# Investigation of Rieske Non-Heme Iron-Dependent Oxygenase-Catalyzed Oxidative Carbocyclization Reactions in Prodiginine Alkaloid Biosynthesis

**Christopher James Perry**

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry

**University of Warwick**

**Department of Chemistry**

**September 2019**

# List of contents

## List of figures

**Figure 2.10**: (*left*) Agarose gels highlighting DNA bands resulting from screening pCAP1000-*mcp*BGC for the presence of the entire gene cluster by PCR; M = molecular size marker; lanes 1-4 = DNA bands resulting from using *S. longispororuber* gDNA; lanes 5-8 = DNA bands from using pCAP1000-*mcp*BGC plasmid DNA; (*right*) Agarose gel highlighting PCR products arising from primers designed to screen for four sections of the *mcp* cluster from DNA extracted from *E. coli* ET12567 transconjugant; M = molecular size marker; lanes 1-4 = PCR products from using *S. longispororuber* gDNA as template DNA; lanes 5-8 = PCR products using pCAP1000-*mcp*BGC as template DNA.

**Figure 2.11:** Images of *S. albus* transconjugants patched onto solid SFM media and incubated at 30 $^{o}$C for 2 weeks; **A.** Under-plate view of prodiginine producing *S. albus* transconjugant; **B.** Front view showing two transconjugants producing a prodiginine metabolite (pink) and three non-producing (white); **C.** An image of a prodiginine producing *S. albus* transconjugant (pink) and a non-producing transconjugant (white).

**Figure 2.12:** EICs at *m/z* = 392.27 ±0.01; 394.29 ±0.01 corresponding to the calculated $[M+H]^+$ of metacycloprodigisoin (**25**) and undecylprodigiosin (**23**) respectively from LC-MS analysis of the mycelial extract of; **A.** *S. albus* WT; **B.** *S. albus*/pCAP1000-*mcp*BGC, incubated at 30 $^{o}$C on solid SFM media for two weeks.

**Figure 2.13:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigisoin (**25**) respectively from LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-*mcp*BGC grown on solid SFM membrane (black trace); *S. albus*/pCAP1000-*mcp*BGC grown on solid R5 media overlaid with a permeable membrane (blue trace:); *S. albus*/pCAP1000-*mcp*BGC grown on solid SFM media overlaid with a semi-permeable membrane (red trace); *S. albus*/pCAP1000-*mcp*BGC grown on solid R5 media (green trace).

**Figure 2.14:** Schematic outlining the procedure for generating an in-frame deletion using TAR cloning in yeast; highlighting the digestion of the vector and co-transformation of yeast with a PCR product consisting of DNA that is homologous to sequences that flank the gene to be deleted.

**Figure 2.15: A.** Agarose gel highlighting PCR products corresponding to the segments of *S. longispororuber* gDNA that flank the *mcpLK* gene; M = molecular size marker; lane 1 = lane 2 = left flanking region; lane 3 = lane 4 = right flanking region; **B.** Agarose gel highlighting overlap PCR product resulting from combining both PCR products from gel **A**.

**Figure 2.16:** Agarose gel highlighting 603 bp PCR products from reactions using screening primers and the extracted DNA from six different yeast transconjugants.

**Figure 2.17:** (*left*) EICs at *m/z* = 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of metacycloprodigisoin (**25**) from the LC-MS of the mycelial extract of **A.** *S. albus*/pCAP1000-*mcp*BGC; and **B.** *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* (*right*) Images of the mycelia of

*S. albus*/pCAP1000-*mcp*BGC **C.** and *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* **D**; grown on solid SFM media at 30 $^{o}$C for two weeks.

**Figure 2.18:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (blue) (**23**) and metacycloprodigisoin (**25**) (red) respectively from the LC-MS analysis of the mycelial extract of **A.** *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* fed with synthetic 2-UP (**28**); **B.** *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*.

**Figure 2.19:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) respectively from LC-MS analysis of the mycelial extracts of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* fed with 4 mg synthetic 2-UP (**28**) (green); 1 mg synthetic 2-UP (**28**) (red) and 2 mg synthetic 2-UP (**28**) (blue).

**Figure 2.20:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigisoin (**25**) respectively from the LC-MS analysis of the mycelial extract after; Incubation for two days, post-feeding with synthetic 2-UP (**28**) (red); Incubation for one day, post-feeding with synthetic 2-UP (**28**) (blue).

**Figure 2.21:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) respectively from the LC-MS analysis of the mycelial extract of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) (black); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) + 0.25 mg MBC (**21**) (green); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) + 0.50 mg MBC (**21**) (blue); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) + 1 mg MBC (**21**) (red).

**Figure 2.22:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) respectively from the LC-MS analysis of the mycelial extract of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) (blue); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) + 1 mg MBC (**21**) (red); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* no feeding (black).

**Figure 2.23:** (*Top*) EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) (blue) and metacycloprodigiosin (**25**) (red) respectively from the LC-MS analysis of the mycelial extract of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) + 1mg MBC (**21**) **A**; *S. longispororuber* WT **B**; (*bottom*) MS/MS spectra of the $[M+H]^+$ metabolite at *m/z* = 392 corresponding to metacycloprodigiosin (**25**); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) + 1mg MBC (**21**) **A**; *S. longispororuber* WT **B**.

