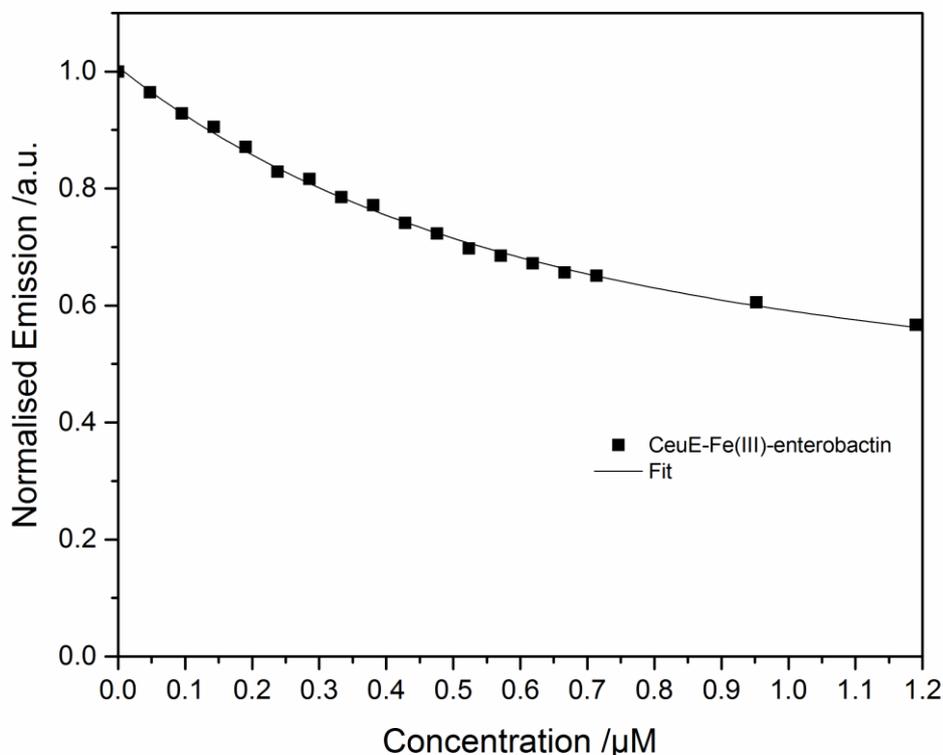


Figure 111: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-enterobactin and their associated non-linear regression from Dynafit.³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-enterobactin⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Higher concentrations of protein and ligand were used in an attempt to drive the equilibrium towards protein-ligand complex formation.⁴¹⁴ 2.4 μM CeuE was titrated in triplicate with a total of seven equivalents of iron(III)-enterobactin and with a final ligand concentration of 16.80 μM , a more pronounced binding curve was achieved (Figure 112), giving a K_d of 3.5 $\mu\text{M} \pm 0.3 \mu\text{M}$ (Table 16). This high K_d , around 10 times larger than the previous estimate of $0.4 \pm 0.1 \mu\text{M}$,⁹⁰ showed that CeuE bound iron(III)-enterobactin around 1000 times more weakly than iron(III)-5-LICAM. The CeuE-iron(III)-enterobactin binding affinity was at least 20 times, and probably over 100 times weaker than the literature values between 30 and 170 nM for iron(III)-enterobactin binding to FepB.^{144, 145}

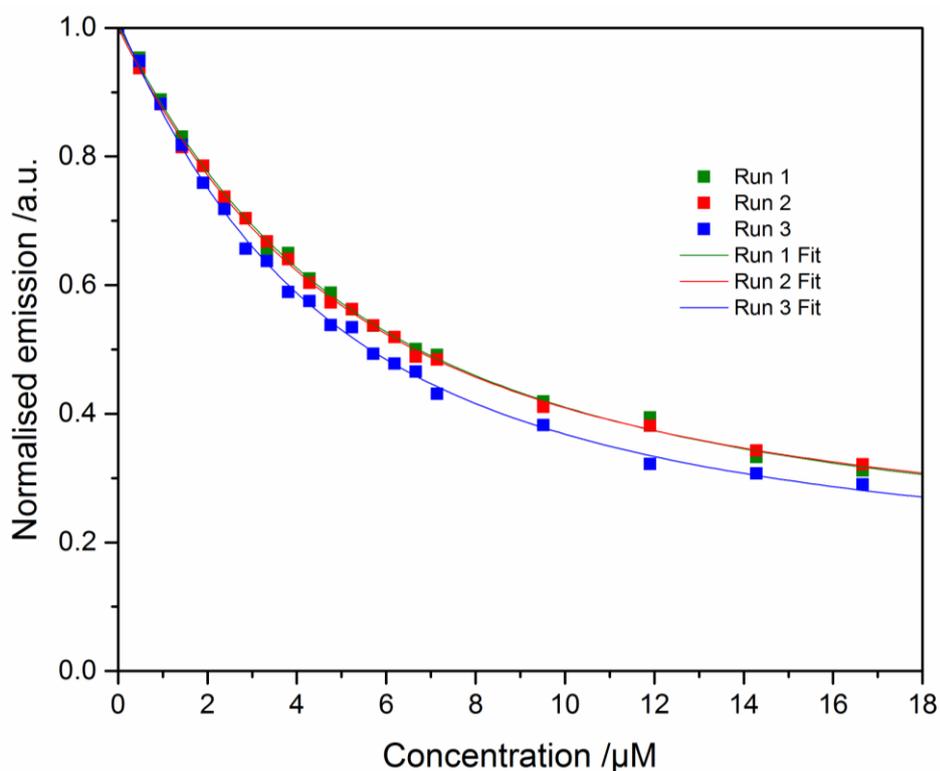


Figure 112: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission *vs* ligand concentration in μM for CeuE-iron(III)-enterobactin and the model from Dynafit.³²⁶ CeuE at a concentration of $2.4 \mu\text{M}$ in $40 \text{ mM Tris-HCl pH } 7.5$, 150 mM NaCl was titrated with aliquots of $120 \mu\text{M iron(III)-enterobactin}$ in $40 \text{ mM Tris-HCl pH } 7.5$, 150 mM NaCl .

Table 16: K_d values calculated for CeuE-iron(III)-enterobactin *via* fluorescence quenching titration and associated errors, giving a weighted average as detailed in Chapter 6.

CeuE-iron(III)-Enterobactin	$K_d / \mu\text{M}$	Error $/\mu\text{M}$
Run 1	3.79	0.19
Run 2	3.66	0.12
Run 3	3.09	0.15
Average	3.5	0.3

These findings confirm that CeuE is in fact not a dedicated enterobactin binding protein, as although it is able to bind this hexadentate siderophore weakly, it is far better adapted to bind the linear hydrolysis product iron(III)-bisDHBS.⁹⁰ No crystal structure of CeuE-iron(III)-enterobactin could be obtained *via* either co-crystallisation or soaking of apo CeuE crystals with an iron(III)-enterobactin solution which yielded only data for the native apo protein. This is probably because the binding affinity was too weak for adequate complex formation.

As previously reported, CeuE is able to bind to the enterobactin mimic iron(III)-MECAM in a 2:2:2 complex as observed *via* crystal structure, and observation of a Λ -configured complex in solution.¹⁸⁷ To complete this study, the binding affinity of CeuE for iron(III)-MECAM was measured. Fluorescence quenching titrations gave binding curves that yielded a K_d of 30 ± 11 nM (Figure 113) (Table 17), more in-keeping with the binding constants observed for iron(III)-n-LICAM (n= 4, 5, 6, 8) and iron(III)-bisDHBS than that of iron(III)-enterobactin.^{90, 186}

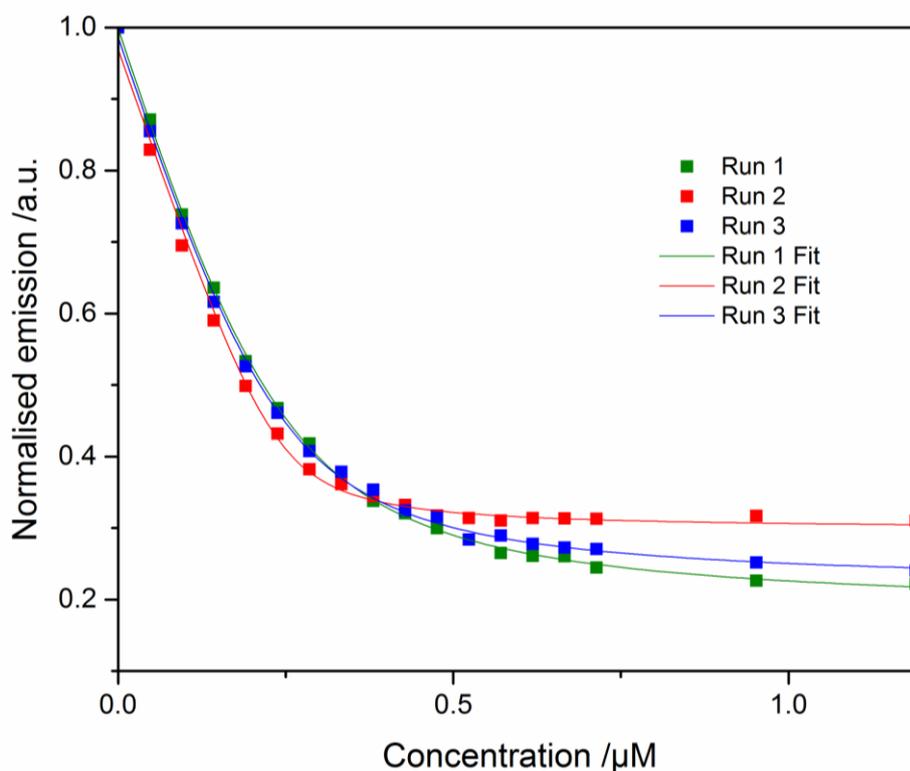


Figure 113: Fluorescence quenching titration plots of normalised fluorescence emission vs ligand concentration in μM for CeuE-iron(III)-MECAM and the model from Dynafit.³²⁶ CeuE at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-MECAM⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Table 17: K_d values for CeuE-iron(III)-MECAM from fluorescence quenching titration, with a weighted average calculated as detailed in Chapter 6.

CeuE-iron(III)-MECAM	K_d /nM	Error /nM
Run 1	44.7	1.8
Run 2	10.1	2.0
Run 3	34.7	2.5
Average	30	11

5.3 FepB Expression and Purification

FepB was produced and purified with the aid of Dr. E. V. Blagova (Chapter 6). The protein sequence for FepB was obtained for *E. coli* K strain MC1061 from Uniprot (code C6EKA3).⁴²⁸ Minimal disorder predictions were used to design a gene construct of 873 base pairs, spanning amino acids 27-318 of the full protein sequence (Figure 11).⁴²⁹ A similar construct of 27-320 was used for FepB in the published crystal structure (PDB ID: 3TLK).¹⁴⁶

```

FepB  1      10      20      30      40      50      60
      MRLAPLYRNALLLTGLLLSGIAAVQAADWPRQITDSRGTHLTLESQPQRIVSTSVTLTGSLS
FepB      70      80      90      100      110      120
      LAIDAPVIASGATTPNNRVADDQGFLRQWSKVAKERKLRQLYIGEPSAEAVAAQMPDLIL
FepB     130      140      150      160      170      180
      ISATGGDSALALYDQLSTIAPTLIINYDDKSWQSLLTQLGEITGHEKQAAERIAQFDKQL
FepB     190      200      210      220      230      240
      AAAKEQIKLPPQPVTIVYTAAAHSANLWTPESAQQQMLEQLGFTLAKLPAGLNASQSQG
FepB     250      260      270      280      290      300
      KRHDIIQLGGENLAAGLNGESLFLFAGDQKDADAIYANPLLAHLPAVQNKQVYALGTETF
FepB     310
      RLDYYSATQVLERLKALF

```

Figure 114: Protein sequence for FepB from *E. coli* K strain MC1061 from Uniprot (code C6EKA3, residues 1-26 were predicted to be disordered (shown in a black box) and were removed to generate the construct 27-318. Image produced with ESPript.⁴²⁷

Table 18: Primers used for amplification of the FepB gene.

Forward Primer	Reverse Primer
5'TCCAGGGACCAGCAATGGCT GACTGGCCGCGTC3'	5'TGAGGAGAAGGCGCGTTAAAACAGCG CCTTAAGTCGTTCCAGC3'

The desired protein sequence was converted to the necessary DNA sequence, and forward and reverse primers designed for amplification of the DNA *via* PCR (Table 18). The PCR product was inserted into the pET-YSBLIC3C vector, and the resulting plasmid was used to transform Competent Nova Blue cells for colony PCR. Plasmid DNA was harvested and external sequencing gave a 100% match to the target sequence. BL21 expression cells were then transformed with the plasmid DNA and cultured in 1% kanamycin LB media. The cells were induced with 1% IPTG upon reaching an OD₆₀₀ of 0.6 and split into two portions for incubation for 4 hours at 37°C or overnight at 16°C. SDS PAGE analysis showed good overexpression of the target protein in both conditions, with a large proportion of soluble protein (Figure 115). The cultures were scaled to four portions of 500 mL, were induced

with 1% IPTG upon reaching an OD_{600} of 0.6, and incubated overnight at 16°C. Overexpression of FepB was confirmed by SDS PAGE gel (Figure 116).

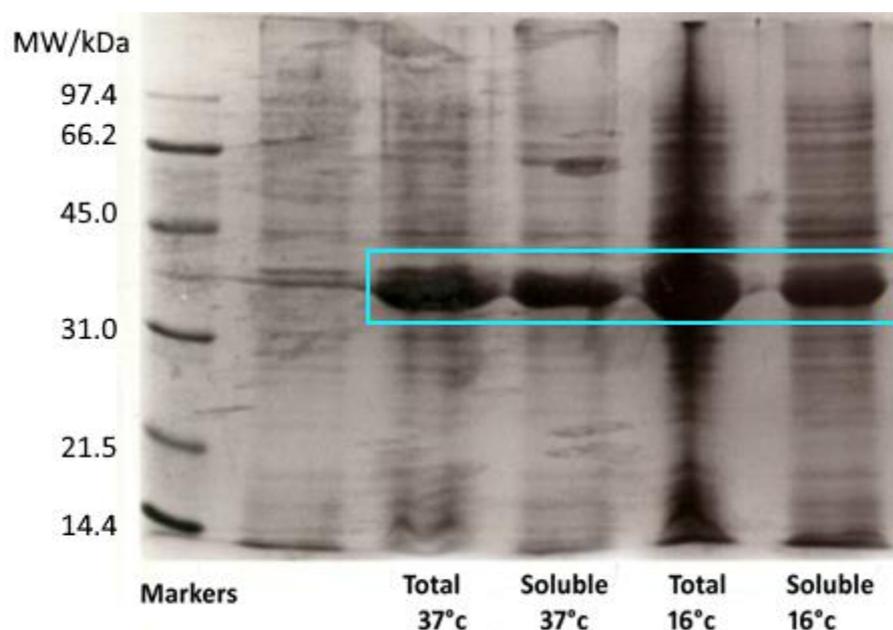


Figure 115: SDS PAGE of test overexpression of FepB in BL21 cells. The bands of overexpressed protein are shown in a light blue box. FepB was present in both total and soluble fractions for cells cultured at both 37 and 16 °C.

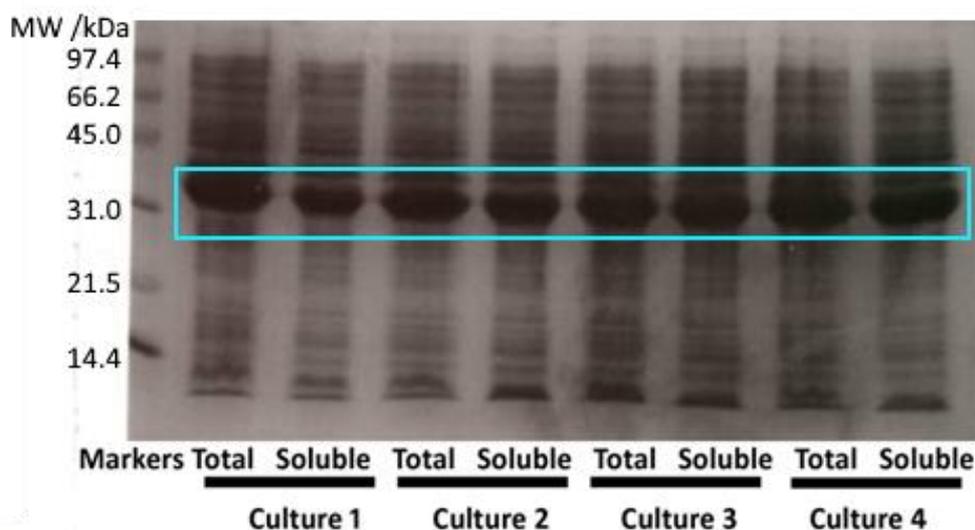


Figure 116: SDS PAGE of overexpression of FepB in 500 mL cultures of BL21 cells. The band of overexpressed protein is shown in a light blue box. FepB was present in both total and soluble fractions for all four cultures.

The cell pellet was collected *via* centrifugation of the cultures, and purified *via* nickel column chromatography. The protein was eluted over an imidazole gradient (Figure 117), and the fractions, confirmed to contain FepB by SDS PAGE (Figure 118) were combined for

histidine-tag cleavage with C3 protease. The resulting protein was purified *via* a second nickel column. The desired protein was collected in the fractions eluted before the imidazole gradient (Figure 118). The presence of FepB with cleaved histidine tag was confirmed *via* SDS PAGE analysis (Figure 120).

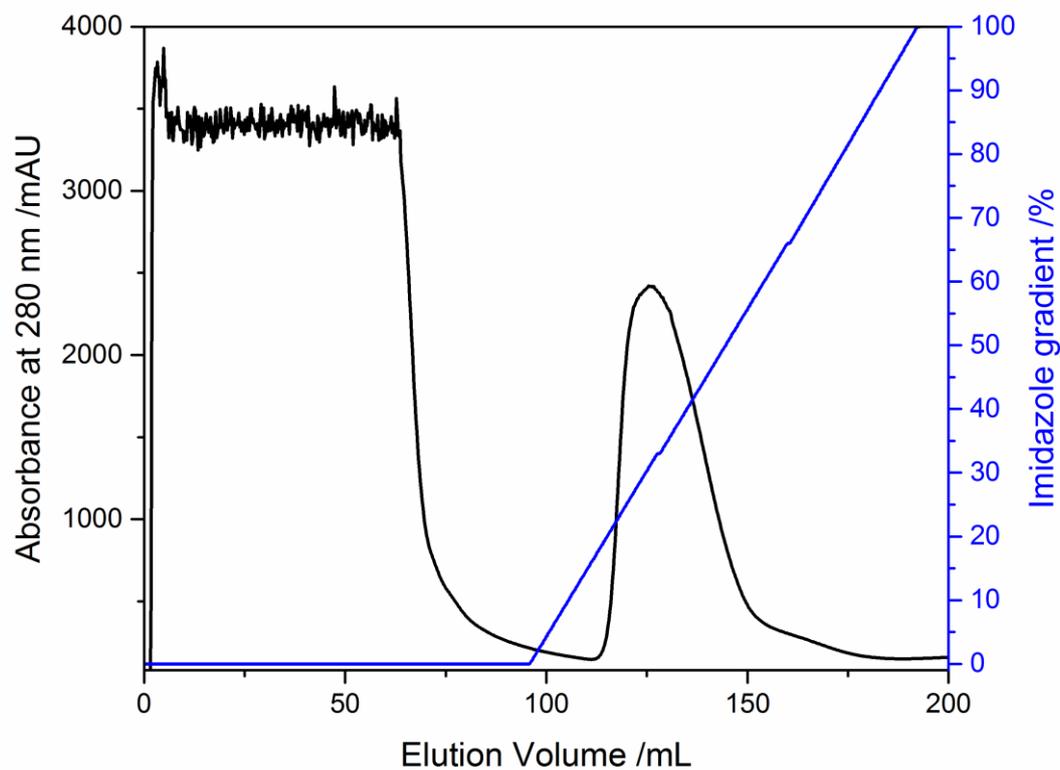


Figure 117: UV-visible absorbance trace at 280 nm (black) for elution of FepB from nickel column over an imidazole gradient (blue) Elution volume from 114 mL to 170 mL was collected for the absorbance band observed at 280 nm.

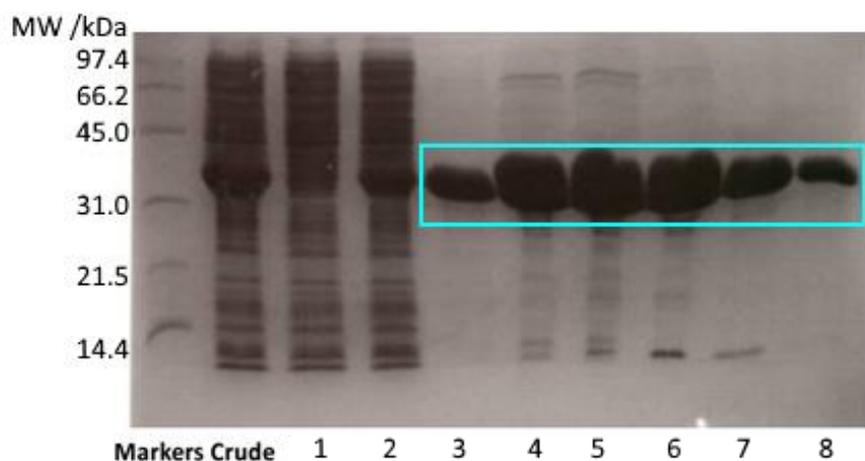


Figure 118: SDS PAGE analysis confirming the presence of FepB in fractions collected after nickel column chromatography. 4 mL fractions shown in the light blue box were collected and loaded in order of total elution volume. Lane 1: 32-36 mL, lane 2: 52-56 mL, lane 3: 114-118 mL, lane 4: 118-122 mL, lane 5: 122-126 mL, lane 6: 130-134 mL, lane 7: 142-146 mL, lane 8: 150-154 mL. All fractions were shown to contain FepB.

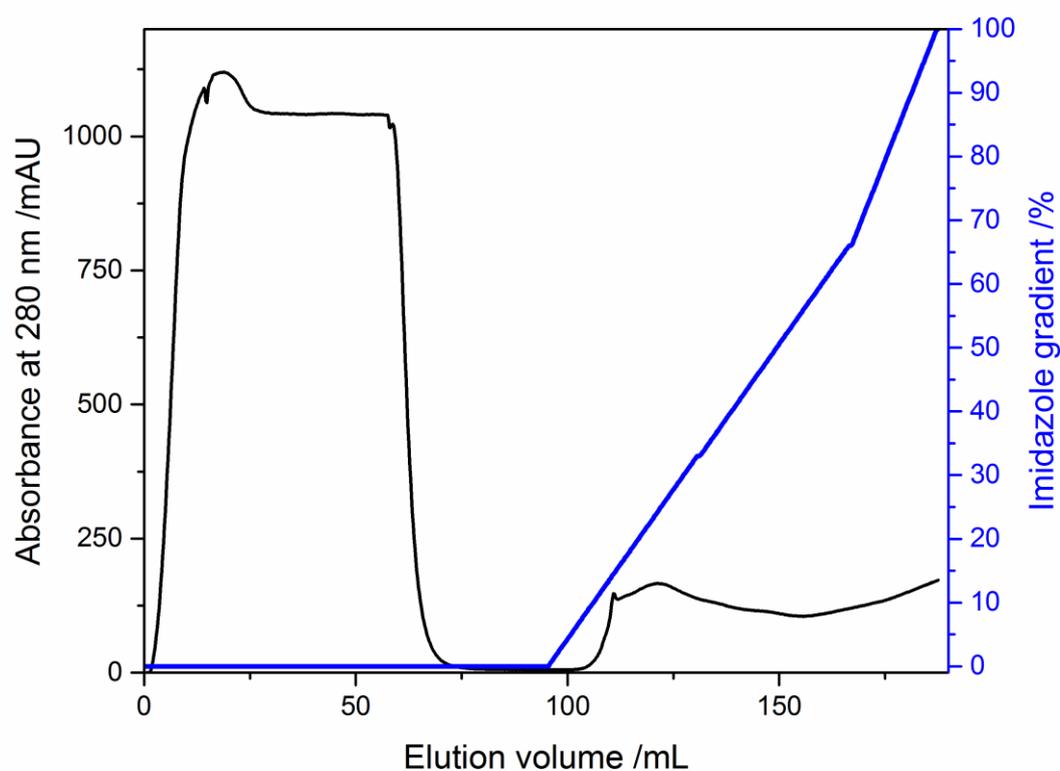


Figure 119: UV-visible absorbance trace at 280 nm (black) for elution of crude FepB from second nickel column over an imidazole gradient (blue). Fractions from 2 mL to 68 mL were collected for the absorbance band observed at 280 nm.

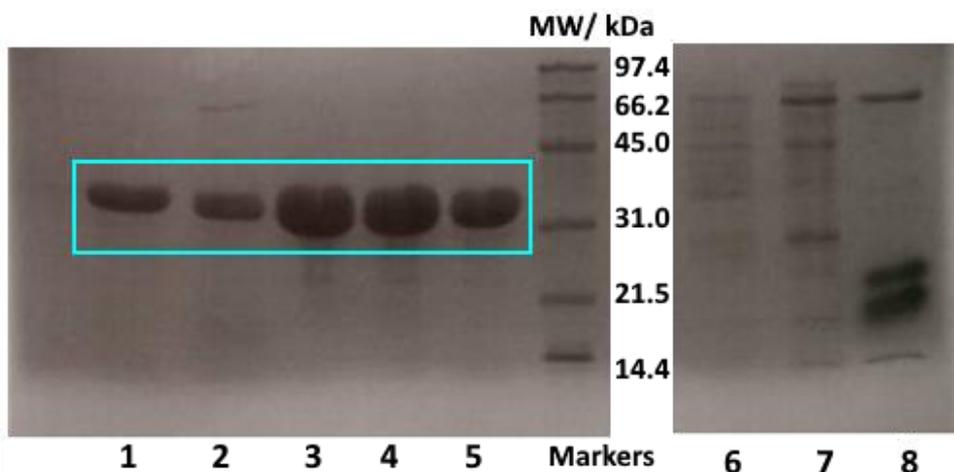


Figure 120: SDS PAGE of fractions after the second nickel column. Lane 1: tagged FepB, lane 2: crude untagged FepB. Untagged protein was confirmed in collected fractions of elution volume from 14 mL to 66 mL, lane 3: 14-18 mL, lane 4: 26-30 mL, lane 5: 62-66 mL. There was no FepB present in the fractions after the imidazole gradient as shown in lanes 6: 108-112 mL, 7: 116-120 mL. FepB fractions are highlighted in the light blue box.

The fractions containing untagged FepB were pooled and purified by size exclusion chromatography. The fractions observed as the major peak in the chromatogram (Figure 121) contained protein with a molecular weight estimated at around 35,000-40,000 Da, expected to be monomeric FepB. These fractions were collected, and the purity of the protein confirmed by SDS PAGE (Figure 122). Correct folding of the protein was established by circular dichroism (Figure 123).^{430,431} The final pure protein, confirmed as monomeric FepB by ESI-MS, was estimated to be a total of 450 mg, and was concentrated to 60 mg mL⁻¹. The concentrated protein was stored at -80°C until required.

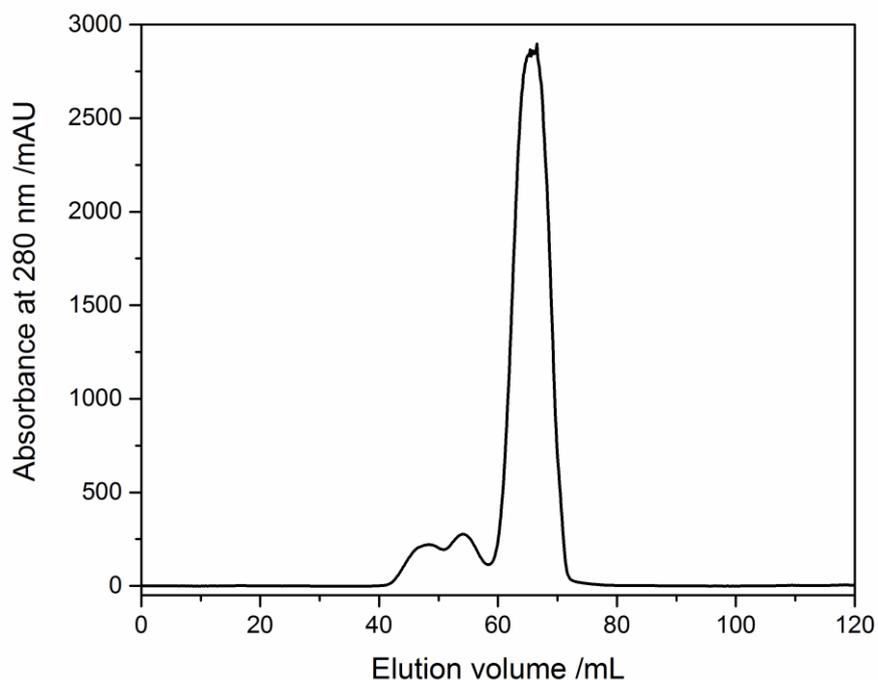


Figure 121: Chromatogram showing absorbance at 280 nm (black) for fractions collected of size exclusion chromatography of FepB. Fractions from elution volume 62 mL to 74 mL were collected, giving the major absorbance peak.

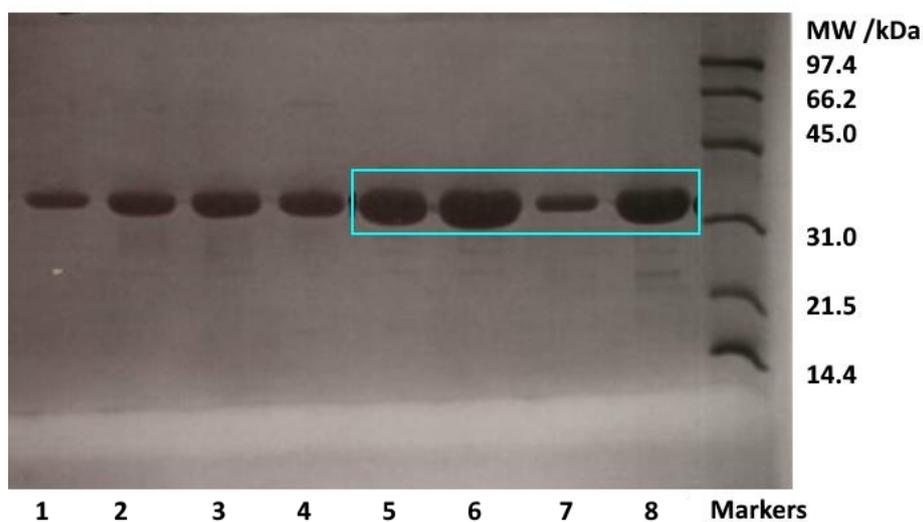


Figure 122: SDS PAGE analysis of the collected fractions of FepB after size exclusion chromatography. Lane 1: crude FepB before size exclusion, pure FepB was confirmed in collected fractions by elution volume from 62 mL to 74 mL, lane 2: 46-50 mL, lane 3: 54-58 mL, lane 4: 58-62 mL, lane 5: 62-66 mL, lane 6: 66-70 mL, lane 7: 70-74 mL, lane 8: 70-74 (twice concentrated). Collected fractions are highlighted in the light blue box.

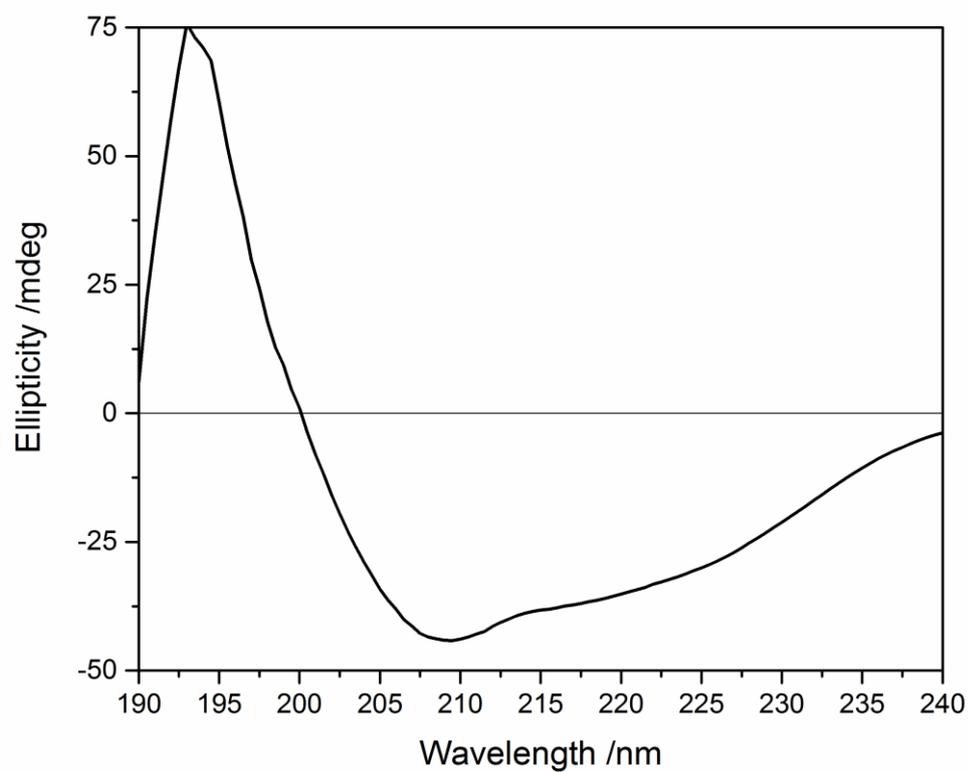


Figure 123: Circular dichroism spectrum of 0.4 mg mL⁻¹ FepB in 5 mM Tris-HCl pH 7.5. to confirm correct folding of the protein.

5.4 Iron(III)-Siderophore Binding of FepB

Binding of FepB was assessed for both hexadentate and tetradentate iron(III)-bound siderophores. As expected, FepB bound to iron(III)-enterobactin with nanomolar affinity, producing three consistent binding curves giving a K_d of 24 ± 2 nM (Figure 124), concordant with literature value of ~ 30 nM (Table 19),¹⁴⁴ providing validation of the fluorescence quenching titration method. FepB was shown to bind the hexadentate siderophore mimic iron(III)-MECAM with comparable affinity to iron(III)-enterobactin with a K_d of 28 ± 6 nM (Figure 125) (Table 19). Both hexadentate species were bound to FepB in the Λ -configuration, as shown by the circular dichroism spectra (Figure 126).

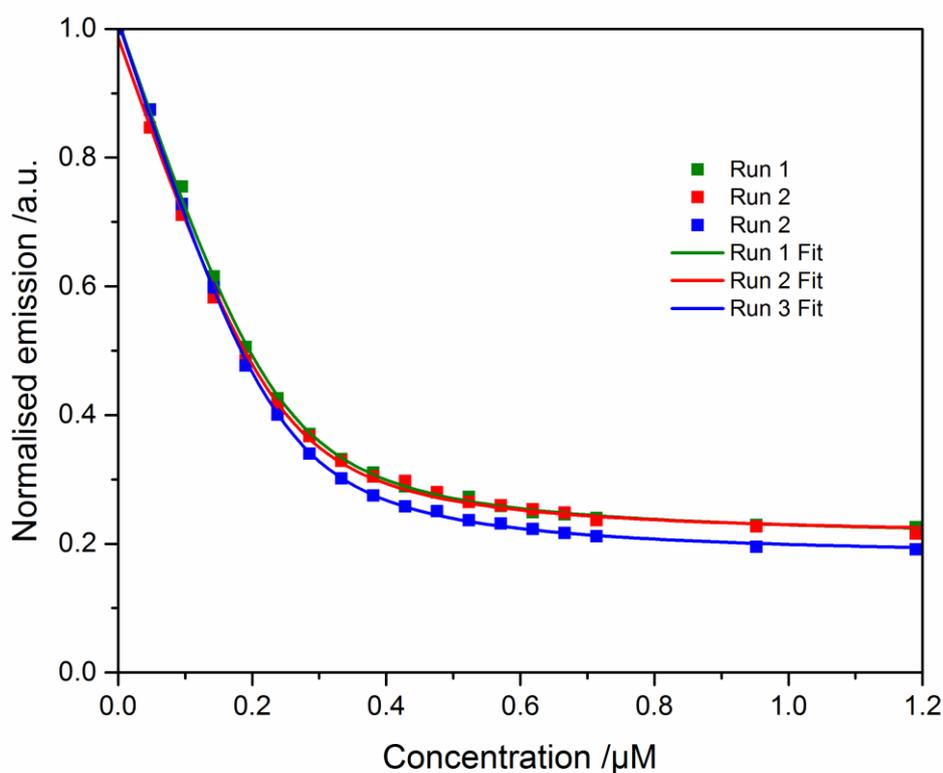


Figure 124: Fluorescence quenching titration curves for FepB iron(III)-enterobactin. Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for FepB-iron(III)-enterobactin and the model from Dynafit.³²⁶ FepB at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-enterobactin⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

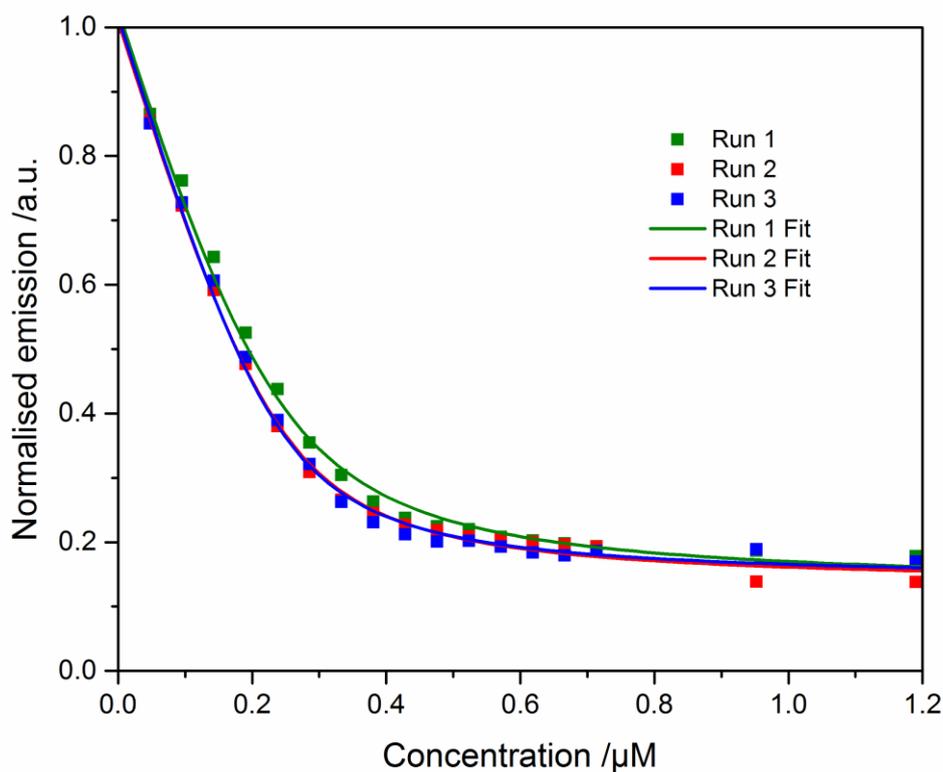


Figure 125: Fluorescence quenching titration curves for FepB-iron(III)-MECAM and the model from Dynafit.³²⁶ FepB at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M iron(III)-MECAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Table 19: K_d values calculated for FepB-iron(III)-enterobactin and FepB-iron(III)-MECAM *via* fluorescence quenching titration and associated errors, with a weighted average, as in Chapter 6.

FepB	Iron(III)-Enterobactin		Iron(III)-MECAM	
	K_d /nM	Error /nM	K_d /nM	Error /nM
Run 1	25.3	1.6	37.5	5.1
Run 2	23.4	2.0	27.0	2.9
Run 3	23.5	1.2	24.5	3.2
Average	24	2	28	6

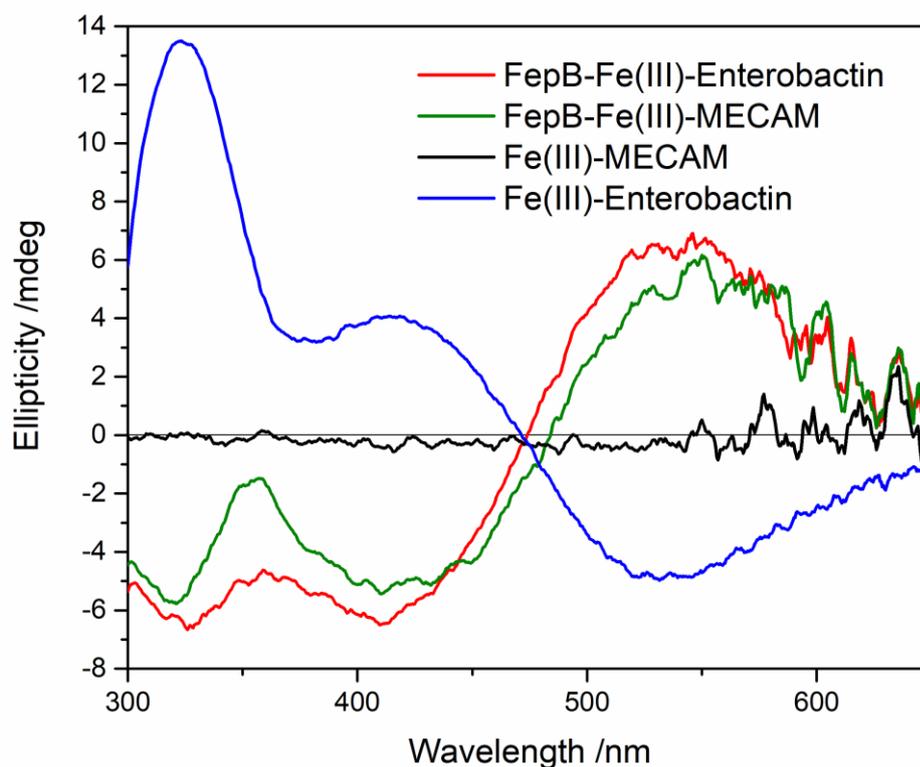


Figure 126: Circular dichroism spectra for FepB plotted as ellipticity in mdeg vs wavelength from 300 to 650 nm. The spectra were run for 50 μ M FepB in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with 50 μ M iron(III)-enterobactin or iron(III)-MECAM as well as iron(III)-enterobactin or iron(III)-MECAM at 50 μ M in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without FepB. The spectra show the induction of the Λ -configuration upon introduction of FepB to each iron(III)-ligand.

In contrast, tetradentate iron(III)-siderophores and siderophore mimics did not bind tightly to FepB. Circular dichroism proved that there was some weak interaction for iron(III)-4-LICAM, iron(III)-5-LICAM and iron(III)-bisDHBS, as solutions of each ligand, with no preference for either Λ or Δ configured complexes, were converted to a Λ -configured

majority in the presence of FepB (Figure 127). The same was true for both salmochelin mimics, iron(III)-Sal-4-LICAM and iron(III)-Sal-5-LICAM (Figure 128).

