Iron Chelators Designed to Study Aspects of Siderophore-Mediated Iron(III) Uptake

Daniel J. Raines

Doctor of Philosophy

University of York
Chemistry

September 2014

Abstract

The investigation of the chemical characteristics of siderophores has been of interest for many years, for both the understanding of their fundamental chemistry and exploiting the design features of these natural compounds for novel applications.

Within the present project, the properties of a biomimetic model of the tetradentate siderophore known as the enterobactin linear dimer was investigated, followed by those of the enterobactin linear dimer itself. Both compounds were chemically synthesised, with an overall yield of 52% in four steps for the biomimetic, and 26% in seven-step convergence synthesis for the linear dimer. Investigation of the coordination chemistry, showing both can form complexes with a ligand-to-metal ratios of 3:2 and 1:1 in solution. This was followed up by X-ray diffraction investigations of the ferric-complexes with the periplasmic binding protein CeuE. Both ferric-complexes bound to CeuE in a novel mode, with conserved tyrosine and histidine residues binding directly to the iron(III) centre to complete the octahedral coordination of the metal. Both complexes have a Λ metal centre configuration when bound to CeuE. The dissociation constants of CeuE with ferric-4-LICAM and ferric-linear dimer were determined to be 29.3 ± 11.7 nM and 8.4 ± 4.3 nM for the ferric-4-LICAM and ferric-linear dimer, respectively.

In addition, selected applications of siderophores were investigated. Firstly, a synthetic route to a novel hexadentate siderophore mimic possessing a chemical linking group was developed, in an eight-step synthesis with an overall yield of 10.5%. Secondly, the natural hydroxamate siderophore ferricrocin was labelled with an Alexa Fluor® 488 fluorophore. Total internal reflection fluorescence microscopy was used to demonstrate that the resulting conjugate binds to *E. coli* cells that possess the ferricrocin outer membrane receptor protein FhuA.

Table of Contents

Abstract 2
Table of Contents 3
Table of Figures, Schemes, Tables and Equations 6
Accompanying Material 19
Acknowledgements 20
Declaration 22

Chapter 1: Introduction 23
1.1 Overview 24
1.2 Iron in Biology 24
1.3 Siderophores 25
Siderophore Mimics 31
1.4 Siderophore Mediated Iron Uptake 35
Gram-negative bacteria 35
Gram-positive bacteria 41
Iron release 41
Regulation of Iron Uptake 42
Campylobacter jejuni 44
1.5 Applications of Siderophores 46
Siderophores in Supramolecular Chemistry 46
Medicinal Applications of Siderophores 49
Other Applications of Siderophores 51
1.6 Project Aims 55

Chapter 2: Tetradentate Siderophores, a Biomimetic Approach 56
Project Aims 57
2.1 Introduction 57
H_4-4-LICAM 58
2.2 Synthesis of H_4-4-LICAM 59
2.3 Ligand-Metal Interactions 60

Stoichiometric Ratio Determination of H_4-4-LICAM and Iron(III) using the Continuous Variation Method 61
Job Plot Analysis of H_4-4-LICAM and Iron(III) Nitrate in the Presence of NTA 66
^{1}H NMR Spectroscopic Analysis of H_4-4-LICAM and Gallium(III) Nitrate . 70
2.4 CeuE-Complex Interactions 77
Expression and Purification 78
Crystallisation of apo-CeuE 81
Crystallisation of the CeuE, Ferric-4-LICAM Structure 85
Circular Dichroism and Fluorescence Biophysical Measurements 98
2.5 Summary of Chapter 2 105

Chapter 3: Tetradentate Siderophores, a Natural Product Approach 106
Project Aims 107
3.1 Introduction 107
Previous Synthetic Approaches 109
The *N*-benzoyl Monomer Approach 110
The Urethane Protection Route 112
3.2 Synthesis 114
Initial Approaches to the Synthesis of the Enterobactin Linear Dimer 114
Synthesis of the Enterobactin Linear Dimer 119
3.3 Ligand-Metal Interactions 122
Investigation of 17^{5-} to Iron(III) Ratios using the Continuous Variation Method 127
Circular Dichroism Spectroscopic Study of the H_4-17 and Iron(III) System 130
3.4 Interactions of the Ferric Complex with CeuE 133
Crystallisation of the CeuE, Ferric-Linear Dimer Complex 138
Circular Dichroism and Fluorescence Biophysical Measurements 153
3.5 Summary of Chapter 3 158

Chapter 4: A Siderophore Mimic for use with Further Functional Components 159
Project Aims 160
4.1 Introduction 160

4.2 Synthesis 166
Synthesis of a Biomimetic Hexadentate Siderophore 166
4.4 Summary of Chapter 4 176

Chapter 5: A Natural Siderophore with a Further Fluorescent Component 177
Project Aims 178
5.1 Introduction 178
5.2 Synthesis 184
5.3 Conjugate-Protein Interactions 192
5.4 Summary of Chapter 5 195

Chapter 6: Conclusions and Future Work 196
6.1 Conclusions and Future Work 197
Interactions of Tetradentate Siderophores with CeuE (Chapters 2 and 3) 197
Development of a Siderophore Mimic for use with Further Functional Components (Chapter 4) 198
Modification of Ferricrocin with a Fluorescent Component (Chapter 5) 199

Chapter 7: Experimental 200
7.1 General Remarks 201
Materials 201
Instrumentation 201
7.2 Synthesis 202
7.3 Protein Production and Crystallisation Trials 245
7.4 Job Plot Method 247
7.5 Circular Dichroism 248
7.6 Fluorescence Quenching Titration 249

Appendix I. Fluorescent Quenching Analysis of CeuE 251
Abbreviations 260
References 263

Table of Figures, Schemes, Tables and Equations

Figure 1 - Representative examples of siderophores with different iron chelating moieties highlighted. *Protonation state of citric acid can differ due to different binding modes available. 26

Figure 2 - Schematic diagram demonstrating the hydrogen bonding network and conformational change of enterobactin upon A) single deprotonation; B) iron complexation. Figure adapted from *Proc. Natl. Acad. Sci.*, 2003, *100*, (7), 3584-3588. Copyright 2003 National Academy of Sciences, USA. 28

Figure 3 - Chemical structures of the enterobactin hydrolysis products. 29

Figure 4 - Chemical structures of a selection of 'low' dentate siderophores. 30

Figure 5 - Λ-configuration and Δ-configuration of a tris-bidentate octahedral complex. 30

Figure 6 - A selection of structures of biomimetic hexadentate catecholamide siderophores and enterobactin, all possessing high affinity for iron(III).[26-29] 32

Figure 7 – Illustration of the preorganised *ababab* configuration of $(Et)_3$MECAM (A) compared to MECAM (B). 33

Figure 8 - Schematic diagram of the iron-uptake mechanism for a Gram-negative bacterium. 36

Figure 9 - Ribbon representation of the ferric enterobactin outer membrane receptor FepA, shown from: A) Side on; B) Top down (from the extracellular space); C) Beneath (from periplasmic space) Key: Barrel domain (light blue), Cork domain (red). PDB code: 1FEP.[38] 37

Figure 10 - Ribbon representation of PBPs: A) VctP (PDB code: 3TEF)[47]; B) FhuD bound to desferral (PDB code: 1K2V)[46]; C) CeuE bound to ferric-MECAM (PDB code: 2CHU)[35]. Key: PBP (light blue), ligands shown as spheres; grey - carbon, blue - nitrogen, red - oxygen, coral - iron. 38

Figure 11 - Ribbon representation of $BtuC_2D_2F$ (PDB code 2QI9)[49]: A) Dimer of $BtuC_2$ (protomers shown in white and grey); B) Dimer of $BtuD_2$ (protomers shown in light blue and coral); C) Complex of $BtuC_2D_2F$. Key: $BtuC_2$ (white

and grey), $BtuD_2$ (blue and coral), BtuF (gold) key docking residues shown as spheres; red - arginine, green - glutamate. 40
Figure 12 - A schematic diagram for a simplified overview of the iron-uptake mechanism for Gram-positive bacteria. 41
Figure 13 - Schematic diagram demonstrating the pH dependence of the catecholate and salicylate binding modes of the catecholamide group. 42
Figure 14 - A schematic diagram demonstrating the general mechanism of action of Fur. 43
Figure 15 - Schematic representation of the three major iron uptake systems of *C. jejuni*. 44
Figure 16 - The structure of the $\{CeuE_2[Fe_2MECAM_2]\}$ complex (PDB code: 2CHU).[35] A) CeuE chain A (light blue), CeuE chain B (gold), ligands shown as spheres; B) The MECAM molecules shown as cylinders. Key: grey - carbon, blue - nitrogen, red - oxygen, coral - iron. 45
Figure 17 - Schematic diagram of a triple helicate and triple mesocate. 46
Figure 18 - Example bis(catechol) ligands that have been shown to form triple helicates.[76] 47
Figure 19 - Structures of the two complexes 2-LICAM^{4-} forms with iron(III), $[Fe_2L_3]^{6-}$ and $[Fe_2L_2OH_2]^{4-}$. 48
Figure 20 - Structures of the natural siderophores rhodotorulic acid and alcaligin and schematic diagrams of their dinuclear six-coordinate iron complexes. 48
Figure 21 - Chemical structures of DGE, one of the salmochelins and H_6-9 designed and synthesised by Zheng *et al.*[112] 52
Figure 22 - Chemical structure of fluorescently-labelled Pyochelin.[121] 54
Figure 23 - Possible iron(III)-4-LICAM complexes that could form in solution. 61
Figure 24 - Job plot for the binding of H_2-**15** to iron (III), obtained by following the absorbance at 512 nm in the presence (red diamonds) and absence (blue triangles) of NTA. The absorbance values are averages of two experiments and the error bars indicate the differences between the runs. Plot recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. 65
Figure 25 – Selected UV-visible absorbance spectra of H_2-**15** and Fe^{III}(NTA), recorded in the presence of NTA. Spectra recorded in 0.1 M TrisHCl pH 7.5,

5% DMSO. Spectra ordered from black (H_2-**15**) to light blue ordered by ligand to metal ratio. 66
Figure 26 - Kinetic traces recorded for 1:2, 1:1 and 3:2 ratios of H_4-4-LICAM and Fe^{III}(NTA), following the absorbance at 498 nm (red diamonds) and 552 nm (blue triangles) over 6 minutes. Spectra recorded every 6 seconds, solution containing 0.1 M TrisHCl pH 7.5, 5% DMSO at 5 °C. Concentration of H_4-4-LICAM; $1.3x10^{-4}$ M; $2.0x10^{-4}$ M; $2.4x10^{-4}$ M for 1:2; 1:1 and 3:2 samples. 67
Figure 27 - Selected UV-visible absorbance spectra of H_4-4-LICAM and Fe^{III}(NTA), used for the Job plot analysis. Spectra recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. Spectra ordered from black (H_4-4-LICAM) to light blue ordered by ligand to metal ratio; (95:5, 85:15, 78:22, 68:32, 60:40, 50:50, 46:54, 30:70, 10:90). [M] + [L] = 0.4 mM. 69
Figure 28 - Job plot for the binding of H_4-4-LICAM and Fe^{III}(NTA), obtained by following the absorbance at both 498 nm (red diamonds) and 552 nm (blue triangles). The absorbance values are averages of two experiments and the error bars indicate the differences between the runs. Plot recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. 69
Figure 29 - Purposed schematic mechanism of formation of different (catechol)iron(III) complexes within samples for Job plot analysis. Actual speciation cannot be determined from these experiments. 70
Figure 30 - Aromatic region of the ^{1}H NMR spectra of 4-LICAM^{4-} in a mixture of d_6-DMSO/D_2O with varying amounts of $Ga(NO_3)_3$ present, with 4 equivalents of NaOH with respect to H_4-4-LICAM; (●) 4-LICAM^{4-}; (○) Complex A; (♦) Complex B. 71
Figure 31 - ^{1}H DOSY NMR spectrum of the 1:1 sample recorded after 1 month of equilibration in 5:1 d_6-DMSO : D_2O. Representations: (○) Complex A; (♦) Complex B. 74
Figure 32 - Models of the proposed 3:2 triple helical $[Fe_2(4\text{-LICAM})_3]^{6-}$ and 1:1 mononuclear $[Fe(4\text{-LICAM})]^-$ complexes, showing the centre of mass calculated by Avogadro and distance to the outer most atom. Key; grey - carbon, blue - nitrogen, red - oxygen, yellow - sulphur, white - hydrogen, coral sphere - iron, light blue - centre of mass. 76
Figure 33 - Full sequence of CeuE with the predicted signalling peptide highlighted as the black triangles. Image produced using ESPript.[45] 77

Figure 34 - SDS PAGE analysis of CeuE protein, after 48 hours at 4 °C with C-protease in a 100:1 ratio. Lane 1: Before treatment with C-protease; Lanes 2 and 3: CeuE after treatment of C-protease; Lanes 4-8: Cut out work not relating to this project; Lane 9: Ladder. 79
Figure 35 – A) UV-visible trace of final gel-filtration column monitoring at 280 nm and 254 nm. B) Final SDS PAGE analysis of CeuE, before and after the gel filtration step. Lane 1: Before Gel filtration column; Lanes 2-8: Fractions from gel filtration column; Lane 9: Molecular weight ladder. Pure CeuE in lanes 6 and 7 are associated with peak at 65 mL in UV-visible trace. 80
Figure 36 - Ribbon representation of apo-CeuE with colours blended from cyan, N-terminus, to dark blue, C-terminus. 81
Figure 37 - Surface electrostatic representation of apo-CeuE, negative charge shown in red and positive charge in blue. 83
Figure 38 - Ribbon representation of apo-CeuE (PDB code 3ZKW) superimposed with the dimeric {$CeuE_2[Fe_2MECAM_2]$} (PDB code 2CHU) structure.[40] ProSMART analysis shows similar regions in yellow and regions of difference in red (white for unaligned).[50] 84
Figure 39 - Ribbon representation of {CeuE[Fe(4-LICAM)]} with colours blended from cyan, N-terminus, to dark blue, C-terminus. His227, Tyr288 and Fe-4-LICAM shown as cylinders. Key: binding residues, green - carbon, blue - nitrogen, red - oxygen; Fe-4-LICAM, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.....86
Figure 40 - Stereoview of the iron chelating ligands in the {CeuE[Fe(4-LICAM)]} with electron density map shown as $2F_{obs}$-F_{calc}, contoured at 1σ and chemical representation showing key interactions of Fe-4-LICAM. Key: binding residues, green - carbon, blue - nitrogen, red - oxygen; Fe-4-LICAM, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral. 87
Figure 41 - Coordination of [Fe(4-LICAM)] in {CeuE[Fe(4-LICAM)]}. Key; CeuE residues, green - carbon, blue - nitrogen, red - oxygen; 4-LICAM, grey - carbon, blue - nitrogen, red - oxygen; Fe - Coral sphere. 89
Figure 42 - Ribbon representation of apo-CeuE (PDB code 3ZKW) superimposed with the dimeric {CeuE[Fe(4-LICAM)]} (PDB code 3ZK3) structure. ProSMART analysis shows similar regions in yellow and regions of

difference in red (white for unaligned).[50] 4-LICAM shown as cylinders- Grey, carbon - blue, nitrogen - red, oxygen; Fe - coral. 91
Figure 43 - Sequence alignment of selected PBPs: CeuE (*C. jejuni*), YclQ (*B. subtilis*), VctP (*V. cholerae*) and FetB (*N. meningitidis*), performed using T-coffee.[57] Red block colours show fully conserved residues; red triangles show the position of basic triad of arginine residues; blue stars indicate position of coordinating histidine and tyrosine residues. Image produced using ESPript.[45] 92
Figure 44 - BLAST+[58] search results of homologous proteins to CeuE. Aligned using T-Coffee[57] and image produced using ESPript.[45] Red block colours show fully conserved residues; red triangles show the position of basic triad of arginine residues; blue starts indicate position of coordinating histidine and tyrosine residues. 96
Figure 45 - Schematic diagram of a potential mechanism for both hexadentate and tetradentate binding in the binding pocket of CeuE. 97
Figure 46 - Metal centre chirality in *cis*-bis(bidenate) complexes in comparison with the {CeuE[Fe(4-LICAM)]} crystal structure. Key; CeuE residues, green - carbon, blue - nitrogen, red - oxygen; 4-LICAM, grey - carbon, red - oxygen; Fe - Coral sphere. 98
Figure 47 - Circular dichroism spectra of CeuE with {CeuE[Fe(4-LICAM)]} in black and ferric 4-LICAM in the absence of CeuE (grey) recorded in 20 mM TrisHCl buffer, 10 mM NaCl, 0.6% DMSO, pH 8. 99
Figure 48 - Emission spectra of CeuE at 240 nM in 40 mM TrisHCl pH 7.5 150 mM NaCl, with various equivalents of Fe-4-LICAM, excitation at 280 nm. Each spectrum was recorded after 1 minute (dark), 5 minutes (medium) and 15 minutes (light), shown as shades of starting colour. Key; Black, 0 eq - blue, 0.4 eq - red, 0.8 eq - green, 2 eq - purple, 5 eq. Emission intensity normalised to initial intensity. 100
Figure 49 - Example fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titrations were carried out with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data points shown as circles; lines give the non-linear least-squares calculated fits (DynaFit).[70] 101
Figure 50 - ^{1}H NMR spectrum of **22** in $CDCl_3$ shown in the regions between 4.85-4.70 ppm and 4.00-3.75 ppm. The J-coupling splitting pattern is shown

below, explaining the splitting for each of the three resonances. *The assignment of H^b and H^c can be interchanged. 116
Figure 51 - Upper - Plot of the absorbance at 590 nm for solutions containing the DHBS monomer and linear dimer in the presence of increasing equivalents of Fe(III). Solutions were buffered at pH 9 and ferric ion was added as the NTA complex. Lower - UV-visible spectra between 370-800 nm recorded during the linear dimer titration. Reprinted with permission from *Inorg. Chem.*, 1991, *30* (5), 900-906. Copyright 1991 American Chemical Society.[8] 123
Figure 52 - Spectrophotometric titration of the 1:1 ferric complex of H_4-**17** in dilution corrected absorbance units. Reprinted with permission from *Inorg. Chem.*, 1991, *30* (5), 900-906. Copyright 1991 American Chemical Society.[8] 124
Figure 53 - Amonabactins. Names refer to the aromatic amino acid constituent and molecular weight. 126
Figure 54 - Upper - Selected UV-visible absorbance spectra recorded during the kinetic run with H_4-**17** (2.0×10^{-4} M) and iron(III) (1:1 ratio). Spectra ordered from black (0 seconds) to light blue (60 seconds) for selected time intervals. Lower - Kinetic trace of upper, following absorbance at 500 nm (red diamonds) and 560 nm (blue triangles) over 2 minutes. Spectra recorded every 3 seconds, solution containing 0.1 M TrisHCl pH 7.5, 5% DMSO at 20 °C. 128
Figure 55 - Selected UV-visible absorbance spectra of H_4-**17** and Fe^{III}(NTA), used for the Job plot analysis Spectra recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. Spectra ordered from black (H_4-**17** excess) to light blue ordered by ligand to metal ratio; 90:10, 70:30, 60:40, 50:50, 40:60, 30:70, 10:90). [M] + [L] = 0.4 mM. 129
Figure 56 - Job plot for the binding of H_4-**17** and Fe^{III}(NTA), obtained by following the absorbance at both 512 nm (red diamonds) and 563 nm (blue triangles). The absorbance values are averages of two experiments and the error bars indicate the differences between the runs. Plot recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. 130
Figure 57 - Circular dichroism spectra of the free ligands (left) and the iron(III) complexes (right) of enterobactin and the hydrolysis products in aqueous solution at pH 7.5, 0.1 M HEPES buffer. Reprinted with permission from *Inorg. Chem.*, 1991, *30* (5), 900-906. Copyright 1991 American Chemical Society.[8] 131

Figure 58 - Circular dichroism spectra of H_4-**17** and ferric-**17** in key ligand-to-metal ratios in aqueous 0.1 M TrisHCl buffer, pH 7.5, 150 mM NaCl, 5% DMSO. [L] + [M] = 0.4 mM. 132
Figure 59 - Ribbon representation of three monomers of Scn co-crystallised with ferric-enterobactin. Key: Chain A (light blue), chain B (gold), chain C (green); ligands shown as cylinders, grey - carbon, blue - nitrogen, red - oxygen, coral - iron. PDB code: 1L6M.[39] 135
Figure 60 - Ribbon representation of three monomers of Scn co-crystallised with ferric-enterobactin. Key: chain A (blue), chain B (gold), chain C (green); ligands shown as cylinders, grey - carbon, blue - nitrogen, red - oxygen, coral - iron. PDB code: 3I0A (unpublished work). 137
Figure 61 - Ribbon representation of the three independent CeuE chains of {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) with colours blended from cyan, N-terminus, to dark blue, C-terminus. His228, Tyr288 and Fe-**17** shown as cylinders. Key: His227 and Tyr288; green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral. 140
Figure 62 - Cylinder representation of the binding pockets of {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) with electron density map shown as $2F_{obs}$-F_{calc}, contoured at 1σ and chemical representation showing key interactions of Fe-**17**. Key: His227 and Tyr288; green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral. 141
Figure 63 - Ribbon representation of the three independent CeuE chains of {$CeuE_3$[Fe(**17**)][Fe (DHBS)$_2$]$_2$} (form III) with colours blended from cyan, N-terminus, to dark blue, C-terminus. His227, Tyr288, Fe-**17** and Fe-$DHBS_2$ shown as cylinders. Key: His228, Tyr288, green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral. 142
Figure 64 - Cylinder representation of the binding pockets of {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} (form III) with electron density map shown as $2F_{obs}$-F_{calc}, contoured at 1σ and chemical representation showing key interactions of the two Fe-DHBS environments. Key: His227 and Tyr288; green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral. 143
Figure 65 - Crystal packing diagram showing symmetry related CeuE molecules in the {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) crystal. Key {$CeuE_3$[Fe(**17**)]$_2$[Fe]} shown

in gold, symmetry related molecules shown in grey. Image created using Coot.[47] 146
Figure 66 - A) Ribbon representation of the crystal packing in the asymmetric unit of {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) with His227, Tyr288, Fe-**17** shown as cylinders. B) Ribbon representation of binding pocket of chain C and surface (electrostatic representation) of N terminus of chain B, with Fe-17 from chain B superimposed into the binding pocket. Key chain A (blue), chain B (gold), chain (dark green); Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.; electrostatic, red - negative, blue - positive. 147
Figure 67 - Coordination environment of the three binding pockets of $CeuE_3$[Fe(**17**)][Fe (DHBS)$_2$]$_2$} (form III). Key; CeuE residues, green - carbon, blue - nitrogen, red - oxygen; DHBS and 4, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral. 151
Figure 68 - Surface electrostatic representation of {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$}, monomers B (translucent surface) and C (solid surface), negative charge shown in red and positive charge shown in blue. Key: DHBS, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.; Hydrogen bond - black dash. 153
Figure 69 - Circular dichroism spectra of ferric-**17** (light grey) with increasing increments of 0.2 eq of CeuE, until {CeuE[Fe(**17**)]} (black) was formed, recorded in 0.1 M TrisHCl buffer, pH 7.5, 150 mM NaCl. 154
Figure 70 - Example fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit).[55] 156
Figure 71 - Schematic design of a biomimetic hexadentate siderophore that possesses a chemical linker. 160
Figure 72 - Dendritic catecholate-based iron chelator developed by Zhou *et al*.[11] 162
Figure 73 - Structures of a biomimetic tris-catecholate siderophore and its aminopenicillin conjugates.[20] 163
Figure 74 - The chemical structure of the designed biomimetic hexadentate siderophore. Key: catecholamide, red - backbone, blue - linker, black. 165

Figure 75 - ^{1}H NMR spectrum of **43** in d_6-DMSO shown between 3.00-2.10 ppm showing the multiplicity of the resonances assigned to protons H^a and H^b. Quintet at 2.50 is the d_6-DMSO residual solvent peak. 169
Figure 76 - Schematic diagram showing the FRAP experiment. 182
Figure 77 - Schematic representation of the TIRFM experimental set-up. A) General overview; B) Schematic representation of the evanescent wave formation. 183
Figure 78 - HPLC traces (absorbance at 420 nm) of ferricrocin (20 mM) in water, injection volume 5 µL (blue) and **53** (0.5 mg in 350 µL of water) injection volume 20 µL (red); 6-40 % (v/v) acetonitrile/water gradient. 186
Figure 79 - HPLC traces of **51** (red) after dialysis and Alexa Fluor® 488 (blue). A) - Absorbance data measured at 420 nm. B) Fluorescence data excited at 500 nm emission measured at 520 nm. 188
Figure 80 - LC-MS of **51**. A) TIC chromatogram; B) UV absorbance (254 nm) chromatogram; C) Mass spectrum associated with the peak eluting after 32 minutes in the chromatogram. Peak at 15 min (dashed black line) is calibrant injected into the ESI source used for accurate mass calibration. 189
Figure 81 - A) UV-visible absorbance spectra of Alexa Fluor® 488 azide ($1.3x10^{-5}$ $moldm^{-3}$), **51** and ferricrocin ($2.0x10^{-4}$ $moldm^{-3}$); B) Emission spectra of Alexa Fluor® 488 azide and **51**, excited at 488 nm; C) Excitation spectra of Alexa Fluor® 488 azide and **51**, monitoring at 525 nm. All spectra recorded in water. 191
Figure 82 - Initial TIFM images. A) 10 nM **51** and *E. coli ΔfhuA* cells (sum of 500 video frames); B) 10 nM **51** and *E. coli ΔtonB* cells (sum of 5 video frames). 193
Figure 83 - Repeat TIRFM images containing 10 nM **51**, *E. coli ΔfhuA, ΔtonB* cells and *ΔtonB* cells with 5 µM ferricrocin. A) sum of 10 video frames; B) sum of 200 video frames; C) Profile plot of grey value of pixels across highlighted cell (yellow). Analysis performed using ImageJ.[29] 194
Figure 84 - Run 1 fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 25.1± 1.3 nM (240 nM), 44.9 ± 2.5 nM (170 nM). 251

Figure 85 - Run 2 fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 30.3± 0.9 nM (240 nM), 40.8 ± 3.8 nM (170 nM). 251
Figure 86 - Run 3 fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 14.9± 2.7 nM (240 nM), 19.9 ± 5.3 nM (170 nM). 252
Figure 87 - Run 1 fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 4.8± 0.9 nM (240 nM), 9.8 ± 1.9 nM (170 nM). 252
Figure 88 - Run 2 fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 16.4 ± 1.8 nM (240 nM), 12.3 ± 2.5 nM (170 nM). 253
Figure 89 - Run 3 fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 14.9± 3.7 nM (240 nM), 14.9 ± 3.6 nM (170 nM). 253
Figure 90 - Script used for Dynafit v3. 258
Figure 91 - Script used for Dynafit v4. 259

Scheme 1 – Synthesis of H_4-4-LICAM. (a) i) Benzyl chloride, K_2CO_3, EtOH ii) $NaClO_2$, H_3NSO_3, acetone/water, 79%; (b) *N*-hydroxysuccimide, DCC, 1,4-dioxane, 71%; (c) 1,4-diaminobutane (0.5 eq), NEt_3, acetone, 94%; (d) H_2, Pd-C 10%, EtOH, 99%. Overall yield 52%. 60
Scheme 2 – Synthesis of H_2-**15**. (a) 2.0 M ethylamine, NEt_3, THF, 72%; (b) H_2, Pd-C 10%, EtOH, 74%. Overall yield 53%. 64

Scheme 3 - Biosynthetic pathway of enterobactin.[2] 108
Scheme 4 - The *N*-benzoyl monomer route when using X = DCC, DCC/HOBt or DCC and pyridine to activate the carboxylic acid for esterification. 110
Scheme 5 - The use of HOBt to reduce the deactivation of the *O*-acylisourea mixed anhydride **20**. 111
Scheme 6 - Oxazolone-mediated racemisation during esterification. 112
Scheme 7 - The urethane protection route for the total synthesis of enterobactin outlined by Rastetter *et al.* 113
Scheme 8 - Initial synthetic route towards the H_4-**17**. (a) NEt_3; (b) NEt_3; (c) THP, pyridinium *p*-toluenesulfonate; (d) NaOH / NMe_4OH; (e) DCC, HOBt, H^+ workup; (f) H_2, Pd-C 10%. 115
Scheme 9 - Second synthetic route towards H_4-**17**. (a) NEt_3; (b) NEt_3; (c) Esterification, various conditions; (d) H_2, Pd-C 10%. 118
Scheme 10 - Third synthetic route towards H_4-**17**. (a) $SOCl_2$; (b) Boc_2O, NEt_3; (c) Boc-*O*-benzyl-L-serine, EDC, HOBt, ; (d) TFA; (e) DIPEA; (f) H_2, Pd-C 10%. 120
Scheme 11 - Synthetic route to H_6-**41**. (a) i) $SOCl_2$, ii) TMS azide, iii) 80 °C, iv) HCl, (b) Boc_2O, NEt_3; (c) H_2, Raney®-Nickel 2800 slurry; (d) Glutaric acid monomethyl ester chloride, NEt_3 ; (e) TFA; (f) **31**, DIPEA; (g) NaOH; (h) H_2, Pd-C 10%. 167
Scheme 12 - Synthetic route to compound **43**. (a) $SOCl_2$; (b) TMS azide; (c) 80 °C; (d) HCl. 168
Scheme 13 - Reaction mechanism of reaction of **42** to form **43** with key intermediates, acyl chloride, acyl azide and isocyanate highlighted within a box. 170
Scheme 14 - Synthesis of **44** using the methodology of Zhao *et al.*[23] 171
Scheme 15 - Synthesis of **45** using commercially available Raney®-Nickel. . 172
Scheme 16 - Synthesis of **46** using the modified methodology of Ji *et al.*[20] 172
Scheme 17 - Synthesis of **47** by Boc deprotection of 46 using trifluoroacetic acid. 173
Scheme 18 - Synthesis of **48** using the methodology outlined in Chapter 3. ... 174
Scheme 19 - Synthesis of **49** via methyl ester hydrolysis using the methodology of Theodorou *et al.*[29] 175

Scheme 20 - Deprotection of **49** via catalytic hydrogenation to yield siderophore mimic H_6-**41.** 176
Scheme 21 - Synthesis of high-molecular weight ferricrocinyl polyethylene glycolyl succinate by Coulton *et al.*[14] 180
Scheme 22 - Synthetic route to **53**. (a) DCC, N_2; (b) Ferricrocin, N_2. 185
Scheme 23 - Synthetic route to **51**. (a) Alexa Fluor® 488 azide, $CuSO_4$, sodium ascorbate. 187

Table 1 - Formation constants reported for the ferric complexes of the indicated tris-catecholamide ligands. * $\log_{10}K_f$ refers to fully deprotonated ligand forming a 1:1 complex with iron(III). ǂ The fully deprotonated charge of the linear trimer includes the six catechol protons and the carboxylic acid on the backbone. 34
Table 2 - Selected examples of Fur regulated genes from *E. coli* with a brief description of their function. 43
Table 3 - Comparison of ^{1}H NMR shifts for the aromatic triplet recorded in 5:1 d_6-DMSO : D_2O. * Charge per catechol calculated assuming complex A has a 3:2 ligand to metal ratio and complex B has a ligand to metal ratio of 1:1, with neutral donors in vacant coordination sites. ǂ H_4-4-LICAM was recorded in d_6-DMSO. 73
Table 4 - Primers used for gene amplification. 78
Table 5- Data collection and refinement statistics of apo-CeuE and {CeuE[Fe(4-LICAM)]} structures. 82
Table 6 - Bond distances between coordinating atoms and the iron centre. 89
Table 7 - A comprehensive list of bond angles between coordinating atoms to the iron centre and their deviation from the ideal octahedral geometry. 90
Table 8 - Calculated dissociation constants from non-linear regression analysis of the fluorescence quenching data of ferric-4-LICAM : CeuE titrations. Error of individual K_d values are calculated from non-linear regression fit, and errors of overall and batch calculated from standard deviation of the different K_d values within that group. 102
Table 9 - Summary of modelled ligands in the binding pockets of both crystal forms, {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (crystal form II) and {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} (crystal form III). 139

Table 10- Data collection and refinement statistics of {$CeuE_3$[Fe(**17**)]$_2$[Fe]}.and {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} structures. 144
Table 11 - Bond distances between coordinating atoms and the iron centres of monomers A and B in the {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) structure. 148
Table 12 - Bond distances between coordinating atoms and the iron centres for the {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} (form III) structure. 149
Table 13 - Calculated dissociation constants from non-linear regression analysis of the fluorescence quenching data of ferric-**17**: CeuE titrations. Error of individual K_d values are calculated from non-linear regression fit, and errors of overall and batch calculated from standard deviation of the different K_d values within that group. 156
Table 14 - Ordering of synthetic experimental details. 203
Table 15 - Raw and normalised fluorescence data for the ferric-4-LICAM CeuE titration at 240 nM. 254
Table 16 - Raw and normalised fluorescence data for the ferric-4-LICAM CeuE titration at 170 nM. *Note missing data point of Fe-4-LICAM concentration at 0.4284. 255
Table 17 - Raw and normalised fluorescence data for the ferric-**17** CeuE titration at 240 nM. 256
Table 18 - Raw and normalised fluorescence data for the ferric-**17** CeuE titration at 170 nM. 257

Equation 1 - Stokes-Einstein equation. Key; η - viscosity, r - radius of a spherical particle, k - Boltzmann constant, T - temperature. 75
Equation 2 - Calculation of weighted average dissociation constant. 102

Accompanying Material

A compact disk with an electronic copy of this thesis, the fluorescence quenching data, the DynaFit scripts, and the PDB, MTZ and CIF files associated with the X-ray diffraction data is attached at the back of this thesis.

Acknowledgements

I would like to thank a number of people, without whom this thesis would not have been possible. Firstly, I would like to thank both of my supervisors, Dr Anne K. Duhme-Klair and Prof Keith S. Wilson, for their continued support in every aspect of my PhD, and so many varied and interesting suggestions for my project. Not only have you both been great supervisors, but both of you have also proved to be great role models to aspire to.

I would also like to extend my recognition to everyone who has been involved in this project, including collaborators, Prof James Coulton and Dr Christoph Baumann, for their specific help and also for providing me with the experience of working within a collaborative network. In particular, I would like to thank Dr Olga Moroz, for her help, patience and expertise in protein expression, purification and crystallography. Importantly, I must also especially thank all the students that I have worked with during my PhD at York: Adam, Andy, Amy, Bethany, Chris, Emily, Jordan, Laura B, Laura W and Will, for their hard work, company and for allowing me to develop my ability as a mentor.

A large thank you to the technical and administration staff in the Department of Chemistry, University of York. So much research activity could not go ahead without their help. I would like to give a special thanks to Heather and Pedro for their NMR assistance, Karl and Helen for the MS service provided, Graeme for CHN analysis, Johan and Sam for all there help with the X-ray crystallography and also the wet lab technicians, in both chemistry and the YSBL.

I would also like to thank all past and present members of the groups I have worked with; Abeda, Aurélien, Chris, Ellis, Jenni, Maria, Stephen and Tom, as well as a number of great friends I have made during my PhD; the E014 lab, the SLUGS, and the pizza night groups; who have been amazing company. A particular thanks must go to Luisa, for your endless supply of Italian grade

coffee, as well as the great many discussions on all different aspects of chemistry throughout the PhD process.

I must thank some of the most important people to me; my family, for all of their unwavering support, not only during my PhD but my entire life. Without the belief that my parents have in me, I would definitely not be in the position I am today. Finally I need to thank my fiancée Kim. You have been amazing during my PhD; without your love and support during this section of our life journey together, I would not have managed it.

Declaration

The research presented within this thesis is, to the best of my knowledge, original and my own. The following work was carried out by or with the assistance of other researchers:

- Synthesis and Job plot analysis of H_2-**15** was carried out by BSc student Miss Bethany Peeters;
- DOSY NMR experiments was conducted by Ms Heather Fish;
- CeuE protein expression and purification was carried out with the assistance of Dr Olga Moroz;
- Initial attempts towards the synthesis of H_4-**17** was carried out by summer placement student Miss Emily Brooke;
- The second batch of CeuE used for the dissociation constant determination (170 nM) in Chapters 2 and 3 was purified by BSc student Mr Adam Hughes with titrations completed with the assistance of Adam.
- Several synthetic procedures for the compounds **43-48** were carried out by summer placement student Miss Amy Scorah;
- The TIRF microscopy analysis outlined in Chapter 5 was performed by Dr Christoph Baumann (Department of Biology, University of York) as part of a collaboration with Prof. James Coulton (McGill University, Montreal);
- All mass spectrometry and LC-MS experiments were carried out by Mr Karl Heaton or Miss Helen Robinson.
- Elemental analyses were carried out my Dr Graeme McAllister.
- Part of the ferric-4-LICAM work outlined in Chapter 2 has been published previously in the following reference: D. J. Raines, O. V. Moroz, K. S. Wilson, A.-K. Duhme-Klair, *Angew. Chem. Int. Ed.*, 2013, **52**, 4595-4598.

Daniel Raines

Chapter 1: Introduction

1.1 Overview

This thesis is concerned with the synthesis, characterisation and investigation of siderophores and analogues, with a view to providing further insight into the mechanism by which iron uptake occurs. This chapter provides an overview of the background to the project and puts the work within this thesis into context with regard to previous research within the field.

Throughout the thesis, the indicated protonation states of compounds only refer to those protons that dissociate upon iron chelation, protons from other functional groups e.g. backbone carboxylic acids, are not included.

1.2 Iron in Biology

Iron is an essential element for nearly all living organisms on Earth.[1] The redox chemistry and high affinity for oxygen make this abundant element ideal for a number of biological processes. In aqueous media, the most common oxidation states are iron(II) and iron(III). The redox potential between these states is such that the metal can perform a wide variety of biologically relevant reactions.[1-2]

While iron is vital for most forms of life, it is also toxic in excess, due to its ability to perform a redox cycle with molecular oxygen, resulting in highly reactive oxygen-derived free radicals.[3-5] Due to this toxicity, biological systems have developed sophisticated solutions for iron uptake, storage, transport and distribution.[3]

Despite iron being one of the most abundant elements in the Earth's crust, it is predominately found as iron(III), due to the oxidising atmosphere.[6] This has the consequence of the free iron concentration in the environment varying from 10^{-18} to 10^{-24} mol dm^{-3}.[7-8] As a result micro-organisms in particular have developed specialised strategies for the uptake of this crucial element from the surrounding

environment.[9] These can be categorised into two different iron uptake mechanisms, direct or indirect. Direct uptake of iron is the ability of an organism to acquire iron from a natural source, e.g. haem when located within the human body. This specialised direct iron uptake mechanism requires a specific receptor for each different iron source which is a disadvantage when the organism is located in different environments.[1] This is thought to be the main reason why indirect iron uptake mechanisms are more widespread and diverse, with the use of siderophores being one of the most common examples.[1-2, 7]

1.3 Siderophores

Siderophores are small strong chelating molecules that are used to bind and solubilise iron(III) to aid iron uptake. They are required to have a selectivity for iron(III) over other abundant metal ions found within the environment, achieved by matching the preferred coordination characteristics of the metal centre with the coordination ability of the siderophore. Iron(III) is a 'hard' ion with a d^5 electronic configuration and thus prefers to coordinate in a octahedral geometry, surrounded by charge dense, negatively charged donor atoms.[10-11] To complement this, the donor atoms of the siderophore chelating groups are dominated by the presence of negatively charged oxygen atoms, usually arranged to form a hexadentate ligand that forms a highly thermodynamically stable complex.[11] While there are not many biologically important 'hard' ions that can compete with iron(III), siderophores must also be selective for iron(III) in the presence of environmentally abundant trivalent cations such as aluminium(III). The binding moieties of siderophores are generally comprised of bidentate units, which are predisposed to have an ideal bite angle for the coordination of iron(III) (ionic radius of 0.65 Å) over other abundant metals (e.g. ionic radius of aluminium(III) is 0.54 Å).[2, 12]

There are over 500 known siderophores, with a wide variety of different structural motifs for the backbone.[2] Despite this, there is a reliance on relatively few coordination groups, presumably due to the limitation of the coordination chemistry of iron(III).[2] Siderophores are classified by these coordinating

groups, with the three most common iron-chelating units being catecholates, hydroxamates and α-hydroxycarboxylates (Figure 1).

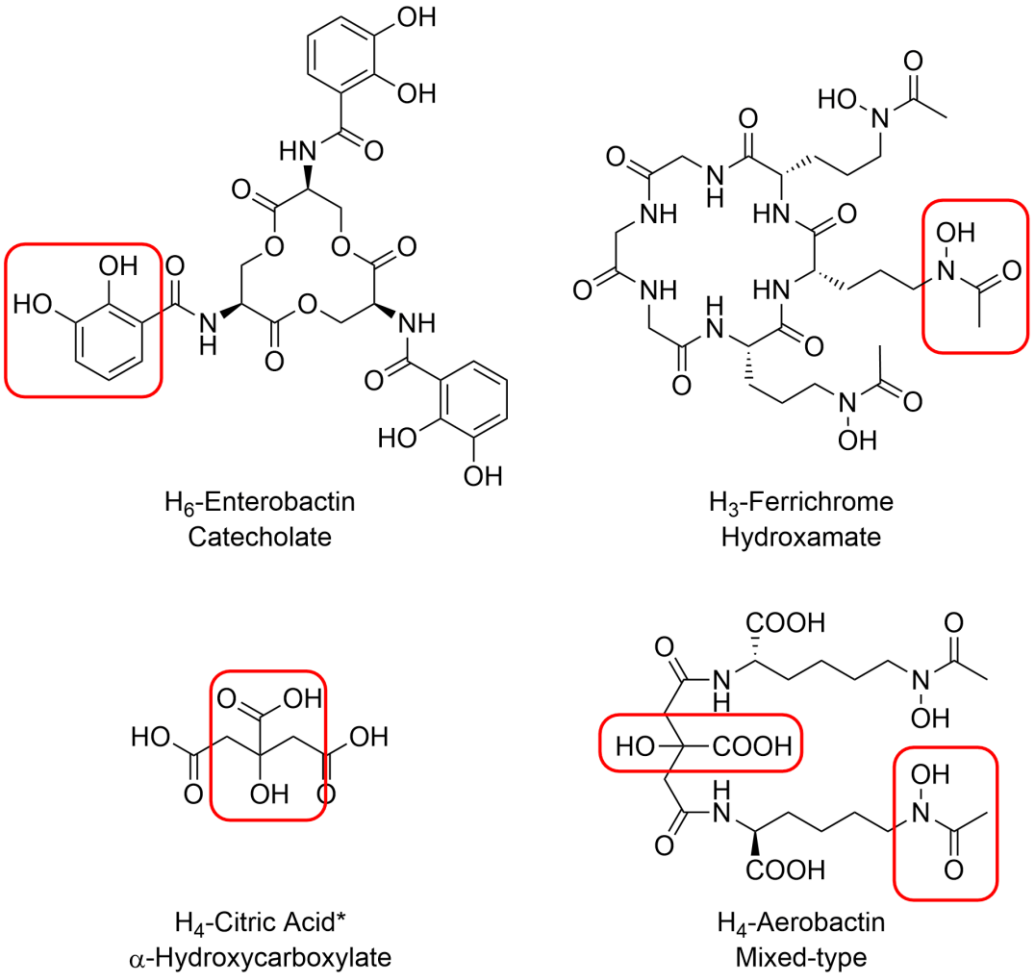


Figure 1 - Representative examples of siderophores with different iron chelating moieties highlighted. *Protonation state of citric acid can differ due to different binding modes available.

Each of these three coordinating groups possess the charge dense oxygen atoms required for strong iron(III) binding. Under neutral conditions, the catechol group deprotonates twice (pK_a values 9.2, 13.0), to form two phenolate atoms which coordinate to the metal centre, whereas the hydroxamate (pK_a 9.0) and α-hydroxycarboxylate (pK_a 3.0, 14.5) groups are only required to deprotonate once, with the latter's carboxylates already deprotonated under neutral conditions.[2] This results in the catecholate complexes being more sensitive to pH conditions.

As a result of this, the catechol group is often observed in siderophores as the related catecholamide group. When the amide is in the adjacent position to provide the 2,3-catecholamide functionality, the resulting hydrogen bonding network reduces the pK_a of the *ortho*-phenolate to 7.2-7.4.[13] The hydrogen bonding network around the catecholamide groups also affects the rotation of the catecholamide with respect to the backbone of the siderophore. For example, in the fully protonated form, the hydrogen bonding network of the *ortho*-phenolate involves the carbonyl of the amide group, whereas upon deprotonation, the network reverses and the proton of the amide group forms a hydrogen bond to the *ortho*-phenolate.[14] This reversal has the consequence of rotating the catecholamide group, which is also observed upon iron(III) chelation. It is postulated that this mechanism is utilised by the tris-catecholate siderophore enterobactin for orientating the catecholamide groups in an 'outwardly' facing direction when fully protonated, and upon iron(III) binding the catecholamide group rotates, pulling the metal centre into the central binding position (Figure 2).[7] For many siderophores the amide functionality has been shown to be important for recognition by the binding proteins during iron uptake.[15]

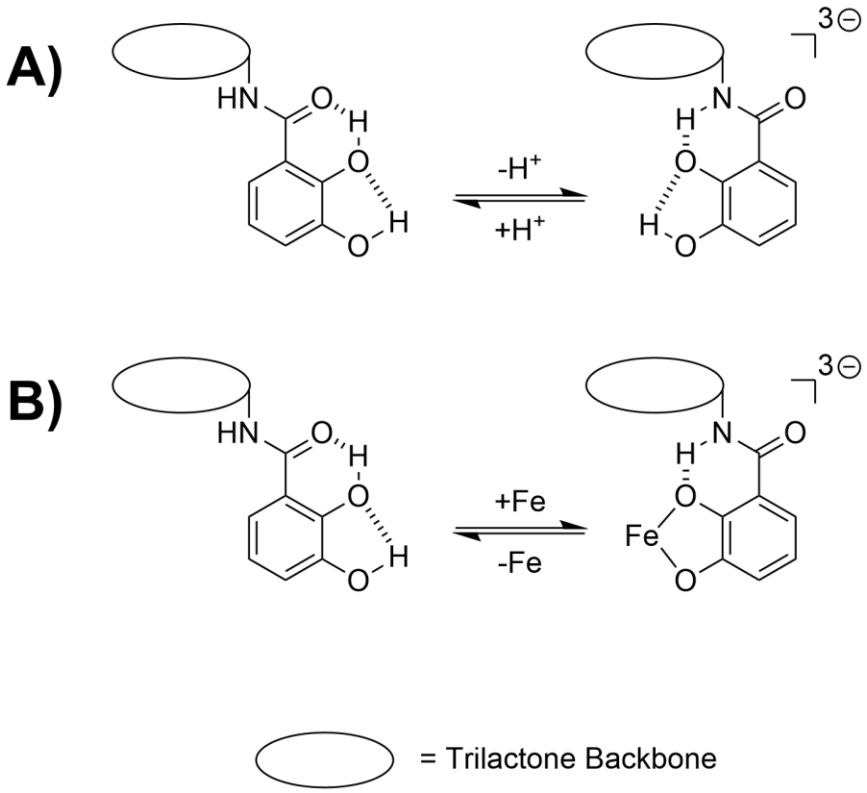


Figure 2 - Schematic diagram demonstrating the hydrogen bonding network and conformational change of enterobactin upon A) single deprotonation; B) iron complexation. Figure adapted from *Proc. Natl. Acad. Sci.*, 2003, *100*, (7), 3584-3588. Copyright 2003 National Academy of Sciences, USA.

Enterobactin is one of the most widely studied siderophores. It contains three catecholamide moieties connected *via* a cyclic tri-L-serine backbone (Figure 1). Enterobactin comprises of three dihydroxybenzoylserine (DHBS) units, which are arranged to form a cyclic ester. As such, enterobactin is prone to hydrolysis under certain conditions and can form three distinct hydrolysis products; the linear trimer, linear dimer and DHBS monomer. Each hydrolysis product has been shown to promote bacterial growth, demonstrating that siderophores are not required to be hexadentate to function as iron scavengers (Figure 3).[16-17]

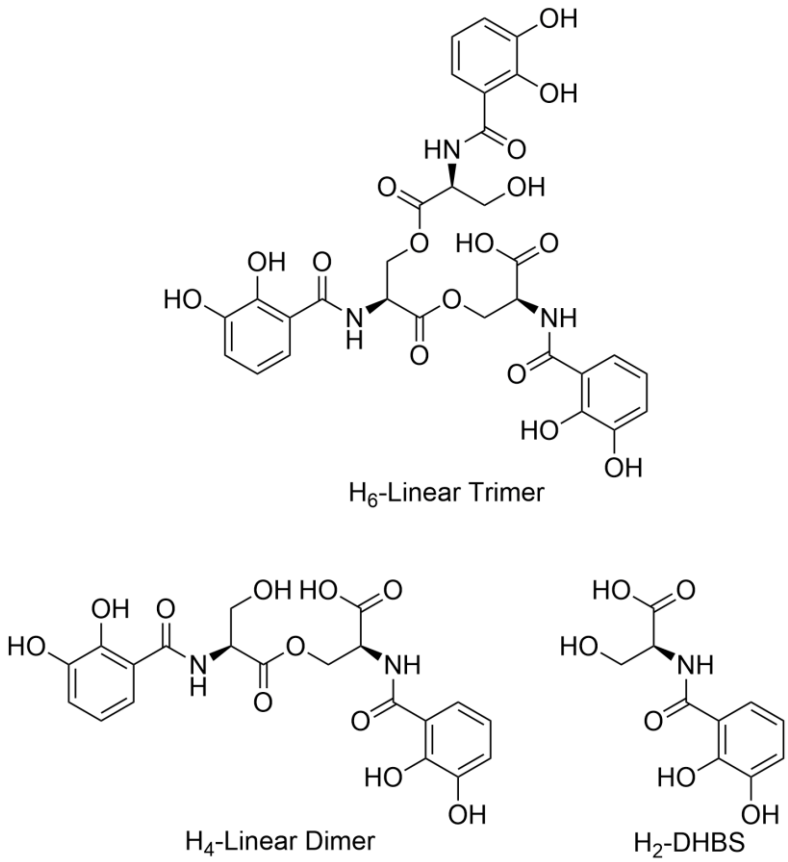


Figure 3 - Chemical structures of the enterobactin hydrolysis products.

There are a number of examples of siderophores with 'low' denticity (Figure 4). While some of these could be perceived as biological precursors to their hexadentate counterparts, e.g. azotochelin and the enterobactin linear dimer are formally fragments of protochelin and enterobactin, respectively, others are only structurally related and are a siderophore in their own right, e.g. rhodotorulic acid and pyochelin.[18] It is still not clear why some organisms purposely synthesise siderophores with a lower denticity. One possibility is their involvement in the uptake of other biologically important ions, and thus the selectivity requirement of the siderophore is less strict.[18-20]

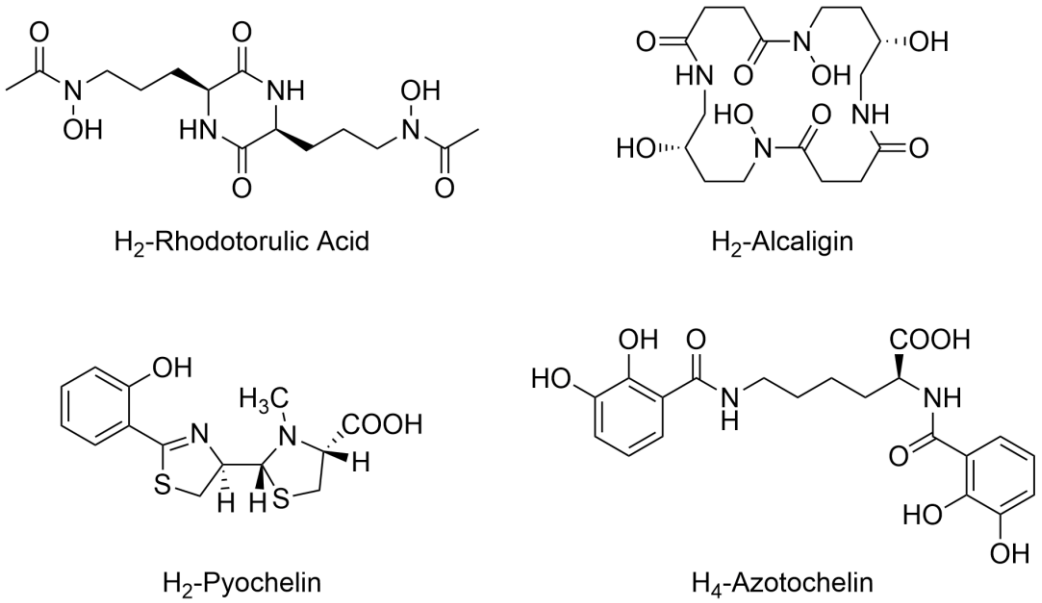


Figure 4 - Chemical structures of a selection of 'low' dentate siderophores.

It has been shown that the stereochemistry of siderophores is greatly important for iron uptake.[21-24] As siderophores are usually comprised of three bidentate units coordinating to the metal centre in a octahedral geometry, they have a metal-centred chirality, in addition to any ligand-based chirality. This allows the possibilty of two alternate configurations; a left-handed propeller (Λ-configuration), or a right-handed propeller (Δ-configuration, Figure 5).[25]

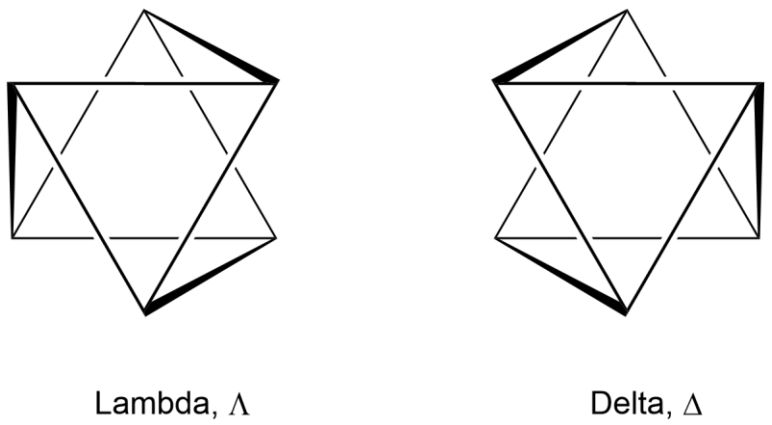


Figure 5 - Λ-configuration and Δ-configuration of a tris-bidentate octahedral complex.

Asymmetric binding moieties also give rise to *cis*- and *trans*-complexes, depending on the orientation of the binding unit with respect to the other

coordinating groups. If there is no chiral or steric induction from the siderophore backbone, these optical isomers form racemic mixtures in solution. As the majority of siderophores are derived from amino acids, and hence are often optically active, racemic siderophore complexes are rare. The most common isomers formed with siderophores are Λ-*cis* complexes, although enterobactin has been shown to have a preference to form a Δ-*cis* complex.[2, 21]

Siderophore Mimics

Due to their remarkable iron(III) chelating chemistry, there has been much focus on the fundamental study of the iron binding ability and physical properties of siderophores. In order to study different aspects of siderophores and their uptake pathways, many siderophore mimics have been synthesised and used in various investigations.

While enterobactin is one of the most widely studied siderophores, due to its hydrolytically unstable tri-lactone backbone, a number of hydrolytically stable mimetic systems have been developed. The most successful were designed to keep the key structural features of enterobactin, namely the three bidentate catecholamide units, positioned on a backbone with three-fold rotational symmetry, and five-atomic unit spacers between the binding moieties (Figure 6).[26-29]

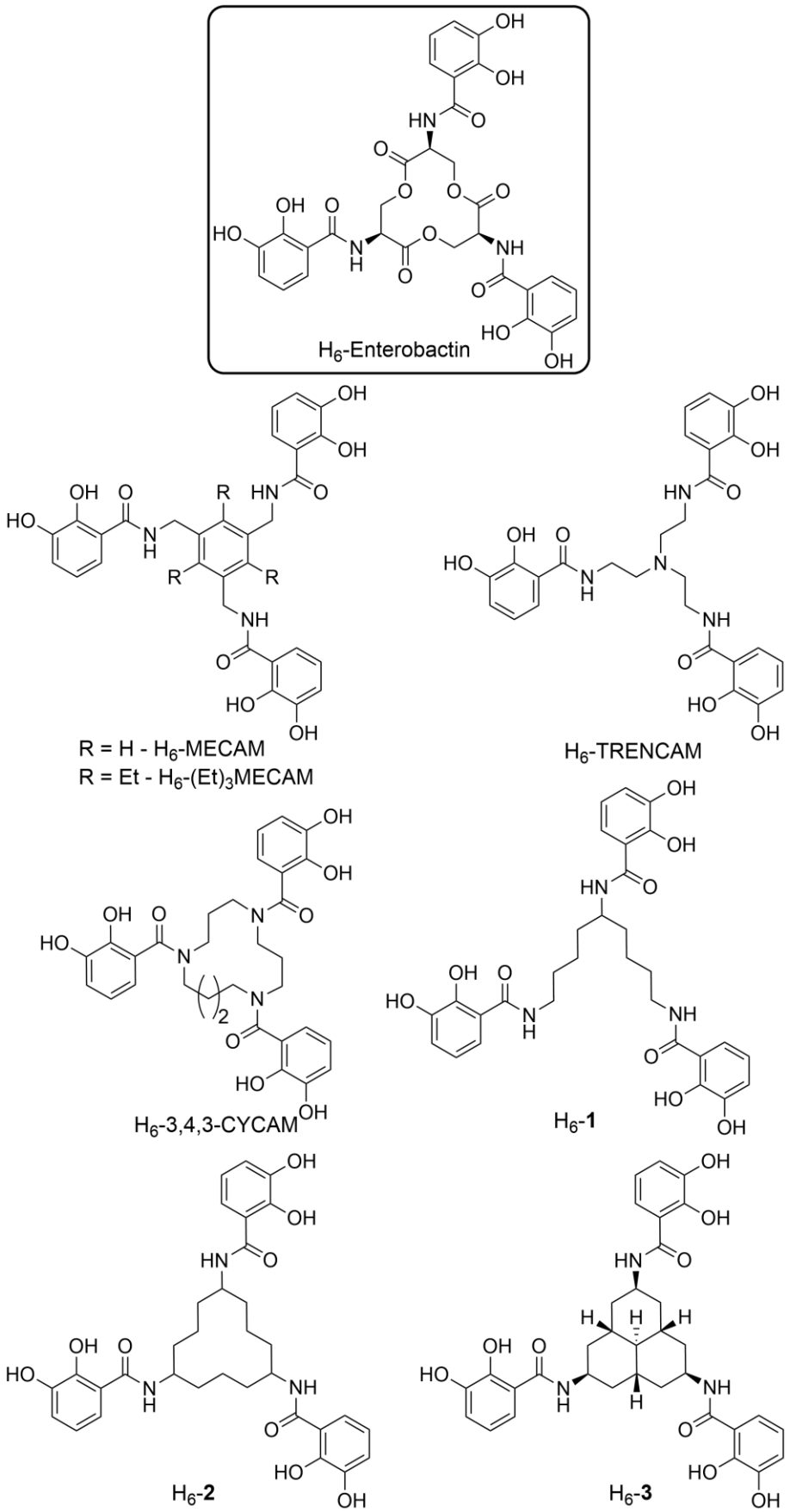


Figure 6 - A selection of structures of biomimetic hexadentate catecholamide siderophores and enterobactin, all possessing high affinity for iron(III).[26-29]

As with natural siderophores, the preorganisation of the catecholamide functional groups correlates with the strength of the iron(III) binding. Enterobactin fulfils this requirement for preorganisation *via* the chirality of the tri-serine backbone, as well as the hydrogen-bonding network formed by the catecholamide functionality.[7]

Catecholamide preorganisation for iron binding has also been demonstrated with siderophore mimics. Stack *et al.* have shown that $(Et)_3MECAM^{6-}$ binds iron(III) with a formation constant of $K_f \approx 10^{47}$ M^{-1}, an increase of 10^4 compared with $MECAM^{6-}$ ($K_f \approx 10^{43}$ M^{-1}).[29] This increase was attributed to the predisposition of the catecholamide rings to be preferentially on one side of the central aromatic ring. This is due the adjacent groups avoiding steric interactions by alternating above (*a*) and below (*b*) the central ring in a (*ababab*) conformation with three-fold symmetry. In $(Et)_3MECAM^{6-}$, the neighbouring groups of the methylene-catecholamide and the ethyl groups are both large enough to form the *ababab* conformation, whereas in MECAM protons occupy the neighbouring positions, hence only minimal steric interactions occur, leading to a weaker preference for the *ababab* conformation, (Figure 7).[29]

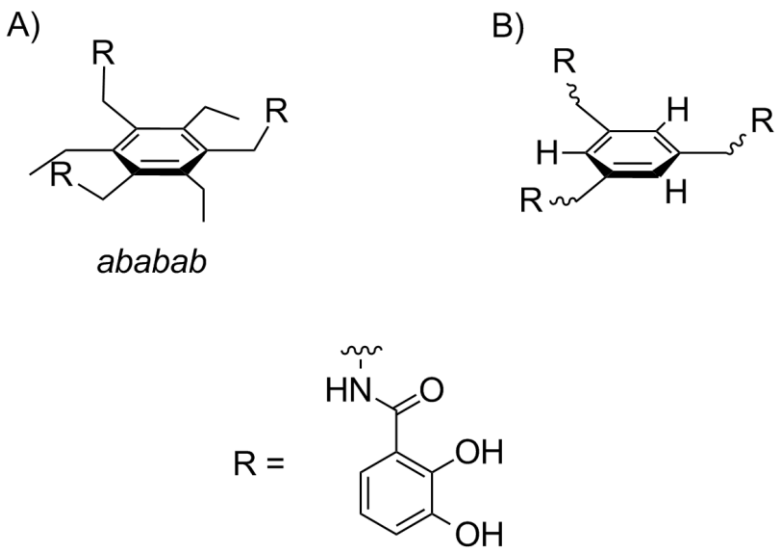


Figure 7 – Illustration of the preorganised *ababab* configuration of $(Et)_3MECAM$ (A) compared to MECAM (B).

The flexibility of the siderophore backbone has also been investigated. Tse *et al.* synthesised compounds H_6-**1**-**3** (Figure 6) and showed that conformational flexibility of the siderophore mimics had a direct impact on the strength of iron(III) binding.[27] Compound H_6-**1** has a linear backbone, allowing rotational freedom, and hence can act as a control compound to highlight the effects of a lack of preorganisation in a hexadentate chelator. Compound H_6-**2** has a cyclic backbone, restricting its rotational freedom, but, its catechol rings are not forced into a particular direction. However, in the most rigid compound, H_6-**3**, the catechol rings are all preorganised to be positioned on one side of the macrocycle. Compound $\mathbf{3}^{6-}$ has a formation constant upon binding to iron(III) of $K_f \approx 10^{49}$ M^{-1}, rivalling that of enterobactin^{6-}, whereas that of the most flexible mimic $\mathbf{1}^{6-}$ was $K_f \approx 10^{45}$ M^{-1}, closer to that of the linear trimer (Table 1).[27]

Compound	**≈log$_{10}K_f$***	**Reference**
Enterobactin^{6-}	49	13
Linear Trimer^{7-} ǂ	43	30
MECAM^{6-}	43	29
Et-MECAM^{6-}	47	29
TRENCAM^{6-}	44	26
3,4,3-CYCAM^{6-}	40	31
$\mathbf{1}^{6-}$	45	27
$\mathbf{2}^{6-}$	47	27
$\mathbf{3}^{6-}$	49	27

Table 1 - Formation constants reported for the ferric complexes of the indicated tris-catecholamide ligands. * $\log_{10}K_f$ refers to fully deprotonated ligand forming a 1:1 complex with iron(III). ǂ The fully deprotonated charge of the linear trimer includes the six catechol protons and the carboxylic acid on the backbone.

As well as being used in investigations of iron(III) chelation, siderophore mimics have also been used to investigate the iron uptake system of various micro-organisms, and as biomimetic systems for the investigation of substrate-protein interactions.[22, 24, 32-37]

Rastetter *et al.* reported a total synthesis of enterobactin, and its enantiomer, enantioenterobactin, which has a D-serine-based backbone. The D-enantiomer has been used in iron(III) uptake studies and was shown not to support bacterial growth.[24] While initially the lack of growth was attributed to stereospecificity of the outer membrane receptor towards the metal-centred chirality,[21] it was later shown that the stereospecificity must be further along the iron uptake or release mechanism.[22]

1.4 Siderophore Mediated Iron Uptake

Gram-negative bacteria

As well as synthesising siderophores, micro-organisms have developed sophisticated iron-uptake systems in order to retrieve iron-siderophore complexes. These uptake systems are specific to each organism. For Gram negative bacteria, such as *E. coli*, there are two cell membranes that the iron-siderophore complex must pass through to enter the cytoplasm of the cell. The outer membrane comprises of a lipid bilayer, which contains outer membrane receptors that internalise the iron-siderophore complex into the volume between the two membranes layers, the periplasm. Within the periplasm, there are periplasmic binding proteins (PBPs) that bind to the iron-siderophore complex, allowing it to be transported to an inner membrane ABC transport complex, which transfers it into the cytoplasm (Figure 8).[9]

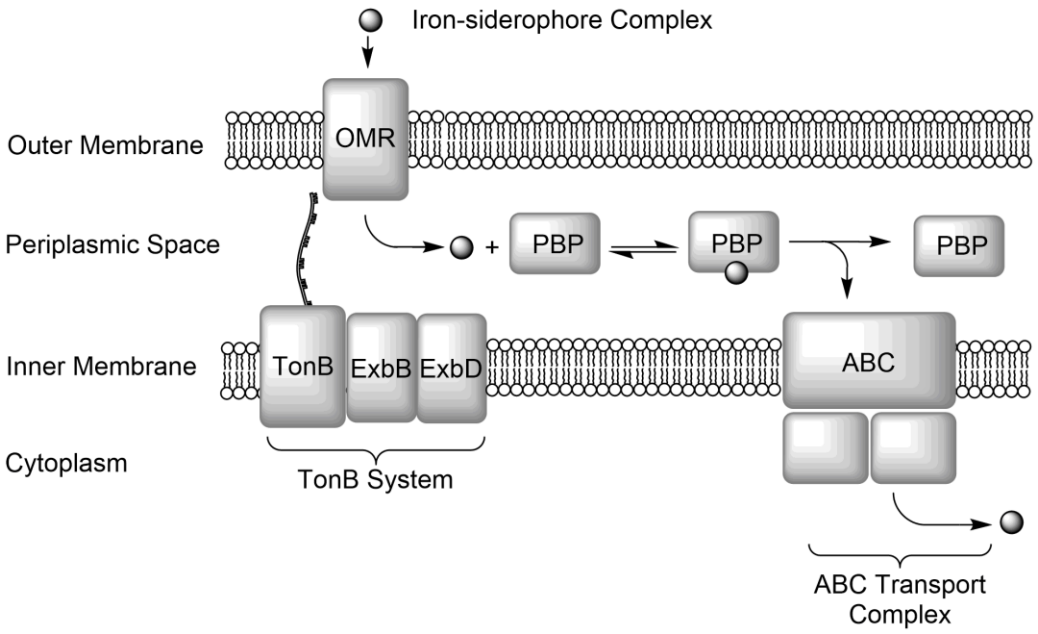


Figure 8 - Schematic diagram of the iron-uptake mechanism for a Gram-negative bacterium.

Outer membrane receptors for iron-siderophore complexes generally consist of 22 β-sheets that form a barrel through which the iron-siderophore complex is transported (Figure 9). The extracellular loops that connect the β-sheets can extend up to 40 Å above the membrane surface and have importance in preventing the binding of unwanted species, as well as providing the initial interaction with the siderophore complex.[38-40] The second key feature of these receptors is the N-terminus occupying the central barrel, known as the 'cork domain'. This domain interacts with the bound complex and the periplasmic spanning protein TonB, which provides the energy for internalisation.[40-42] It is still unclear how the complex circumvents this cork domain, but it is generally accepted that the domain is partially displaced to allow the iron siderophore complex to pass through. Full displacement of the domain is considered unlikely, as this would involve breaking approximately 50 hydrogen bonds, thus having an energy barrier too high to overcome.[38, 43-45]

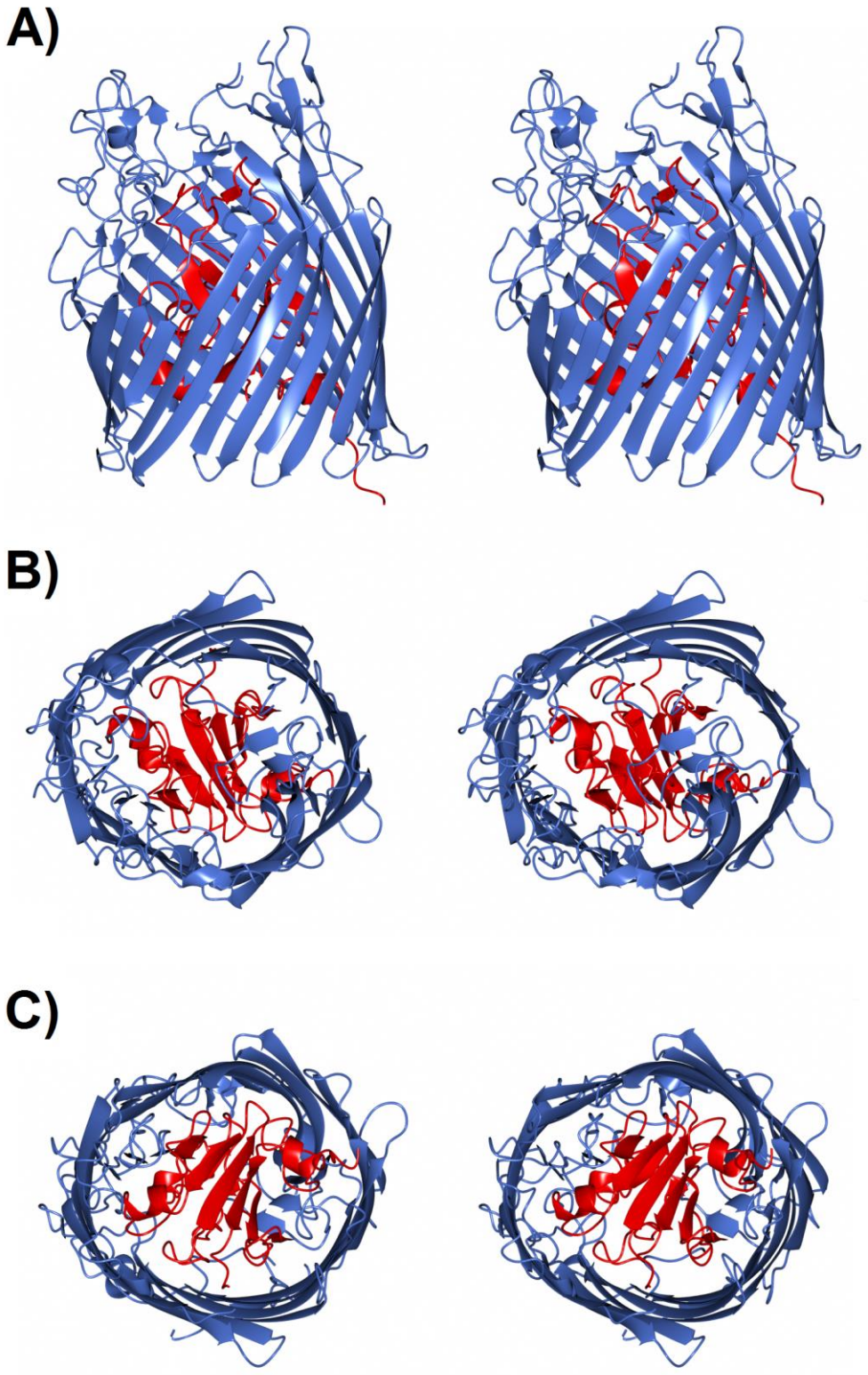


Figure 9 - Ribbon representation of the ferric enterobactin outer membrane receptor FepA, shown from: A) Side on; B) Top down (from the extracellular space); C) Beneath (from periplasmic space) Key: Barrel domain (light blue), Cork domain (red). PDB code: 1FEP.[38]

Once the iron-siderophore complex has been transported into the periplasm, it is bound by a PBP, which escorts the associated siderophores to the correct ABC transport complex. There are many different PBPs, as a result they are classified depending on their ligands. One class binds organic metal ion complexes. The sequence identity of PBPs varies greatly, and can be as low as 10%, yet they retain a similar tertiary structure (Figure 10).[9] PBPs are bi-lobal in structure, with the two domains connected *via* a long α-helix, with the cleft between the two lobes acting as the binding pocket. [9, 35, 46]

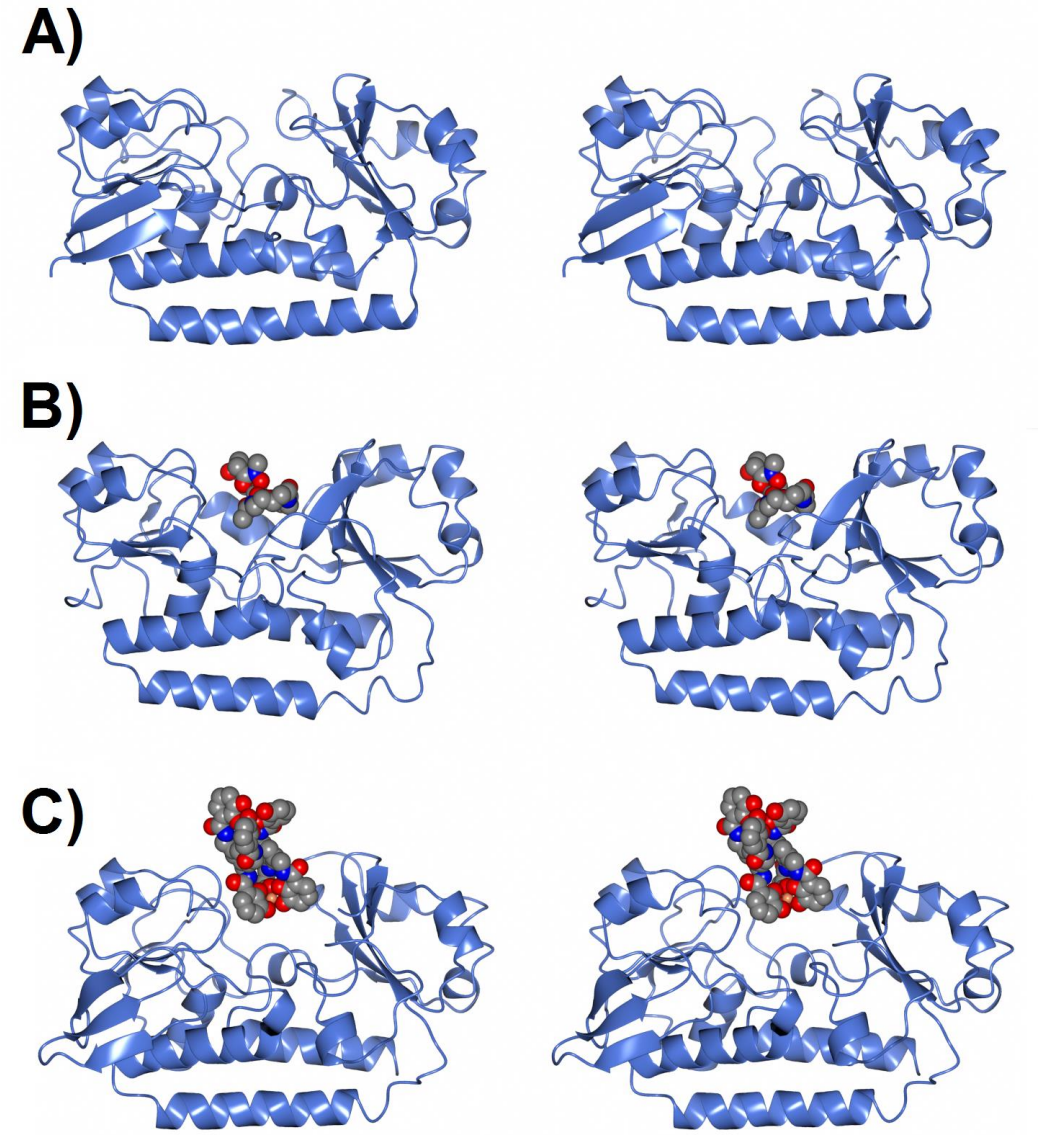


Figure 10 - Ribbon representation of PBPs: A) VctP (PDB code: 3TEF)[47]; B) FhuD bound to desferral (PDB code: 1K2V)[46]; C) CeuE bound to ferric-MECAM (PDB code: 2CHU)[35]. Key: PBP (light blue), ligands shown as spheres; grey - carbon, blue - nitrogen, red - oxygen, coral - iron.

Once the iron-siderophore complex has bound to the PBP, it must be transferred to the ABC transport complex to be shuttled across the inner membrane. Conserved glutamate and arginine residues on the PBP and the ABC transport complex, respectively, allow the two proteins to dock together, allowing exchange of the iron-siderophore complex. Unlike the outer membrane receptors, the ABC transport complexes are usually an assembly of proteins, rather than a single polypeptide. The ferric-enterobactin and the vitamin B_{12} ABC transport complexes in *E. coli* both consist of a dimer of dimers, $FepC_2D_2$ and $BtuC_2D_2$, respectively. In comparison, in the ferric-hydroxamate ABC transport complex in *E. coli*, one of the dimers is fused forming a $FhuBC_2$ complex. The only structurally characterised complexes of these transporters are $BtuC_2D_2$[48-50] and the molybdate transporter $ModB_2C_2$,[51-52] which are assumed to look similar to the ferric-siderophore uptake proteins (Figure 11).

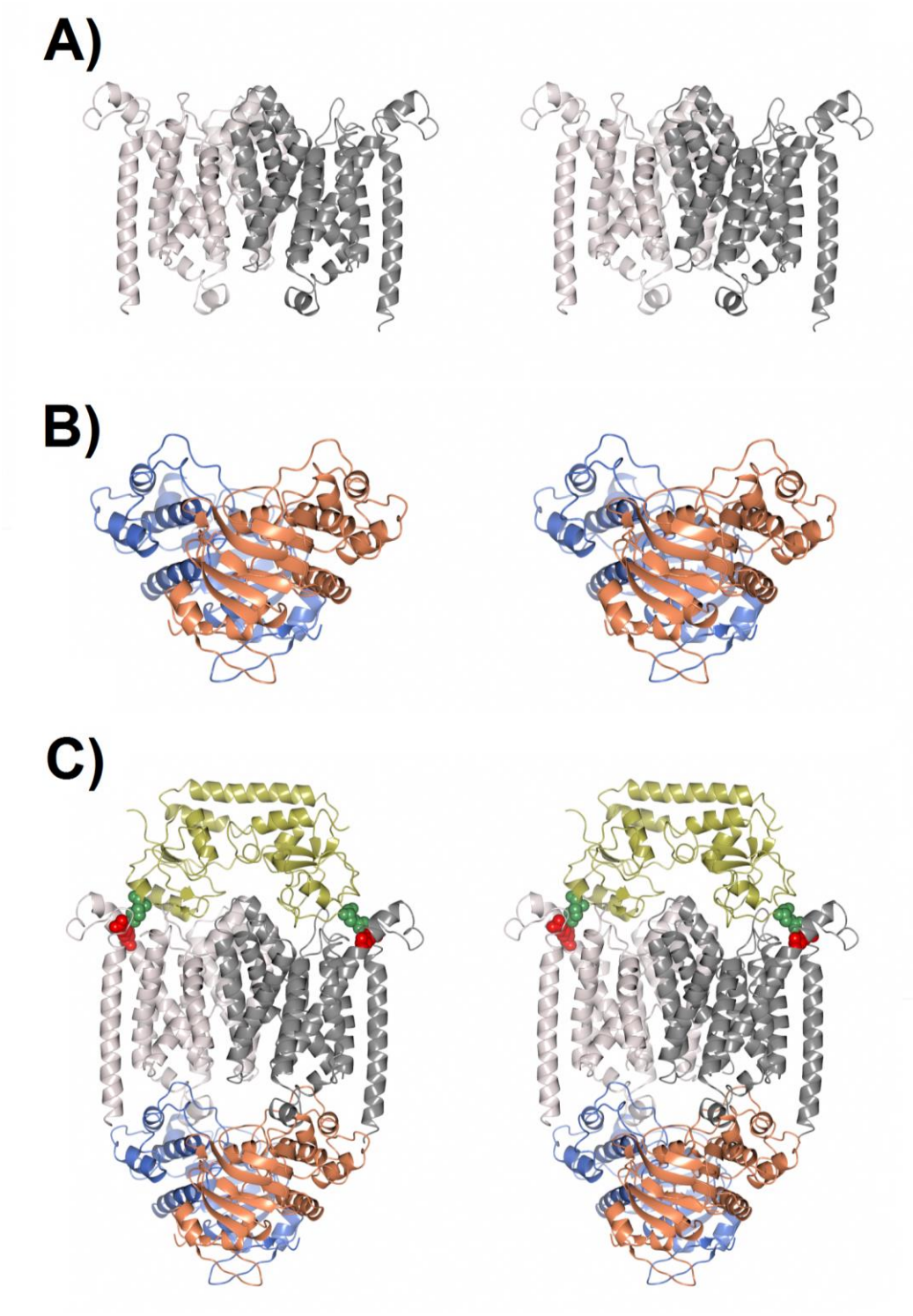


Figure 11 - Ribbon representation of $BtuC_2D_2F$ (PDB code 2QI9)[49]: A) Dimer of $BtuC_2$ (protomers shown in white and grey); B) Dimer of $BtuD_2$ (protomers shown in light blue and coral); C) Complex of $BtuC_2D_2F$. Key: $BtuC_2$ (white and grey), $BtuD_2$ (blue and coral), BtuF (gold) key docking residues shown as spheres; red - arginine, green - glutamate.

Gram-positive bacteria

Compared to their Gram-negative analogues, there is relativity little research published for iron-siderophore uptake mechanisms used by Gram-positive bacteria, with the majority of research focusing on *B. subtilis*.[36, 53-55] Gram-positive bacteria do not possess an outer membrane, and hence contain no periplasm. The uptake system therefore resembles that of the inner membrane transport system for Gram-negative bacteria. The PBP equivalent, a lipoprotein, is bound to the membrane *via* a membrane anchor.[53] These lipoproteins are structurally similarity to PBPs, and are assumed to interact with an ABC transporter in a similar fashion to that of Gram-negative bacteria (Figure 12).

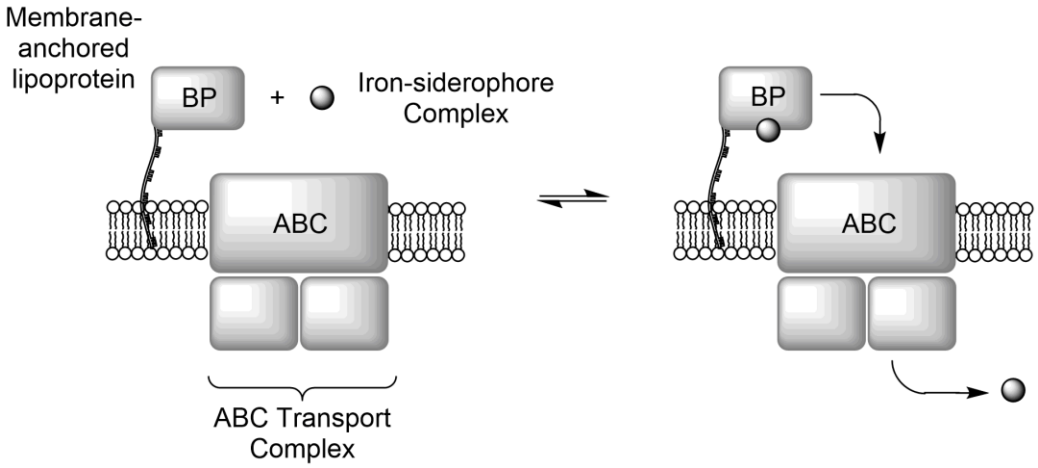


Figure 12 - A schematic diagram for a simplified overview of the iron-uptake mechanism for Gram-positive bacteria.