**Figure 2.24:** The proposed fragmentation pattern of metacycloprodigiosin (**25**).

**Figure 2.25:** A section of the $^{1}$H-NMR spectra of compound **121** and the undesired migratory product **134**.

**Figure 2.26:** Comparison of the $^{13}$C-NMR spectra of undecanol **(A)**, ***R*-117 (B)** and ***S*-117 (C)**.

**Figure 2.27**: Comparison of the $^{1}$H-NMR spectra ($CDCl_3$, 500 MHz) of the four diastereomers of compound **124**, highlighting a key upfield shift of the terminal methyl triplet in the ***R,R* -124** and ***S,S*-124** diastereomers.

**Figure 2.28:** (*Top*) The feeding of ***R*-120** and ***S*-120** to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* resulting in the corresponding undecylprodigiosin analogues ***R*-** and ***S*-125**, and the two possible outcomes for the product of McpG catalysis; (*bottom*) EICs at *m/z* = 395.29 and 393.28 ±0.01 (*top and middle*) corresponding to the calculated $[M+H]^+$ of deuterium-labeled undecylprodigiosin (**125**) (blue) and deuterium-labeled metacycloprodigisoin (**126**) (red) respectively from the LC-MS analysis of the mycelial extract resulting from feeding **A.** *pro-R* labelled 2-UP (***R*-120)**; **B.** *pro-S* labelled 2-UP (***S*-120**); (*bottom*) EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) (blue) and metacycloprodigisoin (**25**) (red) respectively of the mycelial extract from feeding **C.** 2-UP (**28**) (positive control).

**Figure 2.29:** A comparison of the HR-MS spectrum of the cyclic prodiginine compounds eluted at ~19.0 minutes from the LC-MS analysis of the mycelial extracts resulting from feeding *pro-R* labeled 2-UP (***R*-120) (A)**; *pro-S* labeled 2-UP (***S*-120) (B)**; and 2-UP (**28**) **(C)** to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG*

**Figure 2.30:** The mechanism of class B radical SAM methylases.

**Figure 2.31: A.** The oxidative carbocyclization reaction catalyzed by McpG, highlighting the proposed origin of the inversion of stereochemistry; **B.** The proposed reaction orientation of class B radical SAM methylases, highlighting the inversion of stereochemistry.

**Figure 3.1:** EICs from the LC-MS analysis of the mycelial extract resulting from feeding 2-UP (**28**) (or analogue) alongside MBC (**21**) to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **Blue trace:** Feeding nine-carbon chain 2-UP analogue (**151**): EICs at *m/z* = 366.25 and 364.24 ±0.01; **Green trace**: Feeding ten-carbon chain 2-UP analogue (**152**): EICs at *m/z* = 380.27 and 378.25 ±0.01; **Black trace:** Feeding 2-UP (**28**): EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigisoin (**25**) respectively; **Red trace:** Feeding twelve-carbon chain 2-UP analogue (**153**): EICs at *m/z* = 408.30 and 406.29 ±0.01 (**LHS (minor peak)** = product of McpG-catalyzed cyclization; **RHS (major peak)** = product of McpH-catalyzed condensation).

**Figure 3.2:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 422.32 ±0.01 corresponding to the $[M+H]^+$ ion of **32**: No feeding control; **B.** EIC at *m/z* = 420.3009 ±0.01 corresponding to the $[M+H]^+$ ion of **33**: No feeding control; **C.** EIC at *m/z* = 422.32 ±0.01 corresponding to the

$[M+H]^+$ ion of **32**: Feeding 2-UP analogue **154** and MBC **(21)**; **D.** EIC at *m/z* = 420.30 ±0.01 corresponding to the $[M+H]^+$ ion of **33**: Feeding 2-UP analogue **154** and MBC **(21)**.

**Figure 3.3:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **175**: No feeding control; **B.** EIC at *m/z* = 388.24 ±0.01 corresponding to the $[M+H]^+$ ion of **176** No feeding control; **C.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **175**: Feeding acetyline 2-UP **155** and MBC (**21**); **D.** EIC at *m/z* = 388.24 ±0.01 corresponding to the $[M+H]^+$ ion of **176**: Feeding acetyline 2-UP **155** and MBC (**21**).

**Figure 3.4:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 392.27 ±0.01 corresponding to the $[M+H]^+$ ion of **177**: No feeding control; **B.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **178**: No feeding control; **C.** EIC at *m/z* = 392.27 ±0.01 corresponding to the $[M+H]^+$ ion of **177**: Feeding 2-UP analogue **157** and MBC (**21**); **D.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **178**: Feeding 2-UP analogue **157** and MBC (**21**).

**Figure 3.5:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 378.25 ±0.01 corresponding to the $[M+H]^+$ ion of **181**: No feeding control; **B.** EIC at *m/z* = 376.24 ±0. 01 corresponding to the $[M+H]^+$ ion of **182**: No feeding control; **C.** EIC at *m/z* = 378.25 ±0.01 corresponding to the $[M+H]^+$ ion of **181**: Feeding 2-UP analogue **156** and MBC **(21)**; **D.** EIC at *m/z* = 376.24 ±0.01 corresponding to the $[M+H]^+$ ion of **182**: Feeding 2-UP analogue **156** and MBC **(21)**.

**Figure 3.6:** Natural products whose biosynthesis involves oxidation reactions catalyzed by Rieske oxygenases.

**Figure 3.7:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; EIC at *m/z* = 394.25 ±0.01 corresponding to the $[M+H]^+$ ion of **189**; **A.** No feeding control; **B.** Feeding MBC (**21**) and 2-UP analogue **156**.