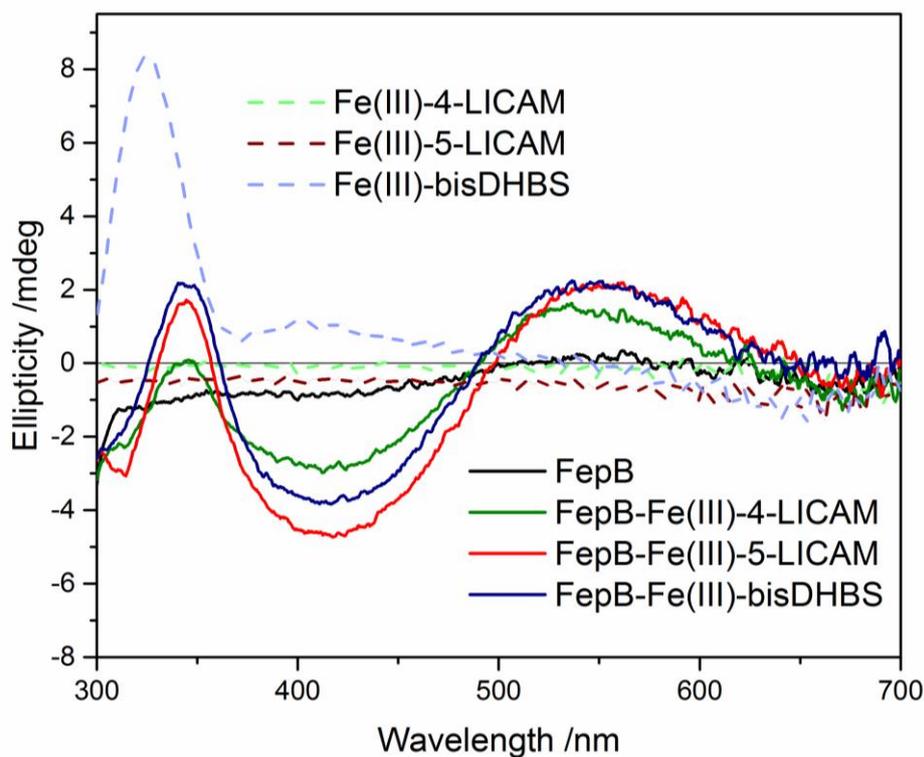


Figure 127: Circular dichroism spectra for FepB plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra are run for 50 μ M FepB in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μ M iron(III)-n-LICAM (n = 4, 5) or iron(III)-bisDHBS ligands as well as the ligands at 50 μ M in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without FepB. The spectra show the induction of Λ -configuration upon introduction of FepB to each iron(III)-ligand.

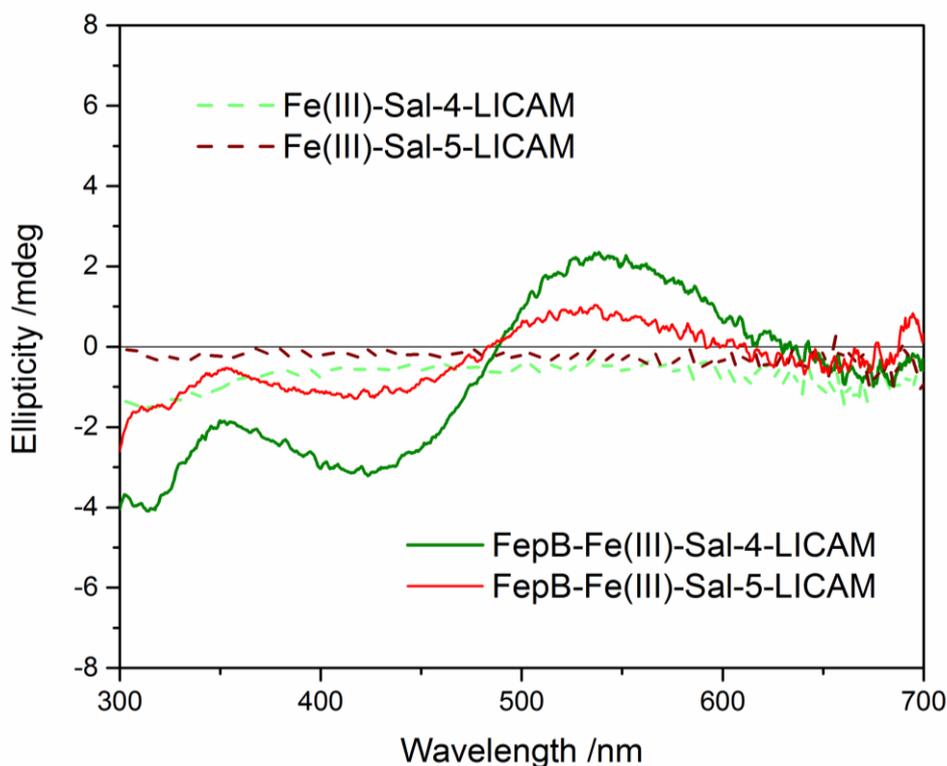


Figure 128: Circular Dichroism spectra for FepB plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm, using 50 μM FepB in 0.11 M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μM iron(III)-Sal- n -LICAM ($n = 4, 5$) ligands as well as the ligands at 50 μM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without FepB. The spectra show the induction of the Λ -configuration upon introduction of FepB to each iron(III)-Sal- n -LICAM ($n = 4, 5$).

The standard fluorescence quenching titration method was employed in an attempt to quantify the binding affinities of each tetradentate siderophore or siderophore mimic with FepB, but shallow incomplete binding curves were obtained for all combinations. A titration was carried out with an equal volume of buffer to establish the decrease in fluorescence of FepB due to dilution and repeated exposure at the excitation wavelength. All five iron(III)-tetradentate complexes gave a binding curve similar to or more shallow than this curve, suggesting that the binding affinities of the tetradentate iron(III)-siderophores were weak, and could not be quantified by this method. It was estimated that the K_d values were likely to be in excess of 1 μM .

To extend the study, salmochelin S4 was purchased in an attempt to establish whether FepB was capable of binding the iron(III)-bound hexadentate stealth siderophore. Unfortunately, the sample was deemed unsuitable for use, as the presence of salmochelin S4 was not confirmed by ESI-MS and many different species were identified by analytical HPLC.

HPLC was run after 1 hour of equilibration of the sample in water and then again from the same sample after 24 hours (Figure 129). The chromatograms were very similar, with almost identical peak ratios, suggesting that salmochelin S4 was not slowly hydrolysing in aqueous solution, but rather that multiple stable species were already present in the sample. The literature suggested that salmochelin S4 was likely to elute after any hydrolysis products, at around 18 minutes, however the major peak eluted at around 7 minutes.²³⁵

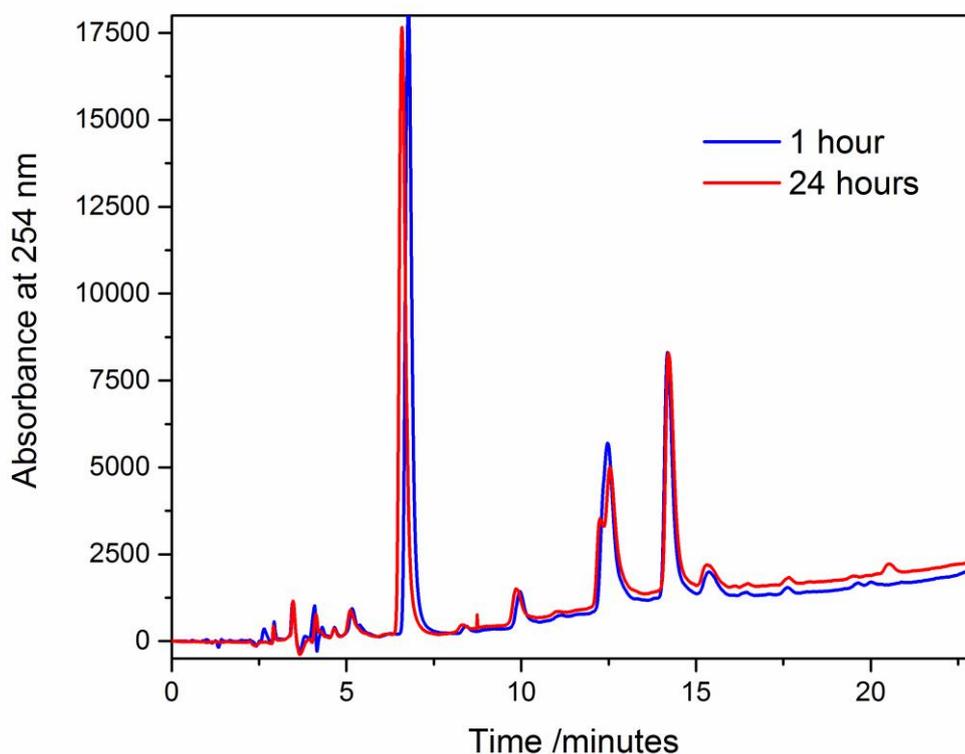
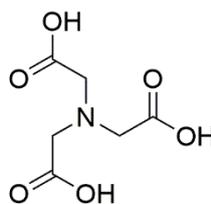


Figure 129: HPLC chromatograms showing multiple species present in the purchased salmochelin S4 sample. The sample in water containing 0.1% formic acid was injected with a 6-40% gradient of acetonitrile in water with 0.1% formic acid over 25 minutes.

5.5 Iron(III)-Siderophore Binding by VctP

The VctPDGC uptake system was reported, *via* growth study analysis, to be capable of uptake of siderophore-free iron.²⁰¹ It was decided therefore to assess whether VctP was capable of binding iron(III) from iron(III)-NTA the source of iron(III) used to form siderophore complexes in the fluorescence assay. The NTA (nitrilotriacetic acid) ligand was employed as a weak iron(III) chelating ligand, to prevent the formation of insoluble hydroxides, without affecting the results of the titration.^{95, 186, 432}



NTA

The binding curve upon titration of VctP with iron(III)-NTA showed a similar profile to that of a buffer dilution titration, suggesting that VctP did not bind iron(III) from the iron(III)-NTA source with any measurable affinity. This suggested that the published growth studies may have contained an uncharacterised chelating agent that was able to deliver iron(III) to the VctPDGC system, or that an uncharacterised system is involved in free iron(III) transport in *V. cholerae*.²⁰¹ This implied that any further fluorescence quenching of VctP observed upon addition of iron(III)-siderophores would be due to the binding of the whole iron(III)-siderophore complex, rather than the iron(III)-siderophore complexes merely acting as a source of free iron(III).

A standard binding curve showed quenching of VctP fluorescence by iron(III)-enterobactin, but did not reach a defined end-point, suggesting the binding was too weak to be quantified. Therefore, the concentration of both VctP and iron(III)-enterobactin were increased 10-fold, to promote complex formation.⁴¹⁴ The experiment was repeated in triplicate, achieving binding curves with well-defined end-points (Figure 130) and gave an overall K_d of 369 ± 25 nM (Table 20). This indicated weak binding compared to the previously observed low nanomolar affinities, and was around 12 times weaker than for FepB of around 30 nM.¹⁴⁵ Circular dichroism showed that Δ -configured iron(III)-enterobactin was however converted to an equilibrium containing a majority of Λ -configured complex in the presence of VctP (Figure 131). It must be noted that at 50 μ M, the circular dichroism spectra were run at a concentration well above the recorded dissociation constant.

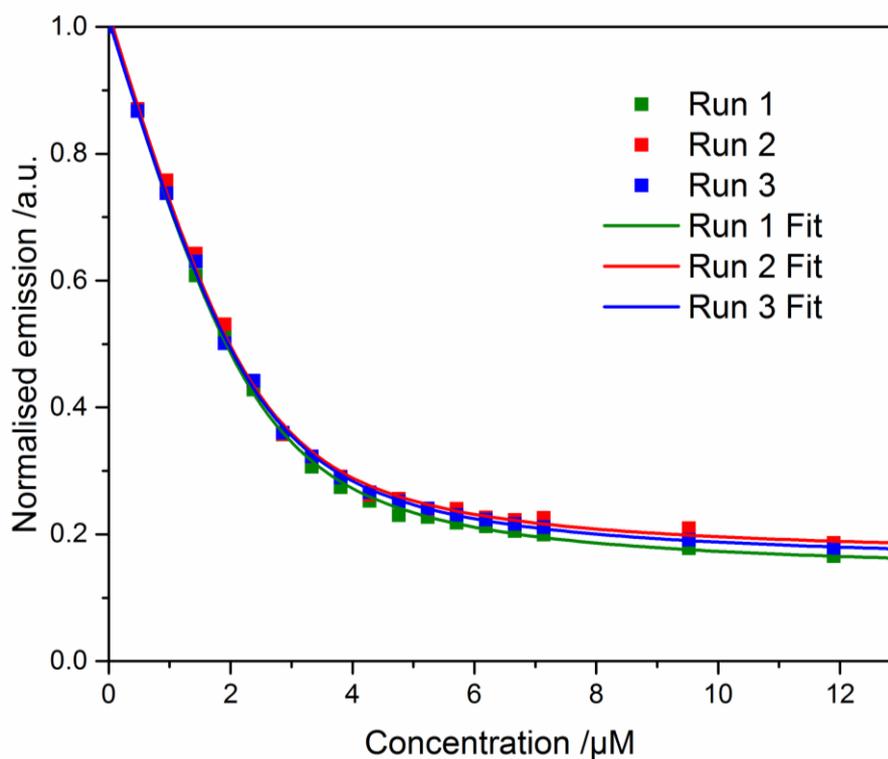


Figure 130: Plot of normalised fluorescence emission vs ligand concentration in μM for VctP-iron(III)-enterobactin and the model from Dynafit.³²⁶ VctP at a concentration of 2.4 μM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl titrated with aliquots of 120 μM iron(III)-enterobactin⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Table 20: K_d values calculated for VctP-iron(III)-enterobactin *via* fluorescence quenching titration and associated errors, with a weighted average, as in Chapter 6.

VctP	Iron(III)-Enterobactin	
	K_d /nM	Error /nM
Run 1	375	25
Run 2	350	38
Run 3	371	23
Average	369	25

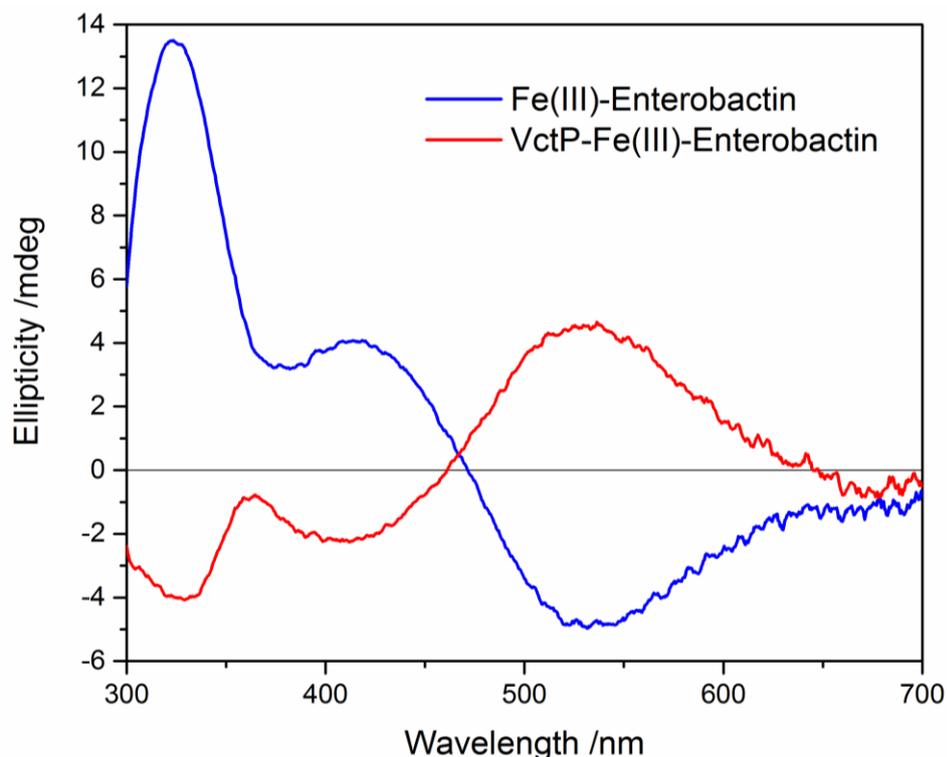


Figure 131: Circular dichroism spectra for VctP plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm for 50 μM VctP in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with 50 μM iron(III)-enterobactin as well as iron(III)-enterobactin at 50 μM in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer without VctP. The spectra show the induction of the Λ -configuration upon introduction of VctP iron(III)-enterobactin, with a transition from a Δ -configured ligand complex to a Λ -configured protein-ligand complex.

Fluorescence quenching titrations were repeated for VctP with iron(III)-MECAM. The binding curve obtained was steep, with full quenching achieved at a 1:1 ratio of VctP to iron(III)-MECAM (Figure 132). As a result, there were not enough data points around the curved region of the plot for the data to provide an accurate K_d , indicating that iron(III)-MECAM bound very tightly. In an attempt to obtain an adequate curve for fitting, the concentration of VctP and iron(III)-MECAM were decreased 5-fold, but this did not improve the curve (Figure 133). The concentrations were decreased by 20-fold, and a Dostal DOSY auto-titrating system was used to pipette 0.25 μL aliquots of ligand solution, to gain more data points around the 1:1 ratio region of the curve. Unfortunately, this method reached the limit of the detector of the fluorescence spectrometer, and the signal to noise ratio was too low to achieve accurate K_d values. Thus, the binding was estimated to be in the picomolar range, but could not be quantified.

The experiments were repeated for VctP with iron(III)-bisDHBS, iron(III)-4-LICAM and iron(III)-5-LICAM and all were estimated to have a K_d in the picomolar range. In addition,

circular dichroism confirmed that iron(III)-bisDHBS, iron(III)-5-LICAM and iron(III)-4-LICAM all bound to VctP with an induced Λ -configuration (Figure 134).

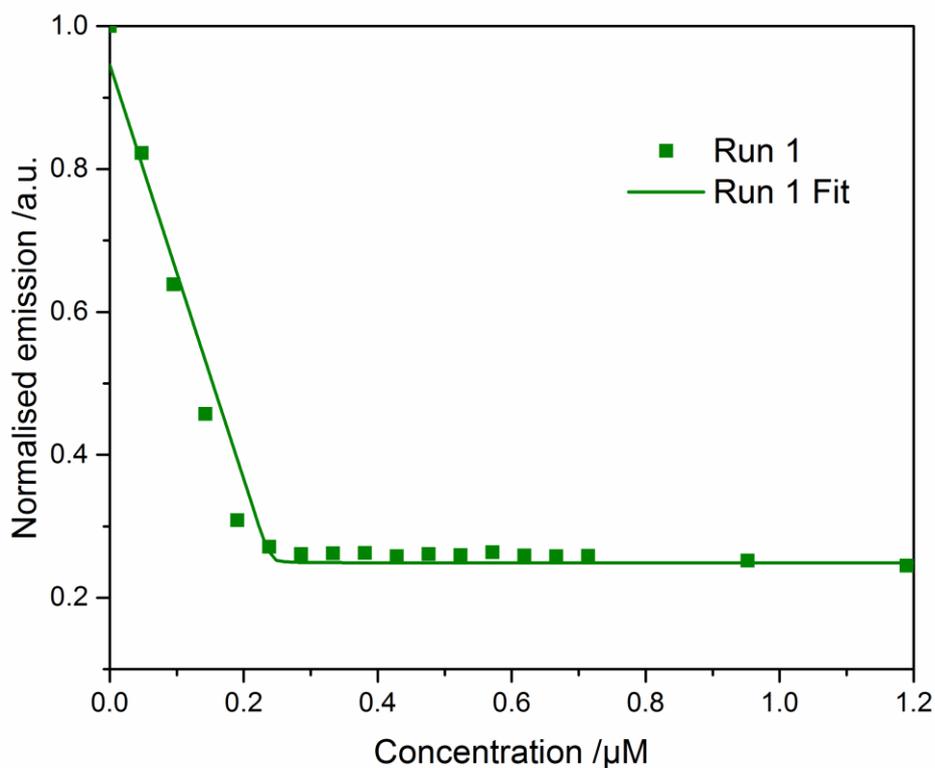


Figure 132: Fluorescence quenching titration curve for VctP-iron(III)-MECAM and the model from Dynafit.³²⁶ VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM iron(III)-MECAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

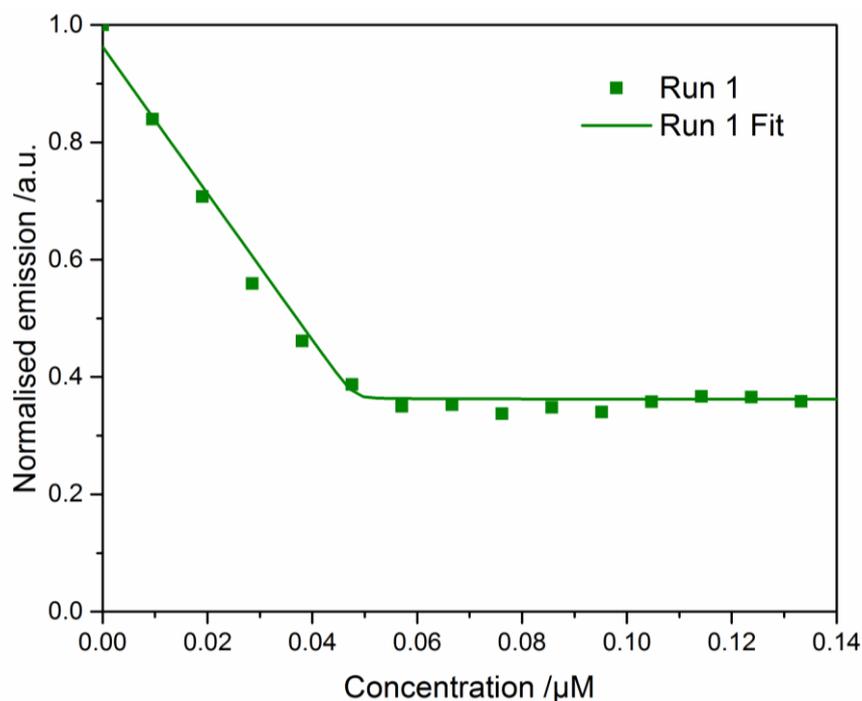


Figure 133: Fluorescence quenching titration curve for VctP-iron(III)-MECAM and the model from Dynafit.³²⁶ 48 nM VctP in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with 2.4 μM iron(III)-MECAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

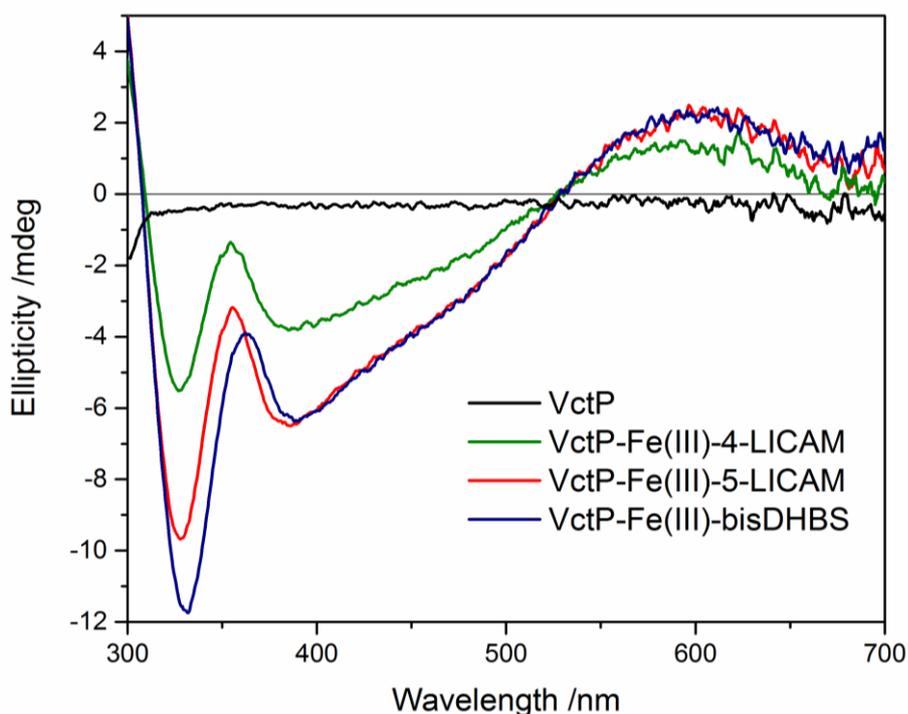


Figure 134: Circular Dichroism spectra for VctP plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra were run for 50 μM VctP in 0.11 M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μM iron(III)-n-LICAM ($n = 4, 5$) and iron(III)-bisDHBS ligands.

VctP was titrated with the salmochelin mimics iron(III)-Sal-4-LICAM and iron(III)-Sal-5-LICAM. The standard method was used for quantification of the binding of iron(III)-Sal-4-LICAM to VctP, and with a small extension to ensure a defined end-point, was used for iron(III)-Sal-5-LICAM. Satisfactory binding curves were achieved in triplicate for both ligands (Figure 135 and Figure 136) giving an average K_d of 7 ± 3 nM for VctP-iron(III)-Sal-4-LICAM and 9 ± 3 nM for VctP- iron(III)-Sal-5-LICAM (Table 21). These results suggest that VctP is likely to be able to bind salmochelin S1 with high affinity, and *V. cholerae* is probably able to use the salmochelin stealth siderophores for iron(III)-uptake via the VctPDGC system.

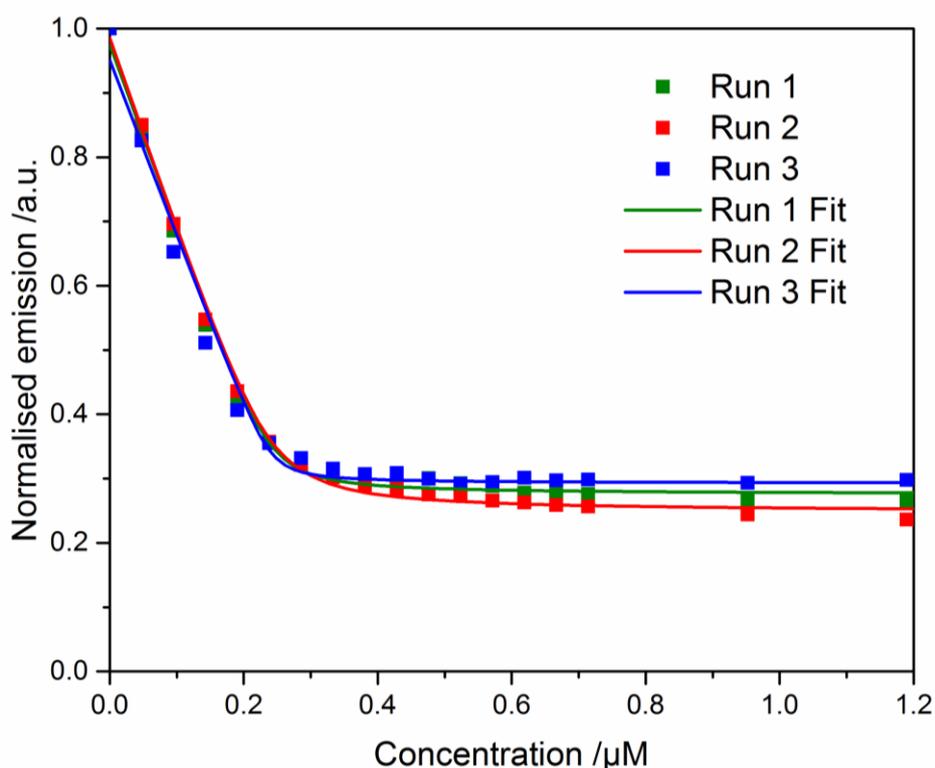


Figure 135: Fluorescence quenching titration curve for VctP-iron(III)-Sal-4-LICAM and the model from Dynafit.³²⁶ VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 µM iron(III)-Sal-4-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

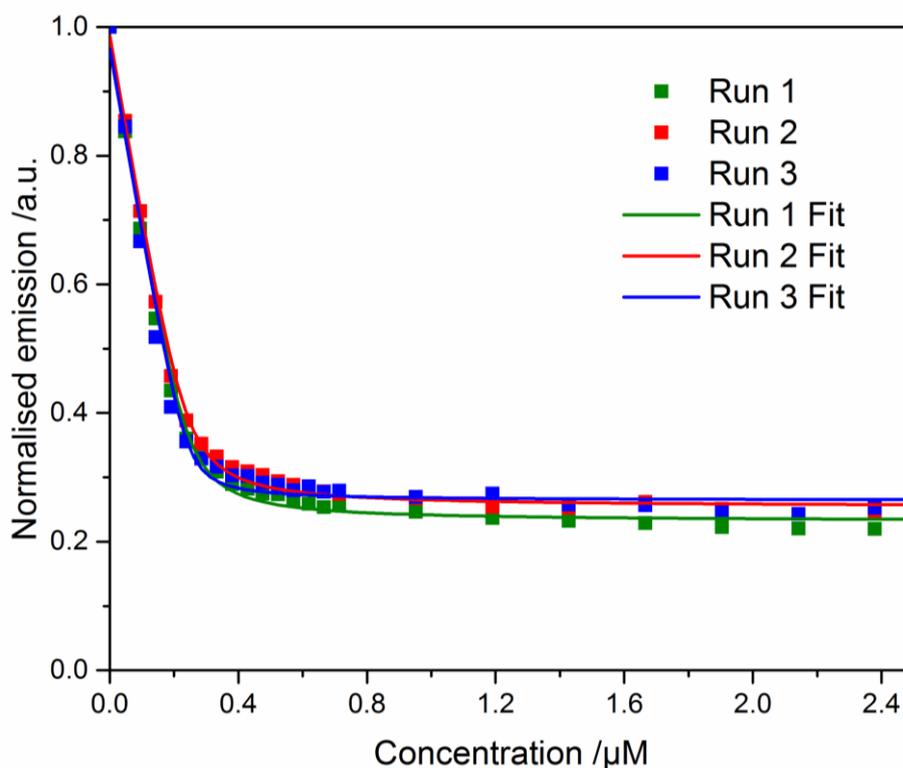


Figure 136: Fluorescence quenching titration curve for VctP-iron(III)-Sal-5-LICAM and the model from Dynafit.³²⁶ VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μ M iron(III)-Sal-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

Table 21: K_d values calculated for VctP-iron(III)-Sal-4-LICAM and VctP-iron(III)-Sal-5-LICAM *via* fluorescence quenching titration and associated errors, with a weighted average, as in Chapter 6.

VctP	Iron(III)-Sal-4-LICAM		Iron(III)-Sal-5-LICAM	
	K_d /nM	Error /nM	K_d /nM	Error /nM
Run 1	3.4	1.5	10.5	1.7
Run 2	6.5	1.4	11.7	1.6
Run 3	10.0	1.3	4.7	1.9
Average	7	3	9	3

The circular dichroism spectra of both complexes showed an induced Λ -configuration upon complexation with VctP (Figure 137). As VctP contains histidine and tyrosine residues equivalent to those in CeuE, it is likely that the Λ -configured binding mode between iron(III)-bound tetradentate siderophores is similar in VctP and CeuE.^{90, 310}

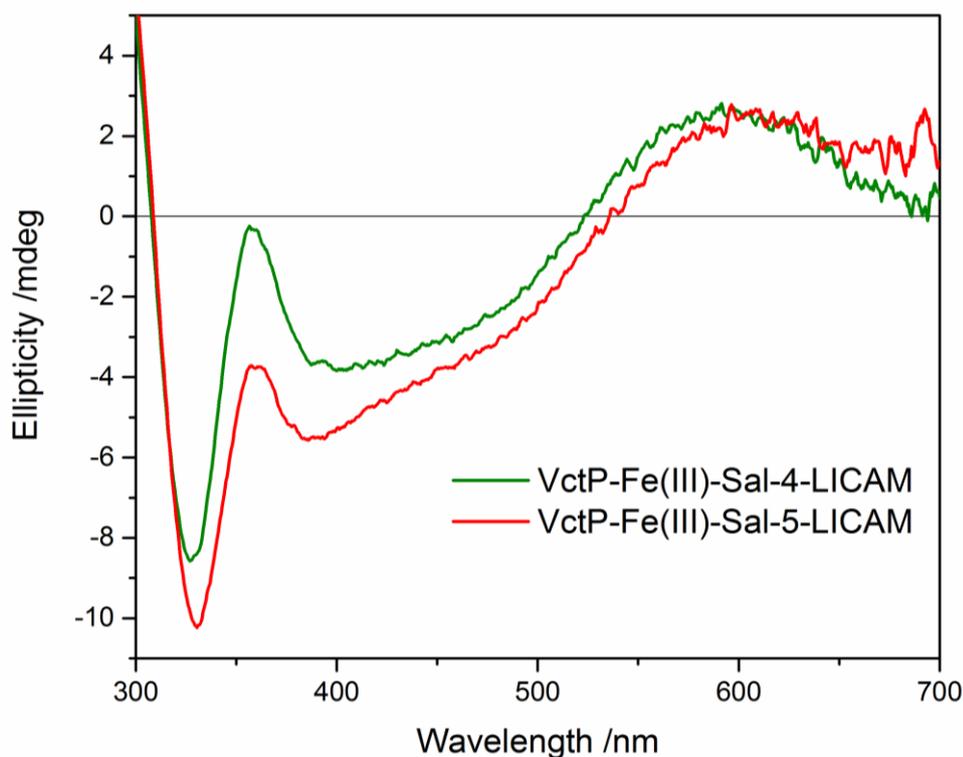


Figure 137: Circular Dichroism spectra for VctP plotted as ellipticity in mdeg vs wavelength from 300 to 700 nm. The spectra were run for 50 μ M VctP in 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer with and without 50 μ M iron(III)-Sal-n-LICAM ($n = 4, 5$) ligands. The spectra show the induction of Λ -configuration upon introduction of VctP to each iron(III)-Sal-n-LICAM ($n = 4, 5$).

Crystal structure determination of iron(III)-Sal-5-LICAM with VctP was unsuccessful. Crystals of apo VctP provided good diffraction data, however as the structure was already solved and deposited (PDB ID: 3TEF), it was decided unnecessary to pursue further structure solution. Although crystals of apo VctP when soaked with iron(III)-Sal-5-LICAM took on a purple colour, indicating uptake of the iron(III)-complex, the crystals did not diffract adequately for structure solution. Further screening of crystal conditions is needed if a 3D structure with a tetradentate ligand is to be achieved.

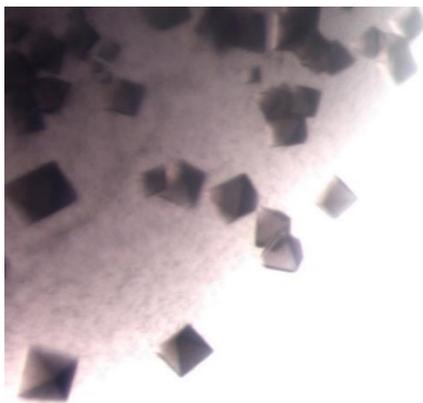


Figure 138: VctP crystals soaked with iron(III)-Sal-5-LICAM.

5.6 Comparison of FepB, CeuE and VctP iron(III)-siderophore binding

Although FepB, CeuE and VctP have all been reported to be enterobactin uptake periplasmic binding proteins,^{115, 144, 192} the binding profiles of iron(III)-siderophores and iron(III)-siderophore mimics, and the associated affinities are quite different (Table 22). Dissociation constants range from micromolar affinities for iron(III)-salmochelin mimics binding to CeuE, to picomolar affinities for iron(III)-MECAM, iron(III)-n-LICAM (n= 4, 5) and iron(III)-bisDHBS binding to VctP. All combinations of ligand and protein do however produce Λ -configured complexes as proven by circular dichroism.

Table 22: Dissociation constants for periplasmic binding proteins with iron(III)-siderophores and iron(III)-siderophore mimics. Dissociation constants taken from the literatures are referenced.

Iron(III)-Ligand	FepB K_d /nM	CeuE K_d /nM	VctP K_d /nM
Enterobactin	24 ± 2	3500 ± 300	369 ± 25
MECAM	28 ± 6	30 ± 11	<1
bisDHBS	>1000	10 ± 4 (lit. ⁹⁰)	<1
4-LICAM	>1000	21 ± 6 (lit. ³¹⁰)	<1
5-LICAM	>1000	<10 (lit. ³¹⁰)	<1
Sal-4-LICAM	>1000	15600 ± 2300	7 ± 3
Sal-5-LICAM	>1000	511 ± 76	9 ± 3

FepB: Hexadentate vs Tetradentate

FepB is optimised for binding of the hexadentate siderophore iron(III)-enterobactin, with nanomolar affinity, as previously reported,¹⁴³⁻¹⁴⁵ and can bind the iron(III)-MECAM siderophore mimic with similar high affinity. In contrast, FepB has only weak affinity for tetradentate siderophores, which suggests that three catechol units are required for tight binding of iron(III)-catecholate complexes. As *E. coli* has the cytoplasmic reductase YqjH, capable of reduction of iron(III)-enterobactin, it is logical that the hexadentate siderophore is taken up into the cytoplasm by the FepBCDG system, given that there is minimal hydrolysis of the trilactone in the extracellular matrix in the timescale before iron(III) sequestration.¹⁶³ The outer membrane receptors Fiu and Cir are known to be involved in the uptake of iron(III)-linear hydrolysis products iron(III)-trisDHBS, iron(III)-bisDHBS and iron(III)-DHBS, so it is likely that a completely separate set of uptake proteins, possibly involving IroC, are required for cytoplasmic transport of tetradentate catecholate siderophores in *E. coli*.^{152, 157, 160}

CeuE: Tetradentate vs Hexadentate

In contrast to FepB, CeuE is optimised for tight binding of the tetradentate siderophore iron(III)-bisDHBS, and binds the iron(III)-n-LICAM siderophore mimics with nanomolar affinity (Table 22).^{90,310} It is not optimised for binding of salmochelin mimics as much larger dissociation constants were achieved with these tetradentate iron(III)-siderophore mimics, and is unlikely to be involved in salmochelin S1 uptake. Interestingly CeuE binds iron(III)-MECAM with comparable nanomolar affinity to FepB, but has only micromolar affinity for iron(III)-enterobactin. It would be of interest to take this study further to establish why CeuE can bind tightly to iron(III)-MECAM but not iron(III)-enterobactin, and whether the hexadentate binding mode of CeuE in solution resembles the binding mode established for FepB.

VctP: Tetradentate and Hexadentate

VctP has strikingly high affinity for almost all iron(III)-siderophores and siderophore mimics tested, with several tetradentate complexes having picomolar dissociation constants (Table 22). The weakest binding affinity was for iron(III) enterobactin, 369 ± 25 nM, in the mid-nanomolar range, but still of possible biological relevance.^{185, 308} That VctP binds both hexadentate and tetradentate iron(III)-siderophores may explain the confusion in the literature as to whether the VctACDGP system is able to uptake iron(III)-enterobactin or its hydrolysis products.^{161, 200, 201} VctP has very high affinity for the linear hydrolysis product iron(III)-bisDHBS, and weaker affinity for iron(III)-enterobactin, but may ultimately be able to use either siderophore. With picomolar affinity, it would be of interest to see whether iron(III)-bisDHBS can indeed be released from VctP for cytoplasmic uptake of the iron(III)-complex, or whether the siderophore is too tightly bound to the protein. *V. cholerae* is able to survive in aquatic environments, in fish, crustaceans, zooplankton and as biofilms, as well as colonising humans, and an adaptable iron(III)-uptake system may enable survival in such a wide range of environments.^{190, 191, 200, 433} The fact that VctP is able to bind a wide range of siderophores with high affinity may be related to the persistence and virulence of *V. cholerae*.^{132, 190, 193}

Features Favouring Tetradentate Siderophore Binding

E. coli is the only species of the three that synthesises enterobactin. While *C. jejuni* and *V. cholerae* do not synthesise enterobactin, it is apparent that their periplasmic binding proteins can bind the iron(III)-bound linear hydrolysis products, but are less optimised for uptake of the full trilactone.

The differences in binding affinities of FepB, CeuE and VctP can be explained by differences in the siderophore binding clefts. FepB lacks the histidine and tyrosine residues conserved in CeuE and VctP (Section 5.1) that are fundamental to chelation of iron(III) to complete hexadentate coordination when a tetradentate siderophore is present.^{90, 186, 310} In solution, iron(III) complexes of tetradentate siderophore complexes are either in the biscatecholate 1:1 form with two coordination sites filled with coordinating solvent, or in the triscatecholate 2:3 triple helicate form.^{90, 260, 322} The 2:3 complex is too large to be accommodated in the FepB binding pocket, however, the 1:1 complex with only two catecholate units around the iron(III) centre, would need to employ water molecules from the surrounding solvent to fulfil octahedral co-ordination. Without direct coordination from the protein by a histidine and tyrosine (or equivalent side chains), the tetradentate siderophores can only bind weakly in the FepB binding pocket. Inspection of the published crystal structure of FepB-iron(III)-enterobactin (PDB ID: 3TLK), reveals that key hydrogen bonding residues Gly126, Arg242 and Arg301 chelate the catecholate oxygens of enterobactin with multiple hydrogen bonds in a specific orientation, allowing the accommodation of the full triscatecholate as described in section 5.1 (Figure 106). Without the potential for direct iron(III)-coordination from FepB, it is clear that the binding cleft is not optimised for tetradentate siderophores.

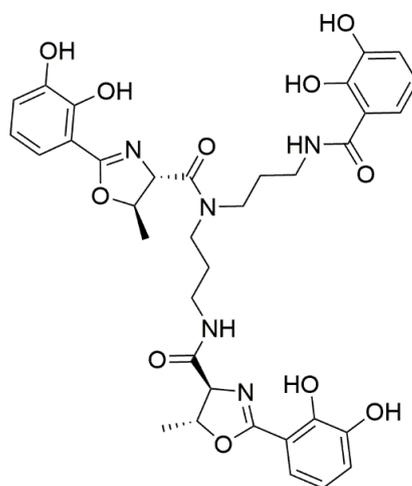
Scavenger Versus Exporter Species

For *C. jejuni* and *V. cholerae* to take advantage of enterobactin that is exported by species such as *E. coli*, these scavenger species may not encounter fresh enterobactin unless they are in close proximity to the exporter species. In addition, direct competition with the exporter species for one specific siderophore is a much greater challenge than utilising the full range of hydrolysed siderophore products.

The preference of CeuE for iron(III) tetradentate catecholate siderophores over hexadentate enterobactin may exist in *C. jejuni* as the species relies entirely on exogenous siderophores, and produces no siderophores of its own.^{115, 168, 172} Enterobactin is known to possess a degree of aqueous instability, and it may be advantageous for *C. jejuni* to acquire the more stable tetradentate hydrolysis product as well as enterobactin.^{86, 87} In addition, iron(III) release *via* a reductase in the cytoplasm from the linear hydrolysis products is likely to be more efficient than from iron(III)-enterobactin which has a very low redox potential of $E_{1/2} = -750$ mV.^{28, 55, 56, 72, 163, 434} If the majority of available iron(III)-enterobactin is first hydrolysed in the periplasm of *C. jejuni*, by the known trilactone esterase Cee, then the hydrolysis product iron(III)-bisDHBS will readily bind to CeuE for transport into the cytoplasm.^{90, 159} This means that only one ABC transporter system CeuBCDE is required for uptake of iron(III)

into the cytoplasm from all enterobactin-derived sources.¹⁵⁹ As bisDHBS is a lower denticity, less entropically favoured siderophore for iron(III)-chelation than enterobactin, iron(III) release should be more efficient from this complex when in the cytoplasm.^{73, 163, 261}

V. cholerae, as both a siderophore exporter and scavenger species, has the periplasmic binding protein VctP with similarities to both FepB and CeuE – which is able to bind both iron(III)-enterobactin and iron(III)-bisDHBS with biologically relevant affinities, allowing for maximum iron(III) sequestration. Based on the crystal structure of VctP (PDB ID: 3TEF) with very high structural similarity to CeuE, it is likely that His 248 and Tyr 310 directly coordinate the iron(III) centre when binding tetradentate complexes, and this provides very high binding affinity. The fact that VctP also seems to bind iron(III)-enterobactin with a medium-strength affinity may be due to the fact that *V. cholerae* produces its own hexadentate catecholate siderophore vibriobactin as well as scavenging iron(III)-enterobactin.²⁰² Vibriobactin is comprised of an amide bonded backbone, so cannot be hydrolysed by an esterase in the same way as enterobactin.⁸⁴ Hence, the full hexadentate siderophore is taken up into the cytoplasm, and the iron released by the action of the reductase ViuB, which means that *V. cholerae* is already tuned for the uptake of hexadentate siderophores into the cytoplasm.