Iron release

As ferric-siderophore complexes are highly thermodynamically stable, specific mechanisms for iron release are required. Bacteria can employ a number of different strategies to access siderophore bound iron. Most involve transporting the iron-siderophore complex into the cytoplasm which gives access to enzymes capable of removing the iron from the siderophore. However, the release mechanisms for some siderophores, such as pyoverdine, has been shown to occur in the periplasm.[56] One of the most widely used mechanisms is the reduction of the coordinated iron(III) to iron(II).[57] This greatly reduces the

stability of the iron complex, allowing iron transfer. This mechanism is not available for the tris-catecholate siderophores, due to the extremely low redox potential (enterobactin $E_{½}$ = -750 mV).[2] As a result, the release mechanisms from these siderophores has been widely debated. There is now a consensus that the major release pathway of the iron is *via* siderophore degradation.[2, 7, 34, 37, 58-61] Studies with siderophore mimics such as H_6-MECAM, that are not prone to the action of an esterase, have shown that iron is still delivered to the cell, though at reduced rates (~5%), thus indicating the possibilty of a secondary release pathway.[33]

The secondary pathway has been proposed to be an acid-mediated release *via* change in coordination from the catecholate to the salicylate mode of binding (Figure 13).[59] This mode of binding weakens the affinity of the siderophore for the metal centre, and hence raises the redox potential, allowing reduction of the iron(III) to iron(II) and extraction of the metal as before. Details of this secondary release pathway are still unclear.

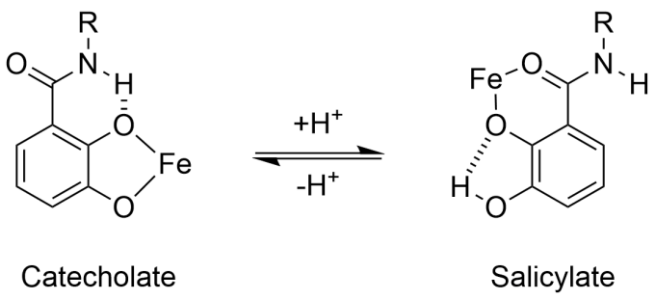


Figure 13 - Schematic diagram demonstrating the pH dependence of the catecholate and salicylate binding modes of the catecholamide group.

Regulation of Iron Uptake

The accumulation of excess iron can be just as hazardous to biological life as iron starvation. Consequently, the production of iron uptake and related proteins is highly regulated by the repressor protein Fur (ferric-iron uptake regulation).[62-64] Over 90 genes from different *E. coli* strains have been shown to be regulated

by Fur including siderophore biosynthesis, transporter systems and also oxidative-stress response systems (Table 2).[62]

Regulated Genes	Function	Reference
fepA-entD	Enterobactin synthesis and transport	65
fes-entF	Enterobactin synthesis and release	65
fhuACDB	Ferrichrome transport	62
iucBCD	Biosynthesis of aerobactin	64
sodB	Iron-dependent superoxide dismutase	66
acnA	Biosynthesis of isocitrate	67

Table 2 - Selected examples of Fur regulated genes from *E. coli* with a brief description of their function.

Fur comprises of two domains; a helix-turn-helix domain, and the carboxy-terminus domain which contains two metal-binding sites. In *E. coli*, one of these sites is occupied by a zinc(II) ion, while the other binds to iron(II) and regulates activity.[68] A dimer of iron loaded Fur binds to a specific 19 bp DNA sequence known as the 'Fur Box', which is located within the promoter of the regulated genes. The affinity of Fur towards iron(II) is balanced so as to allow the accumulation of iron within the cell, but if the intracellular concentration becomes too high, Fur becomes loaded with iron, and binds to the Fur Box which represses further iron uptake, helping to prevent toxic concentrations of iron accumulating within the cell (Figure 14).[63]

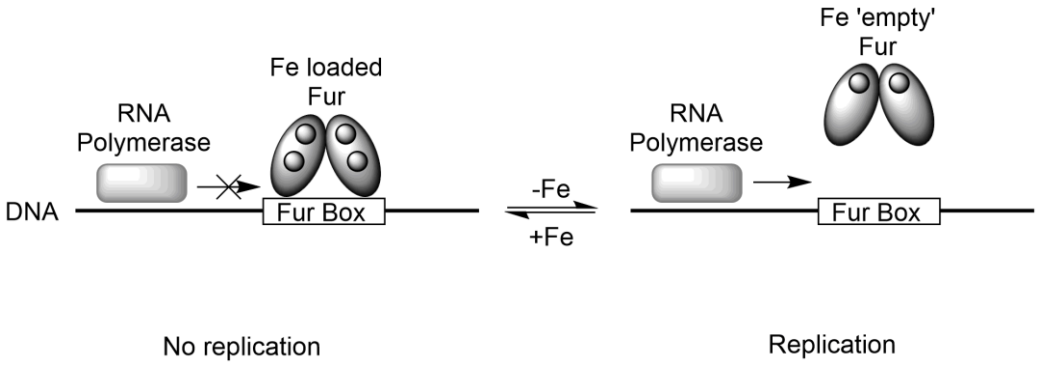


Figure 14 - A schematic diagram demonstrating the general mechanism of action of Fur.

Campylobacter jejuni

C. jejuni is a Gram-negative, microaerobic bacterium.[69] It is the most prevalent cause of gastroenteritis worldwide, and is a major cause of food-borne diarrhoeal illness in humans,[70] with iron acquisition demonstrated to be an important virulence determinant.[71-73] The importance of iron for *C. jejuni* is suggested by the relatively large number of genes associated with iron uptake and regulation in the genome.[74] As a Gram-negative bacterium, it has an iron uptake system similar to that discussed above. *C. jejuni* is thought to be incapable of synthesising its own siderophores, however, most strains have been shown to utilise siderophores such as enterobactin, produced by other bacteria.[69, 75] The major iron uptake pathways for *C. jejuni* are based on enterobactin, haem and transferrin proteins (Figure 15).[71, 74]

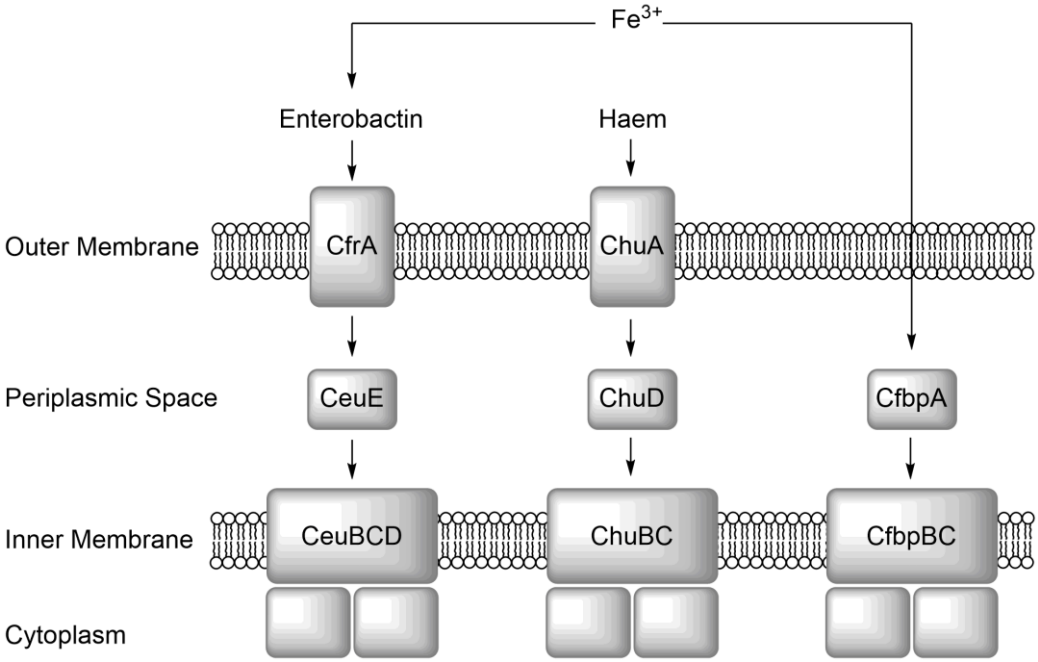


Figure 15 - Schematic representation of the three major iron uptake systems of *C. jejuni*.

Currently, the only structurally characterised protein from the enterobactin uptake system of *C. jejuni* is the PBP CeuE.[35] Müller *et al.* co-crystallised CeuE with ferric-MECAM and found it to form an unusual dimer complex. The

interactions between the two proteins are minimal, suggesting that the driving force for dimerisation is due to the ligand. The overall complexation is stabilised by favourable hydrophobic interactions, as the methylene group of one of the MECAM ligands is directly above the other MECAMs aromatic ring (Figure 16).

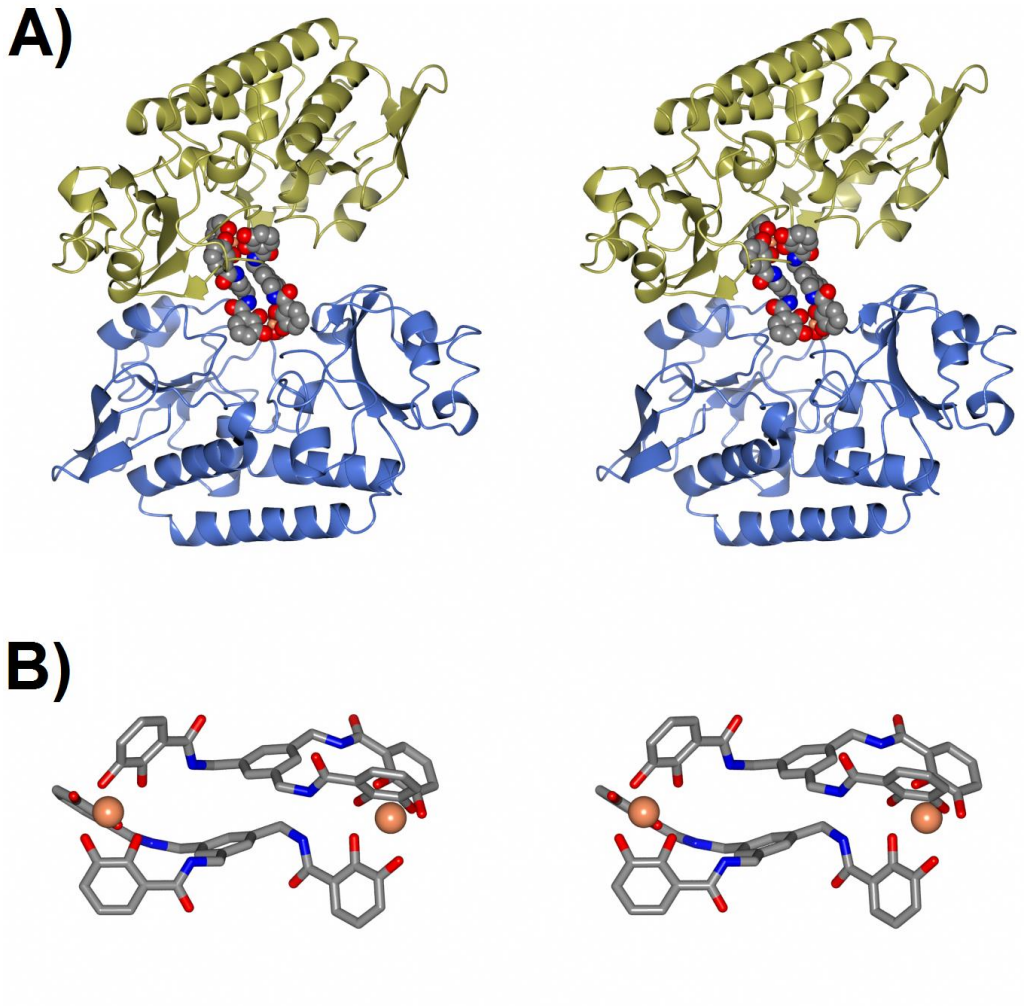


Figure 16 - The structure of the {$CeuE_2[Fe_2MECAM_2]$} complex (PDB code: 2CHU).[35] A) CeuE chain A (light blue), CeuE chain B (gold), ligands shown as spheres; B) The MECAM molecules shown as cylinders. Key: grey - carbon, blue - nitrogen, red - oxygen, coral - iron.

The formation of dimerised PBP using ferric-MECAM also offers an explanation for ferric-MECAM reportedly becoming trapped within the periplasm.[33] Ferric-MECAM has also been co-crystallised with the Gram-

positive binding protein FeuA, and was found to form the 'usual' 1:1:1 complex, thus demonstrating that ferric-MECAM does not dimerise all binding proteins.[36]

1.5 Applications of Siderophores

Siderophores in Supramolecular Chemistry

As stated above, not all siderophores are hexadentate and hence different stoichiometries with metal ions are possible. By utilising the preferred coordination requirements of metals and ligands it is possible to use metal-ligand interactions to drive the formation of molecular assemblies.[76-78] One of the smallest assemblies is the triple helicate (M_2L_3), which forms when three ligands bridge between two metal centres of the same chirality. If the chirality is reversed at the metal centres, then it is referred as a triple mesocate (Figure 17).[79]

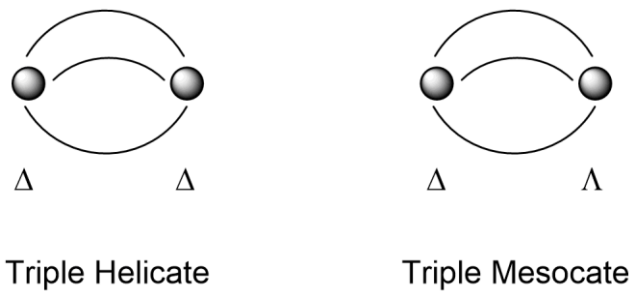
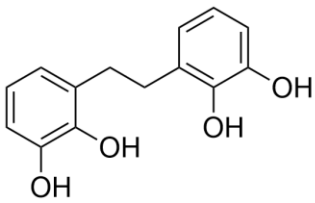
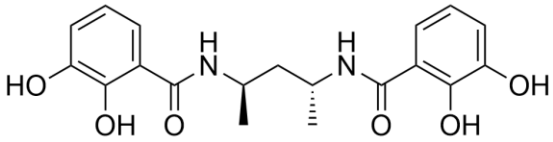


Figure 17 - Schematic diagram of a triple helicate and triple mesocate.

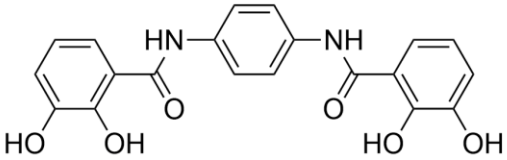
Both triple helicate and mesocate systems with catechol containing ligands have been reported with gallium(III), cobalt(III), aluminium(III), iron(III) and titanium(IV).[76, 80-85] The key design features are a C_2 symmetric bis(bidentate) ligand and a metal ion with pseudo-octahedral coordination. The ligands usually possess a rigid linker that directs the formation of an M_2L_3 system, although it is still possible to form these systems with flexible linkers (Figure 18).[76]



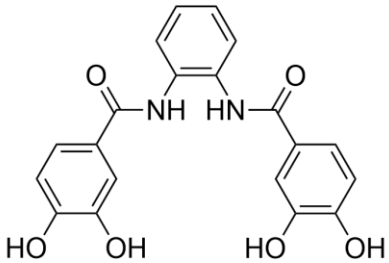
H_4-**4**



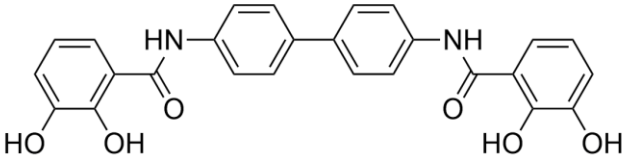
H_4-**5**



H_4-**6**



H_4-**7**



H_4-**8**

Figure 18 - Example bis(catechol) ligands that have been shown to form triple helicates.[76]

Simple bis(catecholamide) ligands with short spacer units such as H_4-2-LICAM, which possesses an ethylene, was found to form a dinuclear 3:2 helicate, as well as a dinuclear 2:2 hydroxo-bridged dimer (Figure 19).[86] The naturally occurring bis(catecholamide) siderophore, which possesses a spermidine linker, was also investigated and shown to also form a 3:2 helicate. There was no evidence of formation of a hydroxo-bridged dimer with this ligand, showing linker length and type has a profound affect on complexation.[86]

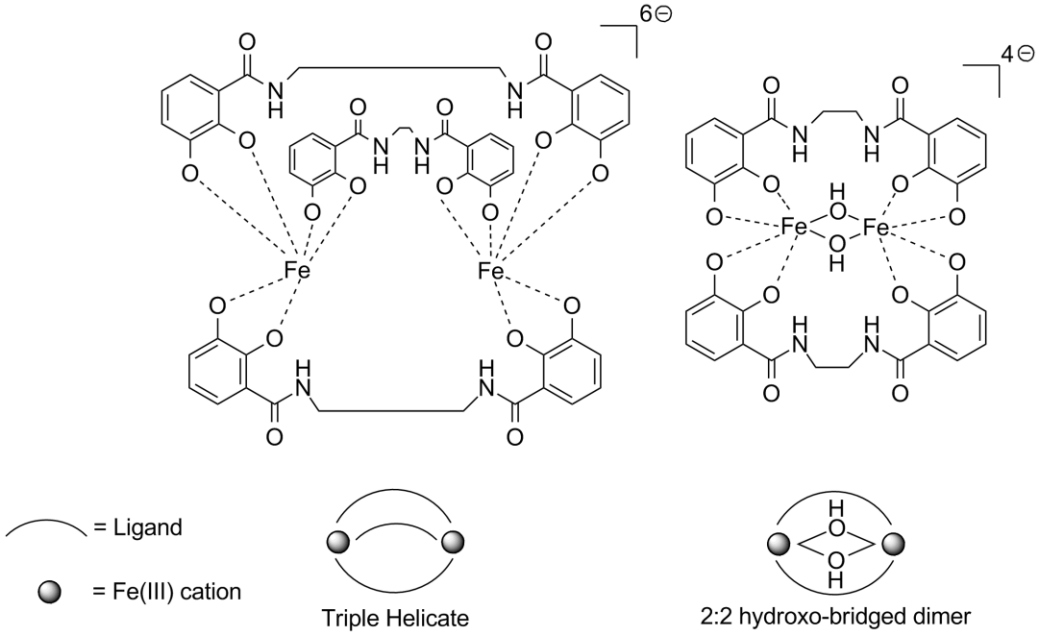


Figure 19 - Structures of the two complexes 2-LICAM^{4-} forms with iron(III), $[Fe_2L_3]^{6-}$ and $[Fe_2L_2OH_2]^{4-}$.

Studies of similar tetradentate ligands with hydroxamate binding groups have identified both mononuclear complexes with a ligand-to-metal ratio of 1:1, dinuclear, 2:2 complexes and dinuclear 3:2 triple stranded helicates forming with a dependence on linker length.[87] Even 3:2 complexes that adopt a mono-bridged dinuclear structure have been reported with the bis-hydroxamate siderophore alcaligin (Figure 20).[88]

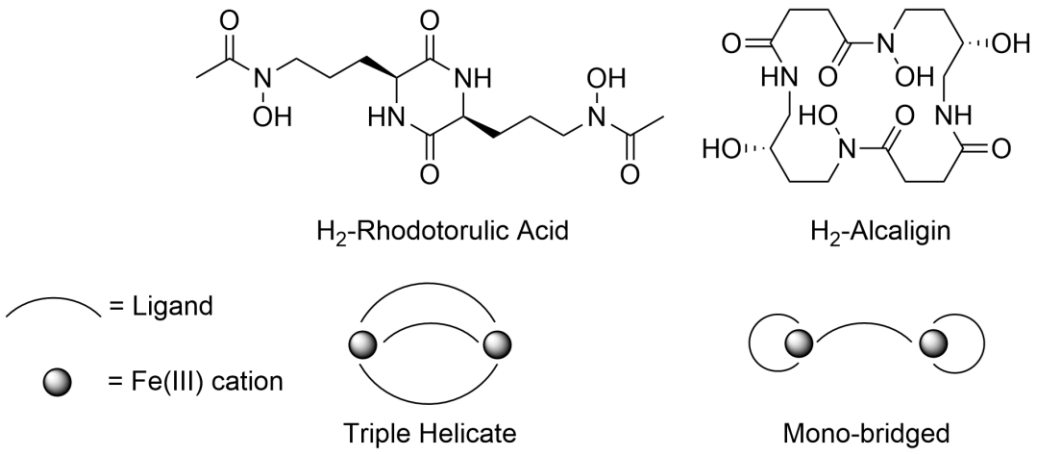


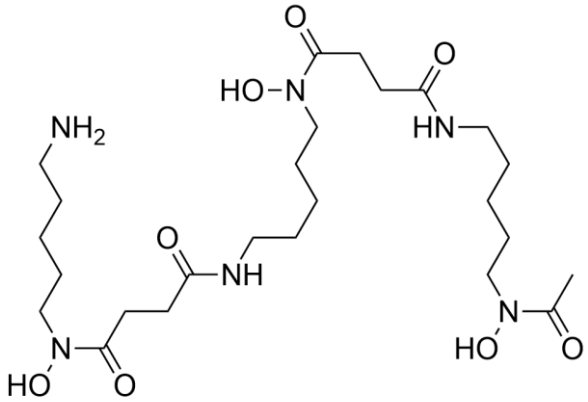
Figure 20 - Structures of the natural siderophores rhodotorulic acid and alcaligin and schematic diagrams of their dinuclear six-coordinate iron complexes.

The ligand H_4-3-LICAM was studied with gallium(III) using ^{1}H NMR spectroscopy, which suggested 3-LICAM^{4-} coordinates to gallium(III) in a 3:2 triple stranded fashion.[80] H_4-5-LICAM was crystallised with Ni^{2+} forming a mononuclear 1:1 complex in a square planar geometry.[89] H_4-5-LICAM was also crystallised with a *cis*-dioxomolybdenum(VI) unit (MoO_2^{2+}), which also requires only four coordinating atoms. However, unlike nickel, the molybdenum formed a dinuclear complex, where two 5-LICAM^{4-} ligands bridged two MoO_2 centres.[90] A follow-up study using a xylene linker found that a dinuclear system could again be formed, which can be converted to the mononuclear system upon heating.[91] Many other ligands have been shown to form triple stranded helicates,[92-93] and larger molecular cages,[76, 79, 94] which have been widely observed and reviewed.[77-78, 95-97]

Medicinal Applications of Siderophores

A major research area for applications of siderophores is the development of new chelation therapeutics, particularly for the treatment of iron overload.[3, 28, 98-99] Iron overload is the accumulation of iron within the body, which can impair organ function and if left unchecked, can lead to multiple organ failure. This condition is either caused by genetic defects, known as haemochromatosis, or as an indirect result from other conditions, such as treatment for β-thalassaemia, known as siderosis.[100-101]

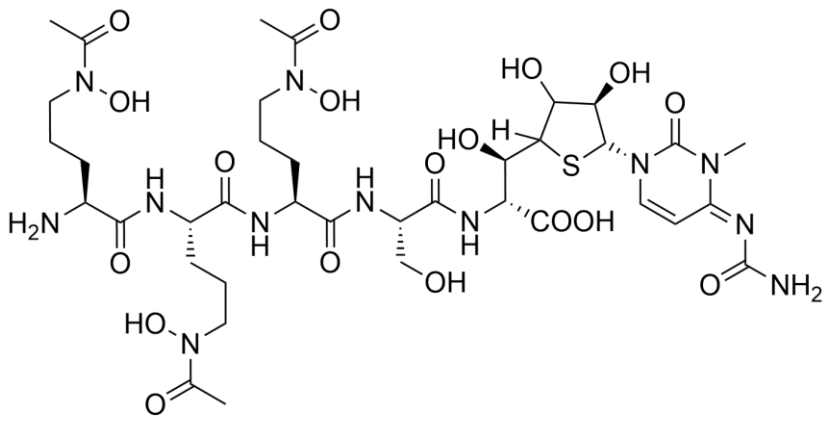
In designing suitable chelators for clinical applications, many of the design principles can be mimicked from siderophores, such as metal selectivity and complex stability. For chelators targeting iron(III), the same binding moieties can be used to achieve selectivity over other biologically important ions. For chelation therapies targeting other ions, the coordination chemistry of the ligand can be tailored to suit the requirements of the metal centre. For example, Weitl *et al.* developed actinide chelators based on four catecholamide units, which show selectivity for the larger actinide ions over iron(III).[102] One of the most widely used chelators for iron(III) currently in use, is the natural hydroxamate siderophore desferrioxamine-B (DFO).[101]



H_3-Desferrioxamine-B (DFO)

Another consideration required for chelation therapy, is the stability of the iron-complex towards the uptake systems of micro-organisms. An adverse side effect in the use of siderophores as iron-chelating therapeutics is the increase in micro-organism growth, resulting in an increased risk of infection.[103-104] Therefore, it would be advantageous to use siderophore mimics which cannot transport iron into micro-organisms.

Other areas of research rely on siderophores for targeting, rather than exclusively for their metal chelating properties. The Trojan Horse Strategy, for example, involves covalently attaching known anti-microbial agents to siderophores to produce siderophore-conjugates. These conjugates can coordinate iron(III), and hence the siderophore uptake system of the targeted micro-organism should actively transport the conjugate into the cell.[105-108] It is purposed that this strategy would evade certain aspects of anti-microbial resistance mechanisms. This strategy was inspired by a range of natural compounds, known as sideromycins, which operate by the same method.[108] Albomycin is the most commonly studied of the sideromycins, and consists of a tris-hydroxamate siderophore unit linked *via* an amide bond to a thioribosyl pyrimidine tRNA synthetase inhibitor.[108-109] In the cytoplasm of *E. coli*, the amide linkage is cleaved by peptidase N, releasing the antibiotic required for anti-microbial activity.



H_3-Albomycin

Since the discovery of sideromycins, synthetic and semi-synthetic Trojan Horse conjugates have been made, and their anti-microbial activity evaluated with varying degrees of success.[110-116] Key design features required for all Trojan Horse siderophore systems are; recognition of the siderophore group, to allow active transport through the iron uptake system, and the availability of a chemical linking group to allow the covalent linkage to the anti-microbial group.

Other Applications of Siderophores

An area related to the Trojan Horse strategy is the delivery of chemical 'cargo' into micro-organisms using siderophores. Enterobactin was modified by Zheng *et al.* for evaluation of delivering different cargos.[112] Interestingly, the point of modification to attach the cargo was the C5 position on the catechol ring. This position has been modified by nature through the addition of a glucose ring, forming a class of siderophores known as salmochelins.[117] Salmochelins (mono- or di-substituted glucosylated enterobactin and associated hydrolysis products) evolved in response to the human immune system recognising and binding ferric-enterobactin, using proteins known as siderocalins.[118-120] While the addition of the glucose ring in the C5 position has no adverse effect on substrate binding for the iron(III) uptake proteins, it inhibits binding by the siderocalins. The structure of these salmochelins inspired the addition of chemical linkers on the C5 position, for the study developing H_6-**9** (Figure 21).[112]

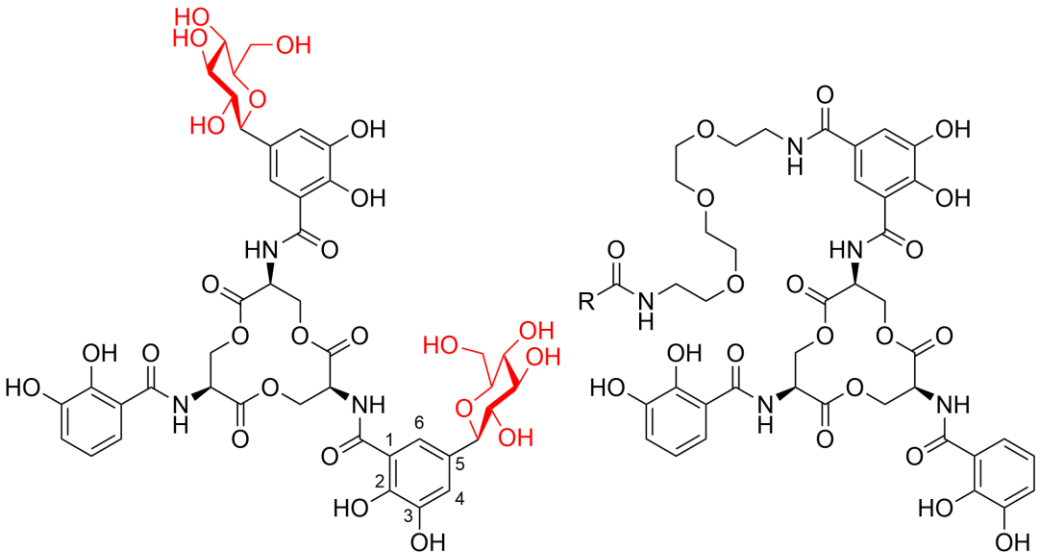
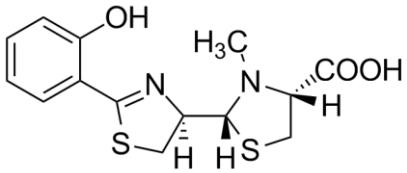


Figure 21 - Chemical structures of DGE, one of the salmochelins and H_6-**9** designed and synthesised by Zheng *et al.*[112]

The modified enterobactin H_6-**9** was recognised and transported into *E. coli* and *P. aeruginosa*. Small cargos, such as Boc, cyclohexyl, napthyl and phenylmethylbenzyl groups, were shown to be recognised and transported into the cell. Conjugates with larger cargos, such as coumarin, were actively transported into *P. aeruginosa* but not *E. coli*, presumably due to the different iron uptake systems.[112]

Modification of a natural siderophore with a fluorescent tag for the development of targeted probes has been demonstrated by Noël *et al.*[121] Pyochelin is one of two major siderophores produced by *P. aeruginosa*, an opportunistic Gram-negative bacterium, which cause severe lung infections that can be fatal for cystic fibrosis patients.[121] Pyochelin can either be a tridentate or tetradentate siderophore which was shown to bind to not only iron(III) with high affinity but also zinc(II), copper(II), cobalt(II), nickel(II) vanadium(IV), molybdenum(VI), gallium(III) and terbium(III).[18-20, 45, 122-123] These metal-chelates are predominantly bound in a 2:1 ligand-to-metal ratio.[18] Only metal complexes of iron(III), cobalt(II), nickel(II), gallium(III) and molybdenum(VI) are transported

through the pyochelin uptake pathway but at a reduced rate compared to the iron(III) complex.[19-20]



H_2-Pyochelin

The ferric-pyochelin complex is recognised by the outer membrane protein, FptA which transports it into the periplasm of the cell.[18] Due to the wide range of metals that pyochelin is able to chelate and deliver to the bacterium, the exact role of this siderophore during infection and its interactions with other biological ions remains unclear.

In response to this challenge, Noël *et al.* designed and synthesised two pyochelin conjugates as molecular tools for the investigation of the pyochelin-dependent iron uptake pathway.[121] Pyochelin was covalently linked to the (4-nitro-benzo[1,2,5]oxadiazole) fluorophore using two different linkers, a short succinic linker and a longer spacer to minimise steric clashes with uptake proteins. The conjugates were shown to be recognised by the pyochelin-specific outer membrane receptor, by selectivity labelling cells that expressed the receptor over cells that did not (Figure 22).

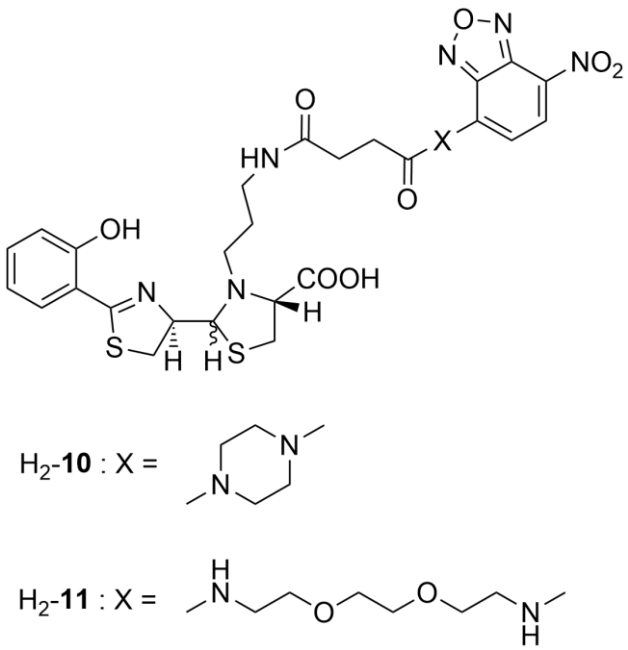
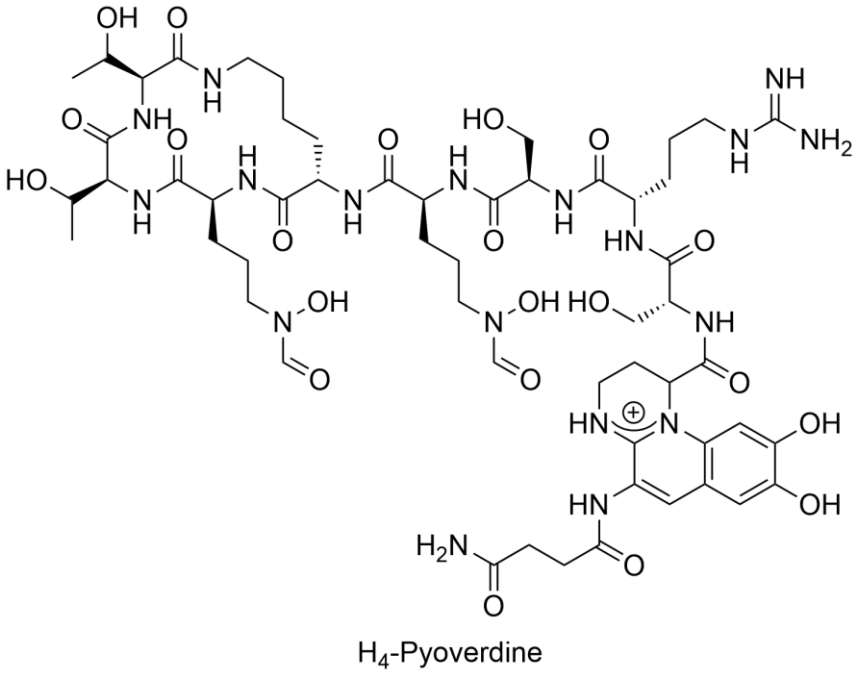


Figure 22 - Chemical structure of fluorescently-labelled Pyochelin.[121]

Another important application of siderophore chemistry is the development of novel sensors. The most obvious example is the detection of iron(III) utilising naturally emissive siderophores such as pyoverdine.[124]



Pyoverdine is produced by all fluorescent pseudomonads, with the exact structure varying between strains. Pyoverdine is a mixed-type siderophore,

forming hexacoordinate complexes with iron(III), binding through one catecholate and the two hydroxamate groups. The unusual catechol group is incorporated as part of the dihydroxyquinoline group, which acts as the chromophore, with the rest of the siderophore made up of both D- and L-amino acids.[125] Immobilisation of pyoverdine in a porous structure of either silica or sol-gel glass, allowed the development of fluorescent sensors for the determination of trace iron(III) levels.[126-128]

1.6 Project Aims

The overall aim of this project was the investigation of interactions between iron-siderophore complexes with selected binding proteins involved in iron(III) uptake. The investigations can be divided into three distinct areas. The first involves the study of tetradentate siderophores interacting with iron(III) in the absence and presence of the periplasmic binding protein CeuE (Chapters 2 and 3). The second is the development of a novel hexadentate siderophore mimic to be used for the attachment of secondary chemical functional groups (Chapter 4). The third involves the modification of a natural siderophore with a fluorescent unit, for use as a novel fluorescent probe for investigating integral membrane proteins and their movements with the outer cell membrane (Chapter 5).

Chapter 2: Tetradentate Siderophores, a Biomimetic Approach

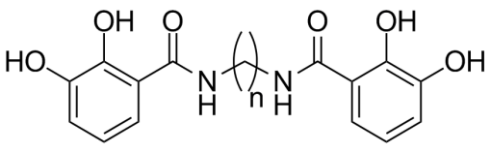
Project Aims

The aim of this chapter was the investigation of a biomimetic model of the enterobactin linear dimer interacting with iron(III) in the absence and presence of the periplasmic binding protein CeuE. Characterisation of possible binding stoichiometries of the iron(III)-ligand complex under biologically relevant conditions; and the determine of the type and strength of interactions of this complex system with CeuE.

2.1 Introduction

As outlined in Chapter 1, enterobactin is comprised of a hydrolytically unstable tri-lactone backbone. Hence, enterobactin is prone to degradation in aqueous medium and can form three distinct enterobactin fragments, the linear trimer, linear dimer and DHBS, each of which mediates iron uptake into bacteria.[1-2] It was therefore of interest to investigate how hydrolysis products, such as the tetradentate linear dimer, interact with iron. In addition, the binding of the resulting iron-complex(es) with their cognate iron-uptake proteins was investigated.

Due to the chemical complexity of the linear-dimer derived from enterobactin, a simplified biomimetic model was used initially to study the potential interactions of the linear dimer with metal ions and the periplasmic binding protein CeuE. The LICAM (LInear Catechol AMide) series of compounds was selected for this purpose due to their simple chemical structure, which consists of two catechol amide groups attached via a linear alkanediyl linker.

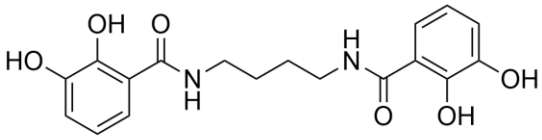


H_4-n-LICAM

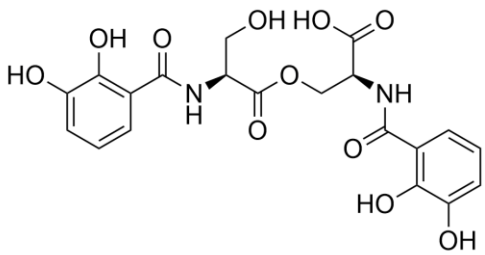
A number of ligands of this series, 2-,4-,6-LICAM, was prepared previously by Weitl *et al.*[3] In addition, 3-,5-,8-,9-,10-,12-LICAM were prepared by Bhargava *et al.*[4], focusing on metal coordination for chelation therapy. Subsequently, the LICAM series of compounds was used for the investigation of binding modes in combination with divalent metal centres, such as nickel(II), cobalt(II) and also the *cis*-dioxomolybdenum(VI) unit (MoO_2^{2+}),[5-7] as well as the trivalent metal centres gallium(III) and iron(III).[8-9] In order to mimic the enterobactin linear dimer successfully, the alkanediyl linker between the two catecholamide functional groups should have a similar length as the linear dimer; therefore a butylene linker was used first.

H_4-4-LICAM

The bis-catecholamide ligand H_4-4-LICAM, is comprised of two catechol-amides linked via a butylene linker. Compared with the enterobactin linear dimer, H_4-4-LICAM avoids the chemically complexity of the functional side chains, the chiral centres and the chemically unstable ester backbone.



H_4-4-LICAM



H_4-Linear Dimer

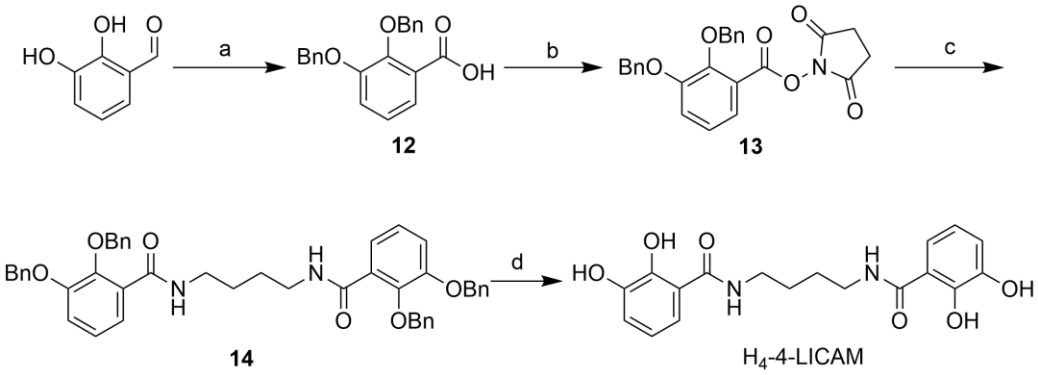
H_4-4-LICAM was first synthesised by Weitl *et al*. as a precursor for the investigation of actinide sequestering agents based on sulfonated catechol ligands.[3] Subsequently, it was investigated as part of a series of tetradentate catechol ligands that form five-coordinate organoarsenic complexes.[10-11] H_4-4-LICAM, along with other similar ligands, was also investigated for chloride binding to demonstrate the importance of the hydrogen bonding network around the catechol functionality towards anion binding.[12]

Previous synthetic routes towards H_4-4-LICAM involve linking 2,3-dimethoxybenzoyl chloride to 1,4-diaminobutane, followed by deprotection of the methyl ethers using boron tribromide. In some cases, the deprotection was reported to work in excellent yields,[3] however, yields have also been reported to be as low as 15%, suggesting that deprotection and/or purification is not straightforward.[12] In addition to the uncertain yields, boron tribromide is a toxic reagent. To avoid these issues a different synthetic route was devised.

2.2 Synthesis of H_4-4-LICAM

The synthesis was carried out according to literature procedures with the exception of altering the catechol protecting groups from methyl ethers to benzyl ethers.[3, 13-15] This change was made due to the ease of benzyl deprotection utilising benzyl hydrogenolysis compared to methyl ether deprotection requiring boron tribromide.

The catechol-amide groups were synthesised from 2,3-dihydroxybenzaldehyde. First the phenolate oxygen atoms were protected as benzyl ethers, using benzyl chloride,[13] followed by oxidation of the aldehyde group to the corresponding carboxylic acid with sodium chlorite and sulfamic acid.[14] The carboxylate functionality was then activated as a succinic ester for the amide coupling, as previously reported.[15] Compound **13** was then reacted with half an equivalent of 1,4-diaminobutane, yielding the benzyl protected 4-LICAM. This novel compound was characterised using ^{1}H, ^{13}C NMR and infra-red spectroscopy, mass spectrometry, and elemental analysis. The benzyl ether protecting groups were removed using hydrogen over a palladium catalyst, resulting in the formation of the final compound H_4-4-LICAM, the characterisation data of which are consistent with the literature (Scheme 1).[3, 12]



Scheme 1 – Synthesis of H_4-4-LICAM. (a) i) Benzyl chloride, K_2CO_3, EtOH ii) $NaClO_2$, H_3NSO_3, acetone/water, 79%; (b) *N*-hydroxysuccimide, DCC, 1,4-dioxane, 71%; (c) 1,4-diaminobutane (0.5 eq), NEt_3, acetone, 94%; (d) H_2, Pd-C 10%, EtOH, 99%. Overall yield 52%.

2.3 Ligand-Metal Interactions

Due to the mismatch between the preferred coordination number of iron(III) and the denticity of H_4-4-LICAM (six and four respectively), it was of interest to investigate the speciation of the complexes formed in solution (see Chapter 1).

The literature suggests that many of these ligands can form a range of complexes with different binding stoichiometries. Based on these literature reports, there would appear to be three possible structures in which H_4-4-LICAM could interact with a iron(III) centre. These are a mononuclear 1:1 complex, a dinuclear 2:2 complex, both with solvent molecules filling the vacant coordination sites on the metal centre, or a dinuclear triple standard 3:2 complex (Figure 23).

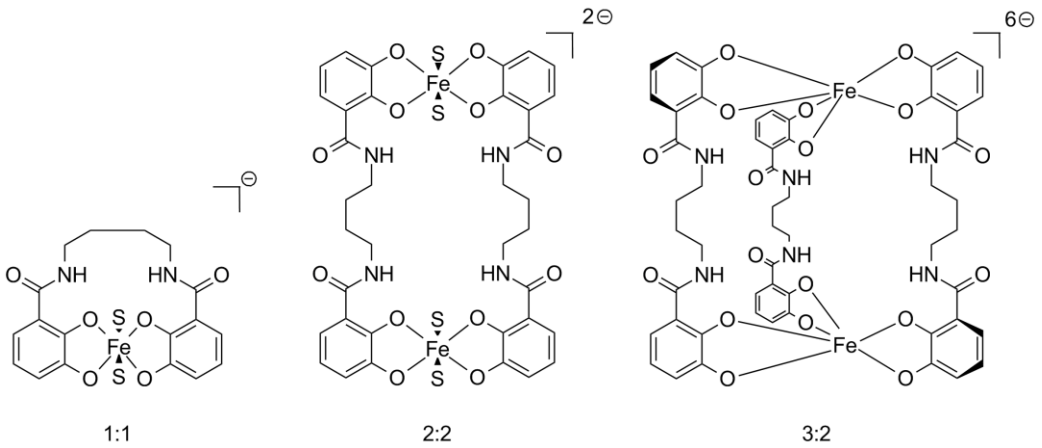


Figure 23 - Possible iron(III)-4-LICAM complexes that could form in solution.

Stoichiometric Ratio Determination of H_4-4-LICAM and Iron(III) using the Continuous Variation Method

In order to determine the composition of the ligand-metal complexes formed in solution, the method of continuous variation (Job plot) was applied to H_4-4-LICAM and iron. This method varies the concentrations of metal ([M]) and ligand ([L]) whilst allowing the sum of their concentrations to remain constant. The evaluation of a unique signal arising from complexation allows the determination of the ratio of components within a complex.[16]

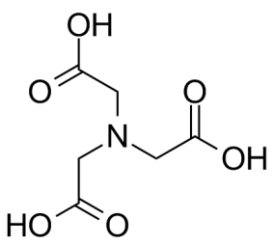
The development of a Job plot based on UV-visible spectra, allows the monitoring of catecholamide-iron(III) speciation *via* the ligand-to-metal charge transfer (LMCT) band arising from the catechol-iron(III) interaction, under biologically relevant conditions once the system has reached equilibrium.[17-19] The measurements required for the Job plot were carried out in buffered aqueous solution at pH 7.5. Due to the limited solubility of H_4-4-LICAM in water, DMSO was added. To ensure consistency, the concentration of DMSO was kept to 5% for all samples under investigation.

Initial studies showed that the ferric iron that was added to the samples resulted in the formation of insoluble ferric hydroxide species within eight hours. Only in the presence of excess ligand was the iron held in solution by complexation.

However, with excess of iron, a visible precipitate would form, which had to be prevented as it would alter the final concentration of iron in solution, and hence affect the result of the Job plot. Measuring the UV-visible absorbance of freshly prepared samples, before precipitate formation occurred, would not reflect the species distribution at equilibrium.

In aqueous conditions at neutral pH, iron(III) readily hydrolyses and precipitates, initiated by the formation of small inorganic polymers. These polymers are able to grow into larger clusters, eventually forming colloidal sized hydroxide solids of various compositions (Fe^{3+}, $Fe(OH)^{2+}$, $Fe(OH)_2^{+}$, $Fe(OH)_{3(aq)}^{0}$ and $Fe(OH_4)^{-}$).[20]

To avoid iron precipitate formation in the samples where an excess of iron is required, a second iron chelating ligand had to be introduced to compete with the hydroxide ligand. Nitrilotriacetic acid, (NTA), has previously been used to solubilise iron(III) in solution and was reported not to effect the results obtained from spectrophotometric titrations.[21-22] The first investigation, by Scarrow *et al.*, examined the complex formation between iron and enterobactin hydrolysis products. The absorbance arising from the LMCT band due to the iron-catechol interaction was plotted against the equivalents of iron per ligand. The resulting plot showed that of the enterobactin hydrolysis products DHBS and the linear dimer could coordinate to one and two iron centres per ligand respectively.[21]



NTA

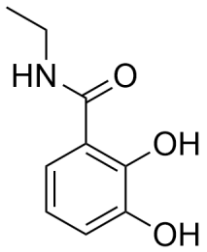
The second investigation by Ma *et al.* was focused on developing a fluorescence-based method for pFe^{3+} determination. The fluorophore chelated iron(III) with a well characterised binding constant, while competing ligands

were introduced to remove the iron(III) from the fluorophore. The removal of iron(III) would lead to an increase of the fluorescent signal, which could be used to calculate the affinity of the ligands for iron(III).[22] Both studies reported that NTA did not interfere to a significant extent with the results, instead it rapidly and completely delivered iron to the ligands with catecholate and 3-hydroxypyridin-4-ones binding moieties.[21-22]

The formation constants of NTA with various metal cations have been previously determined.[23] The reaction of Fe^{3+} with NTA gives FeL, with only limited formation of a FeL_2 species. The 1:1 complex was reported to have a logK value of 16, while for the formation constant of the 1:2 complex logK was 8.5. However, this value was been questioned due to the formation of ferric hydroxide under the experimental conditions.[23] The formation constant of a single catechol towards Fe^{3+} is reported to have a logK of 20. Therefore it would be expected that NTA should have minimal competition with catechol binding; however, it would still be able to compete with any hydroxide ions to stop precipitate formation.[24]

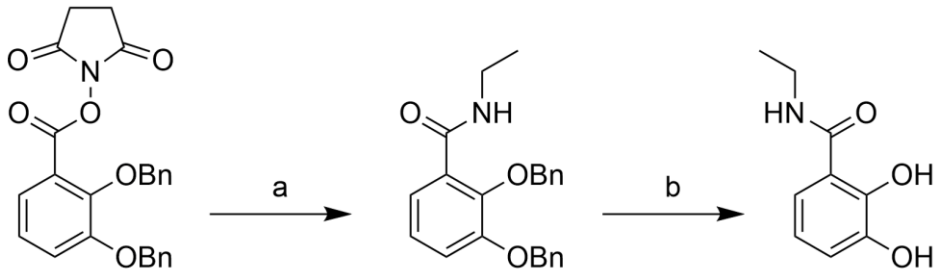
Additional control experiments were carried out to experimentally establish the effect of NTA under the experimental conditions used for the Job plot analysis. A catecholamide control compound was used to investigate if NTA is competing to a significant extent with the catechol functionality for iron and to ensure that NTA inhibits iron precipitation under the chosen conditions.

Thus *N*-ethyl-2,3-dihydroxybenzamide H_2-**15** was synthesised to act as the control. The bidentate ligand H_2-**15** was chosen to investigate if a single catecholamide unit can displace NTA under the conditions used to form the coordinatively saturated mononuclear 3:1 complex, plus potentially, coordinatively unsaturated 2:1 and 1:1 complexes. The presence of the amide functionality should keep the electronic and hydrogen bonding properties of the catechol ring similar to that of H_4-4-LICAM and other natural catecholate siderophores.



H_2-**15**

The synthesis utilised the benzyl protected activated succinic ester catechol unit (**13**), which was coupled to ethylamine.[25] The benzyl protecting groups were removed using catalytic hydrogenation. The product was purified by sublimation, to give product H_2-**15** in an overall yield of 53%. The characterisation data are consistent with literature data (Scheme 2).[26] The synthesis, characterisation and Job plot analysis of H_2-**15** was completed by BSc project student Bethany Peeters under my supervision (2012-2013).



13 **16** H_2-**15**

Scheme 2 – Synthesis of H_2-**15**. (a) 2.0 M ethylamine, NEt_3, THF, 72%; (b) H_2, Pd-C 10%, EtOH, 74%. Overall yield 53%.

The stoichiometric ratio of compound H_2-**15** and iron(III) was then determined in the presence and absence of NTA, using either a 10 mM $Fe(NO_3)_3$ + 10 mM NTA aqueous stock solution, or a 10 mM $Fe(NO_3)_3$ aqueous stock solution with no NTA present as the iron source. The samples were prepared with the iron solution being introduced last to minimise precipitation for the samples with no NTA. The samples were thoroughly mixed, and the absorbance reading taken within a minute to minimise precipitate formation in the samples with no NTA. Data points are the average of duplicate runs, with the standard deviation used as the error of the data points (Figure 24).

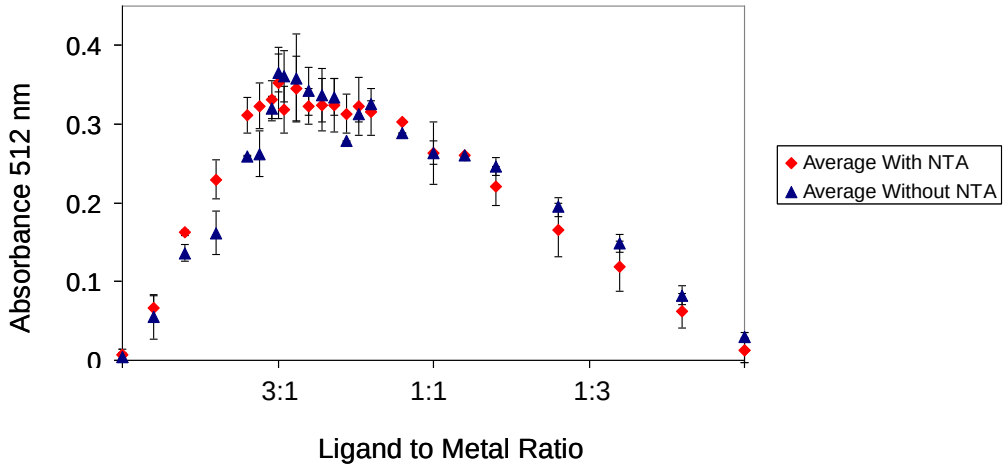


Figure 24 - Job plot for the binding of H_2-**15** to iron (III), obtained by following the absorbance at 512 nm in the presence (red diamonds) and absence (blue triangles) of NTA. The absorbance values are averages of two experiments and the error bars indicate the differences between the runs. Plot recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO.

The maximum of the Job plot obtained indicates a ligand-to-metal ratio of 3:1 in both cases, as expected for a bidentate ligand that coordinates to a metal centre with a preferred coordination number of six. Visual inspection of the samples confirmed that the presence of NTA prevented the iron from precipitating, which together with the similarity between the two plots indicates that NTA does not interfere with the coordination of **15**$^{2-}$ to a ferric metal centre. It was therefore concluded that NTA is a suitable additive.

The formation of more than one species becomes apparent when examining the whole set of UV-visible spectra obtained (Figure 25). The maximum initially observed at 512 nm first increases in intensity, then shifts to 560 nm. The maximum at 560 nm is indicative of a bis(catechol)iron(III) species, as reported previously.[17, 24, 27] The bathochromic shift continues until the maximum is located at 592 nm, which is still within the quoted range for bis(catechol)iron(III) complexes. There is no evidence of a mono(catechol)iron(III) complex, which has a reported λ_{max} of 680-700 nm.[27-28]

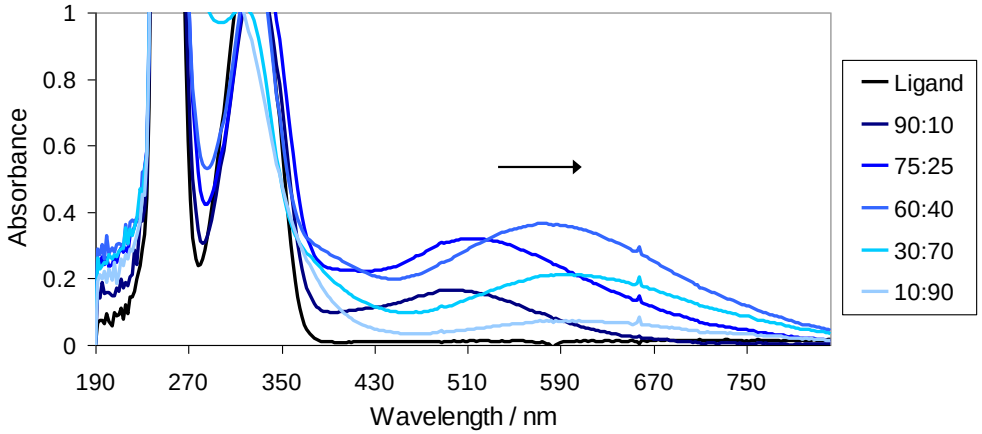


Figure 25 – Selected UV-visible absorbance spectra of H_2-**15** and Fe^{III}(NTA), recorded in the presence of NTA. Spectra recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. Spectra ordered from black (H_2-**15**) to light blue ordered by ligand to metal ratio.

Job Plot Analysis of H_4-4-LICAM and Iron(III) Nitrate in the Presence of NTA

A Job plot investigation with H_4-4-LICAM was carried out under analogous conditions as H_2-**15**. However, this time NTA was always present in equimolar amounts to iron to allow the solutions time to equilibrate without the formation of iron precipitate. When preparing the samples for the UV-visible spectroscopic investigation, it was observed that at close to a 1:1 metal-to-ligand ratio the solutions were initially red but then the colour darkened progressively over time, giving a final colour of purple. To characterise the colour change in more detail, a kinetic experiment was conducted, monitoring the change in absorbance at 498 nm and 552 nm over time. Samples with H_4-4-LICAM to iron(III) ratios of 3:2, 1:1 and 1:2 were monitored over 6 minutes at 5 °C (Figure 26).

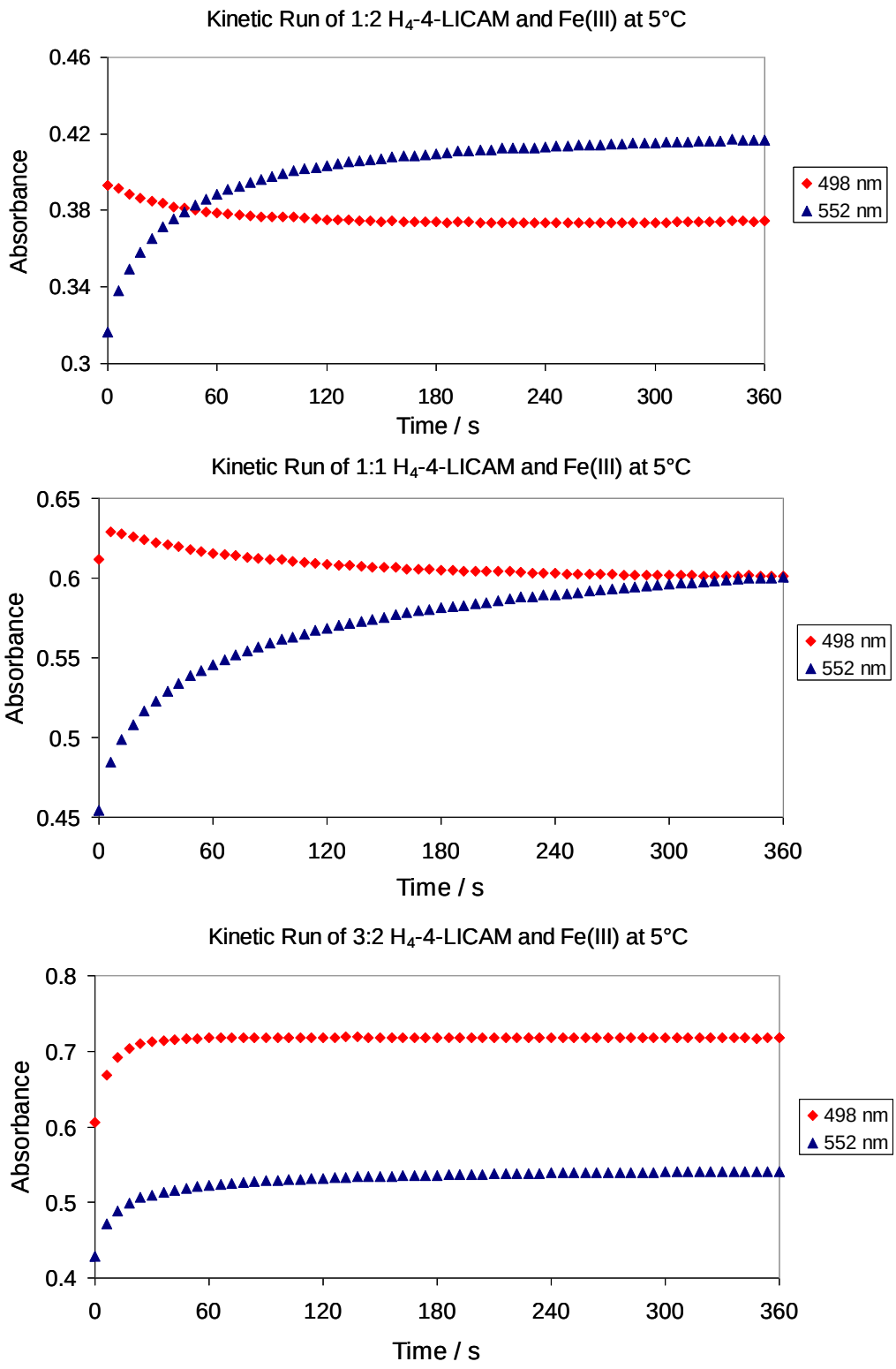


Figure 26 - Kinetic traces recorded for 1:2, 1:1 and 3:2 ratios of H_4-4-LICAM and Fe^{III}(NTA), following the absorbance at 498 nm (red diamonds) and 552 nm (blue triangles) over 6 minutes. Spectra recorded every 6 seconds, solution containing 0.1 M TrisHCl pH 7.5, 5% DMSO at 5 °C. Concentration of H_4-4-LICAM; 1.3×10^{-4} M; 2.0×10^{-4} M; 2.4×10^{-4} M for 1:2; 1:1 and 3:2 samples.

Solutions with ratios of 1:1 and 1:2 showed a change from red to purple over time, due to the increase of absorbance at 552 nm and a decrease in the absorbance at 498 nm, whilst the solution with a ratio of 3:2 showed no colour change. These kinetic runs show that the equilibration of the solution is essentially complete after 6 minutes at 5 °C, yet there are clearly at least two different chemical species within the solution, possessing different λ_{max} absorbance values. Therefore the Job plot samples are required to stand for a minimum of 5 minutes at room temperature in order to ensure that the samples are at equilibrium before UV-visible absorbance spectra are recorded.

The Job plot of H_4-4-LICAM and iron(III) in the presence of NTA was then conducted as previously stated, with the samples now left to reach equilibrium before the absorbance values were recorded. The UV-visible data revealed that there are two species present in solution with characteristic λ_{max} values of 498 nm and 552 nm (Figure 27). Job plot analysis carried out at these two wavelengths indicates two binding ratios. The absorbance at 498 nm indicates a 3:2 binding ratio, whilst the absorbance at 552 nm shows 1:1 binding (Figure 28). This is consistent with the maximum absorbance values of approx 500 nm for a tris(catechol)iron(III) and 550 nm for the bis(catechol)iron(III) complexes.[17, 24, 27, 29] The kinetic data suggests that in samples between the 3:2 and 1:1 ligand-to-metal ratio that a tris(catechol)iron(III) complex forms prior to the bis(catechol)iron(III) complex, before reaching equilibrium (Figure 29).

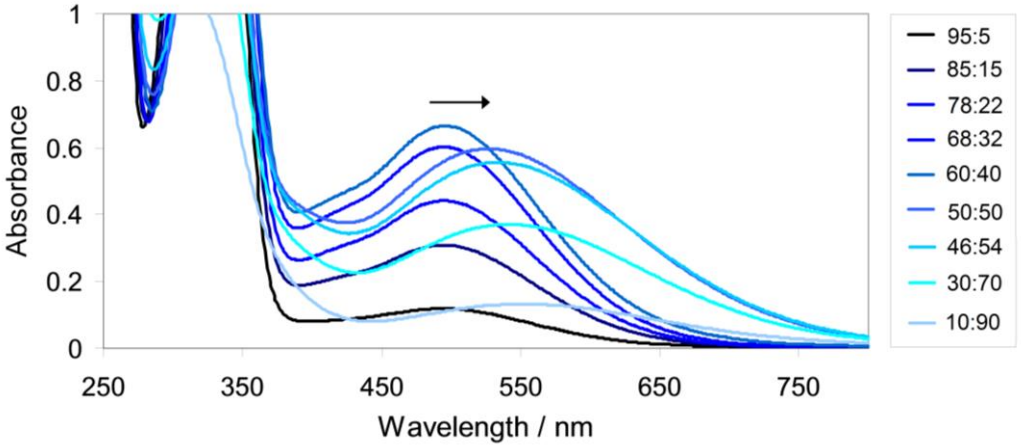


Figure 27 - Selected UV-visible absorbance spectra of H_4-4-LICAM and Fe^{III}(NTA), used for the Job plot analysis. Spectra recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. Spectra ordered from black (H_4-4-LICAM) to light blue ordered by ligand to metal ratio; (95:5, 85:15, 78:22, 68:32, 60:40, 50:50, 46:54, 30:70, 10:90). [M] + [L] = 0.4 mM.

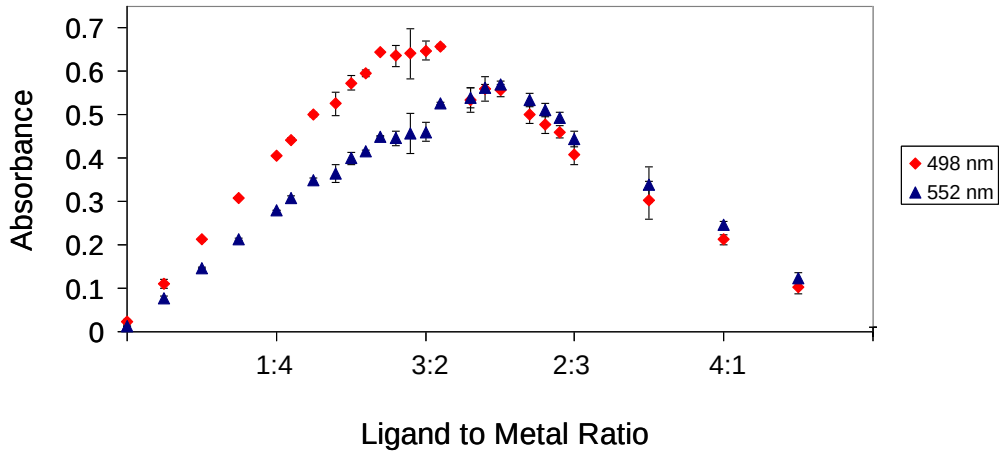


Figure 28 - Job plot for the binding of H_4-4-LICAM and Fe^{III}(NTA), obtained by following the absorbance at both 498 nm (red diamonds) and 552 nm (blue triangles). The absorbance values are averages of two experiments and the error bars indicate the differences between the runs. Plot recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO.

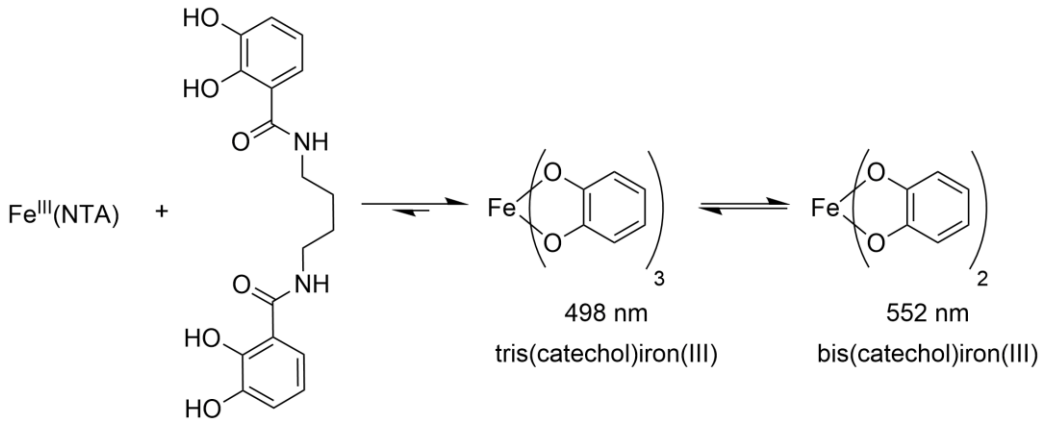


Figure 29 - Purposed schematic mechanism of formation of different (catechol)iron(III) complexes within samples for Job plot analysis. Actual speciation cannot be determined from these experiments.

In conclusion these results suggest that two species can form in solution and this process is dependant on relative concentrations of H_4-4-LICAM and iron(III). One has a ligand to metal ratio of 3:2, with a maximum absorbance for the LMCT band at 498 nm, consistent with a tris(catechol)iron(III) complex; the other has a ligand to metal ratio of 1:1 and a maximum absorbance value for the LMCT band of 552 nm, consistent with a bis(catechol)iron(III) complex. These results only show that the species present have a ligand to metal ratio of 3:2 and 1:1, and the speciation cannot be distinguished between species with the same stoichiometry, such as an *m*(M_2L_3) or an *n*(M_2L_3) structure.

^{1}H NMR Spectroscopic Analysis of H_4-4-LICAM and Gallium(III) Nitrate

To study speciation in solution in more detail, ^{1}H NMR spectroscopy was utilised. Due to the paramagnetic nature of iron(III), gallium(III) was used as a diamagnetic substitute, as it has similar coordination chemistry and physical properties to iron(III).[30] By preparing samples at known ratios of H_4-4-LICAM and $Ga(NO_3)_3$ and by comparing the aromatic region of the ^{1}H NMR spectrum of H_4-4-LICAM, different species were identifiable. To avoid competition from the protons that dissociate upon metal binding, four equivalents of sodium hydroxide with respect to H_4-4-LICAM were added to each sample. Due to the low sensitivity of the NMR technique, higher concentrations of H_4-4-LICAM

were required than for the UV-visible experiments. This required the percentage of d_6-DMSO to be increased, hence the experiments were conducted in a solvent mixture of 5:1 d_6-DMSO:D_2O. In order to ensure equilibrations, the spectra were recorded after 2 days and again after 30 days from sample preparation; the samples were kept at room temperature (Figure 30).

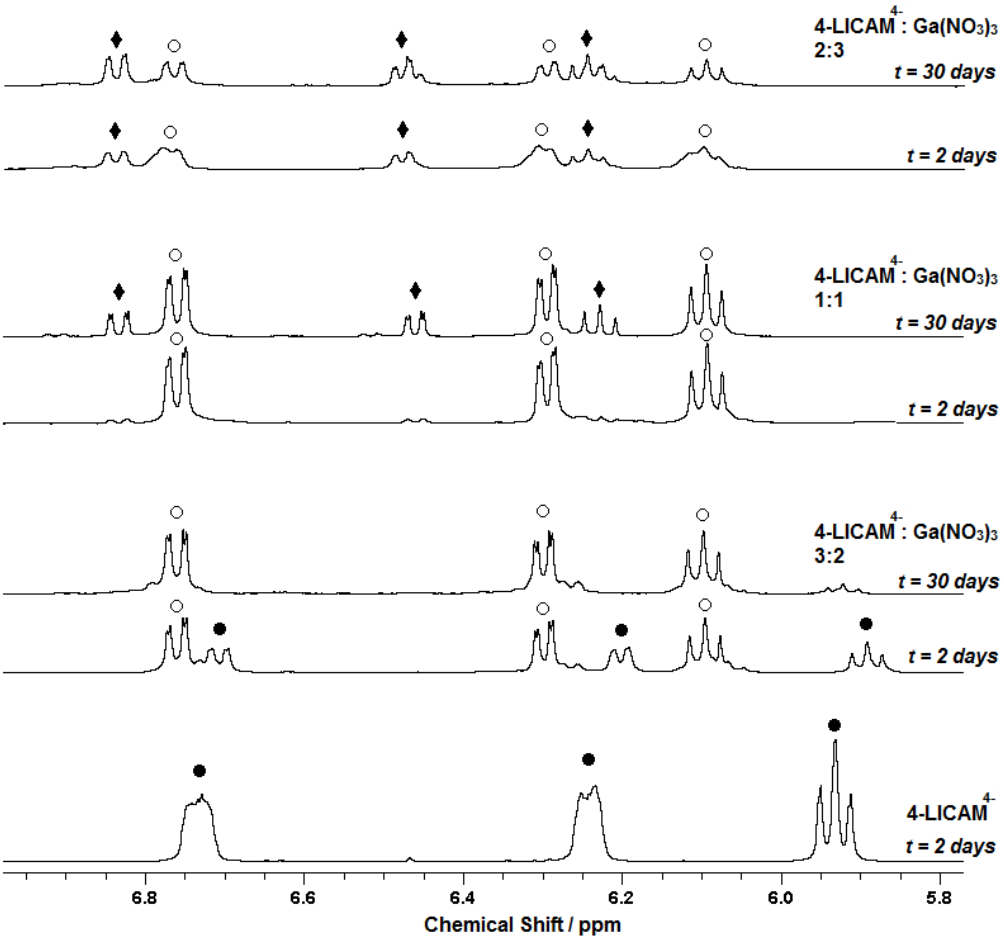


Figure 30 - Aromatic region of the ^{1}H NMR spectra of 4-LICAM^{4-} in a mixture of d_6-DMSO/D_2O with varying amounts of $Ga(NO_3)_3$ present, with 4 equivalents of NaOH with respect to H_4-4-LICAM; (●) 4-LICAM^{4-}; (○) Complex A; (♦) Complex B.

The aromatic region of the ^{1}H NMR spectrum of 4-LICAM^{4-} shows the triplet (5.93 ppm) and the two doublets (6.24 ppm and 6.73 ppm). The doublets have lost the defined splitting pattern and are observed as two broad signals, potentially as a result of delocalisation of charge due to the deprotonated catechol moiety. The ^{1}H NMR spectrum of the solution of H_4-4-LICAM and gallium(III) present in a 3:2 ratio, recorded after two days, show the formation of species A (complex A, ○), with the signals arising from free 4-LICAM^{4-} (●)

still present. The signals for complex A are shifted downfield compared to those of 4-LICAM^{4-}, showing that the aromatic protons are deshielded. After 30 days, the signals due to free 4-LICAM^{4-} are absent, with only the signals due to complex A remaining. In comparison, a solution containing a 1:1 ratio of H_4-4-LICAM and gallium(III) show only signals for complex A after 2 days whereas 30 days later, a second set of signals appear, which is assigned as complex B (♦). The resonances are found further downfield to those of both complex A and free 4-LICAM^{4-}. In the solution containing a 2:3 ratio of H_4-4-LICAM and gallium(III) solution, the signal for complex A is still present, and the signals associated with complex B are also present, with little change occurring over 30 days.

The aromatic triplet signals arising from complexes A and B appear upfield compared to H_4-4-LICAM, and downfield when compared to 4-LICAM^{4-}. This is consistent with a shielding effect on the aromatic protons due to additional electron density on the ring system, when comparing to H_4-4-LICAM.[31-32] Therefore it can be concluded that complex A has a higher negative charge per catechol than complex B, as it is further upfield compared to complex B. This suggests that complex A is the species with 3:2 ligand to metal ratio and complex B is the species with a 1:1 ratio, due to the higher charge density of the 3:2 complex compared to the 1:1 complex (Table 3).

^{1}H NMR shift of aromatic triplet / ppm	Species	Formal Charge per Catechol
5.93	4-LICAM^{4-}	-2
6.09	Complex A, $m(M_2L_3)^{6-}$	-1*
6.22	Complex B, $m(ML)^-$	-0.5*
6.67	H_4-4-LICAM♮	0

Table 3 - Comparison of ^{1}H NMR shifts for the aromatic triplet recorded in 5:1 d_6-DMSO : D_2O. * Charge per catechol calculated assuming complex A has a 3:2 ligand to metal ratio and complex B has a ligand to metal ratio of 1:1, with neutral donors in vacant coordination sites. ♮ H_4-4-LICAM was recorded in d_6-DMSO.

The hypothesis that complexes A and B have a 3:2 and a 1:1 ligand to metal ratio, respectively, is also supported by the change in the relative ratio of complexes A and B over time in both the 3:2 and 1:1 samples. Complex A is shown to slowly equilibrate to form a clean spectrum in the 3:2 sample, suggesting that it possesses a 3:2 stoichiometry.

The 1:1 appears to yield a clean spectrum, however, after allowing the sample to equilibrate, signals arising from complex B also become apparent. The resulting clean spectrum in the 1:1 sample after 2 days suggests that all the 4-LICAM must be bound in the form of complex A as there are no other signals. This suggests there must be 'free' gallium(III) within the sample, however, as the gallium source is $Ga(NO_3)_3$ it possesses no detectable protons to be observed in the ^{1}H NMR spectrum. This 'free' gallium(III) is presumably the driving force for the slow equilibration, dissociating some of the complex A to form complex B. The fact that the 2:3 sample possesses signals for both complex A and complex B, rather than forming solely a 1:1 complex as the UV-visible data would suggest, shows that the equilibrium between complex A and B is clearly altered from that of the UV-visible experiment, likely due to the change in the solvent and pH of the system.

To confirm the assignment of complex A possessing a 3:2 ligand to metal ratio and to determine whether complex B is the 1:1 complex or the 2:2 complex, a diffusion-ordered spectroscopy (DOSY) experiment was carried out (Figure 31). The DOSY NMR experiment is a pseudo 2D technique, which presents chemical shifts against diffusion coefficients of the chemical species in solution. The diffusion coefficients can be determined from the NMR signal intensity decay, in a sequence of 1D spectra recorded with increasing amplitudes of pulsed field gradients. This can be used to map the translational behaviour of the species within solution, which can then be used to calculate the diffusion coefficients of the species within solution.[33-34]

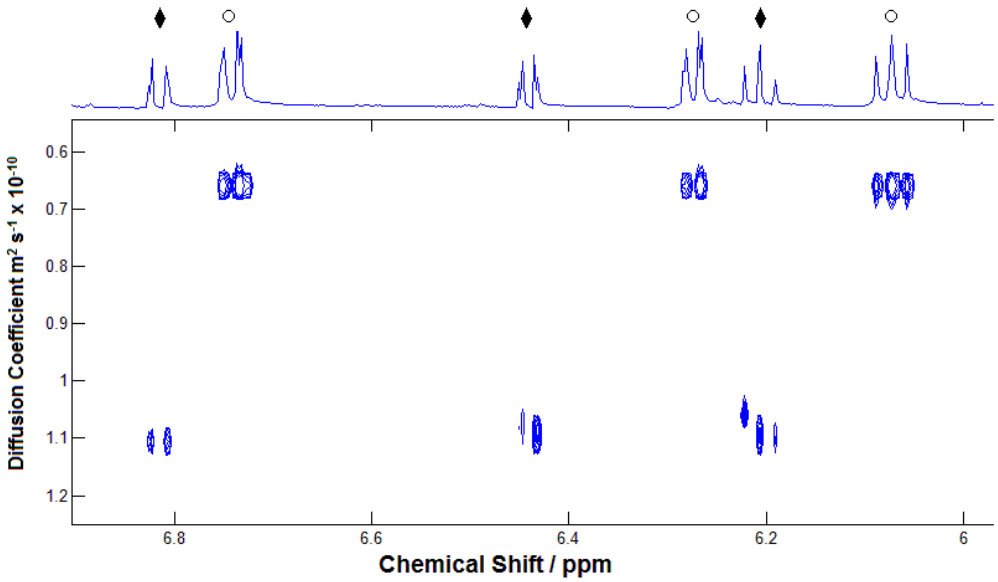


Figure 31 - ^{1}H DOSY NMR spectrum of the 1:1 sample recorded after 1 month of equilibration in 5:1 d_6-DMSO : D_2O. Representations: (○) Complex A; (♦) Complex B.

The DOSY NMR shows that there are only two species that have a defined hydrodynamic radius. The larger set of peaks, which originates from complex A, has a diffusion coefficient of 0.65 x 10^{-10} m^2 s^{-1}, which is approximately half the value of the smaller set of signals with a diffusion coefficient of 1.1 x 10^{-10} m^2 s^{-1}. As the diffusion coefficient is related to the size of the complex via the Stokes-Einstein equation (Equation 1), the diffusion coefficients indicate that complex A is approximately double the size of complex B.[35]

$$D = (1/6\pi\eta r)kT$$

Equation 1 - Stokes-Einstein equation. Key; η - viscosity, r - radius of a spherical particle, k - Boltzmann constant, T - temperature.

The DOSY spectrum was recorded at 295 K, in a DMSO : D_2O 5:1 mixture, which has a viscosity (η) close to 2.6 x 10^{-3} Pa s.[36] Due to the hydroscopic nature of d_6-DMSO, the actual ratio of DMSO to D_2O may differ from the quoted 5:1 ratio and therefore an error of ± 10% on the viscosity value has been calculated and the estimated values for the radii of complexes A and B can be calculated to be 11.6-14.2 Å and 6.8-8.4 Å, respectively.

For comparison, an estimation of the radii of the proposed complexes can be obtained through molecular modelling using the computer program Avogadro (Figure 32).[37] The models were built and energy minimised to get an estimation of molecular size. It is uncertain which ligands are occupying the available coordination sites in the 1:1 complex. DMSO molecules have been modelled in to represent coordinating solvent. Avogadro was then used to calculate the centre of mass for the modelled complexes. The centre of mass can be used to estimate the radius of the complexes by measuring the distance from the centre of mass to the outer atom, which will correspond to the overall radii of the complexes when tumbling in solution. The estimations for a mononuclear 1:1 complex and a 3:2 triple helicate are 6.0 Å and 10.5 Å, respectively. These radii are slightly lower than those calculated from the DOSY experiment, as expected since the modelling does not account for solvation effects, which have an important effect on diffusion coefficients.[38] The 2:2 complex has an estimated radius of 9.2 Å when modelled using Avogadro, which is larger than the DOSY experiment for complex B. This does not include any solvation effects, which would increase the complex radius, suggesting that complex B is a 1:1, rather than a 2:2 complex.

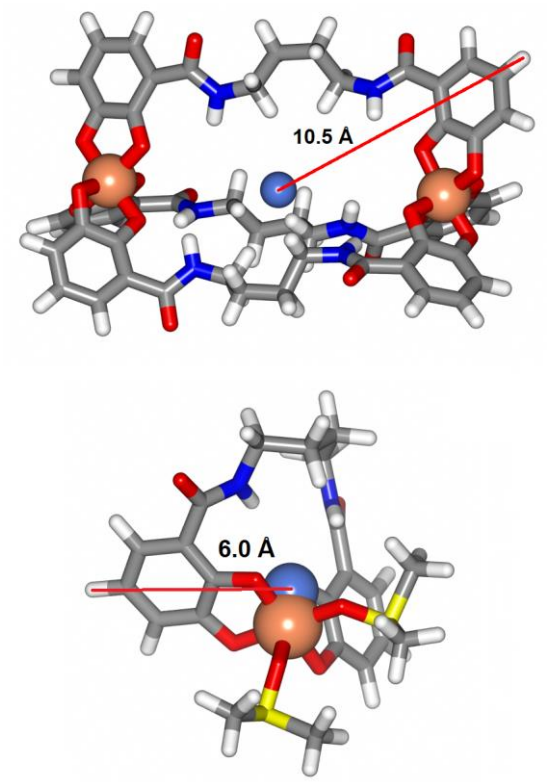


Figure 32 - Models of the proposed 3:2 triple helical $[Fe_2(4\text{-LICAM})_3]^{6-}$ and 1:1 mononuclear $[Fe(4\text{-LICAM})]^-$ complexes, showing the centre of mass calculated by Avogadro and distance to the outer most atom. Key; grey - carbon, blue - nitrogen, red - oxygen, yellow - sulphur, white - hydrogen, coral sphere - iron, light blue - centre of mass.

In conclusion, the Job plot analysis shows that H_4-4-LICAM can adopt two binding stoichiometries with iron(III), depending on the relative concentration of H_4-4-LICAM and iron(III) in solution. A similar dependence on relative concentrations can be seen in the ^{1}H NMR spectra with gallium(III), with two complexes observed under different conditions. These complexes are likely to be a complex with 3:2 ligand to metal ratio and a complex with 1:1 ligand to metal ratio. The DOSY NMR experiment indicates that the 3:2 complex is approximately twice the size of the 1:1 complex, with analysis of the diffusion coefficients suggesting that the approximate size of the complexes fit the assignment of the complexes as a 3:2 triple helicate and a 1:1 monomeric complex.