**Figure 3.8:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; EIC at *m/z* = 406.29 ±0.01 corresponding to the $[M+H]^+$ ion of **198**; **A.** No feeding control; **B.** Feeding 2-UP analogue **197** and MBC **(21)**.

**Figure 3.9:** Base peak chromatogram from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* arising from feeding cyclopropane 2-UP analogue **190** and MBC (**21**) (*red*) and the no feeding negative control (*black*).

**Figure 3.10:** EIC at *m/z* = 364.24 ±0.01 corresponding to the $[M+H]^+$ ion of **199** from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** Feeding 2-UP analogue **197** and MBC (**21**); **B.** Feeding 2-UP analogue **202** and MBC (**21**); **C.)** Mixture of both extracts.

**Figure 3.11:** EIC at *m/z* = 422.28 ±0.01 corresponding to the $[M+H]^+$ ion of ring opened intermediate **203** from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-

**Figure 4.16: A.** SDS-PAGE analysis of the purification of recombinant $His_6$-SUMO-RedG using a pETSUMO expression vector + pDB1282, *E. coli* BL21 (DE3) cells, LB medium and anaerobic conditions; **B.** UV-Vis spectrum of the purified protein.

**Figure 4.17: A.** Purified $His_6$-SUMO-RedG post PD-10 column; **B.** Purified $His_6$-SUMO-RedG post PD-10 column + DTT + Fe (II) + $S^{2-}$ under anaerobic conditions; **C.** SDS-PAGE analysis of purified and reconstituted recombinant $His_6$-SUMO-RedG.

**Figure 4.18:** UV-Vis spectrum of anaerobically purified recombinant $His_6$-SUMO-RedG in which the FeS cluster has been reconstituted.

**Figure 4.19:** UV-Vis spectrum of anaerobically purified recombinant $His_6$-SUMO-RedG in which the FeS cluster has been reconstituted and exposed to atmospheric oxygen for 2 hours.

## List of schemes

## List of tables

## Acknowledgments

First and foremost I would like to thank my supervisor Professor Gregory L Challis for the opportunity to work on such an interesting multidisciplinary project. His humor, guidance and endless generosity means I leave my PhD studies with added confidence and self-belief, both in and out of the lab. Shame that century break never came in time….

I would like to thank all members of the Challis group, old and new, for making me look forward to coming to the lab every day and for invaluable advice and support; Dr. Dani Zabala, Dr. Joleen Masschelein, Dr. Jade Ronan, Dr. Shanshan Zhou, Dr. Gideon Idowu, Dr. Daniel Griffiths, Dr. Richard Gibson, Dr. Joshua Cartwright, Dr. Maz Costa, Dr. Rakesh Saroay, Dr. Emzo De Los Santos, Helen Smith, Ioanna Nakou, Jacob Sargeant, Dr. Yousef Dashti, Matt Beech, Rebin Salih, Dr. Chris Fage and Dr. Simone Kosol. I would like to give a special thank you to Dr. Chuan Huang and Dr. Xinyun Xian for their patience in teaching me all things biology and for being their happy and encouraging selves, every day.

I would like to thank Dr. Douglas Roberts for his 'excellent tutelage,' outstanding contribution to quizzes and for his friendship from the very first day of my MChem project. Thank you to Dr. Matthew Jenner for providing accidental and invaluable moments of hilarity, often based around his predictable attire. Thank you to Dr. Lona Alkhalaf for her support, for always humoring me, keeping me calm and joining me for impromptu breaks. I would like to give special thanks to Christian Hobson, who throughout my undergraduate degree, MChem and PhD has become like a family member. Thanks are also owed to the EPSRC for funding.

I would like to thank my brother and best friend, Matthew, who has bravely overcome hurdles and inspires me every day. I would also like to thank Sam Guest whose support has been without question and throughout childhood, university life and across different continents, remains my oldest and dearest friend.

Finally, and most importantly, I would like to thank my parents. They are my best friends, my favourite people, my biggest fans and my absolute inspiration.

## Declaration

The experimental work in this thesis is original research carried out by the author, unless otherwise stated, in the Department of Chemistry, University of Warwick, between September 2015 and August 2019. No material has been submitted for any other degree or at any other institution.

Results from authors are referenced in the usual manner throughout the text.

…… Date:…….