Vibriobactin

Although the ViuACDGP system is known for vibriobactin uptake, it has been shown that the VctACDGP system can also be employed, and the same is true for iron(III)-enterobactin derivatives. Thus, there is a degree of flexibility in the catecholate iron uptake machinery in *V. cholerae*. Examination of the structure of ViuP in complex with vibriobactin reveals structural similarities to VctP, with an r.m.s.d. of 3.21 Å over 228 residues.²⁰⁸ There are a number of residues that are involved in vibriobactin binding, including arginines and hydrophobic phenylalanines. Interestingly, histidine and tyrosine residues are located on two

nearby flexible loops which may function in the binding of lower denticity iron(III)-siderophore complexes. Further work is needed to study the mechanisms of VctACDGP and ViuACDGP for internalisation of iron(III)-enterobactin and the linear hydrolysis products.²⁰⁸

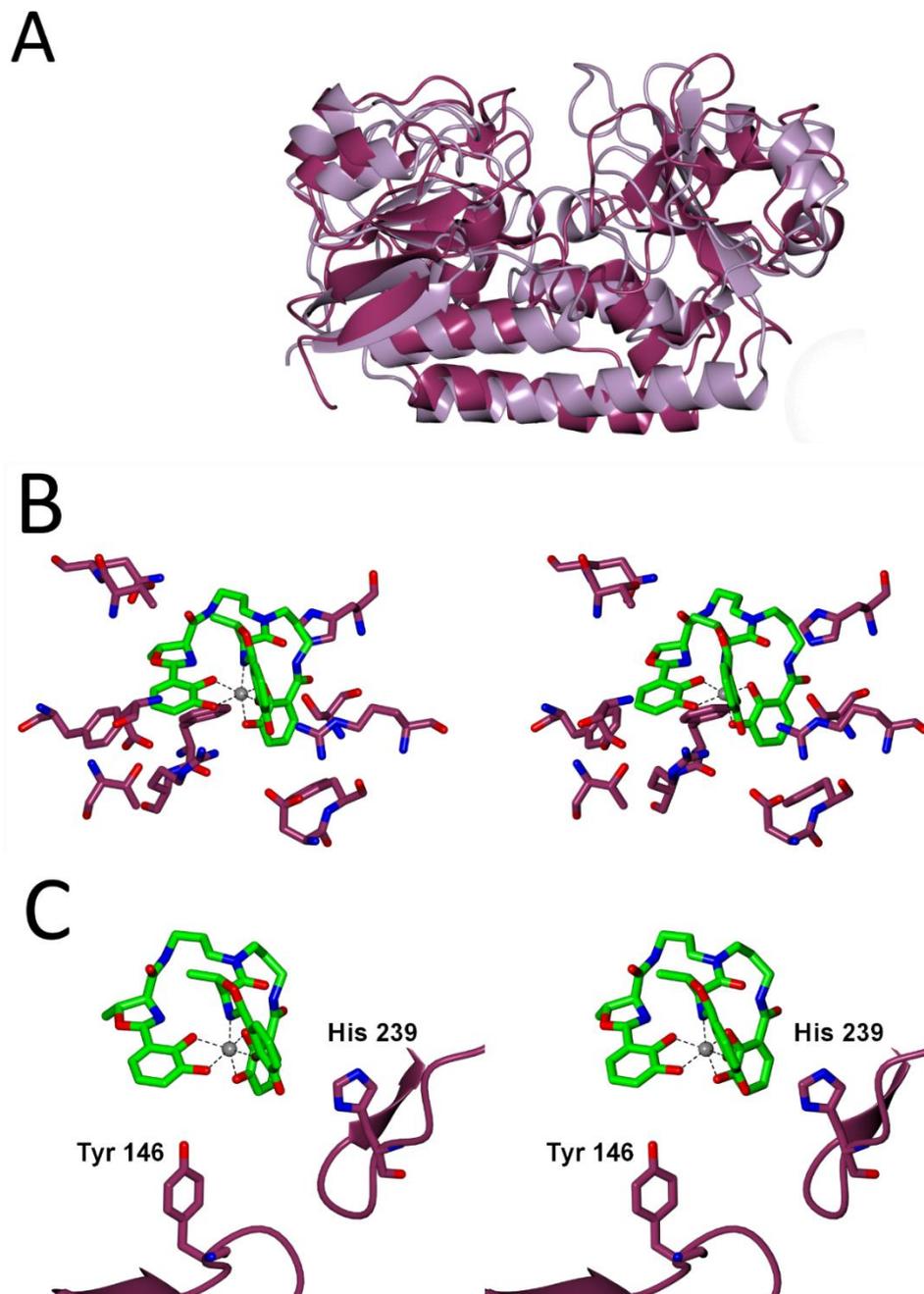


Figure 139: A: Overlay of VctP (lilac ribbons) with ViuP (dark purple), the vibriobactin binding PBP from *V. cholerae*. B: Stereo view of vibriobactin (cylinders coloured by atom type) in the binding cleft, and residues involved in binding (cylinders coloured by atom type with carbon in dark purple). C: Stereo view of neighbouring His 239 and Tyr 146 (cylinders coloured by atom type with carbon in dark purple) that may be involved in binding of lower denticity siderophore complexes (PDB ID: 3R5T).²⁰⁸ Figure produced using CCP4mg.

MECAM as a Mimic for Enterobactin

The fact that CeuE and VctP are able to bind iron(III)-MECAM with high affinity, but have a comparatively weak affinity for iron(III)-enterobactin suggests that MECAM is not a suitable mimic for enterobactin in these cases. Overlaying the coordinates for iron(III)-enterobactin from the FepB-iron(III)-enterobactin crystal structure (PDB ID: 3TLK) into the CeuE binding pocket of the CeuE-iron(III)-MECAM crystal structure (PDB ID: 2CHU) shows no structural reason why iron(III)-enterobactin has weak binding affinity for CeuE.^{146, 187} The catecholamides overlay well with the hexadentate iron(III)-MECAM arrangement (r.m.s.d of 1.83 over 33 atom positions), and the trilactone backbone would not cause steric clashes with the protein scaffold (Figure 140).

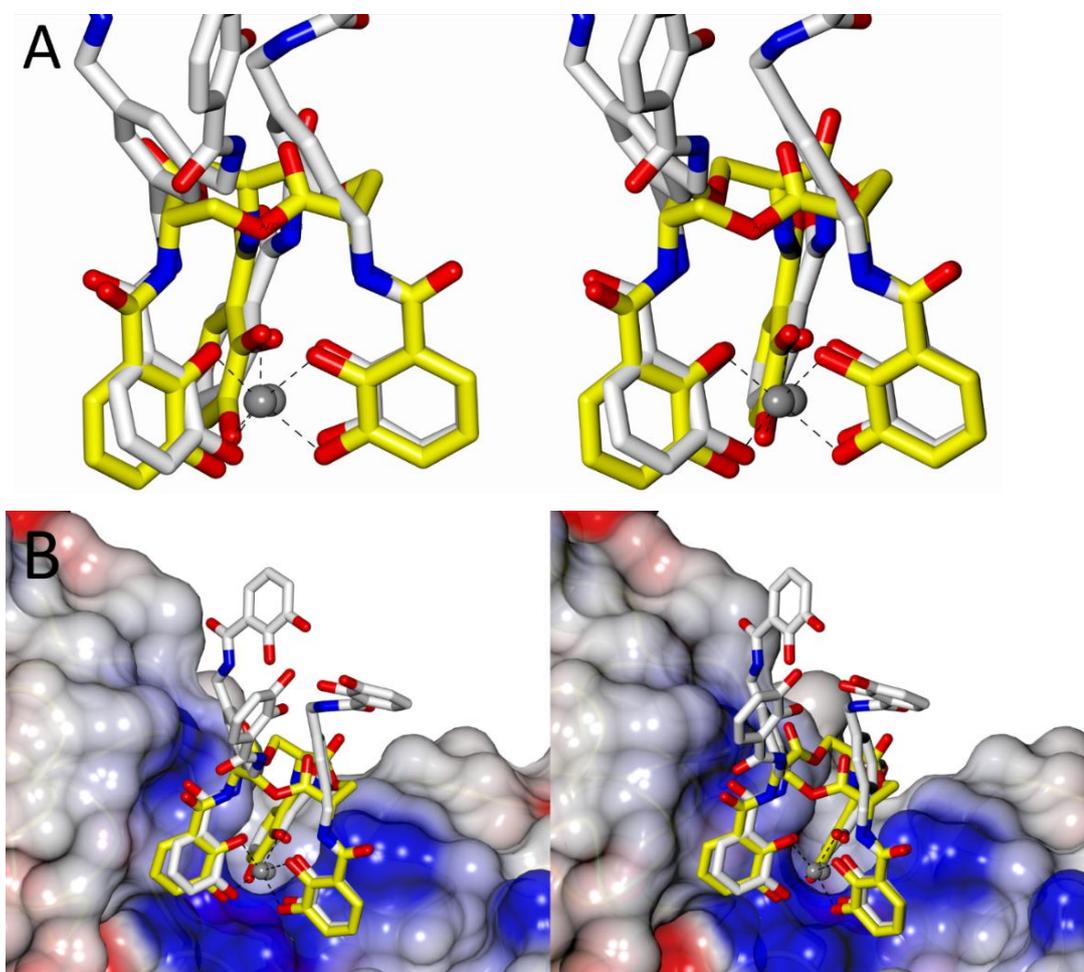
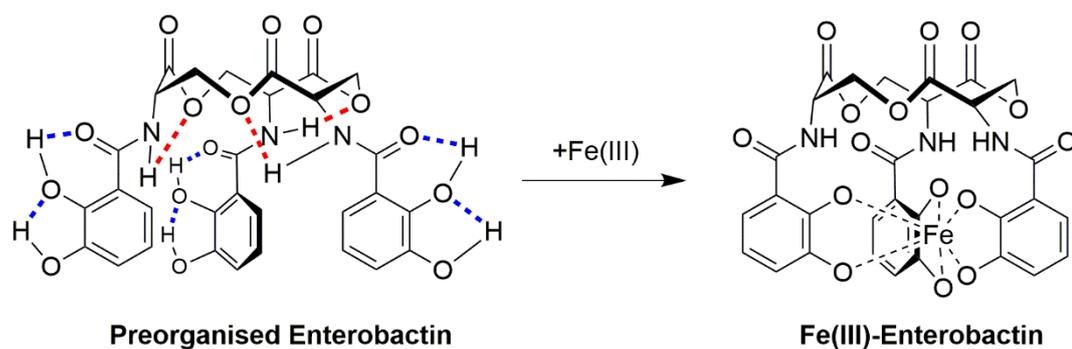


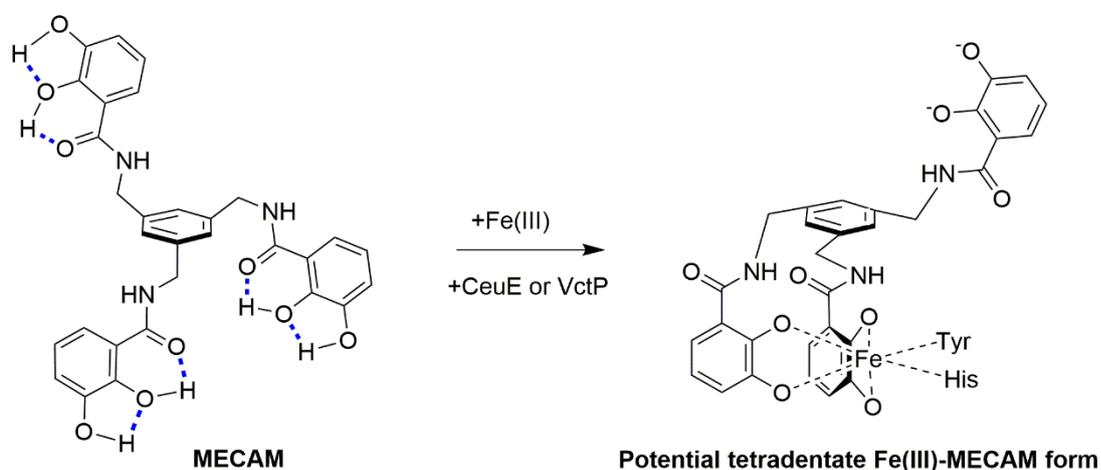
Figure 140: Stereo view of the overlay of CeuE-iron(III)-MECAM crystal structure (PDB ID: 2CHU) with coordinates for iron(III)-enterobactin from FepB-iron(III)-enterobactin crystal structure (PDB ID: 3TLK). Iron(III)-enterobactin is shown in cylinders coloured by atom type with carbon in yellow. Iron(III)-MECAM is shown in cylinders coloured by atom type with carbon in white. A: The catecholamide units of each ligand superposed atom by

atom. B: CeuE is depicted in surface view, coloured by electrostatic potential.^{146, 187} Figure produced using CCP4mg.

As there is no obvious structural reason why CeuE would oppose hexadentate iron(III)-enterobactin binding, it may be that a different binding mode is preferred in solution. There are two possible suggestions for how a hexadentate iron(III)-complex may bind with CeuE or VctP. Either the protein binds the hexadentate structure by folding the histidine and tyrosine residues out of the way of the complex, or the complex may give up one bidentate unit to allow for the histidine and tyrosine to coordinate the iron(III) centre. The reason for such differences in affinity between iron(III)-enterobactin and iron(III)-MECAM are therefore likely to be entropic, and may lie in the structure of each hexadentate compound before iron(III) binding. From computational modelling, it has been suggested that enterobactin has all three catechol units organised on the same face of the trilactone backbone, due to the presence of a hydrogen bond network (Scheme 16).^{62, 64, 435-437} This preorganised structure enables a very high affinity iron(III) complex as discussed in Chapter 1. For MECAM, there are fewer hydrogen bonds, and the planar aromatic backbone does not force pre-orientation of all three catechols towards one face. This property results in a weaker iron(III) binding affinity, but may allow MECAM to form a tetradentate binding mode, with one catechol unit folded away from the iron(III) centre, as is observed in the CeuE-iron(III)-MECAM crystal structure (PDB ID: 2CHU), where a 2:2:2 complex is achieved with each MECAM forming one tetradentate interaction and one bidentate interaction with the iron(III) centres (as discussed in Chapter 1).¹⁸⁷ This could mean that the iron(III)-MECAM complex is able to allow for His and Tyr to coordinate to the iron(III) centre more readily than iron(III)-enterobactin, which may explain the differences in binding affinity for each complex with both CeuE and VctP (Scheme 17). Further crystal structure determination may help to explain whether this hypothesis is accurate.



Scheme 16: Enterobactin has all catechol units organised on one face of the trilactone backbone before iron(III)-binding. Potential hydrogen bonds are shown in blue, and those involved in preorganisation are shown in red.



Scheme 17: MECAM does not have a preorganised structure before iron(III)-binding (hydrogen bonds shown in blue) so may be able to adopt a tetradentate binding mode more easily in the CeueE or VctP binding clefts, by folding one catechol unit away from the iron(III) centre, allowing coordination of His and Tyr.

Salmochelin S1 mimics

Neither CeueE or FepB bound the salmochelin mimics iron(III)-Sal-4-LICAM or iron(III)-Sal-5-LICAM tightly, with the lowest K_d achieved by CeueE with iron(III)-Sal-5-LICAM at 511 ± 76 nM: at least 50 times weaker than the non-functionalised iron(III)-5-LICAM equivalent. The fact that FepB does not have high affinity for the salmochelin mimics is probably because the mimics are tetradentate, and FepB is optimised for the binding of hexadentate siderophores. It would be of interest to establish whether FepB can bind salmochelin S4 or a suitable hexadentate mimic. As *E. coli* is known to synthesise salmochelin S4, there must be dedicated iron(III)-salmochelin uptake machinery, whether this is the FepBCDG system, the IroC system or another.^{157, 235, 236} As CeueE is optimised for

binding tetradentate siderophores, it is more likely that the salmochelin mimics cannot be bound by CeuE with high affinity due to a steric clash of the glucose unit with protein residues adjacent to the siderophore binding cleft, indicating CeuE is unlikely to be optimised for uptake of salmochelin S1. As salmochelin S4 is known to support growth of *C. jejuni*, this suggests there is another yet uncharacterised system capable of iron(III)-salmochelin import.¹⁸²

Interestingly, VctP bound the tetradentate salmochelin mimics with high affinity, indicating that the appended glucose unit does not cause significant steric clashes within the binding cleft, suggesting that the VctCDGP system may be capable of cytoplasmic uptake of iron(III)-salmochelin S1. This is the first indication that *V. cholerae* may be able to take advantage of salmochelin stealth siderophores. Salmochelins have been proven not to bind to siderocalin, as detailed in Chapter 1, and are synthesised and exported by *E. coli* for this purpose, to provide a method of iron uptake that evades the immune response of the host.^{218, 236} As the main siderophores of *V. cholerae*, enterobactin and vibriobactin, are both sequestered by siderocalin,^{84, 220, 223} the potential use of salmochelin S1 by *V. cholerae* may be key to virulence of the species.^{158, 225}

5.7 Summary and Conclusion

E. coli synthesises and exports enterobactin, and then imports the intact cyclic iron(III)-bound enterobactin *via* the Fep system, where it is bound with nanomolar affinity by FepB, in a Λ -configured binding mode. FepB is unable to bind the tetradentate iron(III)-bisDHBS linear hydrolysis product of enterobactin with high affinity, and has similar weak affinities for tetradentate siderophore mimics of both iron(III)-bisDHBS and iron(III)-salmochelin S1. This suggests that the Fep system is optimised for uptake of hexadentate iron(III)-enterobactin rather than the linear hydrolysis products. The linear hydrolysis products are likely to be taken up by a separate set of uptake proteins, starting with the outer membrane receptors Fiu and Cir.^{152, 160}

C. jejuni appears to utilise enterobactin in a different way to *E. coli*. Instead of taking up the full hexadentate iron(III)-enterobactin, into the cytoplasm, CeuE is much better optimised for the binding of iron(III)-n-LICAM (n= 4, 5) and iron(III)-bisDHBS.⁹⁰ It is likely that iron(III)-bisDHBS is acquired by *C. jejuni* from its environment, or iron(III)-enterobactin is first hydrolysed in the periplasm by the esterase Cee.¹⁵⁹ This strategy can be rationalised by the fact that *C. jejuni* does not produce its own source of enterobactin, but scavenges from its environment. To maximise iron(III)-uptake from enterobactin sources, it is most efficient to be able to process the hydrolysis products as well as the full trilactone. Interestingly, CeuE is also able to bind the hexadentate siderophore mimic iron(III)-MECAM with low nanomolar affinity. This suggests that MECAM is not an adequate mimic of enterobactin, as the interaction of the two species with CeuE must be significantly different to produce such different dissociation constants. This raises a question as to whether CeuE may be able to bind other hexadentate catecholate siderophores or siderophore mimics with high affinity, and whether this finding is of biological relevance. Results from salmochelin mimic studies suggests that CeuE may not be optimised for salmochelin S1 binding.

V. cholerae appears to be able to use enterobactin in both hexadentate and tetradentate forms. Since *V. cholerae* does not produce but scavenges enterobactin, and VctP contains the conserved histidine and tyrosine residues present in CeuE for iron(III)-chelation of tetradentate siderophore complexes, it is likely that VctP and CeuE are both optimised for tetradentate siderophore binding in a similar manner. Studies with salmochelin mimics make the first suggestion that the VctPDGC uptake system may be involved in the uptake of iron(III)-salmochelin S1. Such an adaptable iron(III)-siderophore binding periplasmic binding protein as VctP is likely a factor in the success of *V. cholerae* as a particularly persistent pathogen.

5.8 Future Work

To further the salmochelin studies, the binding affinities of each periplasmic binding protein with the natural salmochelin S1 and salmochelin S4 should be measured. FepB may bind to iron(III)-salmochelin S4, since the glucosylated siderophore is known to be synthesised by *E. coli*, and although the salmochelins are known to support growth, there is currently no known dedicated cytoplasmic uptake system for salmochelin S4.^{138, 232, 235} It is necessary to establish whether the Fep system is responsible for uptake of iron(III)-salmochelin S4, whether there is an uncharacterised alternative set of uptake proteins, or whether salmochelin uptake first relies on periplasmic hydrolysis by IroE before cytoplasmic uptake of iron(III)-salmochelin hydrolysis products. It would then be important to establish the binding affinity of iron(III)-salmochelin S4 with CeuE to determine whether the binding affinity is weak like that of iron(III)-enterobactin, or whether binding is tighter, similar to that of iron(III)-MECAM. Extending the study to VctP would reveal whether the protein can bind both hexadentate and tetradentate salmochelin complexes, equivalent to its binding of iron(III)-enterobactin and iron(III)-bisDHBS. Finally, studying each periplasmic binding protein with Salmochelin S1 would prove whether Sal-4-LICAM and Sal-5-LICAM make good mimics of this siderophore. If it again proved impossible to source salmochelin S4, then a synthetic mimic could be designed and synthesised. The same chemistry as employed in Chapter 4 could be used for the installation of the necessary C5 β -D-glucose unit onto a catechol unit, and then amide bonds to a scaffold based on that of MECAM could be used to produce a hexadentate salmochelin S4 mimic.

Growth studies with the tightest binding siderophores would establish whether the binding of iron(III)-siderophores to the periplasmic binding proteins are biologically relevant. It would be useful to establish whether the species could use the synthetic siderophore mimics as effective sources of iron(III) for uptake and growth. Since the dissociation constants of the tetradentate siderophores are in the picomolar range for VctP, this interaction may be too strong for effective delivery of iron(III) into the cytoplasm for release. If these siderophores instead block the periplasmic binding protein from being able to transport iron(III), this would be of interest for development of antimicrobial agents.

Further efforts are required to co-crystallise iron(III)-siderophores and mimics with each periplasmic binding protein. As the binding mode of iron(III)-siderophores with VctP is currently unknown, with only an apo structure of the protein known (PDB ID: 3TEF), these structures would be of greatest interest. The structural explanation as to how VctP binds tetradentate n-LICAM and bisDHBS siderophores, as well as how the protein can

accommodate a C5 appended glucose unit in an adjacent binding cleft, would be particularly informative. Comparison of these structures with that of VctP bound to iron(III)-enterobactin, would establish the differences in binding modes, and suggest structural rearrangements of the protein that might be required to switch between tetradentate and hexadentate binding modes. These binding modes could then be compared to those of FepB and CeuE. It would also be important to co-crystallise FepB with MECAM to establish whether it binds with a similar binding mode to iron(III)-enterobactin. Those iron(III)-siderophore complexes that proved to have weak interactions with the associated periplasmic binding protein, such as FepB with the tetradentates, or CeuE with iron(III)-enterobactin would also be of interest to study *via* crystallography, however the weaker binding affinities may continue to prove challenging for obtaining the necessary co-crystals.

The range of siderophores studied could be expanded. For example, it may be of interest to study other tetradentate catecholate siderophore such as serratiochelin or the amonabactins, to establish whether *C. jejuni* and/or *V. cholerae* are capable of poaching siderophores from a wider range of bacterial siderophore exporter species. The hexadentate linear enterobactin hydrolysis product trisDHBS would be a key siderophore to study, as this would allow the determination of whether the linear or cyclic nature of the siderophore or the denticity is key to recognition by each periplasmic binding protein. In addition, it would be interesting to establish the role of vibriobactin in *V. cholerae* by assessing the binding of iron(III)-vibriobactin with VctP, as well as studying ViuP with enterobactin and the hydrolysis products. A study of whether *C. jejuni* can poach vibriobactin would require investigation if CeuE can bind iron(III)-vibriobactin. As vibriobactin has structural similarities to the hexadentate siderophore mimic MECAM, it could be that CeuE binds this hexadentate siderophore with high affinity. In addition, bidentate analogues such as DHBS could be studied to establish their role in bacterial iron(III) chelation and uptake.

The project could be expanded to include more periplasmic binding proteins. Those identified by sequence alignment in Chapter 3 from pathogenic or thermophilic species could be expressed and purified. Studying many periplasmic binding proteins from a wide range of pathogenic species may be useful for a greater understanding of pathogen virulence or for the design of antibiotics. Thermophilic periplasmic binding proteins may be useful for the development of artificial metalloenzymes, and it is important to first establish the types of siderophores that each periplasmic binding protein scaffold can accommodate before designing the attached catalyst.

It is known that human and other mammalian siderocalins sequesters iron(III)-enterobactin, as explained in Chapter 1.^{219, 220} It is not known whether the linear hydrolysis products are also sequestered by siderocalin, however if their affinity with siderocalin is weaker than that of iron(III)-enterobactin, then they could prove to be a more effective source of iron(III), if iron(III)-enterobactin is taken up readily by the host immunoproteins. To test this idea, siderocalin fluorescence quenching binding assays could be performed to assess the dissociation constants of the tetradentate siderophores and siderophore mimics iron(III)-bisDHBS, and iron(III)-n-LICAM. This study could then be extended to the linear salmochelin mimics iron(III)-Sal-n-LICAM.

6 Experimental

6.1 General Remarks

Materials

Materials were obtained from the following commercial suppliers: Acros Organics, Alfa-Aesar, Cambridge Isotope Laboratories Inc., Fluorochem, Fisher Scientific, Sigma-Aldrich, Tokyo Chemical Industry UK Ltd. Natural siderophores including enterobactin and salmochelins were obtained from EMC Microcollections. H₄-bisDHBS was obtained from Dr. D. J. Raines. H₄-n-LICAM (n= 4, 5, 6, 8) compounds were obtained from Dr. D. J. Raines and Prof. A.-K. Duhme-Klair. CeuE was obtained from Dr. E. V. Blagova, Dr D. J. Raines and Mr. A. Hughes. CeuE variant proteins and VctP were obtained from Dr. E. V. Blagova. All solvents and reagents were used as supplied, unless otherwise stated.

Instrumentation

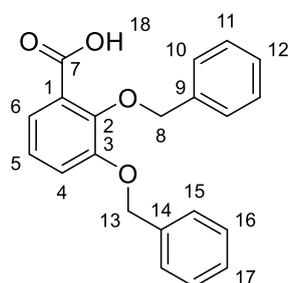
¹H, COSY, ¹³C, DEPT 135, HMQC, and HMBC NMR spectra were recorded on Jeol EX and ES 400 MHz instruments. ¹H experiments were run at 399.78 MHz. ¹³C experiments were run at 100.53 MHz and were proton decoupled. Data are reported as follows: chemical shift (multiplicity, proton environment assignment, coupling constants, number of protons). Chemical shifts are reported relative to residual solvent peaks. Chemical shifts were measured in ppm and are quoted to the nearest 0.01 ppm. Multiplicity is reported as the following: s = single, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. Coupling constants (*J*) are given in Hz, and are quoted to the nearest 0.1 Hz. All ¹H and ¹³C assignments were confirmed with COSY, DEPT 135 and HMQC experiments, and supplemented with HMBC experiments as required. All NMR spectra were processed with ACD/NMR Processor Academic Edition.

Positive and negative ion electrospray ionisation mass spectrometry (ESI-MS) was performed on a Bruker microTOF electrospray mass spectrometer by Mr. K. Heaton, Ms. H. Robinson and Ms. R. Cercola. Elemental analysis was performed on an Exeter CE-440 elemental analyser by Mr. G. McAllister and results are reported within $\pm 0.4\%$. Infrared (ATIR) spectra were recorded on a Perkin Elmer FT-IR spectrum two spectrophotometer at ambient temperature. Melting points were determined using a Stuart Scientific SMP3 melting point apparatus. Specific rotation was recorded on a Jasco DIP-370 digital polarimeter. Fluorescence spectra were recorded on a Hitachi F-4500 fluorescence spectrophotometer at ambient temperature. Electronic absorption spectra were recorded on a Shimadzu UV-1800 spectrophotometer at ambient temperature. Circular dichroism was

performed on a Jasco J810 CD spectropolarimeter at 20 °C under a constant flow of nitrogen. Analytical thin layer chromatography was performed using Merck silica gel 60 F253 aluminium-backed plates using the specified solvent system and visualised under an ultraviolet lamp. Flash column chromatography was carried out using Fluka Silica, pore size 60 Å, 220-440 mesh, 35-75 µm. Analytical HPLC was performed using a Shimadzu Prominence LC 20AD setup, with C18 column and SPD M20A diode array detector under the direction of Ms. A. Dixon.

6.2 Synthesis

2,3-Bis(benzyloxy)benzoic acid (Compound B)



Molecular Formula: C₂₁H₁₈O₄

Molecular Mass: 334.37 g mol⁻¹

2,3-Bis(benzyloxy)benzoic acid was prepared based on a preparation from the literature.¹⁵

2,3-Dihydroxybenzaldehyde (4.058 g, 29.30 mmol) was dissolved in 90 mL dry ethanol, to which benzyl chloride (8.33 mL, 72.40 mmol) and potassium carbonate (10.135 g, 72.40 mmol) were added. The reaction mixture was refluxed at 82 °C under anhydrous conditions for 18 hours. The reaction mixture was cooled and the solid filtered and washed with water. This solid was dissolved in 150 mL ethyl acetate and washed with 2 portions of water (50 mL) one of brine (30 mL) and one of 3:2 water brine mixture (50 mL). The organic layer was dried over magnesium sulfate and volatiles removed *in vacuo*. Without further purification, the crude aldehyde was dissolved in 40 mL acetone and 40 mL water. Sodium chlorite (3.169 g, 35.1 mmol) and sulfamic acid (4.229 g, 42.80 mmol) were each dissolved in 30 mL water and added alternately over 30 minutes. The mixture was stirred in air for one hour before removal of acetone *in vacuo*. The remaining solid in water was cooled and filtered. The resulting solid was recrystallised in a minimum hot ethanol. Filtration and vacuum desiccation for 48 hours resulted in a yellow crystalline solid (5.521 g, 16.50 mmol, 56%) R_f = 0.41 (4:1 chloroform: ethyl acetate) M.P = 122.7-122.9 °C.

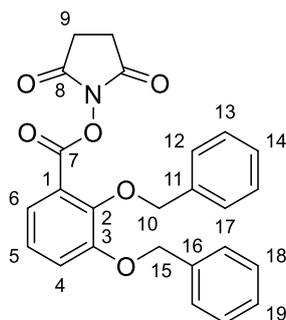
¹H NMR: (400 MHz, DMSO-d₆) δ: 12.95 (br s, *H*18, 1H); 7.51 (d, *J*= 8.0 Hz, *H*6, 1H); 7.23-7.29 (m, *H*10-12, *H*15-17, 10H); 7.26 (d, *H*4, *J*= 8.0 Hz, 1H); 7.14 (dd, *H*5, *J*= 8.0 Hz, 1H); 5.19 (s, *H*8/13, 2H); 5.03 (s, *H*8/13, 2H).

¹³C NMR: (100 MHz, DMSO-d₆) δ: 168.17 (*C*7); 152.9 (*C*1) 147.22, 138.01 (*C*4,6); 137.28, 128.93, 128.67, 128.26, 124.66, 122.04, 117.50 (*C*1-3, *C*5, *C*9-12, *C*14-17); 74.92, 70.37 (*C*8,9).

HRMS: Calcd. [M+H]⁺ (C₂₁H₁₉O₄) *m/z* = 335.1278; Obs. [M+H]⁺ *m/z* = 335.1271, Mean err 1.9ppm. Calcd. [M+Na]⁺ (C₂₁H₁₈NaO₄) *m/z* = 35.1097; Obs. [M+H]⁺ *m/z* = 357.1101, Mean err -0.9ppm.

IR ATIR (cm⁻¹): 3031 w br (O-H), 2873 w br (C-H), 1683 s (C=O), 1576 m (C=C ar).

2,3-Bis(benzyloxy)benzoic acid N-hydroxysuccinimide ester (Compound R)



Molecular formula: C₂₅H₂₁O₆

Molecular weight: 431.44 g mol⁻¹

2,3-Bis(benzyloxy)benzoic acid N-hydroxysuccinimide ester was prepared based on a preparation from the literature.⁴⁰⁵

2,3-Bis(benzyloxy)benzoic acid (2.790 g, 8.35 mmol) and N-hydroxysuccinimide (1.026 g, 8.91 mmol) were dissolved in dioxane (20 mL). DCC (2.066 g, 10.00 mmol) was added and stirred at 0 °C for 5 hours. The reaction mixture was left to stand overnight, and the resulting precipitate removed by filtration. The remaining filtrate was reduced *in vacuo* to yield an oily residue in a minimum dioxane, to which 3 mL isopropyl alcohol was added. A solid was obtained after one hour, which was isolated by filtration and washed with isopropyl alcohol (5 mL) and water (5 mL). The solid was recrystallised from a minimum amount of hot ethyl acetate. The resulting precipitate was isolated and dried under vacuum for 12 hours yielding a yellow crystalline solid (2.342 g, 5.43 mmol, 65%) R_f = 0.65 (4:1 chloroform: ethyl acetate) M.P = 110-112 °C.

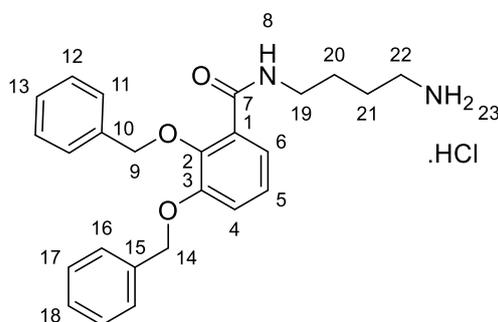
^1H NMR: (400 MHz, DMSO- d_6) δ : 7.62-7.28 (m, H_{4-6} , H_{12-14} , H_{17-19} , 13H); 5.26 (s, $H_{10/15}$, 2H); 5.05 (s, $H_{10/15}$, 2H); 2.89 (br s, H_9 , 4H).

^{13}C NMR: (100 MHz, DMSO- d_6) δ : 171.06 (C_7); 161.52 (C_8); 153.19 (C_1); 148.63 (C_9); 137.20, 136.89 (C_2 , C_3); 128.98, 128.92, 128.63, 128.60, 125.43, 122.61, 121.01, 120.48 (C_4-6 , C_{11-14} , C_{16-19}); 75.31, 70.60 (C_{10} , C_{15}).

HRMS: Calcd. $[\text{M}+\text{Na}]^+$ ($\text{C}_{25}\text{H}_{21}\text{NNaO}_6$) $m/z = 454.1261$; Obs. $[\text{M}+\text{H}]^+$ $m/z = 454.1248$, Mean err 3.3ppm.

IR ATIR (cm^{-1}): 2936 m br (C-H), 1764 m (C=O), 1732 s (C=O), 1578 m (N-O).

1-Amino, 4-(2,3-dibenzyloxybenzamide) butane (Compound C)



Molecular formula: $\text{C}_{25}\text{H}_{29}\text{ClN}_2\text{O}_3$

Molecular weight: $440.97 \text{ g mol}^{-1}$

1-Amino, 4-(2,3-dibenzyloxybenzamide) butane was prepared based on a procedure from the literature.⁸⁵

2, 3-Bis(benzyloxy)benzoic acid (0.300 g, 0.90 mmol) was dissolved in dry THF (5 mL) and carbonyldiimidazole (0.194 g, 0.90 mmol) was added. 1,4-diaminobutane (0.210 g, 1.79 mmol) was dissolved in dry THF (5 mL) and the solution was stirred vigorously. The bis(benzyloxy)benzoic acid solution was then added dropwise to the 1,4-diaminobutane solution over 60 minutes. The resulting reaction mixture was stirred overnight at room temperature. The solvent was removed *in vacuo* and the residue taken up in chloroform (30 mL), washed with saturated NaHCO_3 (15 mL), brine (15 mL), 1M HCl (15 mL), and brine (15 mL). The organic portion was then dried over MgSO_4 and partially reduced *in vacuo*. Ethyl acetate (10 mL) was added and the resulting solution was concentrated *in vacuo* to yield a white solid. The solid was recrystallised from a minimum hot ethyl acetate: chloroform 2:1 to yield a white solid (0.254 g, 0.58 mmol, 72%) $R_f = 0.1$ (1:1 chloroform: methanol) M.P = 129-131 °C.

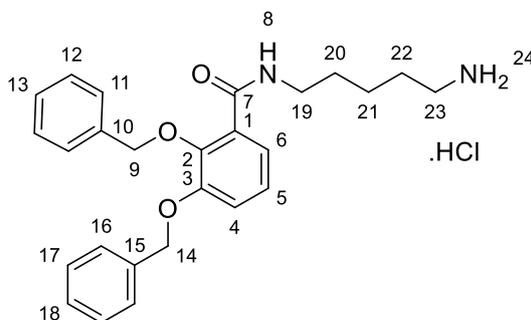
¹H NMR: (400 MHz, MeOD-d₄) δ: 7.50 (d, *J* = 7.8 Hz, *H*_{4/6}, 2H); 7.39-7.29 (m, *H*₁₁₋₁₃, *H*₁₆₋₁₈, 10H); 7.16 (t, *J* = 7.8 Hz, *H*₅, 1H); 5.20 (s, *H*_{9/14}, 2H); 5.11 (s, *H*_{9/14}, 2H); 3.28 (m, *H*₁₉, 2H); 2.86 (t, *J* = 7.6 Hz, *H*₂₂, 2H); 1.60 (m, *H*_{20/21}, 2H); 1.49 (m, *H*_{20/21}, 2H).

¹³C NMR: (100 MHz, MeOD-d₄) δ: 153.56 (*C*₇); 138.48, 138.30, 130.23, 129.97, 129.79, 129.70, 129.64, 129.15, 125.69, 122.71, 118.28 (*C*₁₋₆, *C*₁₀₋₁₃, *C*₁₅₋₁₈); 77.32, 72.37 (*C*_{9,10}); 40.44 (*C*_{20/21}); 39.91(*C*_{20/21}); 27.47, 26.04 (*C*_{18,19}).

HRMS (ESI): Calcd. [M+H]⁺ (C₂₅H₂₉N₂O₃) *m/z* = 405.2173; Obs. [M+H]⁺ *m/z* = 405.2163, Mean err 3.6 ppm.

IR ATIR (cm⁻¹): 3365 m (N-H), 3075 m br (N-H), 2801 m br (C-H), 1641 m (C=O), 1571 m (C=C ar).

1-Amino, 5-(2,3-dibenzyloxybenzamide) pentane (Compound S)



Molecular formula: C₂₆H₃₇ClN₂O₃

Molecular weight: 455.00 g mol⁻¹

1-Amino, 5-(2,3-dibenzyloxybenzamide) pentane was prepared by student Ross P. Thomas.

1, 5-Diaminopentane (1.427 g, 14.00 mmol) and triethylamine (1.396 g, 13.80 mmol) were dissolved in THF (120 mL). A solution of 2, 3-bis(benzyloxy)benzoic acid N-hydroxysuccinimide ester (3.0076 g, 6.97 mmol) in THF (60 mL) was prepared and added dropwise over 2 hours. The mixture was left to stir overnight. The solvent was removed *in vacuo* yielding an off-white solid which was taken up in chloroform (120 mL) and washed with NaHCO₃ (100 mL), brine (100 mL) and 2.25:1 1M HCl: brine (130 mL). The organic layer was then dried over MgSO₄, filtered and the solvent removed *in vacuo* yielding a yellow oil. The oil was purified *via* silica column chromatography (90:10:1 CHCl₃: MeOH: NH₃(aq)). Diethyl ether was added to afford a white solid that was dried *in vacuo* (1.8592 g, 58.5 %) *R*_f = 0.18 (90:10:1 CHCl₃: MeOH: NH₃(aq)) M.P = 133-134 °C.

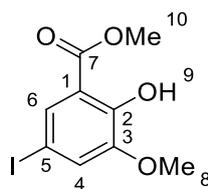
$^1\text{H NMR}$: (400 MHz, CDCl_3) δ : 7.99 (t, $J = 5.5$ Hz, H_8 , 1H); 7.73-7.71 (m, $H_4/6$, 1H); 7.49-7.35 (m, H_{11-13} , H_{16-18} , 10H); 7.16 (d, $J = 2.3$ Hz, $H_6/4$, 2H); 7.14 (t, $J = 4.6$ Hz, H_5 , 1H); 5.16 (s, $H_9/14$, 2H); 5.08 (s, $H_9/14$, 2H); 3.27 (dd, $J = 6.0$ Hz $J = 13.3$ Hz, H_{19} , 2H); 2.77 (t, $J = 6.6$ Hz, H_{23} , 2H); 1.48 (tt (app quint.), $J = 7.3$ Hz, H_{20} , 2H); 1.37-1.20 (m, H_{21} , H_{22} , 4H).

$^{13}\text{C NMR}$: (100 MHz, CDCl_3) δ : 164.98 (C_7); 151.66, 146.67, 136.35, 128.73, 128.64, 128.622, 127.66 (C_{1-3} , C_{10-13} , C_{15-18}); 124.43 ($C_4/6$); 123.26 ($C_4/6$); 116.74 (C_5); 76.31, 71.21 ($C_{9,14}$); 41.96 (C_{23}); 39.56 (C_{19}); 33.29, 29.08, (C_{20} , C_{22}); 24.23 (C_{21}).

HRMS (ESI): Calcd. $[\text{M}+\text{H}]^+$ ($\text{C}_{25}\text{H}_{29}\text{N}_2\text{O}_3$) $m/z = 405.2173$; Obs. $[\text{M}+\text{H}]^+$ $m/z = 405.2168$, Mean err 3.5 ppm.

IR ATIR (cm^{-1}): 3365 m (N-H), 3078 m br (N-H), 2803 m br (C-H), 1641 m (C=O), 1571 m (C=C ar).

Methyl-5-iodo-3-methoxysalicylate (Compound E)



Molecular formula: $\text{C}_9\text{H}_9\text{IO}_4$

Molecular weight: $308.07 \text{ g mol}^{-1}$

Methyl-5-iodo-3-methoxysalicylate was prepared based on a preparation from the literature.^{438,401}

AgNO_3 (2.774 g, 16 mmol) was dissolved in chloroform (30 mL) and pyridine (20 mL) was added resulting in a clear yellow solution. ICl (2.673 g, 16 mmol) was dissolved in chloroform (25 mL) and added dropwise to the reaction mixture over 35 minutes. The combined pale green grey solution was stirred for 45 minutes. Methyl-3-methoxysalicylate (2.000 g, 11 mmol) was dissolved in chloroform (30 mL) and added dropwise to the reaction mixture over 30 minutes. The resulting reaction mixture was stirred overnight at room temperature. The reaction mixture was diluted with diethyl ether (100 mL) and the silver salts removed by filtration. The filter cake was washed with 1:1 chloroform: diethyl ether mixture (200 mL). The organic portions were combined and the solvent reduced *in vacuo* to yield a brown residue. The residue was dissolved in chloroform (100 mL) and washed twice with HCl (0.1 M, 100 mL), once with HCl (1M, 100 mL) resulting in a neutral aqueous

portion. The organic portion was then washed twice with saturated $\text{Na}_2\text{S}_2\text{O}_3$ (100 mL) and H_2O (100 mL). The organic portion was then dried over MgSO_4 , filtered and the solvent removed *in vacuo*, yielding a brown solid. The solid was recrystallised from a minimum amount of hot ethanol. The resulting orange crystals were collected by filtration. Cold H_2O was added to the filtrate resulting in the formation of a yellow precipitate. The yellow powder was collected by filtration. Both the crystals and powder were dried *in vacuo* and combined. (2.379 g, 7.72 mmol, 73%) $R_f = 0.5$ (1:4 ethyl acetate : petroleum ether 40-60°C) M.P = 109-111 °C.