2.4 CeuE-Complex Interactions

As shown previously a range of different ferric-siderophore complexes can be formed in solution. The aim of this study was to clarify which of these structures interact with the ferric-siderophore transport PBP's, and to characterise these ferric-siderophore-protein interactions. As outlined in Chapter 1, the PBP for the enterobactin uptake system in *C. jejuni* is CeuE.[39-40]

The initial sequence identification of CeuE was taken from Uniprot (code Q0P8Q4).[41] As CeuE is located in the periplasm it has a signal peptide at the start of the sequence in order to transport the protein into the periplasm of the organism. The signal peptide by signal peptide prediction programs; SignalP 4.0[42], Signal-3L[43] and LipoP[44] suggest the signalling peptide is 20 amino acids long. Therefore the CeuE sequence numbering relates to the mature protein, omitting the first 20 amino acids from the full sequence ($Cys_{full}21 \rightarrow Cys_{mature}1$) (Figure 33).

```
      1         10         20         30         40         50
CeuE  MKKSLVFAFFAFFLSLILTACNSNSNENNASSTTKTNTATVKVLPISMSDE
      ▼▼▼▼▼▼▼▼▼▼▼▼▼▼▼▼▼▼▼
              60         70         80         90        100
CeuE  GDSFLVKDSLGENKIPKNPSKVVILDLGILDTFDALKLNDKVVGVPAKNLP

            110        120        130        140        150
CeuE  KYLQQFKNKPSVGGVQQVDFEAINALKPDLIIISGRQSKFYDKLKEIAPTL

           160        170        180        190        200
CeuE  FVGLDNANFLSSFENNVLSVAKLYGLEKEALEKISDIKNEIEKAKSIVDED

          210        220        230        240        250
CeuE  KKALIILTNSNKISAFGPQSRFGIIHDVLGINAVDENIKVGTHGKSINSEF

         260        270        280        290        300
CeuE  ILEKNPDYIFVVDRNVILGNKERAQGILDNALVAKTKAAQNKKIIYLDPEY

        310        320        330
CeuE  WYLASGNGLESLKTMILEIKNAVK
```

Figure 33 - Full sequence of CeuE with the predicted signalling peptide highlighted as the black triangles. Image produced using ESPript.[45]

Expression and Purification

In previous work conducted by Axel Müller, it was shown that the first 23 amino acids of the mature CeuE are disordered and removal of this disordered region *via* limited proteolysis was required before crystals were able to be grown.[39] Therefore a new construct of CeuE was designed to start at Leu24 and was successfully amplified from genomic *C. jejuni* DNA. The primers used to perform PCR are shown in Table 4.

Forward Primer	**Reverse Primer**
5'CCAGGGACCAGCAATGTTGCCTATTAGTATGAGCGATGAG3'	5'GAGGAGAAGGCGCGTTATTATTTTACAGCGTTTTTGATTTC3'

Table 4 - Primers used for gene amplification.

The PCR product was then inserted into the pET-YSBLic system yielding a recombinant plasmid, which was subsequently transferred and grown in the expression strain BL21 *E. coli*. Once expressed, purification was completed over three steps. First, CeuE was purified using nickel affinity column chromatography utilising the cleavable His-tag that was introduced by the pET-YSBLic vector. The His-tag was then removed using C-protease in a 100:1 ratio in a dialysis bag for 48 hours at 4 °C to remove any imidazole in the buffer, as well as removing the His-tag. SDS PAGE analysis revealed that only half of the His-tags had been removed (Figure 34).

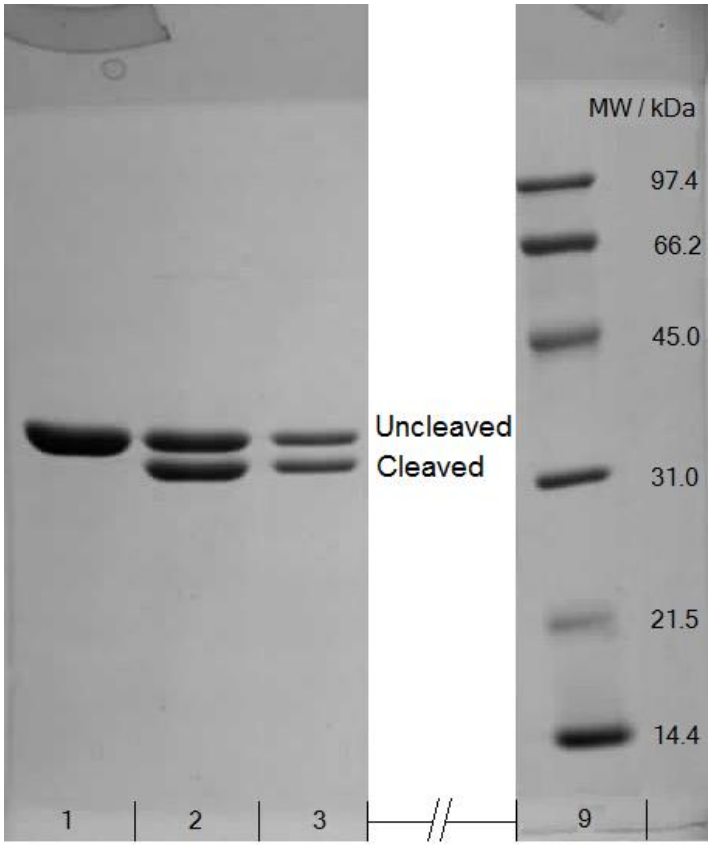


Figure 34 - SDS PAGE analysis of CeuE protein, after 48 hours at 4 °C with C-protease in a 100:1 ratio. Lane 1: Before treatment with C-protease; Lanes 2 and 3: CeuE after treatment of C-protease; Lanes 4-8: Cut out work not relating to this project; Lane 9: Ladder.

The protein was therefore subjected to a second C-protease digestion, this time in a 50:1 ratio for 18 hours at 4 °C. SDS PAGE analysis then confirmed the removal of almost all of the histidine tags. This allowed the second step of purification to be completed, using nickel affinity column chromatography to separate the His-tag free CeuE from other nickel binding proteins. The final step was size exclusion chromatography to yield pure CeuE. Fractions were checked by SDS PAGE analysis and pooled (lanes 6 and 7 in Figure 35). The sample was concentrated to 18 mg/mL, and divided into 50 µL samples and flash frozen in liquid nitrogen and stored at -80 °C until required with the total protein produced being 19.8 mg from one litre of culture (see section 7.3).

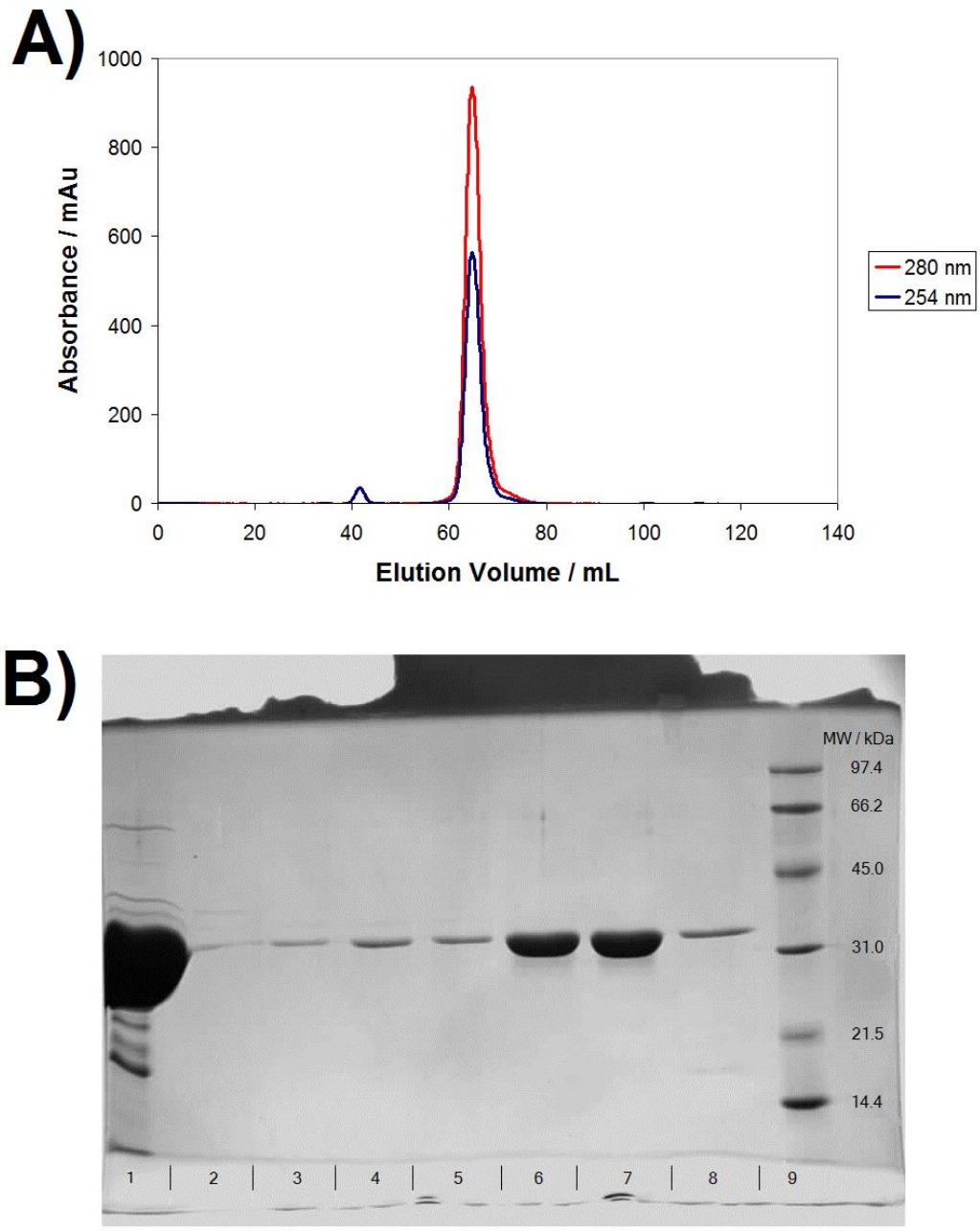


Figure 35 – A) UV-visible trace of final gel-filtration column monitoring at 280 nm and 254 nm. B) Final SDS PAGE analysis of CeuE, before and after the gel filtration step. Lane 1: Before Gel filtration column; Lanes 2-8: Fractions from gel filtration column; Lane 9: Molecular weight ladder. Pure CeuE in lanes 6 and 7 are associated with peak at 65 mL in UV-visible trace.

Crystallisation of apo-CeuE

Initial crystallisation trials were conducted on apo-CeuE, using 96 well, sitting drop, PACT and Hampton commercial crystallisation screens. A suitable crystal was selected from well B2 in the PACT screen (0.1 M MIB buffer, pH 5, 25% (w/v) PEG 1500), and a data set was collected at Diamond synchrotron beamline i03. The structure was solved by means of molecular replacement with MOLREP[46] using one of the CeuE peptide chains from the ferric-MECAM structure as a model (PDB code: 2CHU)[40]. Refinement was completed using REFMAC5[47] and Coot.[48] The crystal was in the space group P1, with three CeuE monomers in the asymmetric unit. The structure is that of a typical PBP with two domains linked by a α-helical backbone (Figure 36).

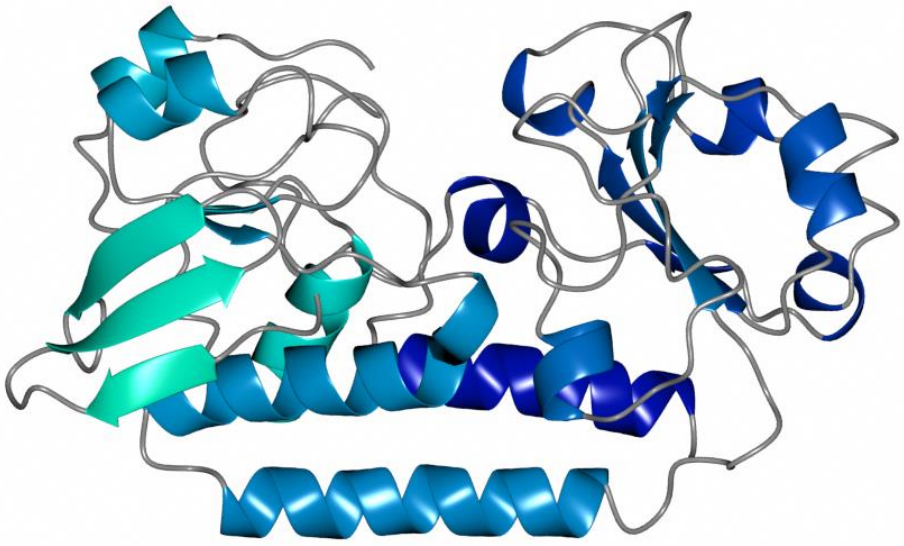


Figure 36 - Ribbon representation of apo-CeuE with colours blended from cyan, N-terminus, to dark blue, C-terminus.

	CeuE	{CeuE[Fe(4-LICAM)]}
Data collection		
Diffraction Source	DLS Beamline i03	DLS Beamline i24
Wavelength / Å	0.976	0.978
Temperature / K	100	100
Crystal Data		
Space Group	P1	$P2_12_12_1$
a, b, c / Å	56.95, 62.74, 67.98	60.42, 66.89, 67.30
α, β, γ / °	82.19, 76.74, 75.96	90, 90, 90
Resolution / Å	65.93 - 1.69	44.96 - 1.61
Unique Reflections	95173	28929
Completeness / %	96.8	99.5
Redundancy	2.4	5.7
R_{merge}	0.04	0.05
Mean [I/σ (I)]	11.1	18.3
Refinement Statistics		
Reflections, working	88200	20945
Reflections, test	4647	1146
Resolution Range / Å	65.01 - 1.71	33.65-1.89
R-factor	0.195	0.261
R_{free}	0.227	0.309
Mean B Factors / $Å^2$	29.4	19.5
RMS deviation from ideal		
Bond length / Å	0.019	0.016
Bond angles / °	2.064	2.011
PDB Code	3ZKW	3ZK3

Table 5- Data collection and refinement statistics of apo-CeuE and {CeuE[Fe(4-LICAM)]} structures.

As discussed in Chapter 1, a common feature of PBPs is the presence of two domains that are connected through either a long α-helix or two or three β-strands, producing a bi-lobal structure. The apo-CeuE model fits this description and has a long α-helical backbone. The cleft between the two lobes of the model is the known binding pocket of CeuE, which is dominated by a central positively charged electrostatic region (Figure 37). The long connective α-helical backbone allows PBP's to move the two domains towards each other for substrate binding.[49] However, superimposition of the apo-CeuE structure with {$CeuE_2[Fe_2MECAM_2]$} (PDB: 2CHU) shows little movement in CeuE upon substrate binding (RMS 0.72), possibly due to the main driving force for substrate binding being electrostatic, rather than coordinative (Figure 38).

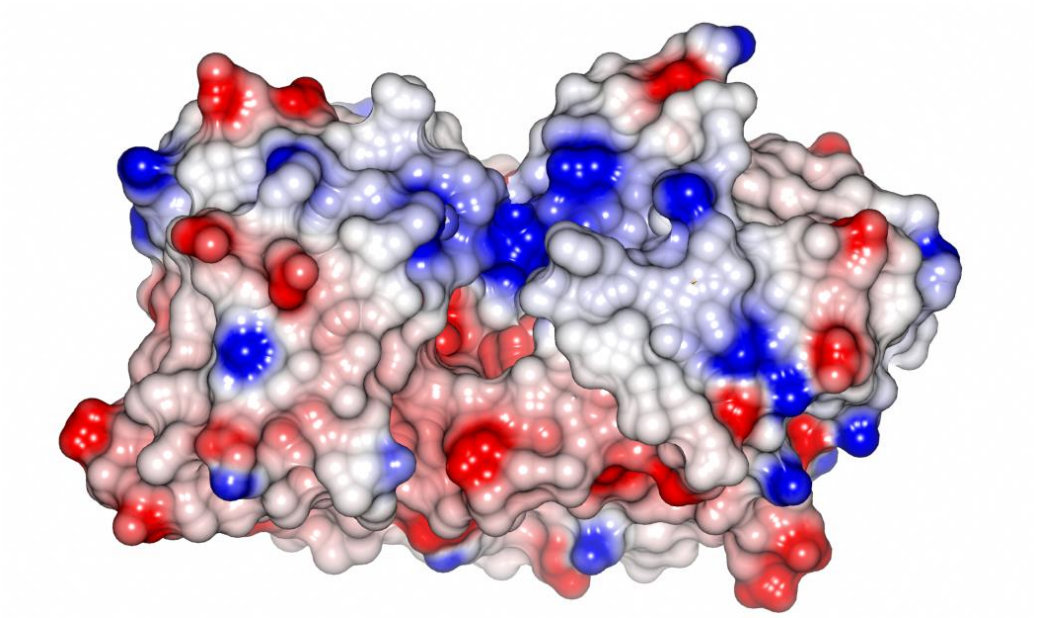


Figure 37 - Surface electrostatic representation of apo-CeuE, negative charge shown in red and positive charge in blue.

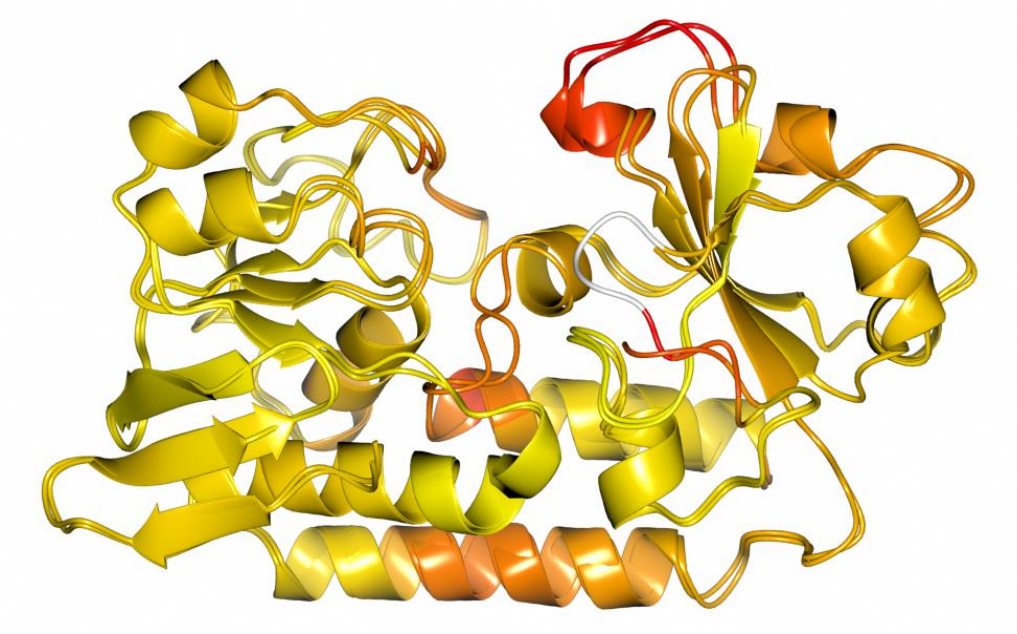


Figure 38 - Ribbon representation of apo-CeuE (PDB code 3ZKW) superimposed with the dimeric {$CeuE_2[Fe_2MECAM_2]$} (PDB code 2CHU) structure.[40] ProSMART analysis shows similar regions in yellow and regions of difference in red (white for unaligned).[50]

Crystallisation of the CeuE, Ferric-4-LICAM Structure

The growth of apo-CeuE crystals and X-ray diffraction analysis to afford a 3D model structure provided the basis to determining how ferric 4-LICAM interacts with CeuE. Co-crystallisation screens were set up in order to obtain single crystals for X-ray diffraction analysis. A solution of ferric 4-LICAM was prepared from H_4-4-LICAM and an aqueous solution of $FeCl_3$ in methanol. The solvent was removed to yield a purple solid, which subsequently was dissolved in DMF to produce a concentrated stock solution of ferric 4-LICAM. This stock solution was mixed with the CeuE protein, and a purple precipitate formed. The precipitate was removed by centrifugation, leaving a clear lightly purple solution, which was used for crystallisation using commercially available PACT and Hampton screens.

Crystals grew in many of the crystallisation wells but few produced single crystals suitable for X-ray diffraction analysis. A crystal found in well G5 of the PACT screen (0.2 M sodium nitrate, 0.1 M Bis Tris propane pH 7.5, 20% (w/v) PEG 3350) diffracted to approximately 2 Å resolution upon in-house testing on a Rigaku MicroMax 007HF generator with an MAR 345 imaging plate detector. The crystal was sent to the Diamond light source (DLS) synchrotron beamline i24 for full data collection. The structure was solved as described for apo-CeuE (Figure 39). The crystal was in the space group $P2_12_12_1$, with one CeuE monomer in the asymmetric unit, however, the crystal was found to be twinned.

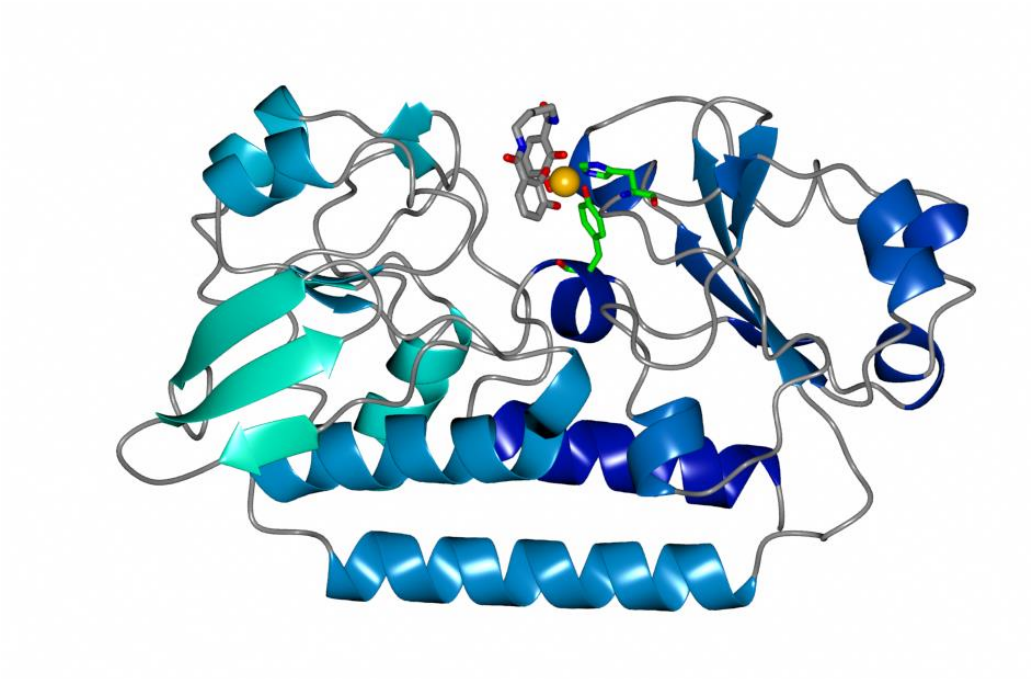


Figure 39 - Ribbon representation of {CeuE[Fe(4-LICAM)]} with colours blended from cyan, N-terminus, to dark blue, C-terminus. His227, Tyr288 and Fe-4-LICAM shown as cylinders. Key: binding residues, green - carbon, blue - nitrogen, red - oxygen; Fe-4-LICAM, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.

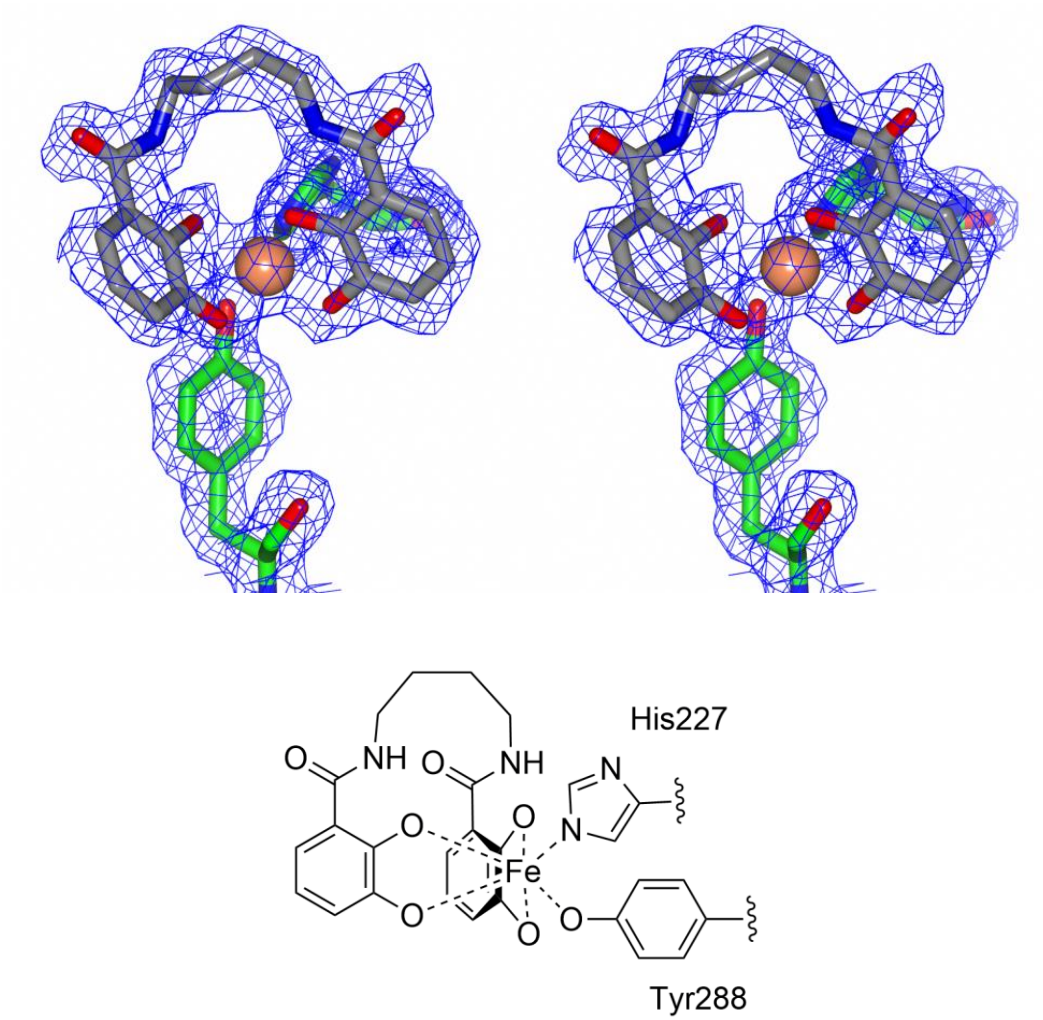


Figure 40 - Stereoview of the iron chelating ligands in the {CeuE[Fe(4-LICAM)]} with electron density map shown as $2F_{obs}$-F_{calc}, contoured at 1σ and chemical representation showing key interactions of Fe-4-LICAM. Key: binding residues, green - carbon, blue - nitrogen, red - oxygen; Fe-4-LICAM, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.

The {CeuE[Fe-4-LICAM]} complex is the first to offer structural insights into the binding of a tetradentate siderophore to a PBP. The structure shows that ferric-4-LICAM has a 1:1 metal to ligand stoichiometry and binds to a single CeuE monomer. This is in a similar fashion to the binding of ferric-bacillibactin, ferric-enterobactin and ferric-MECAM to FeuA from *B. subtilis*.[51-52]

The remaining two free coordination sites of the iron centre are occupied by the side chains of two amino acid residues, His227 and Tyr288. The iron is bound in a distorted octahedral geometry (Figure 41). The bond lengths are shown in Table 6 and range from 1.9 Å to 2.3 Å, consistent with previous known catechol iron bond lengths both in the absence of protein,[9] and when bound to their respective proteins.[40, 51-53] The iron-tyrosine and the iron-histidine bond lengths are 1.9 Å and 2.2 Å, respectively, which are typical for iron-complexes[54] and match those in the structure of the ion transporter cFbpA from *C. jejuni*, in which Fe(III) is bound to four tyrosine and one histidine residues[55] and also FbpA from *H. influenzae* in which Fe(III) is bound to two tyrosine and one histidine residues (PDB codes: 1Y4T, 3OD7).[56]

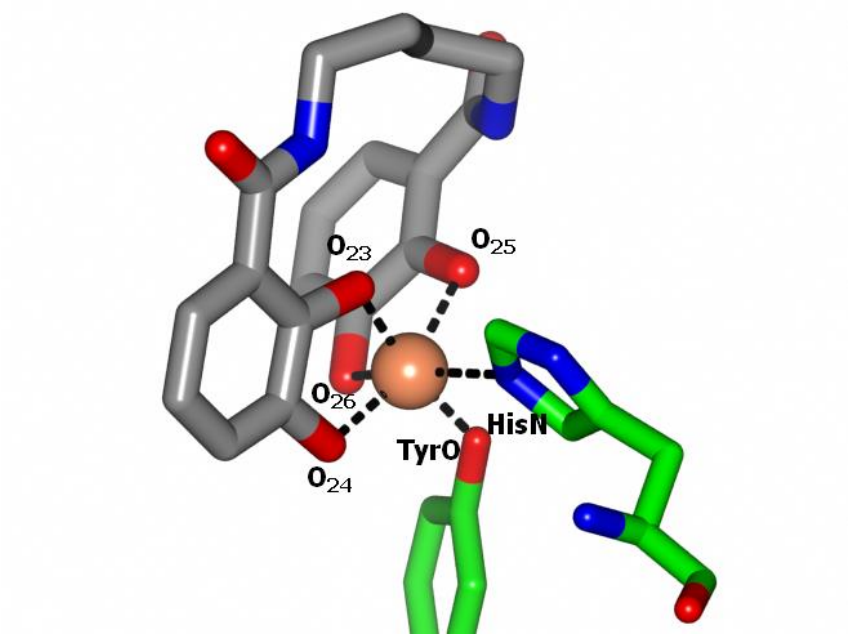


Figure 41 - Coordination of [Fe(4-LICAM)] in {CeuE[Fe(4-LICAM)]}. Key; CeuE residues, green - carbon, blue - nitrogen, red - oxygen; 4-LICAM, grey - carbon, blue - nitrogen, red - oxygen; Fe - Coral sphere.

Atom	Bond Length / Å
$CatO_{23}$ - Fe	1.9
$CatO_{24}$ - Fe	2.0
$CatO_{25}$ - Fe	2.2
$CatO_{26}$ - Fe	2.3
TyrO - Fe	1.9
HisN - Fe	2.2

Table 6 - Bond distances between coordinating atoms and the iron centre.

The bond angles in the {CeuE[Fe(4-LICAM)]} crystal, Table 7, deviate from a perfectly octahedral geometry, consistent with other ferric-siderophores bound to their respective proteins.[40, 52]

Bond	Bond Angle / °	Deviation from Ideal / °
$CatO_{23}$-Fe-$CatO_{24}$	113	23
$CatO_{23}$-Fe-$CatO_{25}$	54	-36
$CatO_{23}$-Fe-$CatO_{26}$	104	14
$CatO_{23}$-Fe-HisN	92	2
$CatO_{24}$-Fe-$CatO_{26}$	93	3
$CatO_{24}$-Fe-HisN	101	11
$CatO_{25}$-Fe-$CatO_{26}$	77	-13
$CatO_{25}$-Fe-HisN	94	4
TyrO-Fe-$CatO_{24}$	105	15
TyrO-Fe-$CatO_{25}$	88	-2
TyrO-Fe-$CatO_{26}$	77	-13
TyrO-Fe-HisN	84	-6
$CatO_{23}$-Fe-TyrO	170	10
$CatO_{24}$-Fe-$CatO_{25}$	161	-19
$CatO_{26}$-Fe-HisN	159	-21

Table 7 - A comprehensive list of bond angles between coordinating atoms to the iron centre and their deviation from the ideal octahedral geometry.

Superimposition of the apo-CeuE model onto {CeuE[Fe(4-LICAM)]}, shows that there is minimal structural change upon ligand binding (RMS 0.88), (Figure 42). The area of the largest structural change is around His227, which is the residue that binds to the iron centre. His227 is located on a loop, which is not surrounded by major α-helices or β-sheets.

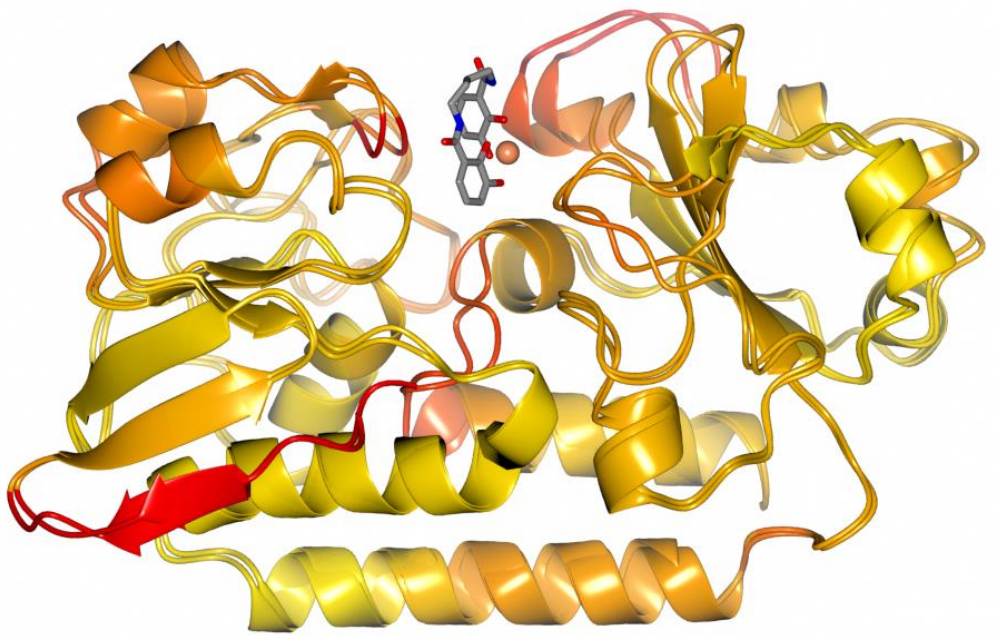


Figure 42 - Ribbon representation of apo-CeuE (PDB code 3ZKW) superimposed with the dimeric {CeuE[Fe(4-LICAM)]} (PDB code 3ZK3) structure. ProSMART analysis shows similar regions in yellow and regions of difference in red (white for unaligned).[50] 4-LICAM shown as cylinders- Grey, carbon - blue, nitrogen - red, oxygen; Fe - coral.

In the {$CeuE_2[Fe_2MECAM_2]$} structure the His227 loop is also disordered and was therefore excluded from the final model. Unlike the histidine residue, which seems to be flexible, the tyrosine residues of the binding pockets of the two structures are found in the same location. The main difference between the two is that in {$CeuE_2[Fe_2MECAM_2]$}, Tyr288 donates a hydrogen bond to the complex, while in {CeuE[Fe(4-LICAM)]}, it coordinates directly to the iron centre.

Sequence alignment of CeuE with the related ferric catecholate siderophore uptake proteins; YclQ, VctP and FetB, using T-coffee[57] shows that the coordinating histidine and tyrosine residues are conserved among these proteins (Figure 43).

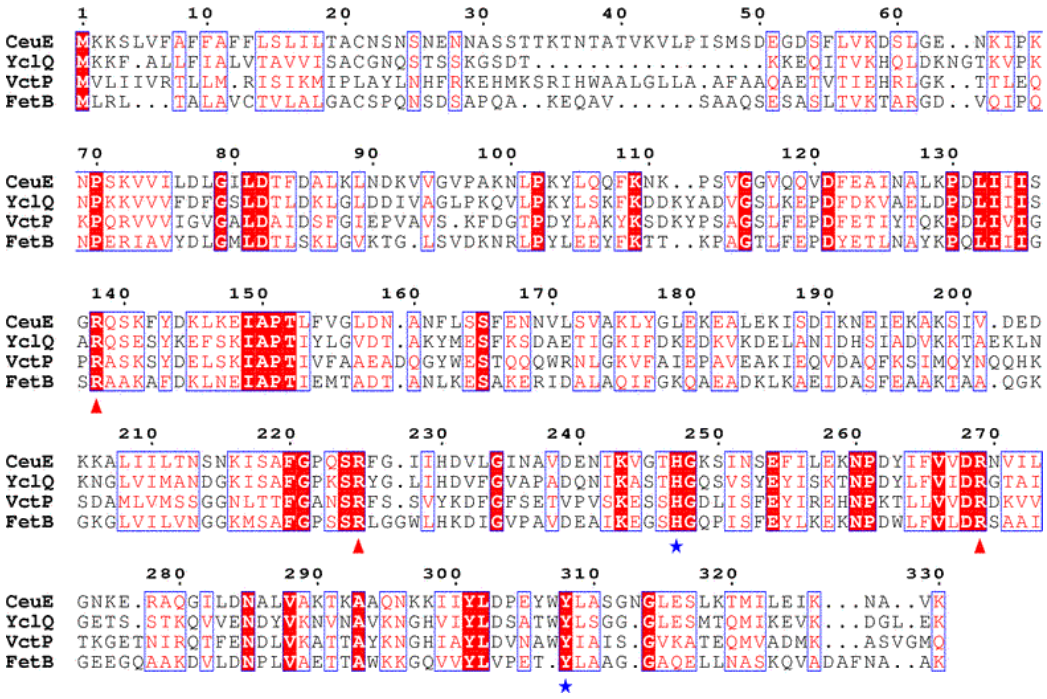


Figure 43 - Sequence alignment of selected PBPs: CeuE (*C. jejuni*), YclQ (*B. subtilis*), VctP (*V. cholerae*) and FetB (*N. meningitidis*), performed using T-coffee.[57] Red block colours show fully conserved residues; red triangles show the position of basic triad of arginine residues; blue stars indicate position of coordinating histidine and tyrosine residues. Image produced using ESPript.[45]

The full CeuE sequence can also be queried using a BLAST+[58] search, which results in a large number of homologous proteins from a wide range of bacteria (Figure 44). Both alignments show the histidine and tyrosine residues and the basic triad of arginine residues fully conserved. The glutamic acid residues proposed for docking onto the inner membrane protein complex (E123 and E254), as outlined in Chapter 1, are present in all but one sequence (*B. subtilis*, YclQ). In YclQ, the glutamic acid (E123) has mutated to an aspartic acid (D123), which presumably should be able to complete a similar role to the glutamic acid.

```
                          1                                        10         20
C.jejuni                  M.............KKSL................VFAFFA.FFLSLILTA
B.sonorensis              MYLFFIKKDEVITLKKLSL................CFMAI..IAVFVIAA
E.faecium                 M.............KKLV................VGILIL.A..SFGLAA
S.thermophilum            M.............KRKVS................LVSMLL.A.LALAVVG
W.virosa                  MK............KLIH................LSILT...TLLLVGA
E.faecalis                M.............KKKF................LAMMAVSMMGLLMLSA
B.subtilis                M.............KKFAL................LFIAL..VTAVVISA
B.cecembensis             M.............KKF.A................FITMLM.A.FILFLGA
A.jilinensis              M.............KKI.M................LFMLLV.M.LAVFAVA
W.chitiniclastica         M.............KSYKK................AALAS..AFALMLAA
C.botulinum               M.............KNISK................KTIII..VAAIIVGL
M.odoratus                M.............SKFMI................SCLAA...SVFLFVS
C.lentocellum             M.............SKKL................LLTLMT..AAAITITG
A.xylanus                 M.............KKL.S................LFIFML.I.FASFIVA
B.cellulosilyticus        M.............KKSL................LLVILA.FSMMMLLAA
L.buccalis                M.............IKKTG...............KILVLIA.L.SIIFIVG
E.brevis                  MNIT.........EKKFVIILIYNTLIIYMKNWIKISLLT...GIAFLSS
C.phytofermentans         M.............KKFSFN..KSIFSGAKKK.VLAGILLGVMATSVLTG
E.casseliflavus           ..............................................
B.clausii                 M.............KKP.F................MWALTA.F.ATVSLAA
L.sphaericus              M.............KKWKL...............LTILMA.M.MLLVLAA
T.bispora                 M.............RSQPLV..........RR.LFALMLV.P.ALLVLAA
A.muciniphila             M.............RTTI................TA.FLA.A.LILP..G
C.botulinum               M.............NKKVW...............TIVGI..LIVFIVG.
M.succiniciproducens      M.............IDKENS.........MKK.TLITLA.....AGLVAA
V.stercoraria             M.............KR...............L.IPALLAA..TVAVALSA
A.baumannii               M.............LF..............KK.TLPVMML....SVALAA

                                  30             40         50
C.jejuni                  CNSNSNEN.NASSTTKTNT....ATVKVLPISMSDEGDSF.....LVK..
B.sonorensis              CGSKDAQ.......................S..DAKKKETV.....AIQ..
E.faecium                 CGNSNNTSQADTKK.................SSTQTETTL.....TIT..
S.thermophilum            CGGGTSTGANNLST.GG.....STSPVQTAEPSPAAQDTI.....VIT..
W.virosa                  CKENEQH......................H...NANIDTI.....EVV..
E.faecalis                CQTNKKTA..DSAT................TETTAKTEV.....TVK..
B.subtilis                CGNQSTS......................SKGSDTKKEQI.....TVK..
B.cecembensis             CGNDNAAETEE.NS.G......STSG.EATEETTEDSKSV.....TIE..
A.jilinensis              CGSDTEDTTGTTDN.GE.....VQEDSNDQEEVDEGPSEI.....EVE..
W.chitiniclastica         CGNEQPKTETVVIETQTET..TTNAPAASAEKAPEAAKTM.....TIK..
C.botulinum               FGIVEATKF....................AGKGSNDNNDKI.....TIT..
M.odoratus                CKETKNE......................EQSSATETAVL.....TVD..
C.lentocellum             CGQTTTNEGTGAVNTEAPV....NTEASATPEEGQEATTV.....KVV..
A.xylanus                 CGDTSKREADNT.................SKEAAATV.....EVV..
B.cellulosilyticus        CGSEDSEG.SESSEPVVEEEQEGQEQQEEVEEETEEPSEL.....TIT..
L.buccalis                CSKSGNEK.....................AGNGKQL.....TIS..
E.brevis                  CNKENKI......................D..TIDAENQI.....EVI..
C.phytofermentans         CAKKQGDTPENNANKGEES..ATTATPAPTPTTAAEETRF.....SLT..
E.casseliflavus           ..............................................
B.clausii                 CGTSSDKNEQTST..GE............GTSTEATEETI.....EVE..
L.sphaericus              CGSKDEAKEGDKGT.ST.......DKPAEEKNEASAAYPITIPGSTIEGR
T.bispora                 CGGERSGSRD.TAA.TP.......QQSAE.........QV.....TVA..
A.muciniphila             LSM...AA.....................ADEAKTI.....TAT..
C.botulinum               .GTLFLKSK....................EKKEDVATSVDM.....KIT..
M.succiniciproducens      FGVVSA.......................QAADI.....GLE..
V.stercoraria             CGGNTEQTANTTAA.AS.......A..PATASNTASAQQI.....QLQ..
A.baumannii               CGNSDTPSANQD.................SHTAATQDTL.....SFN..
```

Continued…

```
                                60        70        80        90        100
C.jejuni                ...DSL..GE.NKIPKNPSKVVILNLGILDTFDALKLNDRVAGVPA.KNL
B.sonorensis            ...HSS..GS.TDVPKNPKKIVVFNFGVLDTLDQLGLGDRVVGLPK.QNI
E.faecium               ...DSN..GDQIEVPNNPEKVVVFDNGSLDTMDALGVGDKVVGAAT.SSL
S.thermophilum          ...HQL..GE.TPVPVNPQKVVVFDFGILDTLDLLGVE..VAGLPQ.GNV
W.virosa                ...HEL..GT.TQVKENPERVVVFDIGSLETLDQLGVK..IVGIPK.DFL
E.faecalis              ...DTN..GQ.LTVPKNPKKVVVFDNGSLDTMDALGVGDRVVGAPT.KNI
B.subtilis              ...HQLDKNG.TKVPKNPKKVVVFDFGSLDTLDKLGLDDIVAGLPK.QVL
B.cecembensis           ...HEL..GE.ATIEGTPEKVVVFDFGILDTLDELGVD..VAGLPQ.ASV
A.jilinensis            ...HEL..GT.TTVPTNPETVVVFDFGTVDTLDYLGVD..VAAVPQ.ANI
W.chitiniclastica       ...HLS..GE.TVVPMNPQRTVVFDFGSLDTLEALGVN..VLGIPK.ANI
C.botulinum             ...HKL..GK.TEVSKNPERVVVFDYGVLDSLDALGIN..VVGVPQ.DGI
M.odoratus              ...HLS..GT.SEVKVNPKNPVVLSYGILDTFDELGIP..VKGVPG.SNL
C.lentocellum           ...HAL..GE.TEAPKNPKNIVVLELGILDALDALEIT..PAGLPLSGSL
A.xylanus               ...HNL..GT.TVVPKNPETVVVFDFASLDILNYFGVE..VAALPK.ASV
B.cellulosilyticus      ...HEL..GD.ITVPVNPEKVVVFDMNALETLDKMGVN..VTGVPQ.ANI
L.buccalis              ...TVK..GD.VKVPENPKKVAVFDYAALDVIGTLGVKP.ELALPA.SNV
E.brevis                ...HKL..DT.TNVVKNPQRVVVFDMGALESLKELGIK..PAGIPK.DHI
C.phytofermentans       ...HPL..GT.VTLEGAAKKVVVFDMGALDTIDALSFDG.ELAVPH.GNI
E.casseliflavus         ..............................MVVFDNSALDTMDALGLGDRIVAAAT.SNL
B.clausii               ...SMN..GEVVKVPVNPEVVVSFDHGITDSIRALGGE..VSGIPKANNI
L.sphaericus            DGKTTF..EE.VKLDKIPEKVVVFDNGFLDTLDALGVN..PTAVVQ.DSL
T.bispora               ...HAQ..GS.TSVPKNPQKVIVFDVGVLATLDELGVK..VAGVPSVENL
A.muciniphila           ...HTL..GT.VELPVNPKRVAVLDYGSLETIGELGVTP.ALALPK.KFL
C.botulinum             ...HKL..GE.ATLKKNPKKVVVFDYGTLDSLDKMGIE..IKGLPK.SNI
M.succiniciproducens    ...TFG..GK.QIVPENPKRVVVLDFAALDTIREIGAKETVVGISK.GRI
V.stercoraria           ...GKL..GT.TLVPANVARIAALDYTTIDTLMALGATDKIIGLPLASNV
A.baumannii             ...SKA..GQ.ISIPANVSRIAVLDTSALNTIHALGADDKVVALPKGTPL

                                110       120       130       140
C.jejuni                PKYLQQFKNK...PSVGGVQQVDFEAINALKPDLIIISGRQSKFYDKLKE
B.sonorensis            PSYLKQYKDK.KYESVGGLKEPDFEKIAELAPDLIIIEGRQAEQYDEFKR
E.faecium               PEYLSSYKK...VESAGGIKEPDLEKINQLQPDLIIISGRQRDFQKDLSQ
S.thermophilum          PPYLAQYTHS.KYANVGTLQEPDFEALSALDPDLIIISGRTSTHYEALSE
W.virosa                PEHLQQYAQDSSIENVGSVKEANYEKVNALNPDLIIISDRLQDAYKELSS
E.faecalis              PAYLKKYQK...VESAGGIKEPDLEKINQLKPDLIIISGRQQDYQEQLKA
B.subtilis              PKYLSKFKDD.KYADVGSLKEPDFDKVAELDPDLIIISARQSESYKEFSK
B.cecembensis           PGYLEQYASD.DYTNVGSLKEPDFEAIHALQPDVIFISGRQAQLYDQFAE
A.jilinensis            PSYLDQFEAD.TYENAGTLFEPDFETLANLDPDLIIISGRTSEAYDDLTD
W.chitiniclastica       PAYLSEYKDP.KYLDFGALKEPDFEAIHAAKPDFMIIAIRQAPLYEQFAE
C.botulinum             PSYLGEYSSD.QYISVGGIKEPNFEKVHEAKPDLIIISGRQEESYDEFSK
M.odoratus              PVYLEKYSAA.EYDNVGGIKEPNAEKVNAADTELIIIAGRTAAMYDEFKK
C.lentocellum           PSYLAHLEDESLYTNVGSLKELDMEKIYEIEPELIIIGGRQADYYEELSK
A.xylanus               PSYLSQYEAD.QYENAGTLFEPDFELLANLNPDLIIISGRTQESYDDLSE
B.cellulosilyticus      PAYLSQYEDA.SYENVGTLFEIDYEKLAEINPDLIITGGRQAEQYDDLSE
L.buccalis              PEGLKQYAQG..AVDAGDVKEPNLEKISEFKPDLIIISGRQTKFYDELSK
E.brevis                PEHLKVLGNDPSIEDVGSLVEPNLERINALNPDLIIISDRQNKFYDELSK
C.phytofermentans       PTYLSKYEES..TVNAGGLFEPDMEAIFTFEPDVIIIGGRQVDFYDELNE
E.casseliflavus         PDYLSAYAD...VESAGGIKEPDLEKINQLQPDLIIISGRQRDFQEELSA
B.clausii               PDYLSEFESD.DYEDVGSLFEPNFELIHSMEPDVIFISGRASDNYEELSE
L.sphaericus            PSYLSKYKDS.TYVNAGTLFEPDYEKLSEINPDIIFISGRASAAYAELSK
T.bispora               PDNLAKYAGD.EYPKVGSLFEPDYEKVNSLSPDLIIVAARSAAAYGELTK
A.muciniphila           PPYLSKFSGE.QYTDLGTVKEFNIETINAFKPDLIIISGRQQDYYQKLSS
C.botulinum             PSYLSKYKDD.KYIDVGTLFEPNFEKLNEIKPDVIFISARQSKAYQELNK
M.succiniciproducens    PQYLAEFDTD.KYANAGTMPEPAFEKINEMSPDLIIASARQKKVLARLKE
V.stercoraria           PSSLDSFKNG.KYIDFGAVNNINLETLASNQPELIISADRLEKQNDSFKQ
A.baumannii             PKVLQNFSEG.KYLDVGTVKEPNLEKIAASQPELILMSGRMENLVGQIKQ
                                                            ▲
```

Continued…

```
                        150       160       170       180       190
C.jejuni                IAPTLFVGLDNANFLSSFENNVLSVAKLYGLEKEALEKISDIKNEIEKAK
B.sonorensis            IAPTINIELDDQNYMKSFKENTESIGKIFGKEDAVKDALAKIDAQVEKTN
E.faecium               IAPTIFLSLDAKNPWESFQQNVTALGEIFGKQEEAKTQLEELSSAIDQTK
S.thermophilum          LAPTIYLAVDTSNYMESFRKNMETIGAIFQKEEQVAAELKEIEDKIAEVR
W.virosa                IAPTIYMQVDYKDYLNSFEKNTLTLGKIFGKEKETEEKVKQIKDKIQTEK
E.faecalis              IAPTIYLAVDAKNPWASTKQNIETLGTIFDKEEVAKEKITGLEKEIADVK
B.subtilis              IAPTIYLGVDTAKYMESFKSDAETIGKIFDKEDKVKDELANIDHSIADVK
B.cecembensis           IAPTVYIPIDTTNYVESFKENMNTVAEIFGKEDEMNTELEDIDSRIAAIQ
A.jilinensis            LAPTIYMGVDESDYMSSFADNVSLLGEIFNKQDESQEAIEQVNADLEDLV
W.chitiniclastica       IAPTINYNVDDQNYMASLTENATTLGKLFGKEDEAKAQLAELNQEIAKVK
C.botulinum             IAPTIYLGIENESYIESFTNNMNTLGQIFDKEDEINKTLQGINSKIADLN
M.odoratus              IAPTINLDVDAKDYMNSFKNNQRIIGKLFGKEEQVEQELKDIDARIAKIK
C.lentocellum           IAPTVNLAVDNADYLTSFETNMNYLGEIFDKEEEVTAKVKEIKDAVAEVS
A.xylanus               LAPTIFLGIDTADYLGSFKTNVSILAEIFDKEEVATEAIAQIESDIADLV
B.cellulosilyticus      LAPVLHMSTDYDNFYESFQSNMTTLGQVFNQEAFIAKELQAIDDAIAALN
L.buccalis              IAPTIFLETDTKNFISSSNENAKKIASIFGKEKEADEQIAKINSEINDIK
E.brevis                IAPTISMQVDPLDYMNSFEKNTLTLGKIFDKEEEMKKEVEAIKAEVKTDK
C.phytofermentans       IAPTIYVEINADTYMEDFKRNTDNIAKAIGKENEAKTAIDGIMNKVEEVK
E.casseliflavus         IAPTVFLSVDNSDTWASIEANITTIGQIFGKEEEAQEQLASLQGQIDTLK
B.clausii               IAPTVYMAVDNENFMESFETNMRTLGDIFEAQDEVEEQLAGIHETIDNVK
L.sphaericus            IAPTVYIGVDNKNFLESFKANTELAGKIFGKEKEAADAIAAYEAKVEEVK
T.bispora               IARTVDLSIDFSDFLGSFRQRMEAIGAIFGKEAEVKQRLDAIDAAVQETK
A.muciniphila           IAPVYLVDTLAKDQLGEAKKNIRLLGDVFGVPEKAAEAVKNIDEAVTRTK
C.botulinum             IAPTIHFNTENGKYMESVKSNLEKLGKVFDKEDFVRDEIKKLDDSVKDIN
M.succiniciproducens    IAPVFYMENDYENYYPSFEQNLLALGKIFNKESAVKEKLAQLDNRMTALA
V.stercoraria           IAPTYQYHIDTANYWPSFKEQTLNIAKVVGKTAEAEQKLQTLETQVSEIA
A.baumannii             IAPTYFVDVDYNDQFNSFKQQTLNIAEMVGKKTEAEQQLQTLEKDIAALK

                        200       210       220       230       240
C.jejuni                SIVD.EDKKALIILTNSNKISAFGPQSRFGIIHDVLGINAVDENIKVGTH
B.sonorensis            KLAKEKGGKGLVVLTNDKHISAYGKGSRFGLVHDVLGVAPADQNIKASLH
E.faecium               KKAEATDKKALVTLVNEGQLSAYGSGSRFGFIHDLFGFEQADDQIEASTH
S.thermophilum          TKVAGGSEKALILLSNDKSISAYGPGSRFALIHDVLGFPAVDPTIEVSTH
W.virosa                DILSKDDKKALIVLYNNGKFSAYGKGSRFGFIHDVLGVQQATENLEVSTH
E.faecalis              KQAEASANNALVVLVNEGQLSAYGKGSRFGLIHDTFGFKAADDKIEASTH
B.subtilis              KTAEKLNKNGLVIMANDGKISAFGPKSRYGLIHDVFGVAPADQNIKASTH
B.cecembensis           EKTSASEEKALIVLGTQSKVSAYGPASRFGLIHDVFGVKPADEGIESSTH
A.jilinensis            GDV.PEDEKGLIVLTNDGALNAYGPGSRFGIIHDDFGIEPVALDIEAATH
W.chitiniclastica       AKTENMPEKALIILANDGKISAYGPGSRFGFIHDVLGFKPADETIEVSRH
C.botulinum             KKTSESQKNGLIILANDGSFSVYGKESRFGIIHQNFGVAPVDNNVESSTH
M.odoratus              SITENSDKKGLIVLANEGRMSSYGKGSRFGIIHDVFGLKPADENIEVATH
C.lentocellum           AKAAEKDVNALIALANGDAFSVYGKGSRFGIIHNEFGIKPVDETIEVSTH
A.xylanus               DTV.PTDKTGLILLTSDGSLSAYGPGSRFGVIHDEFNVTPAVTDVEAATH
B.cellulosilyticus      EKVS.SEETALVVLANSGNVNAYGTGSRFGIIHEKFGFTPVDESIEVTTH
L.buccalis              NIAKNYNKKALVTLTNDGKISAFGANSRFGFIFTDLGLKNIDDSIKASTH
E.brevis                ENIDKIDKKGLIILYNNGKFSAYGKNSRFGFIHDVMGIKPVSENLEVAVH
C.phytofermentans       QITNSSDKKGLIIMTNDGSISAYGLGSRFGLIHDVLEVKPADETIEVSKH
E.casseliflavus         EQAQNADEKALVVMVNEGNLSAFGPGSRYSILYDTFGFPAVDEELEASTH
B.clausii               ERADNTDKNALILSADEGSASAFGAGSRFGIIHDVLGIKTAD.DIEAENH
L.sphaericus            GKATASEEKALIVLGSEGALSAYGSGSRFGVIHDVFGVKAADENVKVGTH
T.bispora               AK..AGDRTGLIVLTTGGKMSAFGPGSRFGLIHDVIGVKPADPNLKVDTH
A.muciniphila           DKAQASGKKALVLLTNDGKVSAYGSGSRFGLVHDALGVAQADKNIKVGVH
C.botulinum             KKASEGGKKALVILANDGALSAYGKGSRFGIIHEELGFPLSDEHIDTAVH
M.succiniciproducens    KLT..AGKSALVTIVNESRISAFGDKSRYALVYQKFGFTPIDKNLSSSTH
V.stercoraria           AKT..QGKTALIVLVNNDKLMAFGAGSRFGLIHEKMGFTPVDPSIKVGTH
A.baumannii             AKT..KGKTALVVLVNNSKIAAYGPGSRFGNIHDLYGFTPADPSIKVGLH
                                                  ▲                      ★
```

Continued…

```
                                250       260       270       280       290
C.jejuni                        GKSINSEFILEKNPDYIFVIDRNIIVG..NKERAQGILDNALVAKTKAAQ
B.sonorensis                    GQSVSSEYISEKNPDYLFVIDRSAAIG..EKGSAKEVIENEIVKSTNAYK
E.faecium                       GQSVSYEYVLEKNPDILFVVDRTKAIG..GDDSKDDISANELVAQTNAGK
S.thermophilum                  GQQVSFEYVKEKNPDYIFVIDRAAVVSG.GGQPAKDTFDNDLVKQTNAYK
W.virosa                        GQPISNEFIEHKNPDYLFIVDRSAMIN..RVETNKKEIENVLIQRTNAYK
E.faecalis                      GQSVSYEYVLEKNPGILFVVDRTKAIG..GDDSKDNVAANELIQKTDAGK
B.subtilis                      GQSVSYEYISKTNPDYLFVIDRGTAIG..ETSSTKQVVENDYVKNVNAVK
B.cecembensis                   GQNITFEYILDTNPDILYVIDRDAAIG..EDSSVKESFENDLVQKTNAYK
A.jilinensis                    GQSISFEYVAEMDPDYLFVMDRNAVTG..GDYTASTTLDNALINQTTAAQ
W.chitiniclastica               GQSVSFEYVMNLDPDYIFVVDRTAVINE.GGTTAKQTVENDLVKNTKAFK
C.botulinum                     GQKASFEYLVEKNPDYLFVVDRSAVA.G.GQVSAQALLDNDLVKSTNAYK
M.odoratus                      GQVISNELIKELNPDYIFVIDRGAAIK..RATLSKEEFANALVQQTNAYK
C.lentocellum                   GQNASFEYILEQNPDYLFVVDRSAVVA..GEGSAPALFDNELIKKTAASE
A.xylanus                       GQNISFEFVADVNPDYLFVLDRNAVTG..GEYDAASTLDNELMNQTTAVQ
B.cellulosilyticus              GMNISFEYIVDKDPDYIFVVDRDLVVT..GEPAAQDTIENELVQNTKAYQ
L.buccalis                      GQEINYEYIAEKNPDIIFYVDRTKIVG..GSKAGADVLKNELVANTNAGK
E.brevis                        GQPVSNEFIESVNPDYLFVVDRSAMMN..KVQTNKNDIENVLIKQTNAYK
C.phytofermentans               GMDASYEYIAKVDPDILFVVDRTVIAG..GSTTAEKTLDNELVLGTKAAK
E.casseliflavus                 GQSISYEYLLEKNPDIIFVIDRTKAIG..GNESESEFVANEIVQETNAGQ
B.clausii                       GETVSFEYIADVDPDYIFMIDRGASIG..NEATASQVIDNELVARTKAAQ
L.sphaericus                    GDNVSFEYVRDTNPDILFVVDRDAVVNENGESGTKGAIENEIVSATNAVK
T.bispora                       GEAITAEFIAKTDPDLLYVIDRDAAIGEE.GQPAQKVLDNALVNGTKAAK
A.muciniphila                   GQQVTYEYIAMLNPDIIFAVDRSVAIG..RSAKNPNVLNNVLVNKTKAAK
C.botulinum                     GQKISFEYVVEKNPDYIFVVDRGAVVQG.GHKSVNKILENDLIKTTKAYK
M.succiniciproducens            GNSVGFEYIAEKNPDYLLVVDRTAAITDK.ANNAQTVLDNALIKPTKAAK
V.stercoraria                   GQSISYEFITEKNPDYLFVLDRAASVTGK.TGGAEAAMNNDIIKQSKAFK
A.baumannii                     GMAVSHEFIAQKNPDYLFVIDRGAAITEN.QQGAKQVLNNQIINQSKAAK

                                  300       310       320       330
C.jejuni                        NKKIIYLDPEYWYLASGNGLESLKTMILEIKNAV...K
B.sonorensis                    KGHIIYLDPGYWYLS.GGGLESVSEMVKEIENGL...K
E.faecium                       NQQIISLEPDVWYLS.GGGLESMKLMIEDVNQAF...K
S.thermophilum                  NGKIIYLDPNYWYLS.GGGLVSVSAMIDEVAAAI...Q
W.virosa                        NGKIIYLNPQVWYVS.GGGLTSTQMMVDEVKNAL...K
E.faecalis                      NDKVIMLQPDVWYLS.GGGLESMHLMIEDVKKGL...E
B.subtilis                      NGHVIYLDSATWYLS.GGGLESMTQMIKEVKDGL..EK
B.cecembensis                   NEKIHYLNGEYWYLS.GGGLQSMKEMISEIEGS....L
A.jilinensis                    EDQIVYLTPDYWYLS.SGGVESVTQMIAEIKQGI...E
W.chitiniclastica               NGKIIYLDPNVWYLS.GGGLKSTKVMIEDASKAL..MQ
C.botulinum                     NNNIVYLDANVWYIA.TGGINTTLKMIEEVDNA....F
M.odoratus                      NDKIIFLNPETWYLS.GGGLKSIKMMIDEVENAV..TK
C.lentocellum                   NDRIIYLDPTVWYTA.GGGFTSTMTMVNEIAAAI..EK
A.xylanus                       EEQVTYLTPDYWYLS.SGGIESVTQMIAEIKATF...K
B.cellulosilyticus              NGNIHYLNPDYWYIA.GGGIVSVSEMVKAIEEAL...Q
L.buccalis                      TNKVVALDAENWYLV.SGGLATIQKQLEEIKTAI..TK
E.brevis                        NGKIIYLNPQVWYIS.GGGITSTKQMIEDIKKAI...Q
C.phytofermentans               NNVIISLDSECWYLS.SGGINSMQVMVKEVEGAF..NK
E.casseliflavus                 NDQVIELDAAVWYLA.GSGLESIQLMMDDVKPAV...E
B.clausii                       NDDIYVLDGEVWYLS.GSGLESVKLIVDEIDSMF.ETE
L.sphaericus                    NGKVFYLDPEVWYLS.GGGLQSETLKVEDVLKAF...N
T.bispora                       NNKIIYLDPFTWYIA.PNALSSVENMVKAVGDSL...S
A.muciniphila                   NGAIILLDPQVWYLS.GGGIISLNKMISEVEAAFTDKK
C.botulinum                     NNKIISLNPELWYIS.SGGIVSTTEMLKEIKDS...IK
M.succiniciproducens            NNHIVYLNAENWYLA.FGGLQSMDTMISEIESAV...K
V.stercoraria                   DGKIVYLNTNNWYLM.NGGLNAMQEMIVEVQNVV...K
A.baumannii                     NGNIVYLDSSNWYLM.NNGLDGMKAMLKEVADAV...K
```

Figure 44 - BLAST+[58] search results of homologous proteins to CeuE. Aligned using T-Coffee[57] and image produced using ESPript.[45] Red block colours show fully conserved residues; red triangles show the position of basic triad of arginine residues; blue starts indicate position of coordinating histidine and tyrosine residues.

The His227 loop and Tyr288 positioning can also be compared with the structurally related proteins, VctP and YclQ, for which the apo-forms have also been structurally characterised.[59-60] These structures both show the histidine and tyrosine in the same topological location as CeuE, with the YclQ structure showing reduced electron density around the histidine loop, which also supports that this histidine is flexible.

The conservation of the histidine and tyrosine suggests a biological role in these proteins and hence could be part of a mechanism that enables certain PBPs to bind more than one type of ferric siderophore. The flexible loop of the His227 would enable this residue to coordinate if a tetradentate siderophore binds to the protein, but would also allow it to fold away from the binding pocket if a hexadentate siderophore is bound. The Tyr288 residue is well placed for either hexadetante or tetradentate binding, either acting as a hydrogen bond donor or as a directly coordinating ligand towards the metal centre according to the type (hexa- or tetradentate) of ligand present (Figure 45). This adaptor mechanism may give a biological advantage if it could enable the uptake of multiple iron-complexes *via* a singular system.

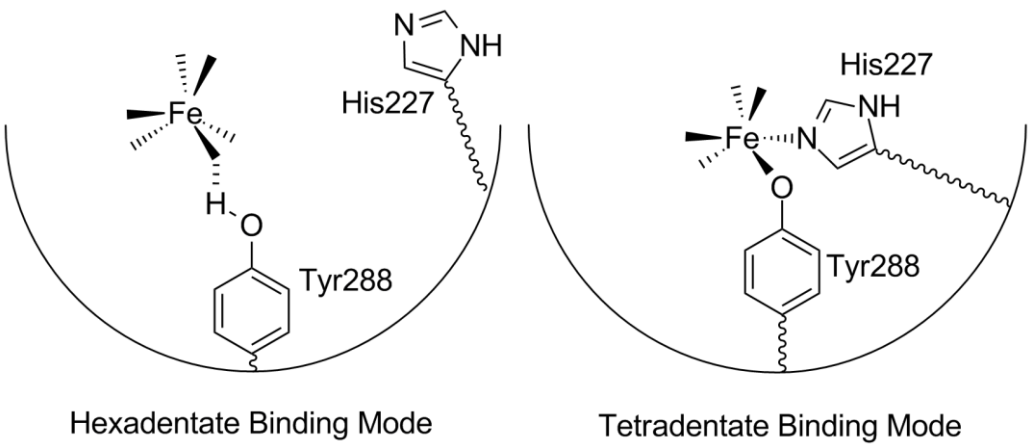


Figure 45 - Schematic diagram of a potential mechanism for both hexadentate and tetradentate binding in the binding pocket of CeuE.

Circular Dichroism and Fluorescence Biophysical Measurements

Octahedral complexes with two *cis*-positioned bidentate ligands are chiral.[61] The chirality at the iron centre in the {CeuE[Fe(4-LICAM)]} crystal structure is Λ, as seen in other ferric siderophore bound proteins, such as {$CeuE_2[Fe_2MECAM_2]$}, {FeuA[Fe(MECAM]}, {FeuA[Fe(enterobactin)]} and {FeuA[Fe(baillibactin)]} (Figure 46).[40, 52, 62]

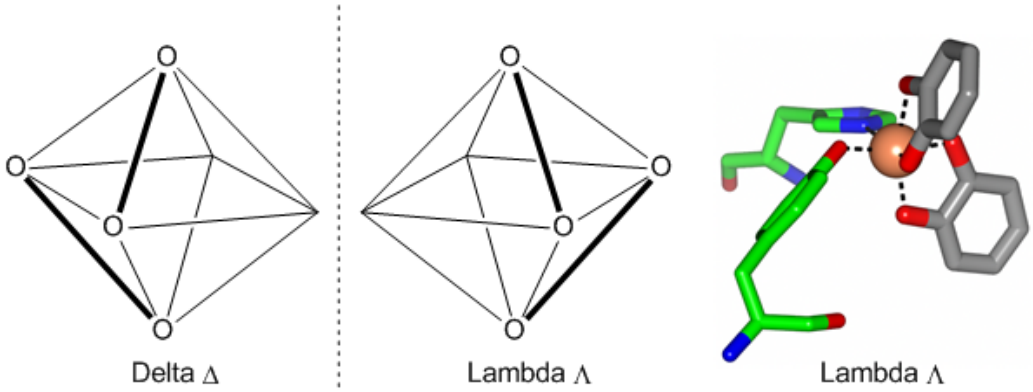


Figure 46 - Metal centre chirality in *cis*-bis(bidenate) complexes in comparison with the {CeuE[Fe(4-LICAM)]} crystal structure. Key; CeuE residues, green - carbon, blue - nitrogen, red - oxygen; 4-LICAM, grey - carbon, red - oxygen; Fe - Coral sphere.

To confirm that the Λ-configuration is preserved in solution, circular dichroism (CD) was used. The CD spectra were recorded in the wavelength range of the LMCT band arising from the catechol-iron interaction, in aqueous TrisHCl buffer solution at pH 8 (Figure 47). As expected, ferric-4-LICAM in the absence of CeuE does not preferentially absorb left or right polarised light, showing that there is a racemic mixture of both Λ and Δ-configured complexes in solution. When CeuE is present the CD spectrum contains a broad negative band with a minimum at 400 nm and the positive band with a maximum at 595 nm, giving an overall positive Cotton effect, a feature previously shown for catechol systems as being indicative of the Λ-configuration.[31, 63-64] There is also a large negative feature in the spectrum at 320 nm, which is often part of the CD spectrum for iron-bound catecholamides, known to be a ligand based, n→ π*

transition.[21] Due to the CD spectra of ferric-4-LICAM showing a racemic mixture, the preference of the Λ-configuration in the presence of CeuE must be due to the environment within the binding pocket of the protein.

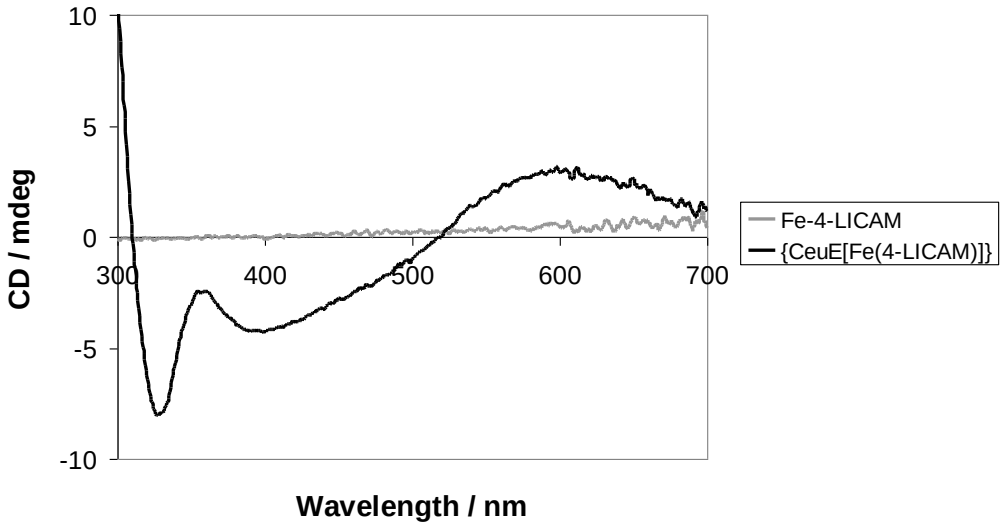


Figure 47 - Circular dichroism spectra of CeuE with {CeuE[Fe(4-LICAM)]} in black and ferric 4-LICAM in the absence of CeuE (grey) recorded in 20 mM TrisHCl buffer, 10 mM NaCl, 0.6% DMSO, pH 8.

Taking account of the direct coordination of His227 and Tyr288, it is important to characterise the affinity of ferric-4-LICAM towards CeuE. If binding affinity is very strong, it could indicate that this binding mode is not biologically relevant, as ferric-4-LICAM needs to dissociate from CeuE, and complex with the CeuBCD to be transported through the cytoplasmic membrane.

The binding affinity between ferric-4-LICAM and CeuE was determined through the measure of intrinsic fluorescence quenching upon ligand binding, as previously for the ferric complexes of various catechol siderophore ligands binding to FeuA from *B. subtilis*[51-52, 65-67], and FepB from *E. coli*.[68]

CeuE contains a single tryptophan (Trp287) adjacent to the coordinating Tyr288, which can be exploited for fluorescence quenching analysis to determine a dissociation constant. The experiments were all conducted in aqueous 40 mM TrisHCl pH 7.5, with 150 mM NaCl. The sodium chloride ensured biologically relevant ionic strength of the solution, which could be an important factor in the

binding affinity, especially in the presence of electrostatic contributions. To determine an accurate CeuE concentration, UV-vis absorbance was utilised, measuring the absorbance at 280 nm, and comparing it to the predicted molar absorbance coefficient (ε = 15930 mol^{-1} dm^3 cm^{-1}) of the mature proteins sequence, using ProtParam.[69]

The tryptophan was excited at 280 nm and the emission signal measured from 295 to 415 nm. The ferric-4-LICAM was complexed *in situ* by combining Fe(NTA) with H_4-4-LICAM in equimolar amounts in a buffered sample to make a concentrated stock solution of complex which was allowed to stand 30 minutes before titration. Control experiments showed that the buffer and the ferric-4-LICAM stock solution had no detectable fluorescence signal within the relevant wavelength range. A previous study involving the binding of ferric-siderophores to FepB, showed that binding of a ligand to FepB occurs within seconds.[68] To confirm that CeuE shows a similar kinetic behaviour, an initial titration was completed recording each data point at 1, 5 and 15 minutes (Figure 48). This showed there is little difference on leaving the solution to stand for longer than one minute.

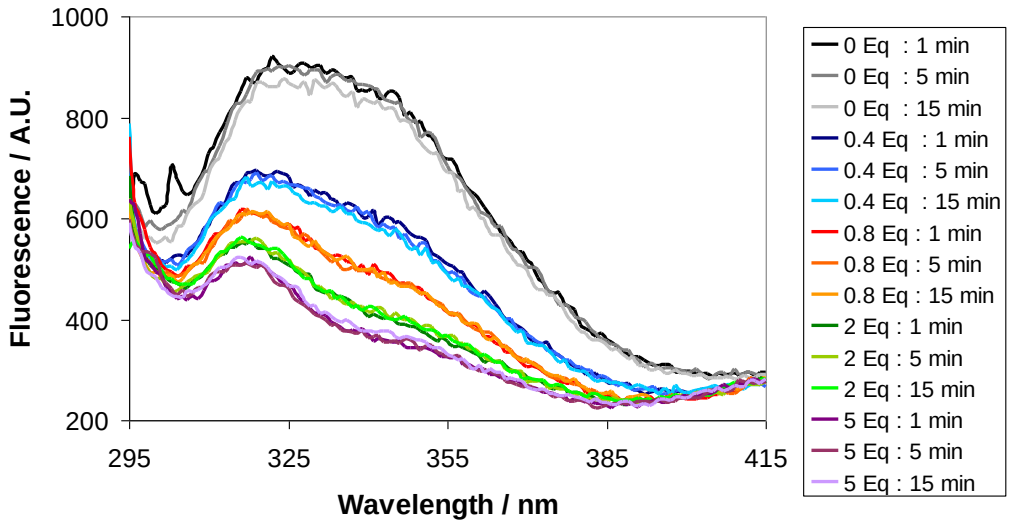


Figure 48 - Emission spectra of CeuE at 240 nM in 40 mM TrisHCl pH 7.5 150 mM NaCl, with various equivalents of Fe-4-LICAM, excitation at 280 nm. Each spectrum was recorded after 1 minute (dark), 5 minutes (medium) and 15 minutes (light), shown as shades of starting colour. Key; Black, 0 eq - blue, 0.4 eq - red, 0.8 eq - green, 2 eq - purple, 5 eq. Emission intensity normalised to initial intensity.

A small hypsochromic shift in the emission maximum at 330 nm can be observed upon Fe-4-LICAM addition. This indicates a change in the environment of the tryptophan residue, which can be expected due to its vicinity to the binding site. To evaluate the emission quenching effect of Fe-4-LICAM, the integrated peak area of the emission band was determined from 305 to 380 nm, in order to minimise the effect of the hypsochromic shift. The data were recorded in triplicate and at two different concentrations, with each concentration originating from a different batch of CeuE. The result of each titration experiment was plotted and fitted using DynaFit to calculate a dissociation constant *via* non-linear regression (Figure 49).[70] The average dissociation constants were calculated using a weighted average using the method reported by Lyons (Equation 2, Table 8).[71]

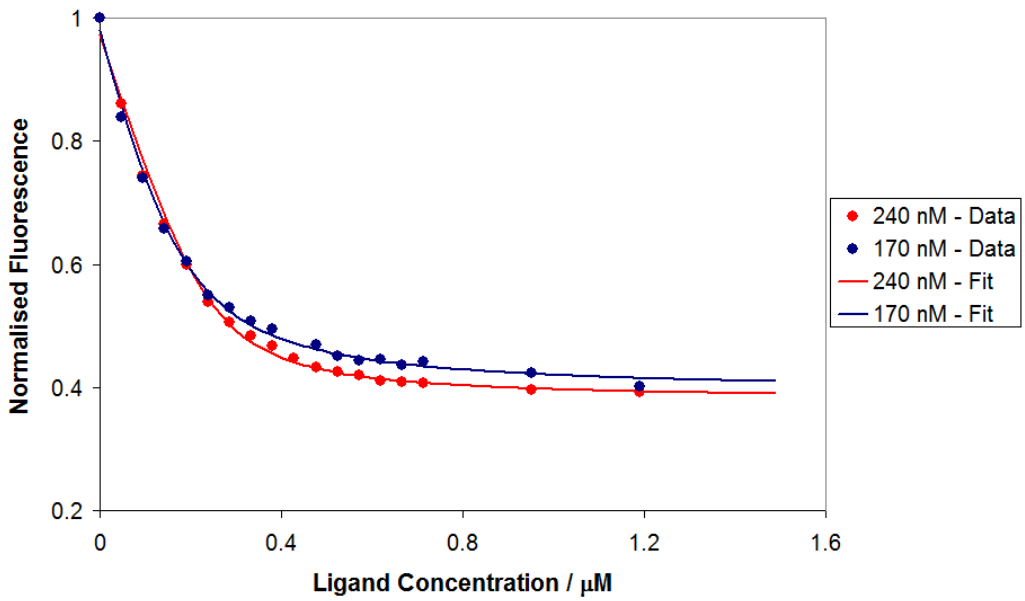


Figure 49 - Example fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titrations were carried out with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data points shown as circles; lines give the non-linear least-squares calculated fits (DynaFit).[70]

$$a = \frac{\sum\left(\frac{a_i}{\sigma_i^2}\right)}{\sum\left(\frac{1}{\sigma_i^2}\right)}$$

Equation 2 - Calculation of weighted average dissociation constant.

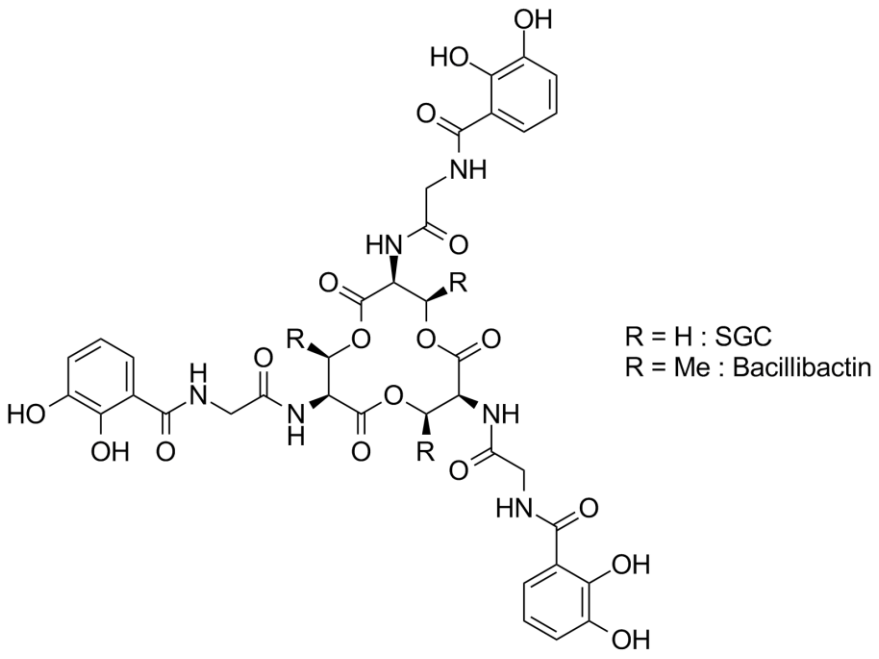
		Dissociation Constant / nM		
		Individual	Average of Batch	Overall Average
Batch 1 (240 nM)	Run 1	25.1 ± 1.3	27.7 ± 7.8	29.3 ± 11.7
	Run 2	30.3 ± 0.9		
	Run 3	14.9 ± 2.7		
Batch 2 (170 nM)	Run 1	44.9 ± 2.5	40.5 ± 13.4	
	Run 2	40.8 ± 3.8		
	Run 3	19.9 ± 5.3		

Table 8 - Calculated dissociation constants from non-linear regression analysis of the fluorescence quenching data of ferric-4-LICAM : CeuE titrations. Error of individual K_d values are calculated from non-linear regression fit, and errors of overall and batch calculated from standard deviation of the different K_d values within that group.

The calculated dissociation constants of the two batches of CeuE were found to be 27.7 ± 7.8 nM and 40.5 ± 13.4 nM. The values are not significantly different but have larger error values associated with them compared to most other dissociation constants found in the literature for other PBP systems. However, when comparing the same dissociation constant experimentally determined by different research groups, there is an inconsistency between the results and the level of error that is quoted. For example, the dissociation constant of ferric-bacillibactin and FeuA has been independently shown by separate groups to be 15 ± 4 nM (Abergel *et al.*)[65], 26.5 ± 1.4 nM (Miethke *et al.*)[66] and 29 ± 22.5 nM (Peuckert *et al.*).[51] All three of these results were obtained from different batches of FeuA, in separate laboratories, using fluorescence quenching to determine the dissociation constant. The dissociation constants obtained by Abergel *et al.* and Miethke *et al.* both have small errors, however, the

dissociation constant obtained by Peuckert *et al.* shows a larger error. Clearly there must be a larger error value than quoted by Abergel *et al.* and Miethke *et al.* due to the differences in the values obtained through basically the same experimental approach. Peuckert *et al.* do not state whether the results obtained in their experiments were results of difference batches of FeuA expressed and purified separately, but their dissociation constant appears to have a more realistic approximation of the error, given the other two results. The discrepancy between batches is likely to be due to errors in the estimation of the true protein concentration, and these differences may provide a more realistic estimation of the error in the binding constants. Therefore the dissociation constant of ferric-4-LICAM and CeuE was calculated from the average of all six titrations and is found to be 29.3 ± 11.7 nM.

The dissociation constant of 29.3 ± 11.7 nM is comparable to those reported for other periplasmic binding proteins and their associated ferric-siderophore complexes. Ferric-enterobactin has a dissociation constant of 19 ± 5 nM and 29 ± 1.4 nM for FeuA and FepB, respectively.[65, 68] Ferric-bacillibactin has a dissociation constant of 15 ± 4 nM[65], 26.5 ± 1.4 nM[66] and 29 ± 22.5 nM,[51] to FeuA and a structurally similar serine-glycine-catechol (SGC) found to be 52 ± 8 nM.[65] Interestingly, the enantiomers of enterobactin and SGC, which were found to form the Λ-configuration in solution rather than the Δ-configuration, have also been used for binding studies and show that these bind with a slightly increased strength. Enantioenterobactin was shown to have a dissociation constant of 10 ± 2 nM for FeuA and 14.5 ± 1 nM for FepB, whereas D-serine-glycine-catechol, was found to have a dissociation constant of 32 ± 4 nM, to FeuA, suggesting that both FeuA and FepB, may have a preference for the Λ-configuration.[65, 68]



In conclusion, the CD spectroscopy shows that the ferric-4-LICAM is bound in the Λ-configuration in solution state as observed in the {CeuE[Fe(4-LICAM)]} crystal structure. The observation that ferric-4-LICAM is achiral in the absence of CeuE shows that CeuE selectively binds the Λ-configuration. The fluorescence quenching titration shows that the dissociation constant for ferric-4-LICAM towards CeuE is within the biologically relevant range of similar proteins, suggesting that the coordination of the tyrosine and histidine residues observed in the {CeuE[Fe(4-LICAM)]} crystal structure may be of biological relevance.