Christopher James Perry

## Abbreviations

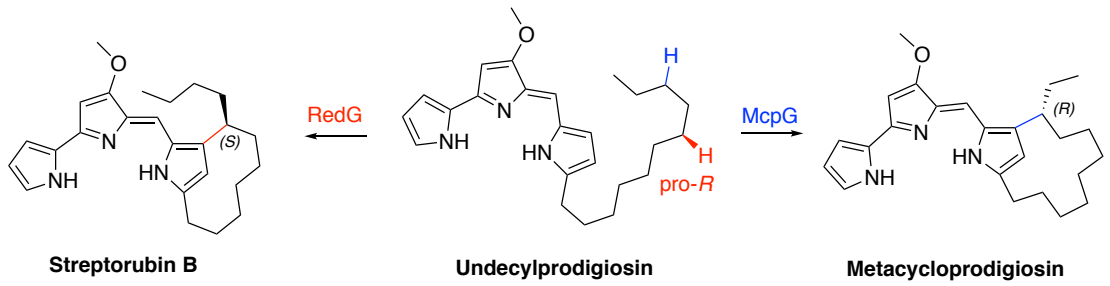
| | |
|---|---|
| 2-UP | 2-undecylpyrrole |
| 5-ALA | 5-Aminolevulinin acid |
| 5-FOA | 5-Fluoroorotic acid |
| ACP | Acyl carrier protein |
| Amp | Ampicillin |
| Asp | Aspartate |
| ATP | Adenosine triphosphate |
| BGC | Biosynthetic gene cluster |
| Bn | Benzyl |
| b.p | Base pairs |
| BOC | *tert*-Butyloxycarbonyl |
| CoA | Coenzyme A |
| COSY | Correlation spectroscopy |
| Chlo | Chloramphenicol |
| $^{13}$C-NMR | Carbon-13 nuclear magnetic resonance |
| CS | Clavaminate synthase |
| Da | Dalton |
| DCM | Dichloromethane |
| DEPT | Distortionless enhancement by polarization transfer |
| DMAP | 4-Dimethylaminopyridine |
| DMF | Dimethylformamide |
| DMSO | Dimethylsulphoxide |
| DNA | Deoxyribonucleic acid |
| DTT | Dithiothreitol |
| *E. coli* | *Escherichia coli* |
| EDC | 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide |
| EDTA | Ethylenediaminetetraacetic acid |
| e.e | Enantiomeric excess |
| EI | Electron ionization |
| EIC | Extracted ion chromatogram |
| EPR | Electron paramagnetic resonance |
| Eqv. | Equivalents |
| ESI | Electrospray ionization |
| FAD | Flavin adenine dinucleotide |

| | |
|---|---|
| FAS | Fatty acid synthase |
| FeS | Iron-sulphur |
| gDNA | Genomic DNA |
| Glu | Glutamic acid |
| HBC | 4-hydroxy-2,2'-bipyrrole-5-carbaldehyde |
| HBM | 4-hydroxy-2,2'-bipyrrole-5-methanol |
| HMBC | Heteronuclear multiple bond coherence |
| $^{1}$H-NMR | Proton nuclear magnetic resonance |
| HPLC | High performance liquid chromatography |
| HR MS | High resolution mass spectrometry |
| HSQC | Heteronuclear single quantum coherence |
| Hyg | Hygromycin |
| Hz | Hertz |
| IBX | 2-Iodoxybenzoic acid |
| IPNS | Isopenicillin-*N*-synthase |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| ISC | Iron-sulphur cluster |
| ISP2 | International *Streptomyces* Project 2 |
| Kan | Kanamycin |
| k.b | Kilobase |
| kDa | Kilodalton |
| LB | Luria-Bertani |
| LCMS | Liquid chromatography-mass spectrometry |
| LRMS | Low resolution mass spectrometry |
| MAP | 2-methyl-3-*N*-amylpyrrole |
| MBC | 4-hydroxy-2,2'-bipyrrole-5-carboxaldehyde |
| mcp | Metacycloprodigiosin |
| MS | Mass spectrometry |
| NADH | β-Nicotinamide adenine dinucleotide |
| NADPH | β-Nicotinamide adenine dinucleotide 2'-phosphate |
| NDX | Naladixic acid |
| NDO | Napthalene dioxygenase |
| OAS | *O*-acetylserine(thiol)lyase |
| OD | Optical density |
| *O*-PFBA.HCl | Pentafluorobenzoic acid |
| PCP | Peptidyl carrier protein |
| PCR | Polymerase chain reaction |

| | |
|---|---|
| PKS | Polyketide synthase |
| PLP | Pyridoxal-5'-phosphate |
| PPDK | Pyruvate phosphate dikinase |
| ppm | Parts per million |
| RE | Restriction enzyme |
| RO | Rieske oxygenase |
| RT | Room temperature |
| SAM | S-adenosyl methionine |
| SDS PAGE | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| SFM | Soya flour medium |
| SUMO | Small ubiquitin-like modifier |
| TAR | Transformation-associated recombination |
| TBAF | Tetrabutylammonium fluoride |
| TBS | *tert*-butyl-dimethylsilyl |
| *t*-BDMSCl | Tetrabutyldimethylsilyl chloride |
| *t*-BDPSCl | Tertabutyldiphenylsilyl chloride |
| THF | Tetrahydrofuran |
| TLC | Thin layer chromatography |
| TOF | Time of flight |
| UDP | Undecylprodigiosin |
| UV | Ultraviolet |
| UV-Vis | Ultraviolet-visible |
| WT | Wild-type |

# Abstract

Streptorubin B and metacycloprodigiosin are specialized metabolites belonging to the prodiginine family produced by *S. coelicolor* A3(2) and *S. longispororuber* respectively. The final step in their biosynthesis is an oxidative carbocyclization reaction of undecylprodigiosin (scheme 1). RedG is a Rieske non-heme iron dependent oxygenase that catalyzes the oxidative carbocyclization reaction at C-7' of the undecylprodigiosin alkyl chain to furnish streptorubin B. McpG is 75 % similar in sequence to RedG and catalyzes an analogous cyclization at C-9' of undecylprodigiosin to form metacycloprodigiosin.




**Scheme 1:** The Rieske oxygenase catalyzed oxidative carbocyclization reactions of undecylprodigiosin to form streptorubin B and metacycloprodigiosin.