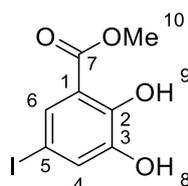
^1H NMR: (400 MHz, CDCl_3) δ : 10.99 (s, H9, 1H); 7.76 (d, $J = 2.0$ Hz, H4/6, 1H); 7.25 (d, $J = 2.0$ Hz, H4/6, 1H); 3.96 (s, H10, 3H); 3.89 (s, H8, 3H).

^{13}C NMR: (100 MHz, CDCl_3) δ : 169.57 (C7); 151.97, 149.32, 114.15 (C1, 2, 3); 129.46, 124.86 (C4,6); 56.37 (C8); 52.66 (C10).

HRMS (ESI): Calcd. $[\text{M}+\text{H}]^+$ ($\text{C}_9\text{H}_9\text{IO}_4$) $m/z = 308.9618$; Obs. $[\text{M}+\text{H}]^+$ $m/z = 308.9629$, Mean err 3.7 ppm. Calcd. $[\text{M}+\text{Na}]^+$ ($\text{C}_9\text{H}_9\text{INaO}_4$) $m/z = 330.9438$; Obs. $[\text{M}+\text{Na}]^+$ $m/z = 330.9446$, Mean err 2.9 ppm.

IR ATIR (cm^{-1}): 3090 br w (C-H), 2943 w (C-H), 1674 m (C=O), 1607 m (C=C ar), 1471 s (C-H), 1343 m (O-H).

Methyl-5-iodo-3-hydroxysalicylate (Compound F)



Molecular formula: $\text{C}_8\text{H}_7\text{IO}_4$

Molecular weight: $294.04 \text{ g mol}^{-1}$

Methyl-5-iodo-3-hydroxysalicylate was prepared based on a preparation from the literature.²⁶⁸

Methyl-5-iodo-3-methoxysalicylate (2.000 g, 6.49 mmol) was dissolved in anhydrous dichloromethane (10 mL) and the mixture was stirred whilst the reaction flask was purged with N_2 . BBr_3 (1M in dichloromethane, 12mL, 12.00 mmol) was added dropwise with vigorous stirring. Resultant gas evolution was passed through a trap containing aqueous NaOH (1M). The reaction flask was then purged again with N_2 . The mixture was stirred overnight at room temperature. The reaction mixture was opened to air, cold H_2O (20 mL)

was carefully added and the reaction mixture stirred for 1 hour. The resulting pale pink turbid mixture was dissolved in methanol (20 mL) and the solvent removed *in vacuo* yielding a red residue. Methanol (10 mL \times 3) was added to the residue, the residue dissolved, and the solvent removed *in vacuo*. The resulting peach coloured solid was redissolved in methanol (50 mL) and concentrated H₂SO₄ (2 mL) and heated under reflux overnight. The reaction mixture was cooled to room temperature and the solvent removed *in vacuo*, yielding a colourless oil and a white solid. The residue was dissolved in ethyl acetate (150 mL) and was washed three times with saturated NaHCO₃ (90 mL) and twice with brine (90 mL). The organic portion was dried over MgSO₄, filtered and the solvent removed *in vacuo* yielding a white solid (1.710 g, 5.82 mmol, 89%) R_f = 0.47 (1:4 ethyl acetate : petroleum ether 40-60°C) M.P = 136-138 °C (Lit. 133-134 °C).³⁹⁶

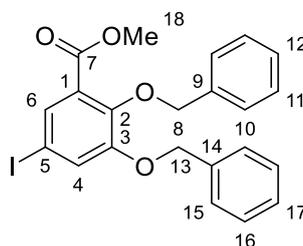
¹H NMR: (400 MHz, CDCl₃) δ : 10.89 (s, H₉, 1H); 7.69 (d, *J* = 1.6 Hz, H₄/6, 1H); 7.40 (d, *J* = 1.6 Hz, H₄/6, 1H); 5.70 (br s, H₈, 1H); 3.97 (s, H₁₀, 3H).

¹³C NMR: (100 MHz, CDCl₃) δ : 169.50 (C₇); 148.79, 145.91, 113.99 (C₁, 2, 3); 129.07, 128.25 (C₄, 6); 52.74 (C₁₀).

HRMS (ESI): Calcd. [M+Na]⁺ (C₉H₉INaO₄) *m/z* = 316.9281; Obs. [M+Na]⁺ *m/z* = 316.9282, Mean err 0.5 ppm. Calcd. [M+2Na]⁺ (C₉H₉INa₂O₄) *m/z* = 338.9101; Obs. [M+2Na]⁺ *m/z* = 338.9099, Mean err 0.2 ppm.

IR ATIR (cm⁻¹): 3450 s (O-H), 3087 w (C-H), 2956 w (C-H), 1665 s (C=O), 1595 m (C-H), 1464 s (C-H), 1358 m (O-H).

Methyl-5-iodo-3,4-benzyloxysalicylate (Compound G)



Molecular formula: C₂₂H₁₉IO₄

Molecular weight: 474.29 g mol⁻¹

Methyl-5-iodo-3,4-benzyloxysalicylate was prepared based on standard preparations from the literature.⁴³⁹

Methyl-5-iodo-3-hydroxysalicylate (1.651 g, 5.61 mmol) was dissolved in anhydrous dimethylformamide (30 mL) and K₂CO₃ (3.413 g, 24.70 mmol) and NaI (0.059 g, 0.39

mmol) were added, forming a suspension. Benzyl bromide (6.140 g, 4.30 mL, 35.91 mmol) was added and the mixture was stirred at room temperature overnight. The solvent was then removed *in vacuo* yielding a sandy brown residue. The residue was dissolved in ethyl acetate (100 mL) and H₂O (60 mL) and the layers separated. The aqueous layer was washed twice with ethyl acetate (100 mL) and the organic layers combined. The organic portion was dried over MgSO₄, filtered and the solvent removed *in vacuo*. The residue was dissolved in 1:1 chloroform : ethyl acetate (10 mL) and purified by flash column chromatography (1:9 ethyl acetate : petroleum ether 40-60°C) followed by recrystallisation in a minimum amount of chloroform and an excess of petroleum ether 40-60°C. The white needle crystals were collected and dried *in vacuo* (1.613 g, 3.40 mmol, 61%) R_f = 0.37 (1:9 ethyl acetate : petroleum ether 40-60°C) M.P = 103-105 °C (Lit. 105-106 °C).³⁹⁶

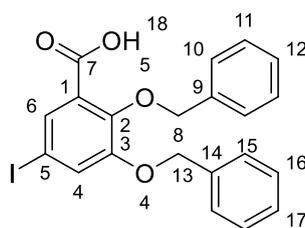
¹H NMR: (400 MHz, CDCl₃) δ: 7.69 (d, *J* = 2.0 Hz, H₄/6, 1H); 7.45-7.31 (m, H₄/6, H₁₀-12, H₁₅-17, 11H); 5.11 (s, H₈/13, 2H); 5.07 (s, H₈/13, 2H); 3.84 (s, H₁₈, 3H).

¹³C NMR: (100 MHz, CDCl₃) δ: 165.11 (C₇); 153.43, 148.14, 136.85 135.76, (C₂, 3, 9, 14); 131.43 (C₄/6); 128.58, 128.55, 128.30, 128.26, 128.22, 128.02, 127.56 126.44(C₄/6, C₁₀-12, C₁₅-17); 86.23 (C₁); 75.64, 71.44, (C₈, 13); 52.29 (C₁₈).

HRMS (ESI): [M+K]⁺ (C₉H₉INKO₄) *m/z* = 512.9960; Obs. [M+K]⁺ *m/z* = 512.9957, Mean err 0.1 ppm.

IR ATIR (cm⁻¹): 3028 w (C-H), 1716 s (C=O), 1564 m (C-H), 1497 s (C-H).

5-Iodo-3,4-benzyloxysalicylate (Compound H)



Molecular formula: C₂₁H₁₇IO₄

Molecular weight: 460.2675 g mol⁻¹

The synthesis of 5-iodo-3,4-benzyloxysalicylate was designed based on a preparation from the literature.^{396, 397}

Methyl-5-iodo-3-hydroxysalicylate (0.130 g, 0.27 mmol) was dissolved in THF (1.5 mL) and methanol (0.5 mL) and NaOH (5 M, 1 mL) was added. The reaction mixture was stirred vigorously overnight. The reaction mixture was acidified to pH 1 with HCl (1 M) and the

resulting white precipitate was collected by filtration and dried *in vacuo* (0.108 g, 0.23 mmol, 86%) $R_f = 0.0$ (1:9 ethyl acetate : petroleum ether 40-60°C) M.P = 168-180 °C.

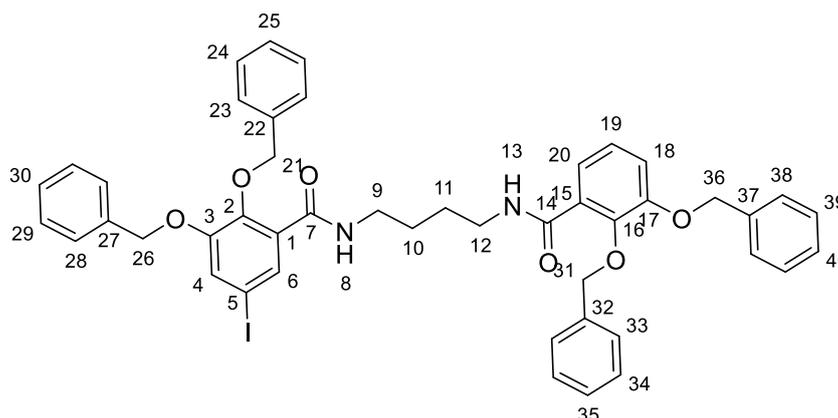
$^1\text{H NMR}$: (400 MHz, MeOD- d_4) δ : 7.64 (d, $J = 2.2$ Hz, $H_{4/6}$, 1H); 7.58 (d, $J = 1.8$ Hz, $H_{4/6}$, 1H); 7.49-7.25 (m, H_{10-12} , H_{15-17} , 10H); 5.14 (s, $H_{8/13}$, 2H); 5.04 (s, $H_{8/13}$, 2H).

$^{13}\text{C NMR}$: (100 MHz, MeOD- d_4) δ : 154.98 (C_7); 138.56, 137.84 ($C_{2/3}$) 132.31, 132.28, 130.01, 129.79, 129.49, 129.37, 129.27, 129.24 (C_1 , C_4 , C_5 , C_6 , C_{9-12} , 14-17); 76.83 ($C_{8/13}$); 72.62 ($C_{8/13}$).

HRMS (ESI): Calcd. $[\text{M-H}]^-$ ($\text{C}_{21}\text{H}_{16}\text{IO}_4$) $m/z = 459.0099$; Obs. $[\text{M-H}]^-$ $m/z = 459.0078$, Mean err 4.9 ppm.

IR ATIR (cm^{-1}): 3102 w (O-H), 2884 w br (C-H), 1688 s (C=O), 1562 m (C-H), 1497 m (C-H), 1375 m (O-H).

5-Iodo-bis(2,3benzyloxy)-4-LICAM (Compound J)



Molecular formula: $\text{C}_{46}\text{H}_{43}\text{IN}_2\text{O}_6$

Molecular weight: $846.76 \text{ g mol}^{-1}$

The synthesis of 5-iodo-bis(2,3benzyloxy)-4-LICAM was designed based on a preparation from the literature.⁹⁴

5-Iodo-3-hydroxysalicylate (0.050 g, 0.11 mmol) was dissolved in anhydrous THF (1.0 mL) and CDI (0.021g, 0.13 mmol) was added. The reaction mixture was stirred under dry conditions for 3 hours. 1-Amino, 4-(2, 3-dibenzyloxybenzamide)butane (0.048 g, 1.09 mmol) was dissolved in anhydrous THF (1 mL). The 5-iodo-3-hydroxysalicylate was added dropwise to the 1-amino, 4-(2, 3-dibenzyloxybenzamide)butane solution and the resulting reaction mixture was stirred under dry conditions overnight. The reaction mixture was reduced *in vacuo* and the residue taken up in chloroform (10 mL). The organic layer was washed with HCl (1 M, 10 mL), HCl (1 M) : brine 1:1 mixture (10 mL) and brine (2×10

mL). The organic layer was dried over MgSO_4 , filtered and solvent removed *in vacuo* to yield a white solid (0.076 g, 0.09 mmol, 82%) $R_f = 0.84$ (1:4 methanol : chloroform) M.P = 98-100°C.

$^1\text{H NMR}$: (400 MHz, MeOD-d_4) δ : 8.07 (d, $J = 1.8$ Hz, $H_{4/6}$, 1H); 7.89 (t, $J = 5.5$ Hz $H_{8/13}$, 1H); 7.76-7.74 (m, $H_{8/13}$, $H_{4/6}$, 2H) 7.49-7.29 (m, $H_{18/20}$, H_{23-25} , H_{28-30} , H_{33-35} , H_{38-40} , 21H); 7.16 (d, $J = 1.8$ Hz, $H_{20/18}$, 1H); 7.15 (t, $J = 4.1$ Hz, H_{19} , 1H); 5.17 (s, $H_{21/26/31/36}$, 2H); 5.12 (s, $H_{21/26/31/36}$, 2H); 5.06 (s, $H_{21/26/31/36}$, 2H); 5.01 (s, $H_{21/26/31/36}$, 2H); 3.19-3.14 (m, H_9 , H_{12} , 4H) 1.23-1.20 (m, H_{10} , H_{11} , 4H).

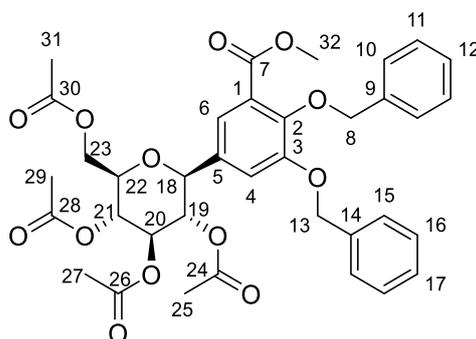
$^{13}\text{C NMR}$: (100 MHz, CDCl_3) δ : 164.95 ($C_{7/14}$); 163.42 ($C_{7/14}$); 152.30 ($C_{1/15}$); 151.64 ($C_{1/15}$); 146.74 (C_5), 136.36, 136.30, 135.86, 135.68 ($C_{2/3/16/17}$); 132.12 ($C_{4/6}$); 128.85, 128.73, 128.69, 128.67, 128.64, 128.48, 128.22, 127.77, 127.62, 127.23, (C_{22-25} , C_{27-30} , C_{32-35} , C_{37-40}); 125.48, 124.37 ($H_{20/18}$); 123.26 ($C_{4/6}$); 116.86 (C_{19}); 76.42, 76.32, 71.54, 71.22 ($C_{21/26/31/36}$); 39.34, 39.18 ($C_{9/12}$); 26.67, 26.61 ($C_{10/11}$).

HRMS (ESI): Calcd. $[\text{M}+\text{Na}]^+$ ($\text{C}_{46}\text{H}_{43}\text{IN}_2\text{NaO}_6$) $m/z = 869.2058$; Obs. $[\text{M}+\text{Na}]^+$ $m/z = 869.2075$, Mean err 1.9 ppm.

IR ATIR (cm^{-1}): 3413 w (N-H), 3287 w br (N-H), 2942 w (C-H), 1666 s (C=O), 1635 s (C=O), 1550 m (N-H), 1498 m (C-H).

Elemental Analysis: Calculated for $[\text{C}_{46}\text{H}_{43}\text{IN}_2\text{O}_6]$: %C 65.25, %H 5.12, %N 3.31; Measured for $[\text{C}_{46}\text{H}_{43}\text{IN}_2\text{O}_6]$: %C 65.09, %H 5.15, %N 3.28.

Methoxy-5-aceto- β -D-glucosyl-3, 4-benzyloxysalicylate (Compound M)



Molecular formula: $\text{C}_{36}\text{H}_{38}\text{O}_{13}$

Molecular weight: $678.69 \text{ g mol}^{-1}$

Methoxy-5-aceto- β -D-glucosyl-3, 4-benzyloxysalicylate was prepared based on a preparation from the literature.^{396, 397}

Acetobromo- α -D-glucose (0.750 g, 1.82 mmol) was purified by column chromatography (100% ethyl acetate) before use. Solvent was removed *in vacuo*, and petroleum ether 40:60 (3×10 mL) was added and removed *in vacuo* to afford a white solid. All glassware was flame dried *in vacuo* before use. Zinc (0.260 g, 2.97 mmol) and lithium chloride (0.112 g, 2.65 mmol) were weighed into a Schlenk tube in a glovebox, sealed under nitrogen, and heated *in vacuo* at 70°C for 90 minutes. Iodine (0.034 g, 1.34 mmol) and methyl-5-iodo-3,4-benzyloxysalicylate (1.153 g, 2.43 mmol) were weighed into a Schlenk tube, DMF (5.3 mL) added, and the resulting brown solution stirred under nitrogen for 15 minutes. The solution was then added to the zinc and lithium chloride, and the resulting mixture heated under nitrogen at 70°C for 90 minutes. In a glovebox, Ni(COD)₂ (0.043 g, 0.16 mmol) and 4,4',4''-tri-tert-Butyl-2,2':6',2''-terpyridine (0.094 g, 0.23 mmol) were weighed into a Schlenk tube, and DMF (3 mL) was added. The resulting solution was stirred for 15 minutes, which darkened from pale grey to deep blue. Acetobromo- α -D-glucose (0.640 g, 1.55 mmol) was added to the Ni(COD)₂ and 4,4',4''-tri-tert-Butyl-2,2':6',2''-terpyridine mixture and the resulting pale green turbid solution was stirred for 15 minutes. The methyl-5-iodo-3,4-benzyloxysalicylate solution was cooled, and then added to the acetobromo- α -D-glucose mixture. The resulting dark green solution was then stirred at room temperature under nitrogen overnight. The reaction mixture was opened to air, and the solvent removed *in vacuo*, yielding a brown residue. The residue was purified by column chromatography (1:1 ethyl acetate : petroleum ether 40-60°C) to afford an off-white solid product (0.701 g, 65%) $R_f = 0.38$ (1:1 ethyl acetate : petroleum ether 40-60°C) M.P = 120-121 °C (Lit. 128-129 °C).³⁹⁶

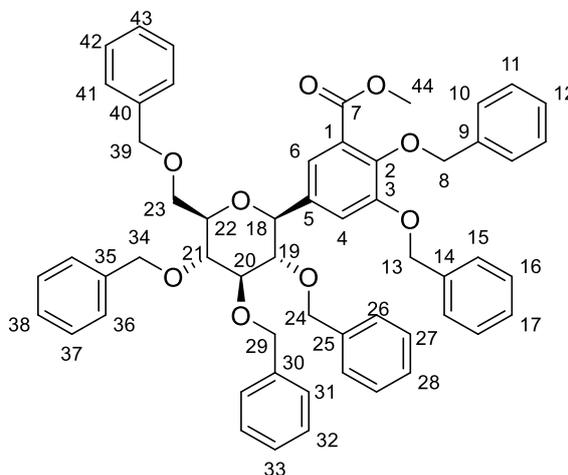
¹H NMR: (400 MHz, CDCl₃) δ : 7.48-7.30 (m, H_{4,8-17}, 11H); 7.23 (d, H₆, J = 1.8 Hz, 1H); 5.37 (t, H₂₀, J = 9.2 Hz, 1H); 5.24 (t, H₂₁, J = 10.1 Hz, 1H); 5.16-5.13 (m, H₁₉, H_{8/13}, 3H); 5.09 (d, H_{8/13}, J = 3.7 Hz, 2H); 4.39 (d, H₁₈, J = 10.1 Hz, 1H); 4.30 (dd, H₂₃, J = 12.4 Hz, J = 4.6 Hz, 1H); 4.18 (dd, H₂₃, J = 12.4 Hz, J = 2.3 Hz, 1H); 3.87-3.85 (m, H₂₂, 1H); 3.84 (s, H₃₂, 3H); 2.11 (s, H_{25/27/29/31}, 3H); 2.08 (s, H_{25/27/29/31}, 3H); 2.02 (s, H_{25/27/29/31}, 3H); 1.83 (s, H_{25/27/29/31}, 3H).

¹³C NMR: (100 MHz, CDCl₃) δ : 170.70, 170.31, 169.52, 168.94, (C₂₄, 26, 28, 30); 166.26 (C₇); 153.10 (C₃); 148.54 (C₂); 137.14, 136.26 (C₉, 14) 131.98 (C₅); 128.61, 128.24, 127.95, 127.65, 125.21 (C₁₀, 11, 12, 15, 16, 17); 126.209 (C₁); 122.16, 115.71 (C₄, 6); 79.48 (C₁₈); 76.21 (C₂₂); 75.56 (C_{8/13}); 74.07 (C₂₀); 72.23 (C₁₉); 71.32 (C_{8/13}); 68.50 (C₂₁); 62.35 (C₂₃); 52.19 (C₃₂); 20.80, 20.63, 20.37 (C₂₅, 27, 29, 32).

HRMS (ESI): Calcd. $[M+Na]^+$ ($C_{36}H_{38}NaO_{13}$) $m/z = 701.2205$; Obs. $[M+Na]^+$ $m/z = 701.2207$, Mean err 0.3 ppm.

IR ATIR (cm^{-1}): 2962 w (C-H), 1739 s (C=O), 1715 s (C=O), 1483 w (C-H).

Methoxy-5-benzyloxy- β -D-glucosyl-3,4-benzyloxysalicylate (Compound N)



Molecular formula: $C_{56}H_{54}O_9$

Molecular weight: $871.0390 \text{ g mol}^{-1}$

Methoxy-5-benzyloxy- β -D-glucosyl-3, 4-benzyloxysalicylate was prepared based on a preparation from the literature.^{396, 397}

Methoxy-5-aceto- β -D-glucosyl-3, 4-benzyloxysalicylate (1.314 g, 1.94 mmol) was dissolved in dry methanol (100 mL) and Na_2CO_3 (1.026 g, 9.68 mmol) was added. The resulting suspension was stirred under reflux at $65^\circ C$ overnight. When the reaction was deemed complete by thin layer chromatography ($R_f = 0$, 1:1 ethyl acetate : petroleum ether $40-60^\circ C$) and the presence of the intermediate confirmed by ESI-MS, the reaction mixture was cooled and filtered, and the resulting orange solution reduced *in vacuo* to a pale brown solid. The solid was transferred to a Schlenk tube, and NaH (533 mg, 60% in mineral oil, 13.3 mmol) was added. The solids were dried *in vacuo* for two hours, before the addition of DMF (25 mL) and cooling of the solution to $0^\circ C$. Bu_4NI was dried *in vacuo* for 15 minutes, to which, benzyl bromide (3.00 mL, 4.320 g, 25.26 mmol) and DMF (5 mL) were added. The solution was cooled to $0^\circ C$ for 5 minutes, and was then added dropwise to the methyl-5-aceto- β -D-glucosyl-3,4-benzyloxysalicylate solution over 5 minutes at $0^\circ C$. After 5 minutes of stirring at $0^\circ C$, the pale brown reaction mixture was stirred at RT under N_2 overnight. The reaction mixture was opened to air, and deionised water (40 mL) was carefully added. The resulting solution was extracted with ethyl acetate ($3 \times 50 \text{ mL}$). The organic portions were combined and dried over $MgSO_4$, filtered and solvent removed *in vacuo* to yield a white solid. The

solid was purified by column chromatography 20% ethyl acetate in petroleum ether 40-60°C yielding the product as a white solid (0.719 g, 0.82 mmol, 42%) $R_f = 0.20$ (20% ethyl acetate in petroleum ether 40-60°C) M.P = 84-86 °C.

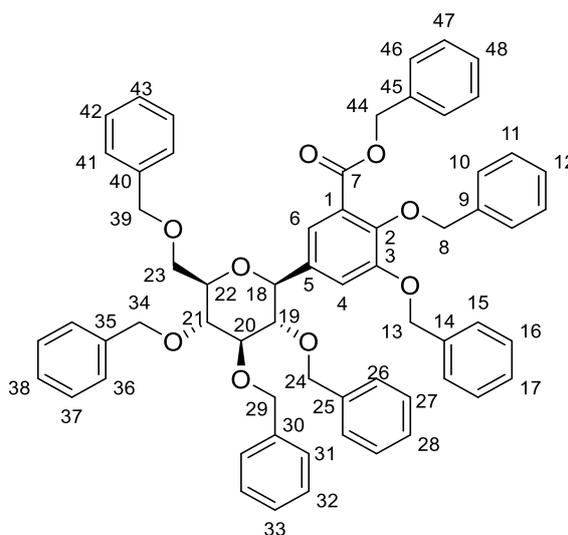
$^1\text{H NMR}$: (400 MHz, CDCl_3) δ : 7.53 (d, *H*6, $J = 1.8$ Hz, 1H) 7.49-7.30 (m, *H*10, *H*15, *H*26-28, *H*31-33, *H*36-38, *H*41-43, 24H); 7.23-7.21 (m, *H*4, *H*11, *H*16, 5H); 6.97-6.93 (m, *H*12, *H*17, 2H) 5.14 (d, *H*8, $J = 1.8$ Hz, 2H); 5.00 (s, *H*13, 2H); 4.98-4.87 (m, *H*24/29, *H*39, 3H) 4.65-4.57 (m, *H*23/*H*24/*H*29, 3H); 4.44 (d, *H*23, $J = 10.1$ Hz 1H); 4.23 (d, *H*18, $J = 9.6$ Hz, 1H); 3.87 (s, *H*44, 3H); 3.81-3.76 (m, *H*20/21/24/29, *H*34, 5H); 3.66-3.59 (m, *H*22, 1H); 3.45 (t, *H*19, $J = 9.2$ Hz, 1H).

$^{13}\text{C NMR}$: (100 MHz, CDCl_3) δ : 166.37 (*C*7); 152.50 (*C*3); 147.86 (*C*2); 138.53 (*C*25); 138.09, 137.98, 137.31, 137.16, 136.36, 135.03, 133.84 (*C*1, *C*9, *C*14, *C*30, *C*35, *C*40); 128.69, 128.45, 128.38, 128.32, 128.30, 128.26 128.20, 128.01, 127.96, 127.87, 127.73, 127.68, 127.61, 127.58, 127.51, 126.89 (*C*5, *C*10-12, *C*15-17, *C*26-28, *C*31-33, *C*36-38, *C*41-43); 121.85 (*C*6); 116.62 (*C*4); 86.56 (*C*20/21); 83.85 (*C*19); 80.80 (*C*18); 79.17 (*C*22); 78.10 (*C*20/21); 75.60 (*C*9); 75.60 (*C*24/29); 73.36 (*C*23); 70.93 (*C*13/39); 10.88 (*C*13/39); 68.87 (*C*34); 65.12 (*C*24/29); 52.05 (*C*44).

HRMS (ESI): Calcd. $[\text{M}+\text{Na}]^+$ ($\text{C}_{56}\text{H}_{54}\text{NaO}_9$) $m/z = 893.3660$; Obs. $[\text{M}+\text{Na}]^+$ $m/z = 893.3655$, Mean err 0.6 ppm.

IR ATIR (cm^{-1}): 3030 w (C-H), 2866 w (C-H), 1728 s (C=O), 1586 w (C-H).

Benzyloxy-5-benzyloxy- β -D-glucosyl-3,4-benzyloxysalicylate (Compound N)



Molecular formula: $\text{C}_{62}\text{H}_{58}\text{O}_9$

Molecular weight: $947.1370 \text{ g mol}^{-1}$

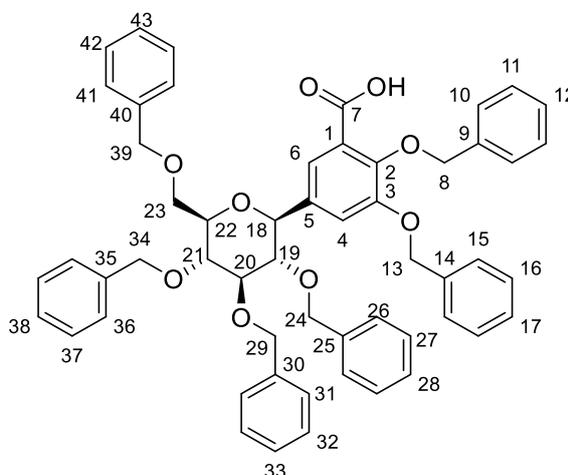
Benzyloxy-5-benzyloxy- β -D-glucosyl-3, 4-benzyloxysalicylate was prepared by the same method as methyl-5-benzyloxy- β -D-glucosyl-3,4-benzyloxysalicylate. The compounds were then separated from each other by column chromatography yielding the product as a white solid. (0.381 g, 0.40 mmol, 21%) $R_f = 0.26$ (1:4 ethyl acetate : petroleum ether 40-60°C) M.P = 104-106 °C.

$^1\text{H NMR}$: (400 MHz, CDCl_3) δ : 7.52 (d, H_6 , $J = 1.7$ Hz, 1H) 7.40-7.28 (m, H_{10} , H_{15} , H_{26-28} , H_{31-33} , H_{36-38} , H_{41-43} , H_{46-48} , 29H); 7.26-7.21 (m, H_4 , H_{11} , H_{16} , 5H) 6.94-6.92 (m, H_{12} , H_{17} , 2H) 5.34 (d, H_{44} , $J = 12.4$ Hz, 2H); 5.07 (d, H_8 , $J = 10.1$ Hz, 2H); 4.98 (d, H_{13} , $J = 2.2$ Hz, 2H); 4.97-4.87 (m, $H_{24/29}$, H_{39} , 3H) 4.66-4.57 (m, $H_{24/29/H_{23}}$, 3H); 4.43 (d, H_{23} , $J = 10.1$ Hz 1H); 4.21 (d, H_{18} , $J = 9.5$ Hz, 1H); 3.82-3.76 (m, $H_{20/21/24/29}$, H_{34} , 5H); 3.61-3.59 (m, H_{22} , 1H); 3.44 (t, H_{19} , $J = 9.0$ Hz, 1H).

$^{13}\text{C NMR}$: (100 MHz, CDCl_3) δ : $^{13}\text{C NMR}$ (100MHz, CDCl_3) δ : 165.88 (C_7); 152.54 (C_3); 147.99 (C_2); 138.60 (C_{25}); 138.23, 138.07, 137.37, 137.17, 136.34, 135.88, (C_1 , 10, 15, 30, 35, 41, 45); 128.62, 128.52, 128.41, 128.36, 128.35, 128.25, 128.23, 128.16, 128.13, 128.06, 127.99, 127.94, 127.91, 127.87, 127.83, 127.77, 127.74, 127.66, 127.63, 127.60, 127.56, 126.44 (C_5 , C_{10-12} , C_{15-17} , C_{26-28} , C_{31-33} , C_{36-38} , C_{42-44} , 46-48); 121.95 (C_6); 116.52 (C_4); 86.63 ($C_{20/21}$); 83.91 (C_{19}); 80.88 (C_{18}); 79.31 (C_{22}); 78.19 ($C_{20/21}$); 75.57 (C_8); 75.12 ($C_{24/29/34}$); 73.23 (C_{23}); 70.90 (C_{39}); 68.94 ($C_{24/29/34}$); 68.11 (C_{13}); 66.88(C_{44}).

HRMS (ESI): Calcd. $[\text{M}+\text{Na}]^+$ ($\text{C}_{62}\text{H}_{58}\text{NaO}_9$) $m/z = 969.3973$; Obs. $[\text{M}+\text{Na}]^+$ $m/z = 969.3982$, Mean err 1.3 ppm.

5-Benzyloxy- β -D-glucosyl-3, 4-benzyloxysalicylate (Compound P)



Molecular formula: $\text{C}_{55}\text{H}_{52}\text{O}_9$

Molecular weight: 857.02 g mol^{-1}

5-Benzyloxy- β -D-glucosyl-3,4-benzyloxysalicylate was prepared based on a preparation from the literature.^{396, 397}

Methoxy-5-aceto- β -D-glucosyl-3, 4-benzyloxysalicylate (0.500 g, 0.57 mmol) was dissolved in THF (7 mL) and methanol (3 mL) and NaOH 5 M in H₂O (2 mL) was added. The mixture was stirred for 21 hours, before acidification to pH 1 with 1 M HCl. Upon acidification, a fine white precipitate formed. The turbid aqueous mixture was extracted with DCM (3 \times 10 mL). The organic portions were combined, dried over MgSO₄, filtered and solvent removed *in vacuo* to yield a colourless oil. Petroleum ether 40-60 °C was added (3 \times 10 mL) and removed *in vacuo* to yield an off white solid residue. The residue was purified by column chromatography (1:4 methanol : chloroform) yielding a white solid product (0.359 g, 0.42 mmol, 73%) R_f = 0.56 (1:9 methanol : chloroform) M.P = 97-99 °C.

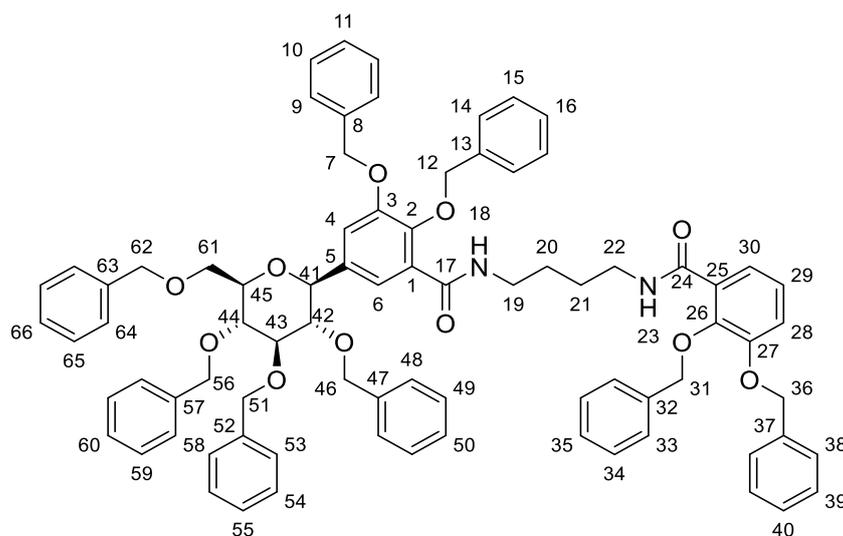
¹H NMR: (400 MHz, CDCl₃) δ : 11.29 (br s, -OH, 1H,); 7.89 (d, H₆, J = 1.8 Hz, 1H); 7.45-7.29 (m, H₁₁, H₁₆, H₂₆₋₂₈, H₃₁₋₃₃, H₃₆₋₃₈, H₄₂₋₄₄, 24H); 7.22-7.20 (m, H₄, H₁₂, H₁₇, 5H); 6.97-6.94 (m, H₁₀, H₁₅, 2H); 5.26 (d, H₈, J = 10.5 Hz, 2H); 4.99 (d, H₁₃, J = 2.3 Hz, 2H); 4.96-4.87 (m, H₃₄, H₂₉, 3H); 4.66-4.63 (m, H_{34/39}, 2H); 4.60-4.57 (d, H₃₉, J = 12.8 Hz, 1H); 4.49 (d, H₂₄, J = 10.5 Hz, 1H); 4.25 (d, H₁₈, J = 9.6 Hz, 1H); 3.86-3.80 (m, H₂₄, H₂₀, 2H); 3.78-3.72 (m, H₂₁, H₂₃, 3H); 3.65-3.62 (m, H₂₂, 1H); 3.44 (t, H₁₉, J = 9.2 Hz, 1H).

¹³C NMR: (100 MHz, CDCl₃) δ : 165.15 (C₇); 150.91 (C₃); 146.40 (C₂); 138.52 (C₃₀); 138.14 (C_{35/40}); 137.98 (C_{35/40}); 137.41 (C₂₅); 136.39 (C₅); 135.73 (C₉); 134.48 (C₁₄); 129.32, 129.29, 128.78, 128.75, 128.51, 128.42, 128.38, 128.24, 128.13, 128.05, 128.01, 127.81, 127.75, 127.70, 127.63, 127.60, 126.95 (C₁₁₋₁₃, C₁₆₋₁₈, C₂₆₋₂₈, C₃₁₋₃₃, C₃₆₋₃₈, C₄₂₋₄₄); 123.16 (C₆); 122.76 (C₁); 117.79 (C₄); 86.69 (C₂₀); 83.66 (C₁₉); 80.57 (C₁₈); 79.29 (C₂₂); 78.20 (C₂₁); 76.90 (C₈); 75.59 (C₂₉); 75.141 (C₃₄); 74.90 (C₂₄); 73.39 (C₃₉); 71.25 (C₁₃); 68.98 (C₂₃).

HRMS (ESI): Calcd. [M+Na]⁺ (C₅₅H₅₂NaO₉) m/z = 879.3504; Obs. [M+Na]⁺ m/z = 879.3529, Mean err 3.1 ppm.

IR ATIR (cm⁻¹): 3063 w (C-H), 3030 w (C-H), 2867 w (C-H), 1740 m (C=O), 1697 w (C=O), 1606 (C=C ar).

5-Benzyloxy- β -D-glucosyl-bis(3, 4-benzyloxy)-4-LICAM (Compound Q)



Molecular formula: $C_{80}H_{78}N_2O_{11}$

Molecular weight: $1243.51 \text{ g mol}^{-1}$

5-Benzyloxy- β -D-glucosyl-3, 4-benzyloxysalicylate (0.107 g, 0.13 mmol) was dissolved in DMF (10 mL) to which HATU (0.114 g, 0.30 mmol) was added. The mixture was stirred for one hour before the addition of DIPEA (78 μ L, 0.058 g, 0.45 mmol) and 1-amino, 4-(2,3-dibenzyloxybenzamide) butane (0.066 g, 0.15 mmol). The mixture was then heated to 50 $^{\circ}$ C and stirred overnight. The resulting brown solution was reduced in vacuo to a brown residue. The residue was purified twice by column chromatography (1:2 ethyl acetate : chloroform) (1:3 acetonitrile : chloroform) yielding a white solid product (0.071 g, 0.057 mmol, 46 %) $R_f = 0.47$ (1:2 ethyl acetate : chloroform) M.P = 84-86 $^{\circ}$ C.

$^1\text{H NMR}$: (400 MHz, CDCl_3) δ : 7.94 (d, H_6 , $J = 2.3 \text{ Hz}$, 1H); 7.91 (t, $H_{18/23}$, $J = 5.5 \text{ Hz}$, 1H) 7.86 (t, $H_{18/23}$, $J = 5.5 \text{ Hz}$); 7.78 (dd, $H_{28/30}$, $J = 6.2 \text{ Hz}$, $J = 3.2 \text{ Hz}$, 1H); 7.51-7.17 (m, H_4 , H_{29} , H_{9-11} , H_{14-16} , H_{33-35} , H_{38-40} , H_{48-50} , H_{53-55} , H_{58-60} , H_{64-66} , 42H); 6.98 (d, $H_{28/30}$, $J = 2.7 \text{ Hz}$, 2H); 5.18 (s, H_{12} , 2H); 5.10 (d, $H_{31/36}$, $J = 10.5 \text{ Hz}$, 1H); 5.08 (s, $H_{31/36}$, 2H); 5.03 (d, $H_{31/H36}$, $J = 10.5 \text{ Hz}$, 1H); 5.00 (br s, H_7 , 2H) 4.97-4.88 (m, H_{46} , H_{51} , 3H) 4.67-4.56 (m, H_{51} , H_{62} , 3H); 4.46 (d, H_{61} $J = 10.5 \text{ Hz}$, 1H); 4.28 (d, H_{41} , $J = 9.6 \text{ Hz}$, 1H); 3.88-3.75 (m, H_{61} , H_{43} , H_{56} , H_{45} , 5H); 3.65 (dt, H_{44} , $J = 9.6 \text{ Hz}$, $J = 3.2 \text{ Hz}$, 1H); 3.53 (t, H_{42} , $J = 9.2 \text{ Hz}$, 1H); 3.23-3.21 (m, H_{19} , H_{22} , 4H); 1.29-1.24 (m, H_{20} , H_{21} , 4H).

$^{13}\text{C NMR}$: (100 MHz, CDCl_3) δ : 164.93 ($C_{17/24}$); 164.65 ($C_{17/24}$); 151.64 ($C_{3/27}$); 151.33($C_{3/27}$); 146.76 ($C_{2/26}$); 146.27 ($C_{2/26}$); 138.64, 138.22, 138.06, 137.66, 136.37, 136.29, 136.15, 135.51 (C_8 , 13, 32, 37, 47, 52, 57, 63); 128.74, 128.70, 128.67, 128.65, 128.63, 128.60, 128.36, 128.32, 128.20, 128.16, 128.09, 127.98, 127.73, 127.66, 127.60,

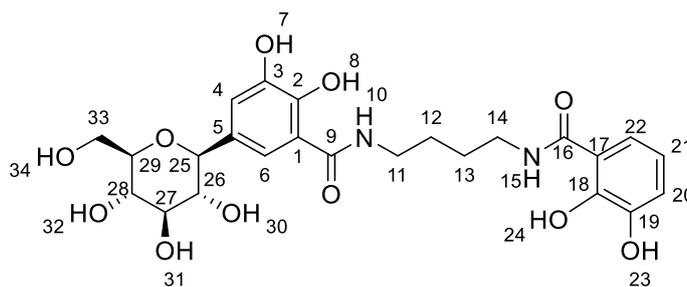
127.53, 127.51, 127.27, 127.13 (C9-11, 14-16, 33-35, 38-40, 48-50, 53-55, 58-60, 64-66); 124.35(C4/28/29/30); 123.29(C4/28/29/30); 122.18(C4/28/29/30); 116.87(C28/30); 115.85 (C6); 86.65 (C43/45); 83.93(C42); 81.08(C41); 79.32 (C44); 78.26 (C43/45); 76.33(C31/36); 75.52 (C31/36); 75.08 (C46/51); 74.81 (C61); 73.39 (C51/62); 71.22 (C33/36); 70.99 (C46/51); 69.11 (C56); 39.31 (C19/22); 39.25 (C19/22); 26.72 (C20/21).

HRMS (ESI): Calcd. $[M+H]^+$ (C₈₀H₇₉N₂O₁₁) $m/z = 1243.5678$; Obs. $[M+H]^+$ $m/z = 1243.5687$, Mean err 0.6 ppm. Calcd. $[M+Na]^+$ (C₈₀H₇₉N₂NaO₁₁) $m/z = 1265.5498$; Obs. $[M+Na]^+$ $m/z = 1265.5488$, Mean err 0.7 ppm.

IR ATIR (cm⁻¹): 3287 w br (N-H), 3026 w (N-H), 2863 w br (C-H), 1639 m (C=O), 1577 m (C=C ar).