2.5 Summary of Chapter 2

The siderophore mimic H_4-4-LICAM was synthesised, characterised and complexed to both iron(III) and gallium(III). 4-LICAM^{4-} forms complexes with 3:2 and 1:1 ligand to metal ratios with these metal cations. Which complex dominates in solution depends on the relative concentrations of both H_4-4-LICAM and metal. For intermediate ligand to metal concentration ratios, between 3:2 and 1:1, the two complexes are in equilibrium. With iron(III), equilibration occurs within minutes in aqueous buffer solution at pH 7.5. In contrast, the equilibration is slower with gallium(III) in 5:1 d_6-DMSO:D_2O in the presence of sodium hydroxide.

The PBP CeuE was over-expressed, purified and crystallised, both in the apo form and also as a co-crystal with ferric-4-LICAM. Ferric-4-LICAM was found to be present in a unique binding mode in the binding pocket of CeuE, with residues His227 and Tyr288 coordinating directly to the metal centre. Sequence alignment revealed that these two residues are conserved among related proteins from several bacteria, suggesting that they are part of a mechanism allowing these proteins to bind ferric-tetradentate as well as ferric-hexadentate siderophores, and facilitate the uptake of both species.

The metal-centred chirality of the {CeuE[Fe(4-LICAM)]} structure was found to be Λ, both in the solid state (x-ray diffraction analysis) and in solution (CD spectroscopy), consistent with the previous {$CeuE_2$[Fe_2MECAM$_2$]} structure. The dissociation constant was determined using intrinsic fluorescence quenching to be 29.3 ± 11.7 nM, consistent with those reported for other ferric-catecholate siderophores. However, this study is the first time a tetradentate catecholate siderophore's binding constant has been determined, and therefore cannot be directly compared to other tetradentate catecholate siderophore binding constants.

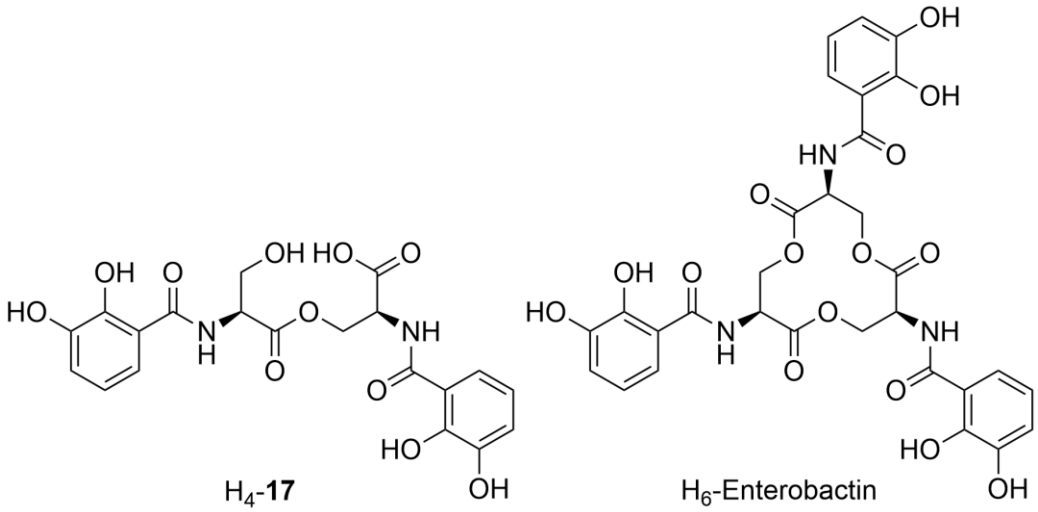
Chapter 3: Tetradentate Siderophores, a Natural Product Approach

Project Aims

The work outlined within this chapter is a continuation of Chapter 2, following on from the biomimetic approach to characterisation of the enterobactin linear dimer, moving onto characterisation of the natural product itself.

3.1 Introduction

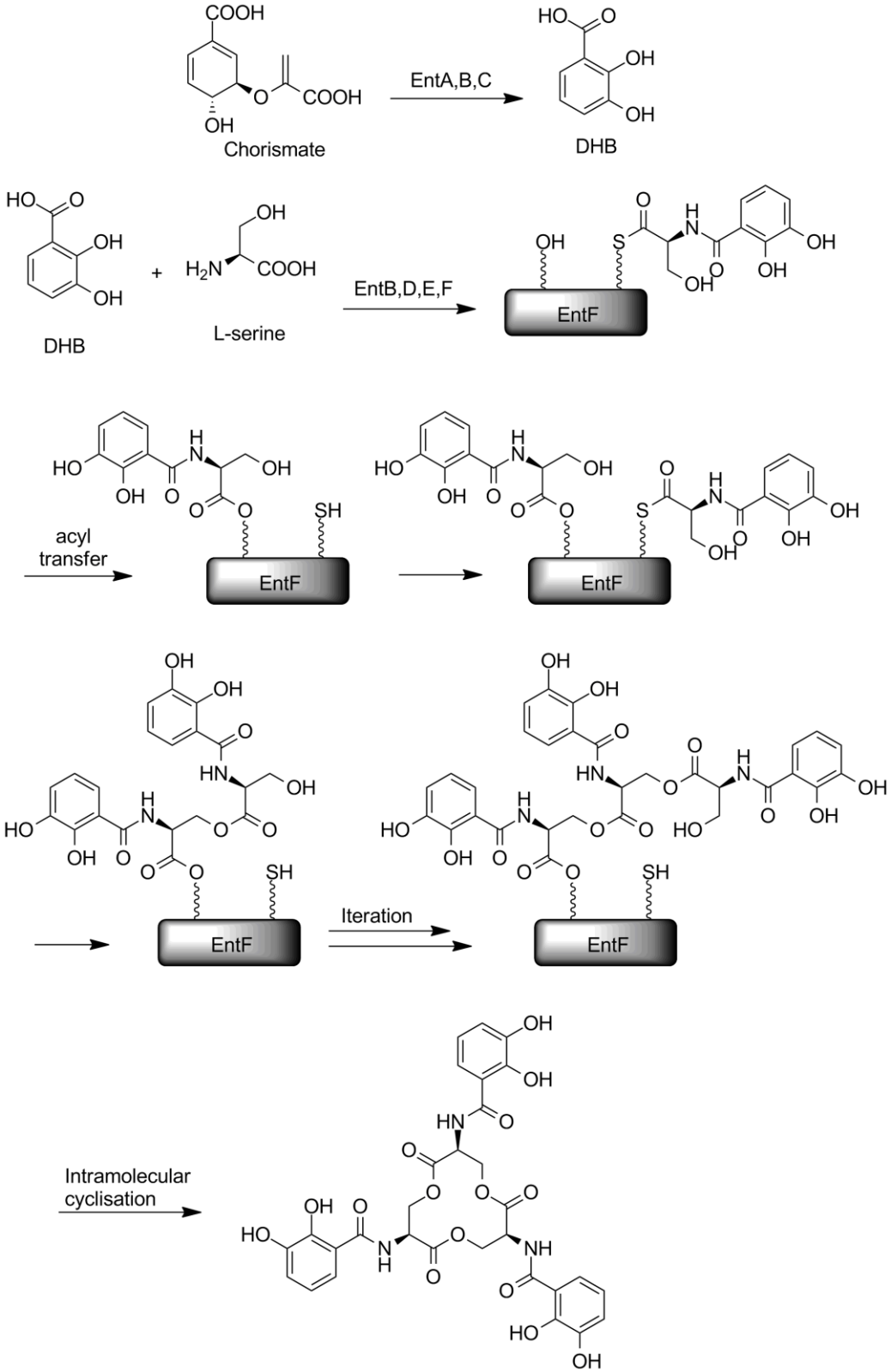
Following the study of the synthetic derivative H_4-4-LICAM, described in Chapter 2, the next step in the study of tetradentate siderophores was to investigate a naturally occurring enterobactin fragment, namely the enterobactin linear dimer H_4-**17** (the quoted protonation state only refers to the catechol protons). A key aim was to establish the coordination chemistry of a biologically relevant tetradentate siderophore both in free solution and in complex with an associated uptake protein.



The linear dimer H_4-**17** can be considered as a segment of the hexadentate enterobactin molecule, with two catecholamide-based chelating groups on a backbone consisting of two L-serine units connected *via* an ester linkage.

The biosynthesis of enterobactin is carried out by six enzymes, EntA-F, coded for by the genes *entA-F*.[1] EntA and EntC synthesise 2,3-dihydroxybenzoic acid

(DHB) from chorismate, while EntD-F are involved in the assembly of the DHB and L-serine units into the trilactone.[1] EntB is a bifunctional enzyme involved in both steps of the biosynthesis (Scheme 3).



Scheme 3 - Biosynthetic pathway of enterobactin.[2]

It is not clear whether the enterobactin fragments are released from the biosynthetic pathway in order to act as siderophores in their own right, or are a result of the hydrolysis of enterobactin.[3] A more detailed investigation of the EntF-catalysed elongation and cyclolactonisation step by Shaw-Reid *et al.* showed that only enterobactin was observed as a product when wild-type EntF was utilised, with no detectable amount of any enterobactin fragments using HPLC.[2] On repeating the experiment with a mutation in EntF (S1138C), the catalytic turnover rate was dramatically reduced (100 min^{-1} to 0.9 min^{-1}), and detectable amounts of the enterobactin fragments were observed using mass spectrometry. The authors concluded that due to the absence of detectable amounts of fragments in the wild-type experiment it is likely that they are primarily hydrolysis degradation products.[2]

After its identification and characterisation,[4-7] the enterobactin linear dimer H_4-**17** was used previously for several studies, including its coordination chemistry with iron(III)[8] and gallium(III)[7], growth studies of *E. coli*[9-10] and various marine bacteria[11], and other biological studies.[2, 12-15] The majority of these investigations involved the isolation of H_4-**17** from growth media, while only one involved its synthesis as part of the total synthesis of enterobactin and enantioenterobactin.[16]

There has been relatively little research involving the coordination chemistry and the biological importance of the linear dimer compared to enterobactin itself. The aim of the work in this Chapter was its synthesis, determination of its coordination chemistry with iron(III) and the interactions of the ferric-complex of **17**$^{4-}$ with CeuE in solution and the crystal.

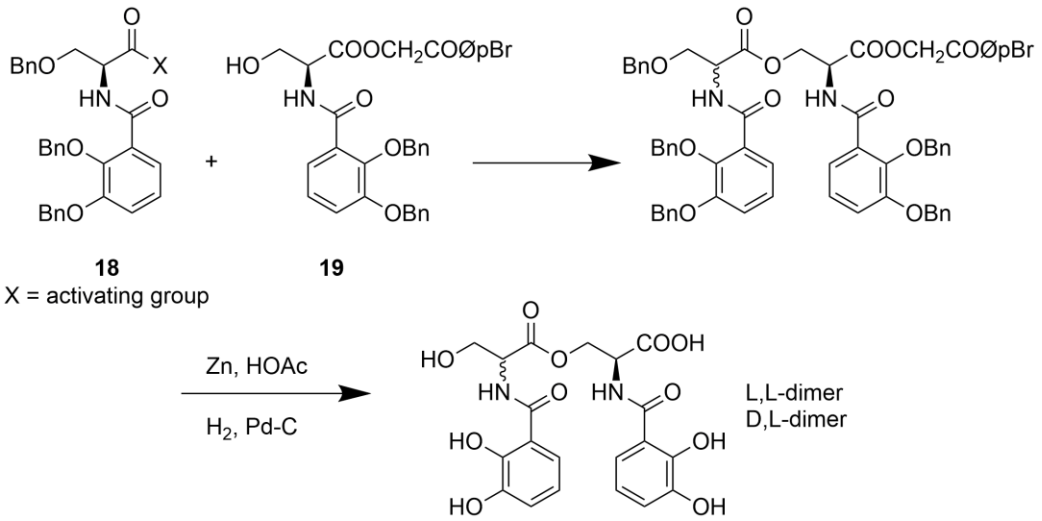
Previous Synthetic Approaches

The synthesis of enantiomerically pure H_4-**17** has not been described previously. In the total synthesis of enterobactin carried out by Rastetter *et al.*, compound H_4-**17** was synthesised in a diastereomeric form.[16] These authors encountered a

number of problems, including the racemisation of the serine backbone and various protecting group issues. They outlined two approaches towards the total synthesis of enterobactin, the *N*-benzoyl monomer approach and the urethane-protection route. These served as the basis of the methodology used in this chapter.

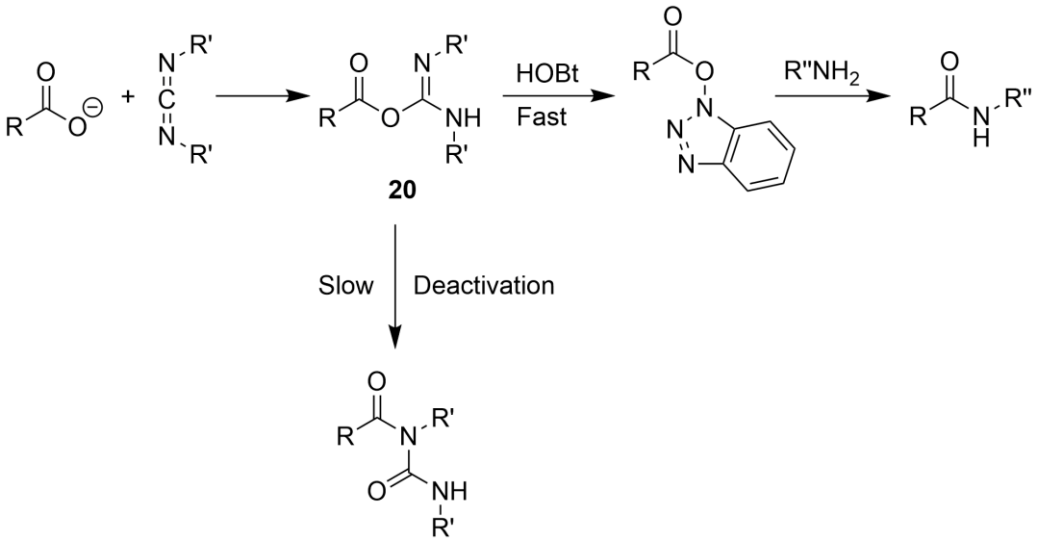
The *N*-benzoyl Monomer Approach

This method involves the synthesis of two differently protected *N*-[2,3-bis(benzyloxy)-benzoyl]serine monomers followed by formation of the trilactone backbone in a step-wise fashion, yielding a protected linear dimer. Controlled deprotection and esterification of this dimer, with another *N*-benzoyl monomer yields the linear trimer, which could be cyclised to form enterobactin. When investigating the *N*-benzoyl monomer approach it was found that the esterification step yields diastereomeric products, due to racemisation of the activated form of the carboxylic acid, **18**, when using DCC, DCC/HOBt or DCC and pyridine (Scheme 4).[16]



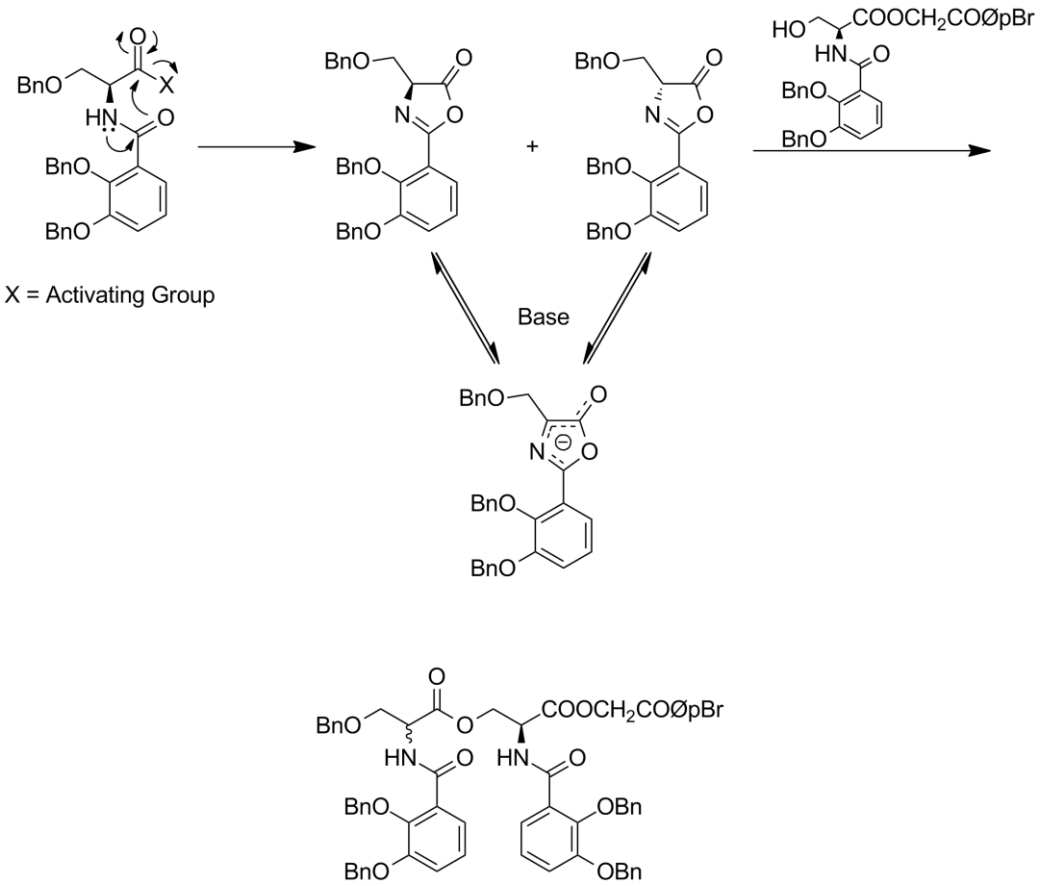
Scheme 4 - The *N*-benzoyl monomer route when using X = DCC, DCC/HOBt or DCC and pyridine to activate the carboxylic acid for esterification.

The additive HOBt usually aids prevention of racemisation and deactivation by rapidly forming an activated aromatic ester with **20** before the competing acyl transfer can occur (Scheme 5).



Scheme 5 - The use of HOBt to reduce the deactivation of the *O*-acylisourea mixed anhydride **20**.

The generated activated aromatic ester is still active enough to couple with a nucleophile, but because of its increased stability compared to the *O*-acylisourea mixed anhydride **20**, slows the rate of oxazolone formation and hence, reduces racemisation (Scheme 6).[18-19] Unusually, racemisation still occurred with the addition of HOBt in the esterification carried out by Rastetter *et al.*, showing that the racemisation must have an extraordinarily high rate, or that the generated activated aromatic ester is still active enough to undergo oxazolone formation.[17] Rastetter *et al.* attributed the high rate of racemisation to the electron-releasing nature of the benzyl groups on the catechol ring, which can promote the nucleophilicity of the benzamide carbonyl oxygen enhancing the rate of oxazolone formation. Various other coupling reagents (Mitsunobu esterification, CDI coupling, *N*-methyl-2-chloropyridinium iodide) lead to dehydration of the monomer alcohol **19**.[16] The synthesis towards the linear dimer was continued and deprotection using zinc and acetic acid, followed by hydrogen over a palladium catalyst removes all the protecting groups, yielding the L, L -dimer and the D, L -dimer.



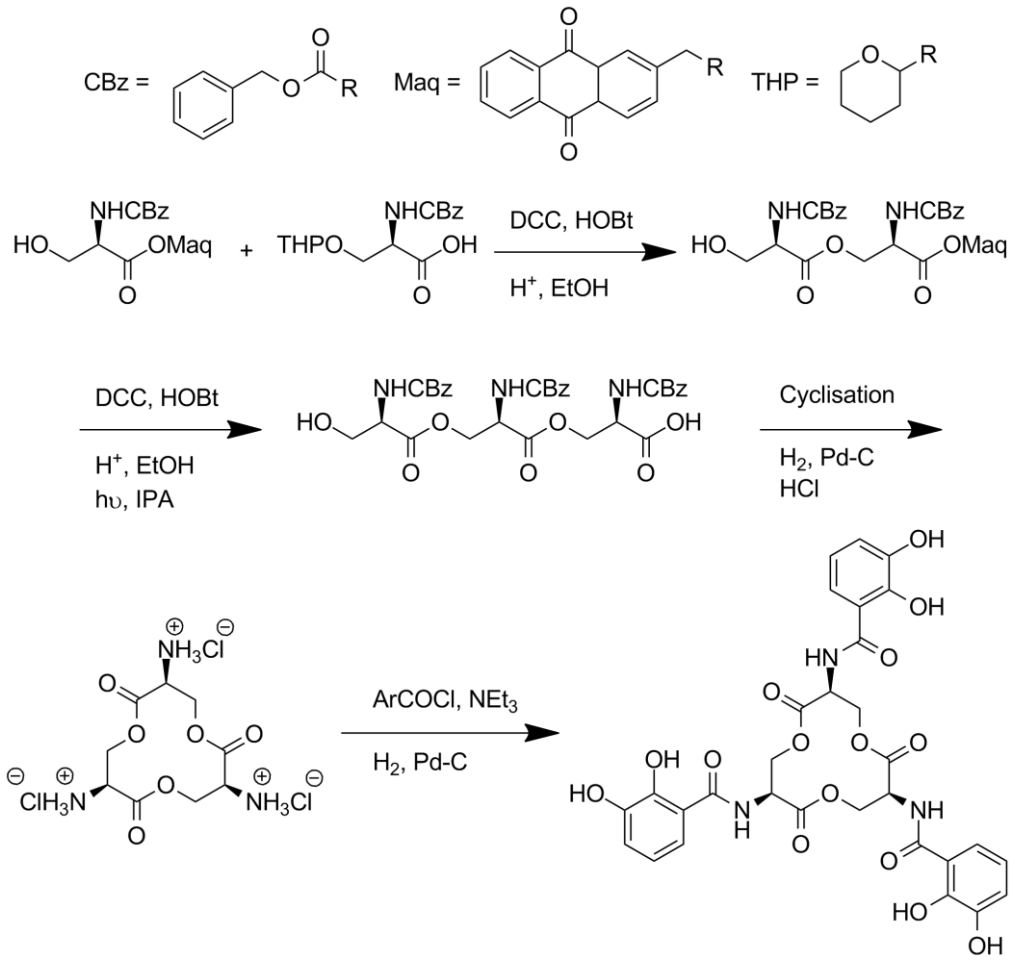
Scheme 6 - Oxazolone-mediated racemisation during esterification.

The *N*-benzoyl monomer approach for the total synthesis of enterobactin was abandoned due to these initial problems.

The Urethane Protection Route

The urethane protection route involves protection of the amine functionality of the serine units as a carbamate, followed by esterification, before urethane deprotection and introduction of the *N*-[2,3-bis(benzyloxy)-benzoyl] monomers. This route leads to successful formation of the ester linkage with no evidence for racemisation. Rastetter *et al.* at this point continued the synthesis towards enterobactin rather than completing the synthesis of the linear dimer in an enantiomerically pure form. The synthesis was completed by coupling the serine containing the free acid with the protected linear dimer to produce the protected linear trimer, which was then cyclised. The benzyl protected catechol

units were added onto the cyclic trimer and subsequently deprotected, yielding enterobactin (L-serine) or enantioenterobactin (D-serine) (Scheme 7).[16]



Scheme 7 - The urethane protection route for the total synthesis of enterobactin outlined by Rastetter *et al.*

The two routes towards the synthesis of the linear dimer both have advantages and disadvantages. The *N*-benzoyl monomer approach is simpler with fewer steps, but it suffers from racemisation. The urethane protection approach has the advantage of being proven to form the ester linkage without racemisation, but requires more synthetic steps. Furthermore, the labile ester linkage has to be maintained during urethane deprotection, catechol coupling and catechol deprotection, whereas in the *N*-benzoyl monomer approach it only has to be maintained during catechol deprotection. Therefore it was decided to investigate the *N*-benzoyl monomer approach first, as there have been advances in the development of peptide coupling reagents since Rastetter *et al.* published their

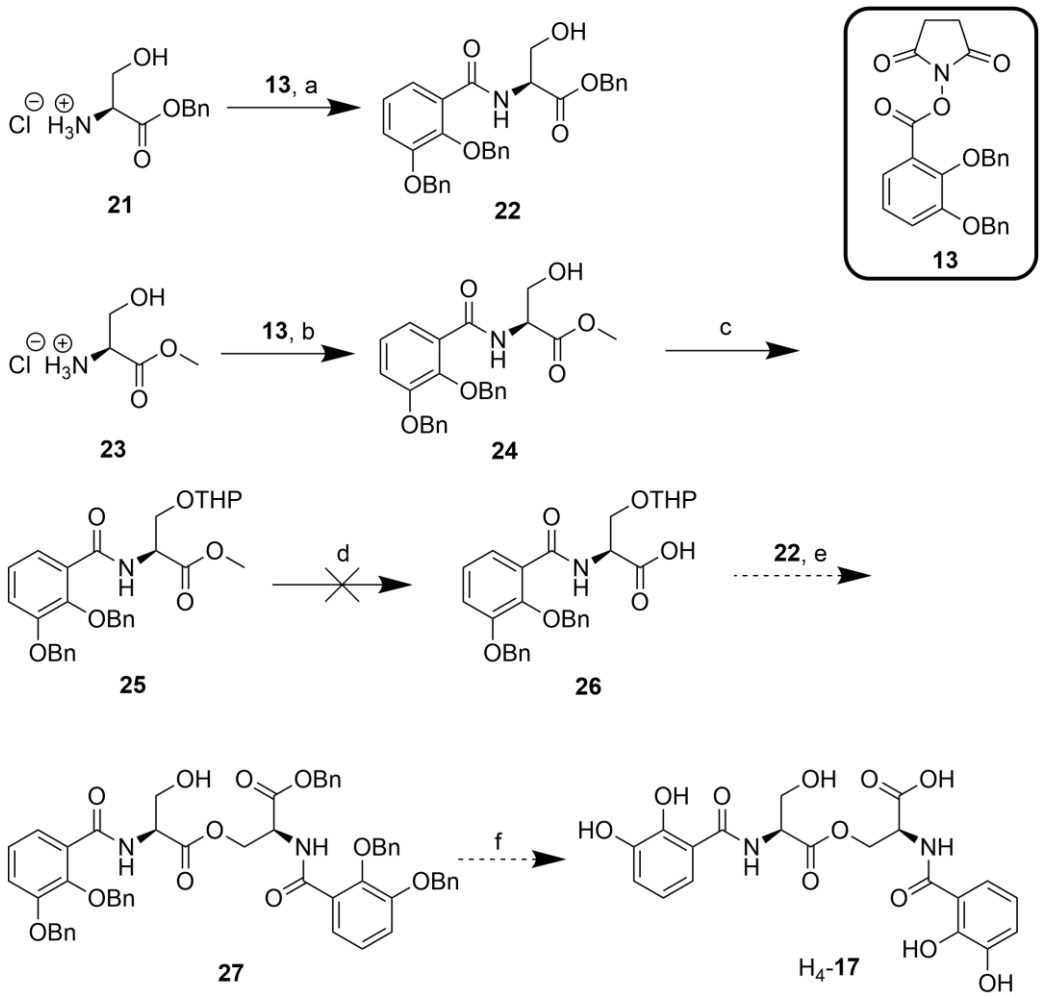
synthesis; the new reagents available were anticipated to prevent racemisation during the esterification step.[18-19]

3.2 Synthesis

Initial Approaches to the Synthesis of the Enterobactin Linear Dimer

The work in this section was completed with the aid of Emily Brooke, a summer placement student working under my supervision (July-September 2012).

As the electron donating properties of the protecting groups were linked to the racemisation of the serine during esterification, it was decided to replace the benzyl with tetrahydropyran (THP) ether, since this was used successfully by Rastetter *et al.* in the urethane-protection route. Due to the limited availability of the 2-bromomethylanthraquinone (Maq-Br) starting material, the carboxylic protecting group had to be changed. The alternative protecting group was required to possess orthogonal deprotection requirements to both the benzyl ethers (catalytic hydrogenation) of the catechol protecting groups and the THP ether (acid deprotection) of the alcohol protecting group; hence a base-labile methyl ester was chosen (Scheme 8).



Scheme 8 - Initial synthetic route towards the H_4-**17**. (a) NEt_3; (b) NEt_3; (c) THP, pyridinium *p*-toluenesulfonate; (d) NaOH / NMe_4OH; (e) DCC, HOBt, H^+ workup; (f) H_2, Pd-C 10%.

Compound **22** was synthesised using commercially available L-serine benzyl ester hydrochloride as the starting material. Compound **13** was used to introduce the protected catechol scaffold, yielding amide **22** in a 91% yield. The product was characterised using ^{1}H and ^{13}C NMR spectroscopy, infra red spectroscopy and ESI mass spectrometry. The mass spectrum shows peaks at *m/z* 512.2063 and *m/z* 534.1867, corresponding to the protonated $[M+H]^+$, and sodiated $[M+Na]^+$, ions, consistent with a molecule of formula $C_{31}H_{29}NO_6$. A characteristic doublet at 8.86 ppm in the ^{1}H NMR spectrum, which could be assigned to the amide proton, confirms formation of the amide bond. The resonances due to the serine backbone consist of a doublet of doublet of doublets due to the proton on the α-carbon (H_a), and two roofed doublet of doublets

which are assigned as the two geminal protons on the β-carbon (H_b and H_c). Due to the chirality of the serine, the two geminal protons, H_b and H_c, are diastereotopic and both couple with H_a, and to the other geminal proton, arising to the two sets of doublet of doublets. As the chemical shifts of these two signals are close, the two signals show a roof effect, which accounts for the profile of the signals (Figure 50).

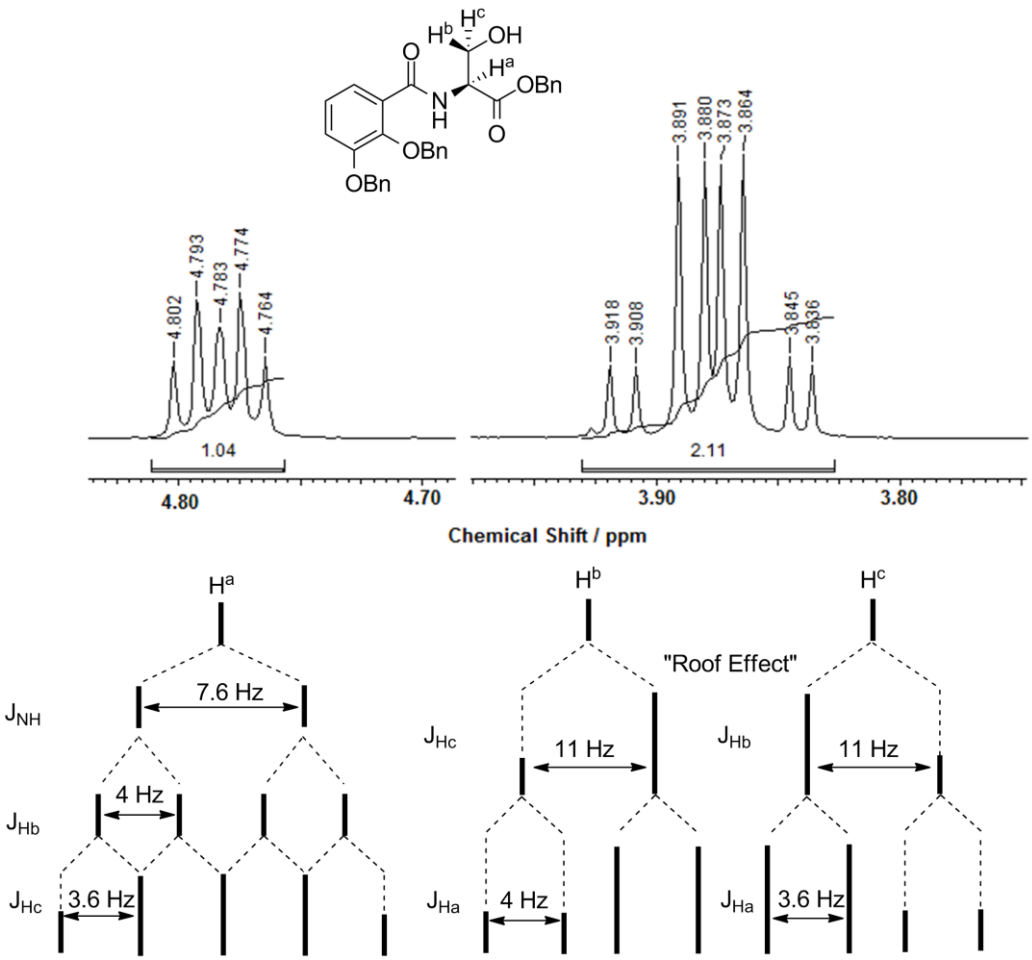


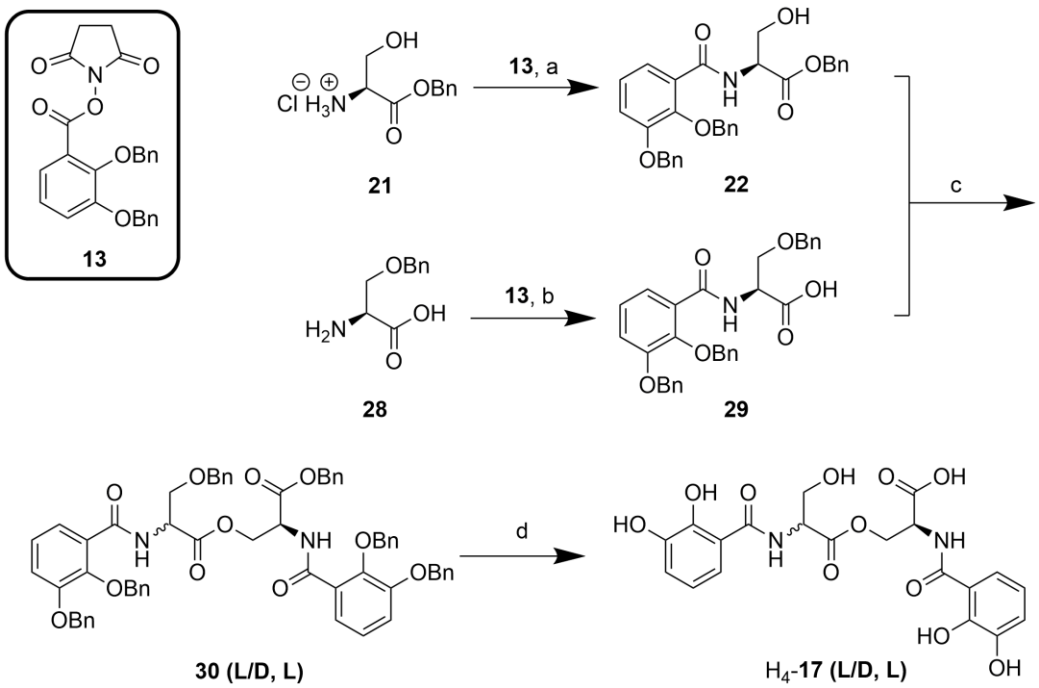
Figure 50 - ^{1}H NMR spectrum of **22** in $CDCl_3$ shown in the regions between 4.85-4.70 ppm and 4.00-3.75 ppm. The J-coupling splitting pattern is shown below, explaining the splitting for each of the three resonances. *The assignment of H^b and H^c can be interchanged.

Compound **24** was synthesised using L-serine methyl ester hydrochloride and the catechol unit introduced and characterised as described for **22** (78% yield). Compound **24** was then treated with 3,4-dihydro-*2H*-pyran, using pyridinium *p*-

toluenesulfonate in dichloromethane to introduce the THP protecting group. The resulting compound **25** was obtained in 86% yield using the same methodology as Rastetter *et al.*[16] Compound **25** was characterised using ^{1}H and ^{13}C NMR spectroscopy, infra-red spectroscopy and ESI mass spectrometry. The mass spectrum showed a peak at *m/z* 542.2139 consistent with a $[M+Na]^+$ ion for a molecule with a formula of $C_{30}H_{33}NO_7$. The ^{1}H and ^{13}C NMR spectra confirmed the addition of THP, and also showed the doubling of all signals associated with the two diastereomers that form upon non-stereospecific introduction of the THP group.

Deprotection of the methyl ester of **25** was attempted with sodium hydroxide in a methanol:dichloromethane mixture. Monitoring the reaction *via* TLC and mass spectrometry indicated formation of the desired de-methylated product, however, the subsequent analysis of the product revealed that the THP protecting group was also lost, either during the reaction or during the work-up procedure. Changes in solvent, reaction conditions, or base (from sodium hydroxide to tetramethylammonium hydroxide) did not result in selective deprotection of the methyl ester whilst maintaining the THP ether. Due to these protecting group compatibility issues an alternative route that avoids THP as protecting group was developed.

The second synthetic route involved protecting the alcohol and acid groups of the two serine backbone units as a benzyl ether or benzyl ester, respectively. The starting materials, L-serine benzyl ester hydrochloride **21** and *O*-benzyl-L-serine **28**, are commercially available (Scheme 9). The *N*-benzoyl monomer **29** was synthesised and characterised as described for compound **22**.



Scheme 9 - Second synthetic route towards H_4-**17**. (a) NEt_3; (b) NEt_3; (c) Esterification, various conditions; (d) H_2, Pd-C 10%.

Various reaction conditions were investigated in order to obtain compound **30** in enantiomerically pure form. Rastetter *et al.* reported the use of DCC/HOBt, CDI, *N*-methyl-2-chloropyridinium iodide[20-21] and ethyl diazodicarboxylate, triphenylphosphine (Mitsunobu reaction).[22-23] In the investigation to form **30**, a range of conditions were applied in order to form the enantiomerically pure ester including DCC, EDC, HBTU, DCC/HOBt, EDC/HOBt, DCC/DMAP, activation using thionyl chloride or formation of an activated ester with *N*-hydroxysuccinimide. Of these, only DCC/HOBt gave an isolatable amount of product in 16% yield as a diastereomeric mixture of **30** (L, L and D,L).

Despite the formation of the diastereomeric mixture of compound **30**, the catalytic hydrogenation for global benzyl deprotection was conducted in order to test the reaction and to form the diastereomeric mixture of compound H_4-**17**. The reaction was performed and the product characterised in the same manner as for H_4-4-LICAM described in Chapter 2. Mass spectrometry showed that the ester linkage had been maintained during the reaction, with the peaks found at *m/z* 465.1142 and *m/z* 487.0985 corresponding to the $[M+H]^+$ and $[M+Na]^+$ ions for a molecular formula of $C_{20}H_{20}N_2O_{11}$ respectively. The ^1H and ^{13}C NMR

spectra again showed the doubling of the signals associated with the two diastereomers in the sample, and the ^{13}C NMR containing the ester linkage carbon signal, confirming the ester had not hydrolysed. There was no evidence of transesterification with ethanol, used as the solvent for the reaction. The reaction yielded a 94% of the diastereomeric mixture of compound H_4-**17** (L, L and D, L).

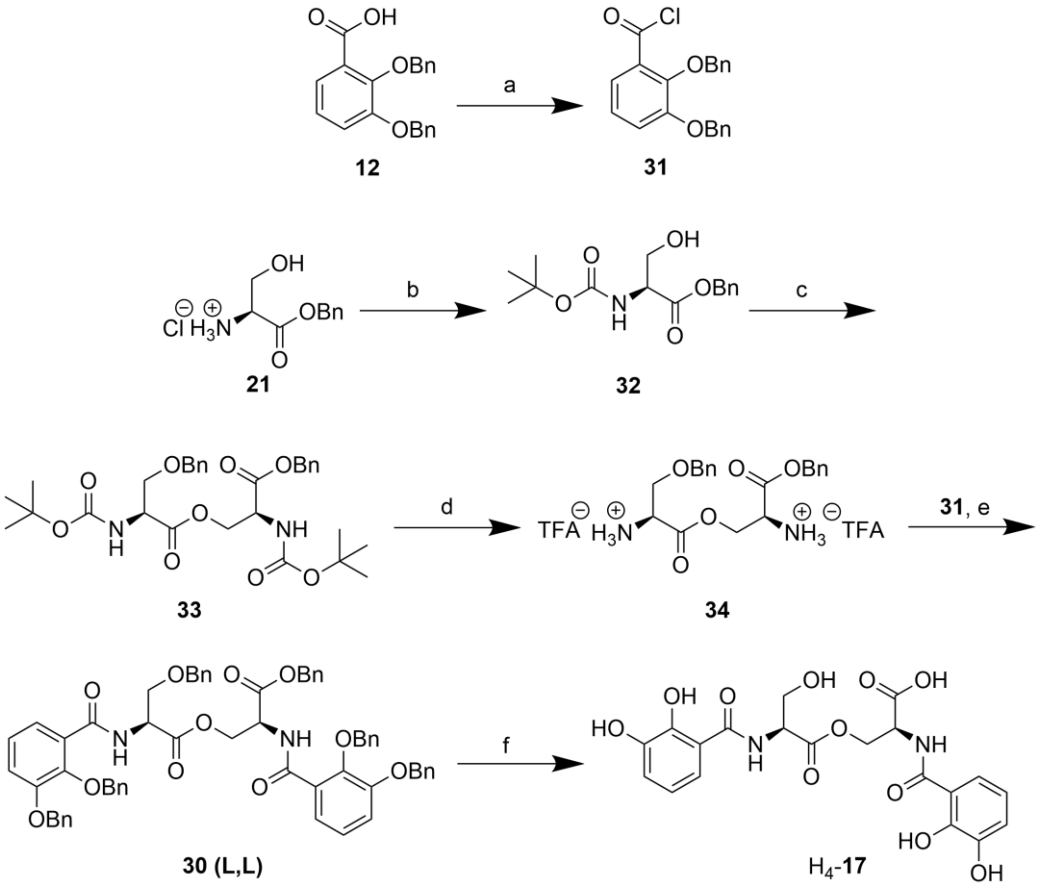
This reaction sequence confirmed that compound H_4-**17** can be synthesised using the outlined synthetic route, although the diastereomers would need to be separated before catalytic hydrogenation. However, due to the low yield of the esterification reaction step, it was concluded that this route is not a viable option for scale-up to access larger quantities. For this reason the route was abandoned.

Synthesis of the Enterobactin Linear Dimer

Due to the problems described above, the urethane protection approach outlined in section 3.2 was attempted. Rastetter *et al.* investigated the use of both the *N-tert*-butoxycarbonyl, and carboxybenzyl groups for the protection of the amino groups. In both cases the esterification step was reported to proceed without racemisation.[16]

Ultimately the carboxybenzyl group was used for amino protection in the synthesis of enterobactin, due to difficulties in the cyclisation step of the linear trimer when the *N-tert*-butoxycarbonyl group was used. However, this issue does not need to be considered for the synthesis of the linear dimer.[16] In the synthesis of the linear dimer, it is necessary for the alcohol and the carboxylic acid of the serine backbone to be protected to control the esterification step. In order to simplify the deprotection procedure once the benzyl protected catechol moieties are coupled to the backbone, and hence to have a global benzyl deprotection, the protecting groups selected for the alcohol and the carboxylic acid were a benzyl ether and benzyl ester, respectively. As the serine monomers have benzyl protecting groups attached, the amino protecting groups must be removed under orthogonal deprotection conditions compared to that of the

benzyl groups. For this reason the carboxybenzyl group cannot be used, and therefore *N-tert*-butoxycarbonyl protecting group were chosen (Scheme 10).



Scheme 10 - Third synthetic route towards H_4-**17**. (a) $SOCl_2$; (b) Boc_2O, NEt_3; (c) Boc-*O*-benzyl-L-serine, EDC, HOBt, ; (d) TFA; (e) DIPEA; (f) H_2, Pd-C 10%.

Compound **32** was synthesised by reacting L-serine benzyl ester hydrochloride with di-*tert*-butyl dicarbonate in the presence of triethylamine (98% yield). Compound **32** showed a characteristic singlet at 1.43 ppm in the ^{1}H NMR spectrum which integrates to nine protons corresponding to the *tert*-butyl group. A broad doublet at 5.52 ppm of relative integration one, can be assigned to the carbamate proton, which couples to the proton on the alpha-carbon of the serine unit. The NMR and mass spectrometry data are consistent with the literature, however, a previous synthesis reports the compound as a colourless oil, whereas a white waxy solid was obtained here (melting point 66.8-68.2 °C).[24]

Initially, compound **33** was synthesised from **32** and commercially available Boc-*O*-benzyl-L-serine using DCC as the coupling reagent and DMAP as an additive,

however, a DCC-derived impurity formed during the reaction. The impurity could not be removed from the product despite numerous purification attempts. Therefore the reaction was repeated using EDC as an alternative coupling reagent, with HOBt as the additive. Compound **33** was easily purified *via* silica column chromatography, with no evidence of racemisation, to give compound **33** in 61% yield (characterisation data consistent with the literature).[25]

Removal of the Boc-protecting groups was then achieved using trifluoroacetic acid in dichloromethane.[26] The ^{1}H NMR spectrum showed the absence of the two *tert*-butyl and the carbamate proton signals, while mass spectrometry and ^{13}C NMR confirmed that the ester linkage had been maintained during Boc-deprotection.

To introduce the catechol functionality onto diamine **34**, compound **13** was initially utilised in the same fashion as in the synthesis of H_4-4-LICAM (see Chapter 2). The reaction, however, was found to proceed in low yields; therefore an alternative method was designed. Compound **12** was instead activated as the acid chloride, compound **31**, prepared in accordance with the literature and used without isolation.[27] Compound **31** was then reacted with **34** in the presence of *N*,*N*-diisopropylethylamine in anhydrous dichloromethane under a dry nitrogen atmosphere in order to limit hydrolysis of the ester under basic conditions. The coupling of the catechol units to **34** proceeded in 44% yield, lower than 61% reported by Rastetter *et al*. in the synthesis of enterobactin.[16] A peak in the mass spectrum at *m/z* 1005.3922 was assigned as the $[M+H]^+$ protonated molecular ion and the ^{1}H and ^{13}C NMR spectra both showed the addition of the benzyl protected catechol units on to compound **34** to form compound **30**. There was no evidence of the formation of diastereomers in the NMR spectra. Elemental analysis was also used to confirm the purity of the sample.

Compound **30** was deprotected using catalytic hydrogenation under an atmosphere of hydrogen in the presence of a 10% Pd-C catalyst, to obtain compound H_4-**17** in a quantitative yield. Mass spectrometry, ^{1}H, ^{13}C NMR and elemental analysis confirmed the successful synthesis and purity of the product.

The characterisation data were consistent with literature data for H_4-**17** isolated from growth media (see section 7.2).[12] To identify whether the chirality of the serine units had been affected during the synthesis, specific rotation of polarised light at the sodium line was measured in methanol, and was found to be $[\alpha]_D$ = +15, in good agreement with the literature value of $[\alpha]_D$ = +13.[12]

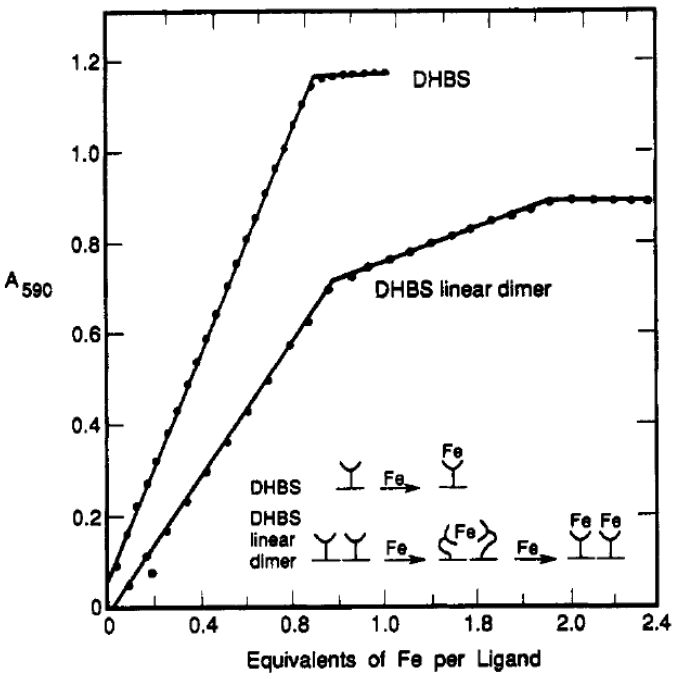
3.3 Ligand-Metal Interactions

As discussed in Chapter 2, the mismatch between the preferred coordination number of iron(III) and the denticity of the ligand required investigation of the speciation of the complexes formed in solution. After the study of the ligand-metal interaction with the biomimetic derivative 4-LICAM, the successful synthesis of the enterobactin linear dimer allowed this investigation to be completed.

Enterobactin is one of the most widely studied siderophores,[28] but little research has been conducted on the coordination chemistry of H_4-**17**. Two studies have investigated ligand-metal coordination chemistry, one by Scarrow *et al.* which investigated the iron(III) coordination of linear dihydroxyserine compounds derived from enterobactin,[8] and the other by Bergstrom *et al.* which focused on NMR studies of 2,3-dihydroxyserine compounds, which also extended into coordination with gallium(III).[7]

Scarrow *et al.* used UV-visible measurements of the LMCT band in a spectrophotometric titration. Upon titrating H_4-**17** with iron(III), the absorbance at 590 nm increased. Plotting the absorbance at this wavelength versus equivalents of iron(III), a linear change of absorbance was observed, with notable breaks at 1.0 and 2.0 equivalents of iron(III). These changes were attributed to the formation of 1:1 and 1:2 ligand-to-metal complexes. However, Scarrow *et al.* did not discuss an apparent bathochromic shift in the UV-visible data shown in the supplementary information.[8] During the titration, the maximum absorption was initially at approximately 500 nm and then shifted to approximately 550 nm, between 0.5 and 1.0 equivalents of iron(III). This

bathochromic shift indicates a change in coordination environment, from a tris(catechol)iron(III) to a bis(catechol)iron(III) complex (Figure 51).



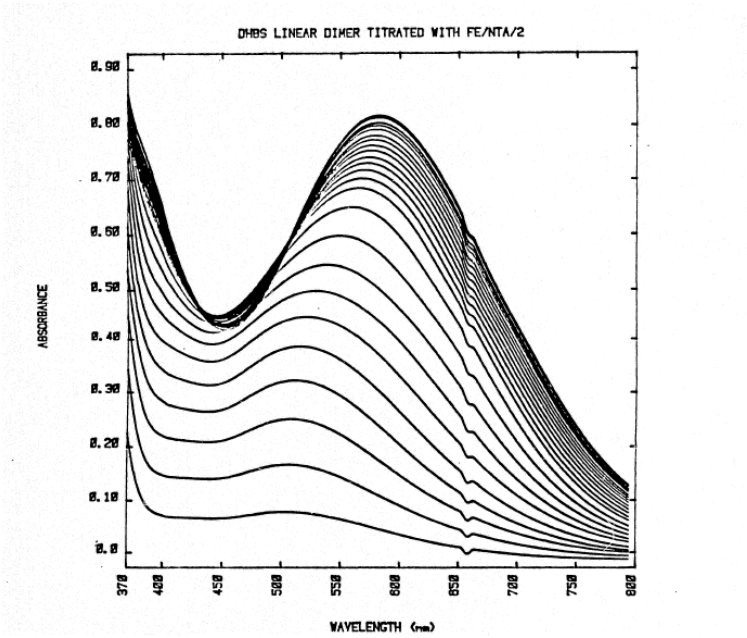


Figure 51 - Upper - Plot of the absorbance at 590 nm for solutions containing the DHBS monomer and linear dimer in the presence of increasing equivalents of Fe(III). Solutions were buffered at pH 9 and ferric ion was added as the NTA complex. Lower - UV-visible spectra between 370-800 nm recorded during the linear dimer titration. Reprinted with permission from *Inorg. Chem.*, 1991, *30* (5), 900-906. Copyright 1991 American Chemical Society.[8]

Scarrow *et al.* also noted that during a spectrophotometric titration of the 1:1 ferric complex of H_4-**17** across pH ranges of 8.0-7.0 another bathochromic shift is observed from λ_{max} of 500 nm at pH 8.0 to a λ_{max} of 540 nm at pH 7.0 (Figure 52). This shift was explained by hydrolysis of one of the two coordinating water molecules assumed to be coordinating in the two remaining iron coordination sites. This shift, which is associated with pH change, cannot be the cause of that seen in Figure 51, as the iron titration was buffered at pH 9.0, which would only allow for minimal pH change during the experiment. The conclusion that H_4-**17** only forms a 1:1 complex in the presence of iron(III) is likely to be an over-simplification of the system.

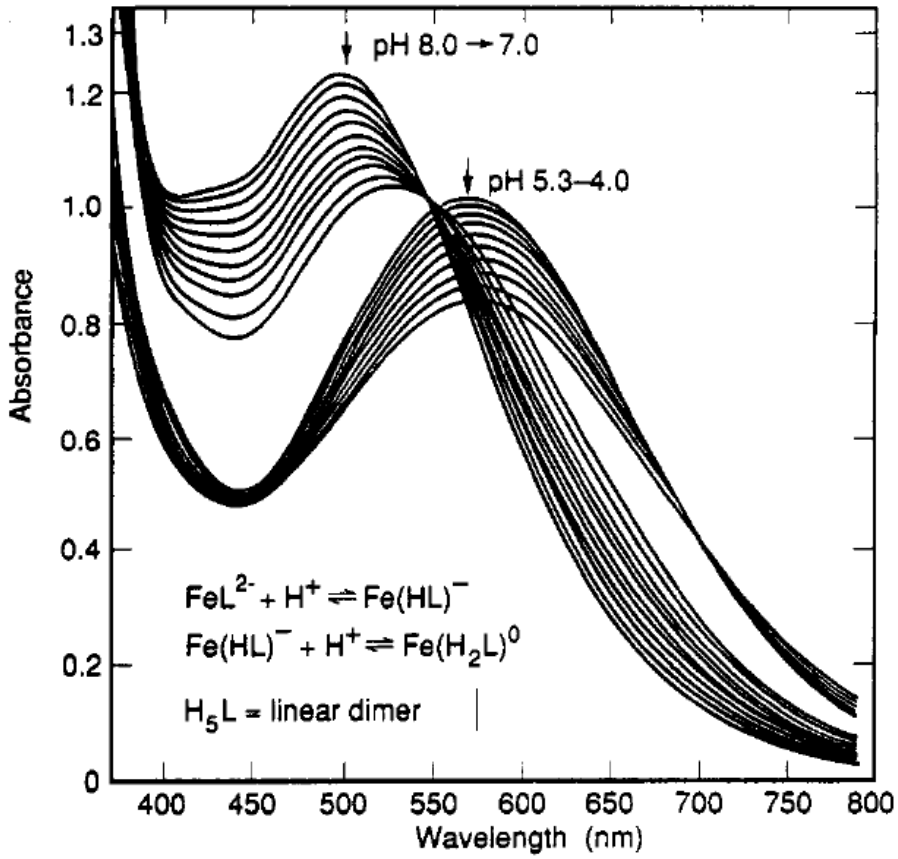


Figure 52 - Spectrophotometric titration of the 1:1 ferric complex of H_4-**17** in dilution corrected absorbance units. Reprinted with permission from *Inorg. Chem.*, 1991, *30* (5), 900-906. Copyright 1991 American Chemical Society.[8]

Another investigation of the coordination chemistry of H_4-**17** was conducted by Bergstrom *et al.*, who explored the change in the ^{1}H and ^{13}C NMR signals of the

linear dimer upon chelation of gallium(III). They showed that upon chelation of gallium by H_4-**17** (ratio not specified), a simple ^{1}H NMR spectrum is observed, where the signals of the aromatic protons and the protons situated on the α-carbon all shift upfield, while the two diastereotopic, geminal protons situated on the β-carbon shift differently to each other, one upfield and the second downfield.[7] The authors stated that it is not clear whether H_4-**17** forms a triple stranded 3:2 complex, similar to that formed by rhodotorulic acid[29], rather than a 3:2 mono-bridged dimer, such as that formed by alcaligin[30] (Chapter 2), although the complex is more likely to have formed the triple stranded complex, due to the observation of a simple spectrum, indicating that the ligands are equivalent.[7] The possibility of H_4-**17** forming a 1:1 complex was not addressed, assuming that H_4-**17** must form a 3:2 complex to satisfy the octahedral binding geometry of iron(III), although this contradicts the conclusion of Scarrow *et al.* who claim that H_4-**17** predominantly forms a 1:1 complex.[8] It is still unclear which coordination complex the tetradentate bis(catecholate) ligand H_4-**17** preferentially forms under biologically relevant conditions.

The coordination chemistry of a structurally similar bis(catecholate) siderophore family known as the amonabactins has also been studied.[31] The amonabactins are a series of bis(catecholate) siderophores that are composed of tri- or tetrapeptides with the general formula (gly)-(L)-lys-(L)-lys-(D)-aro, where the aromatic residue is either a tryptophan or phenylalanine (Figure 53).[32-33]

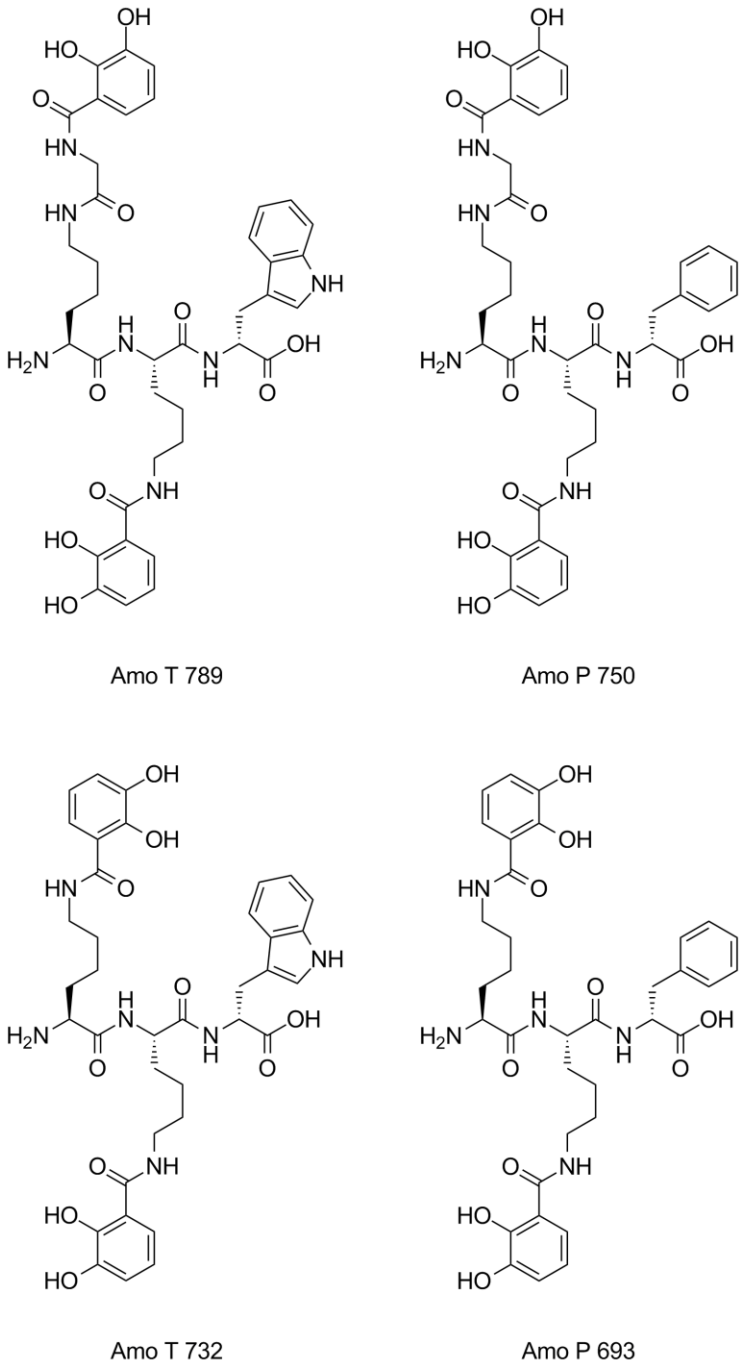


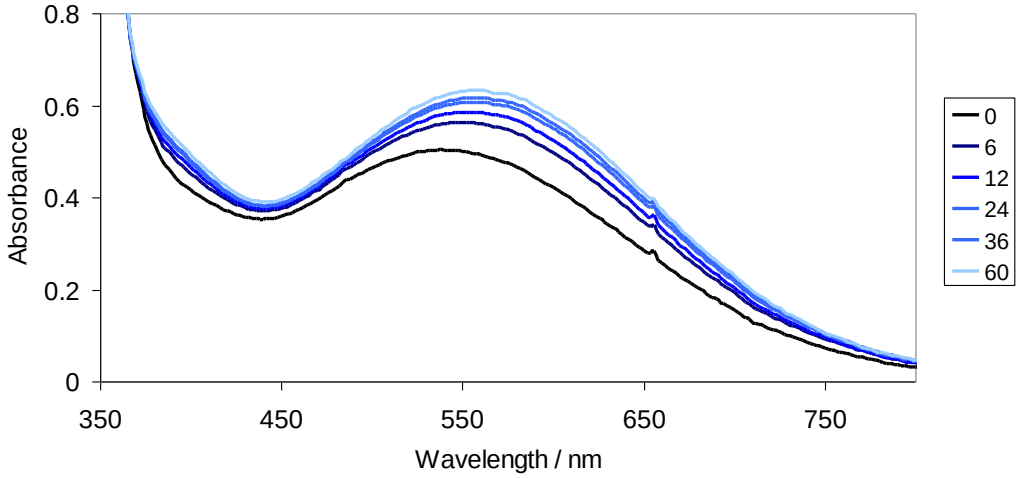
Figure 53 - Amonabactins. Names refer to the aromatic amino acid constituent and molecular weight.

This family of siderophores can interchange their coordination chemistry depending on pH and relative concentration of ligand to metal.[31] At high pH with an excess of ligand a 3:2 ligand to metal complex dominates, whereas at neutral and low pH, they preferentially form 1:1 bis(catechol)iron(III) complexes.[31] These studies by Scarrow *et al.*[8] and Telford *et al.*[31] highlight that the coordination chemistry of these bis(catechol) siderophores is non-trivial, and

a variety of complexes can form under different conditions. This study will focus on the complexes formed by H_4-**17** with iron(III) at neutral pH and under biologically relevant conditions.

Investigation of 17^{5-} to Iron(III) Ratios using the Continuous Variation Method

The investigation of the coordination chemistry of H_4-**17** was carried out under the same conditions as those reported for H_4-4-LICAM (Chapter 2). A kinetic experiment was set up as for H_4-4-LICAM, to investigate the time required for the samples obtained by mixing H_4-**17** with iron(III) in a 1:1 ratio to reach equilibrium. H_4-**17** exhibited different behaviour to H_4-4-LICAM and did not show a notable change in colour over time at close to a 1:1 metal-to-ligand ratio. The experiment showed a bathochromic shift in λ_{max} (542 nm to 563 nm), with the intensity of the LMCT band increasing and then stabilising after 60 seconds (Figure 54). The reason for the bathochromic shift in λ_{max} is not clear. As the system has not reached equilibrium, presumably there are coordination complexes with different speciation (particularly solvation) affecting the observed absorbance. This showed that the equilibrium is reached relatively quickly which allowed the Job plot to be conducted in the same manner as for H_4-4-LICAM.



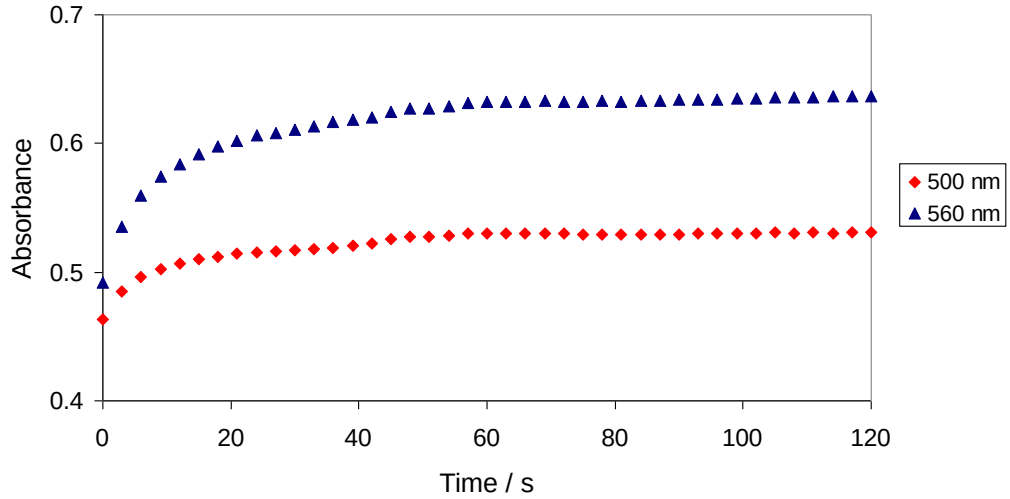


Figure 54 - Upper - Selected UV-visible absorbance spectra recorded during the kinetic run with H_4-**17** (2.0×10^{-4} M) and iron(III) (1:1 ratio). Spectra ordered from black (0 seconds) to light blue (60 seconds) for selected time intervals. Lower - Kinetic trace of upper, following absorbance at 500 nm (red diamonds) and 560 nm (blue triangles) over 2 minutes. Spectra recorded every 3 seconds, solution containing 0.1 M TrisHCl pH 7.5, 5% DMSO at 20 °C.

The UV-visible spectra of ferric-**17** (Figure 55) show two λ_{max} values at 512 nm and 563 nm when varying the ligand to metal ratio, as found in the UV-visible data for ferric-4-LICAM.

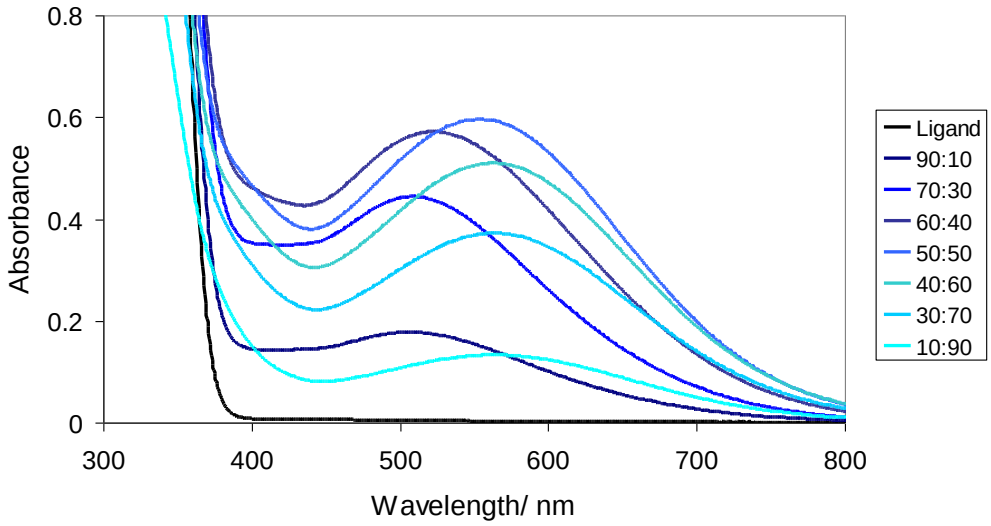


Figure 55 - Selected UV-visible absorbance spectra of H_4-**17** and Fe^{III}(NTA), used for the Job plot analysis Spectra recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO. Spectra ordered from black (H_4-**17** excess) to light blue ordered by ligand to metal ratio; 90:10, 70:30, 60:40, 50:50, 40:60, 30:70, 10:90). [M] + [L] = 0.4 mM.

Job plot analysis carried out at these two wavelengths (512 nm and 563 nm) indicates two different binding ratios, however, unlike ferric-4-LICAM, the ratios for ferric-**17** are closer together (Figure 56). The analysis at 563 nm shows the expected 1:1 ratio, however, the analysis for 512 nm shows a stoichiometry between the 3:2 and 1:1 ratios. This shift suggests that $\mathbf{17}^{5-}$ (extra charge due to deprotonated carboxylate group at pH 7.5) prefers a ligand-to-metal binding ratio of 1:1, but is still capable of coordinating in the 3:2 mode if an excess of ligand is present. The preference for the 1:1 complex could be influenced by a number of factors, such as the stereochemistry of the backbone, prearranging the catechol binding groups away from the linear arrangement required for triple helicate formation. The nature of the backbone could also influence the binding ratio, as the linear dimer possesses a polar linker with the ester which can form hydrogen bonds in the aqueous solution, rather than a hydrophobic one, which could self assemble in order to reduce contact with the polar solvent, potentially favouring the 3:2 assembly. The accumulation of negative charge could also disfavour the 3:2 complex for ferric-**17**, as the free carboxylate group on the backbone will also add negative charge to the complex,

with the resulting 3:2 complex of ferric-**17** formally having a negative charge of nine rather than negative six for ferric-4-LICAM.

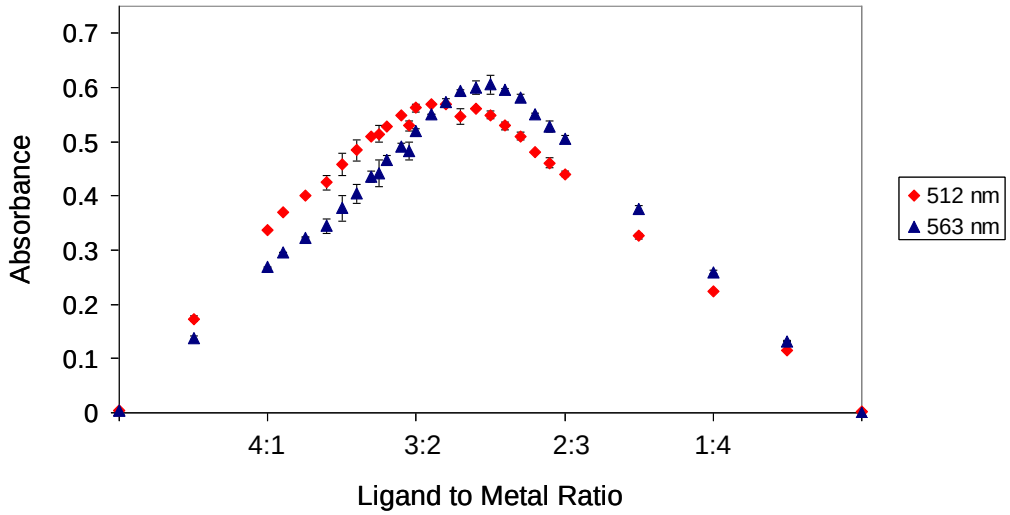


Figure 56 - Job plot for the binding of H_4-**17** and Fe^{III}(NTA), obtained by following the absorbance at both 512 nm (red diamonds) and 563 nm (blue triangles). The absorbance values are averages of two experiments and the error bars indicate the differences between the runs. Plot recorded in 0.1 M TrisHCl pH 7.5, 5% DMSO.

The Job plot shows that H_4-**17** can adopt two binding stoichiometries with iron(III), much like H_4-4-LICAM, depending on the relative concentration of H_4-**17** and iron(III) in solution. Unlike H_4-4-LICAM, the Job plot of H_4-**17** suggests that a 1:1 binding stoichiometry is preferred over 3:2, suggesting it has similar solution behaviour to the amonabactins, although they possess a much longer linker between the coordinating functional groups.

Circular Dichroism Spectroscopic Study of the H_4-17 and Iron(III) System

Due to the chirality of the ligands (amonabactins, enterobactin and the enterobactin hydrolysis products), intramolecular chiral induction influences the metal-centred chirality of the complexes which can be examined using circular dichroism. Enterobactin is known to form mainly Δ-configured complexes, with iron(III),[34] chromium(III),[35] vanadium(IV)[36-37] and rhodium(III).[38]

Scarrow *et al.* investigated the coordination chemistry of the enterobactin hydrolysis products and showed that the linear trimer forms a iron(III) complex with Δ-configuration with a CD spectrum similar to that of ferric enterobactin. The linear dimer (1:1 ratio with iron(III)) was essentially featureless in the visible region, which was attributed to an equilibrium between equal amounts of Λ- and Δ-configured complexes (Figure 57).[8] All three spectra produce a positive band due to a n→π* transition at approximately 320 nm, which was interpreted as a ligand-based transition that increases in intensity upon metal binding.

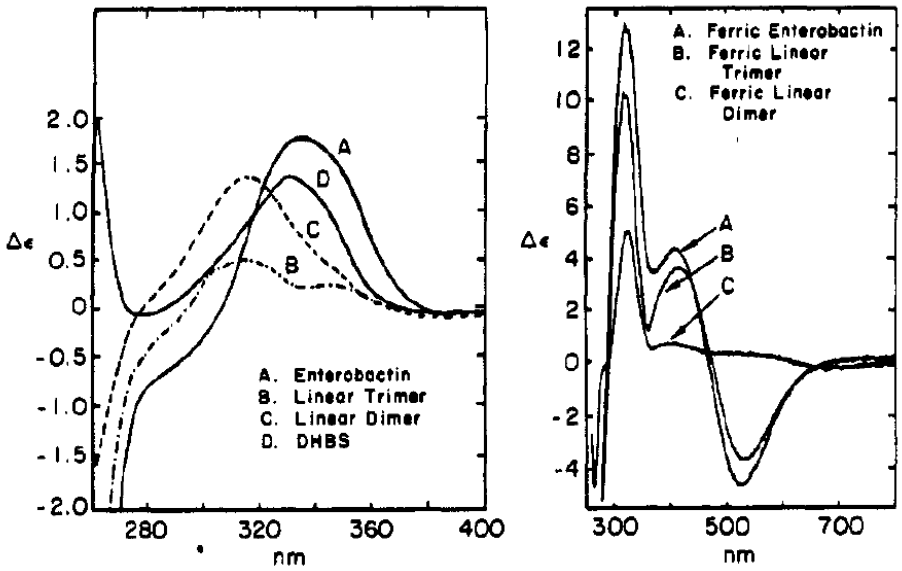


Figure 57 - Circular dichroism spectra of the free ligands (left) and the iron(III) complexes (right) of enterobactin and the hydrolysis products in aqueous solution at pH 7.5, 0.1 M HEPES buffer. Reprinted with permission from *Inorg. Chem.*, 1991, *30* (5), 900-906. Copyright 1991 American Chemical Society.[8]

The CD spectra of the ferric-amonabactins have also been studied and show that the 1:1 ferric-amonabactins are present in a racemic mixture of Λ- and Δ-configured complexes while the 3:2 ligand to metal ratio shows a weak preference for the Δ-configuration.[31]

Job plot analysis suggests that the coordination chemistry behaviour of H_4-**17** is similar to that of the amonabactins, showing a change in coordination geometry

under different conditions, although the original study by Scarrow *et al.* did not record the CD spectrum of the H_4-**17** with iron(III) in a 3:2 ligand to metal ratio.

To determine whether H_4-**17** behaves similarly to the amonabactins, the CD spectra of ferric H_4-**17** were recorded at ligand-to-metal ratios of 3:2, 1:1 and 2:3. Samples were prepared in the same manner as for the continuous variation experiments, (0.1 M TrisHCl pH 7.5, 5% DMSO). The samples were given time to reach equilibrium before spectra were recorded (Figure 58).

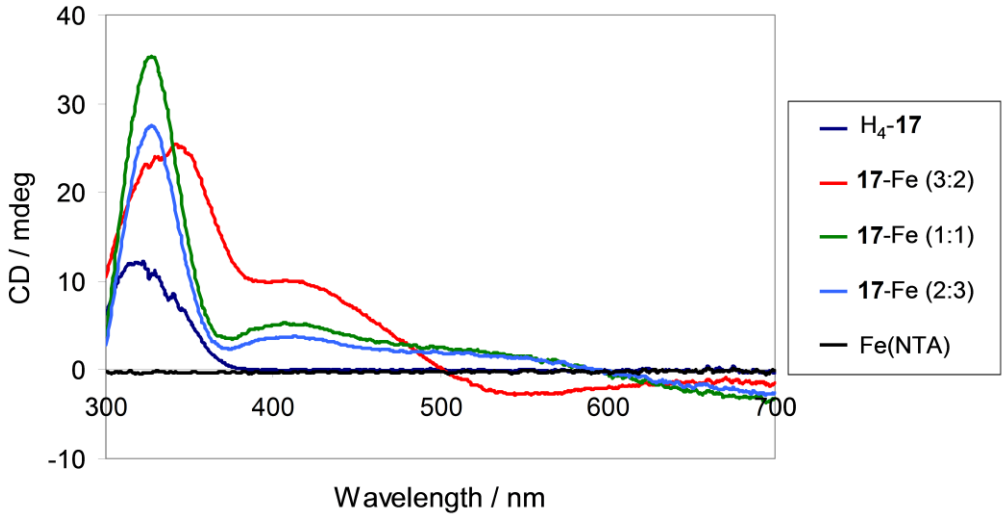


Figure 58 - Circular dichroism spectra of H_4-**17** and ferric-**17** in key ligand-to-metal ratios in aqueous 0.1 M TrisHCl buffer, pH 7.5, 150 mM NaCl, 5% DMSO. [L] + [M] = 0.4 mM.

The CD spectrum of H_4-**17** in the near UV region shows a positive feature arising from an n→π* transition at approximately 320 nm, which grows in intensity upon metal binding. When the ligand to metal ratio is 1:1 or 2:3, the spectra are consistent with that of the ferric linear dimer reported by Scarrow *et al.*, with a large positive feature at 320 nm and a weak, broad, positive feature between 380-610 nm.[8] When the ligand to metal ratio is 3:2, the large positive feature around 320 nm in H_4-**17** is shifted to 340 nm, and the intensity of the weak, broad feature increases, becoming more positive around 415 nm and more negative at 550 nm, with a zero point at 500 nm giving a negative Cotton effect.

This change in CD spectra also coincides with a colour change from purple to red, consistent with the formation of a tris-catecholate system. This indications a change in mode of binding of H_4-**17** towards iron(III), with a clear preference for the Δ-configuration.

The behaviour of the CD spectra at different to ligand to metal ratios resembles that of the amonabactins, but the intensity of the change is more profound with H_4-**17**. This may be due to the distance between the binding catechol groups and the stereo-centre in the ligand. The iron binding catecholamide units of H_4-**17** are directly linked to the stereocentre of the siderophore backbone, whereas in the amonabactins the catecholamide is linked to the side chain amine of the lysine backbone, which has a four carbon spacer between the catecholamide and the stereocentre, potentially reducing the influence of the stereocentre upon metal binding.

In conclusion, H_4-**17** binds both in a 1:1 and a 3:2 fashion, with the relative concentrations of H_4-**17** and iron(III) determining the binding mode. On comparison with H_4-4-LICAM, the 1:1 binding mode seems to be favoured over 3:2, influenced by the backbone of the ligand, due to the stereochemistry, or its electronic character. The CD spectra indicate that ferric complexes of H_4-**17** preferentially bind in the Δ-configuration, however, when bound in a bis-catecholate 1:1 ratio, an interchange between the Δ-configuration and the Λ-configuration is observed. The coordination chemistry of H_4-**17** appears similar to that of the amonabactins, a family of tetradentate catecholamide siderophores.

3.4 Interactions of the Ferric Complex with CeuE

Following the investigation of the solution behaviour of H_4-**17** in the presence of iron, the aim was to examine how ferric-**17** interacts with the ferric-siderophore transport PBP CeuE. No co-crystal structures of enterobactin hydrolysis products with a periplasmic binding protein have been determined. A crystal structure of the human protein siderocalin (Scn), a protein associated with the innate immune response towards siderophores (but has no structural similarity to

CeuE), was co-crystallised with enterobactin and published by Goetz *et al.* (PDB code 1L6M).[39] The tri-lactone backbone of enterobactin had been partially hydrolysed during crystal formation and was therefore modelled as different enterobactin hydrolysis products in each of the binding pockets of the three Scn chains in the asymmetric unit. The iron atoms in the binding pocket associated with chain A and chain B are fully coordinated with three catechol units, modelled as one DHBS and two dihydroxybenzoic acid units, respectively. The binding pocket associated with chain C was modelled as a DHBS and a dihydroxybenzoic acid bound as two independent bidentate ligands, leaving two free coordination sites (Figure 59). The nearest amino acid residues are a tyrosine and a lysine, which are 4.0 Å and 3.9 Å from the iron centre, respectively, therefore too far away from the metal centre for direct coordination. Each of these binding pockets is unique, with different arrangements of the three iron(III)-coordinating ligands found in the binding pockets of chains A and B. The cause of the difference in coordination geometry of the ferric-complexes for the three different chains in the asymmetric unit is unclear.