The mechanism and stereochemical course of the McpG catalyzed cyclization reaction has been investigated *via* mutasynthesis. The entire metacycloprodigiosin biosynthetic gene cluster was captured from the genome of *S. longispororuber* using yeast-mediated transform associated recombination (TAR) cloning. The *mcp* cluster was heterologously expressed in *S. albus* and production of metacycloprodigiosin was confirmed. TAR was then used to create an in-frame deletion of a PKS-reductase fusion gene involved in the biosynthesis of 2-undecylpyrrole (2-UP), resulting in loss of metacycloprodigiosin production, which could be restored by feeding synthetic 2-UP. A range of 2-UP analogues were synthesized bearing various functional groups designed to probe different aspects of the McpG mechanism by feeding to the deletion mutant expressing an extra copy of *mcpH* and *mcpG*. The results showed that McpG catalyzes a direct cyclization reaction and does not catalyze the formation of a hydroxylated undecylprodigiosin intermediate *via* rebound onto the iron-oxo species. McpG was able to cyclize different alkyl chain lengths, but did not accept substrates with additional branching or adjacent π-bonds, and catalyzes cyclization with inversion of stereochemistry at C-9' of the alkyl chain. RedG was over-expressed, purified and the Rieske [2Fe-2S] cluster characterized by UV-Vis spectroscopy and chemically reconstituted *in vitro*.

# Chapter 1: Introduction

## 1.1 *Streptomyces*

*Streptomyces* is the largest genus of Actinobacteria, with more than 500 species currently known.[1] These bacteria inhabit soil and typically grow as a vegetative hyphal mass which differentiates into spores that can rapidly colonize their surroundings. Along with other related Actinobacteria, *Streptomyces* are renowned for their ability to produce bioactive specialized metabolites.

Specialized metabolites produced by Actinobacteria account for approximately two-thirds of the known antibiotics species from microorganisms, and 80 % of these are made by *Streptomyces* species.[1] Bioassay-guided discovery approaches have resulted in the isolation of vast amounts of bioactive metabolites used in pharmaceutical and agricultural industries (figure 1.1). Epoxomycin (**1**), for example, was originally isolated from the actinomycete strain Q996-17 as a selective proteasome inhibitor, exhibits anti-inflammatory activity.[2]

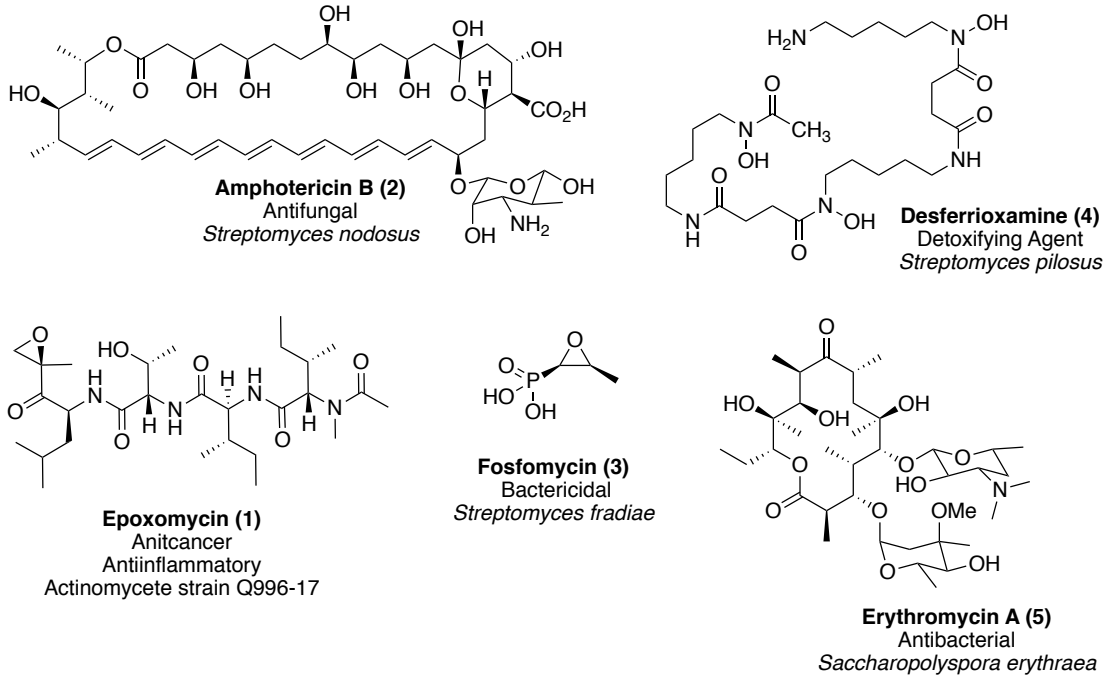



**Figure 1.1:** Bioactive specialized metabolites produced by Actinobacteria.

## 1.2 *Streptomyces coelicolor* A3(2)

*Streptomyces coelicolor* A3(2) is a model *Streptomyces* species that produces structurally diverse bioactive molecules (figure 1.2). In 2002, Bentley and co-workers published the complete genome sequence of *S. coelicolor* A3(2), which led to the identification of several clusters of genes that direct the biosynthesis of previously unknown specialized metabolites.[3] Over 15 families of specialized metabolites have been reported as a result of the genome sequencing project, including coelimycin P1 (**6**), germicidin A (**7**), geosmin (**8**) and the methylenomycin furans (MMFs) (**9-11**).[4–7] A particularly interesting family of specialized metabolites is the prodiginine alkaloids.[3,8] The prodiginines are a family of red-pigmented compounds that contain a common highly conjugated methoxypyrrolyldipyrromethene core and have immunosuppressant, antimalarial, anticancer and antibacterial activity.[9–13] This class of compounds exhibits many structural variations and is biosynthesized by enzymes that undertake challenging chemistries. Consequently, these compounds are the focus of this study.