Elemental Analysis: Calcd. for [C₈₀H₇₈N₂O₁₁]: %C 77.27, %H 6.32, %N 2.25; Measured for [C₈₀H₇₈N₂O₁₁]: %C 77.18, %H 6.31, %N 2.20

5-β-D-glucosyl-4-LICAM



Molecular formula: C₂₄H₃₀N₂O₁₁

Molecular weight: 522.51 g mol⁻¹

5-β-D-Glucosyl-4-LICAM was prepared based on a preparation from the literature.^{396, 397}

5-Benzyloxy-β-D-glucosyl-3, 4-benzyloxy-4-LICAM (0.170 g, 0.33 mmol) was dissolved in 1:1 ethyl acetate: methanol (30 mL) and 3 spatula tips of Pd(OH)₂ 20% on carbon were added. The system was purged with nitrogen before purging with hydrogen for 30 minutes. The reaction mixture was stirred under balloon pressure of hydrogen for 18 hours. The reaction mixture was then purged with nitrogen before opening to air. The catalyst was removed by filtration, and the solvent removed *in vacuo* to yield a pale brown oil. The solid off white product was obtained by cooling the obtained oil in liquid nitrogen and removing all residual solvent *in vacuo* (0.066 g, 0.12 mmol, 93 %) R_f = 0.00 (1:2 ethyl acetate : chloroform) M.P=195-197 °C.

¹H NMR: (400 MHz, MeOD) : 7.31 (d, *H*6, *J* = 1.8 Hz, 1H); 7.21 (dd, *H*20/22, *J* = 8.2 Hz, *J* = 1.4 Hz, 1H); 7.01 (d, *H*4, *J* = 1.8 Hz, 1H); 6.92 (dd, *H*20/22, *J* = 7.8 Hz, *J* = 1.4 Hz, 1H); 7.61 (t, *H*21, *J* = 7.8 Hz, 1H); 4.02 (d, *H*25, *J* = 9.2 Hz, 1H); 3.86 (dd, *H*33, ²*J*_{*H*33*a*/*H*33*b*} = 11.9 Hz, ³*J*_{*H*33*a*/*H*29} = 2.4 Hz, 1H); 3.72 (dd, *H*33, ²*J*_{*H*33*a*/*H*33*b*} = 11.9 Hz, ³*J*_{*H*33*b*/*H*29} = 5.2 Hz, 1H); 3.46-3.43 (m, *H*11, *H*14, *H*26-29, 8H); 1.72-1.69 (m, *H*12-13, 4H).

¹³C NMR: (100 MHz, MeOD) δ: 171.66 (*C*9/16); 171.66 (*C*9/16); 150.39 (*C*3/19); 150.27(*C*3/19); 147.43 (*C*2/18); 147.04 (*C*2/18); 131.31 (*C*1/17); 119.71 (*C*20/22); 119.50 (*C*4); 118.78 (*C*20/22); 118.15 (*C*6); 116.89 (*C*21); 116.19 (*C*5); 83.29 (*C*25); 82.11 (*C*26/27/28/29); 79.82(*C*26/27/28/29); 76.43 (*C*26/27/28/29); 71.86 (*C*26/27/28/29); 63.05 (*C*33); 40.29 (*C*11/14); 40.24 (*C*11/14); 28.03 (*C*12/13); 28.01 (*C*12/13).

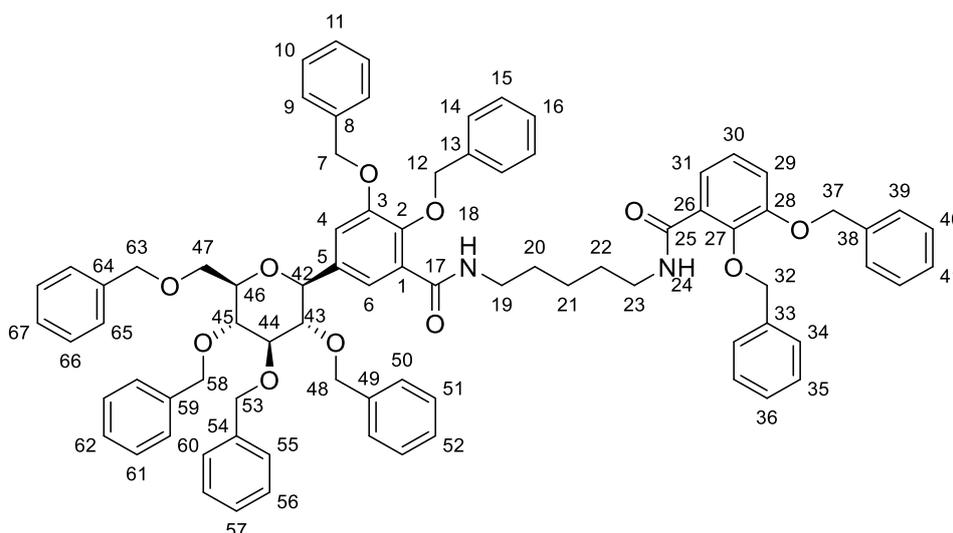
HRMS (ESI): Calcd. [*M*+*H*]⁺ (*C*₂₄*H*₂₉*N*₂*O*₁₁) *m/z* = 521.1777; Obs. [*M*+*H*]⁺ *m/z* = 521.1771, Mean err 0.8 ppm.

IR ATIR (cm⁻¹): 3284 s br (O-H), 2924 m (N-H), 1640 m (C=O), 1587 m (C=C ar), 1452 s (C-O).

Elemental Analysis: Calcd. for [*C*₂₄*H*₃₀*N*₂*O*₁₁·2.1*H*₂*O*]: %C 51.44, %H 6.15, %N 5.00; Measured for [*C*₂₄*H*₃₀*N*₂*O*₁₁·2.1*H*₂*O*]: %C 51.18, %H 5.81, %N 4.96.

Specific Rotation: [*α*]_D (Methanol, conc 0.245 g/100 mL) + 3.2

5-Benzyloxy-β-D-glucosyl-bis(3, 4-benzyloxy)-5-LICAM (Compound T)



Molecular formula: *C*₈₁*H*₈₀*N*₂*O*₁₁

Molecular weight: 1257.53 g mol⁻¹

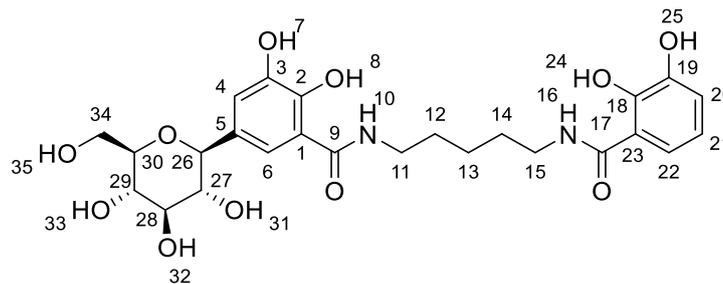
5-Benzyloxy- β -D-glucosyl-3, 4-benzyloxysalicylate (0.252 g, 0.29 mmol) was dissolved in DMF (10 mL) to which HATU (0.268 g, 0.70 mmol) was added. The mixture was stirred for one hour before the addition of DIPEA (184 μ L, 1.06 mmol) and 1-amino, 5-(2, 3-dibenzyloxybenzamide) pentane (0.160 g, 0.35 mmol). The mixture was then stirred overnight. The resulting brown solution was reduced *in vacuo* to a brown residue. The residue was purified twice by column chromatography (1:2 ethyl acetate : chloroform) (1:4 ethyl acetate : chloroform) yielding a white solid product (0.204 g, 0.16 mmol, 55 %) R_f = 0.36 (1:4 ethyl acetate : chloroform) M.P = 96-98 °C.

$^1\text{H NMR}$: (400 MHz, CDCl_3) δ : 7.93 (d, H_6 , J = 1.8 Hz, 1H); 7.90 (m, $H_{18/24}$, 1H); 7.87 (t, $H_{18/24}$, J = 5.5 Hz, 1H); 7.76 (dd, $H_{29/31}$, J = 6.4 Hz, J = 3.2 Hz, 1H); 7.50-7.14 (m, H_4 , H_{30} , H_{9-11} , H_{14-16} , H_{34-36} , H_{39-41} , H_{50-52} , H_{55-57} , H_{60-62} , H_{65-67} , 42H); 6.98 (m, $H_{29/31}$, 2H); 5.17 (s, H_{12} , 2H); 5.11-5.01 (m, $H_{32/37}$, 2H); 5.08 (s, $H_{32/37}$, 2H); 4.98 (br s, H_7 , 2H); 4.95-4.86 (m, H_{48} , H_{53} , 3H); 4.67-4.56 (m, H_{53} , H_{64} , 3H); 4.45 (d, H_{47} J = 10.5 Hz, 1H); 4.26 (d, H_{42} , J = 9.6 Hz, 1H); 3.86-3.73 (m, H_{47} , H_{44} , H_{58} , H_{46} , 5H); 3.63 (dt, H_{45} , J = 9.2 Hz, J = 3.2 Hz, 1H); 3.52 (t, H_{43} , J = 9.2 Hz, 1H); 3.25-3.18 (m, H_{19} , H_{23} , 4H); 1.30-1.23 (m, H_{20} , H_{22} , 4H); 1.16-1.10 (m, H_{21} , 2H).

$^{13}\text{C NMR}$: (100 MHz, CDCl_3) δ : 164.93 ($C_{17/25}$); 164.66 ($C_{17/25}$); 151.67 ($C_{3/28}$); 151.37 ($C_{3/28}$); 146.74 ($C_{2/27}$); 146.27 ($C_{2/27}$); 138.67, 138.25, 137.69, 136.40, 136.34, 136.27, 136.27, 135.54 (C_8 , 13, 33, 38, 49, 54, 59, 64); 128.78, 128.71, 128.67, 128.64, 128.40, 128.35, 128.24, 128.20, 128.15, 128.02, 127.77, 127.71, 127.65, 127.58, 127.54, 127.49, 127.37, 127.24 (C_{9-11} , C_{14-16} , C_{34-36} , C_{39-41} , C_{50-52} , C_{55-57} , C_{60-62} , C_{65-67}); 124.40 ($C_{4/29/30/31}$); 123.31 ($C_{4/29/30/31}$); 122.21 ($C_{4/29/30/31}$); 116.81 ($C_{29/31}$); 115.78 (C_6); 86.68 ($C_{44/46}$); 83.95 (C_{43}); 81.13 (C_{42}); 79.34 (C_{45}); 78.29 ($C_{44/46}$); 76.36 ($C_{32/37}$); 75.56 ($C_{32/37}$); 75.13 ($C_{48/53}$); 74.85 (C_{47}); 73.42 ($C_{53/63}$); 71.24 ($C_{32/37}$); 71.01 ($C_{48/53}$); 69.11 (C_{58}); 39.55 ($C_{19/23}$); 39.48 ($C_{19/23}$); 28.94 ($C_{20/22}$), 28.92 ($C_{20/22}$), 24.46 (C_{21}).

HRMS (ESI): Calcd. $[\text{M}+\text{H}]^+$ ($\text{C}_{81}\text{H}_{81}\text{N}_2\text{O}_{11}$) m/z = 1257.5835; Obs. $[\text{M}+\text{H}]^+$ m/z = 1257.5917, Mean err 4.7 ppm.

IR ATIR (cm^{-1}): 3384 w br (N-H), 3287 w br (N-H), 3063 w (C-H), 3030 w (C-H), 2920, 2859 w br (C-H), 1638 m (C=O), 1605 w (C=O), 1577 m (C=C ar).

5- β -D-glucosyl-5-LICAM

Molecular formula: C₂₅H₃₂N₂O₁₁

Molecular weight: 522.53 g mol⁻¹

5- β -D-Glucosyl-5-LICAM was prepared based on a preparation from the literature.^{396, 397}

5-Benzyloxy- β -D-glucosyl-3, 4-benzyloxy-4-LICAM (0.200 g, 0.1590 mmol) was dissolved in toluene (1 mL) and ethanol added (30 mL) followed by 3 spatula tips of Pd(OH)₂ 20% on carbon. The system was purged with nitrogen before purging with hydrogen for 30 minutes. The reaction mixture was stirred under balloon pressure of hydrogen for 18 hours. The reaction mixture was then purged with nitrogen before opening to air. The catalyst was removed by filtration, and the solvent removed *in vacuo* to yield a pale colourless oil. The solid off white product was obtained by cooling the obtained oil in liquid nitrogen and removing all residual solvent *in vacuo* (0.0855 g, 0.16 mmol, 90 %) R_f = 0.00 (1:2 ethyl acetate : chloroform) M.P = 180-182 °C.

NMR: (400 MHz, MeOD) δ : 7.31 (s, H_{4/6}, 1H); 7.20 (d, H_{20/22}, J = 8.2 Hz 1H); 7.01 (s, H_{4/6}, 1H); 6.91 (d, H_{20/22}, J = 7.8 Hz, 1H); 6.70 (t, H₂₁, J = 7.8 Hz, 1H); 4.02 (d, H₂₆, J = 9.2 Hz, 1H); 3.87 (d, H₃₄, J = 10.5 Hz, 1H); 3.72 (dd, H₃₄, ²J_{H34a/H34b} = 11.9 Hz, ³J_{H34/H30} = 5.0 Hz, 1H); 3.46-3.35 (m, H₁₁, H₁₅, H₂₇₋₃₀, 8H); 1.72-1.64 (m, H₁₂, H₁₄, 4H); 1.50-1.44 (m, H₁₃, 2H).

¹³C NMR: (100 MHz, MeOD) δ : 171.64 (C_{9/17}); 171.61 (C_{9/17}); 150.41 (C_{3/19}); 150.27(C_{3/19}); 147.46 (C_{2/18}); 147.06 (C_{2/18}); 131.32 (C_{1/23}); 119.69 (C_{21/20/22}); 119.44 (C_{21/20/22}); 118.75 (C_{4/6}); 118.12 (C_{20/22}); 116.91 (C_{4/6}); 116.21 (C₅); 83.32 (C₂₆); 82.14 (C_{27/28/29/30}); 79.84 (C_{27/28/29/30}); 76.46 (C_{27/28/29/30}); 71.84 (C_{27/28/29/30}); 63.03 (C₃₄); 40.54 (C_{11/15}); 40.52 (C_{11/15}); 30.33 (C_{12/14}); 30.26 (C_{12/14}), 25.55 (C₁₃).

HRMS (ESI): Calcd. [M+H]⁺ (C₂₅H₃₂N₂O₁₁) m/z = 535.1933; Obs. [M+H]⁺ m/z = 535.1912, Mean err 3.5 ppm.

IR ATIR (cm⁻¹): 3288 s br (O-H), 2930 m br (C-H), 1641 m (C=O), 1589 m (C=C ar).

Elemental Analysis: Calcd. for [C₂₅H₃₂N₂O₁₁.0.9EtOH.1.2H₂O]: %C 53.68, %H 6.70, %N 4.66; Measured for [C₂₅H₃₂N₂O₁₁.0.9EtOH.1.2H₂O]: %C 53.76, %H 6.47, %N 4.52.

Specific Rotation: [α]_D (Methanol, conc. 0.311 g/100 mL) + 5.7

6.3 Sequence Alignments

For known proteins, where structures were identified in the PDB *via* PDBeFold, sequences were obtained by using a BLAST search for the sequence of CeuE against protein structure sequences.⁴⁴⁰ For unknown proteins, sequence alignment was carried out using a BLAST search for the full sequence of CeuE.⁴⁴⁰ Sequence matches from 1000 possible hits were selected for bacterial species of interest. All sequences were aligned using MUSCLE with ClustalW output format,⁴²⁶ and the sequence alignment image generated with ESPript.⁴²⁷

6.4 Protein Production and Purification

FepB was prepared under the guidance of Dr E. V. Blagova as follows.

FepB Protein Sequence

The FepB protein sequence was obtained as follows for *E. coli* K strain MC1061:

MRLAPLYRNALLLTGLLLSGIAAVQAADWPRQITDSRGTHTLESQPQRIVSTSVTL
TGSLLAIDAPVIASGATTPNNRVADDQGFLRQWSKVAKERKLQRLYIGEPSAEAV
AAQMPDLILISATGGDSALALYDQLSTIAPTLIINYDDKSWQSLLTQLGEITGHEKQ
AAERIAQFDKQLAAAKEQIKLPPQPVTIVYTA AAHSANLWTPESAQQQMLEQLG
FTLAKLPAGLNASQSQGKRHDIIQLGGENLAAGLNGESLFLFAGDQKDADAIYAN
PLLAHLPAVQNKQVYALGTETFRLDYYSAMQVLDRLKALF

Primer Design and Cloning

Primers were designed for the gene sequence coding for amino acid residues 27-318 FepB in K Strain MC1061 *E. coli*. Primers were designed using Regional Order Neural Network disorder predictions⁴²⁹, Uniprot including Reference Clusters, T-Coffee, ClustalW2 and Reverse Complement.⁴⁴¹ The designed construct consisted of 873 base pairs. Gradient PCR of the construct was carried out using a 25 cycle procedure of denaturation at 94°C for 30 seconds, annealing at 53 ± 10 °C for 30 seconds, and extension at 72 °C for 50 seconds, followed by a 3-minute final extension. The In-fusion protocol was used to insert the desired construct into the Lic+ vector. This was achieved by mixing the vector and the insert with the In-fusion enzyme and heating for 15 minutes at 50 °C, followed by cooling to 4°C. The resulting solution was diluted in TE buffer and Competent Nova Blue cells transformed

using a heat shock method. The method consisted of cooling 25 μL of cells with 2.5 μL diluted In-fusion mixture on ice for 10 minutes. The cells were then heated at 42 °C for 35 seconds, and cooled on ice for a further 10 minutes. 90 μL of GS96 media was added and the cells were incubated at 37 °C for 1 hour. The culture was plated on agar containing 1% kanamycin and incubated overnight at 37°C.

Colony PCR was carried out on 16 colonies resulting from the transformed Competent Nova Blue cells with the procedure as follows. Heating at 94 °C for 2 minutes, followed by 25 cycles of heating to 94 °C for 30 seconds, cooling to 52 °C for 30 seconds and then heating to 72 °C for 40 seconds. The heating and cooling cycles were then followed by a final extension at 72 °C for 3 minutes, and cooling to 4 °C. 11 out of 16 colonies showed positive inserts of the desired vector. 8 of these colonies were taken forward for culture. Each colony was incubated at 37 °C overnight in 10 mL of sterile LB media, with 1% kanamycin.

Plasmid DNA from the 8 resulting cultures was purified *via* the following protocol. Cells were centrifuged at 5000 rpm for 15 minutes, and the supernatant removed. The pellet was resuspended in buffer P1 (250 μL) and transferred to a microcentrifuge tube where buffer P2 (250 μL) was added and mixed by inversion. After a maximum of 5 minutes, buffer N3 (350 μL) was added and mixed by inversion. The mixture was centrifuged at 13000 rpm for 10 minutes. The supernatant was applied to a QIAprep spin column and centrifuged for 1 minute. The flow-through was discarded and the column was washed with buffer PB (500 μL) and centrifuged for 1 minute. The flow-through was discarded and the column was washed with buffer PE (750 μL) and centrifuged for 1 minute. The column was placed in a clean 1.5 mL collection tube, and DNA eluted with EB buffer (50 μL) by standing for 1 minute followed by centrifugation for 1 minute.

Test restriction was carried out using *Nco*1 and *Nde*1 enzymes and heating to 37 °C for 2 hours. Restriction Cut software⁴⁴² showed no predicted cleavage of the main sequence of the inserted gene. Positive results were produced for DNA obtained from 4 out of the 8 colonies.

External sequencing of 2 of the 4 DNA samples gave positive results for further transformation of expression cells, with 100% sequence similarity to that of the initial K Strain MC1061 gene, as designed.

Test Expression

BL21 expression cells were transformed for test expression using the heat shock method. The method consisted of cooling 30 μL of cells with 1 μL diluted DNA (20 pmol/ μL) on ice for 10 minutes. The cells were then heated at 42°C for 35 seconds, and cooled on ice for a

further 10 minutes. 90 μ L of GS96 media was added and the cells were incubated at 37 °C for 1 hour. The culture was plated on agar containing 1% kanamycin and incubated overnight at 37 °C.

LB media (10 mL) with 1% kanamycin was incubated overnight at 37 °C for two resulting colonies. 200 μ L of each overnight culture was taken for inoculation in LB media (10 mL) with 1% kanamycin. The inoculation cultures were incubated for 100 minutes until the OD₆₀₀ reached 0.653 and 0.563. The cultures were taken forward for induction as well as a portion stored as 25% glycerol stocks at -80 °C for future use.

1% IPTG was used to induce the cells, keeping one culture at 37°C for 4 hours, and the other at 16 °C overnight. 4 stocks of 1 mL were centrifuged at 600 rpm and the supernatant removed. The cells were stored at -80°C. The pellet was defrosted and suspended in 50 mM Tris-HCl, 500 mM NaCl, pH 7.5 (100 μ L) and sonicated for 15 seconds. The suspended cells were aliquoted into 10 equal tubes of 10 μ L and centrifuged at 13000 rpm for 2 minutes to remove the cell pellet. 10 μ L supernatant was deemed the soluble protein. The soluble protein was loaded onto a 12% SDS PAGE gel. The gel was run and stained with Coomassie Brilliant Blue R250. The gel was then soaked in a destaining solution for 24 hours, showing the presence of both soluble and insoluble protein of the correct molecular weight. Protein sequencing gave 100% sequence similarity to the desired product of the initial K Strain MC1061 gene.

Overexpression

Expression was scaled up to 2L BL21 cultures in LB broth. After 105 minutes at 37°C, OD₆₀₀ reached 0.6-0.8. The cells were induced with IPTG and incubated at 16 °C overnight. SDS PAGE gel showed good overexpression.

The 2 L cultures were centrifuged and the cell pellet collected and frozen at -80 °C. The cell pellet (16 mL) was thawed for protein purification and resuspended in 50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 10 mM imidazole (80mL). The suspended cells were sonicated for 55 seconds kept on ice and below 20 °C during sonication. When the temperature had fallen back to 8 °C, after around 15 minutes, a further 55 seconds of sonication were carried out. The process was repeated for 4 sonications. The cell pellet was removed *via* centrifugation in two portions at 4 °C.

Purification

The Lic+ vector provided a histidine tag labelled FepB, meaning the protein could be purified *via* nickel column chromatography.

Nickel column buffer containing 50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 10 mM imidazole was made on a 1 L scale by adding 50 mL 1M Tris-HCl stock, 29.22 g NaCl and 0.6808 g Imidazole to 900 mL of Millipore filtered deionised water. The buffer was adjusted to pH 7.5 with 1 M HCl and made up to 1 L with Millipore filtered deionised water. The buffer was then degassed *via* vacuum pump filtration.

High imidazole nickel column buffer containing 50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 500 mM imidazole was made on a 1 L scale by adding 50 mL 1M Tris-HCl stock, 29.22 g NaCl and 34.04 g Imidazole to 900 mL of Millipore filtered deionised water. The buffer was adjusted to pH 7.5 with 1 M HCl and made up to 1 L with Millipore filtered deionised water. The buffer was then degassed *via* vacuum pump filtration.

Gel filtration buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, was made on a 1 L scale by adding 50 mL 1M Tris-HCl stock and 8.76 g NaCl to 900 mL of Millipore filtered deionised water. The buffer was adjusted to pH 7.5 with 1 M HCl and made up to 1 L with Millipore filtered deionised water. The buffer was then degassed *via* vacuum pump filtration.

The 5 mL volume nickel column was prepared by washing with 7.5 mL NaOH, 15 mL Millipore filtered deionised water, 7.5 mL 0.5M EDTA pH 8.0, 30 mL Millipore filtered deionised water. The column was then charged with 7.5 mL 0.1 M NiSO₄ for 20 minutes, and washed with 10 mL Millipore filtered deionised water to removed excess NiSO₄.

The nickel column was connected to an Akta purifier and equilibrated with nickel column buffer and high imidazole nickel column buffer. A superloop of 50 mL volume was connected and washed with nickel column buffer at a flow rate of 5 mL/min. the nickel column was attached and washed with nickel column buffer at a flow rate of 1 mL/min. The superloop was then loaded with 45 mL crude FepB protein extract.

Akta purifier was prepared for collection of 8 mL fractions in 18 mm tubes. Fractions of FepB in nickel column buffer were collected after passage through the nickel column at a flow rate of 3.5 mL/min and a column pressure limit of 0.5 mPa. Once the superloop was sufficiently emptied, the remaining 25 mL crude FepB protein extract was loaded. UV/Vis spectra were measured at 254, 280 and 420 nm. The column was washed with 35 mL nickel column buffer before introducing high imidazole nickel column buffer to create an imidazole gradient over 3 steps, of 35 mL 33%, 35 mL 66% and 35 mL 100% at a flow rate of 3.5 mL/min and a column pressure limit of 0.5 MPa. The eluted protein was collected in 4 mL fractions in 18 mm tubes. The run was carried out for 67 minutes. The protein eluted in 14 × 4 mL in the gradient fractions was deemed to be the histidine tagged FepB, and the

remaining protein in the flow-through fractions was discarded. The fractions were analysed for the presence of histidine tagged FepB *via* 15% SDS PAGE gel run at 200V for 48 minutes. The gel was stained with Coomassie dye, heated for 70 seconds and shaken for 3 minutes before being rinsed with destain solution and heated for a further 90 seconds. The excess dye was extracted *via* absorption with tissue and further shaking. The gel was imaged to reveal the histidine tagged FepB in the gradient fractions, which were then combined.

The Bradford method was used to estimate the protein concentration using dilutions of 1000, 2000 and 3000 fold. The absorbance at 595 nm of the protein mixed in a 1:1 ratio with Bradford reagent was measured after standing for 10 minutes. The equation $y = 0.0297x + 0.0267$ was used to calculate protein concentration in mg/mL. The dilutions estimated concentration to be 5.6 mg/mL, 8.9 mg/mL and 10.8 mg / mL respectively, averaging to give an estimated concentration of 8.4 mg/mL.

The histidine tag was cleaved *via* C3 protease, added dropwise in a 1:50 ratio (0.94 mL 10mg/mL C3 protease stock, 56 mL, 8.4 mg/mL FepB) followed by the addition of DTT added dropwise (56 μ L, 2.5 M). The mixture was placed into dialysis tubing and dialysed against 4 L 50 mM Tris-HCl, pH 7.5, 150 mM NaCl overnight at 4°C.

The digest was deemed successful *via* 15% SDS PAGE gel of samples of FepB before and after the digest, showing a small shift in band position, indicating a loss of the histidine tag. The digested, dialysed protein was purified *via* a second nickel column to remove any remaining tagged protein. The nickel column was prepared as previously described, and the Akta purifier prepared and protein loaded *via* superloop. Gel filtration buffer was used to elute the protein followed by a gradient of high imidazole nickel column buffer to elute any remaining protein. The untagged protein eluted in the flow through portion, over 17×4 mL portions. UV spectra were measured at 254, 280 and 420 nm and the presence of untagged FepB was confirmed *via* 15% SDS PAGE gel.

The flow-through FepB fractions were collected combined and concentrated in protein filter Falcon tubes at 5000 rpm for 12 minutes. The flow through was checked to be free of protein *via* the Bradford method, and the concentrated protein above the filter collected for size exclusion gel filtration chromatography.

The gel filtration column (Superdex75 S75 Hiload 16/60, 120 mL volume) on the Akta purifier was equilibrated with 150 mL gel filtration buffer at a flow rate of 1 mL/min and pressure limit 0.5 mPa. The loop was washed by syringe with 30 mL gel filtration buffer. Concentrated protein 4 mL was loaded and gel filtration buffer used to elute the protein at a

flow rate of 1.2 mL/min for a total volume of 130 mL. The run lasted 110 minutes. UV spectra were recorded at 256, 280 and 420 nm and 4 mL fractions collected in 18 mm tubes. When the UV spectra indicated the protein had eluted, the loop was washed with 30 mL gel filtration buffer and the gel filtration process repeated for two further portions of concentrated protein. 15% SDS PAGE gel confirmed the presence of FepB in 3 fractions of each gel filtration. The total 9 fractions were combined.

The concentration of purified FepB was estimated *via* the Bradford method detailed above to give an estimated concentration of 13.2 mg/mL. As the total volume was 35 mL, the total protein was estimated to be 460 mg.

The purified protein was concentrated to 7.5 mL *via* centrifugation in a protein filter Falcon tube at 5000 rpm. The final concentration was estimated *via* the Bradford method as 52mg/mL. The total protein was estimated to be 450 mg and was aliquoted into portions and stored at -80°C until required. Correct protein folding was confirmed by circular dichroism, and ESI-MS confirmed a molecular weight of 31932.2 Da.

For further use of the protein, a more accurate concentration was determined using UV-visible spectroscopy. ProtParam was used to calculate an estimated molar absorption coefficient for FepB of $35410 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm in water.⁴⁴³ The absorbance of each protein in 40 mM Tris-HCl pH 7.5, 150 mM NaCl at 280 nm was recorded, and the concentration of the solution determined using the Beer-Lambert Law.

6.5 Preparation of CeuE variant proteins

CeuE Mutants were prepared as follows by Dr. E. V. Blagova. PCR-based site directed mutagenesis was used on the CeuE construct containing residues 24–330. Each desired mutation was incorporated into amplification primers for plasmid DNA in the YSBLic3C vector. Three single CeuE mutants were constructed:

H227L (FWD_H227_Lataaaagtaggcacactcggaaaaagtatcaat; Rev_H227_L
attgatacttttccgagtgctacttttat),

H227A (FWD_H227_Aataaaagtaggcacagccggaaaaagtatcaat; REV_H227_A
attgatacttttccggctgtgctacttttat)

Y288F (FWD_Y288_F gatccagaatactggttttagcaagtggaaat; REV_Y288_F
atttcacttgctaaaaaccagtattctggatc).

The plasmid DNAs of the CeuE single variants were then subjected to a further PCR cycle to yield the double mutants H227L/Y288F and H227A/Y288F. The mutant proteins were

expressed and purified according to the standard procedure detailed for FepB. Pure mutants were concentrated to 20–40 mg/ml and stored at $-80\text{ }^{\circ}\text{C}$.

6.6 Protein Crystallography

General Protein Crystallisation Procedure

All protein crystals were obtained by sitting drop vapour diffusion crystallisation using commercial screens aliquoted into 96 well plates using a Mosquito nanolitre pipetting robot (TTP LabTech, UK).

Crystallisation screens of CeuE and CeuE mutants were performed using the PACT (Molecular Dimensions) screen with each drop contained 150 nL protein complex solution and 150 nL reservoir solution.

Crystals obtained were initially tested for diffraction at 120 K under the direction of Dr E. V. Blagova, using a MarResearch Mar345 Image Plate Detector with Rigaku VariMax HF Confocal X-ray Optics. Alternatively, crystals were tested by Mr S. Hart using a Rigaku Actor Robotic Sample Changer setup with Rigaku VariMax HF Confocal X-ray Optics and an RAXIS IV++ imaging plate detector. Both systems were supplied with X-rays *via* a Rigaku MicroMax-007 HF X-ray Generator, with crystals mounted under an Oxford Cryosystem 700 Series Cryostream. Crystals of acceptable quality were then sent for data collection at the Diamond Light Source.

Preparation of Iron(III)-Ligand Complexes

All iron(III)-n-LICAM complexes were prepared by stirring ~ 3 mg of the relevant n-LICAM⁴⁻ in 1 mL methanol with the relevant FeCl₃ concentration from aqueous stock to produce 1:1 or 2:3 complexes over 1 hour. Solvent was removed *in vacuo* and the Fe-n-LICAM complex taken up in DMF at the desired concentration. Iron(III)-5-LICAM and iron(III)-6-LICAM were each prepared in a 1:1 ratio in a 100 mM stock solution. Iron(III)-8-LICAM was prepared *via* the same method to obtain a 1 M stock solution of a 2:3 complex.

Co-crystallisation of Iron(III)-Ligand Complexes with CeuE and Mutant CeuE H227L

For CeuE-iron(III)-6-LICAM⁻, 57 μL of CeuE diluted to 20 mg mL⁻¹ was mixed with 3 μL of the iron(III)-6-LICAM⁻ solution in a 1:10 molar ratio. The solution was centrifuged at 13000 RPM for 2 minutes to remove any precipitate. The resulting supernatant was separated from the precipitate and used for crystallisation screening. CeuE-iron(III)-8-LICAM was

prepared in the same way as CeuE-iron(III)-6-LICAM but in a ligand: protein molar ratio of 1:100. H227L-CeuE-iron(III)-5-LICAM was prepared in a 1:4 molar ratio.

For CeuE-iron(III)-6-LICAM the best diffraction was observed for those grown from PACT conditions C4 (0.1 M PCB buffer comprising 2:1:2 molar ratio sodium propionate, sodium cacodylate, and BisTrisPropane, pH 7, 25% PEG 1500). A crystal from condition C4 was cryo-protected (0.1 M PCB buffer comprising 2:1:2 molar ratio sodium propionate, sodium cacodylate, and BisTrisPropane, pH 7, 32.5% PEG 1500) and used for data collection. For CeuE-iron(III)-8-LICAM the best diffracting crystals grew in PACT condition H6 0.2 M sodium formate, 0.1 M BisTrisPropane buffer, pH 8.5, 20% PEG 3350. A crystal from condition H6 was cryo-protected in 0.2 M sodium formate, 0.1 M BisTrisPropane buffer, pH 8.5, 32.5% PEG 3350 and used for data collection. For H227L-CeuE-iron(III)-5-LICAM the best diffracting crystals grew in PACT conditions, 0.1 M PCB buffer, pH 8, 20% PEG 1.5 K. The CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM crystals were cryo-protected with 32.5% PEG 1500 and 32.5% PEG 3350 respectively and vitrified at 110 K. H227L-Fe-5-Lic was cryo-protected with 32.5 % PEG 1.5 K vitrified at 110 K.

Crystallisation of CeuE Native Mutants

Crystals were grown, cryo-protected and sent for data collection by Dr E. V. Blagova. Crystals were obtained from the following sitting drop vapour diffusion screening conditions: CeuE-H227L (0.2 M NaBr; 0.1 M BTP, pH 8.5; 20% PEG 3350); CeuE-H227A (0.1 M MMT, pH 9.0; 25% PEG 1.5K) CeuE-Y288F (0.01 M ZnCl₂; 0.1 M MES, pH 6.0; 20% PEG 6 K); CeuE-H227L/Y288F (0.1 M MIB, pH 9.0; 25% PEG 1.5 K); CeuE-H227A/Y288F (0.1 M SPG, pH 9.0; 25% PEG 1.5 K). All crystals were cryo-protected in the relevant well solution with an increased PEG concentration (20% increased to 36%, 25% increased to 41%) before vitrification at 110 K.

6.7 Structure Solution

Data were collected at the Diamond Light Source. Computations were carried out using programs from the CCP4 suite.⁴⁴⁴ Diffraction images were processed with *XIA2*⁴⁴⁵⁻⁴⁴⁷. The structures were solved using *MOLREP*⁴⁴⁸ or *PHASER*⁴⁴⁹, and refined with cycles of *REFMAC*⁴⁵⁰ iterated with manual model building in *COOT*.³⁶⁵ Validation was carried out using *MOLPROBITY*.⁴⁵¹

Spacegroups and Structure Solution for CeuE Co-crystal Structures

CeuE-iron(III)-6-LICAM was in space group $P2_12_12_1$, with a single protein monomer in the asymmetric unit. The CeuE-iron(III)-8-LICAM crystal was also in space group $P2_12_12_1$, but with quite different cell dimensions and crystal packing. H227L-CeuE-iron(III)-5-LICAM was in space group $P1$. The structures were solved starting with CeuE-iron(III)-4-LICAM (PDB ID: 5A1J) as a search model using *MOLREP*⁴⁴⁸ for H227L-CeuE-iron(III)-5-LICAM and using PHASER⁴⁴⁹ for CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM. Restraints for the ligands were modelled using JLigand.⁴⁵²

Spacegroups and Structure Solution for CeuE Variant Proteins

CeuE-H227L, CeuE-H227A, CeuE-H227L/Y288F and CeuE-H227A/Y288F were in space group $P1$ with three protein monomers in the asymmetric unit. CeuE-Y288F was in space group $P3_221$ with one protein monomer in the asymmetric unit. Structures were solved starting with apo wild-type CeuE (PDB ID: 3ZKW) as a search model using *MOLREP*.⁴⁴⁸ The overall protein fold for each mutant was very similar to that of wild-type CeuE. There was no well-defined electron density for a small number of surface loops (Table 23).

Table 23: Disordered regions that were not modelled in protein crystal structures of CeuE variant proteins.

CeuE Variant	Residues not modelled
H227L	None
H227A	A/252-254, B/98, C/222-227
Y288F	None
H227L/ Y288F	A/99, 193-194, 223-225, 254, B/ 96-97, 184, 223-225, C/ 221-225
H227A/ Y288F	A/253-254, C/221-225

There was clear electron density confirming the presence of phenylalanine in place of tyrosine in all three Y288F containing structures. However, residue 227, located on a flexible loop lacked electron density in the majority of structures, resulting in limited modelling of H227A and H227L side chains.

Crystallographic Statistics

Crystallographic statistics for crystal structures of CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM are displayed in Table 24 (values in parentheses are for the highest resolution shell). Refinement statistics for these structures are detailed in

Table 25 (Rouge mean square deviations from ideal geometry are listed with target values in parentheses). Crystallographic statistics for all mutant structures are detailed in Table 26 and Table 28 with related refinement statistics detailed in Table 27 and Table 29, respectively. CC(1/2) is defined as the Pearson correlation coefficient for two half datasets. R_{merge} is defined as $100 \times \frac{\sum |I - \langle I \rangle|}{\sum I}$, where I is the intensity of the reflection.

Table 24: Crystallographic statistics for crystal structures of CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM

Structure	CeuE-iron(III)-6-LICAM	CeuE-iron(III)-8-LICAM
Beamline	Diamond 104	Diamond 104
Wavelength (Å)	0.979	0.979
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell parameters (Å)	a= 61.37	a= 42.98
	b= 66.08	b= 55.98
	c= 68.96	c= 140.08
Resolution range (Å)	44.96-2.04	55.98-1.32
Observations	146372	457991
Unique Reflections	18467	77761
Monomers in AU	1	1
Completeness %	100.0 (100.0)	99.4 (76.0)
I/I(σ)	9.6 (1.7)	11.0 (1.3)
CC(1/2)	0.993 (0.583)	0.998 (0.536)
Average Multiplicity	7.9 (8.2)	5.9 (3.1)
Rmerge (%)	20.6 (123.6)	8.3 (85.1)

Table 25: Refinement statistics for crystal structures of CeuE-iron(III)-6-LICAM and CeuE-iron(III)-8-LICAM

Structure	CeuE-iron(III)-6-LICAM	CeuE-iron(III)-8-LICAM
%R_{cryst} (%)	21.3	14.0
Free R factor (%)	26.3	18.1
Bond distances (Å)	0.014 (0.019)	0.020 (0.019)
Bond angles (°)	1.638 (1.997)	1.960 (1.999)
Chiral centres (Å³)	0.094 (0.200)	0.133 (0.200)
Planar groups (Å)	0.007 (0.021)	0.010 (0.021)
Average B value (Å²)	23.4	13.9
Main chain B (Å²)	23.3	11.0
Side chain B (Å²)	19.4	14.8
No. of waters	80	412
Ramachandran Preferred regions %	96.4	97.8
Ramachandran Allowed regions %	3.2	2.2
Ramachandran Outliers %	0.4	0
PDB code	5A5V	5AD1

Table 26: Crystallographic statistics for crystal structures of CeuE-H227L, CeuE- H227L-iron(III)-5-LICAM and CeuE-H227A.

Structure	H227L	H227L-iron(III)-5-LICAM	H227A
Beamline	Diamond 103	Diamond 102	Diamond 103
Wavelength (Å)	0.976	0.979	0.976
Space group	P1	P1	P1
Cell parameters (Å)	a= 56.92	a= 58.70	a= 56.90
	b= 62.56	b= 62.88	b= 62.61
	c= 67.79	c= 69.87	c=67.79
Resolution range (Å)	65.52-1.52	67.13-1.90	66.02-1.33
Observations	631986	254669	598024
Unique Reflections	189953	70811	192548
Monomers in AU	3	3	3
Completeness %	95.4 (94.6)	95.9 (96.0)	99.8 (41.0)
I/I(σ)	9.5 (2.3)	6.2 (1.3)	11.8 (1.4)
CC(1/2)	0.996 (0.710)	0.991 (0.418)	0.998 (0.604)
Average Multiplicity	3.5 (3.5)	3.4 (3.5)	3.4 (3.0)
Rmerge (%)	6.0 (48.6)	10.2 (95.6)	4.1 (63.9)

Table 27: Refinement statistics for crystal structures of CeuE-H227L, CeuE-H227L-iron(III)-5-LICAM and CeuE-H227A.

Structure	H227L	CeuE- H227L- iron(III)-5-LICAM	H227A
%R_{cryst} (%)	18.5	17.1	14.8
Free R factor (%)	22.2	21.5	19.3
Bond distances (Å)	0.024 (0.019)	0.018 (0.019)	0.033 (0.019)
Bond angles (°)	2.338 (1.983)	2.002 (1.850)	2.736 (1.985)
Chiral centres (Å³)	0.169 (0.200)	0.114 (0.200)	0.200 (0.200)
Planar groups (Å)	0.011 (0.021)	0.009 (0.021)	0.013 (0.021)
Average B value (Å²)	25.7	34.8	25.2
Main chain B (Å²)	25.3	32.5	23.4
Side chain B (Å²)	28.0	37.5	27.1
No. of waters	301	153	526
Ramachandran Preferred regions %	96.1	96.3	96.2
Ramachandran Allowed regions %	3.7	3.2	3.7
Ramachandran Outliers %	0.2	0.5	0.1
PDB code	5LWQ	5TCY	5MBQ

Table 28: Crystallographic statistics for crystal structures of CeuE-Y288F, CeuE-H227L/Y288F and CeuE- H227A/Y288F.

Structure	Y288F	H227L/Y288F	H227A/Y288F
Beamline	Diamond 102	Diamond 103	Diamond 103
Wavelength (Å)	0.979	0.979	0.979
Space group	P3 ₂ 21	P1	P1
Cell parameters (Å)	a= 65.52	a= 56.72	a= 56.90
	b= 65.52	b=62.36	b= 62.61
	c= 145.66	c= 67.71	c= 67.79
Resolution range (Å)	56.74-1.40	60.77-1.80	65.80-1.81
Observations	668348	271678	262644
Unique Reflections	71359	78620	77003
Monomers in AU	1	3	3
Completeness %	99.8(85.2)	96.2 (96.4)	99.6 (96.2)
I/I(σ)	19.2 (0.8)	7.4 (1.2)	8.8 (1.6)
CC(1/2)	1.000 (0.239)	0.995 (0.459)	0.996 (0.599)
Average Multiplicity	9.4 (6.4)	3.5 (3.4)	3.4 (3.3)
Rmerge (%)	4.7 (193.9)	7.6 (81.4)	6.6 (67.5)

Table 29: Refinement statistics for crystal structures of CeuE-Y288F, CeuE- H227L/Y288F and CeuE- H227A/Y288F.

Structure	Y288F	H227L/Y288F	H227A/Y288F
%R_{cryst} (%)	13.6	19.4	20.0
Free R factor (%)	16.7	23.9	23.9
Bond distances (Å)	0.031 (0.019)	0.018 (0.019)	0.019 (0.019)
Bond angles (°)	12.532 (1.994)	1.864 (1.983)	1.961 (1.981)
Chiral centres (Å³)	0.213 (0.200)	0.117 (0.200)	0.123 (0.200)
Planar groups (Å)	0.006 (0.020)	0.008 (0.021)	0.009 (0.021)
Average B value (Å²)	24.4	36.0	35.9
Main chain B (Å²)	22.1	34.1	34.7
Side chain B (Å²)	26.7	38.1	37.4
No. of waters	193	121	109
Ramachandran Preferred regions %	96.4	97.1	96.4
Ramachandran Allowed regions %	3.6	2.8	2.9
Ramachandran Outliers %	0	0.1	0.7
PDB code	5LWH	5MBT	5MBU

6.8 Fluorescence Quenching Titration

Experimental Parameters

Intrinsic fluorescence quenching titrations were carried out at room temperature. An excitation slit width of 10 nm and an emission slit width of 20 nm, scan speed of 60 nm/min with an automatic response used. The detector voltage was set at 950 V for CeuE and FepB titrations, and 700 V for VctP and FepB titrations. Each protein solution was excited at 280 nm and the emission spectrum was recorded from 295 nm to 415 nm.