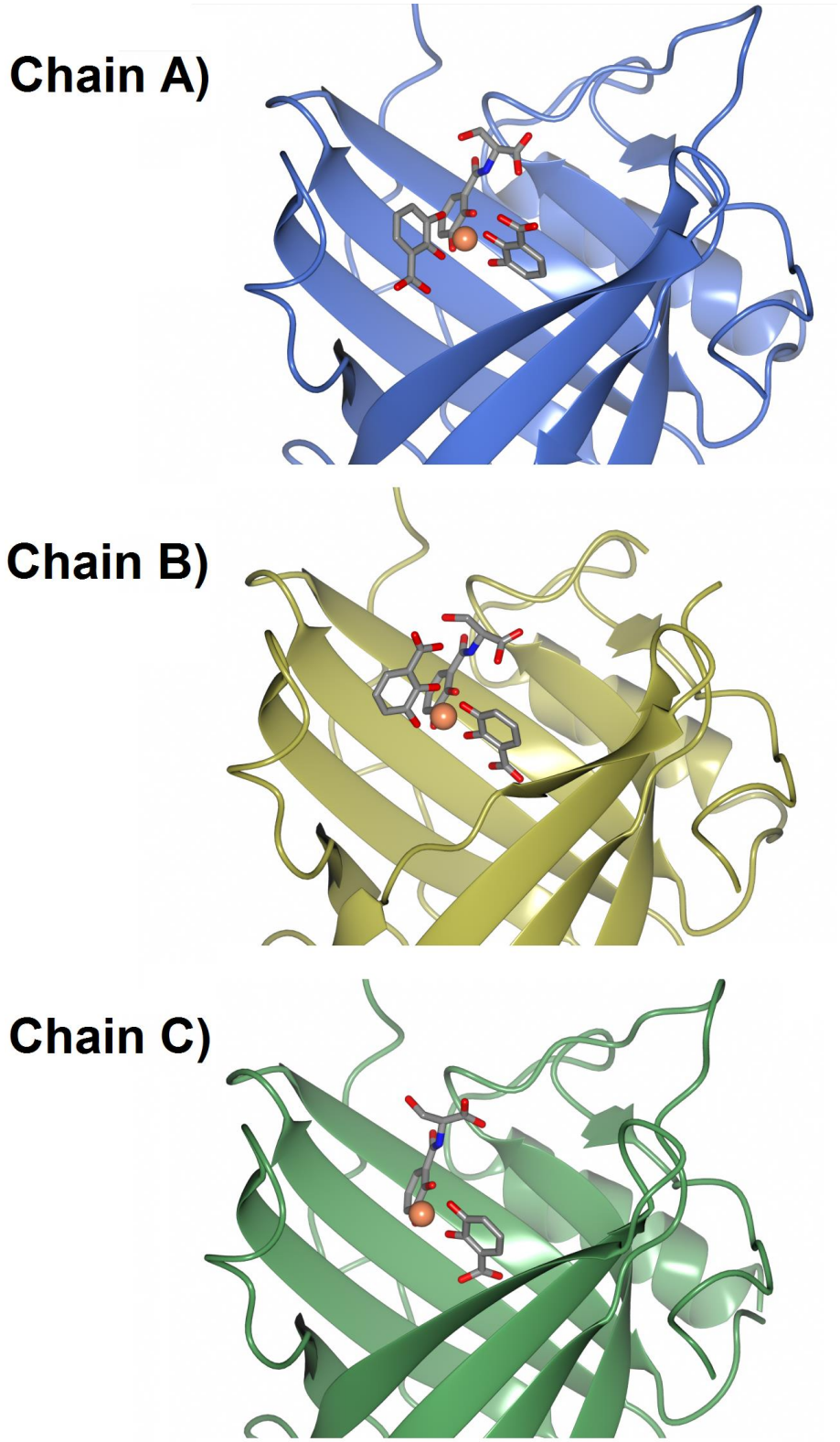


Figure 59 - Ribbon representation of three monomers of Scn co-crystallised with ferric-enterobactin. Key: Chain A (light blue), chain B (gold), chain C (green); ligands shown as cylinders, grey - carbon, blue - nitrogen, red - oxygen, coral - iron. PDB code: 1L6M.[39]

The same research group also deposited a mutant Scn structure in the PDB, however, this structure has never been published (PDB code 3I0A). This structure shows ferric-**17** and other enterobactin hydrolysis products bound to Scn. This structure is similar to the 1L6M structure as it is found in the same space group ($P4_12_12$), with three Scn monomers in the asymmetric unit, each presenting a different ferric-ligand complex in the binding pocket of the protein (Figure 60). To the best of my knowledge, this structure is the only example deposited on the PDB which contains H_4-**17** as a ligand. Scn has also been co-crystallised with vibriobactin,[40-41] carboxymycobatin S and T,[42] ferric-dihydroxybenzoic acid,[43] and ferric catechol,[44] and a complex of Scn with ferric-enterobactin has been solved *via* solution NMR.[45]

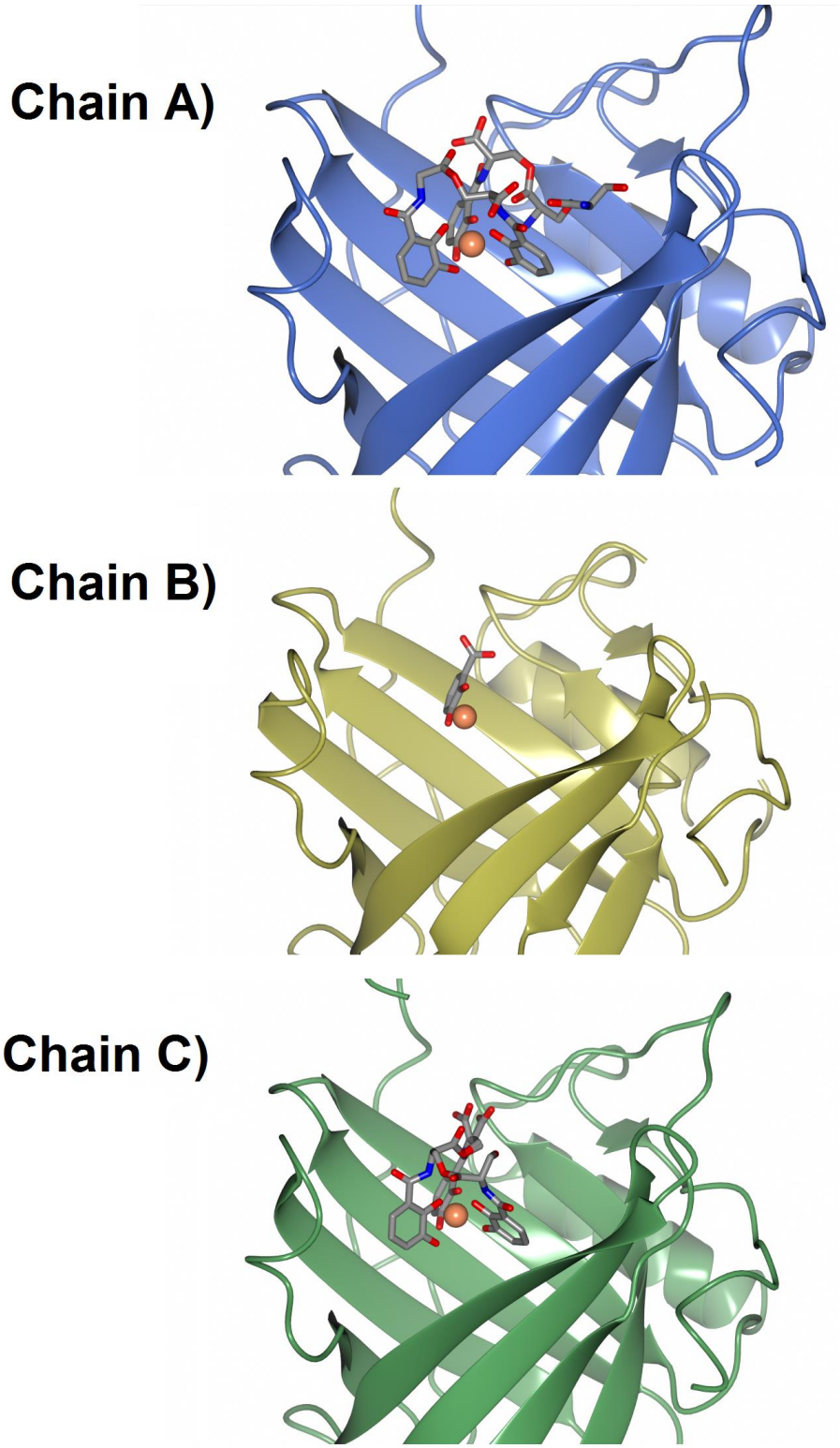


Figure 60 - Ribbon representation of three monomers of Scn co-crystallised with ferric-enterobactin. Key: chain A (blue), chain B (gold), chain C (green); ligands shown as cylinders, grey - carbon, blue - nitrogen, red - oxygen, coral - iron. PDB code: 3I0A (unpublished work).

The co-crystal structures reported by Goetz *et al.* highlight difficulties in co-crystallising ligand prone to hydrolysis under aqueous conditions required for proteins, showing that the crystallisation time is required to be shorter than the rate of hydrolysis of the ligand in the presence of the protein.

Crystallisation of the CeuE, Ferric-Linear Dimer Complex

Initial attempts to crystallise ferric-**17** with CeuE used the same co-crystallisation method as for ferric-4-LICAM, but crystals of X-ray diffraction quality were not obtained. Soaking was therefore used to introduce ferric-**17** into an apo-CeuE crystal, grown from the PACT crystal screen. H_4-**17** and iron(III) were added to the crystallisation well, followed by thorough mixing. 2 μL of the resulting solution containing ferric-**17** was added the sitting drop that contained the crystal. The well was re-sealed and time allowed for the ferric-**17** to diffuse through the crystal. The soaked crystals were collected at a number of time points after the introduction of the crystallisation solution containing ferric-**17**, frozen in liquid nitrogen, and data collected at the DLS.

Crystals were soaked for 90 minutes (form I, 0.2 M sodium fluoride, 0.1 M Bis Tris propane pH 8.5, 20 (w/v) PEG 3350), 24 hours (form II, 0.1 M MMT buffer pH 8.0 25% (w/v) PEG 1500) and 11 days (form III, 0.1 M PCB buffer pH 9.0 25% (w/v) PEG 1500). As expected each of the crystal structures was in the space group P1, with three CeuE chains in the asymmetric unit. The apo-CeuE structure (Chapter 2), excluding water molecules, was used as a starting model.

Form I showed no additional electron density in the CeuE binding pockets, indicating that 90 minutes are not enough time for the complex to fully soak into the lattice of the apo-CeuE. In contrast crystal forms II and III, did show density in the binding pockets of each of the three independent CeuE chains in the asymmetric unit. Crystal forms II and III were refined using REFMAC5[46] and Coot.[47]

The electron density map of crystal form II clearly shows **17**$^{5-}$ bound to the iron(III) centre in a 1:1 ratio in CeuE chains A and B. CeuE chain C was modelled with a partially occupied iron centre with no coordinating ligands, due to poor electron density in this binding (Figure 61 and Figure 62).

Form III was treated as crystal form II, but the electron density for the backbone of H_4-**17** was weaker than in form II, thus it was not clear whether the di-serine backbone had hydrolysed. Hence, the ligands were initially modelled as bidentate DHBS units bound to the iron(III) centres through the catechol functionality. During refinement it was clear that the coordinating groups bound to the three iron(III) centres were different. The binding pocket of chain A showed little electron density linking the two DHBS monomers, and as a result the modelled coordinating ligands were kept as two DHBS monomers representing hydrolysed **17**. Chain B was initially similar to chain A, however during refinement, the DHBS monomers showed electron density linking the two serine units, and the ligand was re-modelled to be ferric-**17**. Chain C was similar to chain A, but the coordination of one of the DHBS monomers was found in the opposite orientation, with the serine backbone pointing inwards into a side pocket of the CeuE chain (Table 9, Figure 63 Figure 64).

The nomenclature of the two crystal forms refers to the number of CeuE molecules found within the unit cell, followed by the ligand species bound by the three CeuE molecules.

Chain	**{CeuE$_3$[Fe(17)]$_2$[Fe]} Form II**	**{CeuE$_3$[Fe(17)][Fe(DHBS)$_2$]$_2$} Form III**
A	Linear Dimer	2 x DHBS
B	Linear Dimer	Linear Dimer
C	Fe	2 x DHBS

Table 9 - Summary of modelled ligands in the binding pockets of both crystal forms, {CeuE$_3$[Fe(**17**)]$_2$[Fe]} (crystal form II) and {CeuE$_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} (crystal form III).

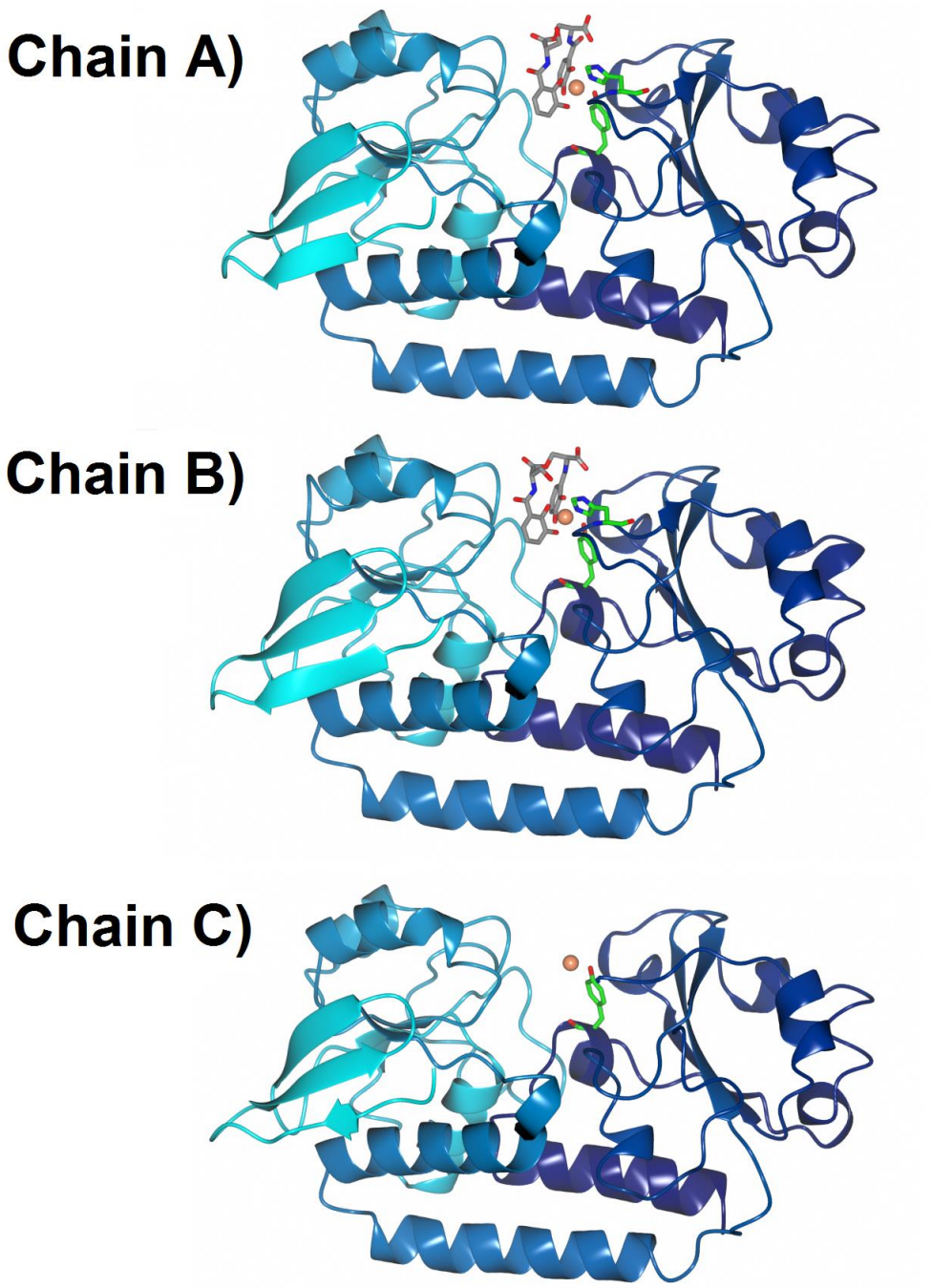


Figure 61 - Ribbon representation of the three independent CeuE chains of {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) with colours blended from cyan, N-terminus, to dark blue, C-terminus. His228, Tyr288 and Fe-**17** shown as cylinders. Key: His227 and Tyr288; green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.

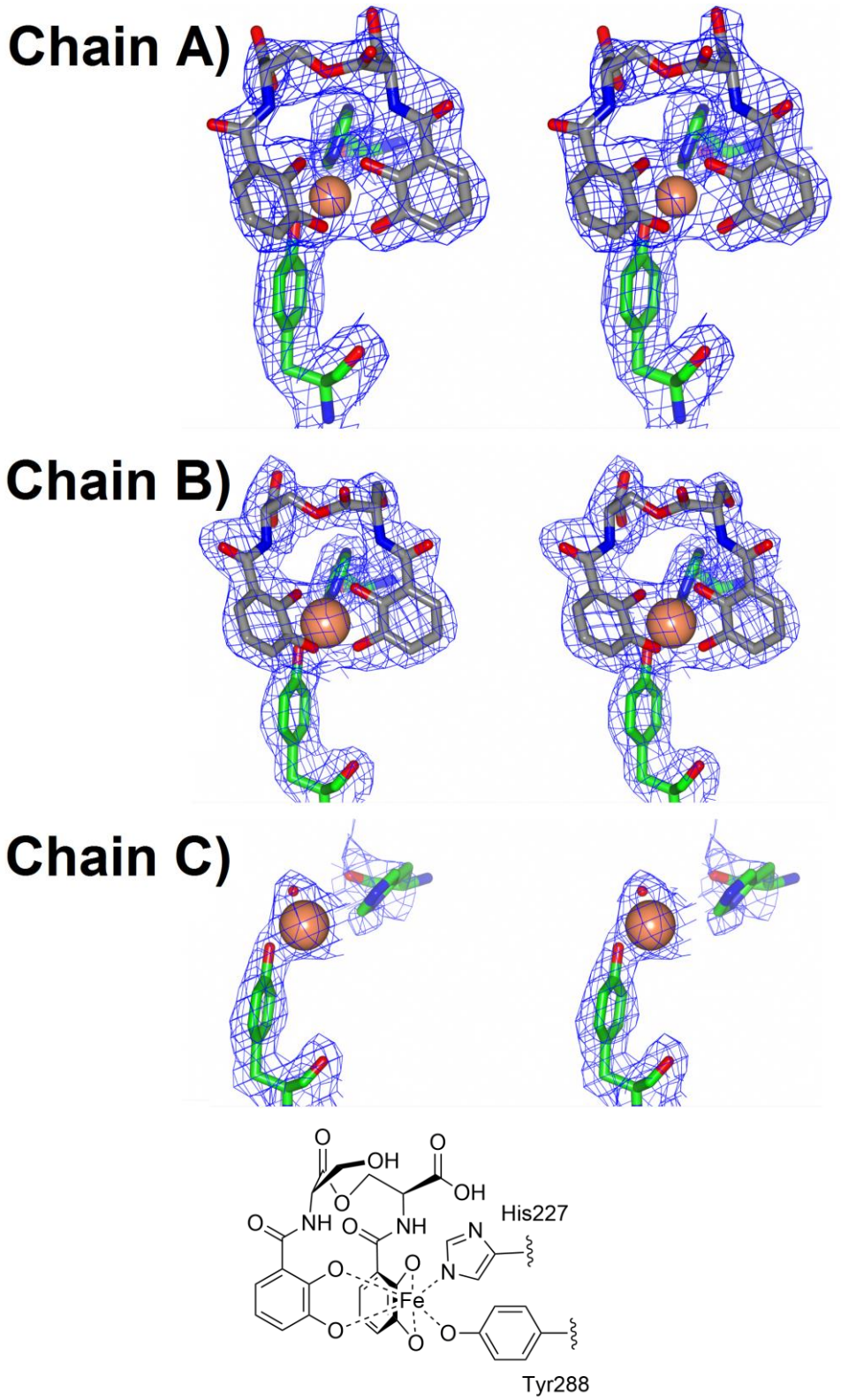


Figure 62 - Cylinder representation of the binding pockets of {CeuE$_3$[Fe(**17**)]$_2$[Fe]} (form II) with electron density map shown as 2F$_{obs}$-F$_{calc}$, contoured at 1σ and chemical representation showing key interactions of Fe-**17**. Key: His227 and Tyr288; green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.

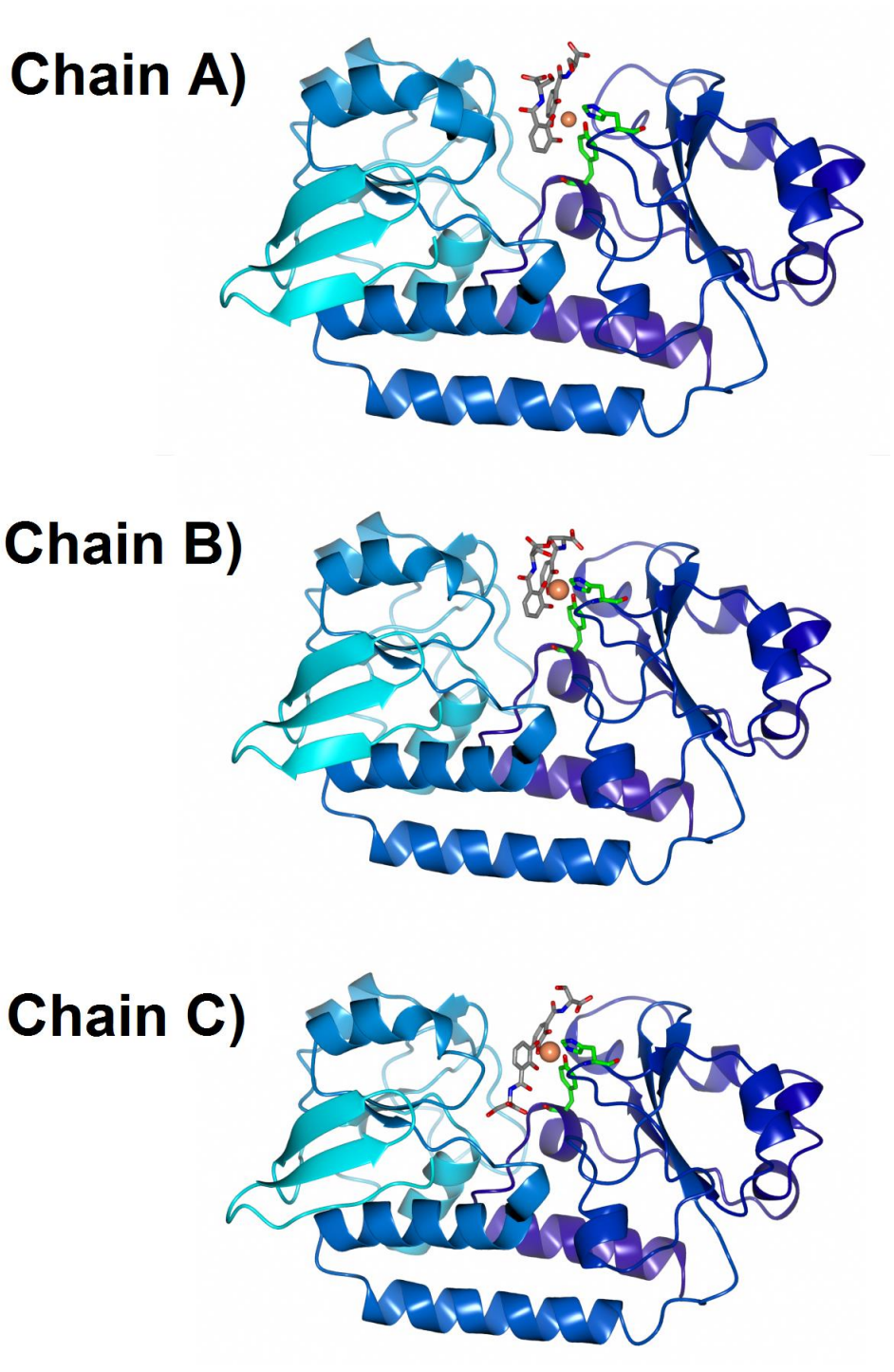


Figure 63 - Ribbon representation of the three independent CeuE chains of {$CeuE_3$[Fe(**17**)][Fe $(DHBS)_2]_2$} (form III) with colours blended from cyan, N-terminus, to dark blue, C-terminus. His227, Tyr288, Fe-**17** and Fe-$DHBS_2$ shown as cylinders. Key: His228, Tyr288, green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.

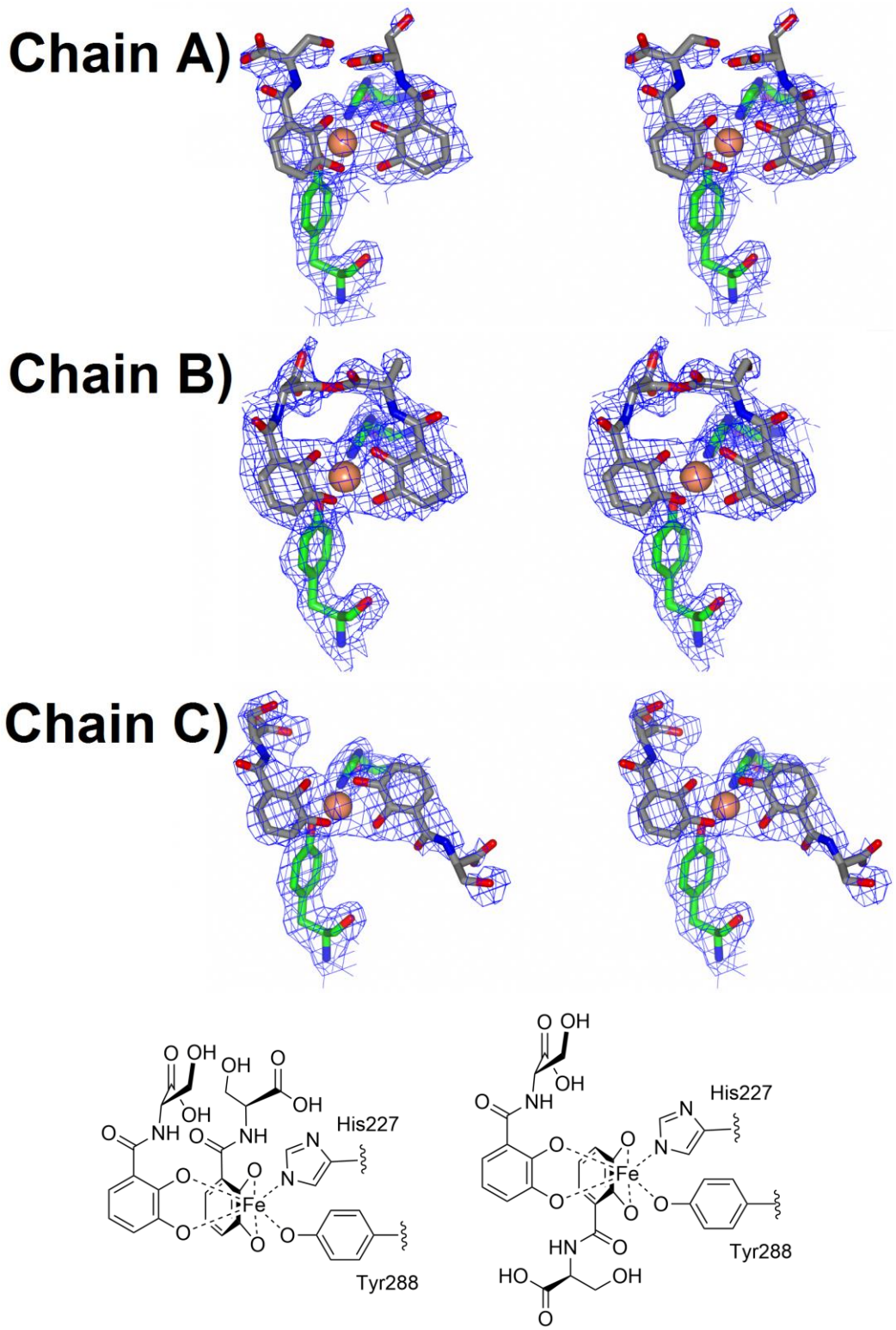


Figure 64 - Cylinder representation of the binding pockets of {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} (form III) with electron density map shown as $2F_{obs}$-F_{calc}, contoured at 1σ and chemical representation showing key interactions of the two Fe-DHBS environments. Key: His227 and Tyr288; green - carbon, blue - nitrogen, red - oxygen; Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.

	{CeuE$_3$[Fe(17)]$_2$[Fe]} Form II	**{CeuE$_3$[Fe(17)][Fe(DHBS)$_2$]$_2$} Form III**
Data collection		
Diffraction Source	DLS Beamline i04	DLS Beamline i02
Wavelength / Å	0.976	0.979
Temperature / K	100	100
Crystal Data		
Space Group	P1	P1
a, b, c / Å	58.46, 63.02, 67.01	58.07, 63.09, 67.16
α, β, γ / °	83.36, 76.57, 78.24	83.09, 76.90, 79.21
Resolution / Å	65.02 - 1.68	65.2 - 1.68
Unique Reflections	70279	58556
Completeness / %	98.1	94.3
Redundancy	4.4	1.9
R_{merge}	0.09	0.09
Mean [I/σ (I)]	17.5	5.2
Refinement Statistics		
Reflections, working	66562	48582
Reflections, test	3495	2510
Resolution Range / Å	65.01 - 1.90	65.19-2.10
R-factor	0.226	0.223
R_{free}	0.267	0.262
Mean B Factors / Å^2	42.3	31.3
RMS deviation from ideal		
Bond length / Å	0.015	0.013
Bond angles / °	1.782	1.688
PDB Code	-	-

Table 10- Data collection and refinement statistics of {CeuE$_3$[Fe(**17**)]$_2$[Fe]}.and {CeuE$_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} structures.

Both crystal form II, {$CeuE_3$[Fe(**17**)]$_2$[Fe]}, and crystal form III, {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$}, show that the binding pocket of CeuE chain C is distinctly different compared to that of chains A and B. In crystal form II, there is a low occupancy of the iron(III) atom and no clear density to model any other coordinating ligands, whereas in crystal form III, two DHBS monomers are bound to the iron(III) centre but in a different orientation than expected. This variation reflects different lattice contacts for the three independent binding pockets within the crystal. A crystal packing diagram shows that the apo-CeuE crystal lattice is tightly packed, which might suggest that the three binding pockets of each of the CeuE chains are likely to have different accessibilities (Figure 65). The binding pockets of chains A and B are open to the surrounding solvent, while access to that of chain C is partially blocked by the neighbouring N terminus of chain B. This can be demonstrated by the superposition Fe-**17** of chain B into the binding pocket of chain C, which demonstrates that encroachment of the N terminus of chain B into the binding pocket of chain C (Figure 66).

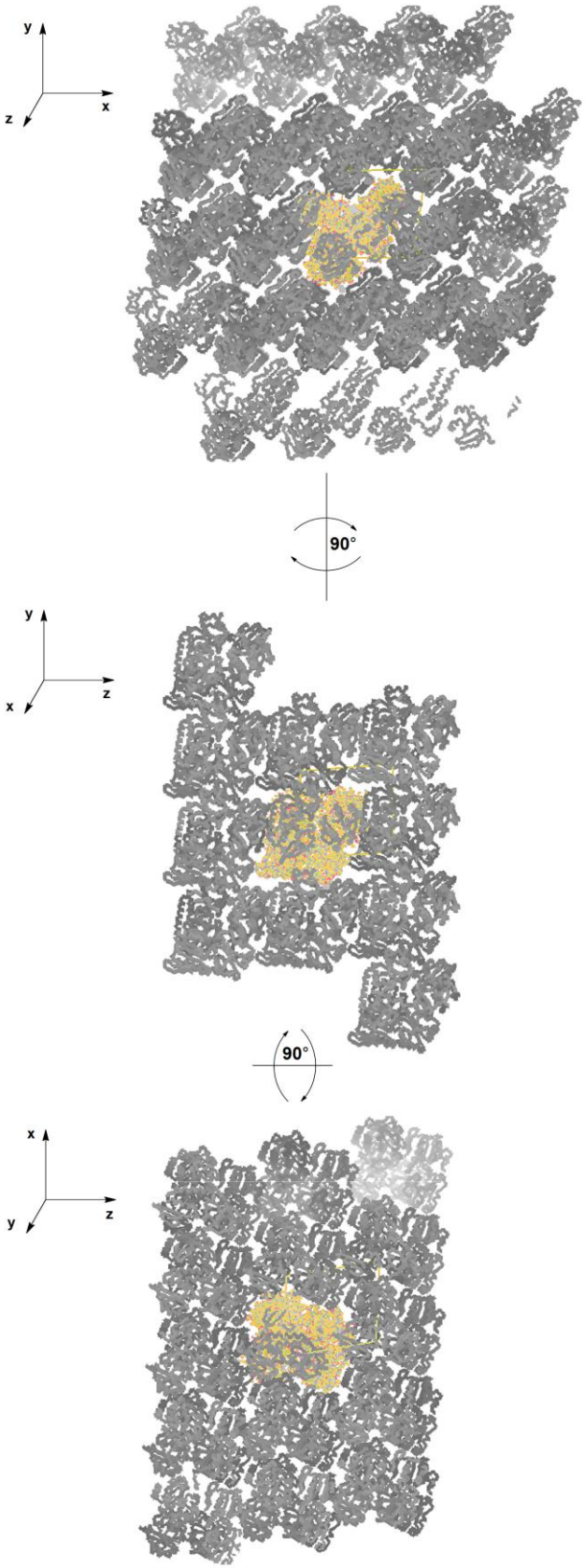


Figure 65 - Crystal packing diagram showing symmetry related CeuE molecules in the {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) crystal. Key {$CeuE_3$[Fe(**17**)]$_2$[Fe]} shown in gold, symmetry related molecules shown in grey. Image created using Coot.[47]

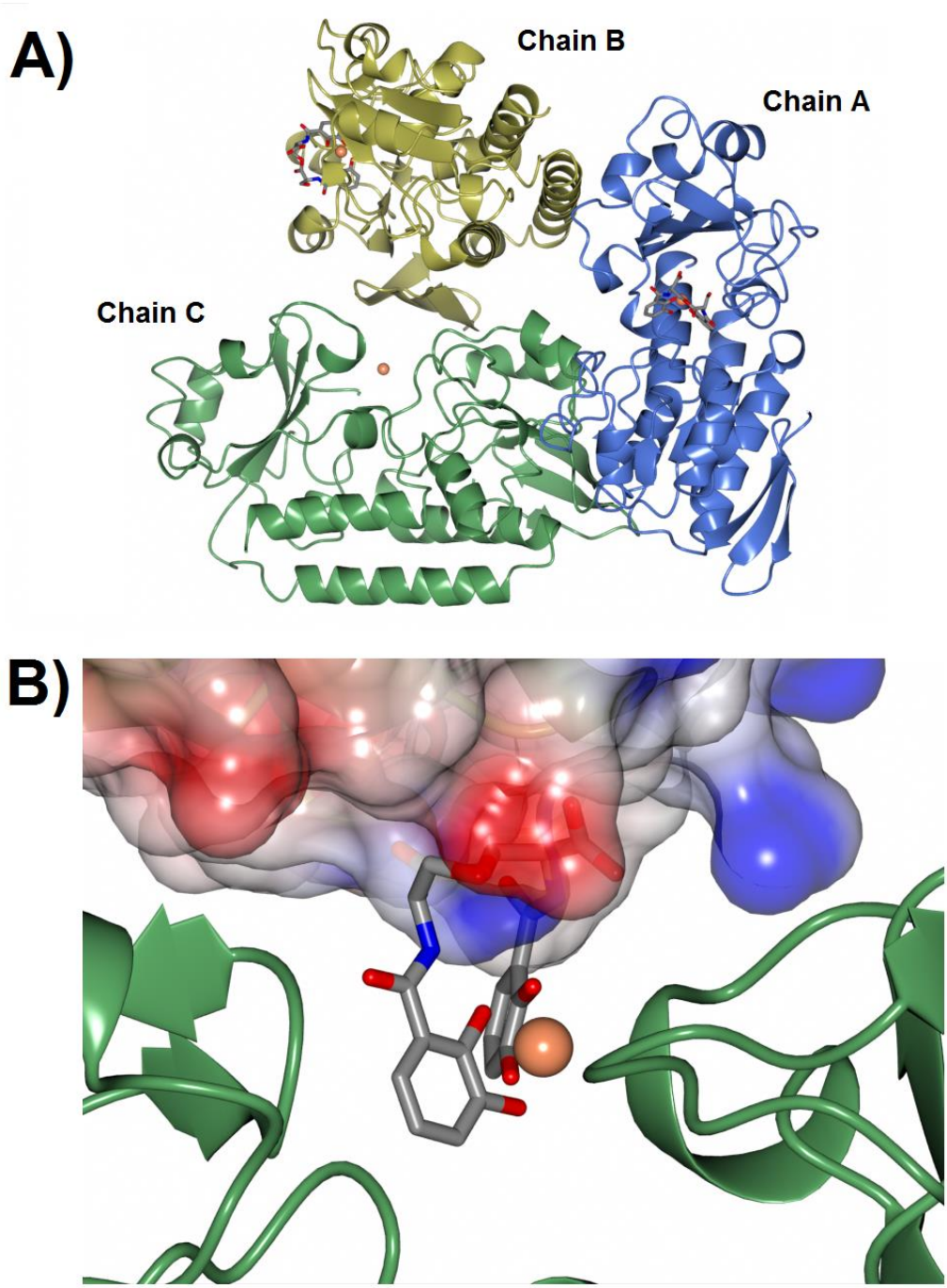


Figure 66 - A) Ribbon representation of the crystal packing in the asymmetric unit of {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) with His227, Tyr288, Fe-**17** shown as cylinders. B) Ribbon representation of binding pocket of chain C and surface (electrostatic representation) of N terminus of chain B, with Fe-**17** from chain B superimposed into the binding pocket. Key chain A (blue), chain B (gold), chain (dark green); Fe-**17**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.; electrostatic, red - negative, blue - positive.

In both structures, {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (chains A and B) and {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} (chain B) the same binding arrangement of the ferric-**17** to CeuE is observed, consistent with the {CeuE[Fe-4-LICAM]} structure. Ferric-**17** is bound in a 1:1 metal to ligand stoichiometry to a single CeuE monomer. In the {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} chains A and B contain the two ferric-DHBS units, which are bound in a 1:2 metal to ligand stoichiometry and mimic a tetracoordinate iron centre. The remaining two free coordination sites of the iron centres are occupied by the side chains of two amino acid residues, His227 and Tyr288, similarly to the {CeuE[Fe-4-LICAM]} structure (Chapter 2).

The bond lengths of the coordinating atoms to the iron centres in CeuE monomers A and B in the {$CeuE_3$[Fe(**17**)]$_2$[Fe]} structure and the bond lengths of the coordinating atoms to the iron centres in the {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} structure are shown in Table 11 and Table 12 respectively.

CeuE Chain	**Atom**	**Bond Length / Å**
A	**17**O_7 - Fe	2.2
	17O_8 - Fe	2.1
	17O_{36} - Fe	2.1
	17O_{37} - Fe	2.0
	TyrO - Fe	2.0
	HisN - Fe	2.6
B	**17**O_7 - Fe	1.8
	17O_8 - Fe	2.3
	17O_{36} - Fe	2.0
	17O_{37} - Fe	2.0
	TyrO - Fe	2.1
	HisN - Fe	2.3

Table 11 - Bond distances between coordinating atoms and the iron centres of monomers A and B in the {$CeuE_3$[Fe(**17**)]$_2$[Fe]} (form II) structure.

The bond lengths in both structures range from 1.8-2.4 Å, with the exception of the HisN-Fe in monomer A of both structures, which is 2.6 Å and 2.7 Å respectively. These bond lengths are too large to form a formal bond, however, the electron density around both histidine residues is poor due to possible restricted movement within the crystal upon binding of ferric-**17**. The poor electron density surrounding the histidine as a result of partial occupancy will increase the error on the bond lengths.

CeuE Chain	Atom	Bond Length / Å
A	$DHBS_1O_1$ - Fe	2.4
	$DHBS_1O_2$ - Fe	2.2
	$DHBS_2O_1$ - Fe	2.2
	$DHBS_2O_2$ - Fe	2.1
	TyrO - Fe	1.9
	HisN - Fe	2.7
B	**17**O_7 - Fe	1.8
	17O_8 - Fe	2.1
	17O_{36} - Fe	2.1
	17O_{37} - Fe	2.1
	TyrO - Fe	2.0
	HisN - Fe	2.4
C	$DHBS_1O_1$ - Fe	2.1
	$DHBS_1O_2$ - Fe	2.4
	$DHBS_2O_1$ - Fe	2.0
	$DHBS_2O_2$ - Fe	2.4
	TyrO - Fe	2.1
	HisN - Fe	2.4

Table 12 - Bond distances between coordinating atoms and the iron centres for the $\{CeuE_3[Fe(\mathbf{17})][Fe(DHBS)_2]_2\}$ (form III) structure.

The ligands in both the A and B chains, in both structures are in a *cis*, *cis* configuration with regards to the O_{ortho}-O_{ortho}, O_{meta}-O_{meta} donor atoms of the two catecholate moeities. The two DHBS monomers in the binding pocket of chain C in the {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} structure, are bound in a *cis*, *trans* configuration, with the serine backbone of DHBS 1 in the opposite direction compared to binding pockets A and B (Figure 67).

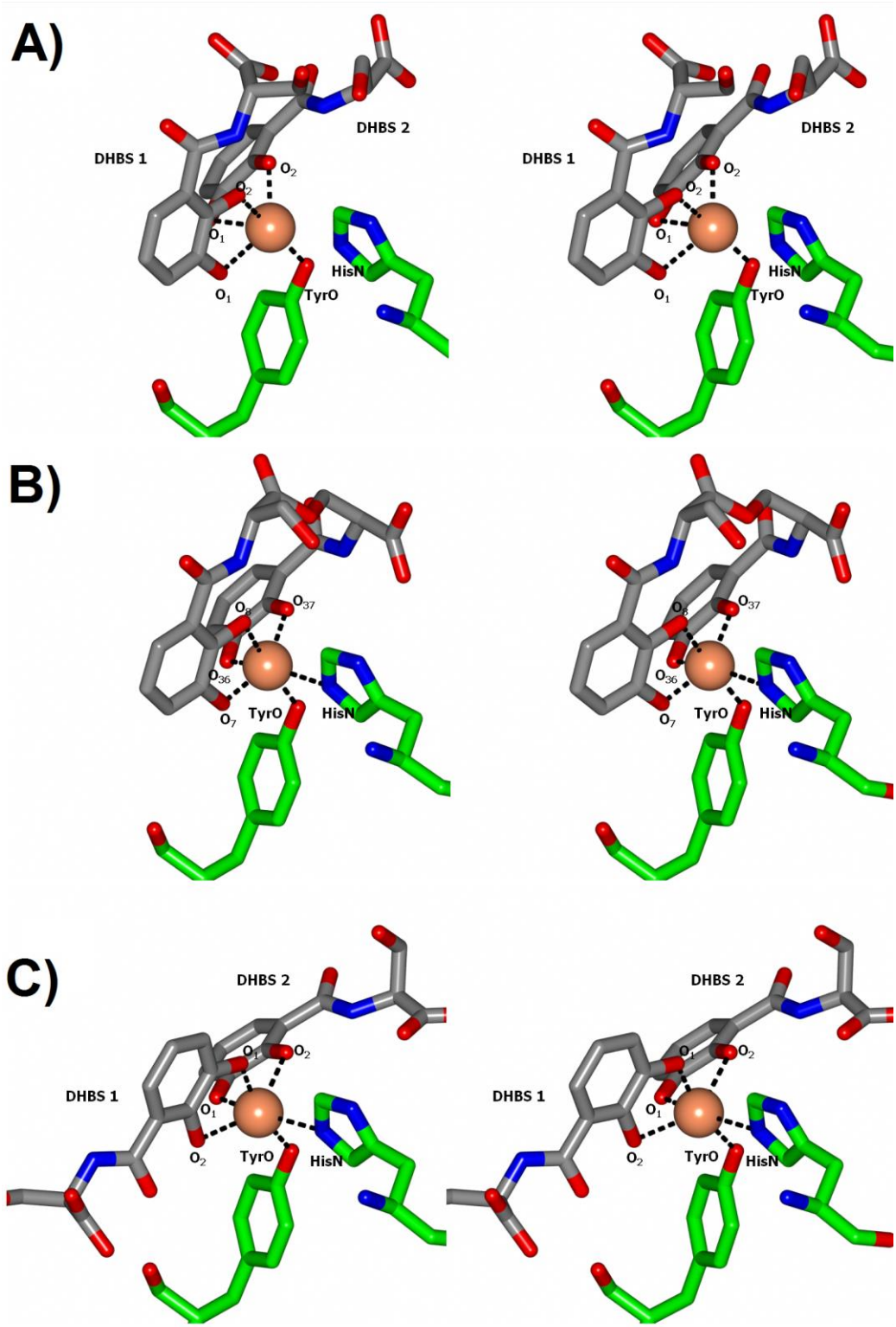


Figure 67 - Coordination environment of the three binding pockets of $CeuE_3[Fe(\mathbf{17})][Fe\ (DHBS)_2]_2\}$ (form III). Key; CeuE residues, green - carbon, blue - nitrogen, red - oxygen; DHBS and **4**, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.

The change in coordination environment of the two DHBS monomers in chain C has the consequence of directing the serine backbone towards the surface of the protein and occupies a pocket adjacent to the binding site of CeuE. The different orientation of the serine backbone in the two DHBS monomers is a clear indication that the hydrolysis must have occurred prior to the binding, as the carboxylic acid and alcohol groups in the monomers are too far apart to linked together by the ester bond.

There are two possible explanations for the change in orientation of the DHBS monomers in the $CeuE_3[Fe(\mathbf{17})][Fe\ (DHBS)_2]_2\}$ structure. The first is the *trans* effect, which can direct certain donor atoms positioning around a metal centre. The *trans* configuration has been shown previously to direct the binding around a metal centre due to the different strengths of the *trans* effect of the *ortho* and *meta* catecholate oxygen donor atoms. The difference between these donor atoms can direct the binding geometry around a metal centre such as MoO_2 with catecholamides[48-49] and 2,3-dihydroxybenzoic acid.[50] These studies showed that the weaker *ortho* catecholate oxygen would be directed to the *trans* position of the strong oxo ligands on the MoO_2 centre in order to minimise the energy of the system. However, the difference in the *trans* effect between the *ortho* catecholate oxygen and the coordinating tyrosine oxygen is likely to be minimal due to their similar pKa values (*ortho* catecholamide pKa ≈ 7.5, tyrosine pKa ≈ 10).[51]

The second explanation is the favourable or unfavourable interactions of the DHBS serine backbone. To examine this hypothesis an electrostatic surface was generated for CeuE chain C, which shows that the DHBS backbone is positioned in a pocket adjacent to the binding site (Figure 68). The electrostatic surface clearly shows the DHBS 1 monomer in the adjacent pocket, with the carbonyl of the catecholamide positioned towards a positively charged protein region, while the alcohol and carboxylic acid groups of the serine backbone are in a neutral region with the alcohol forming a hydrogen bond with the peptide backbone Gly117. As seen in the $\{CeuE_3[Fe(\mathbf{17})]_2[Fe]\}$ structure, the surface of chain B is found close to the binding pocket of chain C with the surface projecting into the region in which the serine backbone of the DHBS monomer would be positioned

if orientated in a *cis*, *cis* configuration. It is more likely that the DHBS monomer has bound in the *cis*, *trans* configuration to avoid this potential steric clash to the N-terminus of CeuE chain B, rather than a energy difference associated with the change in configuration as a result of the *trans* effect the iron centre.

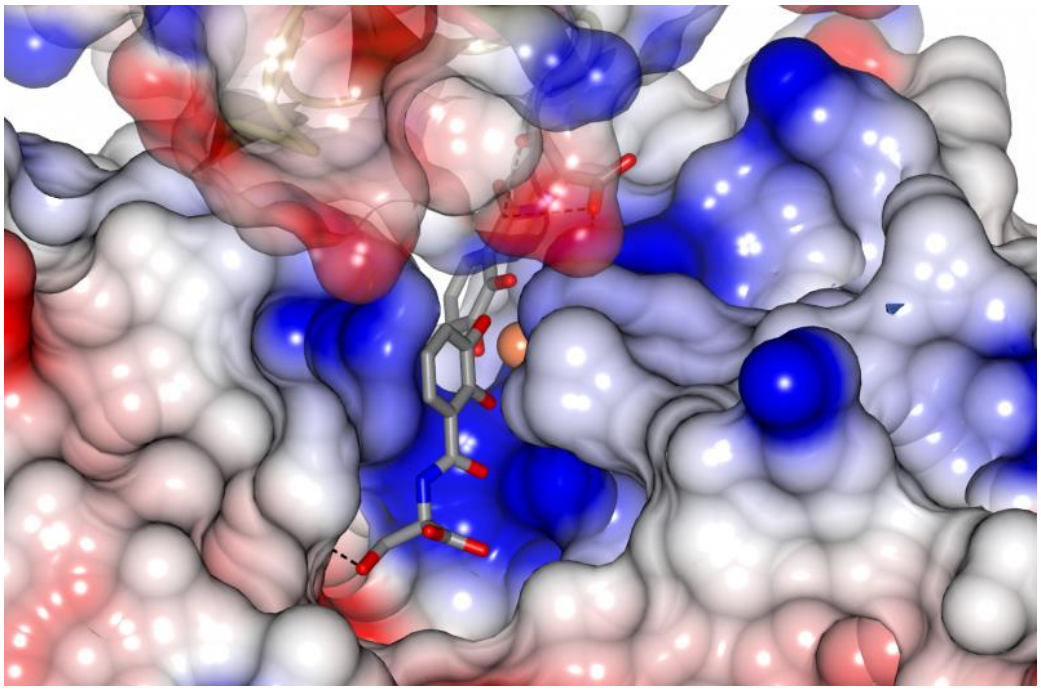


Figure 68 - Surface electrostatic representation of {CeuE$_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$}, monomers B (translucent surface) and C (solid surface), negative charge shown in red and positive charge shown in blue. Key: DHBS, grey - carbon, blue - nitrogen, red - oxygen; Fe - coral.; Hydrogen bond - black dash.

Circular Dichroism and Fluorescence Biophysical Measurements

As outlined in Chapter 2, octahedral complexes with two *cis*-positioned bidentate ligands are chiral. {CeuE[Fe(4-LICAM)]} adopts the Λ-configuration, consistent with the literature for other periplasmic binding proteins and their associated substrates.[52-54] Ferric-**17** has a slight preference towards the Δ-configuration rather than the Λ-configuration, but is able to interchange between the two conformations.[8]

In the {$CeuE_3$[Fe(**17**)]$_2$[Fe]} and {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} in each of the occupied binding pockets the ferric-**17** and the ferric-DHBS complexes are in the Λ-configuration, consistent with the {CeuE[Fe-4-LICAM]} structure. To confirm this behaviour in solution, CD spectra were recorded in the same manner as described for {CeuE[Fe(4-LICAM)]} in Chapter 2. As the previous experiments of ferric-**17** showed it possessed a preference to bind in the Δ-configuration, the addition of CeuE was controlled, with spectra recorded every 0.2 equivalents until 1 equivalent of CeuE was added relative to ferric-**17** (Figure 69).

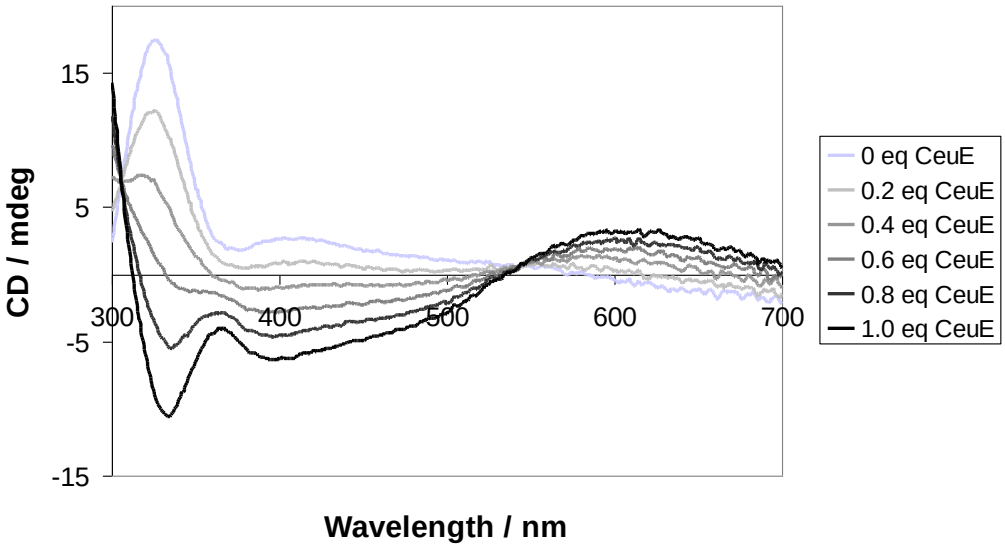


Figure 69 - Circular dichroism spectra of ferric-**17** (light grey) with increasing increments of 0.2 eq of CeuE, until {CeuE[Fe(**17**)]} (black) was formed, recorded in 0.1 M TrisHCl buffer, pH 7.5, 150 mM NaCl.

The spectra recorded show an inversion of configuration and small bathochromic shift of the n→π* transition from a positive peak at 320 nm to a large negative peak at 330 nm upon addition of one equivalent of CeuE, with an isosbestic point at 305 nm. The weak positive feature between 380-610 nm of the LMCT band is also affected, forming a broad negative band with a minimum at 395 nm and a positive band with a maximum at 600 nm, giving an overall positive cotton effect, with another isosbestic point at 540 nm. The isosbestic points suggest that a simple equilibrium is present in the titration involving the

binding of the ferric-**17** to CeuE. The spectrum of {CeuE[Fe(**17**)]} is remarkably similar to that of the {CeuE[Fe(4-LICAM)]} complex. Interestingly, the ligand-based n→π* transition also inverts, from a positive feature to a negative feature upon binding to CeuE. This suggests that despite ferric-**17** being assigned as achiral by Scarrow *et al.* due to the featureless LMCT band showing no clear preference, the strong signal of the n→π* transition can be used as an indicator for the metal-centred chirality. The strong n→π* transition therefore suggests that the complex has a preference for the Δ-configuration, whilst the featureless LMCT band shows that the metal centred configuration is able to convert between Δ and Λ-configurations.[8] The CD spectra show that ferric-**17** is in the Λ-configuration when bound to CeuE, and despite the slight preference for the Δ-configuration of the complex when free in solution, CeuE enforces the Λ-configuration. This is consistent with CD spectra of ferric-enterobactin and ferric-enterobactin bound to FeuA, where the Δ-configuration of free ferric-enterobactin converts to the Λ-configuration when the complex is bound to FeuA.[8, 54]

As outlined in Chapter 2, it is important to determine the affinity of ferric-**17** for CeuE. The dissociation constant for ferric-**17** and CeuE was determined through the same experimental procedure as outlined for ferric-4-LICAM (Chapter 2). The data were plotted and fitted using DynaFit (Figure 70 and Table 13).[55]

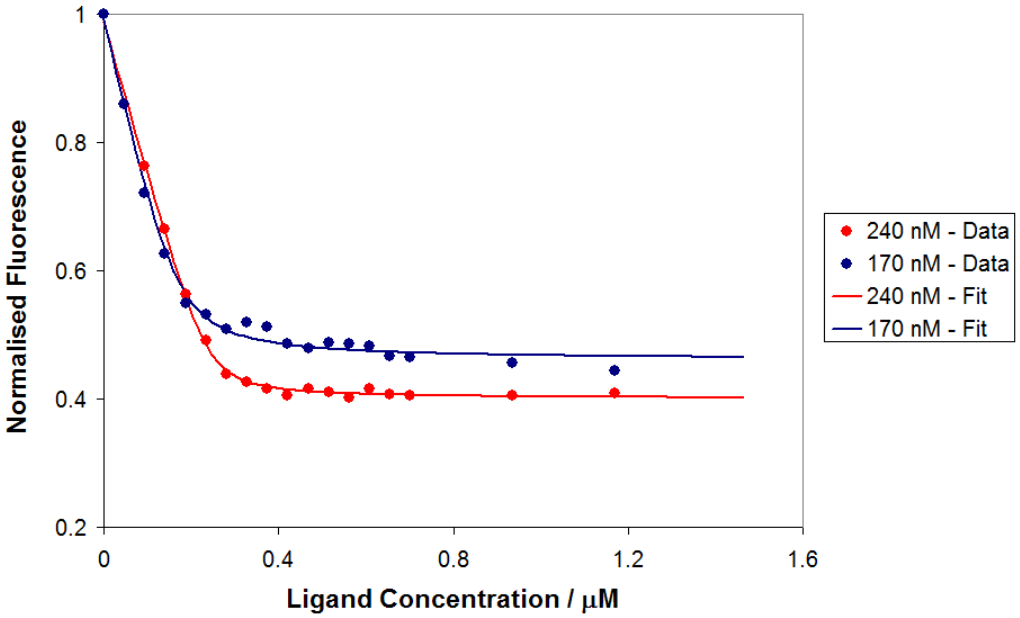


Figure 70 - Example fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit).[55]

		Dissociation Constant / nM		
		Individual	Average of Batch	Overall Average
Batch 1 (240 nM)	Run 1	4.8 ± 0.9	7.5 ± 6.3	8.4 ± 4.3
	Run 2	16.4 ± 1.8		
	Run 3	14.9 ± 3.7		
Batch 2 (170 nM)	Run 1	9.8 ± 1.9	11.3 ± 2.6	
	Run 2	12.3 ± 2.5		
	Run 3	14.9 ± 3.6		

Table 13 - Calculated dissociation constants from non-linear regression analysis of the fluorescence quenching data of ferric-**17**: CeuE titrations. Error of individual K_d values are calculated from non-linear regression fit, and errors of overall and batch calculated from standard deviation of the different K_d values within that group.

The dissociation constant of ferric-**17** was determined for two batches of CeuE as described for ferric-4-LICAM. The first batch, at a concentration of 240 nM, gave an average dissociation constant of 7.5 ± 6.3 nM, whereas the second

batch, at a concentration of 170 nM, gave an average dissociation constant of 11.3 ± 2.6 nM. The second batch of CeuE (170 nM) had a higher normalised fluorescent signal at the end point of the titration than the first batch at 240 nM. The repeat runs within the same batches gave reproducible results; however, the difference of the response of the two CeuE batches was consistently different. The same observation is seen with the ferric-4-LICAM, showing there is a difference in response between batches. As in the case of ferric-4-LICAM, the dissociation constants obtained from the two different batches gave the same result within experimental error.

The average dissociation constant of ferric-**17** (8.4 ± 4.3 nM) is lower than that of ferric-4-LICAM (29.3 ± 11.7 nM) showing that Fe-**17** is bound tighter to the binding pocket of CeuE. This difference in the binding affinity is expected as it is known that ferric-enterobactin is bound to CeuE primarily through electrostatic interactions to the positively charged binding pocket of CeuE. Ferric-4-LICAM, if bound in a 1:1 ligand to metal ratio, would present one negative charge, while ferric-**17** would have a formal negative charge of two, as it has the same binding arrangement through the catechol binding moieties and a free carboxylate attached to the backbone that is hydrogen-bonded in the crystal structure to the neighbouring Arg249 residue. The extra negative charge on the complex may allow the ferric-**17** complex to bind with a greater affinity, and therefore give rise to a lower dissociation constant.

The calculated dissociation constant of ferric-**17** is close to the quoted range of dissociation constants outlined in Chapter 2 for hexadentate catecholate siderophores binding to their associated PBPs, which vary between 15-52 nM.[54, 56-59] The calculated dissociation constant of ferric-**17** is lower than the quoted range in the literature, however, as discussed in Chapter 2 it is likely that there are larger errors within the quoted dissociation constants due to experimental error and different batches of protein expressed and purified, which might suggest that the ferric-**17** has a dissociation constant of similar strength to that of other ferric-catecholate siderophores. This result is somewhat surprising, due to the difference in the binding arrangement between the coordinatively saturated hexadentate siderophores and the two coordinatively unsaturated siderophores

presented here. This suggests that the direct coordination of the histidine and tyrosine residues compensates for the loss of the third catecholamide arm, allowing for the dissociation constant for the tetradentate siderophore binding to be in a biological relevant range.

3.5 Summary of Chapter 3

The enterobactin hydrolysis product H_4-**17** was successfully synthesised, characterised and complexed to iron(III). Ferric-**17** can form complexes with 3:2 and a 1:1 ligand to metal ratios, much like 4-LICAM^{4-}, but has a preference for a 1:1 ratio. The 3:2 complex possesses a preference for the Δ-configuration, whereas the 1:1 complex has a slight preference for Δ, but can interchange between the two configurations.

Structures of {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} and {$CeuE_3$[Fe(**17**)]$_2$[Fe]} were obtained through crystal soaking of ferric-**17** into apo-CeuE crystals for varying lengths of time. The ferric-**17** was bound in the same configuration as ferric-4-LICAM, in a 1:1:1 fashion (metal:ligand:protein), with the His227 and Tyr288 completing the octahedral iron ligation. The ferric-**17** structure in {$CeuE_3$[Fe(**17**)][Fe(DHBS)$_2$]$_2$} was partially hydrolysed, with the DHBS units in CeuE chains A and C bound differently to each other in a *cis, cis,* and a *cis, trans* orientation. The metal-centred chirality of all of the binding pockets of the two crystal structures which had ligands bound was Λ.

Circular dichroism spectroscopy showed an inversion of configuration at the metal centre from Δ to Λ once a stoichiometric amount of CeuE had been added. The dissociation constant was determined for ferric-**17** and CeuE by intrinsic fluorescence quenching and was found to be 8.4 ± 4.3 nM. The complex ferric-**17** shows tighter binding to CeuE compared to ferric-4-LICAM, yet would seem to be within a biologically relevant range.

Chapter 4: A Siderophore Mimic for use with Further Functional Components

Project Aims

The aim of the research outlined in this chapter is the design and synthesis of a biomimetic hexadentate catecholamide siderophore that possesses a chemical functional group to act as a linker. This chemical linker can either be used to add a second chemical function, e.g. a antimicrobial agent or a fluorescence probe, or be left unfunctionalised as a biologically inert group that does not affect the ability of the compound to be recognised as a siderophore, e.g. not affect the coordination of iron(III) (Figure 71).

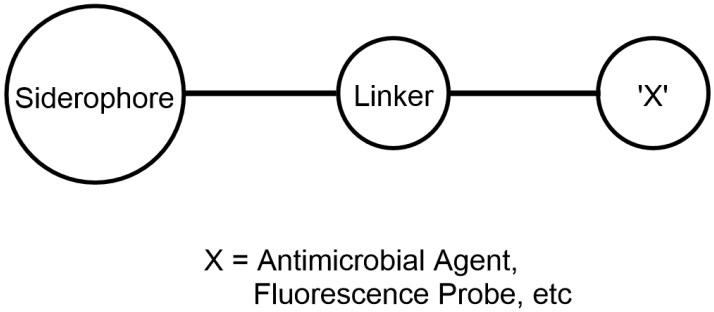


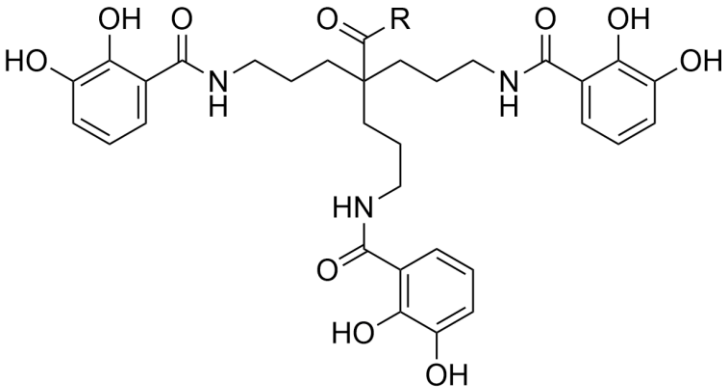
Figure 71 - Schematic design of a biomimetic hexadentate siderophore that possesses a chemical linker.

4.1 Introduction

As outlined in Chapter 1, there are numerous uses for biomimetic models of hexadentate siderophores. Some of these models require chemical linkers, so as to introduce a second chemical moiety (e.g. Trojan Horse strategy, or fluorescent probes),[1-4] whilst others only require the strong iron binding function (e.g. iron overload therapeutics).[5-7]

The most successful previously reported biomimetic hexadentate catecholamide siderophores, were designed as close mimics of enterobactin (see Chapter 1). Only four of the many existing biomimetic hexadentate catecholamide siderophores contain an obvious chemical functionality that can be reacted further in order to covalently attach a chemical tool.

The first was reported by Imbert *et al*. in 2000 and is known as H_6-CacCAM.[8] The structure is similar to H_6-TRENCAM, with an aliphatic three fold symmetric backbone. Instead of connecting through a central tertiary amine, H_6-CacCAM connects through a central quaternary carbon, which is bound to a carboxylic acid. The three iron(III) binding arms possess an extra carbon atom in the chain, compared to H_6-TRENCAM.



R = OH = H_6-**CacCAM**
R = $NHC_{16}H_{33}$ = H_6-**CacCAM(C_{16})**

H_6-CacCAM has been linked to an aliphatic carbon linker (C_{16}) to produce an amphiphilic iron(III) chelator as a mimic of marine bacterial siderophores.[9] The modified H_6-CacCAM(C_{16}) and its ferric complexes are able to self-assemble, the iron-free H_6-CacCAM(C_{16}) forming micellar assemblies, whilst the ferric-CacCAM(C_{16}) complex formed spherical assemblies with diameters of approximately 130 nm. This system mimics the behaviour of marine bacterial siderophores.[9-10]

The second example is an iron binding dendrimer, synthesised by Zhou *et al.* as a novel approach for the treatment of haemochromatosis (a type of iron overload).[11] Dendrimers are synthetic macromolecules with highly branched structures that have been widely studied.[12-17] The structure of the dendrimer allowed the covalent attachment of several hexadentate iron binding moieties to one molecule. For this design, the hexadentate metal binding unit needed a functional group to allow a covalent linkage between the metal binder and the

dendrimer backbone to be formed. This was achieved by using compound **35** as the branched linker, with the carboxylic acids used for both the growth of the dendrimer to the next generation and the attachment of the metal binding catechol moieties to give compound H_{54}-**36** (Figure 72).

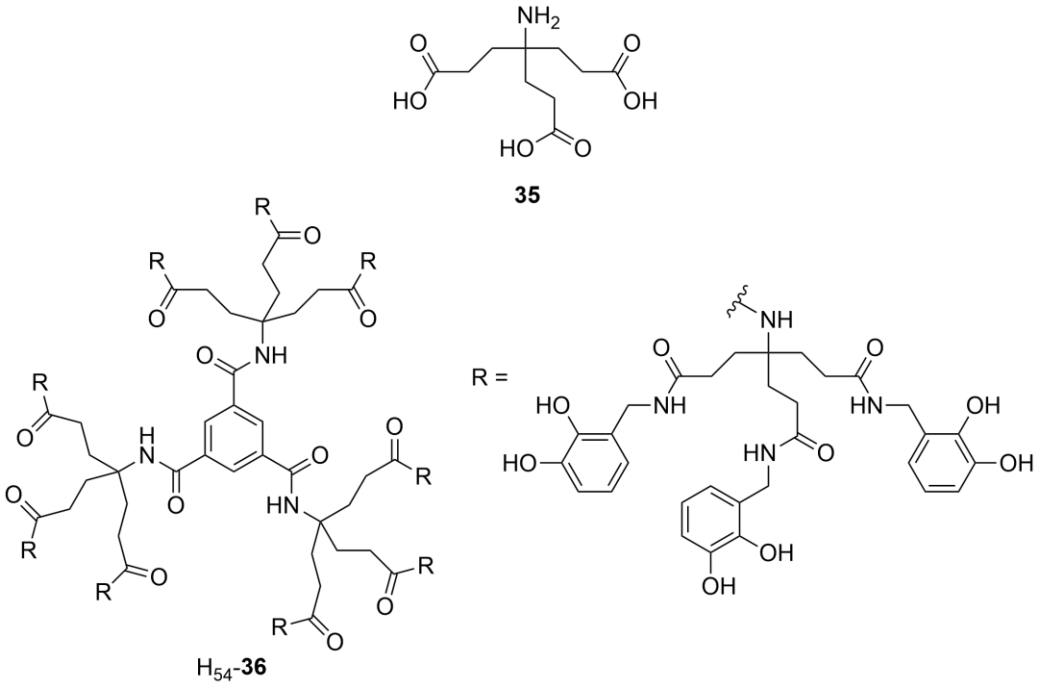


Figure 72 - Dendritic catecholate-based iron chelator developed by Zhou *et al*.[11]

The structure of H_{54}-**36** uses a primary amine, situated on the quaternary carbon atom rather than a carboxylic acid, as seen in H_6-CacCAM for covalent linkage. Consequently, the catechol binding units are not the preferred catecholamide functionality. H_{54}-**36** is still able to bind iron(III), but as described in Chapter 1, is not recognised and therefore evades the iron uptake system.[18-19] This is ideal for use as an iron(III) chelator for the treatment of haemochromatosis but is not suitable for use as a model siderophore.

The third example was recently developed by Ji *et al.* and incorporates design features of both H_6-CacCAM and H_{54}-**36**.[20] The structure of **37** is similar to H_6-CacCAM, with an aliphatic backbone with three-fold symmetry, and seven atom spacers between the catecholamide units. Interestingly, the catechol groups were left protected as acetate groups. *In vivo*, the acylated catechol groups were expected to avoid potential methylation by a catechol *O*-methyl-transferase.

Compound **37** acts as a prodrug, with hydrolysis leading to the required catechol functionality for iron(III) binding. Compound **37** also includes a succinic acid linker, altering the linking functional group from the original amine to a carboxylic acid, which could then be used for covalently linking β-lactam antibiotics, as required for the Trojan Horse strategy, giving compounds **38** and **39** (Figure 73).[20]

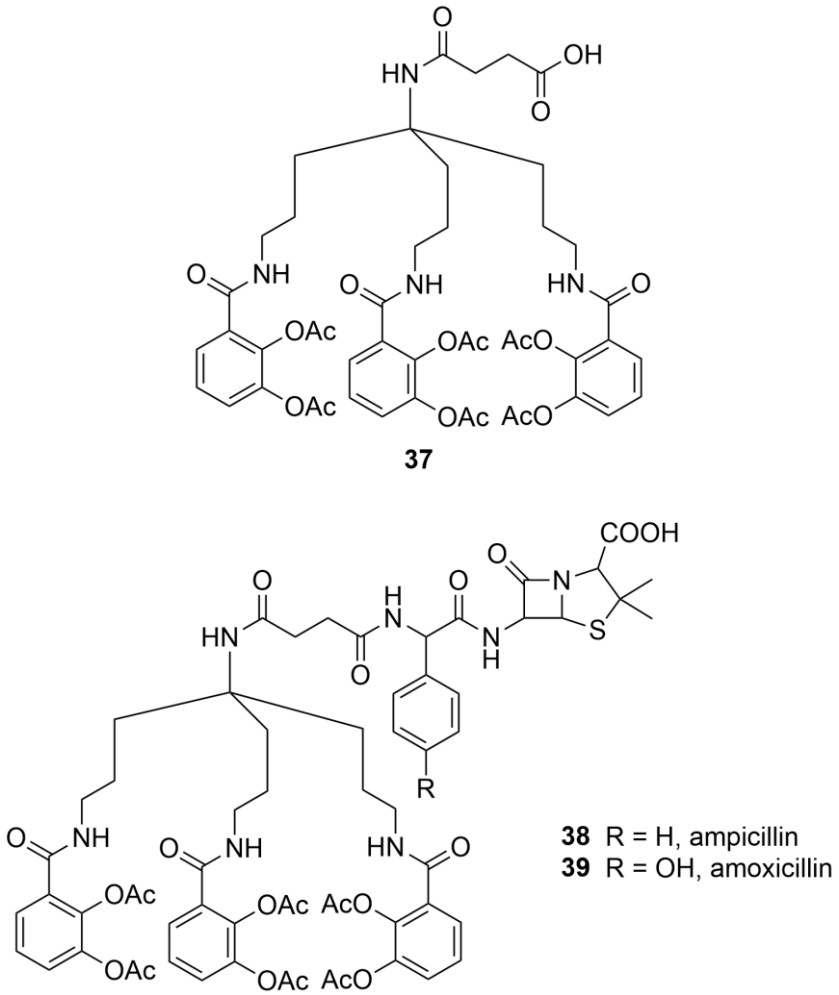
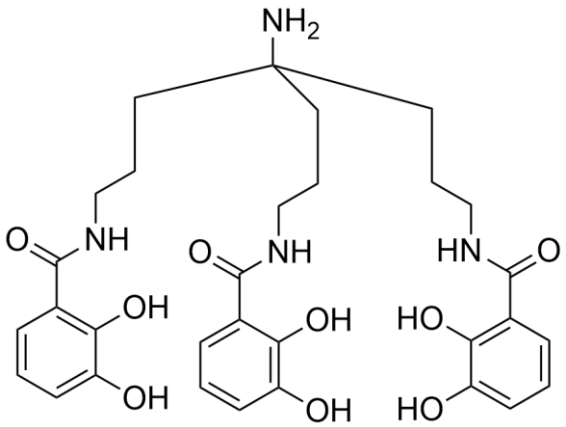


Figure 73 - Structures of a biomimetic tris-catecholate siderophore and its aminopenicillin conjugates.[20]

Both compounds **38** and **39** showed improved antibacterial activity against a selection of strains of *P. aeruginosa* compared to the parent antibiotic.[20] In order to ascertain whether the increased antibacterial activity was as a result of increased uptake through the iron uptake system, the antimicrobial activity was measured in both iron-rich and iron-deficient media. There was an increase in activity when the compounds were tested in iron-deficient medium, where the

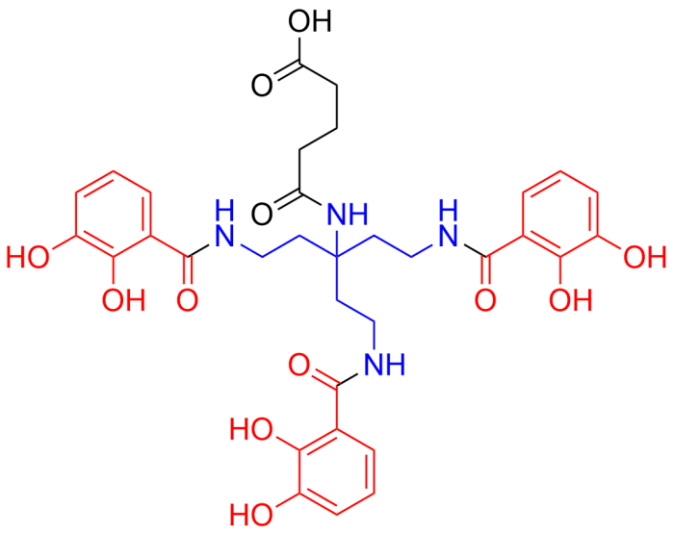
iron uptake system plays a greater role in active nutrient uptake. This suggests that active transport of the conjugates through the siderophore mediated iron uptake mechanism plays a significant role in delivering the drug to the bacteria.[20]

The fourth example, H_6-**40**, was developed by Inomata *et al* and has a similar structure to **37**, but it contains no acetyl protecting groups and the succinic linker is absent.[21] A self assembled monolayer of ferric-**40** was formed and chemically bound to a gold surface. This allowed the adsorption of *E. coli* onto the substrate which was detectable by a quartz crystal micro-balance chip used for detection of micro-organisms with the corresponding siderophore uptake protein.[21]



H_6-**40**

None of these four mimetics possess both the five atomic linker between the catecholamide binding moieties and the correct orientation of the catecholamide required for biological recognition. Thus, the design of the biomimetic hexadentate catecholamide siderophore, forming the rest of this chapter, took into account the key features of enterobactin, and also the strategies of the previously discussed biomimetic siderophores, which possessed a chemical linker. The key features incorporated were: a) the three fold symmetry of the backbone, b) the five atom spacer between the catecholamide units and c) the removal of the hydrolytically unstable trilactone backbone (Figure 74).



H_6-**41**

Figure 74 - The chemical structure of the designed biomimetic hexadentate siderophore. Key: catecholamide, red - backbone, blue - linker, black.

In the design of H_6-**41**, the linker choice was inspired by the linker used by Ji *et al.* (compound **37**), as it could either be incorporated or not, depending on requirements. The inclusion of the linker is determined by the probe to be synthesised, accounting for the requirements of either an amino or carboxylate group for linkage to a functional tool. The position of the linker at the central quaternary carbon has the advantage of directing it away from the iron(III) binding catecholamide moieties, and hence should have minimal impact on the iron(III) binding of the compound. For proof-of-concept, the linker was included in the synthesis of H_6-**41** to demonstrate the feasibility of incorporating this functionality.

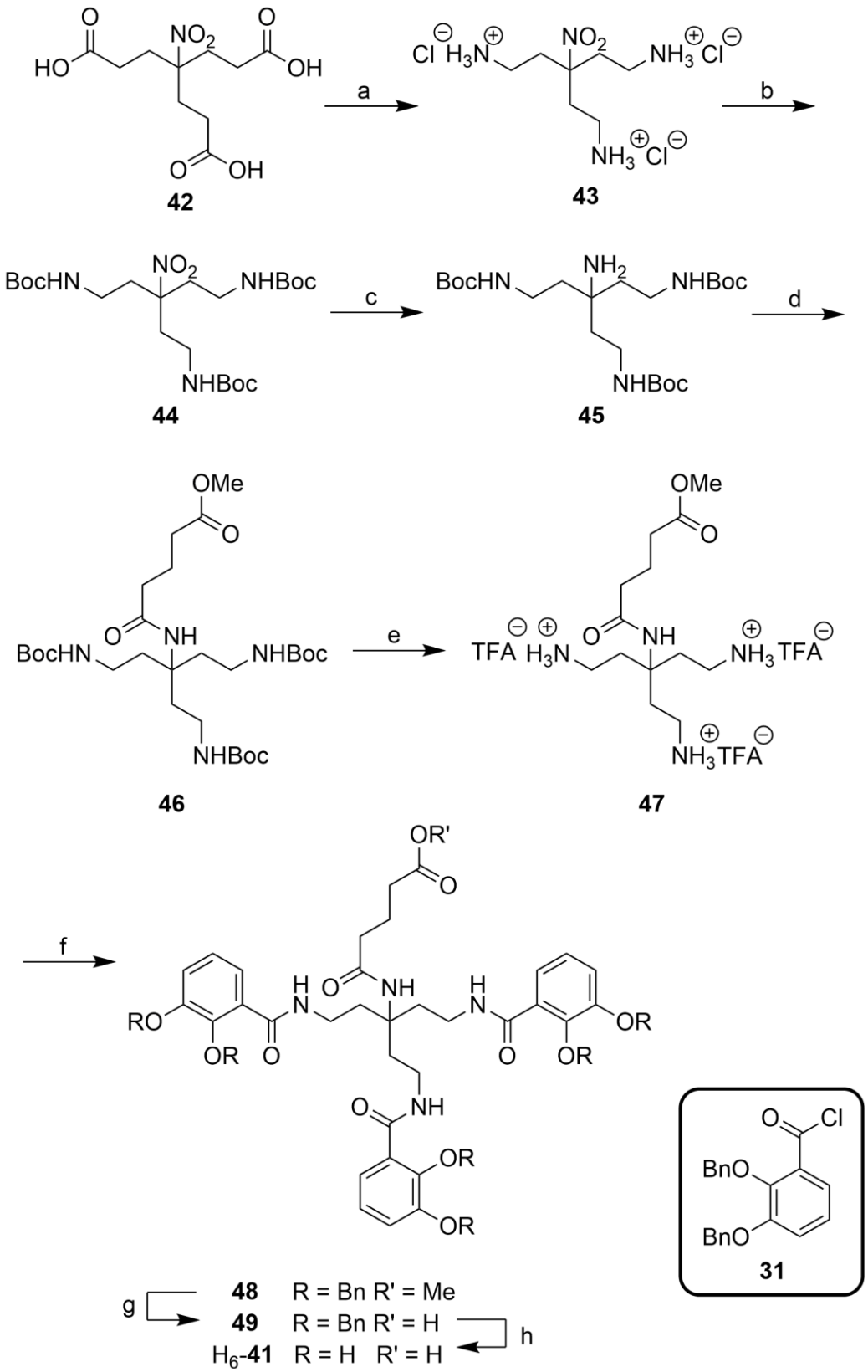
4.2 Synthesis

The synthesis of H_6-**41** can be divided into two sections: firstly the synthesis of the catechol units and secondly of the siderophore backbone. The synthesis of the catechol units has been discussed in Chapters 2 and 3 and therefore the present Chapter will focus on the siderophore backbone.

Synthesis of a Biomimetic Hexadentate Siderophore

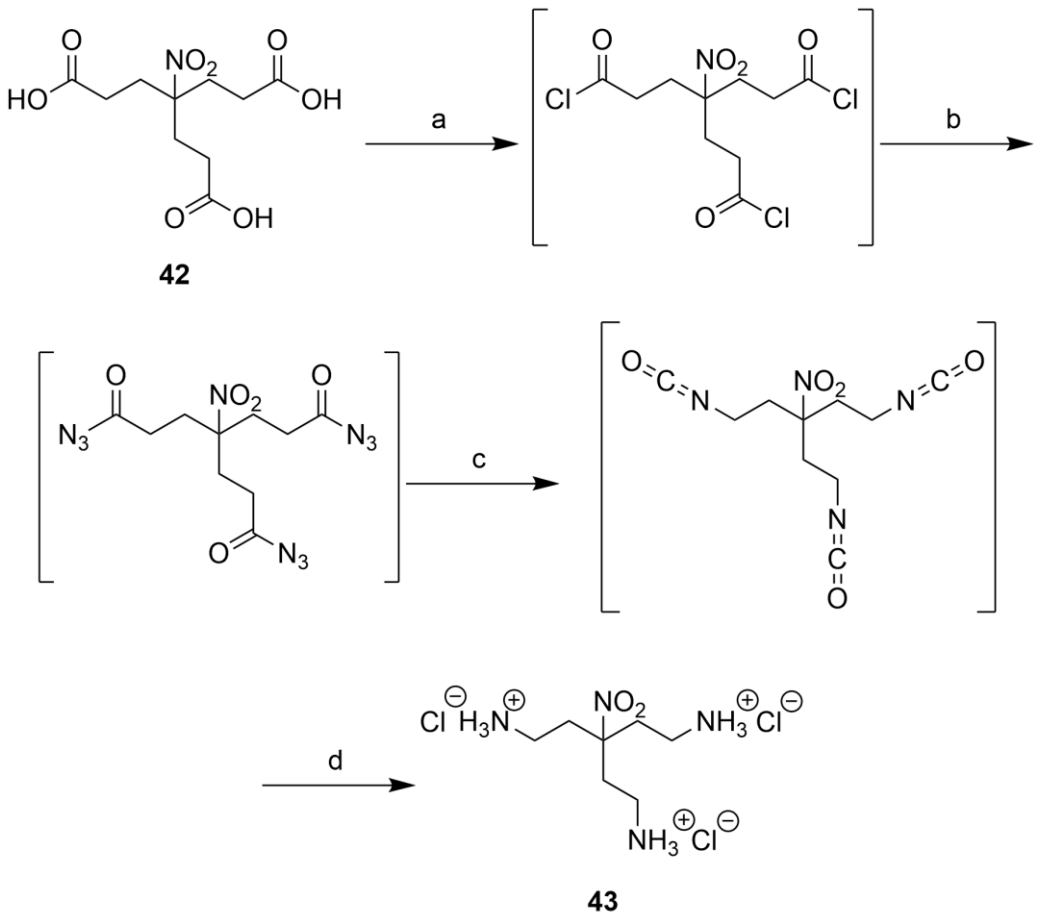
The work outlined within this section was completed with the aid of Amy Scorah, a summer placement student working under my supervision (July-September 2013).

The backbone of H_6-**41** requires four primary amines in two different environments, one located on a central quaternary carbon, whilst the remaining three are located on the terminal of each of the three arms. 4-(2-carboxyethyl)-4-nitroheptanedioic acid, **42**, offered a good starting point for the synthesis, as it possesses much of the carbon skeleton of the backbone. The carboxylic acids can be converted into the required terminal primary amines *via* an isocyanate intermediate using the Curtius rearrangement.[22] Once the terminal amines have been synthesised, standard protecting group chemistry can be employed to enable the reduction of the central nitro group to a primary amine, which can then be used to couple a protected linker. Subsequently, selective deprotection of the terminal primary amines allows the introduction of the catecholamide moieties onto the backbone. Standard deprotection chemistry should then yield H_6-**41** (Scheme 11).



Scheme 11 - Synthetic route to H_6-**41**. (a) i) $SOCl_2$, ii) TMS azide, iii) 80 °C, iv) HCl, (b) Boc_2O, NEt_3; (c) H_2, Raney®-Nickel 2800 slurry; (d) Glutaric acid monomethyl ester chloride, NEt_3 ; (e) TFA; (f) **31**, DIPEA; (g) NaOH; (h) H_2, Pd-C 10%.

The backbone of H_6-**41** was synthesised from **42**, using the methodology reported by Zhao *et al.*[23] The carboxylic acid functional groups were transformed into primary amines by conversion to acyl azides, followed by Curtius rearrangement to give the corresponding isocyanates. Next, acid catalysed hydrolysis of the isocyanate was carried out to give the required amine salt (72% yield) (Scheme 12).



Scheme 12 - Synthetic route to compound **43**. (a) $SOCl_2$; (b) TMS azide; (c) 80 °C; (d) HCl.