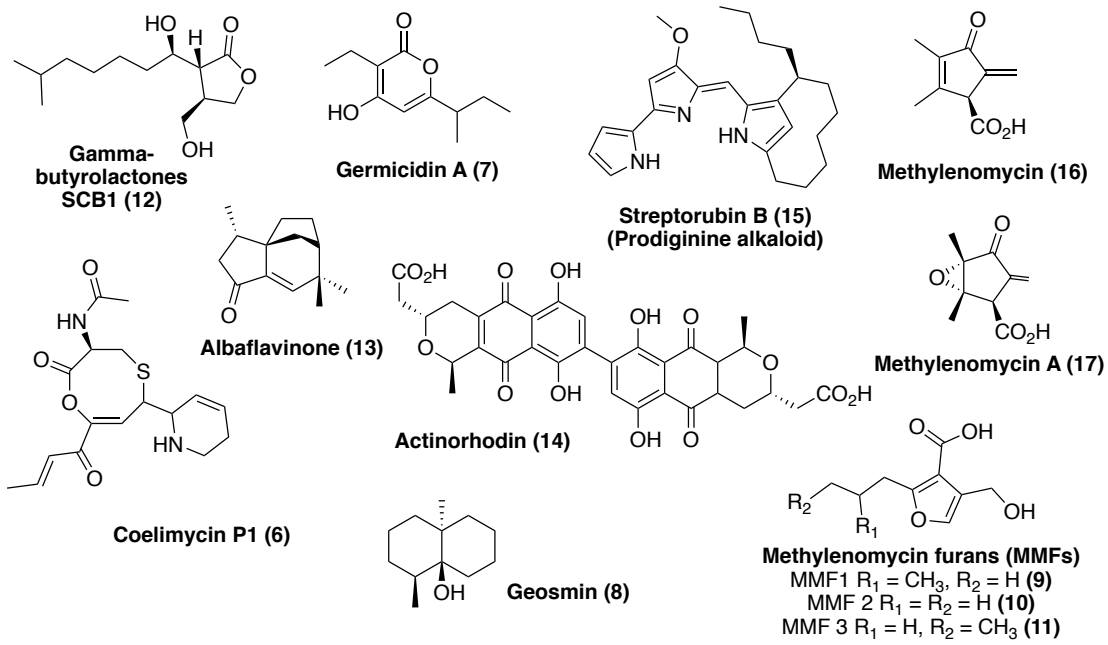



**Figure 1.2:** Specialized metabolites produced by *S. coelicolor* A3(2).

## 1.3 The prodiginine family

### 1.3.1 Structures of prodiginine alkaloids

In 1819 the residents of Padua, Italy observed the spontaneous reddening of the local cornmeal.[14] Quickly the nickname ''bloody polenta'' emerged and the professors of Padua University were appointed to investigate the reports as it was feared that the reddening had occurred as a result of evil spirits. It was soon found that the discolouration was not of supernatural origin but caused by the growth of a microorganism, which was subsequently named *Serratia marcesens*. The red-pigmented compound, prodigiosin (**18**), that was responsible for the discolouration of cornmeal and bread was isolated in pure form in 1929 by Wrede and Hettche.[15] Degradation studies on the isolate suggested three potential structures for prodigiosin (**18**) (figure 1.3).[16]

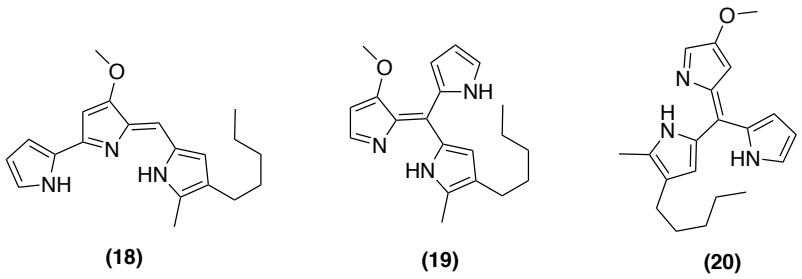



**Figure 1.3:** The originally proposed structures of prodigiosin (**18-20**), with compound **18** later confirmed to be correct.

The structure of prodigiosin (**18**) was not confirmed until 1960, when Vogel and co-workers isolated a compound with a chemical formula $C_{10}H_{10}O_2N_2$ which was later named 4-methoxy-2,2'-bipyrrole-5-carboxaldehyde (MBC) (**21**) from an *S. marcesens* strain that was blocked in prodigiosin (**18**) biosynthesis (figure 1.4).[17] Wasserman and co-workers realized that Wrede's degradation product 2-methyl-3-amypyrrole (MAP) (**22**) and MBC (**21**) can be condensed to form prodigiosin (**18**).[18] Further degradation studies on the compound isolated by Vogel and co-workers implied that the structure contained an aldehyde group and two *α-α'*-linked pyrrole moieties, ruling possible the proposed structures 1**9** and **20** unlikely.

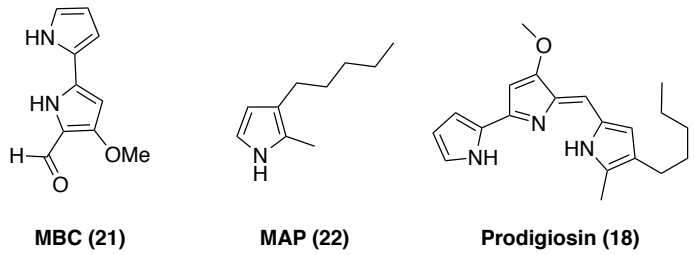



**Figure 1.4**: The structure of MBC (**21**) isolated from *S. marcescens* blocked in prodigiosin (**18**) biosynthesis, MAP (**22**) and the correct structure for prodigiosin (**18**).