Stock Solutions

Tris(hydroxymethyl)aminomethane, (0.02 mol, 2.4228 g) and sodium chloride (0.075 mol, 4.3830 g) was dissolved in approximately 400 mL distilled water. The pH was adjusted to 7.5 with 2 M HCl. The volume of the solution was made up to 500 mL with distilled water.

A stock solution of 10 mM of each ligand was prepared in DMSO.

Nitrilotriacetic acid trisodium salt (0.1 mmol) was dissolved in 0.0179 mol dm⁻³ standard Fe(NO₃)₃ solution (5.587 mL). This solution was then made up to 10 mL with distilled water, leaving a final solution of 0.01 mol dm⁻³ Fe(NO₃)₃ with 0.01 mol dm⁻³ NTA.

A 12 μM stock solution of ferric-ligand was prepared by pipetting 2 μL of the ligand stock solution and 2 μL of the Fe(NTA) stock solution into 1696 μL of 40 mM Tris-HCl pH 7.5, 150 mM NaCl and thoroughly mixed.

Protein Concentration Determination for Fluorescence Titrations

The protein was diluted to 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl, and the concentration verified by and UV-vis absorbance at 280 nm, using the molar absorbance coefficient predicted by ProtParam⁴⁴³ (CeuE: 15930 M⁻¹ cm⁻¹, FepB: 33920 M⁻¹ cm⁻¹, VctP: 34380 M⁻¹ cm⁻¹).

Standard Binding Method

For each measurement a protein solution of 240 nM in 2000 μL, 40 mM Tris-HCl (pH 7.5) NaCl 150 mM was placed in a 1 cm quartz cuvette, and titrated stepwise over 15 8 μL aliquots, with 2 final 40 μL aliquots of concentrated ferric-ligand stock solution. The titration gave a final ligand concentration of 1.19 μM. After each addition the solution was thoroughly mixed and allowed to rest for 1 minute. Integration of the emission between 310 and 410 nm was normalised and used for plotting and binding constant calculation, using the fitting program DynaFit.

Weak Binding Method

For each measurement, a protein solution of 2.4 μM in 2000 μL , 40 mM Tris-HCl (pH 7.5) NaCl 150 mM was placed in a 1 cm quartz cuvette, and titrated stepwise over 15 \times 8 μL aliquots, with 2 final 40 μL aliquots of 120 μM ferric-ligand stock solution. The titration gave a final ligand concentration of 11.9 μM . After each addition the solution was thoroughly mixed and allowed to rest for 1 minute. Integration of the emission between 310 and 410 nm was normalised and used for plotting and binding constant calculation, using the fitting program DynaFit.

VctP Tight Binding Method

For each measurement, a protein solution of \sim 6 nM in 40 mM Tris-HCl (pH 7.5) NaCl 150 mM was placed in a 1 cm quartz cuvette. The solution was subjected to mechanical stirring, and titrated stepwise with 24 \times 0.25 μL aliquots of 6 μM ferric ligand stock solution every 4 minutes using a micro-dosing DOSTAL DOSY and stirred for 1 minutes before spectra were recorded. Spectra were integrated between 305-380 nm and the data normalised to produce a binding curve. Data were fitted with Dynafit, allowing variable VctP concentration, and dissociation constants calculated.

CeuE Fluorescence Control Experiments

Spectra were recorded over 75 minutes to gauge whether CeuE fluorescence might be quenched over time, or *via* repeat exposure to the UV radiation set at the excitation wavelength of 280 nm. Spectral scans lasting 2 minutes each were recorded every 5 minutes over 75 minutes. The fluorescence peak was integrated between 310 nm and 410 nm and the data normalised. Normalised fluorescence was plotted against time, showing that fluorescence intensity was no lower than 94% of the original fluorescence intensity throughout the full 75 minutes. Fluorescence emission did not decrease significantly when CeuE was subjected to repeat exposure to the excitation wavelength of 280 nm. This indicates that the stability of CeuE under the titration conditions was adequate.

To ensure that any fluorescence quenching observed was a direct consequence of the addition of iron(III)-ligand and not due to dilution or mixing, a control titration with buffer (40 mM Tris-HCl pH 7.5, 150 mM NaCl) was carried out. It was observed that the fluorescence of CeuE decreased to around 83% of the intensity of the initial signal with the addition of 200 μL (10% dilution) buffer. Employing a linear dilution correction shows a decrease to around 92% of the original signal. Although there is a small decrease observed, it was decided that this would not impact significantly on the result of the fluorescence quenching titrations. It was decided that for a calculated dissociation constant to be

significant, the final normalised emission must be below 0.4- significantly lower than is observed for dilution effects. All spectra and plotted data is included in Appendix 1.

VctP Fluorescence Control Experiments

VctP in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with 40 mM Tris-HCl pH 7.5, 150 mM NaCl to explore whether the protein fluorescence is quenched *via* dilution or mixing. It was observed over 24 0.25 μ L aliquots of 40 mM Tris-HCl pH 7.5 every 4 minutes using a micro-dosing DOSTAL DOSY and constant overhead stirring, that fluorescence was at around 75% of the original signal. It was decided that titrations should not exceed 24 points, to minimise the decrease in fluorescence of the protein over time. Spectra and plotted data are included in Appendix 1. For the fluorescence quenching titration, 24 aliquots of 0.25 μ L of 6 μ M iron(III)-5-LICAM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl were titrated into 0.012 μ M VctP in 40 mM Tris-HCl pH 7.5, 150 mM NaCl every 4 minutes using a micro-dosing DOSTAL DOSY with constant overhead stirring. The resulting spectra were integrated between 305 nm and 380 nm and the data normalised. A fitting curve and dissociation constant were calculated for each run using Dynafit.³²⁶

Ligand Absorbance Control Experiments

The absorbance of iron(III)-5-LICAM in a 1:1 ratio in 40 mM Tris-HCl pH 7.5, 150 mM NaCl buffer was monitored over a range of concentrations to establish the absorbance of the complex at 280 nm. If ligand absorbance exceeded 0.1, it was deemed unacceptable due to significant contribution of primary inner filter effect. The UV/Vis spectra are shown in Figure 141. Plotting absorbance at 280 nm *vs* concentration produced a linear plot Figure 142.

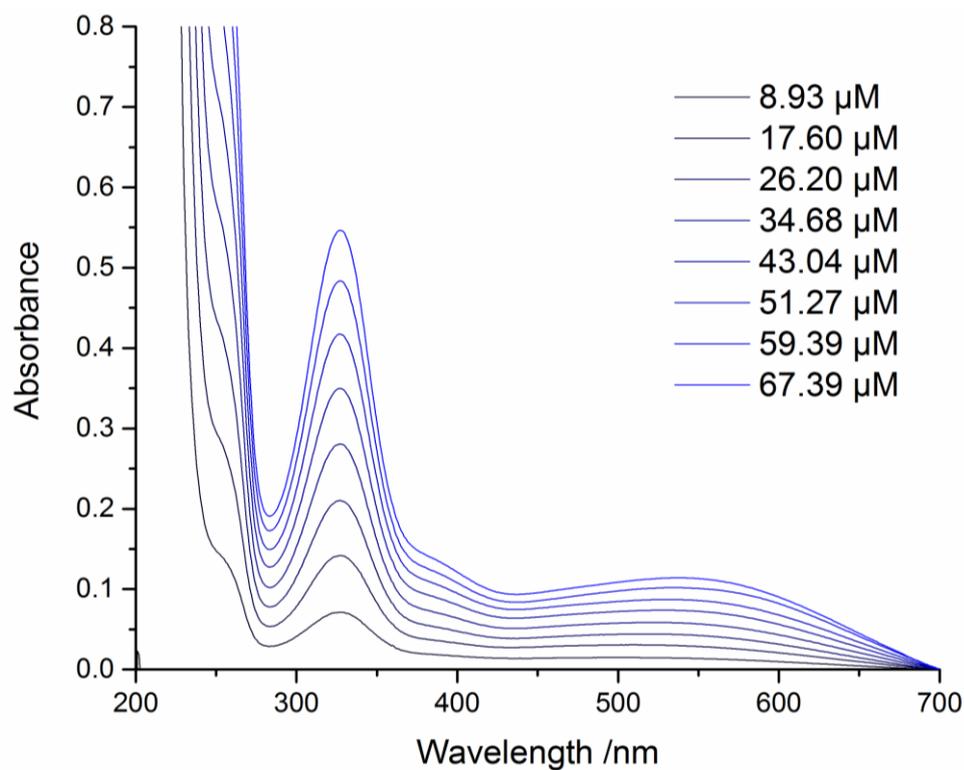


Figure 141: UV/Vis spectra of iron(III)-5-LICAM over a range of concentrations in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

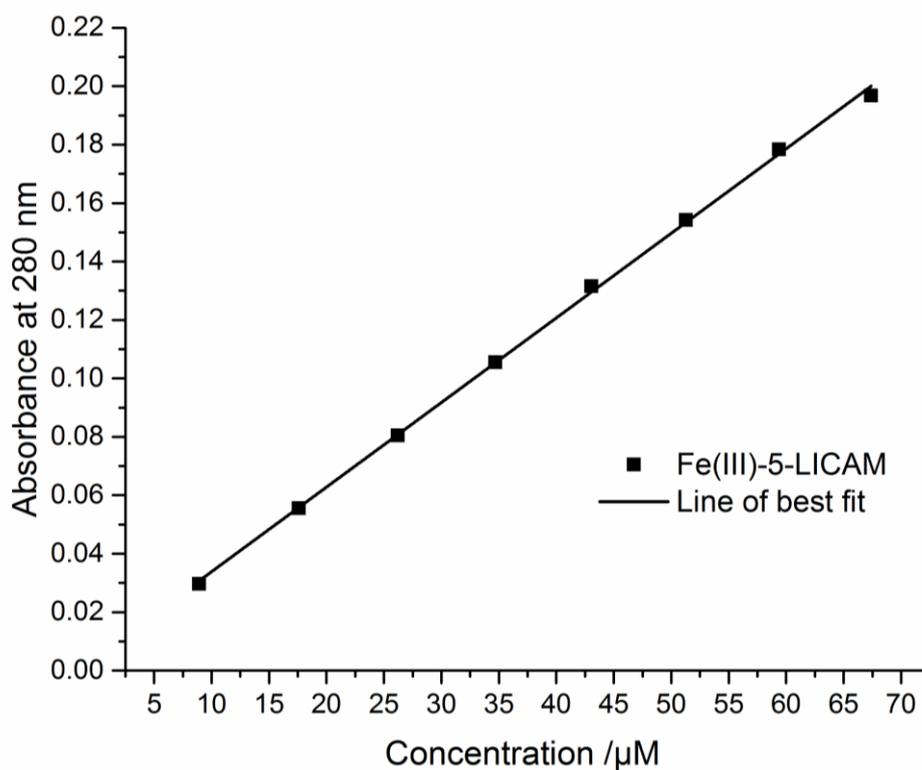


Figure 142: Absorbance of iron(III)-5-LICAM at 280 nm plotted against concentration. $y = 0.0029x + 0.0049$. $R^2 = 0.9992$

The equation of the line of best fit was used to establish the concentration at which absorbance of 0.1 is exceeded. The maximum ligand concentration was calculated as 67.3 μM .

The study was repeated for iron(III)-6-LICAM as well as iron(III)-8-LICAM.

The UV spectra for iron(III)-n-LICAM (n= 6, 8) in a 1:1 ratio were recorded at varying concentrations in 40 mM Tris-HCl pH 7.5, 150 mM NaCl buffer until an absorbance of 0.2 was exceeded (Figure 143 and Figure 144). The resulting data was plotted to produce a linear relationship between absorbance at 280 nm and concentration (Figure 145 and Figure 146).

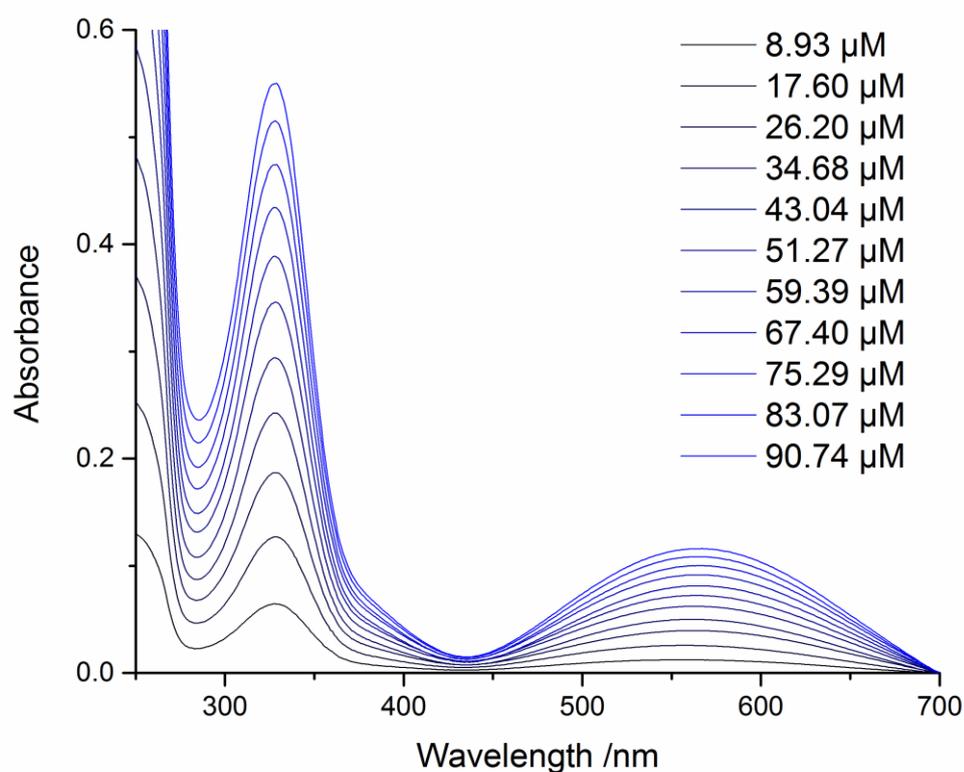


Figure 143: UV/visible spectra for iron(III)-6-LICAM in a 1:1 ratio at varying concentrations.

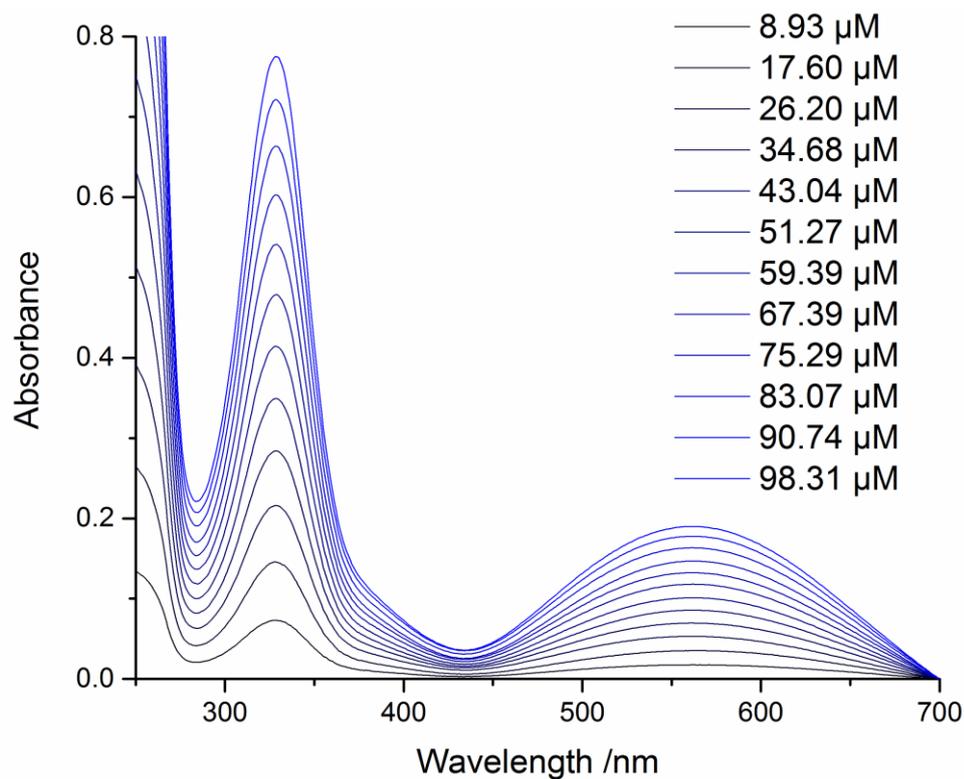


Figure 144: UV/visible spectra of iron(III)-8-LICAM in a 1:1 ratio at varying concentrations.

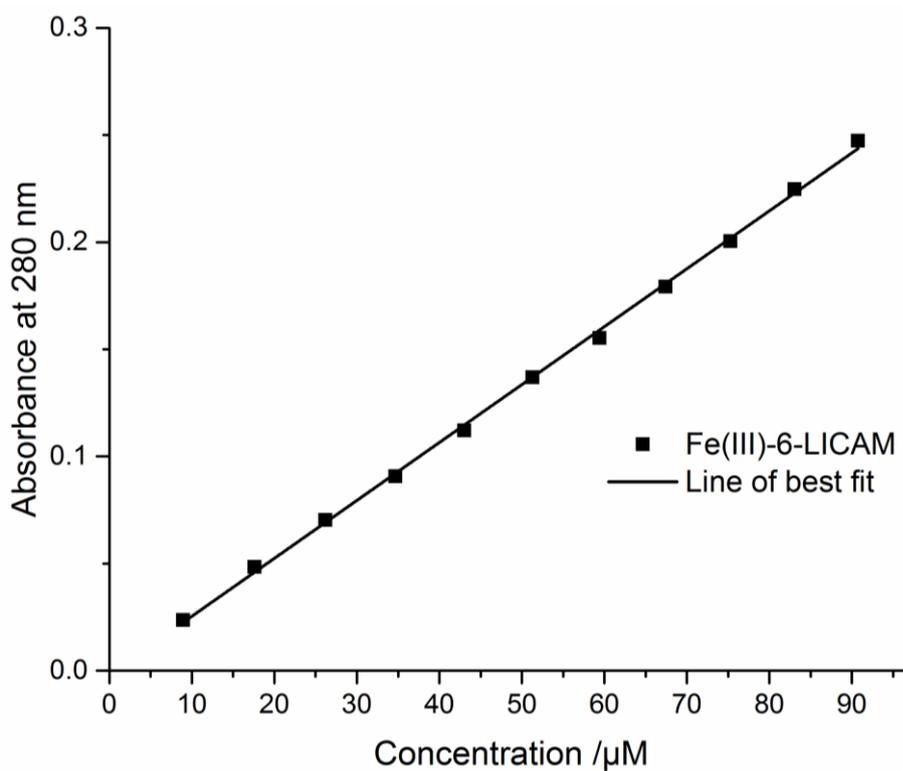


Figure 145: Plot of iron(III)-6-LICAM absorbance at 280 nm vs concentration in μM, showing good adherence to the Beer Lambert law. Equation of the line of best fit: $y = 0.0027x - 0.0016$ $R^2 = 0.9991$.

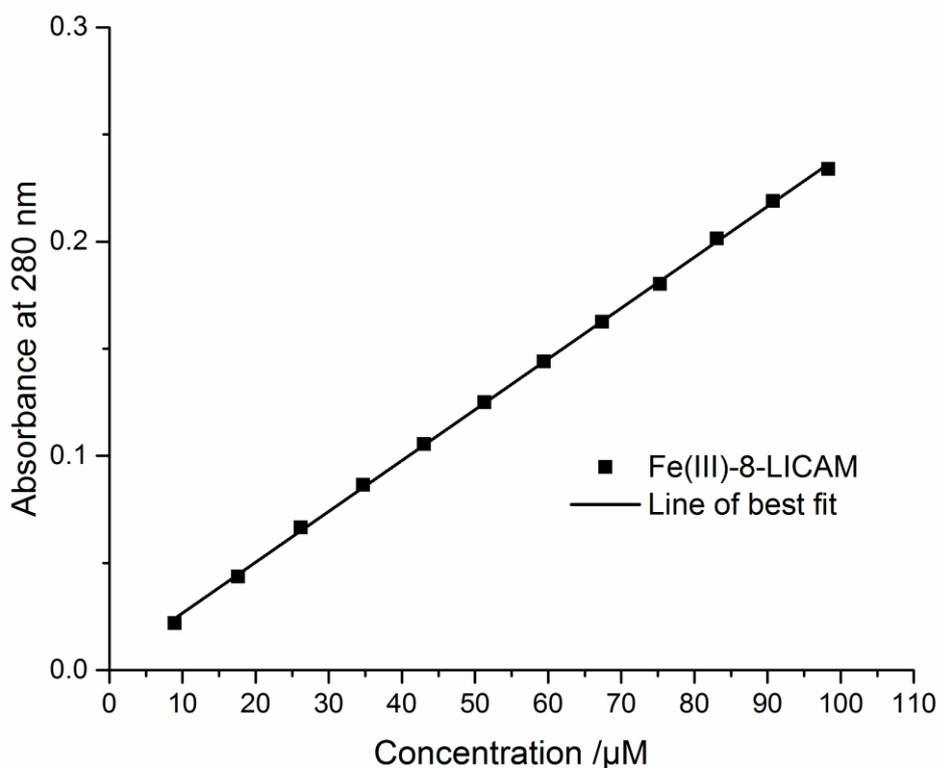


Figure 146: Linear relationship between iron(III)-8-LICAM absorbance at 280 nm vs concentration in μM . Equation of the line of best fit: $y = 0.0024x + 0.0029$ $R^2 = 0.9996$.

In accordance with the Beer-Lambert law, the resulting lines of best fit could be used to determine that at absorbance of 0.1 the concentration of iron(III)-6-LICAM was $37.34 \mu\text{M}$ and iron(III)-8-LICAM was $41.07 \mu\text{M}$ meaning these were the maximum concentrations permissible for this system.

UV spectra of iron(III)-enterobactin were used to confirm the upper iron(III)-enterobactin concentration that could be used in fluorescence quenching titration experiments (Figure 147). The absorbance at 280 nm was plotted against concentration to produce Figure 148. The equation of the line of best fit could then be used to calculate the concentration of iron(III)-enterobactin that provides an absorbance of 0.1, of $18.40 \mu\text{M}$. As five equivalents of ligand are required to provide adequate saturation of the CeuE binding pocket, the maximum protein concentration that could be used in fluorescence quenching titrations of enterobactin was determined as $3.68 \mu\text{M}$.

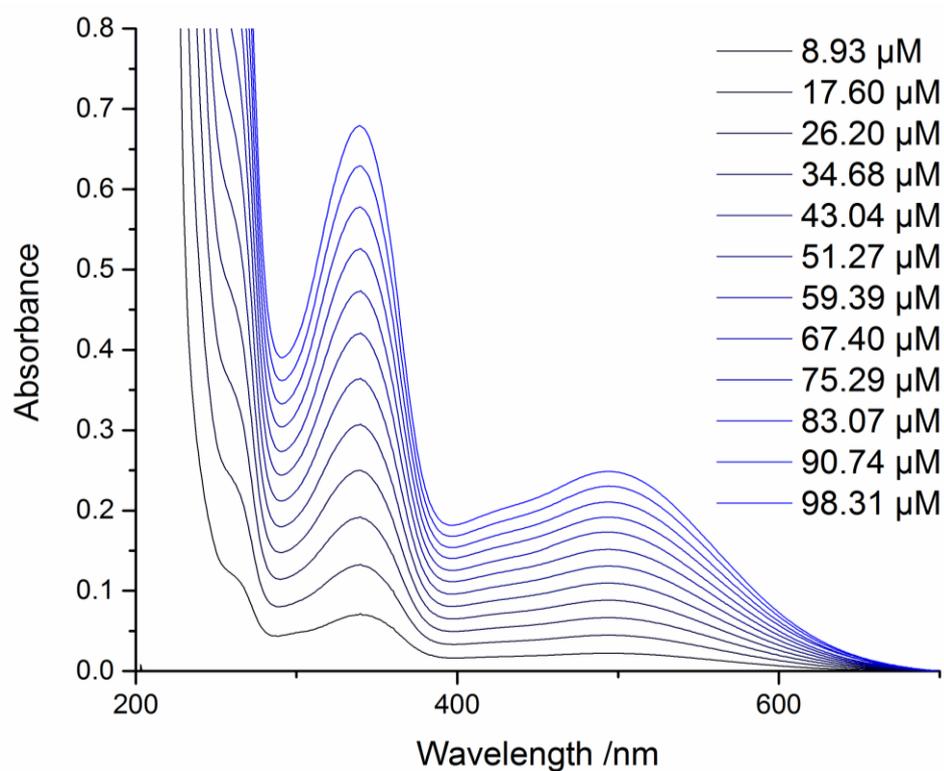


Figure 147: UV/visible spectra of iron(III)enterobactin in a 1:1 ratio at varying concentrations.

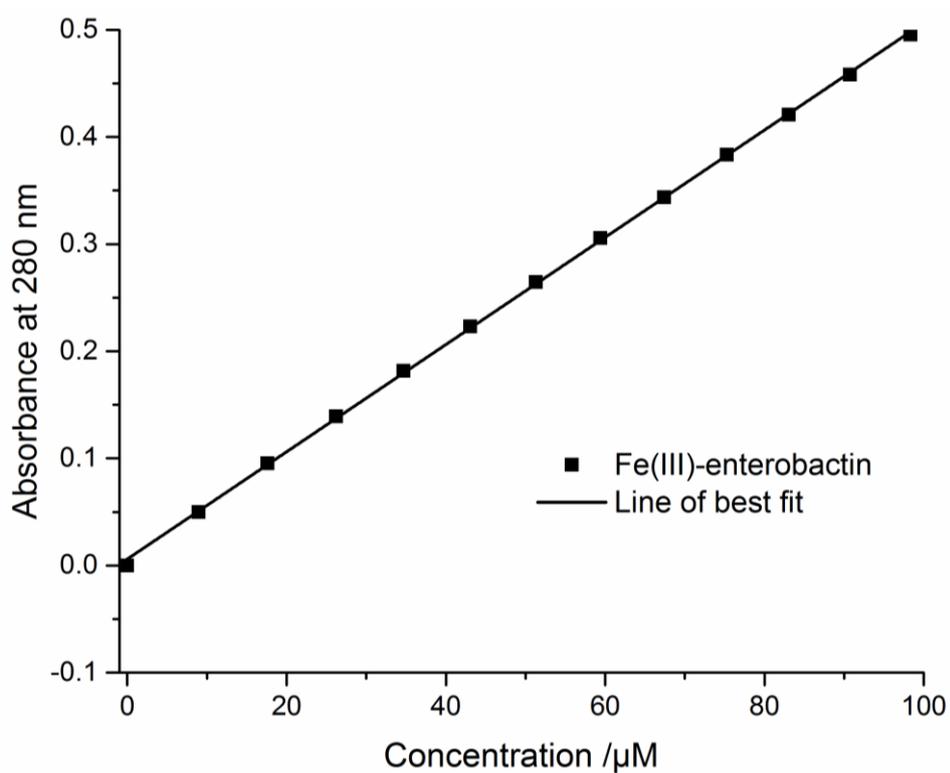


Figure 148: Absorbance at 280 nm of iron(III)-enterobactin vs concentration. $y = 0.005x + 0.008$ $R^2 = 0.9999$.

Averages and Error Calculation

Each titration was performed in triplicate. The average was calculated and an overall error calculated using the uncertainty in average method, classically used instead of standard deviation when a data set is small (Figure 149).⁴⁵³

$$\text{Reported Value} = X_{avg} \pm \Delta X_{avg}$$

$$\text{Weighted Average } (X_{avg}) = \frac{\sum \left(\frac{a_i}{\sigma_i^2} \right)}{\sum \left(\frac{1}{\sigma_i^2} \right)}$$

$$\text{Range (R)} = X_{max} - X_{min}$$

$$\text{Uncertainty } (\Delta X) = \frac{R}{2}$$

$$\text{Uncertainty in Average } (\Delta X_{avg}) = \frac{\Delta X}{\sqrt{N}}$$

Error calculation

Figure 149: Equations used for calculation of weighted average and errors for K_d values. a_i = raw data value, σ_i = raw error, N = number of data values used⁴⁵³

6.9 Circular Dichroism

Protein Folding Experiments

VctP was diluted to 0.4 mg mL⁻¹ in 5 mM Tris-HCl pH 8. FepB was diluted to 0.4 mg mL⁻¹ in 5 mM Tris-HCl pH 7.5. Spectra were recorded from 190-240 nm with data pitch 0.5 nm in continuous scanning mode with speed of 100 nm min⁻¹ and response of 2 seconds. The bandwidth was set to 2 nm, with a path length of 1cm. The spectra were recorded 5 times and averaged.

Ligand Binding Experiments

The spectra were recorded from 300-700 nm with data pitch 0.5 nm in continuous scanning mode with speed of 100 nm min⁻¹ and response of 2 seconds. The bandwidth was set to 2 nm, with a path length of 1cm. The spectra were recorded 5 times and averaged.

Proteins including wild type CeuE and all mutants, FepB and VctP were diluted to 2.5×10^{-3} M in 0.11 M Tris-HCl pH 7.5, 150 mM NaCl buffer. An iron(III)-ligand stock solution containing equimolar NTA was made at a concentration of 5×10^{-4} M. This was done by adding 10 μ L of both 10mM ligand in DMSO and 10 μ L 10mM Fe(III)NTA in H₂O to 180 μ L 0.11M Tris-HCl pH 7.5, 150 mM NaCl buffer.

Spectra were recorded by making a final solution comprising 880 μL 0.11 M Tris-HCl pH 7.5, 150 mM NaCl buffer, 100 μL (iron(III)-ligand NTA stock solution and 20 μL 2.5×10^{-3} M CeuE, resulting in a 1:1 ratio of ligand to protein at 5×10^{-5} M.

The spectrum of buffer was subtracted from all spectra including the spectra for free ligand and apo protein.

CeuE Iron(III)-5- β -D-glucosyl-4-LICAM and Iron(III)-5- β -D-glucosyl-5-LICAM Circular Dichroism Titrations

A 980 μL solution of iron(III)-5- β -D-glucosyl-n-LICAM ($n = 4, 5$) containing equimolar NTA was made at a concentration of 5×10^{-5} M as above. CD spectra were run over a series of additions of CeuE. Additions were made in 0.2 equivalents of protein until 3 equivalents were achieved, after which 1 equivalent additions were made. The final concentration of CeuE was 3×10^{-4} M, equal to a 1:6 ligand to protein ratio, at which point the addition of further protein did not change the spectral signal observed. The ellipticity value in mdeg was recorded at 330 nm for each spectrum, and plotted against CeuE concentration to obtain a binding curve. The binding constant was calculated using Dynafit Software.³²⁶ The titrations were repeated in triplicate and the average and an overall error calculated using the uncertainty in average method as detailed for the above fluorescence titrations.

6.10 Job Plot

5- β -D-Glucosyl-n-LICAM ($n = 4, 5$) were analysed for their iron(III) binding stoichiometries *via* a Job plot method established in the literature.⁹⁰ 5- β -D-glucosyl-n-LICAM ($n = 4, 5$) was combined with varying concentrations of iron(III) to establish preferred metal to ligand binding ratios. The ratios spanned 100% ligand to 100% iron(III), with 5% intervals. Between 60:40 and 50:50 ligand to iron(III) ratios, 2% intervals were used. This protocol resulted in 24 samples over the full range. Stock solutions of 10 mM Fe(III)NTA in H_2O and 10 mM 5- β -D-glucosyl-n-LICAM ($n = 4, 5$) in DMSO were used in the necessary ratios totalling 200 μL to make up a 2 mL solution of each ratio in 1800 μL 0.11 M Tris-HCl pH 7.5, 150 mM NaCl, resulting in an overall iron(III)-ligand concentration of 400 nM. DMSO was kept at 5% v/v for all final solutions. A UV/Vis spectrum from 400 nm to 700 nm was run for each solution after 1 hour of equilibration. The spectra were then rerun after 7 days of equilibration. λ_{max} values were observed at 498 nm and 552 nm for iron(III)-5- β -D-glucosyl-4-LICAM and 495 nm and 555 nm for iron(III)-5- β -D-glucosyl-5-LICAM. The absorbance at these wavelengths was plotted against ligand to protein ratio, and the maximum absorbance for each wavelength recorded at the relevant ligand: protein ratio.

6.11 Analytical HPLC

Purchased enterobactin, salmochelin S1 and salmochelin S4 were examined by analytical HPLC. For enterobactin, the method was based on several reports from literature.^{86, 87, 454}

Enterobactin was dissolved in water containing 0.1% TFA and filtered through a PTFE syringe filter with pore size of 0.22 μm . The column oven temperature was set at 35 $^{\circ}\text{C}$. For enterobactin, the method comprised of a 10-80% gradient of acetonitrile in water with 0.1% TFA throughout, over 35 minutes at a flow rate of 1 mL/min. After 35 minutes the gradient was reduced back to 10% acetonitrile in water with 0.1% TFA for a further 10 minutes.

For salmochelin S1 and salmochelin S4, the method was based on a reports from the literature.²³⁵ Salmochelin S1 and salmochelin S4 were dissolved in water containing 0.1% formic acid and filtered through a PTFE syringe filter with pore size of 0.22 μm . The column oven temperature was set at 35 $^{\circ}\text{C}$. For both samples, the method comprised of a 6-40% gradient of acetonitrile in water with 0.1% formic acid throughout, over 30 minutes at a flow rate of 1 mL/min. After 30 minutes the gradient was increased from 40% to 90% over 15 minutes. The gradient was then reduced back to 6% acetonitrile in water with 0.1% formic acid for a further 9 minutes.

The sample of salmochelin S4 was stored at room temperature for 24 hours before the sample was rerun using the same HPLC setup and conditions.

6.12 ESI-MS for Natural Salmochelins

Enterobactin

HRMS: Calcd. $[\text{M}-\text{H}]^{-}$ ($\text{C}_{30}\text{H}_{26}\text{N}_3\text{O}_{15}$) $m/z = 668.1369$; Obs. $[\text{M}-\text{H}]^{-}$ ($\text{C}_{30}\text{H}_{26}\text{N}_3\text{O}_{15}$) $m/z = 668.1369$, Mean err 0.3 ppm. Calcd. $[\text{M}+\text{OH}]^{-}$ ($\text{C}_{30}\text{H}_{28}\text{N}_3\text{O}_{16}$) $m/z = 686.1475$; Obs. $[\text{M}+\text{OH}]^{-}$ ($\text{C}_{30}\text{H}_{28}\text{N}_3\text{O}_{16}$) $m/z = 686.1475$, Mean err 0.1 ppm.

Salmochelin S1

HRMS: Calcd. $[\text{M}-\text{H}]^{-}$ ($\text{C}_{26}\text{H}_{29}\text{N}_2\text{O}_{16}$) $m/z = 625.1523$; Obs. $[\text{M}-\text{H}]^{-}$ ($\text{C}_{26}\text{H}_{29}\text{N}_2\text{O}_{16}$) $m/z = 625.1510$, Mean err 1.0 ppm.

Appendix 1: Fluorescence Quenching Titration Data and Controls

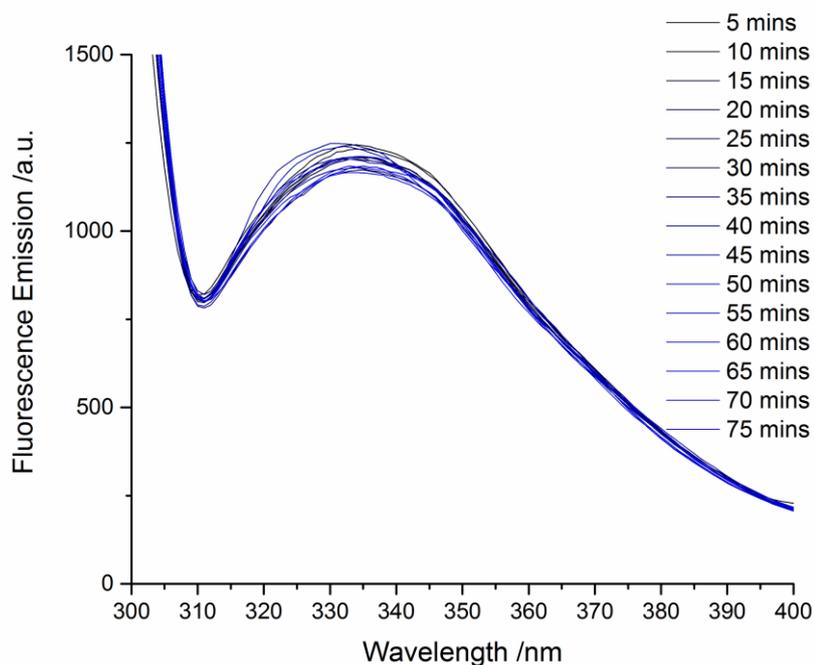


Figure 150: Fluorescence spectra of CeuE recorded over 75 minutes.

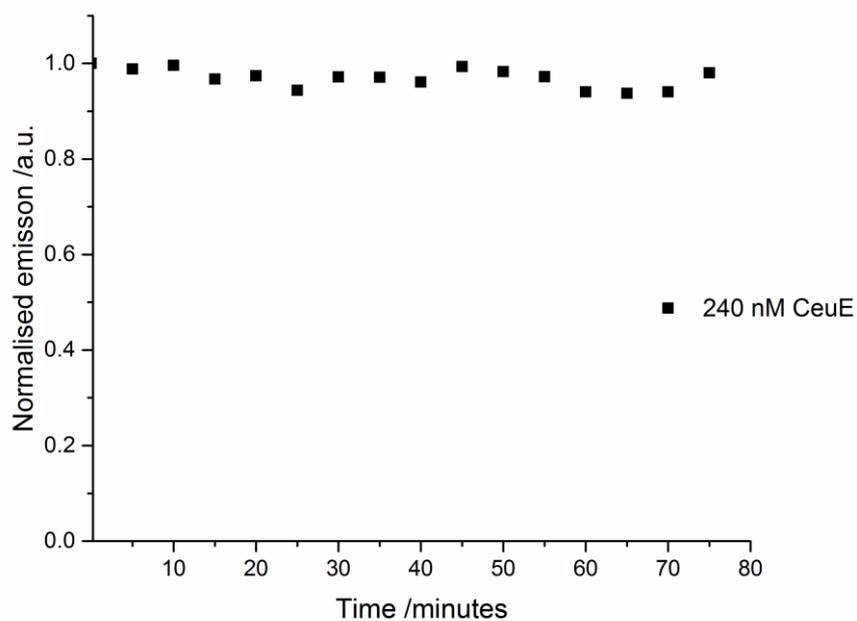


Figure 151: Fluorescence emission of CeuE monitored over time, showing adequate stability of the protein over 75 minutes of exposure to light of the excitation wavelength of 280 nm.

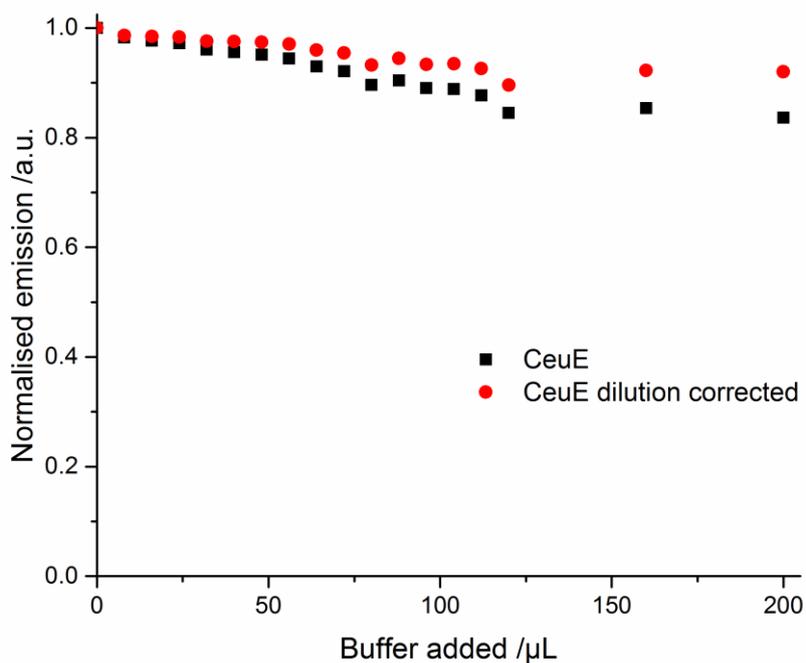


Figure 152: Decrease in fluorescence observed when CeuE is titrated with 40 mM Tris-HCl pH 7.5, 150 mM NaCl, shown as both uncorrected normalised data and a linear dilution correction across the titration.