Due to the reactive nature of the intermediates, no attempt was made to isolate and characterise these compounds. However, compound **43** was isolated and characterised by mass spectrometry, ^1H, ^{13}C NMR and infra-red spectroscopies. Mass spectrometry confirmed the presence of both the $[M+H]^+$ and $[M+Na]^+$ ions for a molecular formula of $C_7H_{19}N_4O_2$ with signals observed at *m/z* 191.1499 and *m/z* 213.1324, respectively. The ^1H and ^{13}C NMR spectra showed that the symmetry of compound **43** had been maintained, with only three signals

in the ^{1}H NMR spectrum for the three proton environments (including amine protons). The multiplicity of the signals does not resemble the standard triplets reported by Zhao *et al.*, possibly due to long range ^{14}N coupling as observed in ammonium ions (Figure 75).[24]

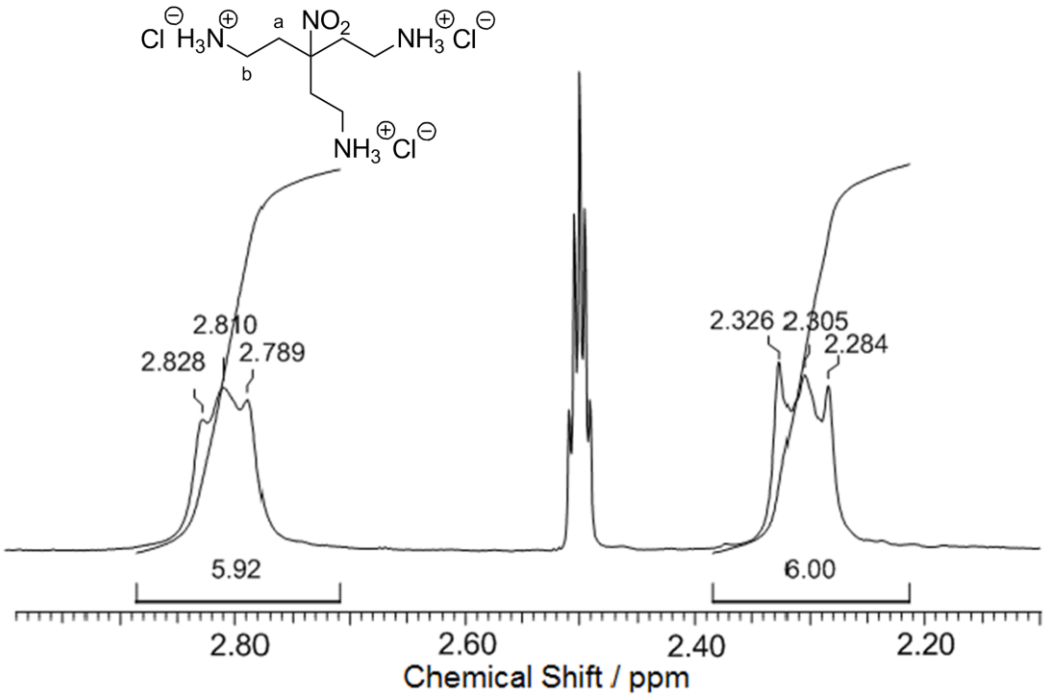
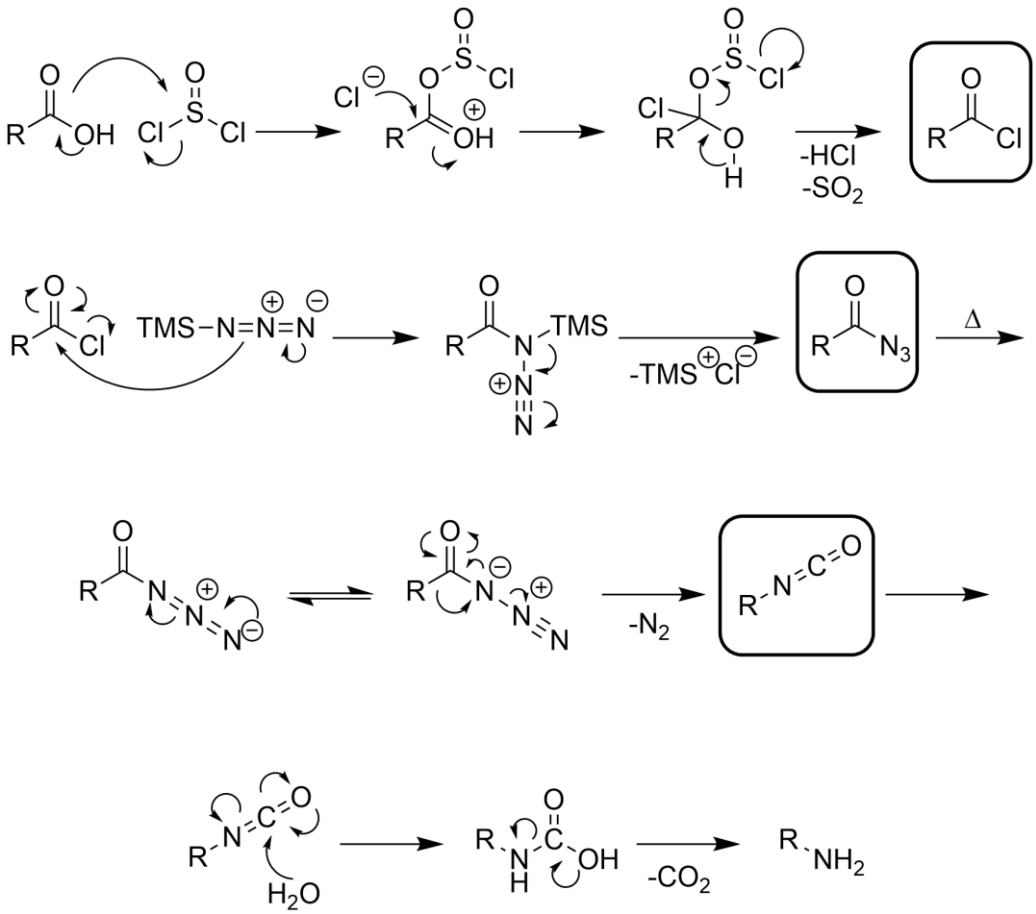


Figure 75 - ^{1}H NMR spectrum of **43** in d_6-DMSO shown between 3.00-2.10 ppm showing the multiplicity of the resonances assigned to protons H^a and H^b. Quintet at 2.50 is the d_6-DMSO residual solvent peak.

The ^{13}C NMR spectrum shows the expected three signals; one at 89.7 ppm consistent with a quaternary carbon atom attached to a nitro group, and two at 33.8 ppm and 31.8 ppm, respectively for the two aliphatic carbon environments. Apart from the unusual ^{1}H NMR multiplicity all other characterisation data matched that of the literature, and therefore it was concluded that compound **43** had been successfully synthesised.[23]

Upon repeating the reaction, it became apparent that the reproducibly of the yield and purity of compound **43** was difficult to maintain, probably due to the nature of the reactive intermediates. The mechanism of the reaction first involves the reaction of the carboxylic acid with thionyl chloride, to give the activated acyl chloride, which then reacts with TMS (trimethylsilyl) azide to

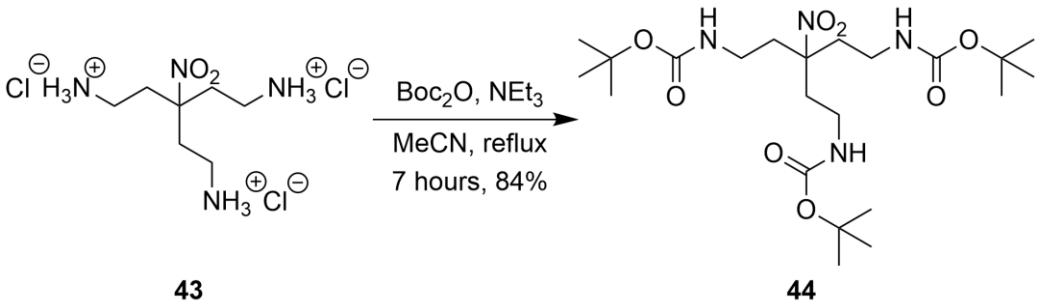
produce the acyl azide. Upon heating, acyl azides undergo Curtius rearrangement, which involves the evolution of nitrogen gas upon rearrangement of the acyl azide into the isocyanate. Once the isocyanate has formed, acid catalysed hydrolysis forms the required amine (Scheme 13).[25]



Scheme 13 - Reaction mechanism of reaction of **42** to form **43** with key intermediates, acyl chloride, acyl azide and isocyanate highlighted within a box.

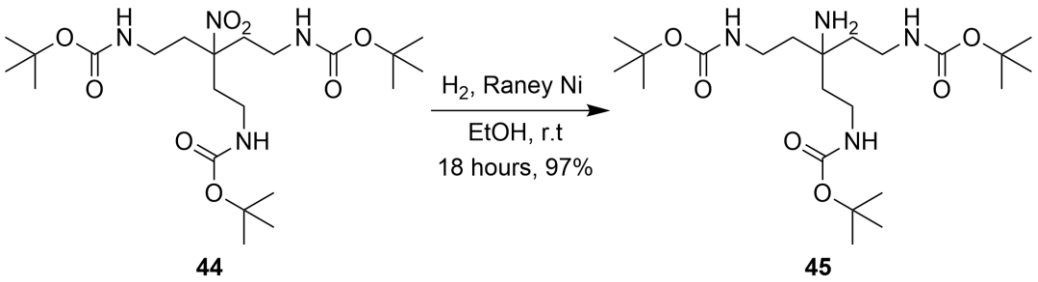
The key intermediates in this reaction, the acyl chloride and acyl azide, are prone to hydrolysis, creating additional competing side reactions, forming undesired by-products. Due to the three-fold symmetry of the starting material, the reaction process must go to completion three times on each molecule to produce the target compound. The procedure was therefore modified to be completed under rigorously dry conditions, with a constant nitrogen atmosphere utilising a Schlenk line. This improved the yield to 92%, with characterisation data consistent with the previous synthesis, and literature data.[23]

Compound **44** was synthesised by reacting **43** with excess di-*tert*-butyl dicarbonate in the presence of triethylamine using the methodology of Zhao *et al.* (84% yield, Scheme 14).[23] Compound **44** showed a characteristic singlet at 1.42 ppm in the ^{1}H NMR spectrum, with a relative integration of 27 corresponding to the three equivalent *tert*-butyl groups. Signals in the mass spectrum were observed at *m/z* 491.3068 and *m/z* 513.2885 consistent with the $[M+H]^+$ and $[M+Na]^+$ ions, respectively, for a molecular formula of $C_{22}H_{42}N_4O_8$. All other characterisation data are consistent with the literature.[23]



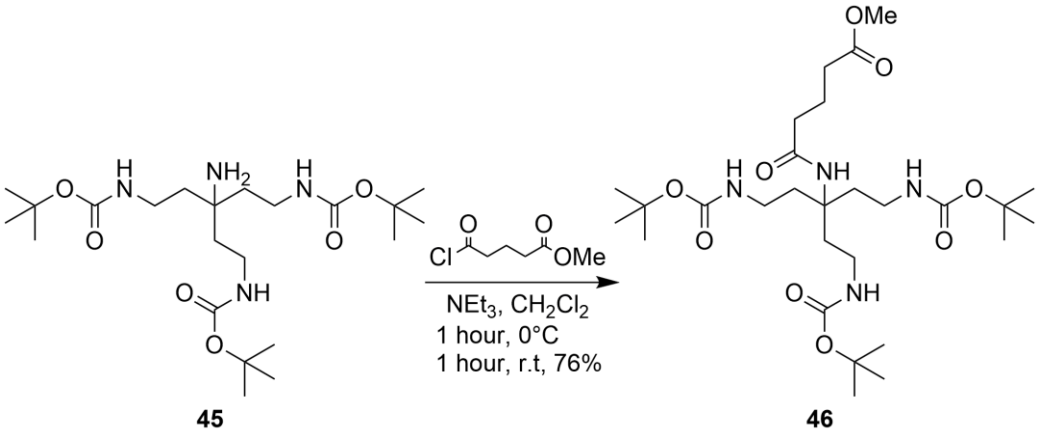
Scheme 14 - Synthesis of **44** using the methodology of Zhao *et al.*[23]

Compound **44** was reduced to the corresponding amine **45** by catalytic hydrogenation using Raney®-Nickel 2800 slurry as the catalyst. There have been reports in the literature that many reducing conditions, such as platinum, palladium and commercially available Raney®-Nickel, give poor yields for aliphatic nitro reduction, and that a specialised T-1 Raney®-Nickel is required to be synthesised and utilised in order to complete the reaction.[26-27] However, others have claimed to use commercially available Raney®-Nickel to successfully reduce a aliphatic nitro group on a quaternary carbon to give a primary amine.[28] Therefore, commercially available Raney®-Nickel (*Raney Nickel 2800 slurry in H_2O active, Sigma-Aldrich*) was tried first and found to be successful at reducing the nitro group to the amine under a pressurised atmosphere of hydrogen (35 bar, Scheme 15). Successful reduction was indicated by mass spectrometry, with signals observed at *m/z* 461.3356 and *m/z* 483.3164 consistent with the $[M+H]^+$ and $[M+Na]^+$ ions, respectively, for a molecular formula of $C_{22}H_{44}N_4O_6$ and an upfield shift in the ^{13}C NMR spectrum of the signal due to the quaternary carbon from 90.4 ppm to 53.8 ppm.



Scheme 15 - Synthesis of **45** using commercially available Raney®-Nickel.

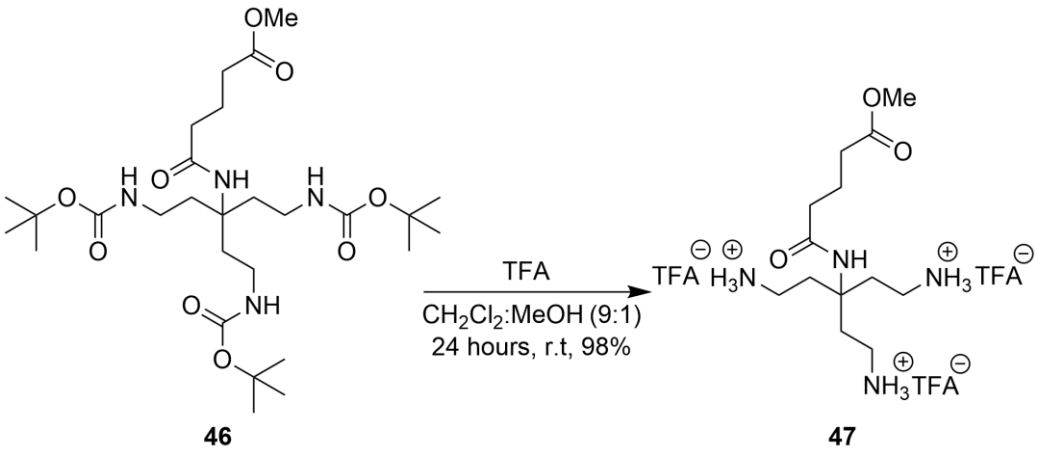
The primary amine of **45** was used to attach the linker unit onto the backbone, by reacting **45** with glutaric acid monomethyl ester chloride in the presence of triethylamine in anhydrous dichloromethane under a dry nitrogen atmosphere to yield **46** (76% yield, Scheme 16). The additional signal due to the amide proton in the ^{1}H NMR spectrum at 6.57 ppm, and the amide and ester carbonyl signals in the ^{13}C NMR spectrum at 174.1 and 173.1 ppm, respectively, confirmed the addition of the monomethyl linker *via* the amide linkage. Mass spectrometry confirmed the presence of both the $[M+H]^+$ and $[M+Na]^+$ ions at *m/z* 589.3807 and *m/*z 611.3645 ppm, respectively, consistent with a molecular formula of $C_{28}H_{52}N_4O_9$.



Scheme 16 - Synthesis of **46** using the modified methodology of Ji *et al*.[20]

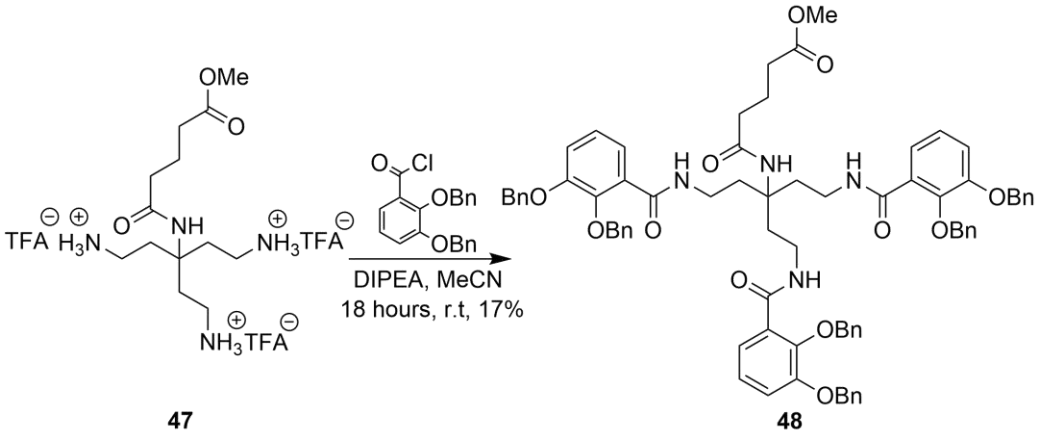
Compound **46** represents the protected backbone of the hexadentate siderophore with a methyl ester protected linker attached. The continuation of the synthesis of H_6-**41** first involved the removal of the Boc-protecting groups of **46**, which was achieved using trifluoroacetic acid in a dichloromethane : methanol (9:1) mixture. The ^{1}H NMR spectrum showed the complete removal of the three *tert*-butyl signals, while ^{13}C NMR confirmed that the ester and amide linkage of **47**,

173.2 and 172.1 ppm, respectively, had not been affected during the reaction. Mass spectrometry gave further confirmation that compound **47** was isolated with a signal observed at *m/z* 289.2247 assigned as the $[M+H]^+$ consistent with for a molecular formula of $C_{13}H_{28}N_4O_3$. Compound **47** was isolated as a white solid in 98% yield (Scheme 17).



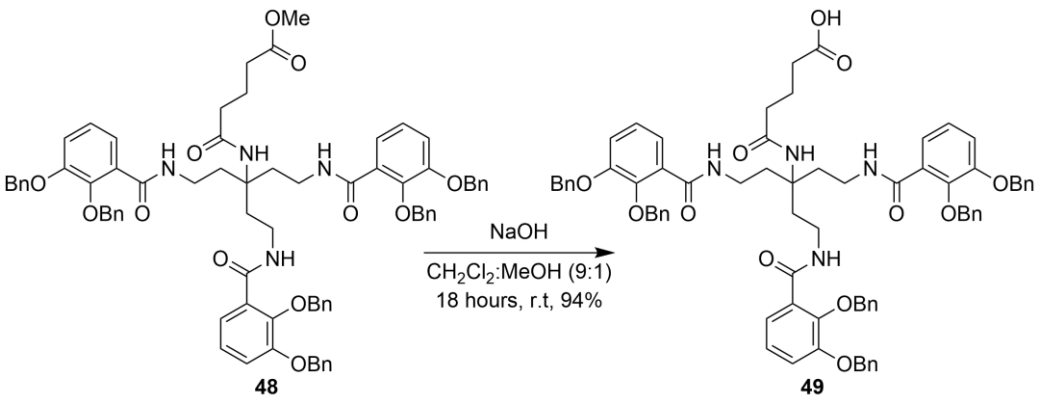
Scheme 17 - Synthesis of **47** by Boc deprotection of **46** using trifluoroacetic acid.

Compound **31** was utilised for the addition of the catechol functionality to **47**, in the same manner as in the synthesis of H_4-**17** (Chapter 3). The reaction proceeded to **48** in 17% yield (Scheme 18), a surprisingly low value, however, it was later found that the starting material, **47**, was highly hygroscopic, forming a clear viscous oil when left open to the atmosphere. It is believed that this had occurred during the set-up of the reaction leading to the introduction of water, hydrolysing reagent **31** in a competing reaction, explaining the low yield. In order to confirm this hypothesis the reaction would need to be repeated using inert conditions, and the preparation and storage of compound **47** would need to be altered to eliminate any water. Despite the low yield, mass spectrometry showed a signal at *m/z* 1259.5357 assigned as the $[M+Na]^+$ ion, consistent with a molecular formula of $C_{76}H_{76}N_4O_{12}$. The ^{1}H NMR spectrum shows the addition of the aromatic signals between 7.71 and 7.11 ppm, with a total relative integration of 39 protons. Furthermore, two additional singlets at 5.15 ppm and 5.09 ppm, which both have a relative integration of six protons, are assigned as the methylene signals of each of the benzyl protecting groups of the catechol moieties.



Scheme 18 - Synthesis of **48** using the methodology outlined in Chapter 3.

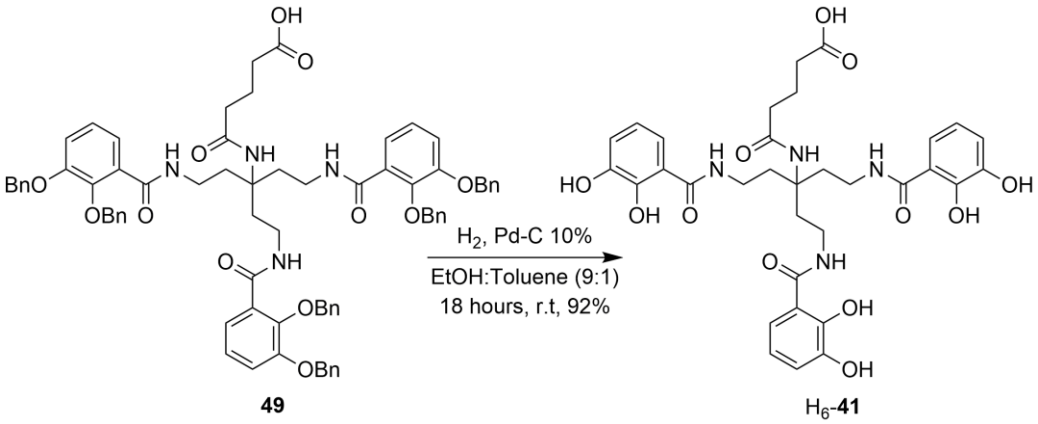
The methyl ester protecting group of **48** was removed using base hydrolysis, employing the methodology outlined by Theodorou *et al.* by reacting compound **48** with sodium hydroxide in a dichloromethane : methanol (9:1) mixture, yielding compound **49** (94% yield, Scheme 19).[29] The ^{1}H and ^{13}C NMR spectra showed the complete removal of the signal due to the methyl protons (^{1}H NMR: 3.65 ppm, ^{13}C NMR: 51.9 ppm), with the rest of the resonances in the spectra unaffected. Mass spectrometry confirmed the amide had not been hydrolysed during the reaction with a signal found at *m/z* 1221.5240 assigned as the $[M-H]^{-}$ ion.



Scheme 19 - Synthesis of **49** via methyl ester hydrolysis using the methodology of Theodorou *et al*.[29]

The free carboxylic acid functionality of **49** at this point could be coupled to a addition functional component to functionalise the compound. However, this is beyond the scope of this project, and it was therefore left as the free carboxylic acid and the reaction scheme was continued towards the target compound, H_6-**41**.

The benzyl ethers of compound **49** were removed using the same methodology as for H_4-4-LICAM and H_4-**17** (Chapters 2 and 3). Catalytic hydrogenation using 10% Pd-C yielded compound H_6-**41** (92% yield). Mass spectrometry confirmed the presence of H_6-**41** with a signal observed at *m/z* 681.2420 assigned as the $[M-H]^-$ ion, consistent with the molecular formula $C_{33}H_{37}N_4O_{12}$. The ^1H and ^{13}C NMR spectra confirmed the complete removal of the benzyl protecting groups by absence of the methylene signals, (^1H NMR, 5.13 and 5.07 ppm, ^{13}C NMR, 76.5 and 71.4 ppm). Furthermore the aromatic region of the ^1H NMR spectrum simplified to three signals (^1H NMR, 7.19, 6.91 and 6.70 ppm). The rest of the resonances in the spectra were unaffected, confirming the successful synthesis of compound H_6-**41**.



Scheme 20 - Deprotection of **49** via catalytic hydrogenation to yield siderophore mimic H_6-**41**.

4.4 Summary of Chapter 4

A novel catecholamide siderophore mimic that possesses a five-atomic spacer, free chemical functionality and a three-fold symmetric axis, was successfully designed, synthesised and characterised. The three carboxylic acid groups **42** were converted to primary amines *via* activation with thionyl chloride, followed by reaction with TMS azide to produce the acyl azide. Subsequent rearrangement of the acyl azide into the isocyanate allowed for acid catalysed hydrolysis to form **43**. Protection of the three terminal primary amines using di-*tert*-butyl dicarbonate, allowed the reduction of the nitro group to the central primary amine. This central amine was subsequently coupled to the linker completed the synthesis of the protected backbone. Deprotection, liberated the terminal primary amines, and allowed the catecholamide moieties to be introduced. This was followed by deprotection of the methyl ester of the linker and benzyl deprotection of the catecholamides to yield the target compound H_6-**41** in an overall yield of 10.5% over eight synthetic steps. All the reactions, with exception of the coupling of the catecholamide moieties to the backbone proceeded in yields >76%. and therefore if this reaction yield could be improved (Section 4.2), to other comparable yields (~60%) then the overall yield to the successful synthesis of H_6-**41** should increase to approximately 30%.

Chapter 5: A Natural Siderophore with a Further Fluorescent Component

The work outlined within this Chapter is an ongoing collaboration with Prof James Coulton (McGill University, Montreal) and Dr Christoph Baumann (Department of Biology, University of York).

Project Aims

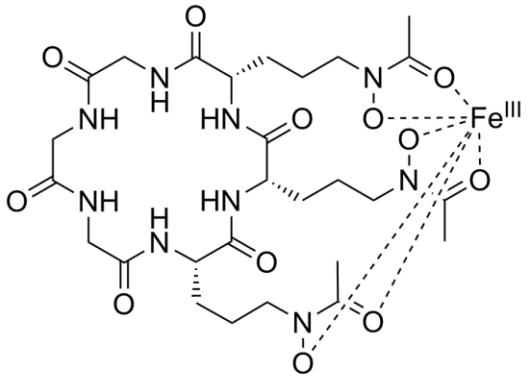
The aim of this collaborative research project was the design and synthesis of a small molecule-fluorophore conjugate by the modification of ferricrocin for the targeting and labelling of the outer membrane receptor, FhuA. This should allow a targeted approach for fluorescently labelling the protein and allow the location and mobility to be and characterised within the outer cell membrane.

5.1 Introduction

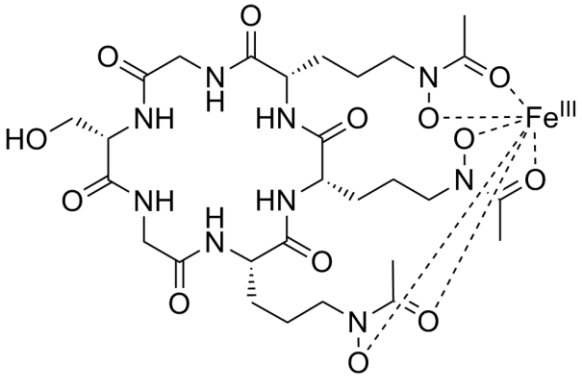
Along with the use of siderophore mimics as chemical tools (see Chapter 4), natural siderophores can also be utilised either for their iron(III) binding ability, e.g. the use of H_3-desferrioxamine B for the treatment of iron overload, or be modified to introduce an additional functional component, e.g. the attachment of antimicrobial agents to natural siderophores for use in Trojan Horse antimicrobials.[1-5] Other applications of functionalised siderophores include the delivery of cargos,[6-7] targeted fluorescent probes,[8-13] and the investigation of iron(III) uptake (see Chapter 1).[14-15]

Ferrichrome is a cyclic hexapeptide, made up of three glycine units and three modified ornithine units that act as the iron binding hydroxamate units. Ferricrocin is a naturally occurring hexadentate hydroxamate siderophore, structurally related to ferrichrome but with the second glycine residue replaced by a serine.[16] Despite *E. coli* not synthesising either ferrichrome or ferricrocin, both siderophores can act as iron suppliers for *E. coli* and a range of other organisms using the hydroxamate uptake system.[17] Modification of ferricrocin to investigate iron(III) uptake in *E. coli* was previously reported by Coulton *et al.* .[14]

The free alcohol functional group on ferricrocin makes it an ideal hydroxamate siderophore for studying the iron uptake mechanism, as the free alcohol functionality allows direct coupling of a fluorescent label to the siderophore backbone, away from the iron binding region.

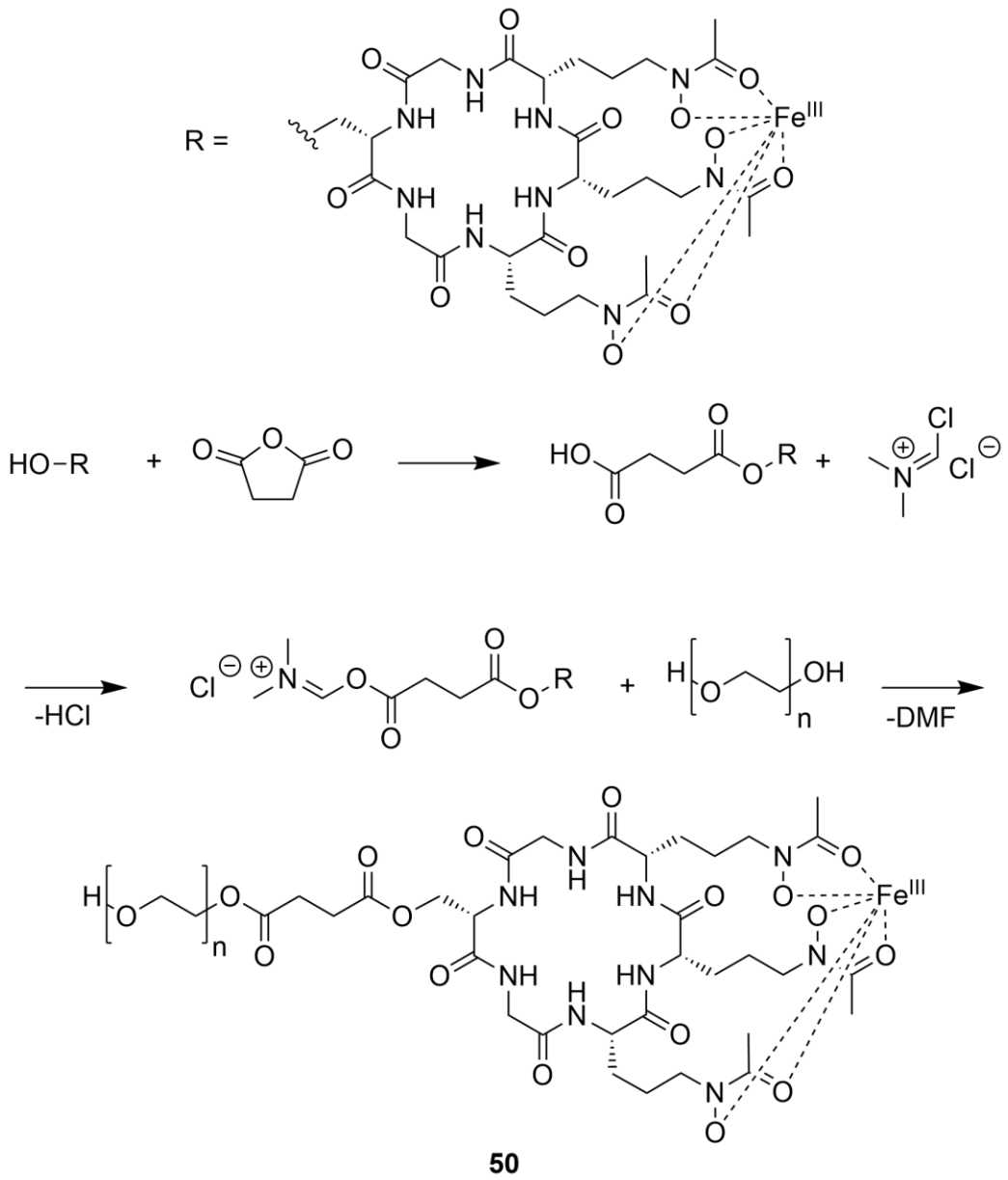


Ferrichrome



Ferricrocin

Coulton *et al.* used this alcohol functional group to covalently link the ferricrocin to poly(ethylene glycol) to form the polymer-ferricrocin conjugate. First the alcohol of the ferricrocin was reacted with succinic anhydride to yield the monoester, ferricrocinyl succinate. The free carboxylate was then activated by the reaction of *N*-(chloromethylene)-*N*-methylmethanaminium chloride followed by the addition of poly(ethylene glycol) (M_r 7000-8500) to yield the covalently linked ferricrocinyl polyethylene glycolyl succinate, **50** (Scheme 21).[14]



Scheme 21 - Synthesis of high-molecular weight ferricrocinyl polyethylene glycolyl succinate by Coulton *et al.*[14]

Coutlon *et al.* showed that the polymer-ferricrocin conjugate could act as an iron source, although at a reduced rate compared to free ferricrocin. When either free ferricrocin or the polymer-linked ferricrocin conjugate was used as the iron source, the TonB system was required for growth. Control experiments indicated that the ester linkage of the polymer-ferricrocin conjugate was not simply hydrolysing and acting as a source of free ferricrocin, rather that the conjugate could itself act as a iron(III) source for the cells.[14] These results

demonstrate the complex nature of the iron(III) uptake mechanism, with many different processes all playing their role in the uptake of this essential element.

The characterisation of the location and mobility of integral membrane proteins has become of great interest since the discovery of the complex and non-uniform nature of cell surfaces. The lipids and proteins of a cell surface are not homogenously distributed in the membrane, but are separated into biochemically and biophysically distinct domains.[18] Highly ordered domains are known as lipid rafts, which have implications for protein distributions, diffusion and trafficking.[18-21] The chemical make-up of the lipid rafts make the characterisation of these systems difficult as they comprise of complex dynamic systems. Fluorescence microscopy is an ideal technique for the investigation of these systems as it can be applied to cell membranes and avoids the fixation of integral membrane proteins which can introduce artefacts in organisation.[18, 22] The analysis of the localisation and dynamics of membrane proteins can offer good insights into the constitution of these complex systems utilising techniques such as fluorescence recovery after photobleaching (FRAP), which allows the direct observation of the diffusion of fluorescently-tagged membrane proteins (Figure 76).

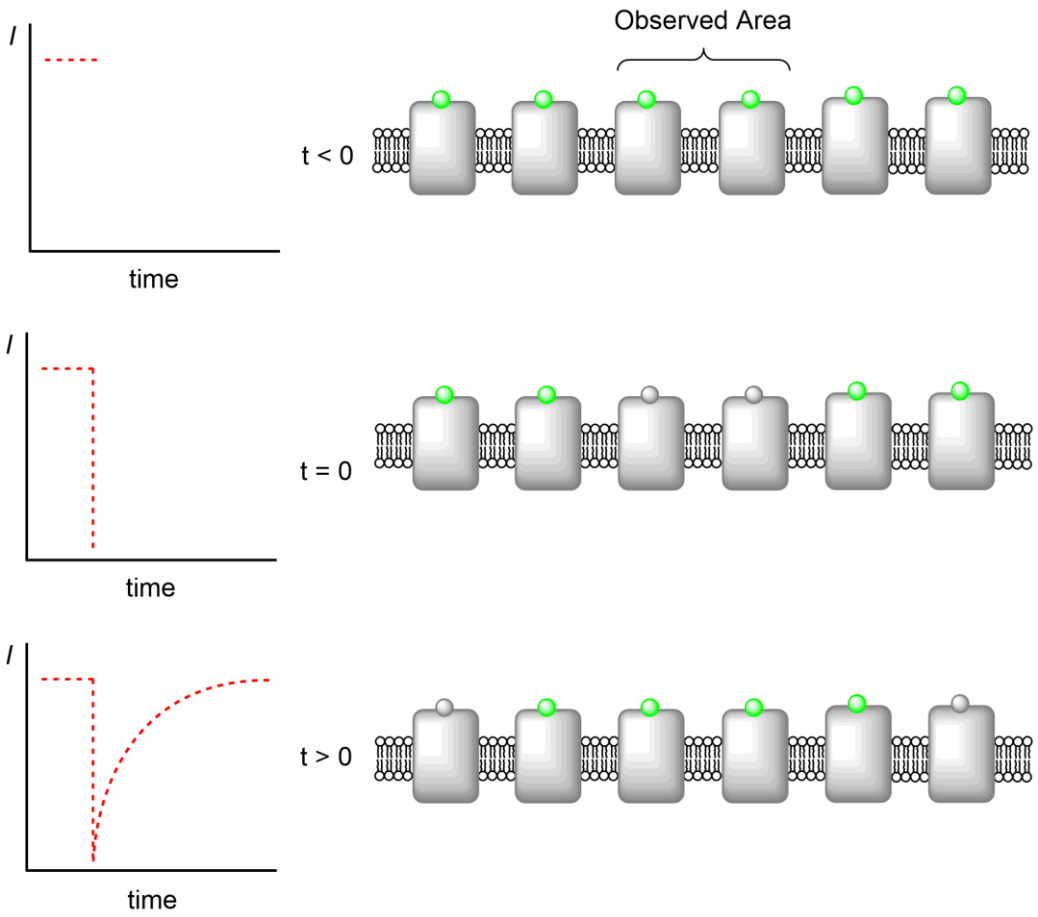


Figure 76 - Schematic diagram showing the FRAP experiment.

Due to the low number of FhuA receptors expressed on the cell surface of *E. coli*, a technique which possesses a high signal to noise ratio is required for visualisation of single receptors. Total internal reflection fluorescence (TIRF) microscopy is an analytical technique that utilises this physical property to observe a thin region of a specimen at the surface. Laser light is used to excite the sample, by shining it onto a quartz prism which refracts it through a quartz cover slip onto the sample. As the beam hits the sample it passes from a material of high optical density (quartz) to a material of low optical density (aqueous medium). As the angle of incidence is large the laser light is reflected back into the quartz prism, forming a evanescent wave at the interface. This wave has a high intensity at the surface, however decays exponentially, generating excited fluorophores close to the TIRF surface. This principle allows the selective excitation of surface fluorophores over bulk fluorophores which enhances the signal to noise ratio, giving a high sensitivity of the technique (Figure 77).[23-24]

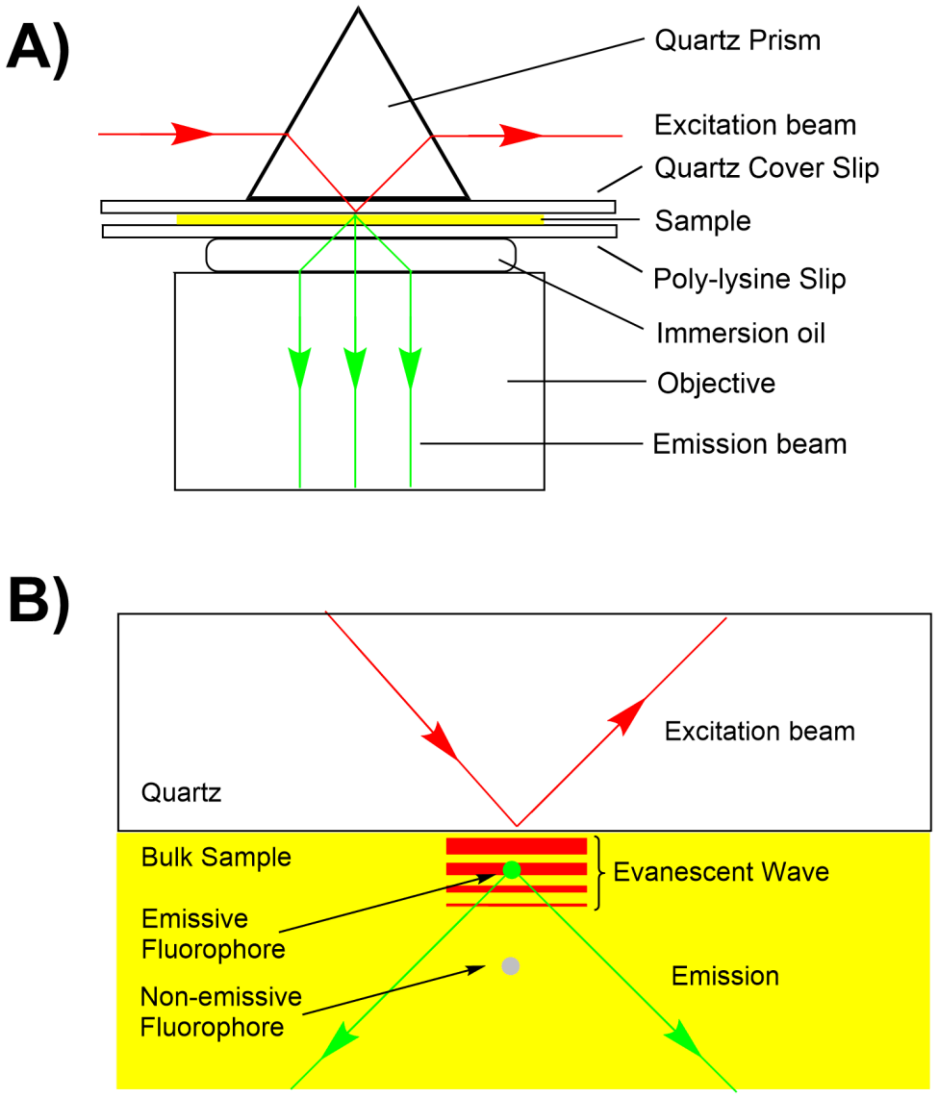
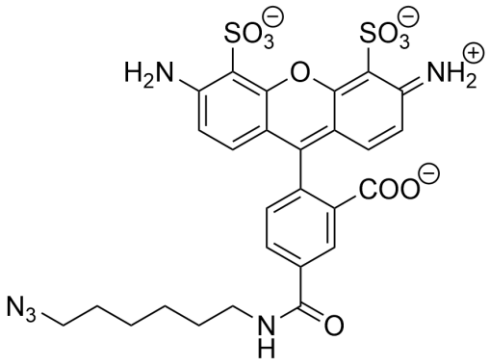


Figure 77 - Schematic representation of the TIRFM experimental set-up. A) General overview; B) Schematic representation of the evanescent wave formation.

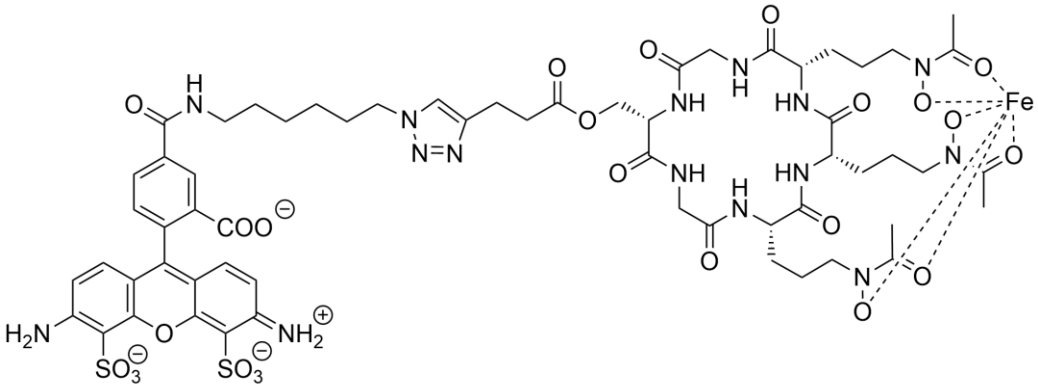
5.2 Synthesis

Due to the TIRF experimental set-up, which utilises an argon ion fixed wavelength laser source (488 nm), the fluorophore chosen for the design of the conjugate was Alexa Fluor® 488 azide (Life Technologies), which possesses an intense absorbance band at 488 nm (ε = 76,000 $mol^{-1}dm^{3}cm^{-1}$) and efficiently emits at 519 nm (ϕ = 0.92). This fluorophore also has a number of other desirable properties such as being water soluble, photostable, and pH independent between pH 4-10.[25]



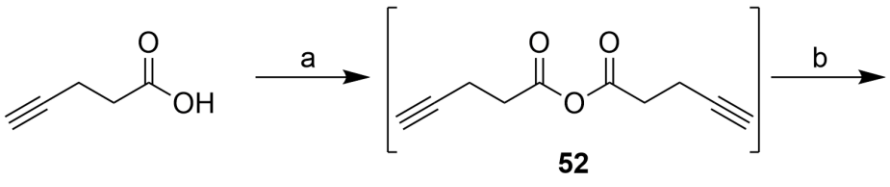
Alexa Fluor® 488 Azide

The fluorophore possesses a six carbon linker to the azide functionality, which reduces the likelihood of adverse steric interactions when the target conjugate is bound to the FhuA receptor. The azide functional group allows reaction with an alkyne using a copper catalysed 1,3-dipolar-cycloaddition reaction ('click-chemistry'). The required alkyne functionality can be introduced onto ferricrocin by forming an ester linkage using the alcohol of ferricrocin and an alkyne containing carboxylic acid. Compound **51** is the purposed structure of the target compound.

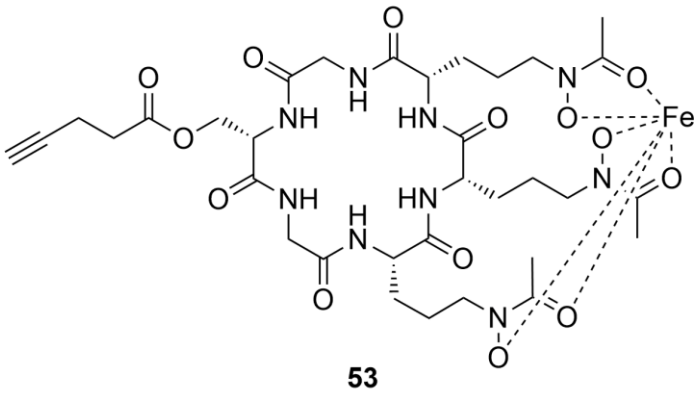


51

The synthesis starts with the addition of the alkyne functionality to ferricrocin (Scheme 22). Following the synthetic strategy employed by Coulton *et al.*[14] it was decided to first activate 4-pentynoic acid as an anhydride. This was synthesised through self condensation using DCC as a coupling reagent, following a modified methodology described by Ledin *et al.*[26] The 4-pentynoic acid was first reacted with DCC in anhydrous dichloromethane under a dry nitrogen atmosphere. The activated 4-pentynoic acid subsequently reacted with another 4-pentynoic acid molecule producing anhydride **52** and a DCU by-product. The DCU was removed whilst remaining under an inert atmosphere *via* cannular filtration. The ferricrocin was introduced dissolved in dry pyridine and allowed to react as described by Coulton *et al.*[14]



52



53

Scheme 22 - Synthetic route to **53**. (a) DCC, N_2; (b) Ferricrocin, N_2.

Due to the small scale of the synthesis (< 1 mg) and the use of ferricrocin which contains iron(III), NMR could not be used to characterise **53**. Therefore semi-preparative HPLC and mass spectrometry were utilised to purify and characterise the product (Figure 78).

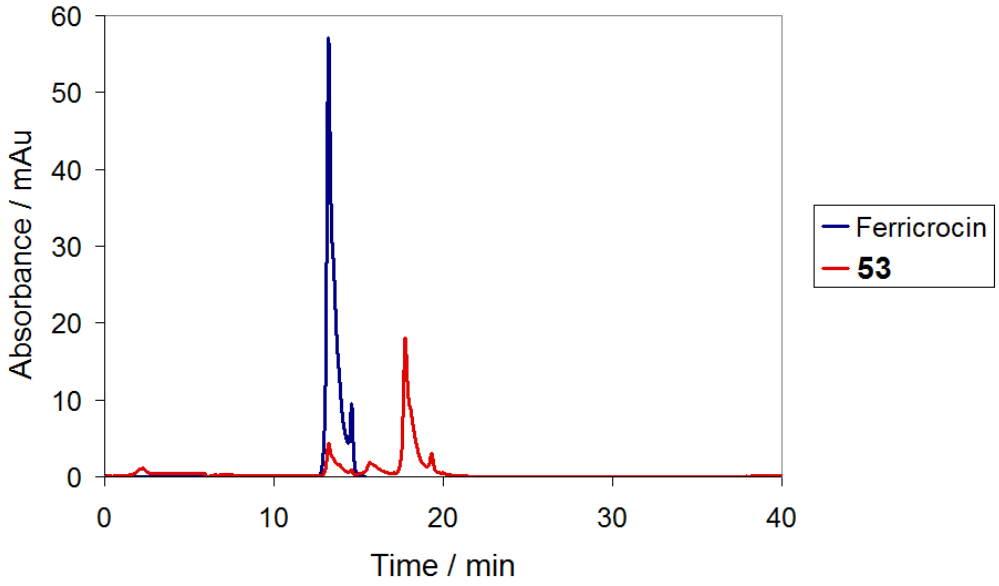
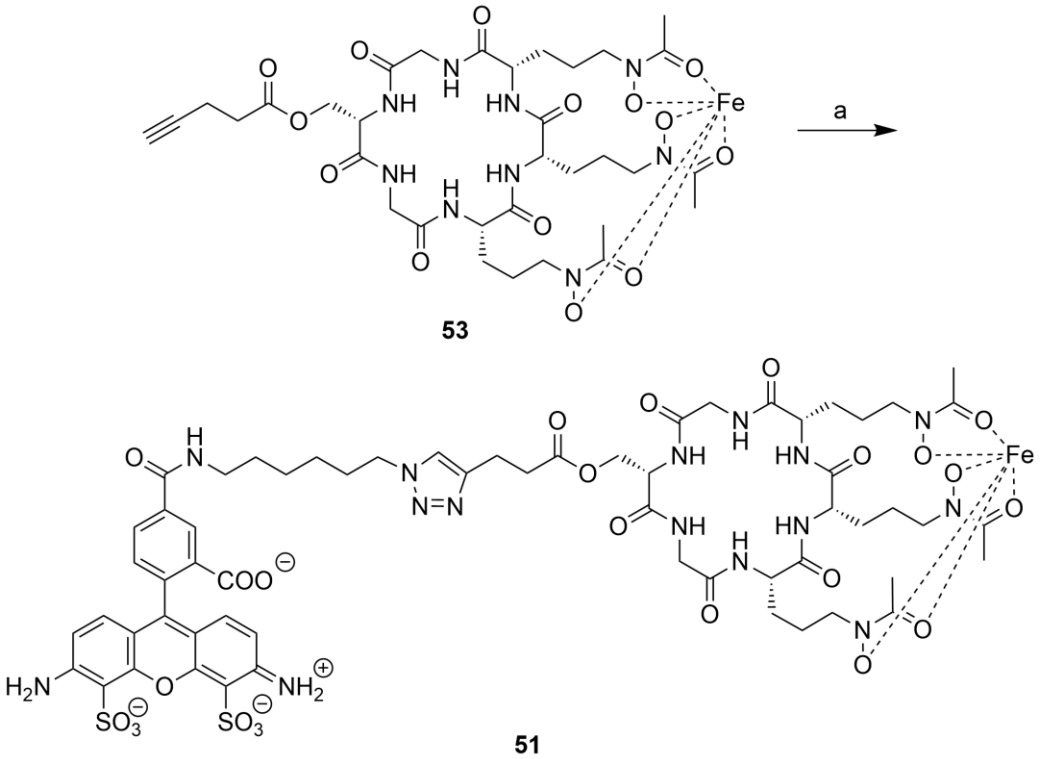


Figure 78 - HPLC traces (absorbance at 420 nm) of ferricrocin (20 mM) in water, injection volume 5 µL (blue) and **53** (0.5 mg in 350 µL of water) injection volume 20 µL (red); 6-40 % (v/v) acetonitrile/water gradient.

The HPLC traces show that ferricrocin has a retention time of 13 minutes under the experimental conditions, whereas the reaction mixture contained both ferricrocin and a more intense new species which eluted at 17 minutes. The sample was purified by HPLC and the fraction eluting at 17 minutes was collected. Mass spectrometry confirmed the presence of **53** with signals observed at *m/z* 851.2705 and *m/z* 873.2541 consistent with $[M+H]^+$ and $[M+Na]^+$ ions, respectively, for a molecular formula of $FeC_{33}H_{48}N_9O_{14}$.

The next step in the synthesis was the addition of the Alexa Fluor® 488 azide fluorophore onto **53**, using a copper-catalysed 1,3 dipolar-cycloaddition ('click'-chemistry) reaction (Scheme 23). This type of reaction is well established in the literature. It was discovered by Rostovtsev *et al.*[27] and has been used

extensively since, due to high yields and the wide range of chemical conditions in which the reaction can take place.[28] The copper acts as a catalyst and is first introduced into the reaction as a copper(II) salt, which is reduced *in situ* by sodium ascorbate to produce the catalytically active copper(I) form. The reaction is reported to work in a wide pH range and in a number of different solvents from aqueous tert-butyl alcohol to neat water.[27-28]



Scheme 23 - Synthetic route to **51**. (a) Alexa Fluor® 488 azide, $CuSO_4$, sodium ascorbate.

The reaction was conducted using water as the solvent, due to the water solubility of the reagents and requirement for the work-up using dialysis. The reaction was allowed to proceed for 6 hours, at room temperature and in the dark in order to reduce potential photobleaching. The reaction mixture was then dialysed using a 0.1-0.5 kDa molecular weight cut off dialysis membrane (*Dialysis Float-A-Lyzer 5 mL, Spectrum labs*) in-order to remove any remaining copper salts. The sample was analysed by HPLC under the same conditions as for **53** (Figure 79).

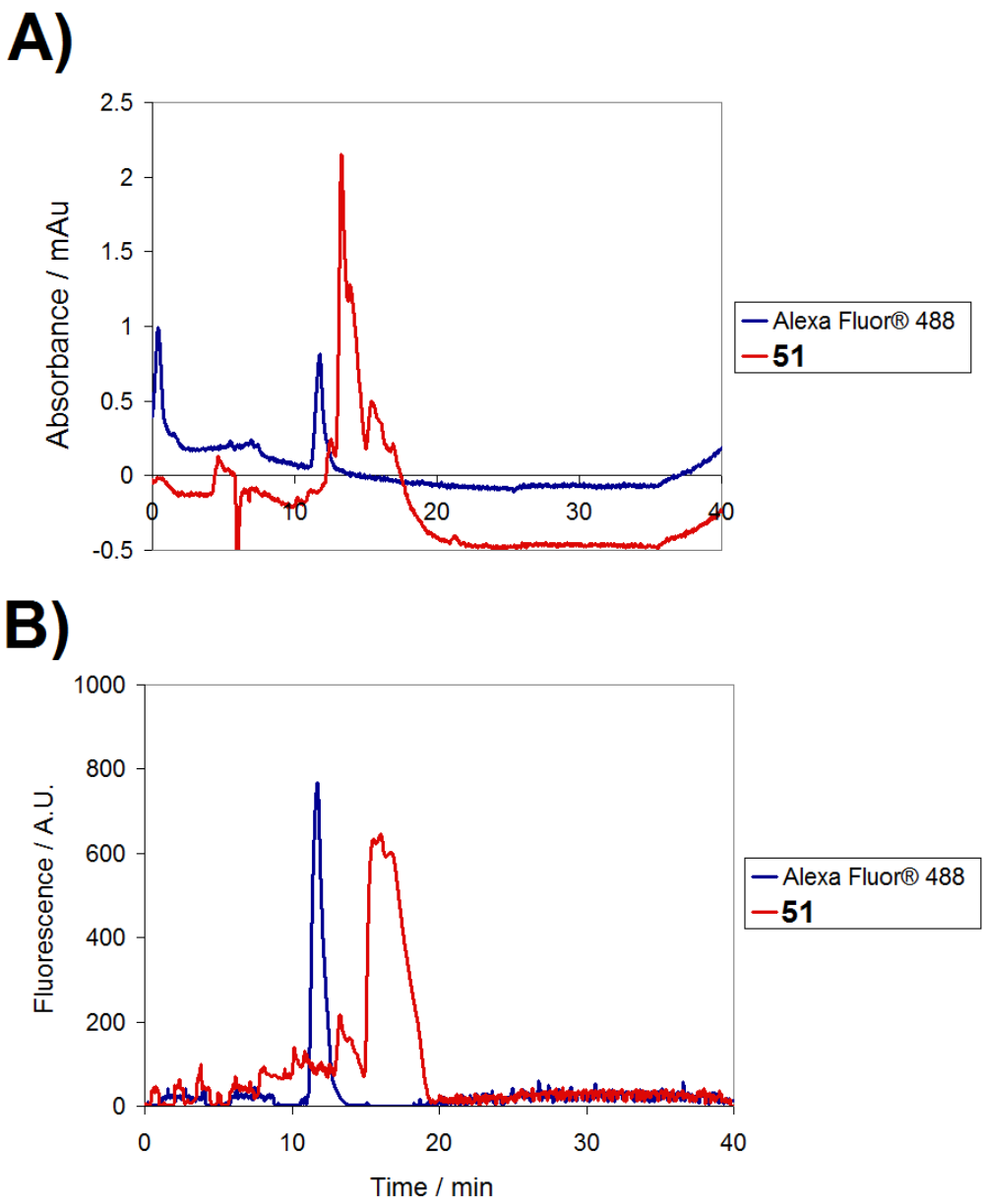


Figure 79 - HPLC traces of **51** (red) after dialysis and Alexa Fluor® 488 (blue). A) - Absorbance data measured at 420 nm. B) Fluorescence data excited at 500 nm emission measured at 520 nm.

The HPLC traces show that free Alexa Fluor® 488 azide has a retention time of 12 minutes observed by both the absorbance and fluorescence detectors. The reaction mixture has a strongly absorbing species with a retention time of 13 minutes, but only shows a weak emission intensity, and a weakly absorbing species with a large emission intensity at 16 minutes. The reaction mixture was purified using the same semi-preparative HPLC methodology as with **53**, collecting the band eluting at 16 minutes and mass spectrometry was utilised to

confirm that presence of the conjugate **51**. The broad nature of the emission peak in the HPLC trace of **51** demonstrated a need to confirm the purity of the conjugate. Therefore LC-MS with an UV-visible detector was used to identify any impurities. Due to the different experimental limitations for the LC-MS, the chromatography procedure was different to that used for the semi-preparative HPLC in the synthesis. Control experiments showed that free ferricrocin could be detected using both the UV-visible absorbance detector (254 nm) and negative mode ESI mass spectrometry. Therefore a sample of **51** was investigated using the same methodology as for free ferricrocin. There was only one peak in the total ion count (TIC) chromatogram, with three features in the UV-visible trace (Figure 80).

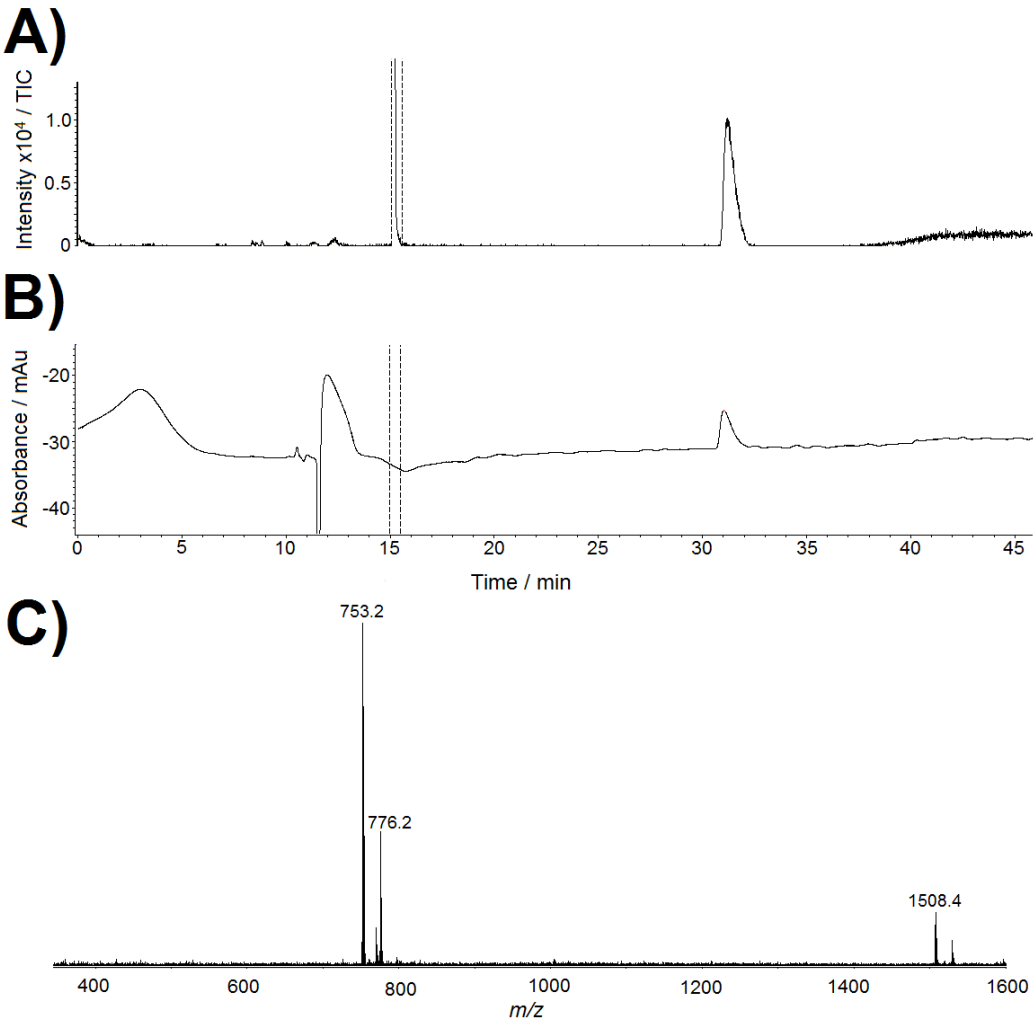


Figure 80 - LC-MS of **51**. A) TIC chromatogram; B) UV absorbance (254 nm) chromatogram; C) Mass spectrum associated with the peak eluting after 32 minutes in the chromatogram. Peak at 15 min (dashed black line) is calibrant injected into the ESI source used for accurate mass calibration.

The band in the chromatograms eluting after 32 minutes is assigned as **51**, due to the ESI signals observed at *m/z* 753.1849 and 1507.3876 consistent with the doubly charged, (M^{2-}) and singly charged (M^{-}) molecular ions for a compound of molecular formula $FeC_{60}H_{73}N_{15}O_{24}S_{2,}$ respectively. The other features in the UV-visible trace are a broad peak between 0-5 minutes and a sharp negative signal followed by a broad positive peak starting at 11 minutes. These features had no detectable ions in the TIC chromatogram, suggesting that the signal is due to the void of the column. This was confirmed by control blank chromatograms, with these features always present. The only species detected by either absorbance at 254 nm or mass spectrometry was conjugate **51** showing that the semi-preparative HPLC methodology for the purification was successful.

To further characterise **51**, the UV-visible, emission and excitation spectra were recorded (Figure 81). The profile of the UV-visible spectrum is similar to that of the Alexa Fluor® 488, as it is the most dominate chromophore in the molecule. Conjugate **51** does possess a larger absorbance than that of the Alexa Fluor® 488 chromophore between 380-420 nm due to the LMCT band of the ferricrocin component of the conjugate. Ferricrocin was found not to be fluorescent and therefore the profiles of the emission and excitation spectra of **51** match that of the Alexa Fluor® 488 fluorophore.

Due to the small scale of the synthesis an accurate mass could not be obtained for **51**. The final yield was therefore calculated from the UV-visible absorbance spectrum using the molar absorption coefficient of the Alexa Fluor® 488 azide ($\lambda_{488,}$ ε = 76,000 $mol^{-1}dm^{3}cm^{-1}$), to be $4.2x10^{-9}$ mol (6.3 μg).

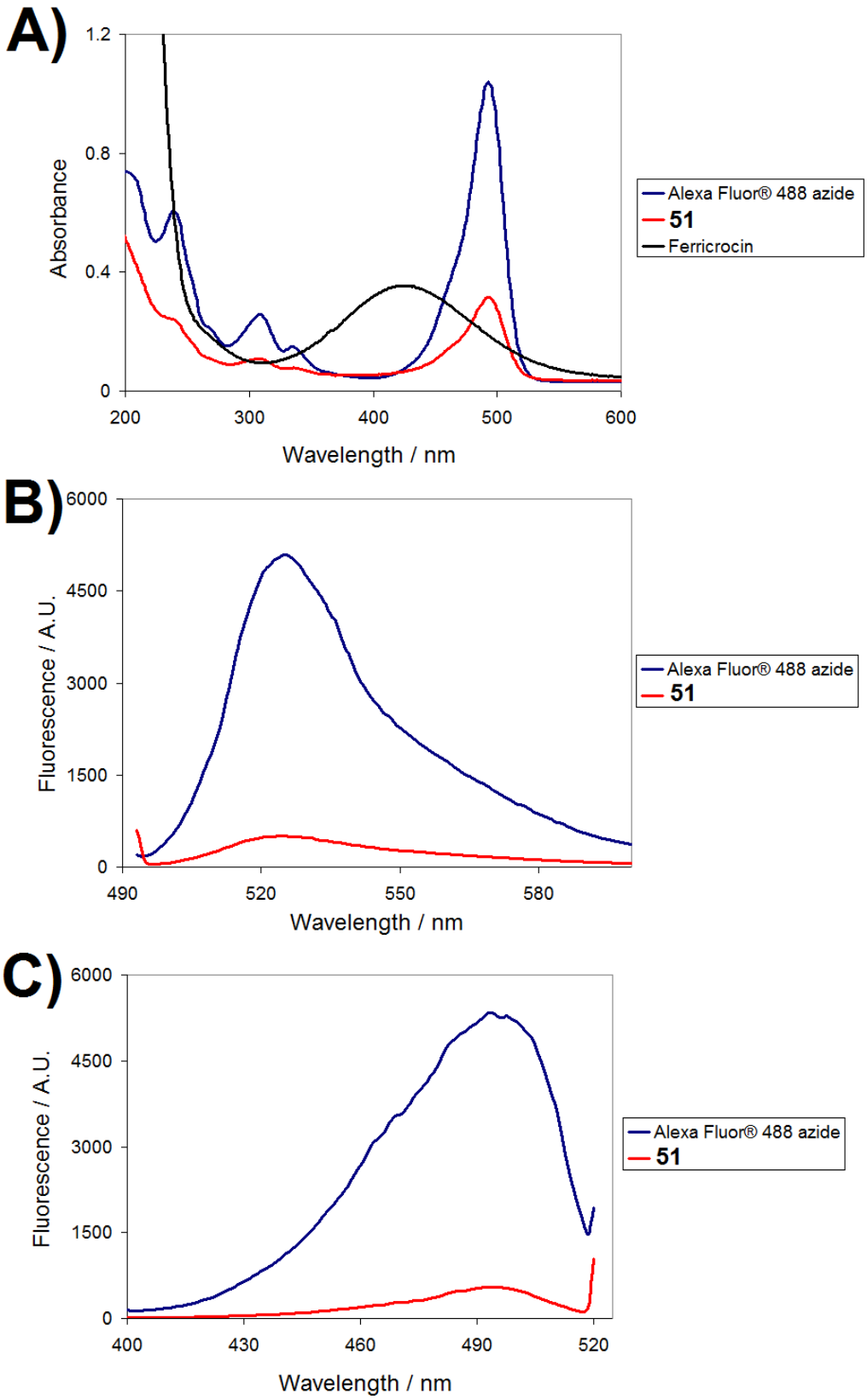


Figure 81 - A) UV-visible absorbance spectra of Alexa Fluor® 488 azide ($1.3x10^{-5}$ $moldm^{-3}$), **51** and ferricrocin ($2.0x10^{-4}$ $moldm^{-3}$); B) Emission spectra of Alexa Fluor® 488 azide and **51**, excited at 488 nm; C) Excitation spectra of Alexa Fluor® 488 azide and **51**, monitoring at 525 nm. All spectra recorded in water.

5.3 Conjugate-Protein Interactions

The work outlined within this section was conducted by Dr Christoph Baumann, Department of Biology, University of York.

To determine if **51** will selectivity bind to the outer membrane receptor FhuA as outlined in the introduction, TIRF microscopy was utilised. Two strains of *E. coli* were grown and used for this experiment, a Δ*fhuA* strain and a Δ*tonB* strain. The Δ*fhuA* strain provided a negative control: due to the lack of FhuA receptors expressed, conjugate **51** should not bind to these cells. The Δ*tonB* strain provided the positive control as **51** should bind to the FhuA receptor but not be internalised into the cell and hence remains detectable under the TIRF microscopy experimental conditions.

A sample of each cell culture was taken and centrifuged, the supernatant was removed and the cells subsequently resuspended in M9 medium containing 10 nM of conjugate **51**. The cells were incubated in the dark for 20 minutes, before being centrifuged to remove any supernatant, before resuspended in fresh M9 medium containing no **51**. The cells where then immobilised onto a poly-lysine slip and images recorded under the TIRF microscope (Figure 82).

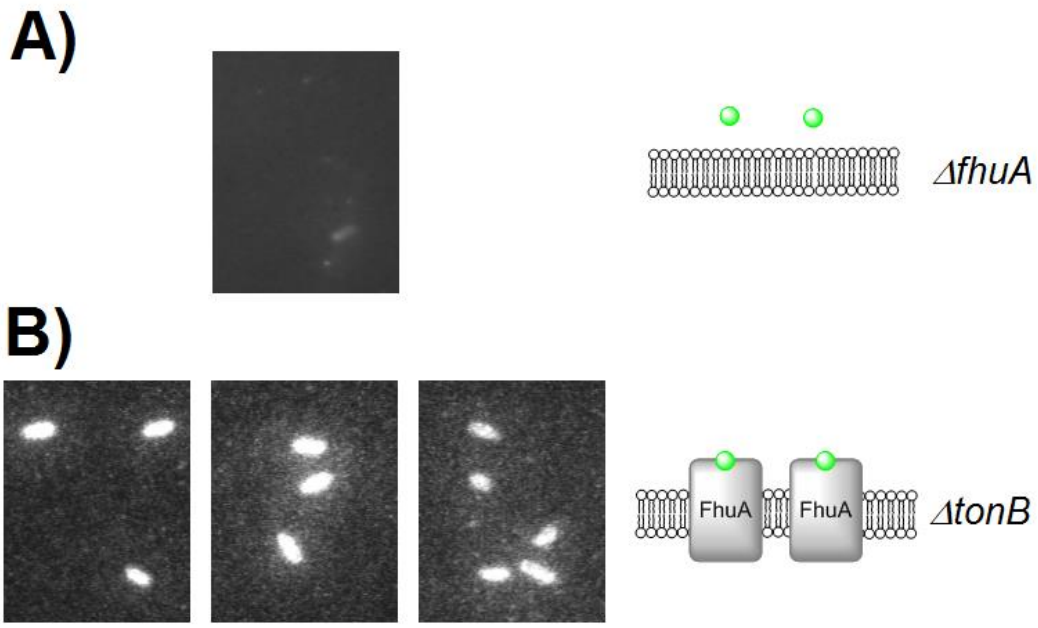


Figure 82 - Initial TIFM images. A) 10 nM **51** and *E. coli ΔfhuA* cells (sum of 500 video frames); B) 10 nM **51** and *E. coli ΔtonB* cells (sum of 5 video frames).

The TIRFM images show the successful labelling of the *E. coli* cells that contained the FhuA receptor. To further test the hypothesis of the specific labelling of the receptor, the experiments were repeated, but with a third experimental condition. This time the Δ*tonB* cells were incubated with 5 μM of ferricrocin, before the incubation of 10 nM of conjugate **51**. The unlabelled ferricrocin, should bind and block the FhuA receptors from **51** resulting in a weak fluorescence signal (Figure 83).

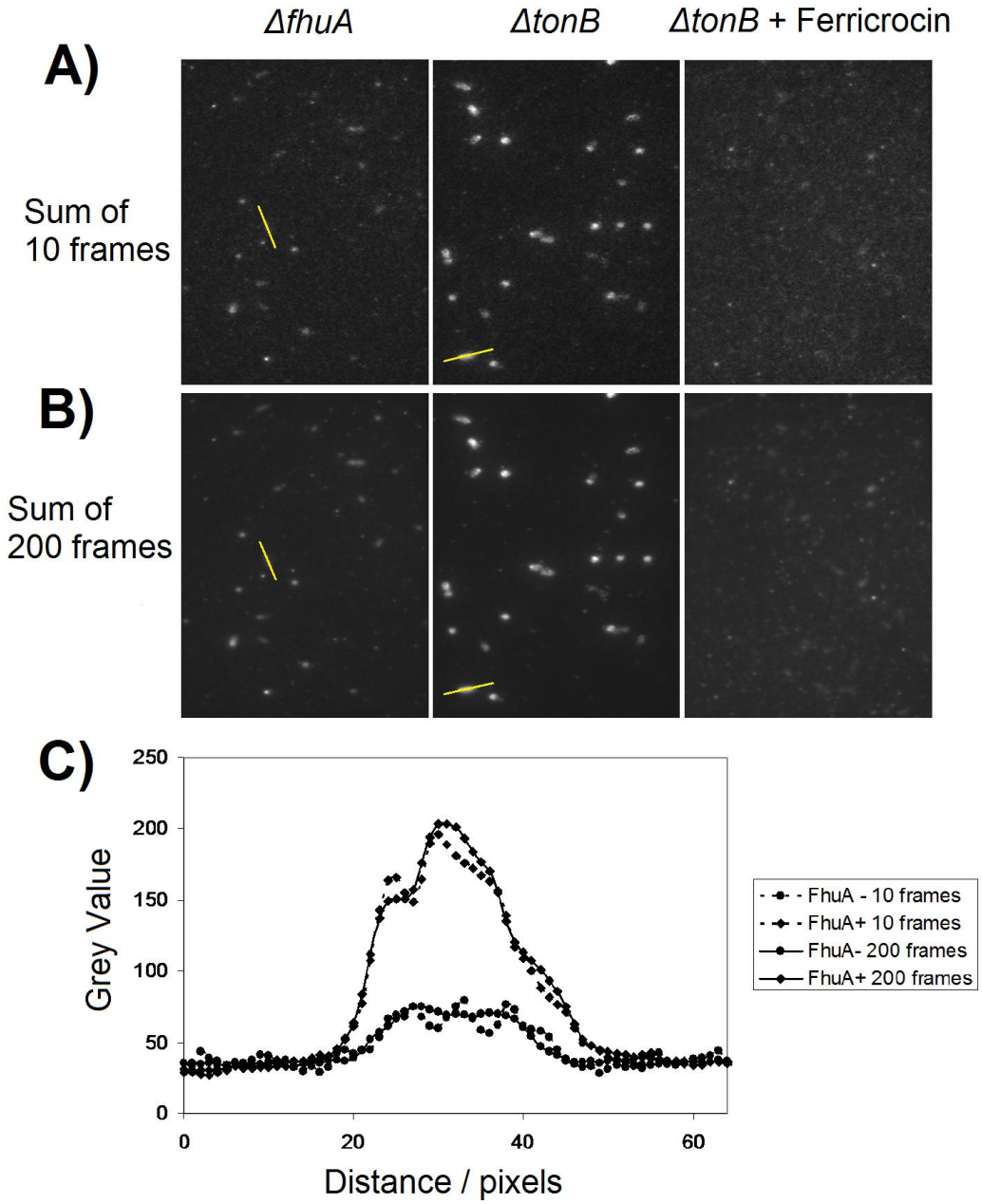


Figure 83 - Repeat TIRFM images containing 10 nM **51**, *E. coli ΔfhuA*, *ΔtonB* cells and *ΔtonB* cells with 5 μM ferricrocin. A) sum of 10 video frames; B) sum of 200 video frames; C) Profile plot of grey value of pixels across highlighted cell (yellow). Analysis performed using ImageJ.[29]

The repeat TIRFM images show the ferricrocin block reduces the fluorescence of the *E. coli* cells, showing that conjugate **51** specifically binds to the FhuA outer membrane receptor. The intensity of the fluorescence signal, however, is weaker than expected, possibly due to quenching of the fluorophore by the ferricrocin moiety. As a result, the planned FRAP work could not be completed

as the conjugate would not be compatible with the confocal microscope equipment required for the experiment.

5.4 Summary of Chapter 5

The ferricrocin-Alexa Fluor® 488 conjugate **51** was successfully synthesised in a two step procedure. First ferricrocin was reacted with 4-pentynoic anhydride, followed by click-chemistry to react the alkyne group with the commerically available Alexa Fluor® 488 azide fluorophore. Characterisation by LC-MS, UV-visible and fluorescence spectroscopies confirmed the successful formation of **51**. Due to the small scale of the synthesis yields were not obtainable. Conjugate **51** was incubated with two *E. coli* strains, Δ*fhuA* and Δ*tonB* and analysed using TIRF microscopy. The Δ*fhuA* strain showed a weak fluorescence, whereas the Δ*tonB* strain showed observable fluorescence under these experimental conditions, indicating compound **51** is selectivity binding to cells which possess the outer membrane transporter FhuA.

Chapter 6: Conclusions and Future Work

6.1 Conclusions and Future Work

Interactions of Tetradentate Siderophores with CeuE (Chapters 2 and 3)

The simple bis(catecholamide) biomimetic model, H_4-4-LICAM was the first tetradentate ligand to be characterised, complexed with iron(III) and co-crystallised with a PBP, leading to the discovery of a novel binding mode of CeuE. Ferric-4-LICAM was found to be present as a 1:1 complex in the binding pocket of CeuE, with residues His227 and Tyr288 coordinating directly to the iron centre, thus completing its coordination sphere. Sequence alignment revealed that these two residues are conserved among a number of related proteins from different bacteria. In the absence of CeuE, it was found that 4-LICAM^{4-} is capable of forming complexes with a ligand to metal ratio of both 3:2 and 1:1 in equilibrium, however, only the 1:1 complex is found bound if the complex is co-crystallised with CeuE. The chirality of the iron centre in the {CeuE[Fe(4-LICAM)]} structure was found to be Λ, both in the solid state (X-ray diffraction analysis) and in solution (CD spectroscopy). The dissociation constant was determined by intrinsic fluorescence quenching to be 29.3 ± 11.7 nM, which is the first time a ferric-tetradentate siderophores dissociation constant has been determined.

To confirm if the binding mode of the mimetic is biologically important, the successful total synthesis of the natural siderophore enterobactin linear dimer was carried out. It was shown that the linear dimer preferentially binds iron(III) in a 1:1 complex in solution, unlike 4-LICAM^{4-}, and that the basic arrangement within the binding pocket of CeuE in the crystal is the same. The dissociation constant was determined to be 8.4 ± 4.3 nM. The similarity between the dissociation constants of PBPs with ferric-4-LICAM, ferric-linear dimer and other characterised hexadentate siderophores, suggest that the tyrosine and histidine residues act as an adapter for the protein, enabling the binding of both tetradentate and hexadentate ligands with minimal impact on the binding mode (e.g. dissociation constant and metal centred chirality).

Future work should first involve determination of the dissociation constant of ferric-enterobactin towards CeuE, which would reveal whether the ferric-4-LICAM and ferric-linear dimer K_d values are within a biologically relevant range for CeuE or not. Selective mutagenesis could be used to confirm the importance of the His227 and Tyr288 residues, with determination of the mutant dissociation constants indicating the relative contribution to the K_d value for each residue. Beyond these experiments, expanding the research to include PBPs from other organisms, as well as other proteins in the uptake system would be of great interest.

Development of a Siderophore Mimic for use with Further Functional Components (Chapter 4)

The second area of research involved the development of a novel catecholamide hexadentate siderophore mimic that possesses a linker for the attachment of a secondary chemical functional group. The successful synthesis used 4-(2-carboxyethyl)-4-nitroheptanedioic acid as a starting material, in eight synthetic steps to produce the final compound, with an overall yield of 10.5%.

Future work would include full optimisation of the synthetic route, particularly the coupling of the catecholamide groups to the backbone, which is likely to be hampered by hydrolysis of the acid chloride. After full optimisation, the iron(III) binding abilities of H_6-**41** could be investigated. In addition, a second functional component could be covalently attached *via* the free carboxylate, to enable evaluation of its effect on the iron(III) binding of the siderophore moiety. Once the ligand-metal interactions are understood, the interactions between the protein and the ferric siderophore could be investigated, using CeuE as a model protein. Determination of the metal-centred chirality and the dissociation constant of the CeuE-ferric-**41** complex would provide useful insight into the protein recognition of ferric-**41**, and how this compares to ferric-enterobactin and other siderophore models. The second functional component attached could be either: a fluorophore, for use as a novel fluorescent probe to target

enterobactin receptors or an anti-microbial agent for use in the Trojan Horse Strategy.

Modification of Ferricrocin with a Fluorescent Component (Chapter 5)

The final area of research involved modification of a natural siderophore, ferricrocin, by the attachment of the Alexa Fluor® 488 fluorophore. This was completed by the attachment of an alkyne to the free alcohol functionality of the ferricrocin, followed by a 1,3 dipolar-cycloaddition ('click'-chemistry) reaction to introduce the Alexa Fluor® 488 fluorophore *via* the azide group. The conjugate was characterised by LC-MS, UV-visible and fluorescence spectroscopies. In collaborative experiments conjugate **51** was incubated with two strains of *E. coli*, Δ*fhuA* and Δ*tonB* and analysed using TIRF microscopy. The Δ*fhuA* strain showed a weak fluorescence, whereas the Δ*tonB* strain showed a stronger fluorescence signal, indicating selective binding of **51** to cells which possess the outer membrane transporter FhuA. However, the fluorescence signal was not considered intense enough for the intended FRAP experiments.

Future work would involve change from the current conjugate **51**. Increasing the link length between the ferricrocin and the fluorophore may reduce the impact of the quenching, or alternative fluorophores could be coupled with ferricrocin using similar synthetic strategies.

Chapter 7: Experimental

7.1 General Remarks

Materials

All materials were obtained from commercial suppliers (Acros, Alfa-Aesar, Fisher Chemicals, Fisher Scientific, Fluka, Sigma-Aldrich) and used as supplied unless otherwise stated.

Instrumentation

^{1}H, ^{13}C{^{1}H}, DEPT 135, HSQC, COSY, HMBC NMR spectra were recorded on a Jeol EX and ES 400 MHz instruments (^{1}H NMR 400 MHz, ^{13}C NMR 100.6 MHz). DOSY spectra were recorded on a Bruker 500 MHz instrument (^{1}H NMR 500 MHz) and the experiments were carried out by Ms. Heather Fish. Multiplicity abbreviations are as follows; s for singlet, d for doublet, t for triplet, q for quartet quin for quintet, dd for double doublet, td triple doublet, ddd for doublet of doublet of doublets, m for multiplet, br for broad and app for apparent. All NMR assignments were aided by DEPT 135, HSQC experiments, with additional COSY and HMBC experiments used as required. Electro-spray ionisation mass spectrometry (ESI-MS) was recorded on a Bruker microTOF Electrospray mass spectrometer by Mr. K. Heaton and Ms. H. Robinson. Infra-red spectra were recorded on a Thermo Nicolet Avatar 370 FT-IR spectrophotometer (KBr), or on a Perkin Elmer FT-IR Spectrum Two spectrometer (ATIR) in the region of 4000-500 cm^{-1}. Elemental analysis was carried out by Dr. G. McAllister on an Exeter CE-440 elemental analyser and results are within ± 0.4 %. TLC was performed on Merck silica gel 60 F_{254} aluminium backed plates and visualised under a Chromato-vue Model CC-10 UV lamp, unless otherwise stated. Melting points were recorded on a Stuart Scientific SMP3 instrument. Electronic absorbance spectra were recorded on a Shimadzu UV-1800 spectrophotometer, with kinetic data recorded on a Hewlett Packard 8452A diode array spectrophotometer with a thermostated cell holder. Fluorescence spectra were recorded on a Hitachi F-4500 fluorescence spectrophotometer. The specific rotation was recorded on a Jasco DIP-370 digital polarimeter. The circular dichromism spectroscopic measurements were

performed on a Jasco J810 CD spectropolarimeter at 20 °C under constant nitrogen flush.

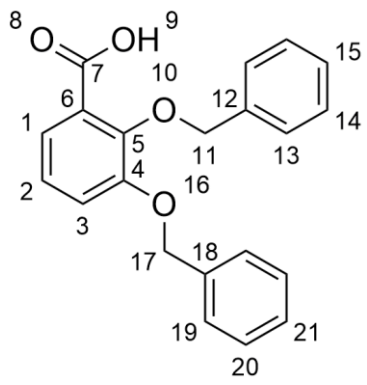
7.2 Synthesis

The order of the experimental details of each molecule is shown in Table 14.

Compound Number (Target Compound)	**Chapter of Thesis**	**Page Number**
12 (Catechol Synthesis)	2	204
13 (Catechol Synthesis)	2	205
16 (H_2-EBA Synthesis)	2	206
15 (H_2-EBA Synthesis)	2	208
14 (H_4-4-LICAM Synthesis)	2	209
H_4-4-LICAM (H_4-4-LICAM Synthesis)	2	211
22 (H_4-**17**)	3	212
24 (H_4-**17**)	3	214
25 (H_4-**17**)	3	215
29 (H_4-**17**)	3	217
30 racemic (H_4-**17**)	3	219
31 (Catechol Synthesis)	3	221
32 (H_4-**17**)	3	221
33 (H_4-**17**)	3	223
34 (H_4-**17**)	3	225
30 L/L (H_4-**17**)	3	226
H_4-**17** (H_4-**17**)	3	228
43 (H_6-**41**)	4	230
44 (H_6-**41**)	4	231
45 (H_6-**41**)	4	232
46 (H_6-**41**)	4	234
47 (H_6-**41**)	4	235
48 (H_6-**41**)	4	237
49 (H_6-**41**)	4	239
H_6-**41** (H_6-**41**)	4	240
53 (**51**)	5	242
51 (**51**)	5	243

Table 14 - Ordering of synthetic experimental details.