Following the discovery of prodigiosin (**18**) a large number of prodiginine compounds were isolated, predominantly from Actinobacteria. Undecylprodigiosin (**23**) is a prodiginine alkaloid produced by *S. coelicolor* A3(2) and its carbocyclic derivative, streptorubin B (**15**) contains an ansa-bridged ring incorporating the alkyl chain and the pyrrole C-ring (figure 1.5).[19–21,22] The structure of streptorubin B (**15**) was originally mis-assigned as the ortho-bridged isomer (**24**), later to be corrected to the meta-bridged isomer (**15**).[22,23] Following streptorubin B (**15**), metacycloprodigiosin (**25**) and cycloprodigiosin (**26**) were the first cyclic analogues of prodigiosin (**18**) to be isolated, from *S. longispororuber* and the *Alteromonas rubra* strain respectively.[24,25]

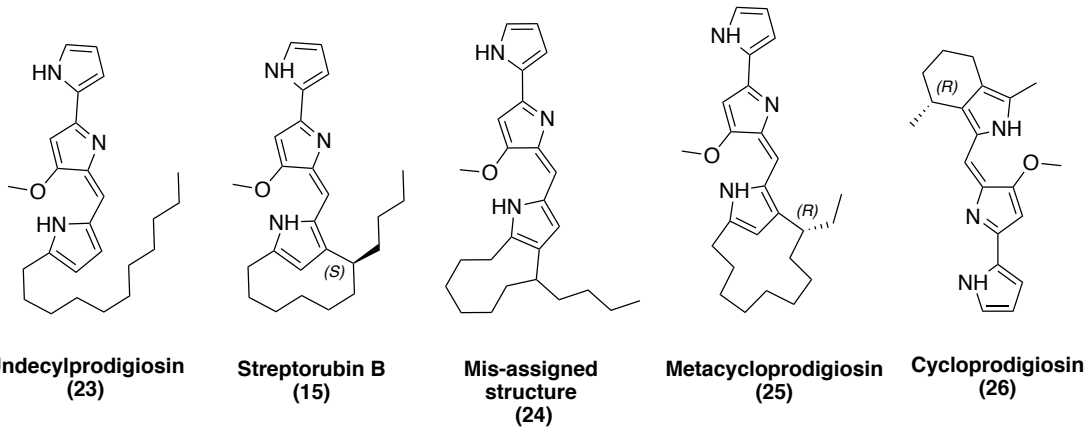



**Figure 1.5:** The structures of undecylprodigiosin (**23**), streptorubin B (**15**), metacycloprodigiosin (**25**) and cycloprodigiosin (**26**) and the mis-assigned structure (**24**) of streptorubin B.

Carbocyclic prodiginine alkaloids display interesting structures due to the various stereocentres and the geometries of the exocyclic double bond. Due to the orientation of the *n*-butyl side-chain and the methoxypyrrolidipyrromethene core relative to the 10-membered ansa-bridge, streptorubin B (**15**) can exist as either the *syn* or the *anti* atropisomer. Challis and co-workers and Thomson and co-workers both undertook independent studies to show that

streptorubin B (**15**) exists as a pair of interconverting atropisomers and that the major structure at equilibrium is the *anti*–atropisomer (figure 1.6).[26,27]

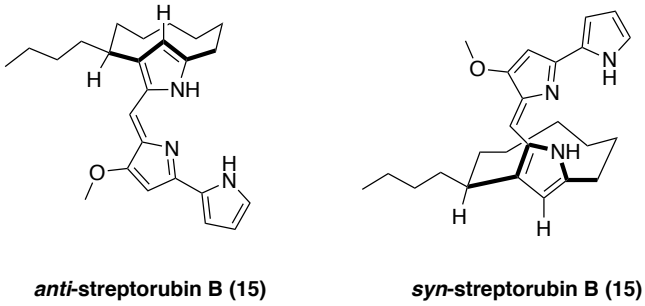



**Figure 1.6:** The structures of *anti* and *syn*-streptorubin B (**15**). The *anti* atropsiomer was shown to be the major carbocyclic derivative produced by *S. coelicolor*.

Both the Challis group and Thomson group then independently confirmed the absolute stereochemistry of streptorubin B (**15**).[26,27] Thomson and co-workers enantioselectively synthesized streptorubin B (**15**) to determined the 7*'S* absolute stereochemistry by X-ray crystallography and by comparing the circular dichroism spectrum with that of natural streptorubin B (**15**).[26] Challis and co-workers utilized synthetic *R* and *S* [4*'*-$^2$H]-2-undecylpyrrole (**27**) fed to a mutant of *S. coelicolor* blocked in 2-undecylpyrrole (2-UP) (**28**) biosynthesis (a late staged intermediate in streptorubin B (**15**) biosynthesis) (figure 1.7).[27]