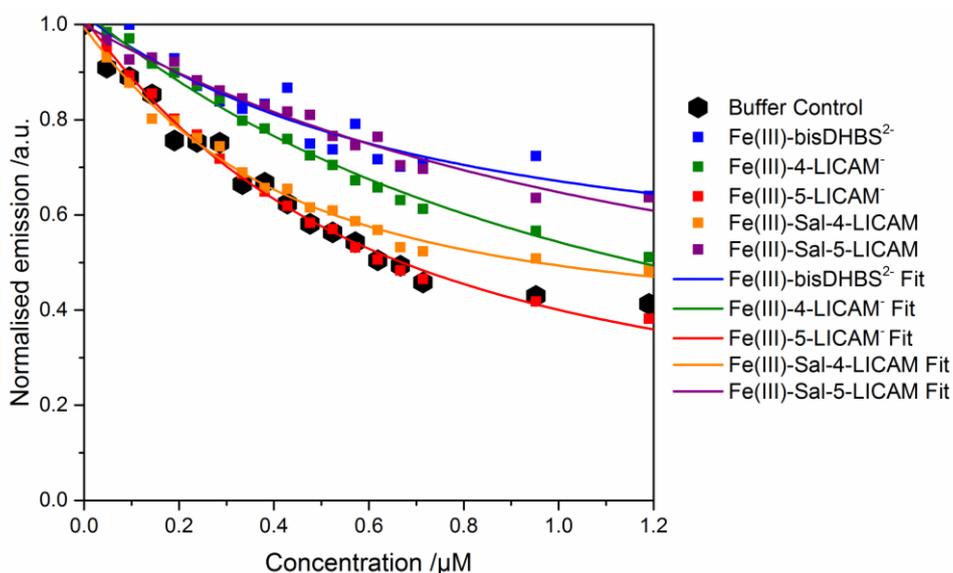


Figure 153: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for FepB-Fe(III)-bisDHBS, FepB-Fe(III)-n-LICAM ($n = 4, 5$) and FepB-Fe(III)Sal-n-LICAM ($n = 4, 5$) and their associated non-linear regression fitting data from Dynafit. 240 nM FepB in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM Fe(III)-ligand⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

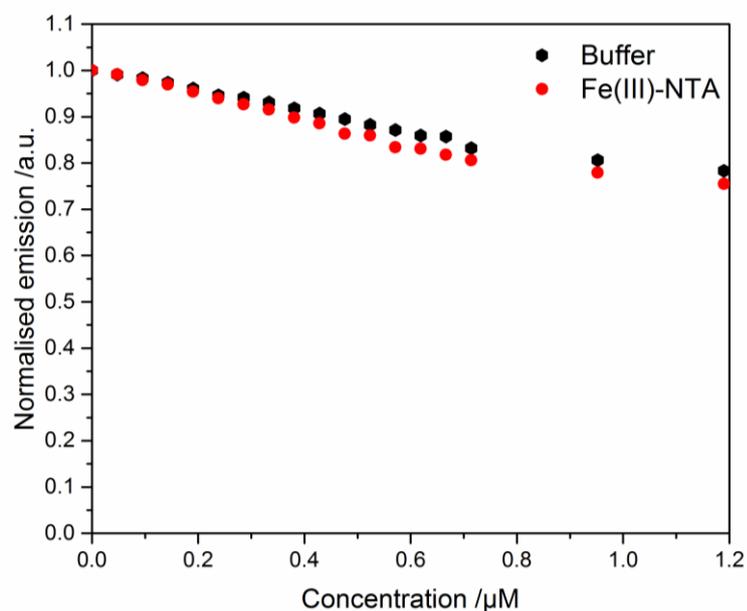


Figure 154: Fluorescence quenching titration curves, as a plot of normalised fluorescence emission vs ligand concentration in μM for VctP titrated with buffer and Fe(III)-NTA. VctP was at concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl and was titrated with aliquots of 12 μM Fe(III)-NTA in 40 mM Tris-HCl pH 7.5, 150 mM NaCl, and the equivalent volume of 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

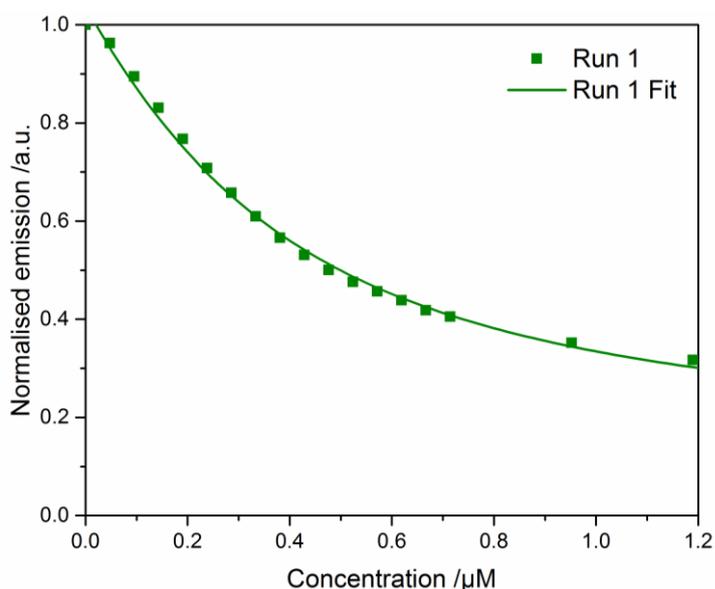


Figure 155: Fluorescence quenching titration curves for VctP-Fe(III)-enterobactin, as a plot of normalised emission vs ligand concentration in μM for VctP-Fe(III)-enterobactin and their associated non-linear regression fitting data from Dynafit. 240 nM VctP in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM Fe(III)-enterobactin⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

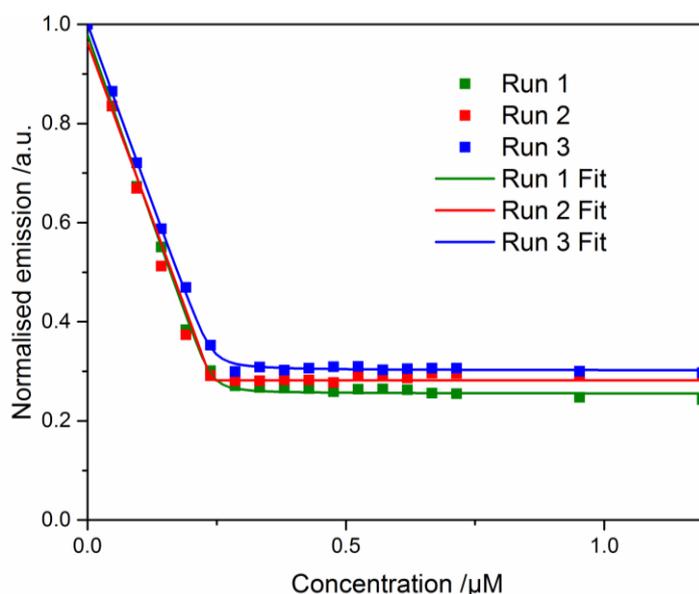


Figure 156: Fluorescence quenching titration curve for VctP-Fe(III)-bisDHBS Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μM for VctP-Fe(III)-bisDHBS and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM Fe(III)-bisDHBS⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

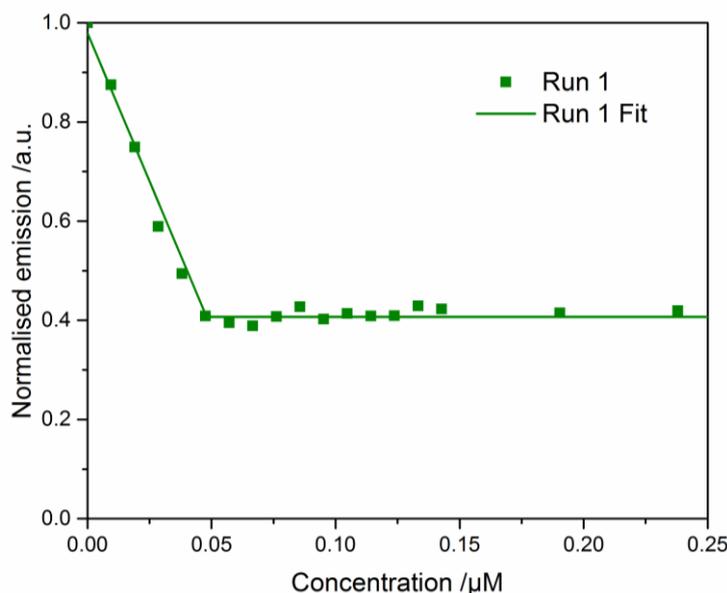


Figure 157: Fluorescence quenching titration curve for VctP-Fe(III)-bisDHBS Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μM for VctP-Fe(III)-bisDHBS and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 48 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 2.4 μM Fe(III)-bisDHBS⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

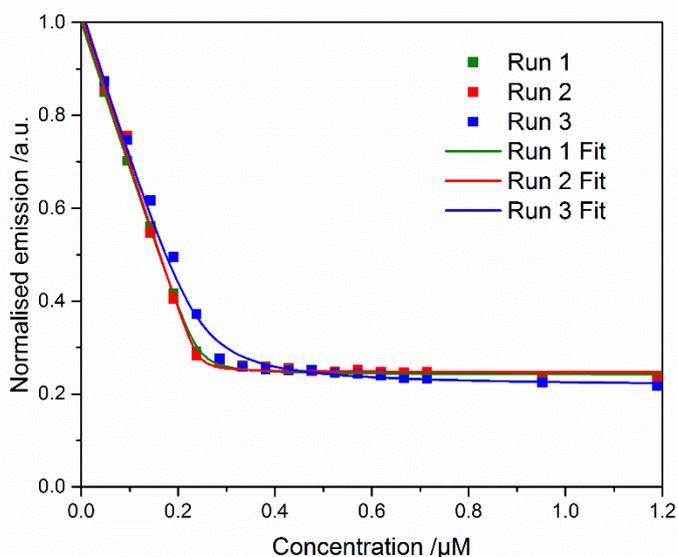


Figure 158: Fluorescence quenching titration curve for VctP-Fe(III)-5-LICAM Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μM for VctP-Fe(III)-5-LICAM and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM Fe(III)-5-LICAM⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

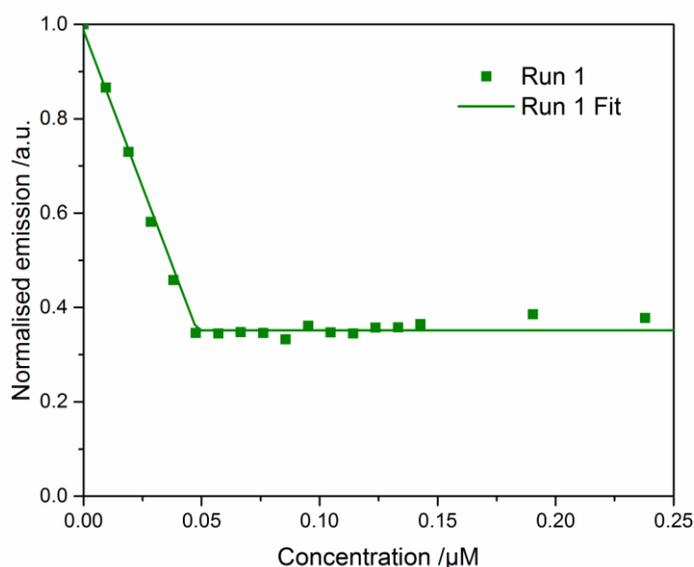


Figure 159: Fluorescence quenching titration curve for VctP-Fe(III)-5-LICAM Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μM for VctP-Fe(III)-5-LICAM and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 48 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 2.4 μM Fe(III)-5-LICAM⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

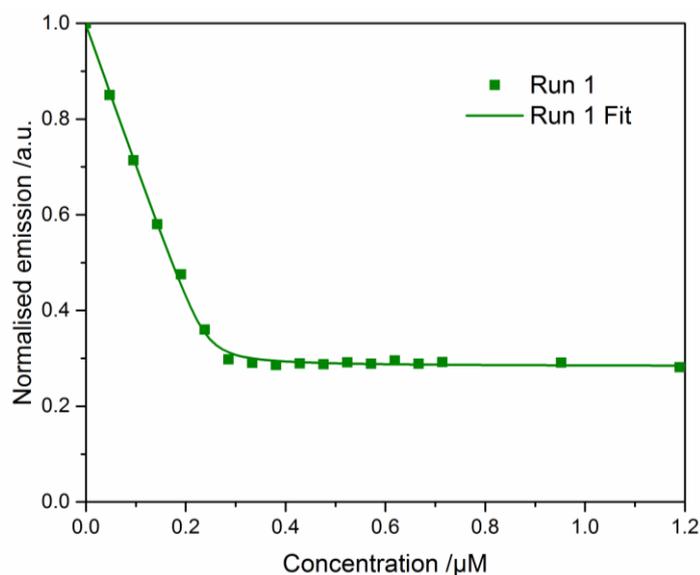


Figure 160: Fluorescence quenching titration curve for VctP-Fe(III)-4-LICAM Fluorescence quenching titration curve, as a plot of normalised fluorescence emission vs ligand concentration in μM for VctP-Fe(III)-4-LICAM and the associated non-linear regression fitting data from Dynafit. VctP at a concentration of 240 nM in 40 mM Tris-HCl pH 7.5, 150 mM NaCl was titrated with aliquots of 12 μM Fe(III)-4-LICAM⁻ in 40 mM Tris-HCl pH 7.5, 150 mM NaCl.

[task]

data = equilibria

task = fit

[mechanism]

PBP + FeL <==> PBP.FeL : Kd dissociation

[constants]

Kd = **V** ?

[concentrations]

CeuE = **W** uM / approx 5.00 mgmL⁻¹ 3 uL

[responses]

CeuE = **X** ?, CeuE.FeLD = **Y** ? ; Initial fluorescence divided by amount of CeuE i.e. **X**=(1/**W**) ; End Fluorescence divided by amount of FeLD.CeuE i.e. **Y**=(**Z**/**W**)

[data]

variable FeLD

file ./Data/PBP/FeL/Run1/Data1.txt

[output]

directory ./Data/PBP/FeL/Output1

[end]

Figure 161: Dynafit script used for a general fluorescence titration. **V** is edited with an estimate of an approximate dissociation constant. **W** is edited for protein concentration used, **X**= 1/**W**, **Y**= **Z**/**W** where **Z** is the final normalised emission for the titration.

Abbreviations

Iron(III)-Siderophore Complex Nomenclature

For simplicity, all iron(III) complexes have been reported in a simplified format: PBP-iron(III)-ligand. This naming system does not necessarily reflect the true stoichiometry of iron(III) to ligand, and there may be multiple iron(III) to ligand ratios in equilibrium, with associated charges on each overall complex. For complexes with proteins, the overall charge of the system cannot be accurately estimated due to the large number of charged protein sidechains. ‘PBP’ refers to CeuE, FepB or VctP. The most likely formulae are listed for each term, but are not exhaustive and do not consider the potential for solvent coordination:

Simplified name	Possible true complex(es)
Iron(III)-n-LICAM (n= 4, 5, 6, 8)	[Fe(III)-n-LICAM] ⁻ [Fe(III)] ₂ {n-LICAM} ₃ ⁶⁻
PBP-iron(III)-n-LICAM (n= 4, 5, 6, 8)	PBP-[Fe(III)-n-LICAM] ⁻
Iron(III)-bisDHBS	[Fe(III)-bisDHBS] ²⁻ [Fe(III)] ₂ {bisDHBS} ₃ ⁹⁻
PBP-iron(III)-bisDHBS	PBP-[Fe(III)-bisDHBS] ²⁻
Iron(III)-enterobactin	[Fe(III)-enterobactin] ³⁻
PBP-iron(III)-enterobactin	PBP-[Fe(III)-enterobactin] ³⁻
Iron(III)-MECAM	[Fe(III)-MECAM] ³ [Fe(III)] ₂ {MECAM} ₂ ⁶⁻
PBP-iron(III)-MECAM	PBP-[Fe(III)-MECAM] ³⁻ PBP ₂ -[Fe(III)] ₂ {MECAM} ₂ ⁶⁻
Iron(III)-Sal-n-LICAM (n= 4, 5)	[Fe(III)-Sal-n-LICAM] ⁻ [Fe(III)] ₂ {Sal-n-LICAM} ₃ ⁶⁻
PBP-iron(III)-Sal-n-LICAM (n= 4, 5)	PBP-[Fe(III)-Sal-n-LICAM] ⁻

°	degrees
°C	degrees Celcius
Å	Angstrom
ABC	ATP-binding cassette
<i>A. fulgidus</i>	<i>Archaeoglobus fulgidus</i>
Ala	alanine
<i>A. muciniphilia</i>	<i>Akkermansia muciniphilia</i>
Ar	aromatic
aq	aqueous
Arg	arginine
Asn	asparagine
ATIR	attenuated total reflectance infrared
atm	atmosphere
ATP	adenosine triphosphate
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
BnBr	benzyl bromide
c (prefix)	centi
Calcd.	calculated
CD	circular dichroism
CDCl ₃	deuterated chloroform
CDI	1,1'-carbonyldiimidazole
CD ₃ OD	deuterated methanol
CHN	carbon hydrogen nitrogen
<i>C. jejuni</i>	<i>Campylobacter jejuni</i>
COSY	correlation spectroscopy
<i>C. trachomatis</i>	<i>Chlamydia trachomatis</i>
d ₆ -DMSO	deuterated dimethyl sulfoxide
Da	dalton
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
deg	degree
DEPT	Distortionless Enhancement by Polarisation Transfer
DHB	2,3-dihydroxybenzoic acid
DHBS	2,3-dihydroxybenzylserine
DHP	2,3-dihydroxyphenylthiazolanyl
DIPEA	N-N'-diisopropylethylamine
DMF	dimethylformamide

DNA	deoxyribonucleic acid
DTT	dithiothreitol
DtxR	diphtheria toxin regulator
E _{1/2}	half cell potential
<i>E. coli</i>	<i>Escherichia coli</i>
<i>E. faecium</i>	<i>Enterococcus faecium</i>
Ent	enterobactin
<i>E. raffinosus</i>	<i>Enterococcus raffinosus</i>
Err.	error
ESI-MS	electrospray ionisation mass spectrometry
Et ₃ N	triethylamine
EtOAc	ethyl acetate
EtOH	ethanol
FAD	flavin adenine dinucleotide
Flu	fluvibactin
Fur	ferric uptake regulator
x g	relative centrifugal force (times gravity)
g	grams
Gly	glycine
Gln	glutamine
<i>G. thermoglucosidasius</i>	<i>Geobacillus thermoglucosidasius</i>
<i>H. influenzae</i>	<i>Haemophilus influenzae</i>
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
His	histidine
HMBC	heteronuclear multiple-bond correlation
HMQC	heteronuclear multiple quantum correlation
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
Hz	hertz
Ile	isoleucine
IPTG	isopropyl β-D-1-thiogalactopyranoside
J	Joule
k (prefix)	kilo
K	Kelvin
K _d	dissociation constant
K _f	formation constant

λ_{\max}	wavelength of maximum absorbance
L	litre
L	ligand in complexation
LB	lysogeny broth
Leu	leucine
LICAM	linear catecholamide
LMCT	ligand to metal charge transfer
Lys	lysine
M (prefix)	mega
m	metre
M	molar
m (prefix)	milli
MECAM	[1,3,5-N,N',N''-tris-(2,3-dihydroxybenzoyl)- triaminomethylbenzene]
MeOH	methanol
min	minute
mol	mole
M.P	melting point
m/z	mass/charge
n (prefix)	nano
NADH	reduced nicotinamide adenine dinucleotide
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
NSPD	norspermidine
NTA	nitriloacetic acid
NVT	Constant number volume and temperature
Obs.	observed
OD ₆₀₀	optical density at 600 nm
p (prefix)	pico
Pa	pascal
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PBP	periplasmic binding protein
PCR	polymerase chain reaction
PDB ID	Protein Data Bank identifier code
PEG	polyethylene glycol
Phe	phenylalanine
ppm	parts per million

PTFE	polytetrafluoroethylene
R _f	retardation factor
r.m.s.d	root mean square deviation
RNA	ribonucleic acid
rpm	revolutions per minute
s	seconds
SSM	secondary structure matching
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SDS PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
<i>S. enterica</i>	<i>Salmonella enterica</i>
Ser	serine
<i>S. pneumoniae</i>	<i>Streptococcus pneumoniae</i>
STEC	shiga toxin-producing <i>E. coli</i>
<i>S. typhimurium</i>	<i>Salmonella typhimurium</i>
TE	tris-EDTA
TFA	trifluoroacetic acid
THF	tetrahydrofuran
Thr	threonine
TRENCAM	β,β',β''-triaminotriethylamine catecholamide
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
Trp	tryptophan
Tyr	tyrosine
μ (prefix)	micro
UV-visible	ultra violet-visible
V	volts
v/v	volume by volume
Val	valine
<i>V. cholerae</i>	<i>Vibrio cholerae</i>
Vib	vibriobactin
<i>X. cellulolytica</i>	<i>Xylanimonas cellulolytica</i>

References

1. L. D. Palmer and E. P. Skaar, in *Annual Review of Genetics, Vol 50*, ed. N. M. Bonini, Annual Reviews, Palo Alto, 2016, vol. 50, pp. 67-91.
2. D. J. Raines, T. J. Sanderson, E. J. Wilde and A. K. Duhme-Klair, in *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier, 2015, DOI: <http://dx.doi.org/10.1016/B978-0-12-409547-2.11040-6>.
3. M. L. Guerinot, *Annu. Rev. Microbiol.*, 1994, **48**, 743-772.
4. P. M. Harrison and P. Arosio, *BBA- Bioenergetics*, 1996, **1275**, 161-203.
5. J. Burgess and M. V. Twigg, in *Encyclopedia of Inorganic Chemistry*, John Wiley & Sons, Ltd, 2006.
6. J. H. Crosa, A. R. Mey and S. M. Payne, *Iron Transport in Bacteria*, ASM press Washington DC:, 2004.
7. P. C. Junk, B. J. McCool, B. Moubaraki, K. S. Murray, L. Spiccia, J. D. Cashion and J. W. Steed, *J. Chem. Soc., Dalton Trans.*, 2002, 1024-1029.
8. T. Morrison, A. Reis Jr, G. Knapp, F. Fradin, H. Chen and T. Klippert, *J. Am. Chem. Soc.*, 1978, **100**, 3262-3264.
9. A. Shanzer, J. Libman, P. Yakirevitch, Y. Hadar, Y. Chen and E. Jurkevitch, *Chirality*, 1993, **5**, 359-365.
10. R. J. Abergel, A. M. Zawadzka, T. M. Hoette and K. N. Raymond, *J. Am. Chem. Soc.*, 2009, **131**, 12682-12692.
11. M. Brorson, T. Damhus and C. E. Schaffer, *Inorg. Chem.*, 1983, **22**, 1569-1573.
12. H. Brunner, *Adv. Organomet. Chem.*, 1980, **18**, 151-206.
13. S. S. Isied, G. Kuo and K. N. Raymond, *J. Am. Chem. Soc.*, 1976, **98**, 1763-1767.
14. N. A. P. Kane-Maguire, T. W. Hanks, D. G. Jurs, R. M. Tollison, A. L. Heatherington, L. M. Ritzenthaler, L. M. McNulty and H. M. Wilson, *Inorg. Chem.*, 1995, **34**, 1121-1124.
15. W. H. Rastetter, T. J. Erickson and M. C. Venuti, *J. Org. Chem.*, 1981, **46**, 3579-3590.

16. T. D. P. Stack, T. B. Karpishin and K. N. Raymond, *J. Am. Chem. Soc.*, 1992, **114**, 1512-1514.
17. M. E. Bluhm, S. S. Kim, E. A. Dertz and K. N. Raymond, *J. Am. Chem. Soc.*, 2002, **124**, 2436-2437.
18. J. L. Pierre and M. Fontecave, *Biometals*, 1999, **12**, 195-199.
19. C. Wandersman and P. Delepelaire, *Annu. Rev. Microbiol.*, 2004, **58**, 611-647.
20. K. D. Krewulak and H. J. Vogel, *BBA - Biomembranes*, 2008, **1778**, 1781-1804.
21. E. R. Frawley and F. C. Fang, *Molec. Microbiol.*, 2014, **93**, 609-616.
22. J. R. Sheldon and D. E. Heinrichs, *FEMS Microbiol. Rev.*, 2015, **39**, 592-630.
23. J. B. Neilands, *Biol. Met.*, 1991, **4**, 1-6.
24. J. R. Chipperfield and C. Ratledge, *Biometals*, 2000, **13**, 165-168.
25. C. Ratledge and L. G. Dover, *Annu. Rev. Microbiol.*, 2000, **54**, 881-941.
26. H. Boukhalfa and A. L. Crumbliss, *Biometals*, 2002, **15**, 325-339.
27. J. E. Cassat and E. P. Skaar, *Cell Host Microbe*, 2013, **13**, 509-519.
28. R. C. Hider and X. Kong, *Nat. Prod. Rep.*, 2010, **27**, 637-657.
29. L. D. Loomis and K. N. Raymond, *Inorg. Chem.*, 1991, **30**, 906-911.
30. M. Miethke and M. A. Marahiel, *Microbiol. Mol. Biol. Rev.*, 2007, **71**, 413-451.
31. K. N. Raymond and C. J. Carrano, *Acc. Chem. Res.*, 1979, **12**, 183-190.
32. M. Ellermann and J. C. Arthur, *Free Radic. Biol. Med.*, 2017, **105**, 68-78.
33. G. A. Snow, *J. Chem. Soc.*, 1954, **1**, 2588-2596.
34. G. A. Snow, *J. Chem. Soc.*, 1954, **1**, 4080-4093.
35. J. B. Neilands, *J. Am. Chem. Soc.*, 1952, **74**, 4846-4847.
36. J. Garibaldi and J. Neilands, *Nature*, 1956, **177**, 526-527.
37. K. N. Raymond, B. E. Allred and A. K. Sia, *Acc. Chem. Res.*, 2015, **48**, 2496-2505.
38. C. Hesselstine, C. Pidacks, A. Whitehill, N. Bohonos, B. Hutchings and J. Williams, *J. Am. Chem. Soc.*, 1952, **74**, 1362-1362.
39. K. Hantke, *Mol. Gen. Genet.*, 1983, **191**, 301-306.

40. B. F. Burnham and J. Neilands, *J. Biol. Chem.*, 1961, **236**, 554-559.
41. M. Hannauer, Y. Barda, G. L. Mislin, A. Shanzer and I. J. Schalk, *J. Bacteriol.*, 2010, **192**, 1212-1220.
42. T. Emery and J. Neilands, *Nature*, 1959, **184**, 1632-1633.
43. B. Schwyn and J. B. Neilands, *Anal. Biochem.*, 1987, **160**, 47-56.
44. M. Sandy and A. Butler, *Chem. Rev.*, 2009, **109**, 4580-4595.
45. V. V. Homann, K. J. Edwards, E. A. Webb and A. Butler, *Biometals*, 2009, **22**, 565-571.
46. A. Butler, H. K. Zane, M. P. Kem, H. Naka and M. G. Haygood, *J. Biol. Inorg. Chem.*, 2014, **19**, S707-S707.
47. M. P. Kem, H. K. Zane, S. D. Springer, J. M. Gauglitz and A. Butler, *Metallomics*, 2014, **6**, 1150-1155.
48. J.-M. Meyer, V. A. Geoffroy, C. Baysse, P. Cornelis, I. Barelmann, K. Taraz and H. Budzikiewicz, *Arch. Biochem. Biophys.*, 2002, **397**, 179-183.
49. Y. L. Qi, H. Hayen and D. A. Volmer, *Biometals*, 2016, **29**, 53-60.
50. M. E. Peek, A. Bhatnagar, N. A. McCarty and S. M. Zughaier, *Interdiscip. Perspect. Infect. Dis.*, 2012, **2012**, 1-10.
51. S. Pahlow, S. Stockel, S. Pollok, D. Cialla-May, P. Rosch, K. Weber and J. Popp, *Anal. Chem.*, 2016, **88**, 1570-1577.
52. O. Cunrath, V. A. Geoffroy and I. J. Schalk, *Environ. Microbiol.*, 2016, **18**, 3258-3267.
53. I. Schalk and L. Guillon, *Amino Acids*, 2013, **44**, 1267-1277.
54. I. J. Schalk and O. Cunrath, *Environ. Microbiol.*, 2016, **18**, 3227-3246.
55. K. N. Raymond and E. A. Dertz, in *Iron Transport in Bacteria*, American Society of Microbiology, 2004, pp. 3-17.
56. J. Harrington and A. Crumbliss, *Biometals*, 2009, **22**, 679-689.
57. I. O'brien and F. Gibson, *BBA- Gen. Subjects*, 1970, **215**, 393-402.
58. J. R. Pollack and J. B. Neilands, *Biochem. Biophys. Res. Commun.*, 1970, **38**, 989-992.

59. J. R. Pollack, B. N. Ames and J. Neilands, *J. Bacteriol.*, 1970, **104**, 635-639.
60. G. Cox, F. Gibson, R. Luke, N. Newton, I. O'brien and H. Rosenberg, *J. Bacteriol.*, 1970, **104**, 219-226.
61. E. R. Marinez, E. K. Salmassian, T. T. Lau and C. G. Gutierrez, *J. Org. Chem.*, 1996, **61**, 3548-3550.
62. K. N. Raymond, E. A. Dertz and S. S. Kim, *Proc. Natl. Acad. Sci.*, 2003, **100**, 3584-3588.
63. B. P. Hay, D. A. Dixon, R. Vargas, J. Garza and K. N. Raymond, *Inorg. Chem.*, 2001, **40**, 3922-3935.
64. T. B. Karpishin, T. M. Dewey and K. N. Raymond, *J. Am. Chem. Soc.*, 1993, **115**, 1842-1851.
65. K. N. Raymond, *Coord. Chem. Rev.*, 1990, **105**, 135-153.
66. T. B. Karpishin, M. S. Gebhard, E. I. Solomon and K. N. Raymond, *J. Am. Chem. Soc.*, 1991, **113**, 2977-2984.
67. R. K. Hocking, S. DeBeer George, K. N. Raymond, K. O. Hodgson, B. Hedman and E. I. Solomon, *J. Am. Chem. Soc.*, 2010, **132**, 4006-4015.
68. M. E. Cass, T. M. Garrett and K. N. Raymond, *J. Am. Chem. Soc.*, 1989, **111**, 1677-1682.
69. R. J. Abergel, J. A. Warner, D. K. Shuh and K. N. Raymond, *J. Am. Chem. Soc.*, 2006, **128**, 8920-8931.
70. S. M. Cohen, M. Meyer and K. N. Raymond, *J. Am. Chem. Soc.*, 1998, **120**, 6277-6286.
71. S. K. Porwal, E. Furia, M. E. Harris, R. Viswanathan and L. Devireddy, *J. Inorg. Biochem.*, 2014, **145**, 1-10.
72. C. W. Lee, D. J. Ecker and K. N. Raymond, *J. Am. Chem. Soc.*, 1985, **107**, 6920-6923.
73. Z. L. Reitz, M. Sandy and A. Butler, *Metallomics*, 2017, **9**, 824-839.
74. G. L. Challis and D. A. Hopwood, *Proc. Natl. Acad. Sci.*, 2003, **100**, 14555-14561.
75. E. A. Dertz, J. Xu, A. Stintzi and K. N. Raymond, *J. Am. Chem. Soc.*, 2005, **128**, 22-23.

76. M. Miethke, O. Klotz, U. Linne, J. J. May, C. L. Beckering and M. A. Marahiel, *Molec. Microbiol.*, 2006, **61**, 1413-1427.
77. R. Annamalai, B. Jin, Z. Cao, S. M. Newton and P. E. Klebba, *J Bacteriol*, 2004, **186**, 3578-3589.
78. Y. Wen, X. Wu, Y. Teng, C. Qian, Z. Zhan, Y. Zhao and O. Li, *Environ. Microbiol.*, 2011, **13**, 2726-2737.
79. S. I. Patzer and V. Braun, *J. Bacteriol.*, 2010, **192**, 426-435.
80. H. K. Zane, H. Naka, F. Rosconi, M. Sandy, M. G. Haygood and A. Butler, *J. Am. Chem. Soc.*, 2014, **136**, 5615-5618.
81. K. H. Li, W. H. Chen and S. D. Bruner, *Biometals*, 2016, **29**, 377-388.
82. S. A. Ong, T. Peterson and J. Neilands, *J. Biol. Chem.*, 1979, **254**, 1860-1865.
83. S. Yamamoto, N. Okujo, Y. Fujita, M. Saito, T. Yoshida and S. Shinoda, *J. Biochem.*, 1993, **113**, 538-544.
84. B. E. Allred, T. Hoette and K. N. Raymond, *Abstr. Pap. Am. Chem. Soc.*, 2010, **239**, 1.
85. A.-K. Duhme, R. C. Hider and H. H. Khodr, *Chem. Ber.*, 1997, **130**, 969-973.
86. I. Berner, M. Greiner, J. Metzger, G. Jung and G. Winkelmann, *Biol. Met.*, 1991, **4**, 113-118.
87. G. Winkelmann, A. Cansier, W. Beck and G. Jung, *Biometals*, 1994, **7**, 149-154.
88. K. H. Li and S. D. Bruner, *Proteins*, 2016, **84**, 118-128.
89. R. Bergeron, J. B. Dionis, G. T. Elliott and S. J. Kline, *J. Biol. Chem.*, 1985, **260**, 7936-7944.
90. D. J. Raines, O. V. Moroz, E. V. Blagova, J. P. Turkenburg, K. S. Wilson and A.-K. Duhme-Klair, *Proc. Natl. Acad. Sci.*, 2016, **113**, 5850-5855.
91. S. Barghouthi, R. Young, M. Olson, J. Arceneaux, L. Clem and B. Byers, *J. Bacteriol.*, 1989, **171**, 1811-1816.
92. J. R. Telford, J. A. Leary, L. M. Tunstad, B. R. Byers and K. N. Raymond, *J. Am. Chem. Soc.*, 1994, **116**, 4499-4500.
93. G. Ehlert, K. Taraz and H. Budzikiewicz, *Z. Naturforsch. C.*, 1994, **49**, 11-17.

94. A.-K. Duhme, *J. Chem. Soc., Dalton Trans.*, 1997, **1**, 773-778.
95. R. C. Scarrow, D. J. Ecker, C. Ng, S. Liu and K. N. Raymond, *Inorg. Chem.*, 1991, **30**, 900-906.
96. B. Kunze, N. Bedorf, W. Kohl, G. Hölfe and H. Reichenbach, *J. Antibiot.*, 1989, **42**, 14-17.
97. A. Proschak, P. Lubuta, P. Grün, F. Löhr, G. Wilharm, V. De Berardinis and H. B. Bode, *ChemBioChem*, 2013, **14**, 633-638.
98. I.-K. Lee, D.-W. Ki, S.-E. Kim, J.-H. Yeom, Y.-S. Kim and B.-S. Yun, *J. Korean Soc. Appl. Biol. Chem.*, 2011, **54**, 312-315.
99. M. R. Seyedsayamdost, S. Cleto, G. Carr, H. Vlamakis, M. João Vieira, R. Kolter and J. Clardy, *J. Am. Chem. Soc.*, 2012, **134**, 13550-13553.
100. E. J. Enemark and T. D. P. Stack, *Inorg. Chem.*, 1996, **35**, 2719-2720.
101. R. C. Scarrow, D. L. White and K. N. Raymond, *J. Am. Chem. Soc.*, 1985, **107**, 6540-6546.
102. J. A. Shapiro and T. A. Wencewicz, *ACS Infect. Dis.*, 2016, **2**, 157-168.
103. W. Y. Song, D. Jeong, J. Kim, M. W. Lee, M. H. Oh and H. J. Kim, *Org. Lett.*, 2017, **19**, 500-503.
104. J. A. Shapiro and T. A. Wencewicz, *Metallomics*, 2017, **9**, 463-470.
105. T. Ito, *Appl. Environ. Microbiol.*, 1993, **59**, 2343-2345.
106. W. J. Page and M. Von Tigerstrom, *J. Gen. Microbiol.*, 1988, **134**, 453-460.
107. A. K. Duhme, R. C. Hider, M. J. Naldrett and R. N. Pau, *J. Biol. Inorg. Chem.*, 1998, **3**, 520-526.
108. A. L. Crumbliss and J. M. Harrington, in *Advances in Inorganic Chemistry, Vol 61: Metal Ion Controlled Reactivity*, eds. R. VanEldik and C. D. Hubbard, Elsevier Academic Press Inc, San Diego, 2009, vol. 61, pp. 179-250.
109. J. R. Telford and K. N. Raymond, *Inorg. Chem.*, 1998, **37**, 4578-4583.
110. I. Spasojević, S. K. Armstrong, T. J. Brickman and A. L. Crumbliss, *Inorg. Chem.*, 1999, **38**, 449-454.

111. M. M. Andrade and A. P. Rauter, in *Carbohydrate Chemistry: Volume 38*, The Royal Society of Chemistry, 2012, vol. 38, pp. 398-415.
112. R. Niehus, A. Picot, N. M. Oliveira, S. Mitri and K. R. Foster, *Evolution*, 2017, **71**, 1443-1455.
113. S. A. West and A. Buckling, *P. Roy. Soc. Lond. B. Bio.*, 2003, **270**, 37-44.
114. A. Leinweber, R. F. Inglis and R. Kummerli, *Isme J.*, 2017, **11**, 1179-1188.
115. K. Palyada, D. Threadgill and A. Stintzi, *J. Bacteriol.*, 2004, **186**, 4714-4729.
116. B. D. Connelly and K. J. Dickinson, *Evolution*, 2017, **71**, 1735-1736.
117. W. Köster, *Res. Microbiol.*, 2001, **152**, 291-301.
118. C. K. Y. Lau, K. D. Krewulak and H. J. Vogel, *FEMS Microbiol. Rev.*, 2016, **40**, 273-298.
119. S. C. Andrews, A. K. Robinson and F. Rodríguez-Quñones, *FEMS Microbiol. Rev.*, 2003, **27**, 215-237.
120. P. I. Higgs, P. S. Myers and K. Postle, *J. Bacteriol.*, 1998, **180**, 6031-6038.
121. L. Moynié, A. Luscher, D. Rolo, D. Pletzer, A. Tortajada, H. Weingart, Y. Braun, M. G. Page, J. H. Naismith and T. Köhler, *Antimicrob. Agents Chemother.*, 2017, **61**, e02531-02516.
122. T. E. Clarke, V. Braun, G. Winkelmann, L. W. Tari and H. J. Vogel, *J. Biol. Chem.*, 2002, **277**, 13966-13972.
123. K. K. Kumar, D. A. Jacques, G. Pishchany, T. Caradoc-Davies, T. Spirig, G. R. Malmirchegini, D. B. Langley, C. F. Dickson, J. P. Mackay and R. T. Clubb, *J. Biol. Chem.*, 2011, **286**, 38439-38447.
124. C.-J. Chen, P. F. Sparling, L. A. Lewis, D. W. Dyer and C. Elkins, *Infect. Immun.*, 1996, **64**, 5008-5014.
125. C. Wandersman and I. Stojiljkovic, *Curr. Opin. Microbiol.*, 2000, **3**, 215-220.
126. V. Braun, *Front. Biosci.*, 2003, **8**, 1409-1421.
127. L. Escolar, J. Pérez-Martín and V. De Lorenzo, *J. Bacteriol.*, 1999, **181**, 6223-6229.
128. M. F. Fillat, *Arch. Biochem. Biophys.*, 2014, **546**, 41-52.

129. J. Butcher, S. Sarvan, J. S. Brunzelle, J.-F. Couture and A. Stintzi, *Proc. Natl. Acad. Sci.*, 2012, **109**, 10047-10052.
130. M. Askoura, S. Sarvan, J. F. Couture and A. Stintzi, *Infect. Immun.*, 2016, **84**, 1287-1300.
131. K. Holmes, F. Mulholland, B. M. Pearson, C. Pin, J. McNicholl-Kennedy, J. M. Ketley and J. M. Wells, *Microbiology*, 2005, **151**, 243-257.
132. A. R. Mey, E. E. Wyckoff, V. Kanukurthy, C. R. Fisher and S. M. Payne, *Infect. Immun.*, 2005, **73**, 8167-8178.
133. J.-W. Lee and J. D. Helmann, *Biometals*, 2007, **20**, 485-499.
134. O. Tenaillon, D. Skurnik, B. Picard and E. Denamur, *Nat. Rev. Micro.*, 2010, **8**, 207-217.
135. T. A. Russo and J. R. Johnson, *Microb. Infect.*, 2003, **5**, 449-456.
136. S. E. Majowicz, E. Scallan, A. Jones-Bitton, J. M. Sargeant, J. Stapleton, F. J. Angulo, D. H. Yeung and M. D. Kirk, *Foodborne Pathog. Dis.*, 2014, **11**, 447-455.
137. R. Mann, D. G. Mediati, I. G. Duggin, E. J. Harry and A. L. Bottomley, *Front. Cell. Infect. Microbiol.*, 2017, **7**, 1-15.
138. M. Valdebenito, A. L. Crumbliss, G. Winkelmann and K. Hantke, *Int. J. Med. Microbiol.*, 2006, **296**, 513-520.
139. J. H. Crosa and C. T. Walsh, *Microbiol. Mol. Biol. Rev.*, 2002, **66**, 223-249.
140. M. G. Gresock, K. A. Kastead and K. Postle, *J. Bacteriol.*, 2015, **197**, 3433-3445.
141. M. G. Gresock and K. Postle, *J. Bacteriol.*, 2017, **199**, e00649-00616.
142. L. Ma, W. Kaserer, R. Annamalai, D. C. Scott, B. Jin, X. Jiang, Q. Xiao, H. Maymani, L. M. Massis, L. C. Ferreira, S. M. Newton and P. E. Klebba, *J. Biol. Chem.*, 2007, **282**, 397-406.
143. D. L. Stephens, M. D. Choe and C. F. Earhart, *Microbiology*, 1995, **141**, 1647-1654.
144. C. Sprencel, Z. Cao, Z. Qi, D. C. Scott, M. A. Montague, N. Ivanoff, J. Xu, K. M. Raymond, S. M. Newton and P. E. Klebba, *J. Bacteriol.*, 2000, **182**, 5359-5364.
145. B. C. H. Chu, R. Otten, K. D. Krewulak, F. A. A. Mulder and H. J. Vogel, *J. Biol. Chem.*, 2014, **289**, 29219-29234.

146. B. Li, N. Li, Y. Yue, X. Liu, Y. Huang, L. Gu and S. Xu, *Biochem. Biophys. Res. Commun.*, 2016, **478**, 1049-1053.
147. B. A. Ozenberger, M. S. Nahlik and M. McIntosh, *J. Bacteriol.*, 1987, **169**, 3638-3646.
148. K. J. Linton and C. F. Higgins, *Molec. Microbiol.*, 1998, **28**, 5-13.
149. S. S. Chenault and C. F. Earhart, *Microbiology*, 1992, **138**, 2167-2171.
150. S. Kalynych, M. A. Valvano and M. Cygler, *Protein Eng. Des. Sel.*, 2012, **25**, 797-802.
151. A. Tocilj, C. Munger, A. Proteau, R. Morona, L. Purins, E. Ajamian, J. Wagner, M. Papadopoulos, L. Van Den Bosch, J. L. Rubinstein, J. Fethiere, A. Matte and M. Cygler, *Nat. Struct. Mol. Biol.*, 2008, **15**, 130-138.
152. J. L. Furrer, D. N. Sanders, I. G. Hook-Barnard and M. A. McIntosh, *Mol. Microbiol.*, 2002, **44**, 1225-1234.
153. G. Grass, *Biometals*, **19**, 159-172.
154. B. Bister, D. Bischoff, G. Nicholson, M. Valdebenito, K. Schneider, G. Winkelmann, K. Hantke and R. Süssmuth, *Biometals*, 2004, **17**, 471-481.
155. M. Caza, F. Lépine, S. Milot and C. M. Dozois, *Infect. Immun.*, 2008, **76**, 3539-3549.
156. M. Caza, F. Lépine and C. M. Dozois, *Molec. Microbiol.*, 2011, **80**, 266-282.
157. M. G. Zhu, M. Valdebenito, G. Winkelmann and K. Hantke, *Microbiology*, 2005, **151**, 2363-2372.
158. M. L. V. Crouch, M. Castor, J. E. Karlinsey, T. Kalthorn and F. C. Fang, *Molec. Microbiol.*, 2008, **67**, 971-983.
159. X. Zeng, Y. Mo, F. Xu and J. Lin, *Molec. Microbiol.*, 2013, **87**, 594-608.
160. K. Hantke, *FEMS Microbiol. Lett.*, 1990, **67**, 5-8.
161. E. E. Wyckoff, B. E. Allred, K. N. Raymond and S. M. Payne, *J. Bacteriol.*, 2015, **197**, 2840-2849.
162. V. A. Bamford, M. Armour, S. A. Mitchell, M. Cartron, S. C. Andrews and K. A. Watson, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, 2008, **64**, 792-796.