2,3-Bis(phenylmethoxy)benzoic acid, 12



Prepared as in accordance to the literature.[1]

Yield:

10.01 g, 29.9 mmol, 80%

Mp:

124.1-125.9 °C

R_f:

Chloroform : MeOH (8:1) : R_f : 0.51

***m/z* (ESI):**

357.1097 ($[M+Na]^+$, 97%), 335.1289 ($[M+H]^+$, 13%)

HRMS (ESI):

Calc for $[C_{21}H_{18}O_4] + H^+$ = 335.1278 Found 335.1289 (-3.3 ppm error)

Calc. for $[C_{21}H_{18}O_4] + Na^+$ = 357.1097 Found 357.1097 (0.1 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

12.93 (1H, s, H-***9***), 7.49 (2H, d, *J* = 7.3 Hz, H_{ar}), 7.42-7.29 (9H, m, H_{ar}), 7.22 (1H, dd, *J* = 7.8, 1.5 Hz, H-***1/3***), 7.15 (1H, t, *J* = 7.8 Hz, H-***2***), 5.19 (2H, s, H-***11/17***), 5.00 (2H, s, H-***11/17***)

^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

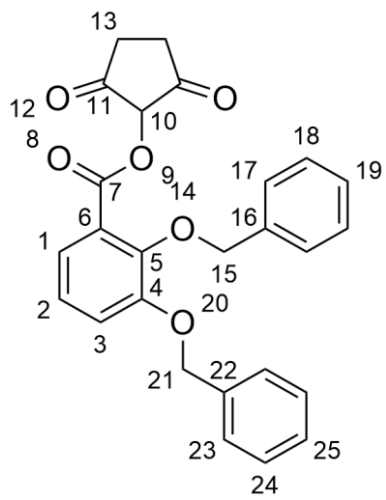
168.2 (C-***7***), 152.9 (C-***4/5***), 147.2 (C-***4/5***), 137.9 (C_{ar}), 137.3 (C_{ar}), 128.9 (C_{ar}), 128.7 (C_{ar}), 128.6 (C_{ar}), 128.5 (C_{ar}), 128.3 (C_{ar}), 128.2 (C_{ar}), 124.7 (C_{ar}), 121.9 (C_{ar}), 117.5 (C_{ar}), 74.9 (C-***11/17***), 70.3 (C-***11/17***)

IR (KBr cm^{-1}):

3431 *weak* (O-H), 3032 *weak* (C-H), 1692 *strong* (C=O, acid), 1577 *strong* (C-C ar)

The characterisation data are consistent with those reported.

2,3-Bis(phenylmethoxy)-benzoic acid 2,5-dioxo-pyrrolidinyl ester, 13



Prepared as in accordance to the literature.[2]

Yield:

2.58 g, 5.98 mmol, 72%

Mp:

115.7 – 116.9 °C

R_f:

Chloroform : Ethyl Acetate (4:1) : R_f : 0.47

m/z (ESI):

454.1248 ([M+Na]$^+$, 100%)

HRMS (ESI):

Calc. for [$C_{25}H_{21}NO_6$] + Na^+ = 454.1261 Found 454.1248 (2.9 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

7.61-7.29 (13H, m, H_{ar}), 5.25 (2H, s, H-***15/21***), 5.04 (2H, s, H-***15/21***), 2.89 (4H, s, H-***13***)

^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

171.2 (C-***11***), 161.5 (C-***7***), 153.2 (C-***4/5***), 148.6 (C-***4/5***), 137.2 (C_{ar}), 136.9 (C_{ar}), 129.0 (C_{ar}), 128.9 (C_{ar}), 128.6 (C_{ar}), 128.6 (C_{ar}), 128.5 (C_{ar}), 128.5 (C_{ar}), 125.4 (C_{ar}), 122.6 (C_{ar}), 121.0 (C_{ar}), 120.5 (C_{ar}), 75.3 (C-***15/21***), 70.6 (C-***15/21***), 25.5 (C-***13***)

IR (KBr cm^{-1}):

1768 *strong* (C=O), 1738 *strong* (C=O), 1499 *medium* (C-C ar)

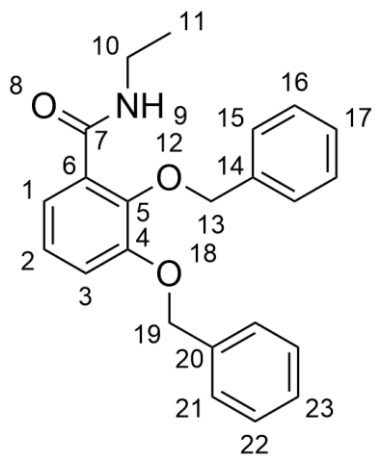
Elemental Analysis:

Calculated for [$C_{25}H_{21}NO_6$]: %C 69.60, %H 4.91, %N 3.25

Measured for [$C_{25}H_{21}NO_6$]: %C 69.21, %H 4.92, %N 3.23

The characterisation data are consistent with those reported.

2,3-Bis(phenylmethoxy)-*N*-ethylbenzamide, 16



Compound **13** (0.203 g, 0.47 mmol) was dissolved in 10 mL of THF. To this stirred solution, triethylamine (85 μL, 0.6 mmol) was added, followed by 300 μL of a 2.0 M solution of ethylamine in THF. The solution was stirred at room temperature for 4 hours. The volatiles were then removed *in vacuo* before the residue was taken up in 50 mL of dichloromethane. The organic layer was washed with 15 mL 0.1 M HCl, followed by 15 mL saturated $NaHCO_3$ and 3 x

20 mL water. The organic layer was dried over $MgSO_4$, before the dichloromethane was removed *in vacuo*, to yield the product as an oil. Characterisation data consistent with literature.[3]

Yield:

0.1243 g, 0.34 mmol, 72 %

R_f:

Chloroform : Ethyl Acetate (4:1) : R_f : 0.64

m/z (ESI):

362.1745 ($[M+H]^+$, 100%), 384.1561 ($[M+Na]^+$, 37%)

HRMS (ESI):

Calc for $[C_{23}H_{23}NO_3]+H^+$ = 362.1751 Found 362.1745 (1.7 ppm error)

Calc. for $[C_{23}H_{23}NO_3]+Na^+$ = 384.1570 Found 384.1561 (2.3 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

7.92 (1H, br t, J = 5.0 Hz, H-***9***), 7.77-7.74 (1H, m, H_{ar}), 7.50-7.14 (12H, m, H_{ar}), 5.16 (2H, s, H-***14/28***), 5.07 (2H, s, H-***14/28***), 3.32 (2H, dq, J = 7.0, 5.0 Hz, H-***10***), 0.99 (3H, t, J = 7.0 Hz, H-***11***)

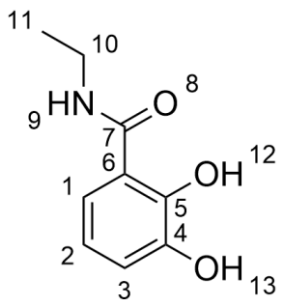
^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

165.3 (C-***7***), 152.1 (C-***4/5***), 147.1 (C-***4/5***), 136.7 (C_{ar}), 136.6 (C_{ar}), 129.1 (C_{ar}), 128.9 (C_{ar}), 128.5 (C_{ar}), 127.9 (C_{ar}), 127.7 (C_{ar}), 124.6 (C_{ar}), 123.5 (C_{ar}), 116.9 (C_{ar}), 76.3 (C-***14/28***), 71.2 (C-***14/28***), 34.2 (C-***10***), 14.0 (C-***11***)

IR (KBr cm^{-1}):

3385 *medium* (N-H), 2932 *weak* (C-H), 1652 *strong* (C=O, amide), 1526 *strong* (C-C ar)

***N*-ethyl-2,3-dihydroxybenzamide, 15**



All glassware used was soaked in 6M HCl for at least 2 hours before use, in order to remove iron adsorbed on the surface of the glassware.

Compound **16** (4.1 mmol, 1.480 g) was dissolved in 35 mL of dry ethanol. To the solution, one small spatula tip of Pd-C 10% was added and a balloon of hydrogen gas was passed through the solution to purge the system. After the system is purged a second balloon of hydrogen gas was added which allowed the reaction to proceed for 3 days under a positive pressure of hydrogen. After the reaction was completed the catalyst was filtered off and the product was isolated by removing the solvent *in vacuo*. The product was purified by sublimation at 10^{-2} mbar at 80 °C to yield a white crystalline solid. Characterisation data consistent with literature.[4]

Yield:

0.547 g, 3.02 mmol, 74 %

Mp:

78.2-78.8 °C

R_f:

Chloroform : Ethyl Acetate (4:1) : R_f : 0.38

m/z (ESI):

182.0810 ($[M+H]^+$, 50%), 204.0624 ($[M+Na]^+$, 100%)

HRMS (ESI):

Calc for $[C_9H_{11}NO_3]+H^+$ = 182.0812 Found 182.0810 (1.2 ppm error)

Calc. for $[C_9H_{11}NO_3]+Na^+$ = 204.0631 Found 204.0624 (3.5 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

13.94 (1H, br s, H-***12/13***), 9.13 (1H, br s, H-***12/13***), 8.80, (1H, t, *J* = 5.2 Hz, H-***9***), 7.27 (1H, dd, *J* = 8.0, 1.5 Hz, H-***1/3***), 6.90 (1H, dd, *J* = 8.0, 1.5 Hz, H-***1/3***), 6.66 (1H, app t, *J* = 8.0 Hz, H-***2***), 3.31(2H, dq, *J* = 7.2, 5.5 Hz, H-***10***), 1.13 (3H, t, *J* = 7.5 Hz, H-***11***).

^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

170.3 (C-***7***), 150.4 (C-***4/5***), 146.8 (C-***4/5***), 119.2 (C_{ar}), 118.3 (C_{ar}), 117.5 (C_{ar}), 115.3 (C_{ar}), 38.6 (C-***10***), 26.3 (C-***11***).

IR (KBr cm^{-1}):

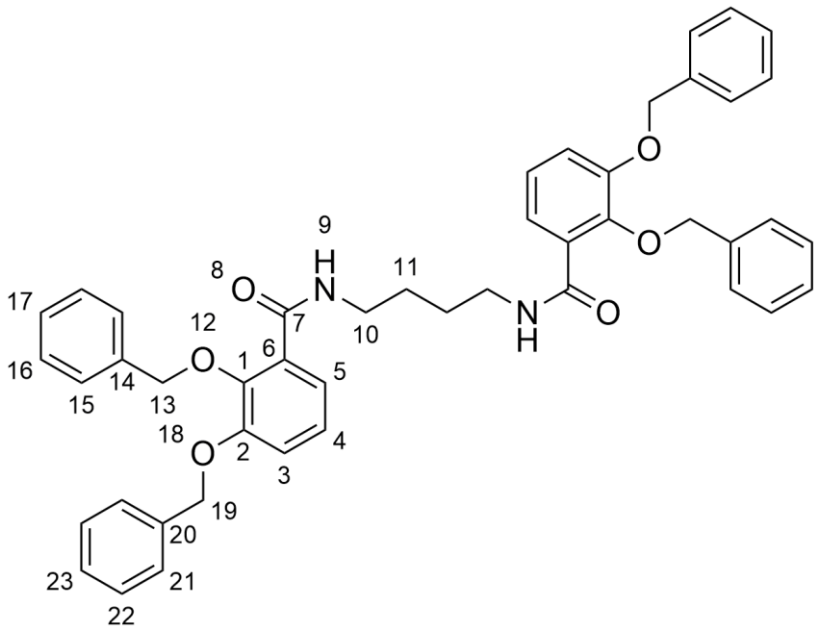
3350 *medium* (O-H), 3223 *medium* (N-H), 2982 *weak* (C-H), 1641 *medium* (C=O, amide), 1575 *medium* (C-C ar)

Elemental Analysis:

Calculated for [$C_9H_{11}NO_3$]: %C 59.66, %H 6.12, %N 7.73

Measured for [$C_9H_{11}NO_3$]: %C 59.62, %H 6.23, %N 7.69

***N,N'*-[bis(2,3-bis(phenylmethoxy)benzoyl]-1,4-diaminobutane, 14**



1,4-diaminobutane, 50 µL (0.5 mmol) was dissolved in 5 mL acetone also containing triethylamine, 350 µL (2.5 mmol) and stirred. To this stirred solution, compound **13** (0.431 g, 1.0 mmol), dissolved in 5 mL of acetone was

slowly added and the resulting mixture was stirred at ambient temperature for 18 hours. The solvent was removed *in vacuo*. The solid formed was dissolved in chloroform and washed with 0.05 M HCl. The organic layer was washed with saturated $NaHCO_3$ and H_2O until the aqueous washes were neutral. The organic layer was dried over $MgSO_4$, filtered, and the solvent removed *in vacuo* to yield the crude product. The crude product was purified *via* flash column chromatography on silica 4:1 $CHCl_3$:EtOAc, with the solvent system changing to 9:1 $CHCl_3$:MeOH to elute the product yielding an off white solid.

Yield:

0.336 g, 0.47 mmol, 94%

Mp:

122.1-124.9 °C

R_f:

Chloroform : Ethyl Acetate (1:1) : R_f : 0.43

***m/z* (ESI):**

721.3268 ($[M+H]^+$, 100%), 743.3090 ($[M+Na]^+$, 60%)

HRMS (ESI):

Calc for $[C_{46}H_{44}N_2O_6]+H^+$ = 721.3272 Found 721.3268 (0.6 ppm error)

Calc. for $[C_{46}H_{44}N_2O_6]+Na^+$ = 743.3092 Found 743.3090 (0.2 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

8.17 (2H, t, J = 6.0 Hz, H-***9***), 7.51 (4H, d, J = 8.0 Hz, H_{ar}), 7.43-7.27 (18H, m, H_{ar}), 7.14-7.12 (4H, m, H_{ar}), 5.20 (4H, s, H-***13/19***), 5.01 (4H, s, H-***13/19***), 3.14 (4H, br q, J = 5.6 Hz H- ***10***), 1.40-1.36 (4H, br m, H-***11***)

^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

166.1 (C-***7***), 152.2 (C-***1/2***), 145.7 (C-***1/2***), 131.5 (C_{ar}), 128.9 (C_{ar}), 128.7 (C_{ar}), 128.6 (C_{ar}), 128.5 (C_{ar}), 128.4 (C_{ar}), 128.3 (C_{ar}), 124.6 (C_{ar}), 121.7 (C_{ar}), 116.2 (C_{ar}), 75.3 (C-***13/19***), 70.4 (C-***13/19***), 38.7 (C-***10***), 26.4 (C-***11***)

IR (KBr cm^{-1}):

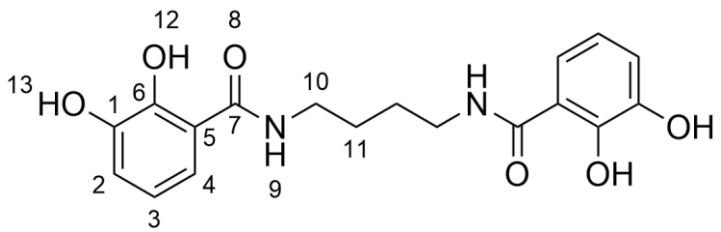
3394 *medium* (N-H), 2933 *weak* (C-H), 1659 *strong* (C=O, amide), 1529 *strong* (C-C ar)

Elemental Analysis:

Calculated for $[C_{46}H_{44}N_2O_6]$: %C 76.64, %H 6.15, %N 3.89

Measured for $[C_{46}H_{44}N_2O_6]$: %C 76.29, %H 6.19, %N 3.83

H_4-4-LICAM



Compound **14**, 0.0564 g (0.078 mmol) was dissolved in 2 mL of dry toluene before 50 mL of dry ethanol was added. To this stirred solution one small spatula tip of Pd-C 10% was added. The mixture was purged with hydrogen and allowed to react for 18 hours under a positive pressure of hydrogen. After the reaction was completed the catalyst was filtered off and the product was isolated by removing the solvent *in vacuo,* to yield a beige solid. Characterisation data consistent with literature.[5]

Yield:

0.0267 g, 0.074 mmol, 95%

Mp:

Decomposed approx 195 °C

***m/z* (ESI):**

361.1391 ($[M+H]^+$, 10%), 383.1216 ($[M+Na]^+$, 24%)

HRMS (ESI):

Calc for $[C_{18}H_{20}N_2O_6]+H^+$ = 361.1394 Found 361.1391 (0.9 ppm error)

Calc. for $[C_{18}H_{20}N_2O_6]+Na^+$ = 383.1214 Found 383.1216 (-0.7 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

12.88 (2H, br s, H-***23/24***), 9.14 (2H, br s, H-***23/24***), 8.81 (2H, t, *J* = 5.5 Hz, H-***9***), 7.29 (2H, dd, *J* = 8.5, 1.2 Hz, H-***2/4***), 6.90 (2H, dd, *J* = 8.0, 1.2 Hz, H-***2/4***), 6.68 (2H, dd, *J* = 8.5, 8.0 Hz, H-***3***), 3.30-3.25 (4H, m, H-***10***), 1.59-1.54 (4H, m, H-***11***).

^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

170.4 (C-***7***), 150.3 (C-***1/6***), 146.8 (C-***1/6***), 119.2 (C_{ar}), 118.3 (C_{ar}), 117.5 (C_{ar}), 115.3 (C_{ar}), 38.6 (C-***10***), 26.3 (C-***11***).

IR (KBr cm^{-1}):

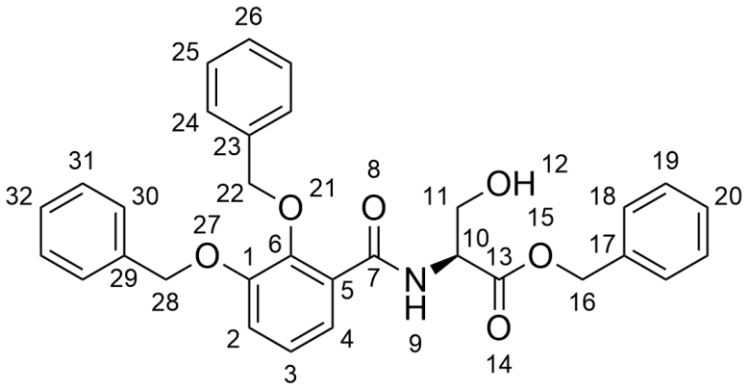
3415 *medium* (O-H), 3390 *medium* (N-H), 1643 *medium* (C=O, amide), 1590 *medium* (C-C ar)

Elemental Analysis

Calculated for [$C_{18}H_{20}N_2O_6$.0.8H_2O]: %C 57.69, %H 5.81, %N 7.47

Measured for [$C_{18}H_{20}N_2O_6$.0.8H_2O]: %C 57.78, %H 5.94, %N 7.46

L-Serine-*N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-(phenylmethyl)-ether, 22



To a stirred suspension of L-serine benzyl ester hydrochloride (0.70 g, 3 mmol) in acetonitrile (75 mL), triethylamine (0.50 mL, 3.6 mmol) was added and allowed to stir for a short time before, a solution of compound **13** (1.29 g, 3 mmol) in acetonitrile (25 mL) was added and the mixture and left to stir overnight. After 18 hours, the solvent was removed *in vacuo* and the resulting oil dissolved in chloroform 100 mL. This was washed with HCl (50 mL, 0.1 M), saturated $NaHCO_3$ (50 mL) and water (4 x 50 mL) until the aqueous layer was at neutral pH. The organic layer was then dried over $MgSO_4$, filtered and the solvent removed *in vacuo*, yielding an off white solid. Characterisation data consistent with literature.[6]

Yield:

1.40 g, 2.74 mmol, 91.0%

Melting point:

76.4 – 79.6 °C

R_f:

Chloroform : Ethyl Acetate (1:1) : R_f : 0.53

***m/z* (ESI):**

534.1874 ($[M+Na]^+$, 100 %), 512.2056 ($[M+H]^+$, 91.5 %)

HRMS (ESI):

Calc for $[C_{31}H_{29}NO_6] + H^+$ = 512.2068 Found 512.2056 (2.3 ppm error)

Calc for $(C_{31}H_{29}NO_6] + Na^+$ = 534.1887 Found 534.1874 (2.4 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

8.86 (1H, d, *J* = 7.0 Hz, H-***9***), 7.71 (1H, dd, *J* = 7.0, 2.5 Hz, H-***2/4***), 7.47-7.15 (17H, m, H_{ar}), 5.21 (1H, d, *J* = 10.0 Hz, H-***16/22/28***), 5.19 (2H, s, H-***16/22/28***), 5.15 (2H, s, H-***16/22/28***), 5.07 (1H, d, *J* = 10.0 Hz, H-***16/22/28***), 4.78 (1H, ddd, *J* = 7.5, 4.0, 4.0 Hz, H-***10***), 3.90 (1H, dd, *J* = 11.2, 4.4 Hz, H-***11/11****), 3.85 (1H, dd, *J* = 11.2, 4.0 Hz, H-***11/11****)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

170.2 (C-***13***), 165.8 (C-***7***), 151.9 (C_{ar}), 147.2 (C_{ar}), 136.4 (C_{ar}), 136.4 (C_{ar}), 135.4 (C_{ar}), 129.2 (C_{ar}), 128.8 (C_{ar}), 128.7 (C_{ar}), 128.7 (C_{ar}), 128.6 (C_{ar}), 128.5 (C_{ar}), 128.4 (C_{ar}), 128.3 (C_{ar}), 127.9 (C_{ar}), 126.7 (C_{ar}), 124.5 (C_{ar}), 123.4 (C_{ar}), 117.7 (C_{ar}), 76.4 (C-***16/22/28***), 71.5 (C-***16/22/28***), 67.5 (C-***16/22/28***), 63.7 (C-***11***), 55.6 (C-***10***)

IR (KBr cm^{-1}):

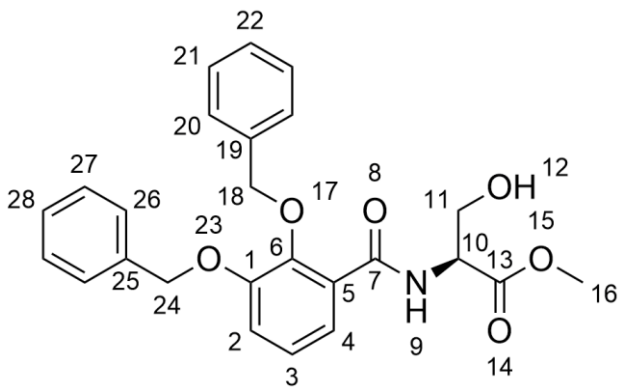
3358 *strong* (N-H), 3036 *medium* (C-H ar), 2956 *medium* (C-H), 1743 *strong* (C=O, ester), 1627 *strong* (C=O, amide)

Elemental Analysis:

Calculated for $[C_{31}H_{29}NO_6.0.2H_2O]$: %C 72.27, %H 5.75, %N 2.72

Measured for $[C_{31}H_{29}NO_6.0.2H_2O]$: %C 72.18, %H 5.66, %N 2.78

L-Serine-*N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-methyl ester, 24



To a stirred solution of L-serine methyl ester hydrochloride (1.74 g, 11.2 mmol) in acetonitrile (75 mL), triethylamine (3.83 mL, 27.5 mmol) was added and allowed to stir for a short time before, a solution of compound **13** (4.31 g, 9.89 mmol) in acetonitrile (25 mL) was added and the mixture and left to stir overnight. After 18 hours, the solvent was removed in vacuo and the resulting oil dissolved in chloroform (100 mL). This was washed with HCl (50 mL, 0.1 M), saturated $NaHCO_3$ (50 mL) and water (4 x 50 mL) until the aqueous layer was at neutral pH. The organic layer was then dried over $MgSO_4$, filtered and the solvent removed in vacuo, yielding an off white solid.

Yield:

3.40 g, 7.8 mmol, 78%

Melting point:

110.3 – 111.2 °C

R_f:

Chloroform : Ethyl Acetate (1:1) : R_f : 0.37

***m/z* (ESI):**

458.1581 ($[M+Na]^+$, 100 %), 436.1766 ($[M+H]^+$, 11.3 %)

HRMS (ESI):

Calc for $[C_{25}H_{25}NO_6] + H^+$ = 436.1755 Found 436.1766 (-2.6 ppm error)

Calc for $[C_{25}H_{25}NO_6] + Na^+$ = 458.1574 Found 458.1581 (-1.6 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

8.86 (1H, d, *J* = 7.3 Hz, H-***9***), 7.72 (1H, dd, *J* = 7.0, 2.5 Hz, H-***2/4***), 7.26-7.48 (10H, m, H_{ar}), 7.14-7.20 (2H, m, H_{ar}), 5.26 (1H, d, *J* = 10.5 Hz, H-***18/18*/24/24****), 5.24 (2H, s, H-***18/24***), 5.19 (1H, d, *J* = 10.5 Hz, H-***18/18/24/24****), 4.85 (1H, ddd, *J* = 7.5, 4.0, 4.0 Hz, H-***10***), 3.98 (1H, dd, *J* = 11.5, 4.5 Hz, H-***11***), 3.92 (1H, dd, *J* = 11.5, 4.0 Hz, H-***11***), 3.80 (3H, s, H-***16***)

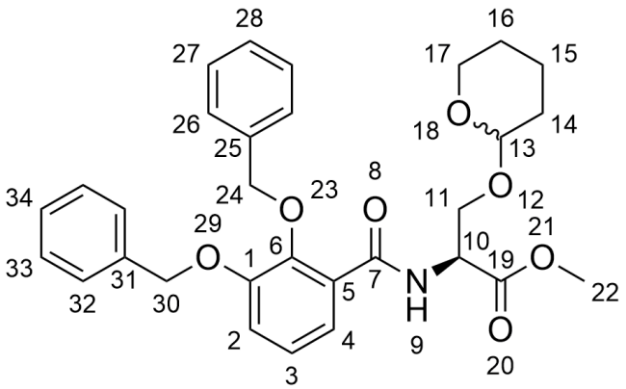
^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

171.1 (C-***13***), 166.0 (C-***7***), 152.1 (C_{ar}), 147.4 (C_{ar}), 136.6 (C_{ar}), 136.6 (C_{ar}), 129.3 (C_{ar}), 128.9 (C_{ar}), 128.8 (C_{ar}), 128.7 (C_{ar}), 128.5 (C_{ar}), 128.0 (C_{ar}), 126.8 (C_{ar}), 124.6 (C_{ar}), 123.5 (C_{ar}), 117.7 (C_{ar}), 76.2 (C-***18/24***), 71.3 (C-***18/24***), 63.4 (C-***11***), 55.1 (C-***10***), 52.4 (C-***16***)

IR (KBr cm^{-1}):

3335 *strong* (N-H), 3064 *medium* (C-H ar), 2935 *medium* (C-H), 1747 *strong* (C=O, ester), 1643 *strong* (C=O, amide)

L-Serine-*N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-(tetrahydro-pyran-2-yloxy)-*O*-methyl ester, 25



Pyridinium *p*-toluenesulfonate (0.75 g, 3 mmol) was added to a stirred solution of **24** (3.34 g, 7.66 mmol) in DCM (100 mL) at room temperature. To this, 3,4-Dihydro-2*H*-pyran (1.75 mL, 19 mmol) was added and the solution was left to stir. After 18 hours, the DCM was removed *in vacuo* to yield an oil which was subsequently dissolved in diethyl ether (150 mL) and washed with water (3 x 50 mL) and saturated $NaHCO_3$ solution (50 mL). After removal of the solvent *in vacuo*, the product is isolated as a viscous oil of diastereomers in an equal ratio.

Yield:

3.44 g, 6.62 mmol, 86%

R_f:

Chloroform : Ethyl Acetate (1:1) : R_f : 0.52

***m/z* (ESI):**

542.2139 ($[M+Na]^+$, 100 %), 520.2344 ($[M+H]^+$, 3.0 %)

HRMS (ESI):

Calc for $[C_{30}H_{33}NO_7] + H^+$ = 520.2330 Found 520.2344 (-2.7 ppm error)

Calc for $[C_{30}H_{33}NO_7] + Na^+$ = 542.2149 Found 542.2139 (1.9 ppmq error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

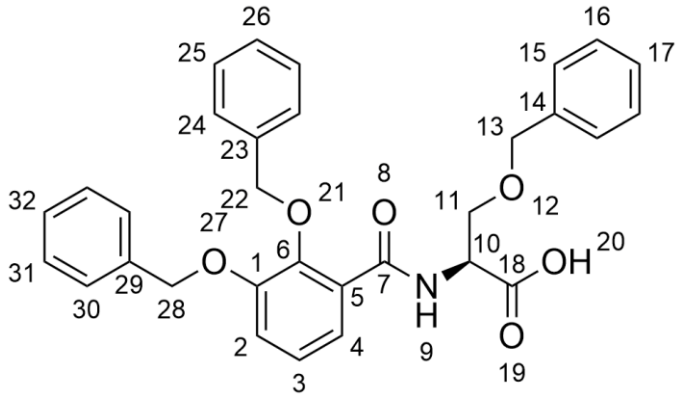
8.80 (2H, d, *J* = 7.5 Hz, H-***9, 9'***), 7.52-7.18 (26H, m, H_{ar}, H_{ar}'), 5.21 (4H, s, H-***24/30, 24'/30'***), 5.12-4.98 (4H, m, H-***24/30/24*/30*, 24'/30'/24*'/30*'***), 4.85-4.79 (1H, m, H-***10/10'***), 4.78-4.73 (1H, m, H-***10/10'***), 4.50-4.47 (1H, m, H-***13/13'***), 4.43-4.41 (1H, m, H-***13/13'***), 4.00 (1H, dd, *J* = 10.5, 4.5 Hz, H-***11/11*/11'/11*'***), 3.89-3.85 (1H, m, H-***11/11*/11'/11*'***), 3.74 (1H, dd, *J* = 10.5, 4.5 Hz, H-***11/11*/11'/11*'***), 3.67 (3H, s, H-***22/22'***), 3.66 (3H, s, H-***22/22'***), 3.61-3.53 (3H, m, H-***11/11*/11'/11*', 17/17'***), 3.46-3.30 (2H, m, H-***17/17'***)ǂ, 1.49-1.28 (12H, m, H-***14/14', 15/15', 16/16'***)

ǂObscured by water signal

^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

170.4 (C-***19/19'***), 170.2 (C-***19/19'***), 164.6 (C-***7/7'***), 163.3 (C-***7/7'***), 151.5 (C_{ar}), 145.7 (C_{ar}), 145.6 (C_{ar}), 136.4 (C_{ar}), 136.4 (C_{ar}), 136.4 (C_{ar}), 136.3 (C_{ar}), 128.4 (C_{ar}), 128.3 (C_{ar}), 128.3 (C_{ar}), 127.9 (C_{ar}), 127.9 (C_{ar}), 127.9 (C_{ar}), 127.9 (C_{ar}), 127.8 (C_{ar}), 127.3 (C_{ar}), 124.3 (C_{ar}), 124.2 (C_{ar}), 121.5 (C_{ar}), 121.3 (C_{ar}), 116.8 (C_{ar}), 116.7 (C_{ar}), 98.2 (C-***13/13'***), 97.2 (C-***13/13'***), 93.0 (C), 75.1 (C-***24/30/24/30'***), 75.0 (C-***24/30/24'/30'***), 70.2 (C-***24/30/24'/30'***), 70.2 (C-***24/30/24'/30'***), 66.3 (C-***11/11'***), 65.9 (C-***11/11'***), 64.9 (C), 64.7 (C), 61.3 (C), 60.6 (C), 52.7 (C-***10/10'***), 52.3 (C-***10/10'***), 51.9 (C-***22/22'***), 51.9 (C-***22/22'***), 29.9 (C), 29.7 (C), 29.3 (C), 24.8 (C), 24.6 (C), 24.5 (C), 18.9 (C), 18.6 (C), 18.4 (C)

L-Serine-*N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-(phenylmethyl)-ether, 29



Triethylamine (0.50 mL, 3.6 mmol) was added to a stirred solution of *O*-benzyl L-serine (0.70 g, 3 mmol) in acetonitrile (75 mL). To this, a solution of compound **13** (1.29 g, 3 mmol) in acetonitrile (25 mL) was added and the mixture left to stir overnight. After 18 hours, the solvent was removed *in vacuo* and the resulting oil dissolved in chloroform 100 mL. This was washed with HCl (50 mL, 0.1 M) and water (4 x 50 mL) until the aqueous layer was at neutral pH. The organic layer was then dried over $MgSO_4$, filtered and the solvent removed *in vacuo*, yielding an off white solid. Characterisation data consistent with literature.[1]

Yield:

1.28 g, 2.50 mmol, 84%

Melting point:

116.2 - 117.9 °C

***m/z* (ESI):**

512.2082 ($[M+H]^+$, 13 %), 534.1880 ($[M+Na]^+$, 100 %)

HRMS (ESI):

Calc for $[C_{31}H_{29}NO_6] + H^+$ = 512.2068 Found 512.2082 (-2.8 ppm error)

Calc for $[C_{31}H_{29}NO_6] + Na^+$ = 534.1887 Found 534.1880 (1.6 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

9.01 (1H, d, *J* = 7.0 Hz, H-***9***), 7.74 (1H, dd, *J* = 7.0 Hz, 2.5 Hz, H-***2/4***), 7.48-7.22 (17H, m, H_{ar}), 5.16 (2H, s, H-***22/28***), 5.11 (1H, d, *J* = 10.5 Hz, H-***22/28/22*/28****), 5.09 (1H, d, *J* = 10.5 Hz, H-***22/28/22*/28****), 4.87 (1H, ddd, *J* = 7.5, 4.0, 4.0 Hz, H-***10***), 4.37 (2H, s, H-***13***), 3.86 (1H, dd, *J* = 9.5, 4.5 Hz, H-***11/11****), 3.66 (1H, dd, *J* = 9.5, 4.0 Hz, H-***11/11****)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

173.6 (C-***18***), 166.2 (C-***7***), 152.0 (C_{ar}), 147.3 (C_{ar}), 137.4 (C_{ar}), 136.4 (C_{ar}), 136.2 (C_{ar}), 129.3 (C_{ar}), 128.9 (C_{ar}), 128.6 (C_{ar}), 128.6 (C_{ar}), 128.5 (C_{ar}), 128.5 (C_{ar}), 128.0 (C_{ar}), 127.9 (C_{ar}), 127.9 (C_{ar}), 126.4 (C_{ar}), 124.6 (C_{ar}), 123.4 (C_{ar}), 117.8 (C_{ar}), 76.4 (C-***22/28***), 73.4 (C-***13***), 71.5 (C-***22/28***), 68.8 (C-***11***), 53.4 (C-***10***)

IR (KBr cm^{-1}):

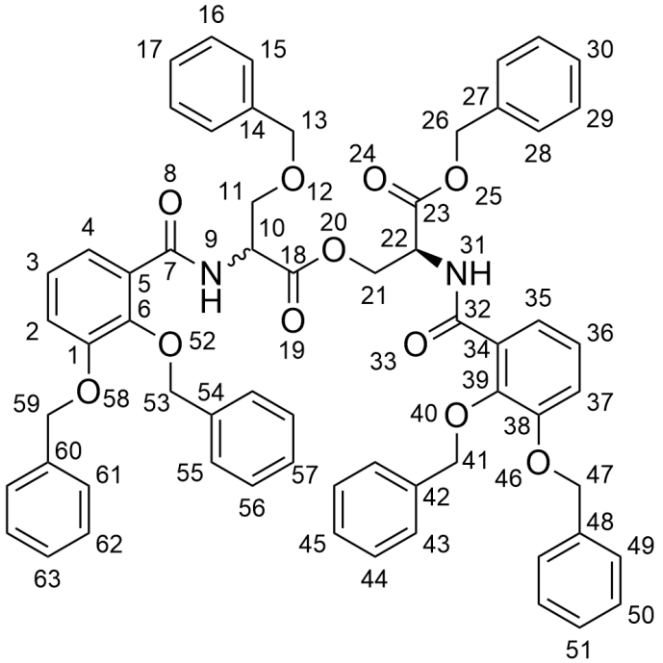
3353 *strong* (N-H), 3060 *medium broad* (O-H acid), 2919 *medium* (C-H ar) 1757 *strong* (C=O, acid), 1636 *strong* (C=O, amide)

Elemental Analysis:

Calculated for [$C_{31}H_{29}NO_6$]: %C 72.78, %H 5.71, %N 2.74

Measured for [$C_{31}H_{29}NO_6$]: %C 72.64, %H 5.69, %N 2.80

L-Serine-*N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-(phenylmethyl)-ester-*O*-L/D-serine, *N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-(phenylmethyl)-ether, 30 (L/D, L/L)



Compound **29** (0.231 g, 0.45 mmol), compound **22** (0.262 g, 0.51 mmol) and HOBt (0.122 g, 0.9 mmol) were suspended in 30 mL dry chloroform. DCC (0.138 g, 0.67 mmol) was added slowly over 10 minutes, and the reaction was stirred overnight at room temperature. The solution was then washed with 20 mL water, 15 mL saturated $NaHCO_3$ solution, 20 mL 0.1 M HCl, followed by 3 x 20 mL water. The organic layer was dried over $MgSO_4$ before the solvent was removed *in vacuo*. The crude product was purified *via* column chromatography, 9:1 dichloromethane : ethyl acetate, followed by a second column eluting with 9:2 dichloromethane : ethyl acetate to yield a colourless oil.

Yield:

0.072 g, 0.07 mmol, 16 %

R_f:

Dichloromethane : Ethyl Acetate (9:2) : R_f : 0.7

***m/z* (ESI):**

1005.3917 ($[M+H]^+$, 100 %), 1027.3743 ($[M+Na]^+$, 6 %)

HRMS (ESI):

Calc for $[C_{62}H_{56}N_2O_{11}] + H^+$ = 1005.3957 Found 1005.3917 (4.0 ppm error)

Calc for $[C_{62}H_{56}N_2O_{11}] + Na^+$ = 1027.3776 Found 1027.3743 (3.2 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

The ^{1}H NMR is comprised of two diastereomers in solution. The ratio between the two diastereomers is 5/4.

8.83-8.76 (2H, m, H-***9/31***), 7.73-7.66 (2H, m, H_{ar}), 7.45-6.92 (34H, m, H_{ar}), 5.22-4.96 (10H, m), 4.65, (1H, ddd, *J* = 7.6, 3.6, 3.2 Hz, H-***10/22***)), 4.57-4.53 (1H, m), 4.34 (1H, dd, *J* = 11.2, 3.6 Hz, H-***11/11*/21/21****), 4.31 (1H, dd, *J* = 11.2, 4.0 Hz, H-***11/11*/21/21****), 4.20 (1H, d, *J* = 12.0 Hz, H_{benzyl}), 4.16 (1H, d, *J* = 12.0 Hz, H_{benzyl}), 3.53 (1H, dd, *J* = 9.6, 3.6 Hz, H-***11/11*/21/21****), 3.44 (1H, dd, *J* = 9.2, 3.2 Hz, H-***11/11*/21/21****)

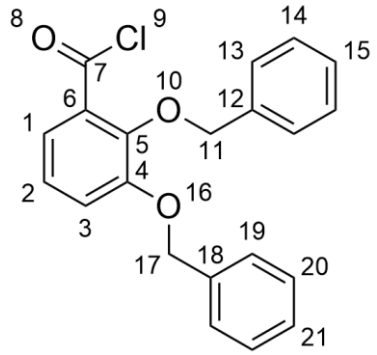
8.83-8.76 (2H, m, H-***9/31***), 7.73-7.66 (2H, m, H_{ar}), 7.45-6.92 (34H, m, H_{ar}), 5.22-4.96 (12H, m), 4.60-4.52 (2H, m), 4.07 (2H, s, H_{benzyl}), 3.27 (1H, dd, *J* = 9.6, 3.6 Hz, H-***11/11*/21/21****), 3.18 (1H, dd, *J* = 9.6, 3.6 Hz, H-***11/11*/21/21****)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

Certain signals in the ^{13}C NMR spectrum overlap

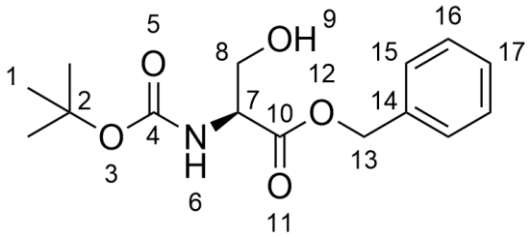
170.1 (C-***18/18*/23/23****), 170.1 (C-***18/18*/23/23****), 169.6 (C-***18/18*/23/23****), 169.5 (C-***18/18*/23/23****), 165.6 (C-***7/7*/32/32****), 165.6 (C-***7/7*/32/32****), 165.4 (C-***7/7*/32/32****), 165.4 (C-***7/7*/32/32****), 152.3 (C_{ar}), 152.2 (C_{ar}), 152.1 (C_{ar}), 147.4 (C_{ar}), 147.4 (C_{ar}), 147.4 (C_{ar}), 147.3 (C_{ar}), 137.7 (C_{ar}), 137.7 (C_{ar}), 136.6 (C_{ar}), 136.6 (C_{ar}), 136.6 (C_{ar}), 136.6 (C_{ar}), 136.5 (C_{ar}), 136.5 (C_{ar}), 136.5 (C_{ar}), 135.5 (C_{ar}), 135.4 (C_{ar}), 129.6 (C_{ar}), 129.4 (C_{ar}), 129.4 (C_{ar}), 128.9 (C_{ar}), 128.9 (C_{ar}), 128.9 (C_{ar}), 128.8 (C_{ar}), 128.8 (C_{ar}), 128.7 (C_{ar}), 128.6 (C_{ar}), 128.6 (C_{ar}), 128.5 (C_{ar}), 128.5 (C_{ar}), 128.5 (C_{ar}), 128.4 (C_{ar}), 128.4 (C_{ar}), 128.4 (C_{ar}), 128.4 (C_{ar}), 128.3 (C_{ar}), 128.2 (C_{ar}), 128.1 (C_{ar}), 128.1 (C_{ar}), 127.9 (C_{ar}), 127.7 (C_{ar}), 127.6 (C_{ar}), 127.6 (C_{ar}), 127.1 (C_{ar}), 127.1 (C_{ar}), 126.9 (C_{ar}), 124.6 (C_{ar}), 124.6 (C_{ar}), 124.5 (C_{ar}), 123.5 (C_{ar}), 123.4 (C_{ar}), 123.4 (C_{ar}), 117.6 (C_{ar}), 117.4 (C_{ar}), 117.3 (C_{ar}), 76.1 (C_{benzyl}), 76.0 (C_{benzyl}), 76.0 (C_{benzyl}), 72.9 (C_{benzyl}), 72.6 (C_{benzyl}), 71.2 (C_{benzyl}), 71.2 (C_{benzyl}), 69.1 (C_{benzyl}), 68.6 (C_{benzyl}), 67.6 (C_{benzyl}), 67.4 (C-***11/11*/21/21****), 64.9 (C-***11/11*/21/21****), 64.5 (C-***11/11*/21/21****), 60.3 (C-***11/11*/21/21****), 52.9 (C-***10/10*/22/22****), 52.8 (C-***10/10*/22/22****), 51.9 (C-***10/10*/22/22****), 51.5 (C-***10/10/22/22****)

2,3-bis(benzyloxy)benzoyl chloride, 31



Prepared as in accordance to the literature and used crude.[7]

L-Serine-*N*-[(1,1-dimethylethoxy)carbonyl]-*O*-(phenylmethyl)-ester, 32



To a stirred suspension of L-serine benzyl ester hydrochloride (1.15 g, 5 mmol) in acetonitrile (20 mL), triethylamine (0.60 mL, 6.0 mmol) was added and the resulting solution was allowed to stir for 10 minutes. Di-*tert*-butyl dicarbonate (1.15 g, 5 mmol) was then added and the mixture and left to stir for three hours at reflux. Once the reaction was completed, the solvent was removed *in vacuo* and the resulting oil dissolved in dichloromethane, 50 mL. The solution was washed with formic acid (15 mL, 0.1 M) and water until the aqueous layer was at neutral pH (3 x 25 mL minimum). The organic layer was then dried over $MgSO_4$, filtered and the solvent removed *in vacuo*, yielding a white solid. Characterisation data consistent with literature.[8]

Yield:

1.382 g, 4.68 mmol, 94%

Melting point:

66.8-68.2 °C

R_f:

Dichloromethane : Methanol (9:1) : R_f : 0.63

***m/z* (ESI):**

318.1305 ($[M+Na]^+$, 80 %), 296.1482 ($[M+H]^+$, 10 %)

HRMS (ESI):

Calc for $[C_{15}H_{21}NO_5] + H^+$ = 296.1492 Found 296.1482 (3.4 ppm error)

Calc for $[C_{15}H_{21}NO_5] + Na^+$ = 318.1312 Found 318.1305 (2.3 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

7.39-7.34 (5H, m, H_{ar}), 5.51 (1H, br, d, *J* = 5.0 Hz, H-***6***), 5.22 (1H, d, *J* = 12.5 Hz, H-***13/13****), 5.20 (1H, d, *J* = 12.5 Hz, H-***13/13****), 4.42 (1H, br s, H-***7***), 3,98 (1H, dd, *J* = 11.0, 4.0 Hz, H-***8***), 3,90 (1H, dd, *J* = 11.0, 4.0 Hz, H-***8***) 1.43 (9H, s, H-***1***)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

171.2 (C-***10***), 156.1 (C-***4***), 135.2 (C_{ar}), 128.8 (C_{ar}), 128.7 (C_{ar}), 128.4 (C_{ar}), 80.2 (C-***2***), 67.5 (C-***13***), 63.3 (C-***8***), 55.7 (C-***7***), 27.9 (C-***1***)

IR (KBr cm^{-1}):

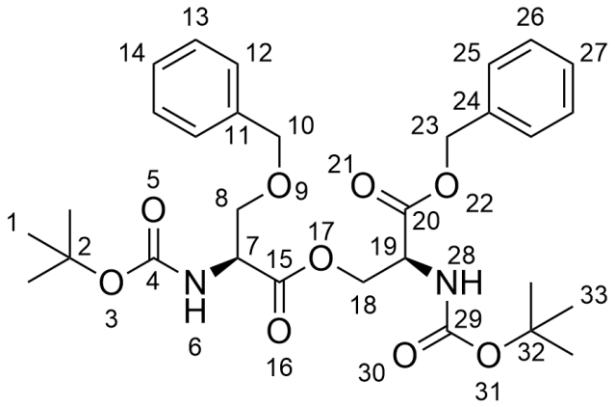
3419 *medium* (O-H), 3365 *medium* (N-H), 2980 *weak* (C-H), 1760 *strong* (C=O), 1669 *strong* (C=O), 1526 *medium* (C-C ar)

Elemental Analysis:

Calculated for $[C_{15}H_{21}NO_5]$: %C 61.00, %H 7.17, %N 4.74

Measured for $[C_{15}H_{21}NO_5]$: %C 61.14, %H 7.12, %N 4.63

L-Serine-*N*-[(1,1-dimethylethoxy)carbonyl]-*O*-(phenylmethyl)-ester-*O*-L-serine, *N*-[(1,1-dimethylethoxy)carbonyl]-*O*-(phenylmethyl)-ether, 33



Boc-*O*-benzyl-L-serine (1.89 g, 6.4 mmol), compound **32** (1.72 g, 5.8 mmol) and HOBt (1.23 g, 9.1 mmol) were suspended in 10 mL dry acetonitrile and the mixture was stirred on ice. Once the mixture was cooled, EDC (1.41 g, 7.4 mmol) was added to the suspension, followed by the addition of 15 mL acetonitrile and allowed to stir for two hours on ice, followed by 48 hours at room temperature. The volatiles were removed *in vacuo*, and the remaining oil dissolved in 100 mL dichloromethane. The organic solution and washed with 50 mL brine and 10 mL sat $NaHCO_3$ solution, followed by 2 x 50 mL brine. The organic layer was dried over $MgSO_4$ before the solvent was removed *in vacuo*. The crude product was purified *via* column chromatography, 9:1 chloroform : acetonitrile to yield a pure white solid. Characterisation data consistent with literature.[6]

Yield:

2.032 g, 3.55 mmol, 60%

Melting point:

80.1-82.2 °C

R_f:

Chloroform : Acetonitrile (4:1) : R_f : 0.47

***m/z* (ESI):**

595.2602 ([M+Na]$^+$, 100 %), 573.2782 ([M+H]$^+$, 62 %)

HRMS (ESI):

Calc for [$C_{15}H_{21}NO_5$] + H^+ = 573.2807 Found 573.2782 (4.4 ppm error)

Calc for [$C_{15}H_{21}NO_5$] + Na^+ = 595.2626 Found 595.2602 (4.0 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

7.37-7.27 (10H, m, H_{ar}), 5.50 (1H, br, d, *J* = 8.8 Hz, H-***6/28***), 5.29 (1H, br, d, *J* = 8.8 Hz, H-***6/28***), 5.21 (1H, d, *J* = 12.4 Hz, H-***10/10****), 5.17 (1H, d, *J* = 12.0 Hz, H-***10/10****), 4.68-4.56 (3H, m, H-***7/19,8/8*/18/18*,23/23****), 4.49-4.43 (2H, m, H-***8/8*/18/18*,23/23****), 4.35-4.31 (1H, m, H-***7/19***), 3.72 (1H, dd, *J* = 9.6, 3.6 Hz, H-***8/8*/18/18****), 3.58 (1H, dd, *J* = 9.6, 3.2 Hz, H-***8/8*/18/18****) 1.43 (9H, s, H-***1/33***), 1.40 (9H, s, H-***1/33***)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

170.6 (C-***15/20***), 170.0 (C-***15/20***), 155.8 (C-***4/29***), 155.7 (C-***4/29***), 137.6 (C_{ar}), 135.4 (C_{ar}), 128.9 (C_{ar}), 128.8 (C_{ar}), 128.7 (C_{ar}), 128.6 (C_{ar}), 128.2 (C_{ar}), 128.0 (C_{ar}), 127.8 (C_{ar}), 80.2 (C-***2/32***), 80.1 (C-***2/32***), 73.2 (C-***8/18/23***), 69.6 (C-***8/18/23***), 67.5 (C-***10***), 65.0 (C-***8/18/23***), 53.8 (C-***7/19***), 53.0 (C-***7/19***), 28.0 (C-***1/33***), 27.9 (C-***1/33***)

IR (KBr cm^{-1}):

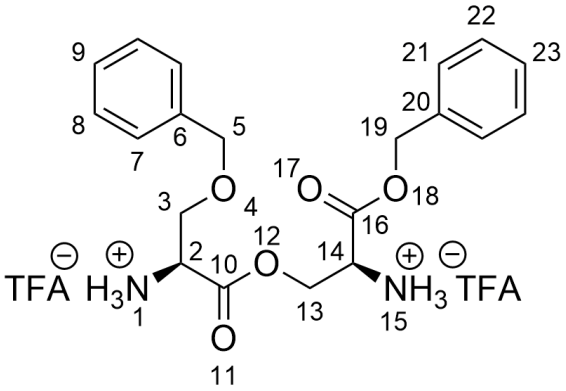
3459 *medium* (N-H), 2978 *weak* (C-H), 1744 *strong* (C=O), 1713 *strong* (C=O), 1505 *medium* (C-C ar)

Elemental Analysis:

Calculated for [$C_{30}H_{40}N_2O_9$]: %C 62.92, %H 7.04, %N 4.89

Measured for [$C_{30}H_{40}N_2O_9$]: %C 63.92, %H 7.04, %N 4.86

L-Serine-*O*-(phenylmethyl)-ester-*O*-L-serine-*O*-(phenylmethyl)-ether trifluoroacetic acid salt, 34



Compound **33** (0.1078 g, 0.18 mmol) was dissolved in 4 mL of dichloromethane and the flask was purged with nitrogen. Trifluoroacetic acid (1 mL) was then added to the stirred solution and the reaction was allowed to proceed at room temperature, monitored by TLC (~2 hours). The volatiles were removed *in vacuo* to yield a pale yellow oil, which was taken up in 5 mL toluene, which was subsequently removed *in vacuo*. This process was repeated two more times to remove traces of TFA and yield a white solid material.

Yield:

0.0908 g, 0.16 mmol, 89%

Melting point:

64.3.1-67.7 °C

***m/z* (ESI):**

373.1753 ($[M+H]^+$, 63 %)

HRMS (ESI):

Calc for $[C_{20}H_{24}N_2O_5] + H^+$ = 373.1758 Found 373.1753 (1.2 ppm error)

^{1}H NMR: (MeOD, 400 MHz) δ_H

7.46-7.27 (10H, m, H_{ar}), 5.35 (1H, d, *J* = 11.5 Hz, H-***5/5****), 5.27 (1H, d, *J* = 11.5 Hz, H-***5/5****), 4.63 (1H, d, *J* = 12.5 Hz, H-***19/19****), 4.54 (1H, d, *J* = 12.5 Hz, H-***19/19****), 4.28 (1H, dd, *J* = 4.0, 3.2 Hz, H-***2/14***), 4.17 (1H, dd, *J* = 4.0, 3.2 Hz, H-***2/14***), 4.02 (1H, dd, *J* = 12.0, 4.4 Hz, H-***3/3*/13/13****), 3.95 (1H, dd, *J* = 12.0, 4.4 Hz, H-***3/3*/13/13****), 3.95 (1H, dd, *J* = 12.0, 4.4 Hz, H-***3/3*/13/13****), 3.83 (1H, dd, *J* = 12.0, 4.4 Hz, H-***3/3*/13/13****)

$^{13}C\{^1H\}$ NMR: (MeOD, 100 MHz) δ_C

168.8 (C-***10/16***), 168.2 (C-***10/16***), 138.9 (C_{ar}), 136.8 (C_{ar}), 130.5 (C_{ar}), 130.5 (C_{ar}), 130.3 (C_{ar}), 130.2 (C_{ar}), 130.1 (C_{ar}), 130.1 (C_{ar}), 129.8 (C_{ar}), 129.8 (C_{ar}), 129.7 (C_{ar}), 74.8 (C-***19***), 69.3 (C-***5***), 68.1 (C-***3/13***), 60.9 (C-***3/13***), 56.3 (C-***2/14***), 54.6 (C-***2/14***)

IR (KBr cm^{-1}):

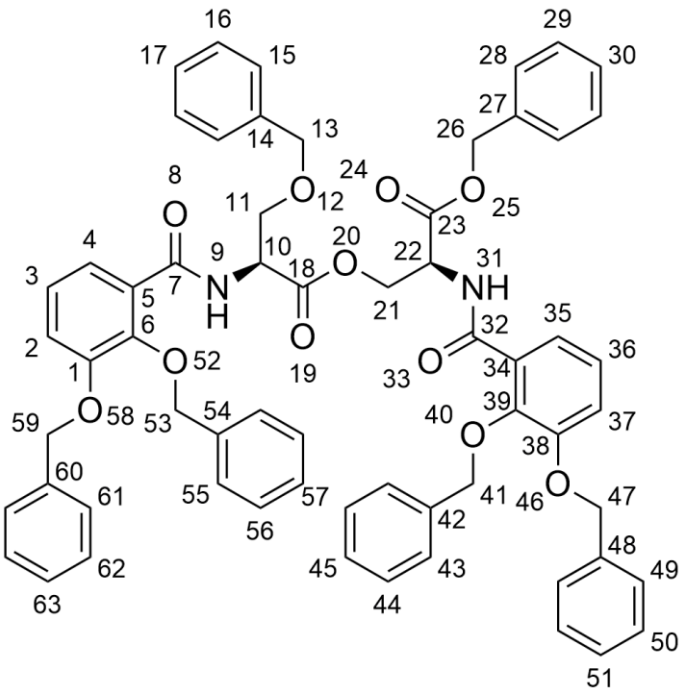
3005 *weak* (C-H), 1764 *medium* (C=O), 1745 *medium* (C=O), 1668 *strong* (C=O), 1538 *medium* (C-C ar)

Elemental Analysis:

Calculated for [$C_{24}H_{24}F_6N_2O_7$.0.5TFA]: %C 47.86, %H 4.29, %N 4.55

Measured for [$C_{24}H_{24}F_6N_2O_7$.0.5TFA]: %C 48.16, %H 3.96, %N 4.49

L-Serine-*N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-(phenylmethyl)-ester-*O*-L-serine, *N*-[2,3-bis(phenylmethoxy)benzoyl]-*O*-(phenylmethyl)-ether, 30 (L/L)



Compound **34** (0.5779 g, 1.0 mmol) was dissolved in 5 mL anhydrous dichloromethane. Alternating amounts of DIPEA (15 mmol. 2.6 mL) and crude compound **31**, (3.0 mmol) dissolved in 10 mL dichloromethane was added. The resulting mixture was stirred for 10 minutes. The mixture was subsequently diluted with 30 mL dichloromethane, followed by the addition of 20 mL 0.1 M

formic acid (aqueous). The organic layer was separated, washed with 2 x 20 mL water, before drying over $MgSO_4$.. The organic layer was then filtered and dried *in vacuo* yielding a crude product. The product was purified *via* column chromatography using 9:1 dichloromethane: acetonitrile as the solvent system to yield an off white solid material

Yield:

0.4575 g, 0.455 mmol, 44%

Melting point:

45.2-46.9 °C

R_f:

Dichloromethane : Acetonitrile (9:1) : R_f : 0.48

***m/z* (ESI):**

1005.3922 ($[M+H]^+$, 100 %)

HRMS (ESI):

Calc for $[C_{62}H_{56}N_2O_{11}] + H^+$ = 1005.3957 Found 1005.3922 (3.4 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

8.82 (1H, d, *J* = 8.4 Hz, H-***9/31***), 8.81 (1H, d, *J* = 7.6 Hz, H-***9/31***), 7.73-7.71 (2H, m, H_{ar}), 7.45-6.91 (34H, m, H_{ar}), 5.24 (1H, d, *J* = 12.0 Hz, H_{benzyl}), 5.22 (1H, d, *J* = 12.0 Hz, H_{benzyl}), 5.15-5.08 (7H, m, H_{benzyl}, ***10/22***), 5.02 (2H, dd, *J* = 10.4, 3.6 Hz, H_{benzyl}), 4.56-4.52 (2H, m, H-***11/11*/21/21****, ***10/22***), 4.35 (1H, dd, *J* = 11.6, 3.6 Hz, H-***11/11*/21/21****), 4.08 (2H, s, H-***26***), 3.28 (1H, dd, *J* = 9.6, 3.2 Hz, H-***11/11*/21/21****), 3.18 (1H, dd, *J* = 9.6, 3.2 Hz, H- ***11/11*/21/21****)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

170.1 (C-***18/23***), 169.6 (C-***18/23***), 165.5 (C-***7/32***), 165.4 (C-***7/32***), 152.2 (C_{ar}), 152.2 (C_{ar}), 147.4 (C_{ar}), 147.3 (C_{ar}), 137.7 (C_{ar}), 136.6 (C_{ar}), 136.5 C_{ar}), 136.4 (C_{ar}), 135.5 (C_{ar}),129.6 (C_{ar}), 129.5(C_{ar}), 129.4 (C_{ar}), 128.9 (C_{ar}), 128.9 (C_{ar}), 128.8 (C_{ar}), 128.8 (C_{ar}), 128.7 (C_{ar}), 128.7 (C_{ar}), 128.6 (C_{ar}), 128.5 (C_{ar}), 128.5 (C_{ar}), 128.4 (C_{ar}), 128.3 (C_{ar}), 128.3 (C_{ar}), 128.2 (C_{ar}), 128.1 (C_{ar}), 127.9 (C_{ar}), 127.6 (C_{ar}), 127.0 (C_{ar}), 126.8 (C_{ar}), 124.6 (C_{ar}), 124.5 (C_{ar}), 123.4 (C_{ar}), 123.3 (C_{ar}), 117.4 (C_{ar}), 117.2 (C_{ar}), 76.0 (C_{benzyl}), 75.9 (C_{benzyl}), 72.6 (C-***26***), 71.2 (C_{benzyl}), 71.1 (C_{benzyl}), 68.6 (C-***11/21***), 67.6 (C_{benzyl}), 64.9 (C-***11/21***), 52.8 (C-***10/22***), 51.4 (C-***10/22***)

IR (ATIR cm^{-1}):

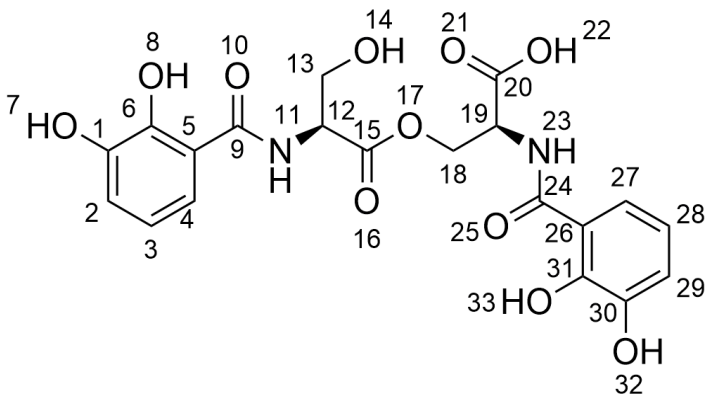
3362 *weak* (N-H), 3032 *weak* (C-H), 1747 *strong* (C=O, ester), 1658 *strong* (C=O, amide), 1575 *strong* (C=C)

Elemental Analysis:

Calculated for [$C_{62}H_{56}N_2O_{11}$.0.25H_2O]: %C 73.76, %H 5.64, %N 2.77

Measured for [$C_{62}H_{56}N_2O_{11}$.0.25H_2O]: %C 73.46, %H 5.51, %N 2.79

Enterobactin Linear Dimer, 17



Compound **30**, 0.359 g (0.35 mmol) was dissolved in 8 mL of dry toluene before 42 mL of dry ethanol was added. To this stirred solution one small spatula tip of Pd-C 10% was added. The mixture was purged with hydrogen and allowed to react for 18 hours under a positive pressure of hydrogen. After the reaction was completed the catalyst was filtered off and the product was isolated by removing the solvent *in vacuo*, to yield a white solid. Characterisation data consistent with literature.[9]

Yield:

0.165 g, 0.32 mmol, 91%

Melting point:

124.4 °C - Foamed; 197.4-200.1 °C - Film melted

***m/z* (ESI):**

463.0094 ([M-H]$^-$, 100%)

HRMS (ESI):

Calc. for [$C_{20}H_{19}N_2O_{11}$]$^-$ = 463.0994 Found 463.0994 (0.1 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

9.00-8.96 (2H, m, H-***11,23***), 7.38 (1H, dd, *J* = 8.4, 1.6 Hz, H-***4/27***), 7.34 (1H, dd, *J* = 8.0, 1.2 Hz, H-***4/27***), 6.96 (1H, dd, *J* = 4.4, 1.2 Hz, H-***2/29***), 6.93 (1H, dd, *J* = 4.8, 1.2 Hz, H-***2/29***), 6.74-6.68 (2H, m, H-***3,28***), 4.81-4.76 (1H, m, H-***12/19***), 4.61 (1H, dd, *J* = 11.2, 4.4 Hz , H-***13/13*/18/18****), 4.58-4.55 (1H, m, H-***12/19***), 4.41 (1H, dd, *J* = 11.2, 6.4 Hz , H-***13/13*/18/18****), 3.83(1H, dd, *J* = 11.6, 6.4 Hz , H-***13/13*/18/18****), 3.77(1H, dd, *J* = 11.2, 4.0 Hz , H-***13/13*/18/18****), 3.44 (2.79H, q, *J* = 7.2 Hz, H-EtOH $\boldsymbol{CH_2}$)♯, 1.05 (2.75H, t, *J* = 7.2 Hz, H-EtOH $\boldsymbol{CH_3}$)

♯Obscured by water signal

^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

171.1 (C-***9/15/20/24***), 170.7 (C-***9/15/20/24***), 169.5 (C-***9/15/20/24***), 169.3 (C-***9/15/20/24***), 149.3 (C-***1/6/30/31***), 148.9 (C-***1/6/30/31***), 146.7 (C-***1/6/30/31***), 146.6 (C-***1/6/30/31***), 119.4 (C_{ar}), 119.3 (C_{ar}), 119.1 (C_{ar}), 118.7 (C_{ar}), 118.7 (C_{ar}), 118.6 (C_{ar}), 116.4 (C_{ar}), 116.0 (C_{ar}), 63.9 (C-***13/18***), 60.9 (C-***13/18***), 56.1 (C-EtOH $\boldsymbol{CH_2}$), 55.4 (C-***12/19***), 51.8 (C-***12/19***), 18.4 (C- EtOH $\boldsymbol{CH_3}$)

IR (ATIR cm^{-1}):

3343 *weak broad* (O-H), 1737 *strong* (C=O, ester), 1639 *medium* (C=O, amide), 1532 *strong* (C=C)

Elemental Analysis:

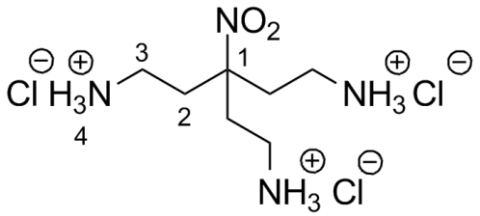
Calculated for [$C_{20}H_{20}N_2O_{11}$0.9EtOH.1.1H_2O]: %C 49.81, %H 5.29, %N 5.33

Measured for [$C_{20}H_{20}N_2O_{11}$.0.9EtOH.1.1H_2O]: %C 49.52, %H 4.94, %N 5.27

Specific Rotation:

$[\alpha]_D$ (Methanol, conc. 0.1 g/100 mL) +15.8

3-(2-Amino-ethyl)-3-nitro-pentane-1,5-diamine-trihydrochloride, 43



A schlenk tube was charged with nitromethane-trispropanoic acid (2.27 g, 8.2 mmol) and placed under a nitrogen atmosphere. To this two drops of DMF was added, followed by thionyl chloride (15 mL, 77 mmol). The suspension was allowed to stir overnight under a postive pressure of nitrogen. Once the solution was clear the excess thionyl chloride was removed *in vacuo*, followed by the replacement of the nitrogen atmosphere. The residue was dissolved in 15 mL anhydrous 1,4-dioxane, followed by the slow addition of trimethylsilyl azide (3.5 mL, 26.6 mmol). This solution was slowly heated over an hour to 70 °C behind a blast shield under an atmosphere of nitrogen. Gas evolution can be observed upon heating (**caution:** after addition of the trimethylsilyl azide the reaction has the capability to form a thermal runaway reaction, producing dinitrogen, potentially becoming expolsive). After the gas evolution slows the solution is allowed to cool to 45 °C. Once at temperature 10 mL of acetone is added, followed by the addition of 6 mL of 6M hydrochloric acid dropwise. The solution turns cloudy and gas evolution can be observed upon the addition of the acid. Once all the acid was added the flask was opened to the atmosphere and the suspension was stirred for an hour. The solid was collected by filtration and washed with ice cold acetone (100 mL) to yield an off white solid. Characterisation data consistent with literature.[10]

Yield:

2.27g, 7.58 mmol, 92%

Melting point:

Decomposed 225 °C

m/z (ESI):

191.1499 ([M+H]$^+$, 100%), 213.1324 ([M+Na]$^+$, 20%)

HRMS (ESI):

Calc for [$C_7H_{19}N_4O_2$]+H$^+$ = 191.1503 Found 191.1499 (1.9 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

8.31 (9H, br s, H-***4***), 2.86-2.77 (6H, m, H-***3***), 2.35-2.28 (6H, m, H-***2***)

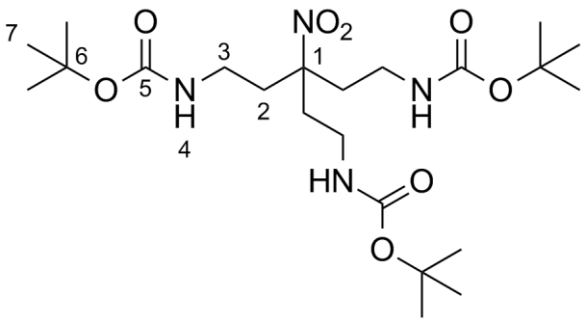
^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

89.7 (C-***1***), 33.8 (C-***2/3***), 31.8 (C-***2/3***)

IR (KBr cm^{-1}):

3437 *strong* (N-H), 3015 *strong* (C-H), 1543 *medium* (NO_{2asymm}), 1457 *weak* (NO_{2symm})

Tert-butyl-*N*-(5-{[(tert-butoxy)carbonyl]amino}-3-(2-{[(tert-butoxy)carbonyl]amino}ethyl)-3-nitropentyl)carbamate, 44



Prepared as in accordance to the literature.[10]

Yield:

0.563 g, 1.15 mmol, 84%

Melting point:

56.8-57.4 °C

R_f:

Ethyl Acetate : R_f : 0.6 (permanganate stain)

m/z (ESI):

491.3068 ([M+H]$^+$, 7%), 513.2885 ([M+Na]$^+$, 100%)

HRMS (ESI):

Calc for $[C_{22}H_{42}N_4O_8]+H^+$ = 491.3075 Found 491.3068 (1.5 ppm error)

Calc. for $[C_{22}H_{42}N_4O_8]+Na^+$ = 513.2895 Found 513.2885 (1.9 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

4.84 (3H, br s, H-***4***), 3.18-3.10 (6H, br m, H-***3***), 2.18 (6H, t, *J* = 7.5 Hz, H-***2***), 1.42 (27H, s, H-***7***)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

156.3 (C-***5***), 90.6 (C-***1***), 79.7 (C-***6***), 35.6 (C-***2/3***), 35.3 (C-***2/3***), 28.1 (C-***7***)

IR (KBr cm^{-1}):

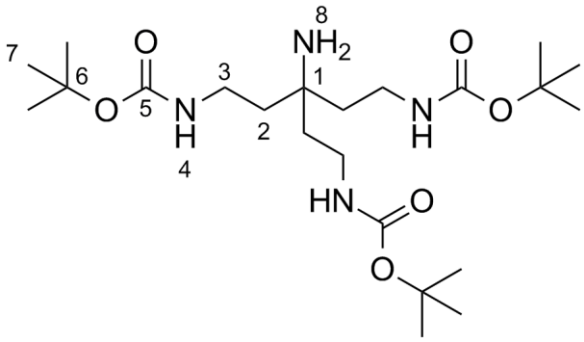
3356 *weak* (N-H), 2978 *weak* (C-H), 1696 *strong* (C=O, carbamate), 1543 *strong* (NO_{2asymm})

Elemental Analysis:

Calculated for $[C_{22}H_{42}N_4O_8]$: %C 53.86, %H 8.63, %N 11.42

Measured for $[C_{22}H_{42}N_4O_8]$: %C 53.46, %H 8.37, %N 11.32

Tert-butyl-*N*-(3-amino-5{[tert-butoxy)carbonyl]amino}-3-(2-{[(tert-butoxy)carbonyl]amino}ethyl)pentyl)carbamate, 45



Compound **44**, 1.651 g (3.36 mmol) was dissolved in 80 mL of ethanol. To this stirred solution one spatula tip of Raney®-Nickel 2800 slurry in water was added. The mixture was first purged with hydrogen before reacting for 18 hours under a hydrogen atmosphere of 35 bar. After the reaction had completed the reaction mixture was filtered through Celite and washed with ethanol, ensuring the catalyst was never allowed to dry. The product was isolated from the filtrate by removing the solvent *in vacuo*, yielding an off white solid. Characterisation data consistent with literature.[10]

Yield:

1.50 g, 3.26 mmol, 97%

Melting point:

61.7-62.9 °C

m/z (ESI):

461.3356 ($[M+H]^+$, 100%), 483.3164 ($[M+Na]^+$, 6%)

HRMS (ESI):

Calc for $[C_{22}H_{44}N_4O_6]+H^+$ = 461.3334 Found 461.3356 (-5.0 ppm error)

Calc. for $[C_{22}H_{44}N_4O_6]+Na^+$ = 483.3153 Found 483.3164 (-2.2 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

5.17-5.12 (3H, br m, H-***4***), 3.18 (6H, q, *J* = 6.5 Hz, H-***3***), 1.62 (6H, t, *J* = 7.5 Hz, H-***2***), 1.42 (27H, s, H-***7***)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

156.3 (C-***5***), 79.3 (C-***6***), 53.3 (C-***1***), (38.8 (C-***2/3***), 35.8 (C-***2/3***), 28.1 (C-***7***)

IR (KBr cm^{-1}):

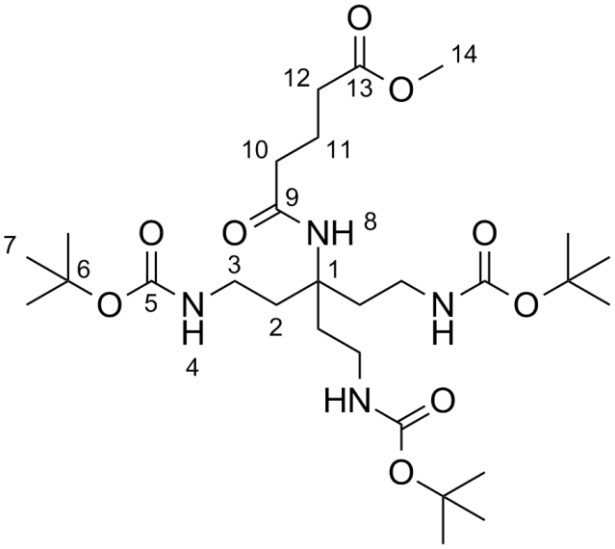
3370 *strong* (N-H), 2978 *weak* (C-H), 1696 *strong* (C=O, carbamate), 1521 *strong* (N-H)

Elemental Analysis:

Calculated for $[C_{22}H_{44}N_4O_6 0.5H_2O]$: %C 56.27, %H 9.66, %N 11.93

Measured for $[C_{22}H_{44}N_4O_6 0.5H_2O]$: %C 56.22, %H 9.40, %N 11.80

Methyl-4-{[1,5-bis({[tert-butoxy)carbonyl]amino})-3-(2-{[tert-butoxy)carbonyl]amino}ethyl)pentan-3-yl]carbamoyl}butanoate, 46



Compound **45** (1.461 g, 3.17 mmol) was dissolved in 15 mL dichloromethane followed by the addition of triethylamine (860 µL, 6 mmol). The solution was cooled on ice before the addition of glutaric acid monomethyl ester chloride (750 µL, 5.4 mmol) and allowed to reaction for 1 hour on ice followed by 1 hour at room temperature. The volatiles were subsequently removed *in vacuo* and the resulting oil dissolved in 50 mL ethyl acetate. This was washed with water (3 x 50 mL). The aqueous layer was then washed with ethyl acetate (2 x 50 mL), and the two organic layers combined. The organic solution was then dried over $MgSO_4$, filtered and the solvent removed *in vacuo*, yielding the crude product. The crude product was purified *via* column chromatography, 4:1 chloroform : ethyl acetate to yield a pure white solid.

Yield:

1.422 g, 2.42 mmol, 76%

R_f:

Chloroform : Ethyl Acetate (4:1) : R_f : 0.39 (permanganate stain)

m/z (ESI):

589.3807 ($[M+H]^+$, 37%), 611.3645 ($[M+Na]^+$, 37%)

HRMS (ESI):

Calc for $[C_{28}H_{52}N_4O_9]+H^+$ = 589.3807 Found 589.3807 (0.1 ppm error)

Calc. for $[C_{28}H_{52}N_4O_9]+Na^+$ = 611.3645 Found 611.3645 (-3.0 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

6.57 (1H, br s, H-***8***), 4.84 (3H, t, *J* = 5.0 Hz, H-***4***), 3.65 (3H, s, H-***14***), 3.09 (6H, q, *J* = 6.5 H-***3***), 2.35 (2H, t, *J* = 7.5 Hz, H-***10/12***), 2.20 (2H, t, *J* = 7.0 Hz, H-***10/12***), 1.93-1.88 (8H, m, H-***2,11***), 1.40 (27H, s, H-***7***)

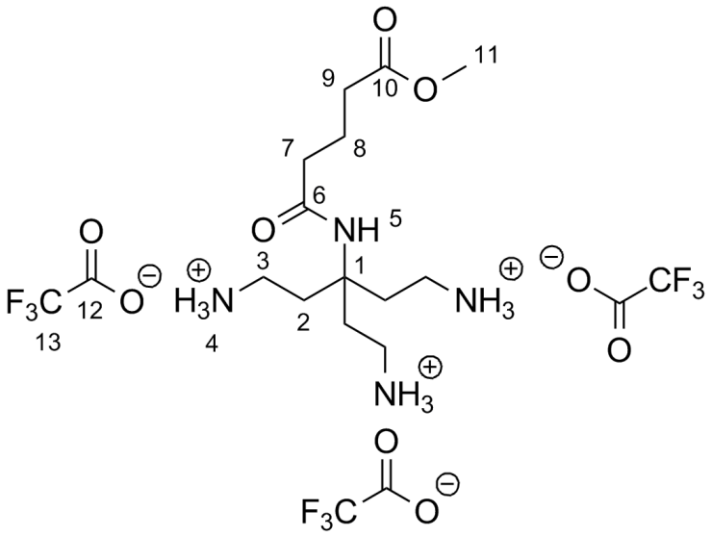
^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

174.1 (C-***9/13***), 173.1 (C-***9/13***) 155.6 (C-***5***), 79.3 (C-***6***), 56.2 (C-***1***), 51.4 (C-***14***), (35.8 (C-***2/3***), 35.8 (C-***10/12***) 35.6 (C-***2/3***), 33.0 (C-***10/12***), 28.1 (C-***7***), 20.5 (C-***11***)

IR (ATIR cm^{-1}):

3331 *medium* (N-H), 2978 *medium* (C-H), 1690 *strong* (C=O, carbamate), 1662 *medium* (C=O, amide), 1513 *strong* (N-H)

Tris(trifluoroacetic acid) methyl 4-{[1,5-diamino-3-(2-aminoethyl)pentan-3-yl]carbamoyl}butanoate, 47



Compound **46** (1.359 g, 2.31 mmol) was dissolved in 50 mL 9:1 dichloromethane : methanol and stirred. To this stirred solution 5 mL of trifluoroacetic acid was added and allowed to react. The reaction was followed *via* TLC (ethyl acetate) and allowed to complete (~24 hours). The volatiles were

removed *in vacuo* to yield a pale yellow oil, which was taken up in 5 mL dry ethanol, which was subsequently removed *in vacuo*. This process was repeated two more times to remove traces of TFA and yield a white hydroscopic solid material.

Yield:

1.440 g, 2.28 mmol, 98%

m/z (ESI):

289.2247 ($[M+H]^+$, 100%)

HRMS (ESI):

Calc for $[C_{13}H_{28}N_4O_3]+H^+$ = 289.2234 Found 289.2247 (-4.5 ppm error)

^{1}H NMR: (d_6-DMSO, 400 MHz) δ_H

7.89 (9H, br s, H-***4***), 7.53 (1H, s, H-***5***), 3.58 (3H, s, H-***11***), 2.78-2.68 (6H, br m, H-***3***), 2.30 (2H, t, *J* = 7.5 Hz H-***7/9***), 2.13 (2H, t, *J* = 7.5 Hz H-***7/9***), 1.96-1.88 (6H, m, H-***2***), 1.73 (2H, app quin, *J* = 7.5 Hz, H-***8***)

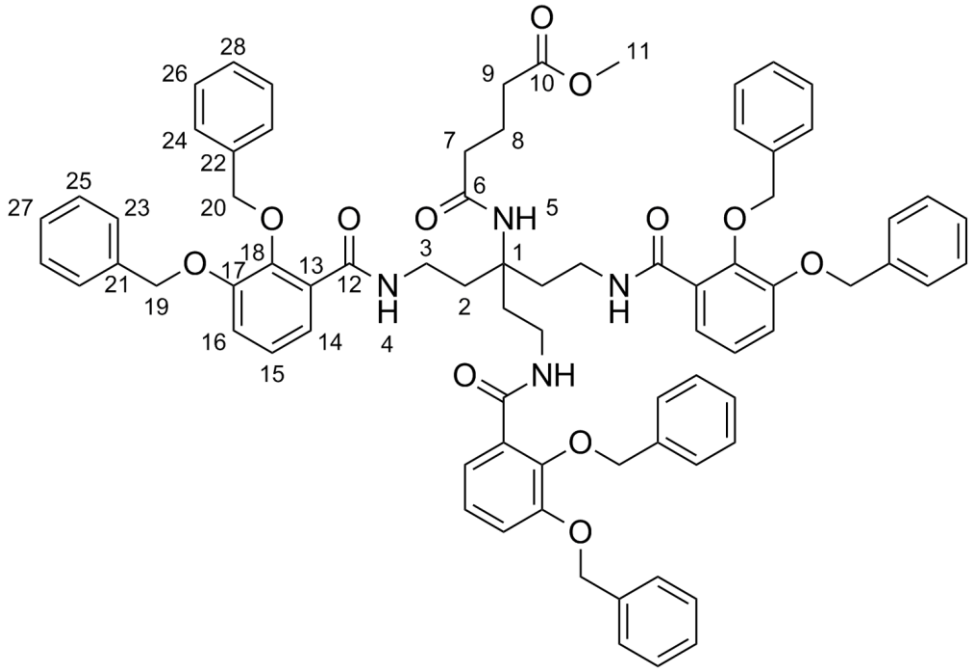
^{13}C{^{1}H} NMR: (d_6-DMSO, 100 MHz) δ_C

173.2 (C-***10***), 172.1 (C-***6***), 158.9 (q, *J* 35 Hz, C-***12***), 115.9 (q, *J* 290 Hz, C-***13***), 54.4.6 (C-***1***), 51.4.1 (C-***11***), 34.7 (C-***7/9***), 34.2 (C-***3***), 32.8 (C-***7/9***), 31.8 (C-***2***), 20.7 (C-***8***)

IR (ATIR cm^{-1}):

3059 *weak, broad* (N-H), 2943 *weak, broad* (C-H), 1668 *strong, broad* (C=O, amide)

4-{[1,5-diamino-3-(2-aminoethyl)-3-amide]butanoate}-*N,N,N*-[-[2,3-bis(phenylmethoxy)benzoyl]-methyl ester, 48



Compound **47** (1 mmol, 0.630 g) was suspended in 10 mL dry acetonitrile and placed on ice. To the stirred solution DIPEA (15 mmol, 2.6 mL) was added, and the system was placed under an atmosphere of nitrogen. Compound **31** was dissolved in 6 mL dry acetonitrile and slowly added and allowed to stir overnight at room temperature. The solution was diluted with 50 mL dichloromethane and washed with 30 mL water followed by 20 mL 0.1M HCl, 20 mL; saturated $NaHCO_3$; and H_2O until the aqueous washes were neutral. The organic layer was dried over $MgSO_4$, filtered, and the solvent removed *in vacuo* to yield the crude product. The crude product was purified *via* flash column chromatography on silica 4:1 EtOAc : dichloromethane, yielding an off white solid.