163. M. Miethke, J. Hou and M. A. Marahiel, *Biochemistry*, 2011, **50**, 10951-10964.
164. S. Wang, Y. Wu and F. W. Outten, *J. Bacteriol.*, 2011, **193**, 563-574.
165. S. Heidinger, V. Braun, V. L. Pecoraro and K. N. Raymond, *J. Bacteriol.*, 1983, **153**, 109-115.
166. D. Acheson and B. M. Allos, *Clin. Infect. Dis.*, 2001, **32**, 1201-1206.
167. M. Skirrow, *Int. J. Food Microbiol.*, 1991, **12**, 9-16.
168. M. Stahl, J. Butcher and A. Stintzi, *Front. Cell. Infect. Microbiol.*, 2012, **2**, 1-10.
169. H. Trigui, A. Thibodeau, P. Fravallo, A. Letellier and S. P. Faucher, *SpringerPlus*, 2015, **4**, 799.
170. I. Indikova, T. J. Humphrey and F. Hilbert, *Front. Microbiol.*, 2015, **6**, 1-6.
171. S. F. Park and P. T. Richardson, *J. Bacteriol.*, 1995, **177**, 2259-2264.
172. C. E. Miller, P. H. Williams and J. M. Ketley, *Microbiology*, 2009, **155**, 3157-3165.
173. C. Pickett, T. Auffenberg, E. Pesci, V. Sheen and S. Jusuf, *Infect. Immun.*, 1992, **60**, 3872-3877.
174. D. Hofreuter, J. Tsai, R. O. Watson, V. Novik, B. Altman, M. Benitez, C. Clark, C. Perbost, T. Jarvie and L. Du, *Infect. Immun.*, 2006, **74**, 4694-4707.
175. L. Field, V. Headley, S. Payne and L. Berry, *Infect. Immun.*, 1986, **54**, 126-132.
176. L. H. Field, V. L. Headley, J. L. Underwood, S. M. Payne and L. Berry, *Infect. Immun.*, 1986, **54**, 118-125.
177. J. Parkhill, B. W. Wren, K. Mungall, J. M. Ketley, C. Churcher, D. Basham, T. Chillingworth, R. M. Davies, T. Feltwell, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Moule, M. J. Pallen, C. W. Penn, M. A. Quail, M. A. Rajandream, K. M. Rutherford, A. H. M. van Vliet, S. Whitehead and B. G. Barrell, *Nature*, 2000, **403**, 665-668.
178. S. A. L. Tom-Yew, D. T. Cui, E. G. Bekker and M. E. P. Murphy, *J. Biol. Chem.*, 2005, **280**, 9283-9290.
179. X. Zeng, F. Xu and J. Lin, *Infect. Immun.*, 2009, **77**, 5437-5448.
180. X. Zeng and J. Lin, *Campylobacter jejuni: Methods and Protocols*, 2017, pp 65-78.

181. A. H. van Vliet, J. M. Ketley, S. F. Park and C. W. Penn, *FEMS Microbiol. Rev.*, 2002, **26**, 173-186.
182. H. Naikare, J. Butcher, A. Flint, J. Xu, K. N. Raymond and A. Stintzi, *Metallomics*, 2013, **5**, 988-996.
183. F. Xu, X. Zeng, R. D. Haigh, J. M. Ketley and J. Lin, *J. Bacteriol.*, 2010, **192**, 4425-4435.
184. R. P.-A. Berntsson, S. H. Smits, L. Schmitt, D.-J. Slotboom and B. Poolman, *FEBS Lett.*, 2010, **584**, 2606-2617.
185. B. C. Chu and H. J. Vogel, *J. Biol. Chem.*, 2011, **392**, 39-52.
186. D. J. Raines, O. V. Moroz, K. S. Wilson and A. K. Duhme-Klair, *Angew. Chem., Int. Ed. Engl.*, 2013, **52**, 4595-4598.
187. A. Müller, A. J. Wilkinson, K. S. Wilson and A.-K. Duhme-Klair, *Angew. Chem., Int. Ed. Engl.*, 2006, **45**, 5132-5136.
188. R. A. Crossley, D. J. Gaskin, K. Holmes, F. Mulholland, J. M. Wells, D. J. Kelly, A. H. van Vliet and N. J. Walton, *Appl. Environ. Microbiol.*, 2007, **73**, 7819-7825.
189. D. A. Sack and S. R. Bradley Sack MD, *N. Engl. J. Med.*, 2006, **355**, 649-651.
190. S. M. Faruque, M. J. Albert and J. J. Mekalanos, *Microbiol. Mol. Biol. Rev.*, 1998, **62**, 1301-1314.
191. J. Reidl and K. E. Klose, *FEMS Microbiol. Rev.*, 2002, **26**, 125-139.
192. A. R. Mey, E. E. Wyckoff, A. G. Oglesby, E. Rab, R. K. Taylor and S. M. Payne, *Infect. Immun.*, 2002, **70**, 3419-3426.
193. C. M. Litwin and S. Calderwood, *Clin. Microbiol. Rev.*, 1993, **6**, 137-149.
194. E. A. Weaver, E. E. Wyckoff, A. R. Mey, R. Morrison and S. M. Payne, *J. Bacteriol.*, 2013, **195**, 4826-4835.
195. E. E. Wyckoff, M. Schmitt, A. Wilks and S. M. Payne, *J. Bacteriol.*, 2004, **186**, 4142-4151.
196. T. Uchida, Y. Sekine, T. Matsui, M. Ikeda-Saito and K. Ishimori, *Chem. Commun.*, 2012, **48**, 6741-6743.
197. X. Liu, J. Gong, T. Wei, Z. Wang, Q. Du, D. Zhu, Y. Huang, S. Xu and L. Gu, *BMC Struct. Biol.*, 2012, **12**, 1-10.

198. D. E. Kunkle, X. R. Bina and J. E. Bina, *mBio*, 2017, **8**, e00126-00117.
199. E. E. Wyckoff, A. R. Mey, A. Leimbach, C. F. Fisher and S. M. Payne, *J. Bacteriol.*, 2006, **188**, 6515-6523.
200. E. E. Wyckoff, A. R. Mey and S. M. Payne, *Biometals*, 2007, **20**, 405-416.
201. E. E. Wyckoff and S. M. Payne, *Molec. Microbiol.*, 2011, **81**, 1446-1458.
202. G. Griffiths, S. Sigel, S. Payne and J. Neilands, *J. Biol. Chem.*, 1984, **259**, 383-385.
203. S. S. Seliger, A. R. Mey, A. M. Valle and S. M. Payne, *Molec. Microbiol.*, 2001, **39**, 801-812.
204. S. M. Payne, A. R. Mey and E. E. Wyckoff, *Microbiol. Mol. Biol. Rev.*, 2016, **80**, 69-90.
205. J. R. Butterton, J. A. Stoebner, S. M. Payne and S. B. Calderwood, *J. Bacteriol.*, 1992, **174**, 3729-3738.
206. J. Rutz, T. Abdullah, S. Singh, V. Kalve and P. Klebba, *J. Bacteriol.*, 1991, **173**, 5964-5974.
207. E. E. Wyckoff, A.-M. Valle, S. L. Smith and S. M. Payne, *J. Bacteriol.*, 1999, **181**, 7588-7596.
208. N. Li, C. Zhang, B. Li, X. Liu, Y. Huang, S. Xu and L. Gu, *J. Biol. Chem.*, 2012, **287**, 8912-8919.
209. X. Liu, Q. Du, Z. Wang, S. Liu, N. Li, Y. Chen, C. Zhu, D. Zhu, T. Wei and Y. Huang, *FEBS Lett.*, 2012, **586**, 1240-1244.
210. M. Nairz, I. Theurl, F. K. Swirski and G. Weiss, *Pflügers Arch.*, 2017, **469**, 397-418.
211. M. Nairz, D. Ferring-Appel, D. Casarrubea, T. Sonnweber, L. Viatte, A. Schroll, D. Haschka, F. C. Fang, M. W. Hentze and G. Weiss, *Cell Host Microbe*, 2015, **18**, 254-261.
212. H. Pi, S. A. Jones, L. E. Mercer, J. P. Meador, J. E. Caughron, L. Jordan, S. M. Newton, T. Conway and P. E. Klebba, *PLOS ONE*, 2012, **7**, e50020.
213. K. Konopka and J. B. Neilands, *Biochemistry*, 1984, **23**, 2122-2127.
214. E. P. Skaar, *PLOS Pathogens*, 2010, **6**, e1000949.
215. E. P. Skaar, *Cell Host Microbe*, 2009, **5**, 422-424.

216. A. L. Nelson, J. M. Barasch, R. M. Bunte and J. N. Weiser, *Cell. Microbiol.*, 2005, **7**, 1404-1417.
217. D. H. Goetz, M. A. Holmes, N. Borregaard, M. E. Bluhm, K. N. Raymond and R. K. Strong, *Mol. Cell*, 2002, **10**, 1033-1043.
218. A. K. Sia, B. E. Allred and K. N. Raymond, *Curr. Opin. Chem. Biol.*, 2013, **17**, 150-157.
219. T. M. Hoette, R. J. Abergel, J. D. Xu, R. K. Strong and K. N. Raymond, *J. Am. Chem. Soc.*, 2008, **130**, 17584-17592.
220. R. J. Abergel, M. C. Clifton, J. C. Pizarro, J. A. Warner, D. K. Shuh, R. K. Strong and K. N. Raymond, *J. Am. Chem. Soc.*, 2008, **130**, 11524-11534.
221. B. E. Allred, P. B. Rupert, S. S. Gauny, D. D. An, C. Y. Ralston, M. Sturzbecher-Hoehne, R. K. Strong and R. J. Abergel, *Proc. Natl. Acad. Sci.*, 2015, **112**, 10342-10347.
222. M. A. Holmes, W. Paulsene, X. Jide, C. Ratledge and R. K. Strong, *Structure*, 2005, **13**, 29-41.
223. B. E. Allred, C. Correnti, M. C. Clifton, R. K. Strong and K. N. Raymond, *ACS Chem. Biol.*, 2013, **8**, 1882-1887.
224. G. Bao, M. Clifton, T. M. Hoette, K. Mori, S.-X. Deng, A. Qiu, M. Viltard, D. Williams, N. Paragas, T. Leete, R. Kulkarni, X. Li, B. Lee, A. Kalandadze, A. J. Ratner, J. C. Pizarro, K. M. Schmidt-Ott, D. W. Landry, K. N. Raymond, R. K. Strong and J. Barasch, *Nat. Chem. Biol.*, 2010, **6**, 602-609.
225. M. Caza, A. Garenaux, F. Lepine and C. M. Dozois, *Molec. Microbiol.*, 2015, **97**, 717-732.
226. M. Valdebenito, S. I. Muller and K. Hantke, *FEMS Microbiol. Lett.*, 2007, **277**, 182-187.
227. M. A. Bachman, J. E. Oyler, S. H. Burns, M. Caza, F. Lépine, C. M. Dozois and J. N. Weiser, *Infect. Immun.*, 2011, **79**, 3309-3316.
228. T. A. Russo, R. Olson, U. MacDonald, J. Beanan and B. A. Davidson, *Infect. Immun.*, 2015, **83**, 3325-3333.
229. K. G. Wooldridge and P. H. Williams, *FEMS Microbiol. Rev.*, 1993, **12**, 325-348.

230. R. J. Abergel, M. K. Wilson, J. E. L. Arceneaux, T. M. Hoette, R. K. Strong, B. R. Byers and K. N. Raymond, *Proc. Natl. Acad. Sci.*, 2006, **103**, 18499-18503.
231. A. M. Zawadzka, Y. Kim, N. Maltseva, R. Nichiporuk, Y. Fan, A. Joachimiak and K. N. Raymond, *Proc. Natl. Acad. Sci.*, 2009, **106**, 21854-21859.
232. M. Valdebenito, B. Bister, R. Reissbrodt, K. Hantke and G. Winkelmann, *Int. J. Med. Microbiol.*, 2005, **295**, 99-107.
233. S. I. Elshahawi, K. A. Shaaban, M. K. Kharel and J. S. Thorson, *Chem. Soc. Rev.*, 2015, **44**, 7591-7697.
234. M. Luo, H. Lin, M. A. Fischbach, D. R. Liu, C. T. Walsh and J. T. Groves, *ACS Chem. Biol.*, 2006, **1**, 29-32.
235. K. Hantke, G. Nicholson, W. Rabsch and G. Winkelmann, *Proc. Natl. Acad. Sci.*, 2003, **100**, 3677-3682.
236. S. Muller, M. Valdebenito and K. Hantke, *Biometals*, 2009, **22**, 691-695.
237. G. Vassiliadis, D. Destoumieux-Garzón, C. Lombard, S. Rebuffat and J. Peduzzi, *Antimicrob. Agents Chemother.*, 2010, **54**, 288-297.
238. M. Sassone-Corsi, S.-P. Nuccio, H. Liu, D. Hernandez, C. T. Vu, A. A. Takahashi, R. A. Edwards and M. Raffatellu, *Nature*, 2016, **540**, 280-283.
239. A. A. Lee, Y. C. S. Chen, E. Ekalestari, S. Y. Ho, N. S. Hsu, T. F. Kuo and T. S. A. Wang, *Angew. Chem. Int. Ed.*, 2016, **55**, 12338-12342.
240. R. Lagos, M. Wilkens, C. Vergara, X. Cecchi and O. Monasterio, *FEBS Lett.*, 1993, **321**, 145-148.
241. G. Vassiliadis, J. Peduzzi, S. Zirah, X. Thomas, S. Rebuffat and D. Destoumieux-Garzón, *Antimicrob. Agents Chemother.*, 2007, **51**, 3546-3553.
242. G. Mercado, M. Tello, M. Marín, O. Monasterio and R. Lagos, *J. Bacteriol.*, 2008, **190**, 5464-5471.
243. M. A. Fischbach, H. Lin, D. R. Liu and C. T. Walsh, *Nat. Chem. Biol.*, 2006, **2**, 132-138.
244. E. Strahsburger, M. Baeza, O. Monasterio and R. Lagos, *Antimicrob. Agents Chemother.*, 2005, **49**, 3083-3086.
245. S. Rebuffat, *Biochem. Soc. Trans.*, 2012, **40**, 1456-1462.

246. T. A. Wencewicz, *Bioorg. Med. Chem.*, 2016, **24**, 6227-6252.
247. I. O'Brien, G. Cox and F. Gibson, *BBA- Gen. Subjects*, 1971, **237**, 537-549.
248. F. L. Weitzl, W. R. Harris and K. N. Raymond, *J. Med. Chem.*, 1979, **22**, 1281-1283.
249. W. H. Harris and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 6534-6541.
250. R. C. Hider, D. Bickar, I. E. G. Morrison and J. Silver, *J. Am. Chem. Soc.*, 1984, **106**, 6983-6987.
251. Z. Hou, T. Stack, C. J. Sunderland and K. N. Raymond, *Inorg. Chim. Acta.*, 1997, **263**, 341-355.
252. F. L. Weitzl and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 2728-2731.
253. T. M. Garrett, T. J. McMurry, M. W. Hosseini, Z. E. Reyes, F. E. Hahn and K. N. Raymond, *J. Am. Chem. Soc.*, 1991, **113**, 2965-2977.
254. T. M. Garrett, P. W. Miller and K. N. Raymond, *Inorg. Chem.*, 1989, **28**, 128-133.
255. M. Albrecht, S. J. Franklin and K. N. Raymond, *Inorg. Chem.*, 1994, **33**, 5785-5793.
256. K. A. Mies, P. Gebhardt, U. Möllmann and A. L. Crumbliss, *J. Inorg. Biochem.*, 2008, **102**, 850-861.
257. M. Meyer, J. R. Telford, S. M. Cohen, D. J. White, J. Xu and K. N. Raymond, *J. Am. Chem. Soc.*, 1997, **119**, 10093-10103.
258. S. Dhungana, S. Heggemann, L. Heinisch, U. Möllmann, H. Boukhalfa and A. L. Crumbliss, *Inorg. Chem.*, 2001, **40**, 7079-7086.
259. S. Dhungana, J. M. Harrington, P. Gebhardt, U. Möllmann and A. L. Crumbliss, *Inorg. Chem.*, 2007, **46**, 8362-8371.
260. W. R. Harris, K. N. Raymond and F. L. Weitzl, *J. Am. Chem. Soc.*, 1981, **103**, 2667-2675.
261. W. R. Harris, C. J. Carrano, S. R. Cooper, S. R. Sofen, A. E. Avdeef, J. V. McArdle and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 6097-6104.
262. W. R. Harris, F. L. Weitzl and K. N. Raymond, *J. Chem. Soc., Chem. Commun.*, 1979, **4**, 177-178.

263. E. Olshvang, A. Szebesczyk, H. Kozlowski, Y. Hadar, E. Gumienna-Kontecka and A. Shanzer, *Dalton Trans.*, 2015, **44**, 20850-20858.
264. D. J. Ecker, L. D. Loomis, M. E. Cass and K. N. Raymond, *J. Am. Chem. Soc.*, 1988, **110**, 2457-2464.
265. B. Birkmann, W. W. Seidel, T. Pape, A. W. Ehlers, K. Lammertsma and F. E. Hahn, *Dalton Trans.*, 2009, **0**, 7350-7352.
266. F. L. Weitz and K. N. Raymond, *J. Am. Chem. Soc.*, 1980, **102**, 2289-2293.
267. K. M. C. Jurchen and K. N. Raymond, *Inorg. Chem.*, 2006, **45**, 2438-2447.
268. A.-K. Duhme, Z. Dauter, R. C. Hider and S. Pohl, *Inorg. Chem.*, 1996, **35**, 3059-3061.
269. N. F. Olivieri and G. M. Brittenham, *Blood*, 1997, **89**, 739-761.
270. Y. Kohgo, K. Ikuta, T. Ohtake, Y. Torimoto and J. Kato, *Int. J. Hematol.*, 2008, **88**, 7-15.
271. Z. D. Liu and R. C. Hider, *Coord. Chem. Rev.*, 2002, **232**, 151-171.
272. H. E. VanOrden and T. M. Hagemann, *Ann. Pharmacother.*, 2006, **40**, 1110-1117.
273. E. J. Neufeld, *Blood*, 2006, **107**, 3436-3441.
274. Z. D. Liu and R. C. Hider, *Med. Res. Rev.*, 2002, **22**, 26-64.
275. C. J. Carrano and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 5401-5404.
276. T. Zhou, X. L. Kong, Z. D. Liu, D. Y. Liu and R. C. Hider, *Biomacromolecules*, 2008, **9**, 1372-1380.
277. Y. J. Zhou, X. L. Kong, J. P. Li, Y. M. Ma, R. C. Hider and T. Zhou, *Med. Chem. Commun.*, 2015, **6**, 1620-1625.
278. J. Qian, B. P. Sullivan, S. J. Peterson and C. Berkland, *ACS Macro Lett.*, 2017, **6**, 350-353.
279. A. Leydier, D. Lecerclé, S. Pellet-Rostaing, A. Favre-Reguillon, F. Taran and M. Lemaire, *Tetrahedron*, 2008, **64**, 6662-6669.
280. A. E. Gorden, J. Xu, K. N. Raymond and P. Durbin, *Chem. Rev.*, 2003, **103**, 4207-4282.
281. Q. Zhang, B. Jin, R. Peng, S. Lei and S. Chu, *Polyhedron*, 2015, **87**, 417-423.

282. Q. Zhang, B. Jin, Z. Shi, X. Wang, Q. Liu, S. Lei and R. Peng, *Sci. Rep.*, 2016, **6**, 1-12.
283. R. J. Bergeron, J. Wiegand, J. S. McManis and N. Bharti, *J. Med. Chem.*, 2014, **57**, 9259-9291.
284. M. Petrik, C. Zhai, Z. Novy, L. Urbanek, H. Haas and C. Decristoforo, *Mol. Imaging Biol.*, 2016, **18**, 344-352.
285. M. Petrik, C. Zhai, H. Haas and C. Decristoforo, *Clin. Transl. Imaging*, 2017, **5**, 15-27.
286. A. Szebesczyk, E. Olshvang, A. Shanzer, P. L. Carver and E. Gumienna-Kontecka, *Coord. Chem. Rev.*, 2016, **327**, 84-109.
287. C. Queiros, A. Leite, G. M. C. M, L. Cunha-Silva, G. Barone, B. de Castro, M. Rangel, M. N. S. A and M. G. S. A, *Chemistry*, 2015, **21**, 15692-15704.
288. D. W. Domaille, E. L. Que and C. J. Chang, *Nat. Chem. Biol.*, 2008, **4**, 168-175.
289. R. Nudelman, O. Ardon, Y. Hadar, Y. Chen, J. Libman and A. Shanzer, *J. Med. Chem.*, 1998, **41**, 1671-1678.
290. K. Ferreira, H. Y. Hu, V. Fetz, H. Prochnow, B. Rais, P. P. Müller and M. Brönstrup, *Angew. Chem., Int. Ed. Engl.*, 2017, **56**, 8272-8276.
291. Y.-J. Zhou, M.-S. Liu, A. R. Osamah, X.-L. Kong, S. Alsam, S. Battah, Y.-Y. Xie, R. C. Hider and T. Zhou, *Eur. J. Med. Chem.*, 2015, **94**, 8-21.
292. I. J. Schalk and G. L. A. Mislin, *J. Med. Chem.*, 2017, **60**, 4573-4576.
293. M. Saha, S. Sarkar, B. Sarkar, B. K. Sharma, S. Bhattacharjee and P. Tribedi, *Environ. Sci. Pollut. Res.*, 2016, **23**, 3984-3999.
294. S. Fardeau, A. Dassonville-Klimpt, N. Audic, A. Sasaki, M. Pillon, E. Baudrin, C. Mullié and P. Sonnet, *Bioorg. Med. Chem.*, 2014, **22**, 4049-4060.
295. T. J. Sanderson, PhD Thesis, University of York, 2016.
296. T. Zheng and E. M. Nolan, *J. Am. Chem. Soc.*, 2014, **136**, 9677-9691.
297. P. Chairatana, T. F. Zheng and E. M. Nolan, *Chem. Sci.*, 2015, **6**, 4458-4471.
298. L. Tan, Y. L. Tao, T. Wang, F. Zou, S. H. Zhang, Q. H. Kou, A. Niu, Q. Chen, W. J. Chu, X. Y. Chen, H. D. Wang and Y. S. Yang, *J. Med. Chem.*, 2017, **60**, 2669-2684.

299. S. J. Milner, A. Seve, A. M. Snelling, G. H. Thomas, K. G. Kerr, A. Routledge and A. K. Duhme-Klair, *Org. Biomol. Chem.*, 2013, **11**, 3461-3468.
300. A. Kim, A. Kutschke, D. E. Ehmman, S. A. Patey, J. L. Crandon, E. Gorseth, A. A. Miller, R. E. McLaughlin, C. M. Blinn, A. Chen, A. S. Nayar, B. Dangel, A. S. Tsai, M. T. Rooney, K. E. Murphy-Benenato, A. E. Eakin and D. P. Nicolau, *Antimicrob. Agents Chemother.*, 2015, **59**, 7743-7752.
301. C. Ji and M. Miller, *Biometals*, 2015, **28**, 541-551.
302. A. Paulen, V. Gasser, F. Hoegy, Q. Perraud, B. Pesset, I. J. Schalk and G. L. A. Mislin, *Org. Biomol. Chem.*, 2015, **13**, 11567-11579.
303. S. K. Boda, S. Pandit, A. Garai, D. Pal and B. Basu, *RSC Adv.*, 2016, **6**, 39245-39260.
304. T. E. Clarke, S.-Y. Ku, D. R. Dougan, H. J. Vogel and L. W. Tari, *Nat. Struct. Mol. Biol.*, 2000, **7**, 287-291.
305. F. Peuckert, Ana L. Ramos-Vega, M. Miethke, Clemens J. Schwörer, Alexander G. Albrecht, M. Oberthür and Mohamed A. Marahiel, *Chem. Biol.*, 2011, **18**, 907-919.
306. A. J. Wilkinson and K. H. Verschueren, *ABC Proteins: From Bacteria to Man*, 2003, 187-208.
307. H. R. Strange, T. A. Zola and C. N. Cornelissen, *Infect. Immun.*, 2011, **79**, 267-278.
308. A. Stintzi, C. Barnes, J. Xu and K. N. Raymond, *Proc. Natl. Acad. Sci.*, 2000, **97**, 10691-10696.
309. D. Cobessi, H. Celia and F. Pattus, *J. Mol. Biol.*, 2005, **352**, 893-904.
310. E. J. Wilde, A. Hughes, E. V. Blagova, O. V. Moroz, R. P. Thomas, J. P. Turkenburg, D. J. Raines, A.-K. Duhme-Klair and K. S. Wilson, *Sci. Rep.*, 2017, **7**, 45941-45955.
311. M. Miethke, *Metallomics*, 2013, **5**, 15-28.
312. D. Hofreuter, *Front. Cell. Infect. Microbiol.*, 2014, **4**, 1-19.
313. W. A. Agger, J. McCormick and M. J. Gurwith, *J. Clin. Microbiol.*, 1985, **21**, 909-913.
314. H. E. Bok, W. H. Holzapfel, E. S. Odendaal and H. J. van der Linde, *Int. J. Food Microbiol.*, 1986, **3**, 273-285.

315. A. du Moulinet d'Hardemare, S. Torelli, G. Serratrice and J.-L. Pierre, *Biometals*, 2006, **19**, 349-366.
316. C. Queirós, A. Leite, A. M. G. Silva, P. Gameiro, B. de Castro and M. Rangel, *Polyhedron*, 2015, **87**, 1-7.
317. E. Krissinel and K. Henrick, *Acta Crystallogr. D.*, 2004, **60**, 2256-2268.
318. W. B. Jensen, *Chem. Rev.*, 1978, **78**, 1-22.
319. H. Edelhoch, *Biochemistry*, 1967, **6**, 1948-1954.
320. S. Agarwal, S. Dey, B. Ghosh, M. Biswas and J. Dasgupta, *Sci. Rep.*, 2017, **7**, 1-15.
321. S. Banerjee, A. J. Weerasinghe, C. J. Parker Siburt, R. T. Kreulen, S. K. Armstrong, T. J. Brickman, L. A. Lambert and A. L. Crumbliss, *Biochemistry*, 2014, **53**, 3952-3960.
322. D. J. Raines, PhD Thesis, University of York, 2014.
323. B. Birdsall, R. W. King, M. R. Wheeler, C. A. Lewis, S. R. Goode, R. B. Dunlap and G. C. Roberts, *Anal. Biochem.*, 1983, **132**, 353-361.
324. L. M. Mayer, L. L. Schick and T. C. Loder I, *Mar. Chem.*, 1999, **64**, 171-179.
325. A. J. Lawaetz and C. A. Stedmon, *Appl. Spectrosc.*, 2009, **63**, 936-940.
326. P. Kuzmič, *Anal. Biochem.*, 1996, **237**, 260-273.
327. I. Jelesarov and H. R. Bosshard, *J. Mol. Recogn.*, 1999, **12**, 3-18.
328. A. Velázquez-Campoy, H. Ohtaka, A. Nezami, S. Muzammil and E. Freire, *Curr. Protoc. Cell Biol.*, 2004, 17.18. 11-17.18. 24.
329. C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun and S. Duhr, *Nat. Commun.*, 2010, **1**, 1-7.
330. T. B. Karpishin, T. D. P. Stack and K. N. Raymond, *J. Am. Chem. Soc.*, 1993, **115**, 6115-6125.
331. C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. v. Streek and P. A. Wood, *J. Appl. Crystallogr.*, 2008, **41**, 466-470.
332. C.-e. A. Chang, W. Chen and M. K. Gilson, *Proc. Natl. Acad. Sci.*, 2007, **104**, 1534-1539.

333. B. Lu and C. F. Wong, *Biopolymers*, 2005, **79**, 277-285.
334. S. J. Irudayam and R. H. Henchman, *J. Phys. Chem. B.*, 2009, **113**, 5871-5884.
335. G. Richarme and T. D. Caldas, *J. Biol. Chem.*, 1997, **272**, 15607-15612.
336. T. Heinisch and T. R. Ward, *Curr. Opin. Chem. Biol.*, 2010, **14**, 184-199.
337. M. A. Dwyer and H. W. Hellinga, *Curr. Opin. Struct. Biol.*, 2004, **14**, 495-504.
338. L. Holm and P. Rosenström, *Nucleic Acids Res.*, 2010, **38**, W545-W549.
339. E. Krissinel and K. Henrick, in *Computational Life Sciences: First International Symposium, CompLife 2005, Konstanz, Germany, September 25-27, 2005. Proceedings*, eds. M. R. Berthold, R. C. Glen, K. Diederichs, O. Kohlbacher and I. Fischer, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 67-78.
340. E. B. Krissinel and K. Henrick, *Software Pract. Exper.*, 2004, **34**, 591-607.
341. E. Krissinel, M. Winn, C. Ballard, A. Ashton, P. Patel, E. Potterton, S. McNicholas, K. Cowtan and P. Emsley, *Acta. Crystallog. D.*, 2004, **60**, 2250-2255.
342. P. T. Richardson and S. F. Park, *Microbiology*, 1995, **141**, 3181-3191.
343. H. Bijlmer, L. Van Alphen, B. Greenwood, J. Brown, G. Schneider, A. Hughes, A. Menon, H. Zanen and H. Valkenburg, *J. Infect. Dis.*, 1990, **161**, 1210-1215.
344. K. A. Fenton, C. Korovessis, A. M. Johnson, A. McCadden, S. McManus, K. Wellings, C. H. Mercer, C. Carder, A. J. Copas and K. Nanchahal, *Lancet*, 2001, **358**, 1851-1854.
345. A. Everard, C. Belzer, L. Geurts, J. P. Ouwerkerk, C. Druart, L. B. Bindels, Y. Guiot, M. Derrien, G. G. Muccioli and N. M. Delzenne, *Proc. Natl. Acad. Sci.*, 2013, **110**, 9066-9071.
346. M. Schneeberger, A. Everard, A. G. Gómez-Valadés, S. Matamoros, S. Ramírez, N. M. Delzenne, R. Gomis, M. Claret and P. D. Cani, *Sci. Rep.*, 2015, **5**, 16643-16657.
347. M. C. Enright, D. A. Robinson, G. Randle, E. J. Feil, H. Grundmann and B. G. Spratt, *Proc. Natl. Acad. Sci.*, 2002, **99**, 7687-7692.
348. R. Leclercq, E. Derlot, J. Duval and P. Courvalin, *N. Engl. J. Med.*, 1988, **319**, 157-161.
349. L. Mundy, D. Sahm and M. Gilmore, *Clin. Microbiol. Rev.*, 2000, **13**, 513-522.

350. V. A. Chirurghi, S. E. Oster, A. A. Goldberg, M. Zervos and R. McCabe, *J. Clin. Microbiol.*, 1991, **29**, 2663-2665.
351. J. S. Brown, A. D. Ogunniyi, M. C. Woodrow, D. W. Holden and J. C. Paton, *Infect. Immun.*, 2001, **69**, 6702-6706.
352. R. Cripps, K. Eley, D. J. Leak, B. Rudd, M. Taylor, M. Todd, S. Boakes, S. Martin and T. Atkinson, *Metab. Eng.*, 2009, **11**, 398-408.
353. B. S. Hartley and G. Shama, *Phil. Trans. R. Soc. A*, 1987, **321**, 555-568.
354. Y. Suzuki, T. Kishigami, K. Inoue, Y. Mizoguchi, N. Eto, M. Takagi and S. Abe, *Syst. Appl. Microbiol.*, 1983, **4**, 487-495.
355. J. Collot, J. Gradinaru, N. Humbert, M. Skander, A. Zocchi and T. R. Ward, *J. Am. Chem. Soc.*, 2003, **125**, 9030-9031.
356. C. M. Thomas and T. R. Ward, *Chem. Soc. Rev.*, 2005, **34**, 337-346.
357. J. Steinreiber and T. R. Ward, *Coord. Chem. Rev.*, 2008, **252**, 751-766.
358. P. P. Lin, K. S. Rabe, J. L. Takasumi, M. Kadisch, F. H. Arnold and J. C. Liao, *Metab. Eng.*, 2014, **24**, 1-8.
359. L. J. Van Zyl, M. P. Taylor, K. Eley, M. Tuffin and D. A. Cowan, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 1247-1259.
360. C. S. Gibbs and M. J. Zoller, *J. Biol. Chem.*, 1991, **266**, 8923-8931.
361. T. M. Savarese and C. M. Fraser, *Biochem. J.*, 1992, **283**, 1-19.
362. G. Dalbadie-McFarland, L. W. Cohen, A. D. Riggs, C. Morin, K. Itakura and J. H. Richards, *Proc. Natl. Acad. Sci.*, 1982, **79**, 6409-6413.
363. D. R. Clemmons, *Endocr. Rev.*, 2001, **22**, 800-817.
364. P. C. Ng and S. Henikoff, *Nucleic Acids Res.*, 2003, **31**, 3812-3814.
365. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta. Crystallog. D.*, 2010, **66**, 486-501.
366. C. Kurth, H. Kage and M. Nett, *Org. Biomol. Chem.*, 2016, **14**, 8212-8227.
367. T. C. Johnstone and E. M. Nolan, *Dalton Trans.*, 2015, **44**, 6320-6339.
368. C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz, *Science*, 1989, **244**, 182-188.

369. J. Ellman, D. Mendel, S. Anthony-Cahill, C. J. Noren and P. G. Schultz, *Method. Enzymol.*, 1991, **202**, 301-336.
370. J. Y. Axup, K. M. Bajjuri, M. Ritland, B. M. Hutchins, C. H. Kim, S. A. Kazane, R. Halder, J. S. Forsyth, A. F. Santidrian and K. Stafin, *Proc. Natl. Acad. Sci.*, 2012, **109**, 16101-16106.
371. T. Schmiederer, S. Rausch, M. Valdebenito, Y. Mantri, E. Mösker, T. Baramov, K. Stelmaszyk, P. Schmieder, D. Butz, S. I. Müller, K. Schneider, M.-H. Baik, K. Hantke and R. D. Süßmuth, *Angew. Chem., Int. Ed. Engl.*, 2011, **50**, 4230-4233.
372. M. A. Fischbach, H. Lin, L. Zhou, Y. Yu, R. J. Abergel, D. R. Liu, K. N. Raymond, B. L. Wanner, R. K. Strong and C. T. Walsh, *Proc. Natl. Acad. Sci.*, 2006, **103**, 16502-16507.
373. R. J. Abergel, E. G. Moore, R. K. Strong and K. N. Raymond, *J. Am. Chem. Soc.*, 2006, **128**, 10998-10999.
374. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187-217.
375. P. G. Hultin, *Curr. Top. Med. Chem.*, 2005, **5**, 1299-1331.
376. T. Bililign, B. R. Griffith and J. S. Thorson, *Nat. Prod. Rep.*, 2005, **22**, 742-760.
377. K. Lalitha, K. Muthusamy, Y. S. Prasad, P. K. Vemula and S. Nagarajan, *Carbohydr. Res.*, 2015, **402**, 158-171.
378. K. W. Wellington and S. A. Benner, *Nucleos. Nucleot. Nucl.*, 2006, **25**, 1309-1333.
379. C. Jaramillo and S. Knapp, *Synthesis*, 1994, **1994**, 1-20.
380. Y. Du, R. J. Linhardt and I. R. Vlahov, *Tetrahedron*, 1998, **54**, 9913-9959.
381. D. Y. Lee and M. He, *Curr. Top. Med. Chem.*, 2005, **5**, 1333-1350.
382. M. J. McKay and H. M. Nguyen, *ACS Catal.*, 2012, **2**, 1563-1595.
383. P. Merino, T. Tejero and E. Marca, *Heterocycles*, 2012, **86**, 791-820.
384. K. Villadsen, M. C. Martos-Maldonado, K. J. Jensen and M. B. Thygesen, *ChemBioChem*, 2017, **18**, 574-612.
385. C. D. Hurd and W. A. Bonner, *J. Am. Chem. Soc.*, 1945, **67**, 1664-1668.

386. P. Verlhac, C. Leteux, L. Toupet and A. Veyrières, *Carbohydr. Res.*, 1996, **291**, 11-20.
387. H. Böshagen, W. Geiger and B. Junge, *Angew. Chem., Int. Ed. Engl.*, 1981, **20**, 806-807.
388. J. Panigot and W. Curley Jr, *J. Carbohydr. Chem.*, 1994, **13**, 293-302.
389. S. Hainke, I. Singh, J. Hemmings and O. Seitz, *J. Org. Chem.*, 2007, **72**, 8811-8819.
390. R. Bihovsky, C. Selick and I. Giusti, *J. Org. Chem.*, 1988, **53**, 4026-4031.
391. G. A. Kraus and M. T. Molina, *J. Org. Chem.*, 1988, **53**, 752.
392. D. E. Levy and C. Tang, *The Chemistry of C-glycosides*, Elsevier, 1995.
393. A. H. Lee and E. T. Kool, *J. Org. Chem.*, 2005, **70**, 132-140.
394. C. R. Johnson and B. A. Johns, *Synlett*, 1997, **12**, 1406-1408.
395. H. Gong, R. Sinisi and M. R. Gagné, *J. Am. Chem. Soc.*, 2007, **129**, 1908-1909.
396. H. G. Gong and M. R. Gagne, *J. Am. Chem. Soc.*, 2008, **130**, 12177-12183.
397. X. Yu, Y. Dai, T. Yang, M. R. Gagné and H. Gong, *Tetrahedron*, 2011, **67**, 144-151.
398. W. J. Barreto, R. A. Ando, P. S. Santos and W. P. Silva, *Spectrochim. Acta Mol. Biomol. Spectrosc.*, 2007, **68**, 612-618.
399. M. J. Sever and J. J. Wilker, *Dalton Trans.*, 2004, 1061-1072.
400. C. Fischer and G. C. Fu, *J. Am. Chem. Soc.*, 2005, **127**, 4594-4595.
401. E. W. Yue, J. M. Gerdes and C. A. Mathis, *J. Org. Chem.*, 1991, **56**, 5451-5456.
402. W. A. Bubb, *Concept Magnetic Res. A*, 2003, **19A**, 1-19.
403. B. G. Davis and A. J. Fairbanks, *Carbohydrate Chemistry*, Oxford University Press, 2002.
404. T. Greene and P. Wuts, *Protective Groups in Organic Synthesis*, John Wiley & Sons, Canada, Second edn., 1991.
405. A. Chimiak and J. B. Neilands, in *Siderophores from Microorganisms and Plants*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1984, pp. 89-96.

406. J. S. Renny, L. L. Tomasevich, E. H. Tallmadge and D. B. Collum, *Angew. Chem., Int. Ed. Engl.*, 2013, **52**, 11998-12013.
407. T. E. Geswindt, W. J. Gerber, H. E. Rohwer and K. R. Koch, *Dalton Trans.*, 2011, **40**, 8581-8588.
408. K. J. Winstanley and D. K. Smith, *J. Org. Chem.*, 2007, **72**, 2803-2815.
409. J. T. Sheff, A. L. Lucius, S. B. Owens and G. M. Gray, *Organometallics*, 2011, **30**, 5695-5709.
410. S. L. Jewett, S. Egging and L. Geller, *J. Inorg. Biochem.*, 1997, **66**, 165-173.
411. F. M. Raymo and J. F. Stoddart, *Chem. Ber.*, 1996, **129**, 981-990.
412. M. L. Saha, S. De, S. Pramanik and M. Schmittel, *Chem. Soc. Rev.*, 2013, **42**, 6860-6909.
413. M. Albrecht, *Chem. Rev.*, 2001, **101**, 3457-3498.
414. T. D. Pollard, *Mol. Biol. Cell*, 2010, **21**, 4061-4067.
415. M. K. Gilson and H.-X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.*, 2007, **36**, 21-42.
416. E. Calva, J. Torres, M. Vázquez, V. Angeles, H. de la Vega and G. M. Ruiz-Palacios, *Gene*, 1989, **75**, 243-251.
417. R. B. Sack, M. Rahman, M. Yunus and E. H. Khan, *Clin. Infect. Dis.*, 1997, **24**, S102-S105.
418. D. J. Bolton, *Food Microbiol.*, 2015, **48**, 99-108.
419. R. N. Hvorup, B. A. Goetz, M. Niederer, K. Hollenstein, E. Perozo and K. P. Locher, *Science*, 2007, **317**, 1387-1390.
420. E. L. Borths, K. P. Locher, A. T. Lee and D. C. Rees, *Proc. Natl. Acad. Sci.*, 2002, **99**, 16642-16647.
421. V. Braun and C. Herrmann, *J. Bacteriol.*, 2007, **189**, 6913-6918.
422. M. T. Sebulsky, B. H. Shilton, C. D. Speziali and D. E. Heinrichs, *J. Biol. Chem.*, 2003, **278**, 49890-49900.
423. E. Fischer, B. Strehlow, D. Hartz and V. Braun, *Arch. Microbiol.*, 1990, **153**, 329-336.

424. E. D. Peng, E. E. Wyckoff, A. R. Mey, C. R. Fisher and S. M. Payne, *Infect. Immun.*, 2016, **84**, 511-523.
425. P. Rice, I. Longden and A. Bleasby, *Trends Genet.*, 2000, **16**, 276-277.
426. R. C. Edgar, *Nucleic Acids Res.*, 2004, **32**, 1792-1797.
427. P. Gouet, E. Courcelle and D. I. Stuart, *Bioinformatics*, 1999, **15**, 305-308.
428. T. U. Consortium, *Nucleic Acids Res.*, 2017, **45**, D158-D169.
429. Z. R. Yang, R. Thomson, P. McNeil and R. M. Esnouf, *Bioinformatics*, 2005, **21**, 3369-3376.
430. S. M. Kelly and N. C. Price, *BBA- Protein Struct. Mol. Enzymol.*, 1997, **1338**, 161-185.
431. N. J. Greenfield, *Nat. Protoc.*, 2006, **1**, 2876-2890.
432. Y. Ma, Y. Xie and R. C. Hider, *Analyst*, 2013, **138**, 96-99.
433. F. J. Reen, S. Almagro-Moreno, D. Ussery and E. F. Boyd, *Nat. Rev. Micro.*, 2006, **4**, 697-704.
434. S. R. Cooper, J. V. McArdle and K. N. Raymond, *Proc. Natl. Acad. Sci.*, 1978, **75**, 3551-3554.
435. T. B. Karpishin and K. N. Raymond, *Angew. Chem., Int. Ed. Engl.*, 1992, **31**, 466-468.
436. T. Stack, Z. Hou and K. N. Raymond, *J. Am. Chem. Soc.*, 1993, **115**, 6466-6467.
437. D. Seebach, H. M. Müller, H. M. Bürger and D. A. Plattner, *Angew. Chem., Int. Ed. Engl.*, 1992, **31**, 434-435.
438. A. V. Joshua, S. K. Sharma and D. N. Abrams, *Synth. Commun.*, 2008, **38**, 434-440.
439. P. C. H. Eichinger, J. H. Bowie and T. Blumenthal, *J. Org. Chem.*, 1986, **51**, 5078-5082.
440. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T. L. Madden, *BMC Bioinformatics*, 2009, **10**, 421-421.
441. W. L. Perry III, *Biotechniques*, 2002, **33**, 1318-1320.
442. T. Vincze, J. Posfai and R. J. Roberts, *Nucleic Acids Res.*, 2003, **31**, 3688-3691.

443. E. Gasteiger, C. Hoogland, A. Gattiker, S. e. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, *Protein identification and analysis tools on the ExPASy server*, Springer, 2005.
444. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie and A. McCoy, *Acta. Crystallog. D.*, 2011, **67**, 235-242.
445. P. Evans, *Acta. Crystallog. D.*, 2006, **62**, 72-82.
446. G. Winter, *J. Appl. Crystallogr.*, 2010, **43**, 186-190.
447. W. Kabsch, *Acta. Crystallog. D.*, 2010, **66**, 125-132.
448. A. Vagin and A. Teplyakov, *J. Appl. Crystallogr.*, 1997, **30**, 1022-1025.
449. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.
450. G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta. Crystallog. D.*, 2011, **67**, 355-367.
451. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, J. Snoeyink and J. S. Richardson, *Nucleic Acids Res.*, 2007, **35**, W375-W383.
452. A. A. Lebedev, P. Young, M. N. Isupov, O. V. Moroz, A. A. Vagin and G. N. Murshudov, *Acta. Crystallog. D.*, 2012, **68**, 431-440.
453. L. Lyons, *A Practical Guide to Data Analysis for Physical Science Students*, Cambridge University Press, 1991.
454. A. Seiffert, K. Goeke, H.-P. Fielder and H. Zähler, *Biotechnol. Bioeng.*, 1993, **41**, 237-244.