Yield:

0.1902 g, 0.154 mmol, 17%

R_f:

Ethyl Acetate : R_f : 0.40

m/z (ESI):

1259.5357 ([M+Na]$^+$, 100%)

HRMS (ESI):

Calc for [$C_{76}H_{76}N_4O_{12}$]+Na$^+$ = 1259.5352 Found 1259.5357 (-0.4 ppm error)

^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

8.04 (3H, t, *J* = 5.5 Hz, H-***4***), 7.73-7.69 (3H, m, H-***14/15/16***), 7.50-7.11 (36H, m, H_{ar}), 5.15 (6H, s, H-***19/20***), 5.09 (6H, s, H-***19/20***), 3.65 (3H, s, H-***11***), 3.18 (6H, dt, *J* = 9.6, 6 Hz, H-***3***), 2.36 (2H, t, *J* = 7.0 Hz, H-***7/9***), 2.26 (2H, t, *J* = 7.5 Hz, H-***7/9***), 1.94 (2H, app quin, *J* = 7.2 Hz , H-***8***), 1.79-1.73 (6H, m, H-***2***)

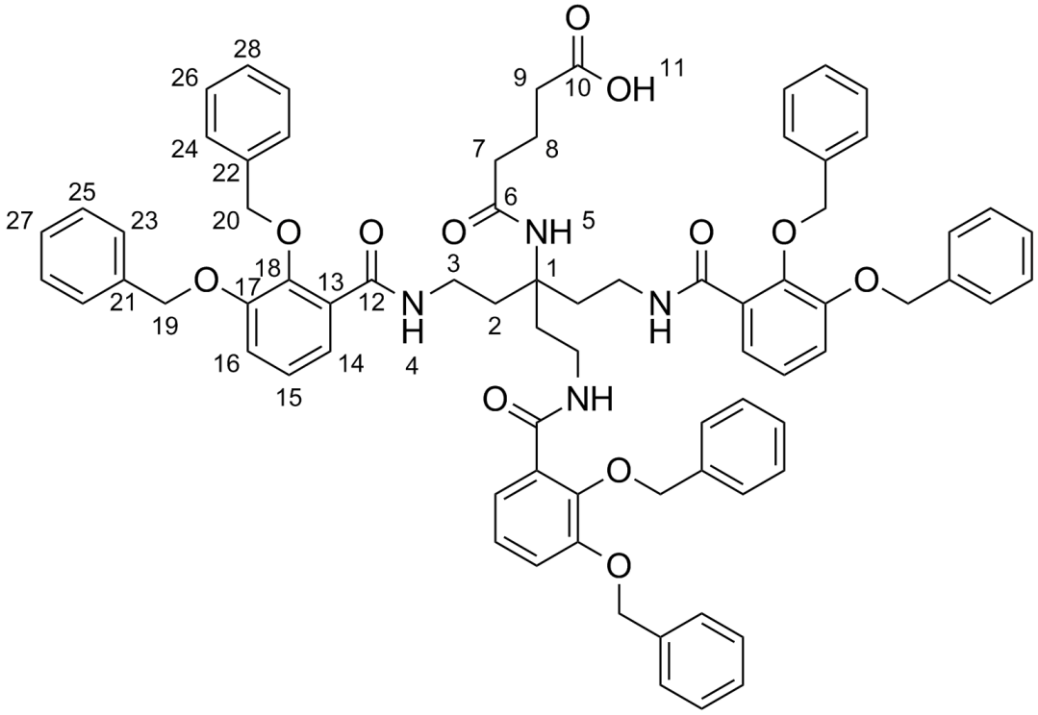
^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

174.1 (C-***6/10***), 172.6 (C-***6/10***), 165.7 (C-***12***), 152.0 (C_{ar}), 147.2 (C_{ar}), 136.7 (C_{ar}), 136.5 (C_{ar}), 129.2 (C_{ar}), 129.0 (C_{ar}), 128.9 (C_{ar}), 128.9 (C_{ar}), 128.5 (C_{ar}), 127.9 (C_{ar}), 127.4 (C_{ar}), 124.5 (C_{ar}), 123.4 (C_{ar}), 117.3 (C_{ar}), 76.3 (C-***19/20***), 71.2 (C-***19/20***), 55.9 (C-***1***), 51.3 (C-***11***), (35.8 (C-***7/9***), 35.1 (C-***3***) 34.9 (C-***2***), 33.1 (C-***7/9***), 20.5 (C-***8***)

IR (ATIR cm^{-1}):

3362 *weak* (N-H), 3031 *weak* (C-H), 2949 weak (C-H), 1734 *medium* (C=O, ester), 1650 *strong* (C=O, amide), 1575 *strong* (C=C)

4-{[1,5-diamino-3-(2-aminoethyl)-3-amide]butanoic}-*N,N,N*-[-[2,3-bis(phenylmethoxy)benzoyl]-acid, 49



Compound **48** (46.8 mg, 0.038 mmol) was dissolved in 9:1 mL of dry dichloromethane : methanol and stirred. To this stirred solution, 150 µL of a 2M methanolic sodium hydroxide solution was added and the solution. The reaction was followed *via* TLC (ethyl acetate) and allowed to complete (~18 hours). The solution was then diluted with 40 mL dichloromethane and washed with 2 x 30 mL 0.1 M formic acid. The organic layer was dried with $MgSO_4$, filtered and dried *in vacuo* yielding a clear glass foamed product.

Yield:

43.4 mg, 0.035 mmol, 94%

Melting point:

68.9-69.7 °C

m/z (ESI):

1221.5240 ($[M-H]^-$, 47%)

HRMS (ESI):

Calc for $[C_{75}H_{73}N_4O_{12}]^-$ = 1221.5230 Found 1221.5240 (-0.8 ppm error)

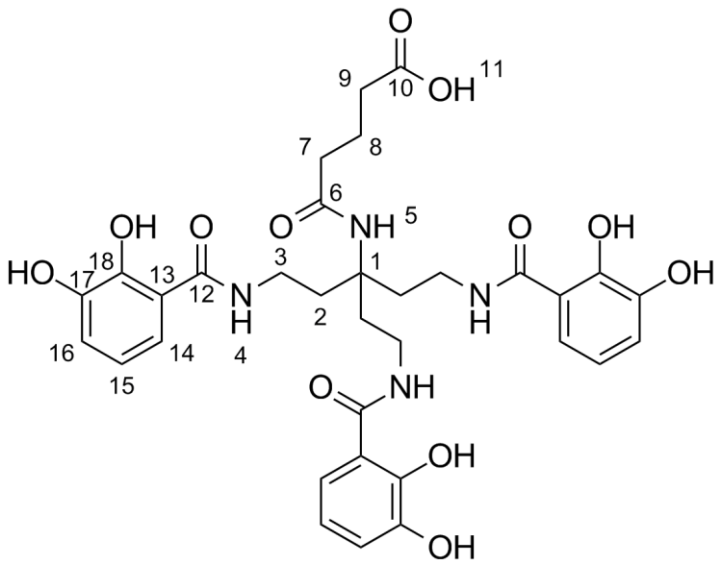
^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

8.08 (3H, t, *J* = 5.5 Hz, H-***4***), 7.70-7.66 (3H, m, H-***14/15/16***), 7.47-7.09 (36H, m, H_{ar}), 6.96 (1H, br s, H-***5***), 5.13 (6H, s, H-***19/20***), 5.07 (6H, s, H-***19/20***), 3.16 (6H, dt, *J* = 9.2, 6 Hz H-***3***), 2.42 (2H, t, *J* = 7.0 Hz, H-***7/9***), 2.29 (2H, t, *J* = 7.0 Hz, H-***7/9***), 1.94 (2H, app quin, *J* = 6.8 Hz , H-***8***), 1.72-1.65 (6H, m, H-***2***)

^{13}C{^{1}H} NMR: ($CDCl_3$, 100 MHz) δ_C

175.4 (C-***10***), 172.9 (C-***6***), 165.7 (C-***12***), 151.8 (C_{ar}), 147.0 (C_{ar}), 136.5 (C_{ar}), 136.3 (C_{ar}), 129.0 (C_{ar}), 128.9 (C_{ar}), 128.9 (C_{ar}), 128.8 (C_{ar}), 128.4 (C_{ar}), 127.8 (C_{ar}), 127.0 (C_{ar}), 124.5 (C_{ar}), 123.3 (C_{ar}), 117.4 (C_{ar}), 76.5 (C-***19/20***), 71.4 (C-***19/20***), 56.4 (C-***1***), 35.9 (C-***7/9***), 35.2 (C-***2/3***) 35.1 (C-***2/3***), 34.0 (C-***7/9***), 21.0 (C-***8***)

4-{[1,5-diamino-3-(2-aminoethyl)-3-amide]butanoic}-*N,N,N*-[2,3-dihydroxybenzamide]-acid, 41



Compound **49**, (32.4 mg, 0.026 mmol) was dissolved in 5 mL of dry toluene before 45 mL of dry ethanol was added. To this stirred solution one small spatula tip of Pd-C 10% was added. The mixture was purged with hydrogen and allowed to react for 18 hours under a positive pressure of hydrogen. After the reaction was completed the catalyst was filtered off and the product was isolated by removing the solvent *in vacuo*, to yield a off-white film.

Yield:

16.6 mg, 0.024 mmol, 92%

m/z (ESI):

681.2420 ([M-H]$^-$, 26%)

HRMS (ESI):

Calc for [$C_{33}H_{37}N_4O_{12}$] = 681.2413 Found 681.2420 (-0.9 ppm error)

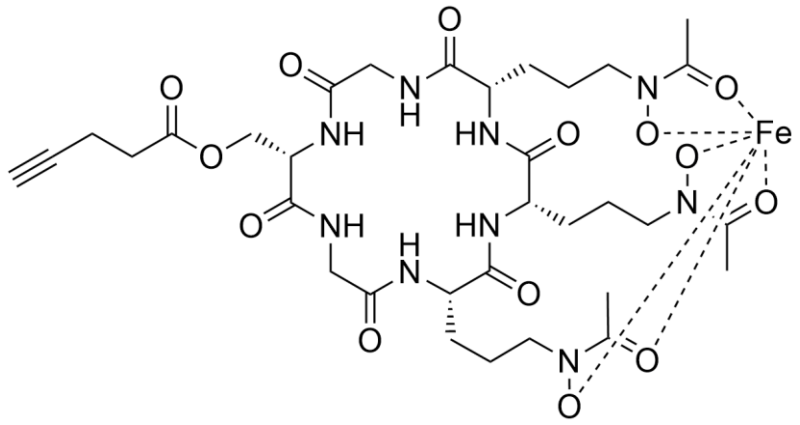
^{1}H NMR: ($CDCl_3$, 400 MHz) δ_H

7.19 (3H, dd, *J* = 8.4, 1.2 Hz, H-***14/16***), 6.91 (3H, dd, *J* = 8.0, 1.2 Hz, H-***14/16***), 6.70 (3H, t, *J* = 8.2 Hz, H-***15***), 3.55-3.48 (6H, m, H-***3***), 2.33 (2H, t, *J* = 7.2 Hz, H-***7/9***), 2.25 (2H, t, *J* = 7.6 Hz, H-***7/9***), 2.20-2.13 (6H, m, H-***2***), 1.88 (2H, app quin, *J* 7.4 Hz, H-***2***)

^{13}C{^{1}H} NMR: (MeOD, 100 MHz) δ_C

176.1 (C-***10***), 174.0 (C-***6***), 170.2 (C-***12***), 149.0 (C-***17/18***), 145.9 (C-***17/18***), 118.3 (C_{ar}), 118.2 (C_{ar}), 117.3 (C_{ar}), 115.3.9 (C_{ar}), 56.4 (C-***1***), 35.5 (C-***7/9***), 34.4 (C-***2/3***) 34.1 (C-***2/3***), 33.2 (C-***7/9***), 21.0 (C-***8***)

Ferric-{8,11,14-tris[3-(*N*-hydroxyacetamido)propyl]-3,6,9,12,15,18-hexaoxo-1,4,7,10,13,16-hexaazacyclooctadecan-2-yl}methyl pent-4-ynoate, 53



To a stirred solution of pent-4-ynoic acid (98.7 mg, 1.01 mmol), in dry dichloromethane (10 mL), DCC (107.6 mg, 0.52 mmol) was added and the flask and placed under an atmosphere of nitrogen. The resulting solution was stirred for three hours at room temperature. The solution was then placed in an ice bath for one hour, and the resulting precipitation was filtered using a cannula filter, to yield crude anhydride in dichloromethane. Ferricrocin (14.6 mg, 0.019 mmol) was dissolved in dry pyridine (3 mL) and added to the solution of crude anhydride dissolved in dichloromethane (5 mL), which was stirred for three hours at room temperature. The solution was then dried *in vacuo* to yield a brown solid. This solid was then purified *via* LH-20 size exclusion chromatography (1:1 $CHCl_3$:MeOH), to yield a brown solid.

HPLC:

HPLC was performed on a reversed phase column, using MeCN:H_2O gradient (6-40 %), from 0-20 minutes. Retention time found to be 17 min.

m/z (ESI):

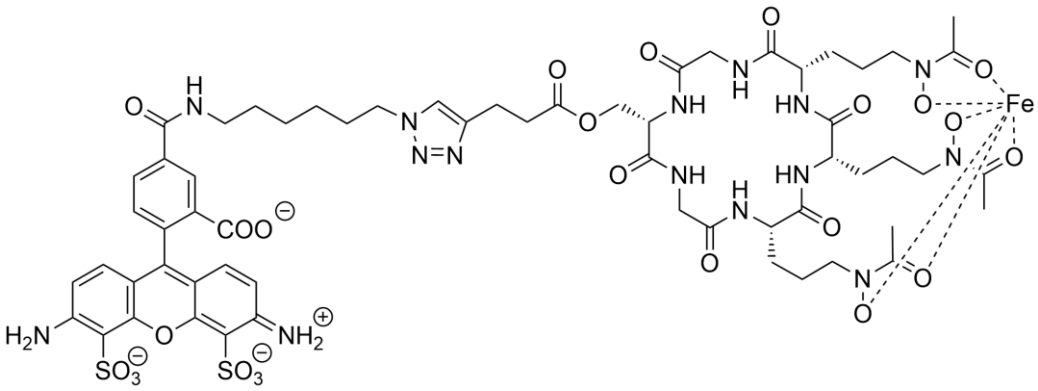
873.2541 ($[M+Na]^+$, 33%), 851.2705 ($[M+H]^+$, 17%)

HRMS (ESI):

Calc for $[FeC_{33}H_{48}N_9O_{14}]+H^+$ = 851.2743 Found 851.2705 (4.4 ppm error)

Calc for $[FeC_{33}H_{48}N_9O_{14}]+Na^+$ = 873.2563 Found 873.2541 (2.5 ppm error)

Ferric-2-(6-amino-3-imino-4,5-disulfo-3H-xanthen-9-yl)-5-[(6-{4-[3-oxo-3-({8,11,14-tris[3-(*N*-hydroxyacetamido)propyl]-3,6,9,12,15,18-hexaoxo-1,4,7,10,13,16-hexaazacyclooctadecan-2-yl}methoxy)propyl]-1H-1,2,3-triazol-1-yl}hexyl)carbamoyl]benzoic acid, 51



The Alexa488 Azide dye was dissolved in HPLC grade water (2 mL) and added to a stirred solution of **53** in HPLC grade water (3 mL). To this reaction mixture, sodium ascorbate (6 μL of 10 mM stock solution) was added, followed by $CuSO_4.5H_2O$ (3 μL of a 5 mM stock solution). The resulting solution was stirred at room temperature for six hours in the dark and then placed in a 5 mL 0.1-0.5 MW dialysis bag and dialysed for 90 minutes. The solution was then dried *in vacuo*.

LC-MS

Synergi Hydro column, 4.6mm x 250mm; Mobile phase solvent A = H_2O solvent B = MeCN; Gradient Profile (0-30 min 94% A, 6% B to 60% A, 40% B, 20-40 min held at 60% A, 40% B; Flow rate 0.3 $mLmin^{-1}$; UV wavelength 295 nm.

Retention time: 31.5 min

Calc for $[C_{60}H_{72}FeN_{15}O_{24}S_2]^{2-}$ = 753.1839 Found 753.1849 (-1.3 ppm error)

Calc for $[C_{60}H_{72}FeN_{15}O_{24}S_2]^{-}$ = 1507.3749 Found 1507.3876 (-8.3 ppm error)

m/z (MALDI:

1478.46111 ($[M-Fe+5H+Na]^+$, 100%), 1456.46872 ($[M-Fe+6H]^+$, 90%), 1494.41644 ($[M-Fe+5H+K]^+$, 65%), 1500.44920 ($[M-Fe+4H+2Na]^+$, 64%), 1516.42003 ($[M-Fe+4H+Na+K]^+$, 50%)

HRMS (MALDI):

Calc for $[C_{60}H_{72}N_{15}O_{24}S_2] + 6H^+ = 1456.46872$

Found 1456.46872 (1.79 ppm mean error)

Calc for $[C_{60}H_{72}N_{15}O_{24}S_2] + 5H^+ + Na^+ = 1478.45995$

Found 1478.46111 (-0.62 ppm mean error)

Calc for $[C_{60}H_{72}N_{15}O_{24}S_2] + 5H^+ + K^+ = 1494.43389$

Found 1494.41644 (10.63 ppm mean error)

Calc for $[C_{60}H_{72}N_{15}O_{24}S_2] + 4H^+ + 2Na^+ = 1500.44190$

Found 1500.44920 (-8.38 ppm mean error)

Calc for $[C_{60}H_{72}N_{15}O_{24}S_2] + 4H^+ + Na^+ + K^+ = 1516.41583$

Found 1516.42003 (-6.79 ppm mean error)

UV-vis:

$\lambda_{max}(H_2O)$ / 494 nm (ε : Not determined: Unable to record a mass)

7.3 Protein Production and Crystallisation Trials

The CeuE coding sequence was cloned from genomic DNA. The PCR amplification of the sequence was performed in a 50 μL reaction containing Kod Hot start polymerase (1 μL), PCR buffer 10x (5 μL), dNTP's (5 μL), $MgSO_4$ (2 μL), genomic DNA (0.5 μL, 500 μg / mL), forward primer, (5'-CCAGGGACCAGCAATGTTGCCTATTAGTATGAGCGATGAG-3') (1 μL, 20 μM), reverse primer, (5'-GAGGAGAAGGCGCGTTATTATTTTACAGCGTTTTTGATTTC-3') (1 μL, 20 μM), made up to 50 μL with mQ grade water. The PCR cycle performed consisted of 95 °C for two minutes and 30 cycles of 95 °C for 20 seconds, 48 °C for 30 seconds and 72 °C for 37 seconds. The PCR product was inserted into pET-YSBLIC 3C vector. The CeuE was then expressed in *E. coli* BL-21 cells. Single colonies were used to inoculate 500 mL cultures of Luria-Bertani broth containing 30 μg/mL of the antibiotic kanamycin. When the OD_{600} reached approximately 0.8, CeuE expression was induced with the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside. The samples were then incubated at 37 °C for four hours and then harvested by centrifugation. The cells were then resuspended in 50 mM TrisHCl pH 7.5, 500 mM NaCl, 10 mM imidazole, with C-complete protease inhibitor cocktail EDTA-free present, and sonicated to disrupt the cells. The insoluble portion was separated from the solution phase using centrifugation, and the crude supernatant was loaded onto a 5 mL nickel-agarose affinity column. The His-tagged CeuE was eluted with an increasing gradient (10-500 mM) imidazole in buffer, and fractions analysed by SDS-PAGE. Fractions containing CeuE were pooled, and subjected to dialysis at 4 °C for 48 hours in the presence of C-protease (100:1). The mixture after dialysis was analysed by SDS-PAGE found that all the His-tags had not been removed. Therefore the sample was subjected to C-protease again (50:1) and mixed at 4 °C overnight. The SDS-PAGE confirmed all of the His-tags had now been removed. The mixture was loaded onto a 5 mL nickel-agarose affinity column, and eluted with the same buffer gradient as previous. The fractions collected were pooled and concentrated, then applied to a gel filtration column

(equilibrated in 20 mM TrisHCl pH 8.0, 150 mM NaCl). The fractions were checked by SDS-PAGE, pooled and concentrated to 18 mg/mL. The protein solution was then divided into 50 μL aliquots and shock frozen in liquid nitrogen and stored at -80 °C until required.

Crystallisation. The CeuE stock as described above was used for the crystal screening. All initial screening was carried out using a Mosquito Crystal Robot and Hampton and PACT commercial crystal screens. These screens used 150 nL of protein and 150 nL of screening solution in the crystallisation drop. The best apo-CeuE crystals were obtained with 0.1 M MIB buffer, pH 5, 25% (w/v) PEG 1500.

For the {CeuE[Fe(4-LICAM)]} complex a stock solution of ferric 4-LICAM was prepared by mixing 3 mg of H_4-4-LICAM in 1 mL of methanol with 35 μL of a 0.16 M stock of $FeCl_3$, and stirring for 2 hours. The methanol was removed *in vacuo* to yield a purple solid. The solid (1.3 mg) was dissolved in 11 μL DMF to yield a 100 mM stock assuming the complex to be a 3:2 triple stranded complex. The CeuE stock (57 μL) was mixed with this solution (3 μL) which caused precipitation. The purple precipitate was removed by centrifugation yielding the pale purple ferric 4-LICAM CeuE stock solution used for crystallisation. Initial screening was carried out as described for apo-CeuE. The best crystals were obtained with 0.2 M sodium nitrate, 0.1 M Bis Tris propane, pH 7.5, 20% (w/v) PEG 3350.

For the {CeuE[Fe(**17**)} crystals were obtained through crystal soaking experiments. A crystal was grown from the PACT crystal screen and Fe-**17** was introduced by the addition of H_4-**17** (2 μL, $1x10^{-2}$ M in DMSO) and Fe(NTA) (2 μL, $1x10^{-2}$ M, aqueous) to the condition reservoir, followed by thorough mixing. 2 μL of the resulting solution was added the sitting drop that contained the crystal. The well was re-sealed and time allowed for the ferric-**17** to diffuse through the crystal.

Data Collection, Structure Solution and Refinement. The {CeuE[Fe(4-LICAM)]} crystal was coated with a cryoprotectant solution consisting of the

mother liquor with 20% PEG 1.5 K before flash-cooling in liquid nitrogen and storage for data collection. All other crystal data collection reported within this thesis were flash-cooled in liquid nitrogen without the use of cryoprotectant. Crystal X-ray diffraction data was collected at the Diamond Light Source synchrotron. The structure was determined by molecular replacement using one domain from the ferric-MECAM structure (PDB code 2CHU) as a search model using MOLREP[11] in the CCP4[12] software package. Model building was carried out using Coot[13] and refinement using REFMAC[14].

7.4 Job Plot Method

A series of aqueous solutions containing the ligand of interest and Fe(NTA) were prepared such that the sum of the concentrations of both remained constant (400 µM) in 0.1 M TrisHCl pH 7.5, 5% DMSO. The ratio of ligand and metal was varied from 100% ligand to 100% metal. The absorbance of interest for the system was plotted against the ligand-to-metal ratio.

110 mM TrisHCl pH 7.5. Tris(hydroxymethyl)aminomethane, (27.5 mmol, 3.334 g) was dissolved in approximately 200 mL distilled water. The pH was adjusted to 7.5 with 2.5M HCl. The volume of the solution was made up t 250 mL with distilled water.

Ligand Solution. A stock solution of 10 mM of each ligand was prepared in DMSO.

Fe(NTA) Solution. Nitrilotriacetic acid trisodium salt (0.1 mmol) was dissolved in 0.0179 mol dm^{-3} standard $Fe(NO_3)_3$ 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} $Fe(NO_3)_3$ with 0.01 mol dm^{-3} NTA.

7.5 Circular Dichroism

Circular dichroism spectra were recorded using a Jasco J-810 spectropolarimeter with a Peltier temperature control unit at 20 °C under constant nitrogen flush. The instrument was operated with the following parameters: range: 300-700 nm, data pitch: 0.5 nm, scanning mode: continuous, scanning speed: 100 nm/min, response: 2 seconds, bandwidth: 2 nm, accumulation: 5, pathlength: 1 cm.

{CeuE[Fe(4-LICAM)]}. The sample consisted of 550 μL of CeuE stock solution (18 mg/mL in 20 mM TrisHCl pH 8.0, 10 mM NaCl), 50 μL of a ferric 4-LICAM (2.55 x 10^{-3} M in 50 mM TrisHCl pH 7.5, 25% DMSO) and 1400 μL 20 mM TrisHCl pH 8.0. The ferric 4-LICAM spectrum was then recorded with a solution containing 50 μL of a ferric 4-LICAM stock solution (2.55 x 10^{-3} M in 50 mM TrisHCl pH 7.5, 25% DMSO) and 1950 μL 20 mM TrisHCl pH 8.0. The blank spectrum was subtracted from both data sets to give the spectra shown.

Fe-17. The samples consisted of a total concentration of [L]+[M] of 0.4 mM, recorded in 100 mM TrisHCl pH 7.5, 150 mM NaCl, 5% DMSO, with ligand to metal ratios recorded at 1:0, 3:2, 1:1, 2:3, 0:1. The blank spectrum was subtracted from all spectra.

{CeuE[Fe(17)]}. The samples consisted of 50 μM Fe-**17**, 0.1 M TisHCl pH 7.5, 150 mM NaCl (1000 μL) titrating in 20 μL aliquots of a CeuE stock solution (17.5 mg/mL in 20 mM TrisHCl pH 8.0, 10 mM NaCl). The blank spectrum was subtracted from all spectra.

7.6 Fluorescence Quenching Titration

Intrinsic fluorescence spectroscopy was carried out at room temperature. The excitation slit width of 10 nm and an emission slit width of 20 nm, scan speed 60 nm/min with an automatic response. The detector voltage was set at 950 V. CeuE was excited at 280 nm and the emission spectrum was recorded from 285 nm to 415 nm. For each measurement a protein solution of 240 nM in 2000 μL, 40 mM TrisHCl (pH 7.5) NaCl 150 mM was placed in a 1 cm quartz cuvette, and titrated stepwise with concentrated ferric-ligand stock solution. After each addition the solution was thoroughly mixed and allowed to rest for one minute. The integrated emission (corrected for PMT response) at from 305 nm to 380 nm was used for plotting and binding constant calculation, using the fitting program DynaFit.[15] The non-linear regression fit assumes there is no pre-equilibrium of the ligand and iron(III), and the ferric-ligand complex is a single stable molecule. The scripts for DynaFit can be found in the appendix and on the compact disc attached to the thesis. For each titration, the task and mechanism fields of the script remain the same. The final concentration of CeuE in the cuvette is required under the concentraton field and a initial esitmate of the K_d value is required in the constant field, both in μM. The responses are required to be changed for each indivual titration, with the CeuE response being the initial normallised fluroscent signal divided by the final concentration of CeuE in the cuvette, and the CeuE.FeLD response is the end normallised fluroscent signal divided by the final concentration of CeuE. Under the equilibria field the location of the .TXT file containing the titration data must be located, and the output field must be filled in in order for DynaFit to write the output data.

40 mM TrisHCl pH 7.5, 150 mM NaCl. Tris(hydroxymethyl)aminomethane, (0.02 mol, 2.423 g) and sodium chloride (7.5 mmol, 4.383 g) was dissolved in approximately 400 mL distilled water. The pH was adjusted to 7.5 with 2M HCl. The volume of the solution was made up to 500 mL with distilled water.

Fe-Ligand Solution. 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 TrisHCl pH 7.5, 150 mM NaCl and thoroughly mixed.

Appendix I. Fluorescent Quenching Analysis of CeuE

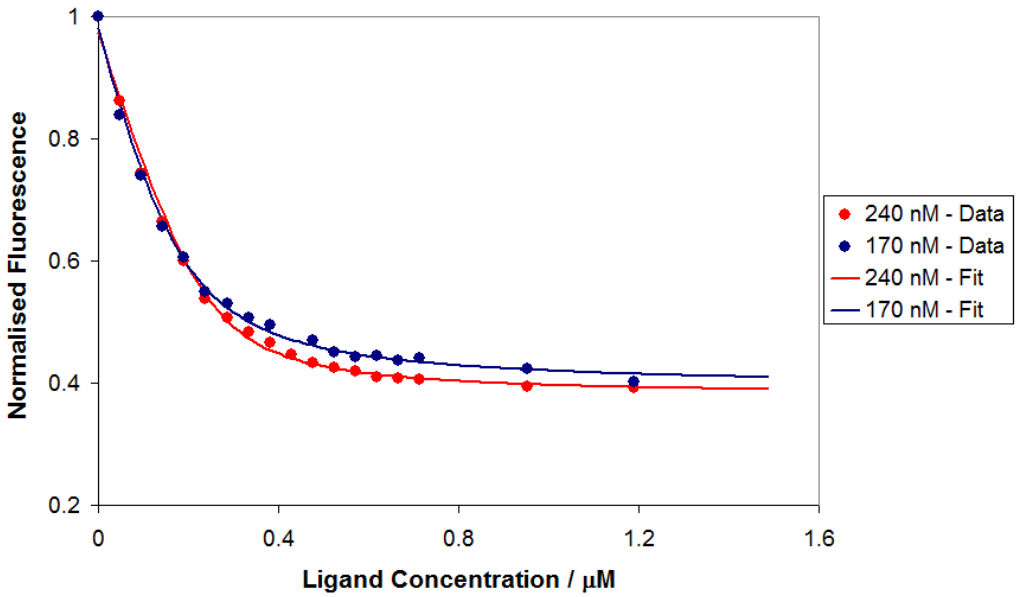


Figure 84 - Run 1 fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 25.1± 1.3 nM (240 nM), 44.9 ± 2.5 nM (170 nM).

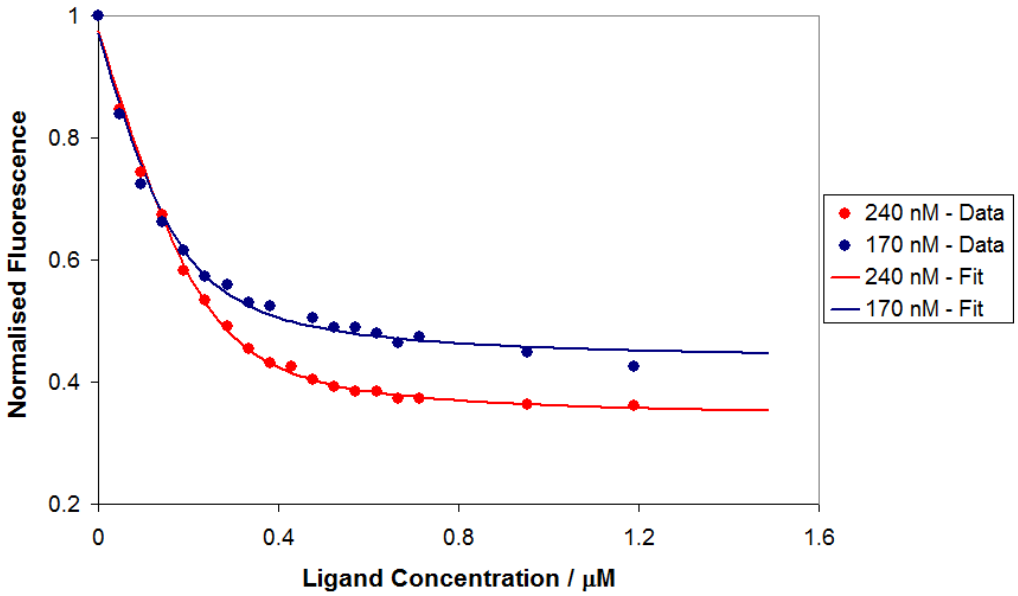


Figure 85- Run 2 fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 30.3± 0.9 nM (240 nM), 40.8 ± 3.8 nM (170 nM).

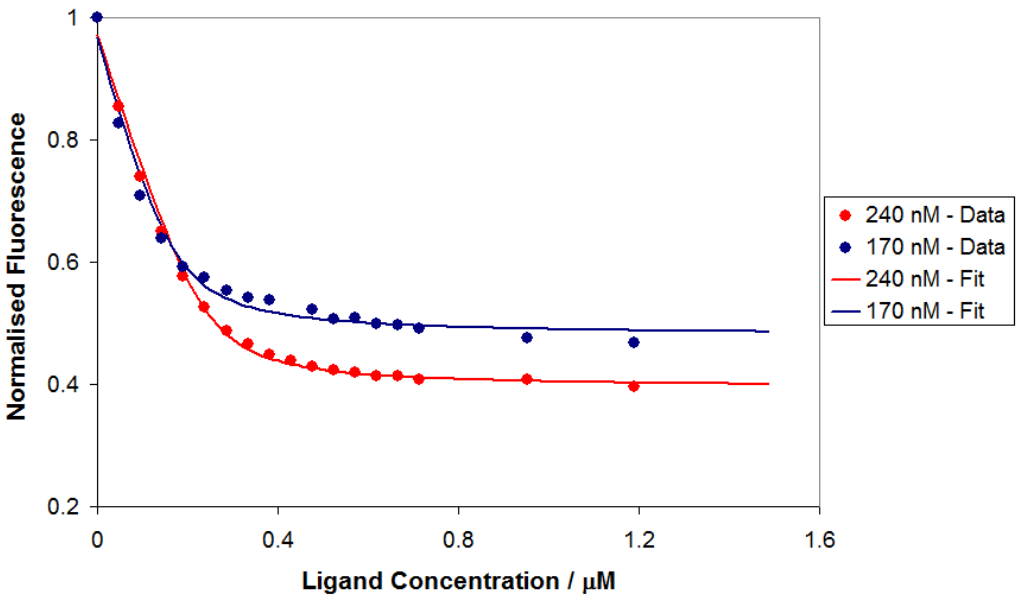


Figure 86 - Run 3 fluorescence quenching analyses of PBP CeuE with ferric-4-LICAM. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 14.9± 2.7 nM (240 nM), 19.9 ± 5.3 nM (170 nM).

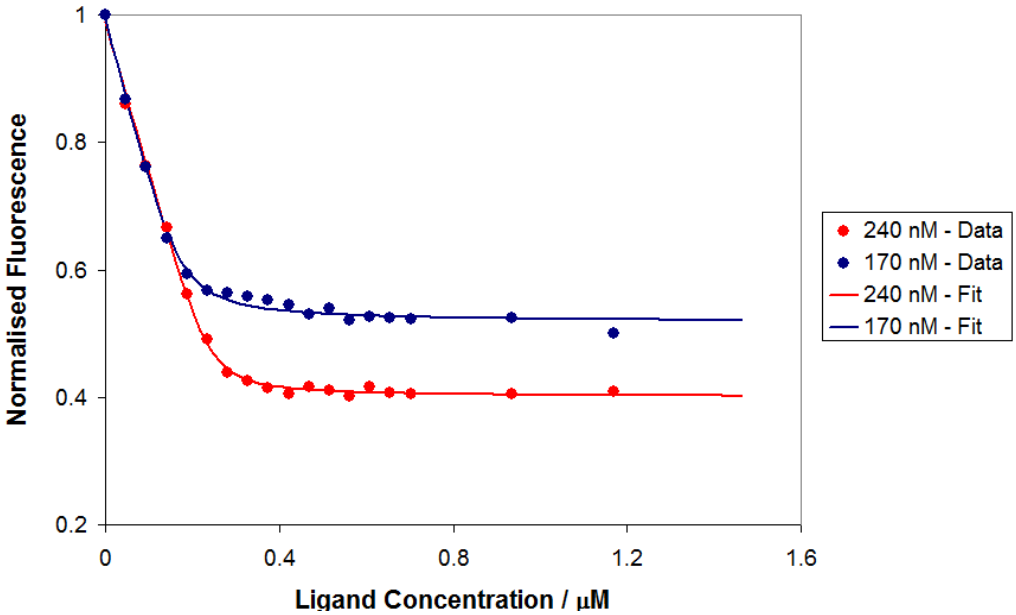


Figure 87 - Run 1 fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 4.8± 0.9 nM (240 nM), 9.8 ± 1.9 nM (170 nM).

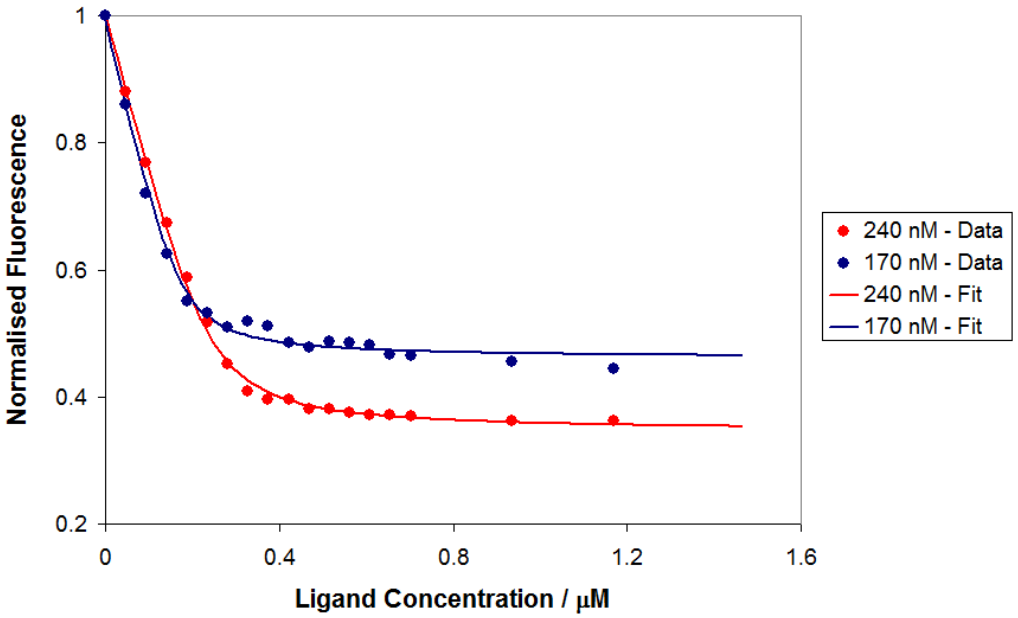


Figure 88 - Run 2 fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 16.4 ± 1.8 nM (240 nM), 12.3 ± 2.5 nM (170 nM).

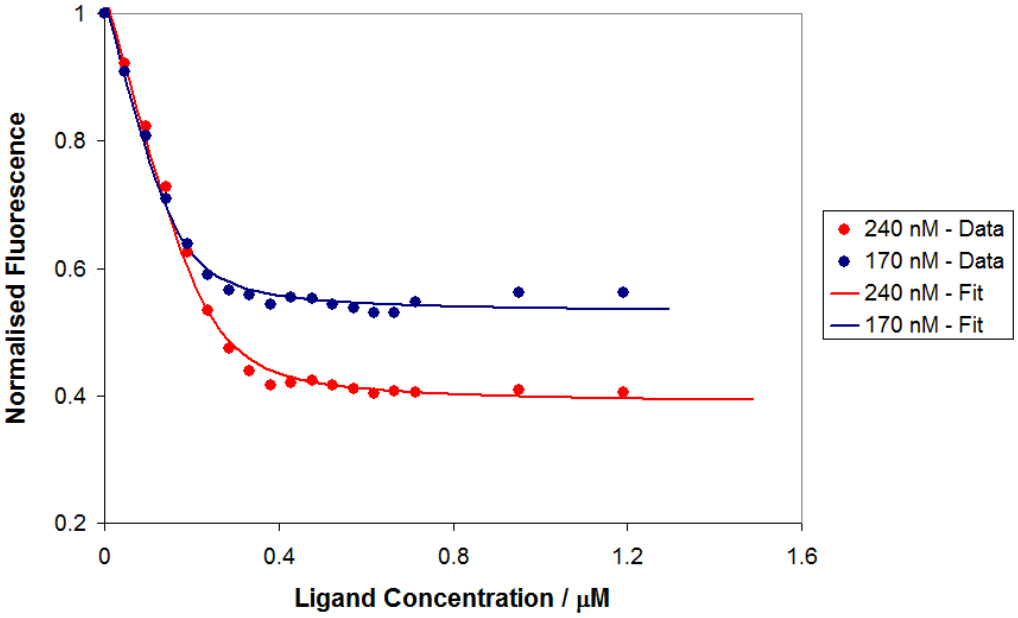


Figure 89 - Run 3 fluorescence quenching analyses of PBP CeuE with ferric-**17**. Titration recorded with 240 nM CeuE (red) and 170 nM CeuE (blue) in 40 mM TrisHCl pH 7.5, NaCl 150 mM. Recorded data shown as circles; lines give the non-linear least-squares calculated fits (DynaFit). K_d = 14.9± 3.7 nM (240 nM), 14.9 ± 3.6 nM (170 nM).

RAW DATA			
Batch 1 - 240 nM			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	1400	1217	857
0.0476	1206	1030	733
0.0952	1042	904	634
0.1428	930	821	558
0.1904	840	708	494
0.238	754	649	451
0.2856	709	597	417
0.3332	678	554	400
0.3808	653	525	385
0.4284	626	517	376
0.476	606	492	368
0.5236	596	478	362
0.5712	587	467	359
0.6188	574	469	355
0.6664	572	454	355
0.714	569	454	350
0.952	553	442	349
1.19	549	440	340

NORMALISED			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	1	1	1
0.0476	0.861428571	0.846343468	0.855309218
0.0952	0.744285714	0.742810189	0.739789965
0.1428	0.664285714	0.674609696	0.651108518
0.1904	0.6	0.581758422	0.576429405
0.238	0.538571429	0.533278554	0.526254376
0.2856	0.506428571	0.490550534	0.486581097
0.3332	0.484285714	0.455217749	0.466744457
0.3808	0.466428571	0.431388661	0.44924154
0.4284	0.447142857	0.424815119	0.43873979
0.476	0.432857143	0.404272802	0.429404901
0.5236	0.425714286	0.392769104	0.422403734
0.5712	0.419285714	0.383730485	0.418903151
0.6188	0.41	0.38537387	0.414235706
0.6664	0.408571429	0.37304848	0.414235706
0.714	0.406428571	0.37304848	0.4084014
0.952	0.395	0.363188168	0.407234539
1.19	0.392142857	0.361544782	0.396732789

Table 15 - Raw and normalised fluorescence data for the ferric-4-LICAM CeuE titration at 240 nM.

RAW DATA			
Batch 2 - 170 nM			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	991	965	634
0.0476	831	809	525
0.0952	733	698	449
0.1428	651	639	405
0.1904	600	594	376
0.238	545	552	365
0.2856	525	539	351
0.3332	503	512	343
0.3808	490	505	341
0.476	465	488	331
0.5236	446	472	321
0.5712	439	473	323
0.6188	441	463	317
0.6664	433	448	315
0.714	437	457	311
0.952	419	433	302
1.19	398	410	297

NORMALISED			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	1	1	1
0.0476	0.838546922	0.838341969	0.82807571
0.0952	0.739656912	0.723316062	0.708201893
0.1428	0.65691221	0.662176166	0.638801262
0.1904	0.605449041	0.615544041	0.593059937
0.238	0.549949546	0.572020725	0.575709779
0.2856	0.529767911	0.558549223	0.55362776
0.3332	0.507568113	0.530569948	0.541009464
0.3808	0.49445005	0.523316062	0.53785489
0.476	0.469223007	0.505699482	0.522082019
0.5236	0.450050454	0.489119171	0.506309148
0.5712	0.442986882	0.49015544	0.509463722
0.6188	0.445005045	0.479792746	0.5
0.6664	0.436932392	0.464248705	0.496845426
0.714	0.440968718	0.47357513	0.490536278
0.952	0.422805247	0.448704663	0.476340694
1.19	0.401614531	0.424870466	0.468454259

Table 16 - Raw and normalised fluorescence data for the ferric-4-LICAM CeuE titration at 170 nM. *Note missing data point of Fe-4-LICAM concentration at 0.4284.

RAW DATA			
Batch 1 - 240 nM			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	1115	1181	813
0.0476	959	1039	750
0.0952	851	909	669
0.1428	742	795	591
0.1904	627	695	509
0.238	547	610	434
0.2856	490	534	386
0.3332	475	482	356
0.3808	463	467	339
0.4284	452	467	341
0.476	464	449	344
0.5236	458	450	339
0.5712	448	444	334
0.6188	464	439	328
0.6664	453	439	331
0.714	451	436	330
0.952	452	428	332
1.19	455	428	329

NORMALISED			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	1	1	1
0.0476	0.860089686	0.879762913	0.922509225
0.0952	0.7632287	0.769686706	0.822878229
0.1428	0.665470852	0.67315834	0.726937269
0.1904	0.562331839	0.588484335	0.626076261
0.238	0.49058296	0.516511431	0.533825338
0.2856	0.439461883	0.452159187	0.474784748
0.3332	0.426008969	0.408128704	0.437884379
0.3808	0.415246637	0.395427604	0.41697417
0.4284	0.405381166	0.395427604	0.419434194
0.476	0.416143498	0.380186283	0.423124231
0.5236	0.410762332	0.381033023	0.41697417
0.5712	0.401793722	0.375952583	0.410824108
0.6188	0.416143498	0.371718882	0.403444034
0.6664	0.406278027	0.371718882	0.407134071
0.714	0.404484305	0.369178662	0.405904059
0.952	0.405381166	0.362404742	0.408364084
1.19	0.408071749	0.362404742	0.404674047

Table 17 - Raw and normalised fluorescence data for the ferric-**17** CeuE titration at 240 nM.

RAW DATA			
Batch 2 - 170 nM			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	865	855	707
0.0476	751	735	643
0.0952	658	616	571
0.1428	561	535	501
0.1904	513	470	451
0.238	490	455	417
0.2856	488	435	400
0.3332	483	444	394
0.3808	478	438	384
0.4284	472	415	392
0.476	459	409	390
0.5236	466	417	384
0.5712	451	415	380
0.6188	455	412	375
0.6664	454	399	375
0.714	452	398	386
0.952	453	390	397
1.19	433	380	397

NORMALISED			
	Run 1	Run 2	Run 3
Conc Fe-4-LICAM/ uM	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)	INTEGRATION (305-380NM)
0	1	1	1
0.0476	0.868208092	0.859649123	0.909476662
0.0952	0.760693642	0.720467836	0.807637907
0.1428	0.648554913	0.625730994	0.708628006
0.1904	0.593063584	0.549707602	0.637906648
0.238	0.566473988	0.532163743	0.589816124
0.2856	0.56416185	0.50877193	0.565770863
0.3332	0.558381503	0.519298246	0.5572843
0.3808	0.552601156	0.512280702	0.543140028
0.4284	0.54566474	0.485380117	0.554455446
0.476	0.530635838	0.478362573	0.551626591
0.5236	0.538728324	0.487719298	0.543140028
0.5712	0.521387283	0.485380117	0.53748232
0.6188	0.526011561	0.481871345	0.530410184
0.6664	0.524855491	0.466666667	0.530410184
0.714	0.522543353	0.465497076	0.545968883
0.952	0.523699422	0.456140351	0.561527581
1.19	0.500578035	0.444444444	0.561527581

Table 18 - Raw and normalised fluorescence data for the ferric-**17** CeuE titration at 170 nM.

```
[task]
 data = equilibra
 task = fit

[mechanism]
 CeuE + FeLD <==> CeuE.FeLD : Kd dissoc.

[concentrations]
 CeuE = 0.24 ; Concentration of CeuE in cuvette in uM / approx 5.00 mgmL-1 3
uL

[constants]
 Kd = 0.02 ?

[Responses]
 CeuE = 4.16 ? ; Initial fluorescence divided by amount of CeuE i.e. (1/0.24)
 CeuE.FeLD = 1.70 ? ; End fluorescence divided by amount of FeLD.CeuE i.e.
(0.408/0.24)

[equilibria]
 variable FeLD
 file ./CeuE/FeLD/FULLDATA/Dan1/FeLD.txt

[output]
 directory ./CeuE/FeLD/FULLDATA/Dan1/Output

[end]
```

Figure 90 - Script used for Dynafit v3.

```
[task]
 data = equilibria
 task = fit

[mechanism]
 CeuE + FeLD <==> CeuE.FeLD : Kd dissoc

[concentrations]
 CeuE = 0.17 ; Concentration of CeuE in cuvette in uM / approx 5.00 mgmL-1 3
uL

[constants]
 Kd = 0.02 ? ; uM

[responses]
 CeuE = 5.88 ? , CeuE.FeLD = 3.3 ? ; Initial fluorescence divided by amount of

CeuE i.e. (1/0.17); End fluorescence divided by amount of FeLD.CeuE i.e.

(0.561/0.17)

[data]
 variable FeLD
 file ./CeuE/FeLD/FULLDATA/Adam4/FeLD.txt

[output]
 directory ./CeuE/FeLD/FULLDATA/Adam4/Output

[end]
```

Figure 91 - Script used for Dynafit v4.

Abbreviations

°	degrees
°C	degrees Celsius
Å	Angstroms
Amo	amonabactin
aro	aromatic residue
B. subtilis	Bacillus subtilis
Boc	*tert*-butyloxycarbonyl
Boc_2O	di-tert-butyl dicarbonate
c (prefix)	centi
C. jejuni	Camoylobacter jejuni
CD	circular dichroism
CDI	1,1'-carbonyldiimidazole
D_2O	deuterium oxide
d_6-DMSO	deuterated dimethyl sulfoxide
Da	Dalton
dATP	deoxyadenosine triphosphate
DCC	dicyclohexylcarbodiimide
DCU	dicyclohexylurea
DFO	desferrioxamine-B
DGE	diglucosylated enterobactin
DHBS	dihydroxybenzolyserine
DIPEA	*N*-*N*'-diisopropylethylamine
DLS	Diamond Light Source
DMAP	4-dimethylaminopyridine
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DOSY	diffusion ordered spectroscopy
E. coli	Escherichia coli
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
Ent	enterobactin
ESI	electrospray ionisation

EtOH	ethanol
FRAP	fluorescence recovery after photobleaching
Fur	ferric-iron uptake regulation
g	grams
gly	glycine
H. influenzae	Haemophilus influenzae
HBTU	*N*,*N*,*N'*,*N'*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate
His	histidine
HOBt	hydroxybenzotriazole
HPLC	high peformance liquid chromatgraphy
IPTG	isopropyl β-D-1-thiogalactopyranoside
K	Kelvin
L	ligand (complexation)
L	litre
LB	lysogeny broth
LC	liquid chromatrography
LICAM	linear Catechol Amide
LMCT	ligand to metal charge transfer
Lys	lysine
M	metal (complexation)
m	metre
M	molar
m (prefix)	milli
m/z	mass/charge
Maq-Br	2-bromomethylanthraquinone
MeOH	methanol
MIB	sodium malonate, imidazole and boric acid
Mr	relative molecular mass
MS	mass spectrometry
MW	molecular weight
n (prefix)	nano
NEt_3	triethylamine

nm	nanometer
NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
OD_{600}	optical density at 600 nm
P. aeruginosa	Pseudomonads aeruginosa
Pa	Pascal
PBP	periplasmic binding protein
PCR	polymerase chain reaction
PEG	polyethylene glycol
ppm	parts per million
RMS	root mean square
RNA	ribonucleic acid
rpm	revolutions per minute
s	second
SDS PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SGC	serine-glycine-catechol
TFA	trifluoroacetic acid
THP	tetrahydropyran
TIC	total ion count
TIRF	total interal reflection fluorescence
TLC	thin layer chromatography
TMS	trimethylsilyl
TrisHCl	tris(hydroxymethyl)aminomethane hydrochloride
Trp	tryptophan
Tyr	tyrosine
UV-visible	ultra violet-visible
V	potential Difference
λ_{max}	wavelength of maximum absorbance
μ (prefix)	micro

References

1 M. Miethke and M. A. Marahiel, *Microbiol. Mol. Biol. Rev.*, 2007, **71**, 413.

2 R. C. Hider and X. Kong, *Nat. Prod. Rep.*, 2010, **27**, 637.

3 Z. D. Liu and R. C. Hider, *Coord. Chem. Rev.*, 2002, **232**, 151.

4 D. Touati, *Arch. Biochem. Biophys.*, 2000, **373**, 1.

5 B. Halliwell and J. M. Gutteridge, *Biochem. J.*, 1984, **219**, 1.

6 K. N. Raymond, *Coord. Chem. Rev.*, 1990, **105**, 135.

7 K. N. Raymond, E. A. Dertz and S. S. Kim, *Proc. Natl. Acad. Sci. USA.*, 2003, **100**, 3584.

8 J. B. Neilands, *J. Biol. Chem.*, 1995, **270**, 26723.

9 K. D. Krewulak and H. J. Vogel, *BBA-Biomembranes.*, 2008, **1778**, 1781.

10 R. G. Pearson, *J. Am. Chem. Soc.*, 1963, **85**, 3533.

11 H. Drechsel and G. Jung, *J. Pept. Sci.*, 1998, **4**, 147.

12 G. Xiao, D. van der Helm, R. C. Hider and P. S. Dobbin, *J. Chem. Soc., Dalton Trans.*, 1992, 3265.

13 L. D. Loomis and K. N. Raymond, *Inorg. Chem.*, 1991, **30**, 906.

14 B. P. Hay, D. A. Dixon, R. Vargas, J. Garza and K. N. Raymond, *Inorg. Chem.*, 2001, **40**, 3922.

15 D. J. Ecker, L. D. Loomis, M. E. Cass and K. N. Raymond, *J. Am. Chem. Soc.*, 1988, **110**, 2457.

16 R. W. Hancock, K. Hantke and V. Braun, *Arch. Microbiol.*, 1977, **114**, 231.

17 K. Hantke, *FEMS Microbiol. Lett.*, 1990, **67**, 5.

18 J. Brandel, N. Humbert, M. Elhabiri, I. J. Schalk, G. L. A. Mislin and A.-M. Albrecht-Gary, *Dalton Trans.*, 2012, **41**, 2820.

19 A. Braud, M. Hannauer, G. L. A. Mislin and I. J. Schalk, *J. Bacteriol.*, 2009, **191**, 3517.

20 P. Visca, G. Colotti, L. Serino, D. Verzili, N. Orsi and E. Chiancone, *Appl. Environ. Microb.*, 1992, **58**, 2886.

21 J. B. Neilands, T. J. Erickson and W. H. Rastetter, *J. Biol. Chem.*, 1981, **256**, 3831.

22 P. Thulasiraman, S. M. C. Newton, J. Xu, K. N. Raymond, C. Mai, A. Hall, M. A. Montague and P. E. Klebba, *J. Bacteriol.*, 1998, **180**, 6689.

23 R. J. Abergel, A. M. Zawadzka, T. M. Hoette and K. N. Raymond, *J. Am. Chem. Soc.*, 2009, **131**, 12682.

24 W. H. Rastetter, T. J. Erickson and M. C. Venuti, *J. Org. Chem.*, 1981, **46**, 3579.

25 M. Brorson, T. Damhus and C. E. Schaffer, *Inorg. Chem.*, 1983, **22**, 1569.

26 S. J. Rodgers, C. W. Lee, C. Y. Ng and K. N. Raymond, *Inorg. Chem.*, 1987, **26**, 1622.

27 B. Tse and Y. Kishi, *J. Org. Chem.*, 1994, **59**, 7807.

28 F. L. Weitl and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 2728.

29 T. D. P. Stack, Z. Hou and K. N. Raymond, *J. Am. Chem. Soc.*, 1993, **115**, 6466.

30 R. C. Scarrow, D. J. Ecker, C. Ng, S. Liu and K. N. Raymond, *Inorg. Chem.*, 1991, **30**, 900.

31 W. H. Harris and K. N. Raymond, *J. Am. Chem. Soc.*, 1979, **101**, 6534.

32 S. Heidinger, V. Braun, V. L. Pecoraro and K. N. Raymond, *J. Bacteriol.*, 1983, **153**, 109.

33 B. F. Matzanke, D. J. Ecker, T. S. Yang, B. H. Huynh, G. Muller and K. N. Raymond, *J. Bacteriol.*, 1986, **167**, 674.

34 T. J. Brickman and M. A. McIntosh, *J. Biol. Chem.*, 1992, **267**, 12350.

35 A. Muller, A. J. Wilkinson, K. S. Wilson and A.-K. Duhme-Klair, *Angew. Chem. Int. Ed.*, 2006, **45**, 5132.

36 F. Peuckert, A. L. Ramos-Vega, M. Miethke, C. J. Schwörer, A. G. Albrecht, M. Oberthür and M. A. Marahiel, *Chem. Biol.*, 2011, **18**, 907.

37 K. N. Raymond, M. E. Cass and S. L. Evans, *Pure Appl. Chem.*, 1987, **59**, 771.

38 S. K. Buchanan, B. S. Smith, L. Venkatramani, D. Xia, L. Esser, M. Palnikar, R. Chakraborty, D. Van der Helm and J. Deisenhofer, *Nat. Struct. Mol. Biol.*, 1999, **6**, 56.

39 P. D. Pawelek, N. Croteau, C. Ng-Thow-Hing, C. M. Khursigara, N. Moiseeva, M. Allaire and J. W. Coulton, *Science*, 2006, **312**, 1399.

40 A. Sauter and V. Braun, *J. Bacteriol.*, 2004, **186**, 5303.

41 A. D. Ferguson, J. W. Coulton, K. Diederichs, W. Welte, V. Braun and H.-P. Fiedler, *Protein Sci.*, 2000, **9**, 956.

42 N. Noinaj, M. Guillier, Barnard, Travis J. and S. K. Buchanan, *Annu. Rev. Microbiol.*, 2010, **64**, 43.

43 A. D. Ferguson, E. Hofmann, J. W. Coulton, K. Diederichs and W. Welte, *Science*, 1998, **282**, 2215.

44 D. Cobessi, H. Celia, N. Folschweiller, I. J. Schalk, M. A. Abdallah and F. Pattus, *J. Mol. Biol.*, 2005, **347**, 121.

45 D. Cobessi, H. Celia and F. Pattus, *J. Mol. Biol.*, 2005, **352**, 893.

46 T. E. Clarke, V. Braun, G. Winkelmann, L. W. Tari and H. J. Vogel, *J. Biol. Chem.*, 2002, **277**, 13966.

47 K. P. Locher, A. T. Lee and D. C. Rees, *Science*, 2002, **296**, 1091.

48 R. N. Hvorup, B. A. Goetz, M. Niederer, K. Hollenstein, E. Perozo and K. P. Locher, *Science*, 2007, **317**, 1387.

49 V. M. Korkhov, S. A. Mireku, R. N. Hvorup and K. P. Locher, *FEBS Lett.*, 2012, **586**, 972.

50 S. Gerber, M. Comellas-Bigler, B. A. Goetz and K. P. Locher, *Science*, 2008, **321**, 246.

51 K. Hollenstein, D. C. Frei and K. P. Locher, *Nature*, 2007, **446**, 213.

52 M. Miethke, O. Klotz, U. Linne, J. J. May, C. L. Beckering and M. A. Marahiel, *Mol. Microbiol.*, 2006, **61**, 1413.

53 J. Ollinger, K.-B. Song, H. Antelmann, M. Hecker and J. D. Helmann, *J. Bacteriol.*, 2006, **188**, 3664.

54 F. Peuckert, M. Miethke, A. Albrecht, L. O. Essen and M. Marahiel, *Angew. Chem. Int. Ed.*, 2009, **48**, 7924.

55 J. Greenwald, F. Hoegy, M. Nader, L. Journet, G. L. A. Mislin, P. L. Graumann and I. J. Schalk, *J. Biol. Chem.*, 2007, **282**, 2987.

56 J. Harrington and A. Crumbliss, *BioMetals*, 2009, **22**, 679.

57 J. S. Lodge, C. G. Gaines, J. E. L. Arceneaux and B. R. Byers, *Biochem. Biophys. Res. Commun.*, 1980, **97**, 1291.

58 S. M. Cohen, M. Meyer and K. N. Raymond, *J. Am. Chem. Soc.*, 1998, **120**, 6277.

59 K. T. Greenwood and R. K. J. Luke, *BBA-Enzymology*, 1978, **525**, 209.

60 I. Schalk and L. Guillon, *Amino Acids*, 2013, **44**, 1267.

61 K. Hantke, *Curr. Opin. Microbiol.*, 2001, **4**, 172.

62 J.-W. Lee and J. Helmann, *BioMetals*, 2007, **20**, 485.

63 L. Escolar, J. Pérez-Martín and V. de Lorenzo, *J. Bacteriol.*, 1999, **181**, 6223.

64 M. D. Hunt, G. S. Pettis and M. A. McIntosh, *J. Bacteriol.*, 1994, **176**, 3944.

65 S. Dubrac and D. Touati, *J. Bacteriol.*, 2000, **182**, 3802.

66 Y. Tang and J. R. Guest, *Microbiology*, 1999, **145**, 3069.

67 E. W. Althaus, C. E. Outten, K. E. Olson, H. Cao and T. V. O'Halloran, *Biochemistry*, 1999, **38**, 6559.

68 C. L. Pickett, T. Auffenberg, E. C. Pesci, V. L. Sheen and S. S. Jusuf, *Infect. Immun.*, 1992, **60**, 3872.

69 World Health Organisation, http://www.who.int/mediacentre/factsheets/fs255/en/, Accessed 08/08/2014.

70 K. Palyada, D. Threadgill and A. Stintzi, *J. Bacteriol.*, 2004, **186**, 4714.

71 H. Naikare, K. Palyada, R. Panciera, D. Marlow and A. Stintzi, *Infect. Immun.*, 2006, **74**, 5433.

72 X. Zeng, Y. Mo, F. Xu and J. Lin, *Mol. Microbiol.*, 2013, **87**, 594.

73 C. E. Miller, P. H. Williams and J. M. Ketley, *Microbiology*, 2009, **155**, 3157.

74 L. H. Field, V. L. Headley, S. M. Payne and L. J. Berry, *Infect. Immun.*, 1986, **54**, 126.

75 D. L. Caulder and K. N. Raymond, *J. Chem. Soc., Dalton Trans.*, 1999, 1185.

76 M. Albrecht, *Chem. Rev.*, 2001, **101**, 3457.

77 C. Piguet, G. Bernardinelli and G. Hopfgartner, *Chem. Rev.*, 1997, **97**, 2005.

78 D. L. Caulder and K. N. Raymond, *Acc. Chem. Res.*, 1999, **32**, 975.

79 E. J. Enemark and T. D. P. Stack, *Angew. Chem. Int. Ed.*, 1995, **34**, 996.

80 M. Albrecht and S. Kotila, *Angew. Chem. Int. Ed.*, 1995, **34**, 2134.

81 B. Kersting, M. Meyer, R. E. Powers and K. N. Raymond, *J. Am. Chem. Soc.*, 1996, **118**, 7221.

82 L. J. Charbonniere, G. Bernardinelli, C. Piguet, A. M. Sargeson and A. F. Williams, *J. Chem. Soc., Chem. Commun.*, 1994, 1419.

83 M. Albrecht and S. Kotila, *Angew. Chem. Int. Ed.*, 1996, **35**, 1208.

84 D. L. Caulder and K. N. Raymond, *Angew. Chem. Int. Ed.*, 1997, **36**, 1440.

85 E. J. Enemark and T. D. P. Stack, *Inorg. Chem.*, 1996, **35**, 2719.

86 M. T. Caudle, R. D. Stevens and A. L. Crumbliss, *Inorg. Chem.*, 1994, **33**, 6111.

87 Z. Hou, C. J. Sunderland, T. Nishio and K. N. Raymond, *J. Am. Chem. Soc.*, 1996, **118**, 5148.

88 S.-P. Huang, K. J. Franz, M. M. Olmstead and R. H. Fish, *Inorg. Chem.*, 1995, **34**, 2820.

89 A.-K. Duhme, Z. Dauter, R. C. Hider and S. Pohl, *Inorg. Chem.*, 1996, **35**, 3059.

90 A.-K. Duhme, *J. Chem. Soc., Dalton Trans.*, 1997, 773.

91 M. Albrecht, I. Janser, S. Kamptmann, P. Weis, B. Wibbeling and R. Frohlich, *Dalton Trans.*, 2004, 37.

92 M. Albrecht, I. Janser and R. Frohlich, *Chem. Commun.*, 2005, 157.

93 M. D. Pluth, R. G. Bergman and K. N. Raymond, *Science*, 2007, **316**, 85.

94 Y. E. Alexeev, B. I. Kharisov, T. C. H. García and A. D. Garnovskii, *Coord. Chem. Rev.*, 2010, **254**, 794.

95 M. Albrecht, *Chem. Soc. Rev.*, 1998, **27**, 281.

96 M. J. Hannon and L. J. Childs, *Supramol. Chem.*, 2004, **16**, 7.

97 M. C. Venuti, W. H. Rastetter and J. B. Neilands, *J. Med. Chem.*, 1979, **22**, 123.

98 W. R. Harris, F. L. Weitl and K. N. Raymond, *J. Chem. Soc., Chem. Commun.*, 1979, 177.

99 A. Pietrangelo, *Gut*, 2003, **52**, ii23.

100 C. Hershko, A. M. Konijn and G. Link, *Br. J. Haematol.*, 1998, **101**, 399.

101 F. L. Weitl and K. N. Raymond, *J. Am. Chem. Soc.*, 1980, **102**, 2289.

102 S. Sheth, *Curr. Opin. Hematol.*, 2014, **21**, 179.

103 J. Van Cutsem and J. R. Boelaert, *Kidney Int*, 1989, **36**, 1061.

104 M. J. Miller, *Chem. Rev.*, 1989, **89**, 1563.

105 G. L. A. Mislin and I. J. Schalk, *Metallomics*, 2014, **6**, 408.

106 U. Möllmann, L. Heinisch, A. Bauernfeind, T. Köhler and D. Ankel-Fuchs, *BioMetals*, 2009, **22**, 615.

107 V. Braun, A. Pramanik, T. Gwinner, M. Köberle and E. Bohn, *BioMetals*, 2009, **22**, 3.

108 A. L. Stefanska, M. Fulston, C. S. V. Houge-Frydrych, J. J. Jones and S. R. Warr, *J. Antibiot.*, 2000, **53**, 1346.

109 S. R. Md-Saleh, E. C. Chilvers, K. G. Kerr, S. J. Milner, A. M. Snelling, J. P. Weber, G. H. Thomas, A.-K. Duhme-Klair and A. Routledge, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 1496.

110 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.

111 T. Zheng, J. L. Bullock and E. M. Nolan, *J. Am. Chem. Soc.*, 2012, **134**, 18388.

112 T. Zheng and E. M. Nolan, *J. Am. Chem. Soc.*, 2014, **136**, 9677.

113 M. Ghosh and M. J. Miller, *Biorg. Med. Chem.*, 1996, **4**, 43.

114 A. Ghosh, M. Ghosh, C. Niu, F. Malouin, U. Moellmann and M. J. Miller, *Chem. Biol.*, 1996, **3**, 1011.

115 A. A. Minnick, J. A. McKee, E. K. Dolence and M. J. Miller, *Antimicrob. Agents Chemother.*, 1992, **36**, 840.

116 K. Hantke, G. Nicholson, W. Rabsch and G. Winkelmann, *Proc. Natl. Acad. Sci. USA.*, 2003, **100**, 3677.

117 M. A. Holmes, W. Paulsene, X. Jide, C. Ratledge and R. K. Strong, *Structure*, 2005, **13**, 29.

118 A. K. Sia, B. E. Allred and K. N. Raymond, *Curr. Opin. Chem. Biol.*, 2013, **17**, 150.

119 D. H. Goetz, M. A. Holmes, N. Borregaard, M. E. Bluhm, K. N. Raymond and R. K. Strong, *Mol. Cell.*, 2002, **10**, 1033.

120 S. Noël, L. Guillon, I. J. Schalk and G. t. L. A. Mislin, *Org. Lett.*, 2011, **13**, 844.

121 C. Baysse, D. De Vos, Y. Naudet, A. Vandermonde, U. Ochsner, J.-M. Meyer, H. Budzikiewicz, M. Schäfer, R. Fuchs and P. Cornelis, *Microbiology*, 2000, **146**, 2425.

122 T. Zheng and E. M. Nolan, *Metallomics*, 2012, **4**, 866.

123 H. Boukhalfa, S. D. Reilly, R. Michalczyk, S. Iyer and M. P. Neu, *Inorg. Chem.*, 2006, **45**, 5607.

124 M. Yoder and W. Kisaalita, *J. Biol. Eng.*, 2011, **5**, 4.

125 G. Renard, M. Mureseanu, A. Galarneau, D. A. Lerner and D. Brunel, *New J. Chem.*, 2005, **29**, 912.

126 J. M. Barrero, C. Camara, M. C. Perez-Conde, C. S. Jose and L. Fernandez, *Analyst*, 1995, **120**, 431.

127 K. K. Bhargava, R. W. Grady and A. Cerami, *J. Pharm. Sci.*, 1980, **69**, 986.

128 A.-K. Duhme, *Z. Anorg. Allg. Chem.*, 1998, **624**, 1922.

129 R. H. Fish and R. S. Tannous, *Organometallics*, 1982, **1**, 1238.

130 P. M. Loiseau, M. Trabelsi, P. Gayral and J. G. Wolf, *Acta Trop.*, 1995, **59**, 237.

131 K. J. Winstanley and D. K. Smith, *J. Org. Chem.*, 2007, **72**, 2803.

132 B. Lindgren, O. and T. Nilsson, *Acta. Chem. Scand.*, 1973, **27**, 888.

133 A. Chimiak and J. B. Neilands, *Struct. Bond.*, 1984, **58**, 89.

134 J. Polster and H. Lachmann, in *Spectrometric Titrations: Analysis of Chemical Equilibria*, VCH, 1989, pp. 256-269.

135 W. R. Harris, K. N. Raymond and F. L. Weitl, *J. Am. Chem. Soc.*, 1981, **103**, 2667.

136 T. B. Karpishin, M. S. Gebhard, E. I. Solomon and K. N. Raymond, *J. Am. Chem. Soc.*, 1991, **113**, 2977.

137 R. C. Scarrow, D. L. White and K. N. Raymond, *J. Am. Chem. Soc.*, 1985, **107**, 6540.

138 A. N. Pham, A. L. Rose, A. J. Feitz and T. D. Waite, *Geochim. Cosmochim. Acta*, 2006, **70**, 640.

139 Y. Ma, Y. Xie and R. C. Hider, *Analyst*, 2013, **138**, 96.

140 G. Anderegg, *Pure Appl. Chem.*, 1982, **54**, 2693.

141 A. Avdeef, S. R. Sofen, T. L. Bregante and K. N. Raymond, *J. Am. Chem. Soc.*, 1978, **100**, 5362.

142 W. Przychodzeń and A. Chimiak, *Phosphorus, Sulfur Silicon Relat. Elem.*, 1998, **143**, 77.

143 D. S. Kemp, S. W. Wang, R. C. Mollan, S. L. Hsia and P. N. Confalone, *Tetrahedron*, 1974, **30**, 3677.

144 Y. Murakami and K. Nakamura, *Bull. Chem. Soc. Jpn.*, 1963, **36**, 1408.

145 R. C. Scarrow, P. E. Riley, K. Abu-Dari, D. L. White and K. N. Raymond, *Inorg. Chem.*, 1985, **24**, 954.

146 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.

147 Y. Tor, J. Libman, A. Shanzer, C. E. Felder and S. Lifson, *J. Am. Chem. Soc.*, 1992, **114**, 6661.

148 M. Llinas, D. M. Wilson and J. B. Neilands, *Biochemistry*, 1973, **12**, 3836.

149 C. S. Johnson Jr, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1999, **34**, 203.

150 M. Nilsson, *J. Magn. Reson.*, 2009, **200**, 296.

151 M. Cappelezzo, C. A. Capellari, S. H. Pezzin and L. A. F. Coelho, *J. Chem. Phys.*, 2007, **126**, 224516.

152 S. A. Schichman and R. L. Amey, *J. Phys. Chem.*, 1971, **75**, 98.

153 M. Hanwell, D. Curtis, D. Lonie, T. Vandermeersch, E. Zurek and G. Hutchison, *J. Cheminform.*, 2012, **4**, 17.

154 E. L. Cussler, in *Diffusion Mass Transfer in Fluid Systems*, Cambridge University Press, 2009, pp. 183-190.

155 A. Muller, in *Department of Chemistry*, University of York, 2006.

156 T. U. Consortium, *Nucleic Acids Res.*, 2012, **40**, D71.

157 T. N. Petersen, S. Brunak, G. von Heijne and H. Nielsen, *Nat. Methods*, 2011, **8**, 785.

158 H.-B. Shen and K.-C. Chou, *Biochem. Biophys. Res. Commun.*, 2007, **363**, 297.

159 A. S. Juncker, H. Willenbrock, G. von Heijne, S. Brunak, H. Nielsen and A. Krogh, *Protein Sci.*, 2003, **12**, 1652.

160 P. Gouet, X. Robert and E. Courcelle, *Nucleic Acids Res.*, 2003, **31**, 3320.

161 A. Vagin and A. Teplyakov, *J. Appl. Crystallogr.*, 1997, **30**, 1022.

162 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta. Crystallogr. D.*, 1997, **53**, 240.

163 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta. Crystallogr. D.*, 2010, **66**, 486.

164 R. A. Nicholls, F. Long and G. N. Murshudov, *Acta. Crystallogr. D.*, 2012, **68**, 404.

165 N. Li, C. Zhang, B. Li, X. Liu, Y. Huang, S. Xu and L. Gu, *J. Biol. Chem.*, 2012, **287**, 8912.

166 K. Hsin, Y. Sheng, M. M. Harding, P. Taylor and M. D. Walkinshaw, *J. Appl. Crystallogr.*, 2008, **41**, 963.

167 S. A. L. Tom-Yew, D. T. Cui, E. G. Bekker and M. E. P. Murphy, *J. Biol. Chem.*, 2005, **280**, 9283.

168 H. K. Khambati, T. F. Moraes, J. Singh, S. R. Shouldice, R. h. Yu and A. B. Schryvers, *Biochem. J*, 2010, **432**, 57.

169 C. Notredame, D. G. Higgins and J. Heringa, *J. Mol. Biol.*, 2000, **302**, 205.

170 S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J. Mol. Biol.*, 1990, **215**, 403.

171 X. Liu, Q. Du, Z. Wang, S. Liu, N. Li, Y. Chen, C. Zhu, D. Zhu, T. Wei, Y. Huang, S. Xu and L. Gu, *FEBS Lett.*, 2012, **586**, 1240.

172 A. M. Zawadzka, Y. Kim, N. Maltseva, R. Nichiporuk, Y. Fan, A. Joachimiak and K. N. Raymond, *Proc. Natl. Acad. Sci. USA.*, 2009, **106**, 21854.

173 F. Peuckert, M. Miethke, A. G. Albrecht, L.-O. Essen and M. A. Marahiel, *Angew. Chem. Int. Ed.*, 2009, **48**, 7924.

174 T. B. Karpishin, T. D. P. Stack and K. N. Raymond, *J. Am. Chem. Soc.*, 1993, **115**, 6115.

175 T. D. P. Stack, T. B. Karpishin and K. N. Raymond, *J. Am. Chem. Soc.*, 1992, **114**, 1512.

176 M. Miethke and A. Skerra, *Antimicrob. Agents Chemother.*, 2010, **54**, 1580.

177 C. Sprencel, Z. Cao, Z. Qi, D. C. Scott, M. A. Montague, N. Ivanoff, J. Xu, K. M. Raymond, S. M. C. Newton and P. E. Klebba, *J. Bacteriol.*, 2000, **182**, 5359.

178 E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, in *The Proteomics Protocols Handbook*, Humana Press, 2005, pp. 571-607.

179 P. Kuzmič, *Anal. Biochem.*, 1996, **237**, 260.

180 L. Lyons, in *A Practical Guide to Data Analysis for Physical Science Students*, Cambridge University Press, 1991, pp. 31-33.

181 A. M. Gehring, K. A. Bradley and C. T. Walsh, *Biochemistry*, 1997, **36**, 8495.

182 C. A. Shaw-Reid, N. L. Kelleher, H. C. Losey, A. M. Gehring, C. Berg and C. T. Walsh, *Chem. Biol.*, 1999, **6**, 385.

183 I. G. O'Brien and F. Gibson, *BBA-Gen. Subjects.*, 1970, **215**, 393.

184 I. Berner, M. Greiner, J. Metzger, G. Jung and G. Winkelmann, *Biol. Met.*, 1991, **4**, 113.

185 G. Winkelmann, A. Cansier, W. Beck and G. Jung, *Biometals*, 1994, **7**, 149.

186 C. P. Bergstrom, M. C. Lu and C. L. Bell, *J. Nat. Prod.*, 1991, **54**, 1003.

187 R. E. Watts, M. Totsika, V. L. Challinor, A. N. Mabbett, G. C. Ulett, J. J. De Voss and M. A. Schembri, *Infect. Immun.*, 2012, **80**, 333.

188 L. L. Guan, K. Kanoh and K. Kamino, *Appl. Environ. Microb.*, 2001, **67**, 1710.

189 Y. Igarashi, T. Iida, T. Fukuda, S. Miyanaga, H. Sakurai, I. Saiki and K. Miyanouchi, *J. Antibiot.*, 2012, **65**, 207.

190 E. Strahsburger, M. Baeza, O. Monasterio and R. Lagos, *Antimicrob. Agents Chemother.*, 2005, **49**, 3083.

191 M. A. Fischbach, H. Lin, L. Zhou, Y. Yu, R. J. Abergel, D. R. Liu, K. N. Raymond, B. L. Wanner, R. K. Strong, C. T. Walsh, A. Aderem and K. D. Smith, *Proc. Natl. Acad. Sci. USA.*, 2006, **103**, 16502.

192 M. Caza, F. Lépine, S. Milot and C. M. Dozois, *Infect. Immun.*, 2008, **76**, 3539.

193 E. Valeur and M. Bradley, *Chem. Soc. Rev.*, 2009, **38**, 606.

194 C. A. G. N. Montalbetti and V. Falque, *Tetrahedron*, 2005, **61**, 10827.

195 G. Windridge and E. C. Jorgensen, *J. Am. Chem. Soc.*, 1971, **93**, 6318.

196 T. Mukaiyama, M. Usui, E. Shimada and K. Saigo, *Chem. Lett.*, 1975, **4**, 1045.

197 K. Saigo, M. Usui, K. Kikuchi, E. Shimada and T. Mukaiyama, *Bull. Chem. Soc. Jpn.*, 1977, **50**, 1863.

198 O. Mitsunobu and M. Yamada, *Bull. Chem. Soc. Jpn.*, 1967, **40**, 2380.

199 D. L. Hughes, R. A. Reamer, J. J. Bergan and E. J. J. Grabowski, *J. Am. Chem. Soc.*, 1988, **110**, 6487.

200 J. A. Sowinski and P. L. Toogood, *J. Org. Chem.*, 1996, **61**, 7671.

201 X. Yu, Y. Dai, T. Yang, M. R. Gagné and H. Gong, *Tetrahedron*, 2011, **67**, 144.

202 B. F. Lundt, N. L. Johansen, A. Vølund and J. Markussen, *Int. J. Pept. Protein Res.*, 1978, **12**, 258.

203 R. Schobert, A. Stangl and K. Hannemann, *Tetrahedron*, 2006, **62**, 7799.

204 J. R. Telford and K. N. Raymond, *Inorg. Chem.*, 1998, **37**, 4578.

205 J. R. Telford, J. A. Leary, L. M. G. Tunstad, B. R. Byers and K. N. Raymond, *J. Am. Chem. Soc.*, 1994, **116**, 4499.

206 J. R. Telford and K. N. Raymond, *J. Biol. Inorg. Chem.*, 1997, **2**, 750.

207 S. S. Isied, G. Kuo and K. N. Raymond, *J. Am. Chem. Soc.*, 1976, **98**, 1763.

208 T. B. Karpishin, T. M. Dewey and K. N. Raymond, *J. Am. Chem. Soc.*, 1993, **115**, 1842.

209 T. B. Karpishin and K. N. Raymond, *Angew. Chem. Int. Ed.*, 1992, **31**, 466.

210 J. V. McArdle, S. R. Sofen, S. R. Cooper and K. N. Raymond, *Inorg. Chem.*, 1978, **17**, 3075.

211 B. E. Allred, C. Correnti, M. C. Clifton, R. K. Strong and K. N. Raymond, *ACS Chem. Biol.*, 2013, **8**, 1882.

212 C. Correnti, Matthew C. Clifton, Rebecca J. Abergel, B. Allred, Trisha M. Hoette, M. Ruiz, R. Cancedda, Kenneth N. Raymond, F. Descalzi and Roland K. Strong, *Structure*, 2011, **19**, 1796.

213 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.

214 N. Coudevylle, M. Hoetzinger, L. Geist, G. Kontaxis, M. Hartl, K. Bister and R. Konrat, *Biochemistry*, 2011, **50**, 9192.

215 W. P. Griffith, H. I. S. Nogueira, B. C. Parkin, R. N. Sheppard, A. J. P. White and D. J. Williams, *J. Chem. Soc., Dalton Trans.*, 1995, 1775.

216 H. Budzikiewicz, *Curr. Top. Med. Chem.*, 2001, **1**, 73.

217 T. Moniz, A. Nunes, A. M. G. Silva, C. Queirós, G. Ivanova, M. S. Gomes and M. Rangel, *J. Inorg. Biochem.*, 2013, **121**, 156.

218 A. Nunes, M. Podinovskaia, A. Leite, P. Gameiro, T. Zhou, Y. Ma, X. Kong, U. Schaible, R. Hider and M. Rangel, *J. Biol. Inorg. Chem.*, 2010, **15**, 861.

219 T. Zhou, Y. Ma, X. Kong and R. C. Hider, *Dalton Trans.*, 2012, **41**, 6371.

220 K. J. Franz, *Curr. Opin. Chem. Biol.*, 2013, **17**, 143.

221 D. Imbert, F. Thomas, P. Baret, G. Serratrice, D. Gaude, J.-L. Pierre and J.-P. Laulhere, *New J. Chem.*, 2000, **24**, 281.

222 M. Apostol, P. Baret, G. Serratrice, J. Desbrières, J.-L. Putaux, M.-J. Stébé, D. Expert and J.-L. Pierre, *Angew. Chem. Int. Ed.*, 2005, **44**, 2580.

223 J. S. Martinez, G. P. Zhang, P. D. Holt, H.-T. Jung, C. J. Carrano, M. G. Haygood and A. Butler, *Science*, 2000, **287**, 1245.

224 T. Zhou, H. Neubert, D. Y. Liu, Z. D. Liu, Y. M. Ma, X. L. Kong, W. Luo, S. Mark and R. C. Hider, *J. Med. Chem.*, 2006, **49**, 4171.

225 F. Zeng and S. C. Zimmerman, *Chem. Rev.*, 1997, **97**, 1681.

226 G. R. Newkome, E. He and C. N. Moorefield, *Chem. Rev.*, 1999, **99**, 1689.

227 D. Astruc and F. Chardac, *Chem. Rev.*, 2001, **101**, 2991.

228 G. M. Dykes, *J. Chem. Technol. Biot.*, 2001, **76**, 903.

229 M. A. Mintzer and M. W. Grinstaff, *Chem. Soc. Rev.*, 2011, **40**, 173.

230 D. K. Smith and F. Diederich, *Chem. Eur. J.*, 1998, **4**, 1353.

231 C. Ji, P. A. Miller and M. J. Miller, *J. Am. Chem. Soc.*, 2012, **134**, 9898.

232 T. Inomata, H. Tanabashi, Y. Funahashi, T. Ozawa and H. Masuda, *Dalton Trans.*, 2013, **42**, 16043.

233 S. Bräse, C. Gil, K. Knepper and V. Zimmermann, *Angew. Chem. Int. Ed.*, 2005, **44**, 5188.

234 X. Zhao and K. S. Schanze, *Chem. Commun.*, 2010, **46**, 6075.

235 J. M. Anderson, J. D. Baldeschwieler, D. C. Dittmer and W. D. Phillips, *J. Chem. Phys.*, 1963, **38**, 1260.

236 E. F. V. Scriven and K. Turnbull, *Chem. Rev.*, 1988, **88**, 297.

237 G. R. Newkome, R. K. Behera, C. N. Moorefield and G. R. Baker, *J. Org. Chem.*, 1991, **56**, 7162.

238 X. Dominguez, I. Lopez and R. Franco, *J. Org. Chem.*, 1961, **26**, 1625.

239 S. Lebreton, S.-E. How, M. Buchholz, B.-E. Yingyongnarongkul and M. Bradley, *Tetrahedron*, 2003, **59**, 3945.

240 V. Theodorou, K. Skobridis, A. G. Tzakos and V. Ragoussis, *Tetrahedron Lett.*, 2007, **48**, 8230.

241 H. Heli, S. Mirtorabi and K. Karimian, *Expert. Opin. Ther. Pat.*, 2011, **21**, 819.

242 J. W. Coulton, H.-U. Naegeli and V. Braun, *Eur. J. Biochem.*, 1979, **99**, 39.

243 C. J. Carrano and K. N. Raymond, *J. Bacteriol.*, 1978, **136**, 69.

244 M. Llinás, M. P. Klein and J. B. Neilands, *Int. J. Pept. Protein Res.*, 1972, **4**, 157.

245 D. Van der Helm, J. R. Baker, D. L. Eng-Wilmot, M. B. Hossain and R. A. Loghry, *J. Am. Chem. Soc.*, 1980, **102**, 4224.

246 J. M. Kwiatek, D. M. Owen, A. Abu-Siniyeh, P. Yan, L. M. Loew and K. Gaus, *PLoS ONE*, 2013, **8**, e52960.

247 K. Simons and M. J. Gerl, *Nat. Rev. Mol. Cell. Biol.*, 2010, **11**, 688.

248 D. Lingwood and K. Simons, *Science*, 2010, **327**, 46.

249 M. F. Hanzal-Bayer and J. F. Hancock, *FEBS Lett.*, 2007, **581**, 2098.

250 K. A. K. Tanaka, K. G. N. Suzuki, Y. M. Shirai, S. T. Shibutani, M. S. H. Miyahara, H. Tsuboi, M. Yahara, A. Yoshimura, S. Mayor, T. K. Fujiwara and A. Kusumi, *Nat Meth*, 2010, **7**, 865.

251 D. Axelrod, T. P. Burghardt and N. L. Thompson, *Annu. Rev. Biophys. Bio.*, 1984, **13**, 247.

252 D. Axelrod, *J. Cell. Biol.*, 1981, **89**, 141.

253 Life Technologies, https://www.lifetechnologies.com/order/catalog/product/A10266/, Accessed 01/07/2014.

254 P. A. Ledin, F. Friscourt, J. Guo and G.-J. Boons, *Chem. Eur. J.*, 2011, **17**, 839.

255 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2002, **41**, 2596.

256 J. E. Moses and A. D. Moorhouse, *Chem. Soc. Rev.*, 2007, **36**, 1249.

257 C. A. Schneider, W. S. Rasband and K. W. Eliceiri, *Nat Meth*, 2012, **9**, 671.

258 Collaborative Computational Project Number 4, *Acta. Crystallogr. D.*, 1994, **50**, 760.