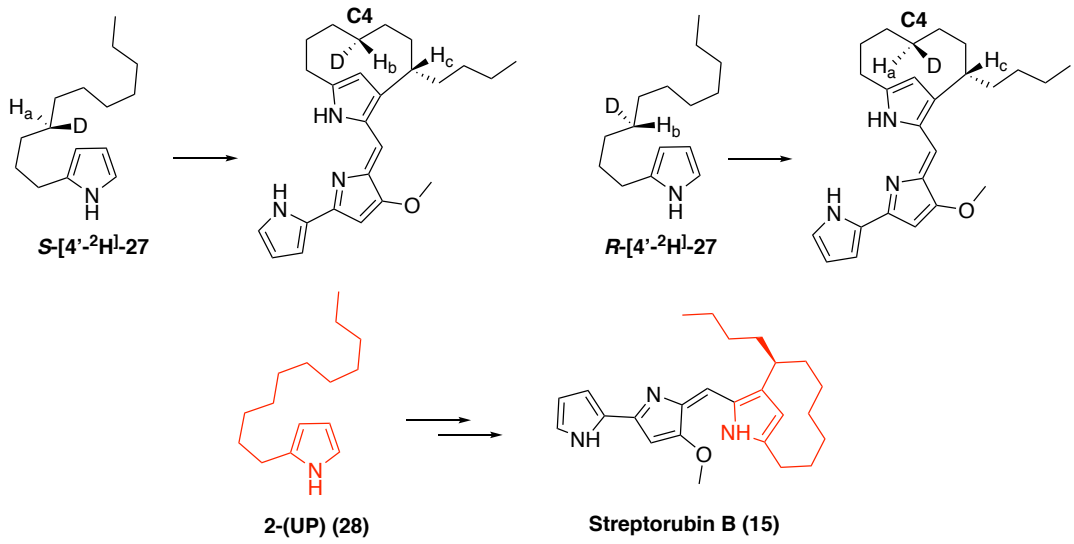



**Figure 1.7**: Feeding experiments using deuterium labelled 2-UP analogues *R*-**27** and *S*-**27** to a *redL S. coelicolor* deletion mutant used to confirm the absolute stereochemistry of streptorubin B (**15**) by $^1$H and $^2$H-NMR comparisons between $H_a$, $H_b$, $H_c$ and the deuterium label and the incorporation of 2-UP (**28**) in the structure of streptorubin B (**15**).

The 7*'S* absolute stereochemistry of streptorubin B (**15**) was confirmed by analyzing the difference in the chemical shift of the diastereotopic protons at C4 of the alkyl chain (which are located over the pyrrole aromatic ring) by using $^{1}$H and $^{2}$H-NMR (figure 1.7). Analysis of isolated natural streptorubin B (**15**) by HPLC using a homochiral stationary phase identified small amounts of the *syn* diasteremoer, and also the 7'*R* enantiomer (and corresponding *syn* diastereomer) suggesting that streptorubin B (**15**) is not biosynthesized with complete stereocontrol.[27]

Comparison of the circular dichroism spectrum of natural metacycloprodigiosin (**25**) isolated from *S. longispororuber* with *R*-metacycloprodigiosin prepared by an enantioselective total synthesis confirmed that the natural compound has 9*'R* absolute stereochemistry. Streptorubin B (**15**) and metacycloprodigiosin (**25**) therefore share an antipodal relationship (but are not enantiomers).[26,28,29]

Cycloprodigiosin (**26**) was shown to be biosynthesized as an 83:17 mixture or *R*- and *S*-enantiomers in *Pseudoalteromonas rubra* (figure 1.8).[24] Another structurally similar cyclic prodiginine is the carbocyclic derivative of nonylprodigiosin (**29**), cyclononylprodigiosin (**30**) (figure 1.8).[30] Both metabolites were reported from the same *Actinomadura madurae* strain, although produced when the strain was grown under different conditions.

Seto and co-workers reported the discovery of roseophillin (**31**) from *S. griseoviridis.* Roseophillin (**31**) is similar in structure to most carbocyclic prodiginines, however contains two C-C bonds between the alkyl chain and the conjugated heterocyclic ring system that creates a tricyclic cyclopentylpyrrolophane. Further differences include a methoxyfuran located at the B-ring position and a chlorine atom on the A-ring.[31] Despite the structural differences of roseophillin (**31**) to the other carbocyclic prodiginines, they appear to be biosynthesized from a similar dimethyl analogue (**32**) of undecylprodigiosin. Evidence for this intermediate was provided by the discovery of prodigiosin R1 (**33**), which contains a dimethyl moeity at the terminus of the alkyl chain (figure 1.8).[32]

Marineosin A (**34**) and B (**35**) are prodiginine compounds with an even more atypical structure. These metabolites were isolated from the marine-derived *Streptomyces*-related strain, CNQ-617.[33] The structure of the marineosins implies that they are also biosynthesized from undecylprodigiosin (**23**) (as is the case for streptorubin B (**15**) and metacycloprodigiosin (**25**)) *via* oxidative cyclization reactions. This was experimentally confirmed by Reynolds and co-workers by heterologous expression of the *mar* biosynthetic

gene cluster in a *S. venezuelae* host.[34,35] It was shown that the preceding steps to the final cyclization likely mirror that in the *red* cluster, minus the addition of the hydroxyl moeity required for spiroaminal formation.

More recently, Ross and co-workers isolated and characterized the macrocylic tambjamine compound tambjamine MYP1 (**37**) from the marine bacterium *Pseudoalteromonas citrea* (figure 1.8).[36] The strain was grown on solid media, extracted, and the metabolite purified by flash Si-chromatography and RP-HPLC. Evidence for the cyclic structure was provided by HR-MS and FTMS-ESI, generating an $[M+H]^+$ that is two mass units lower than that reported for tambjamine YP1 (**36**). The cyclic nature of tambjamine MYP1 (**37**) was subsequently confirmed by 1-D and 2-D NMR experiments.

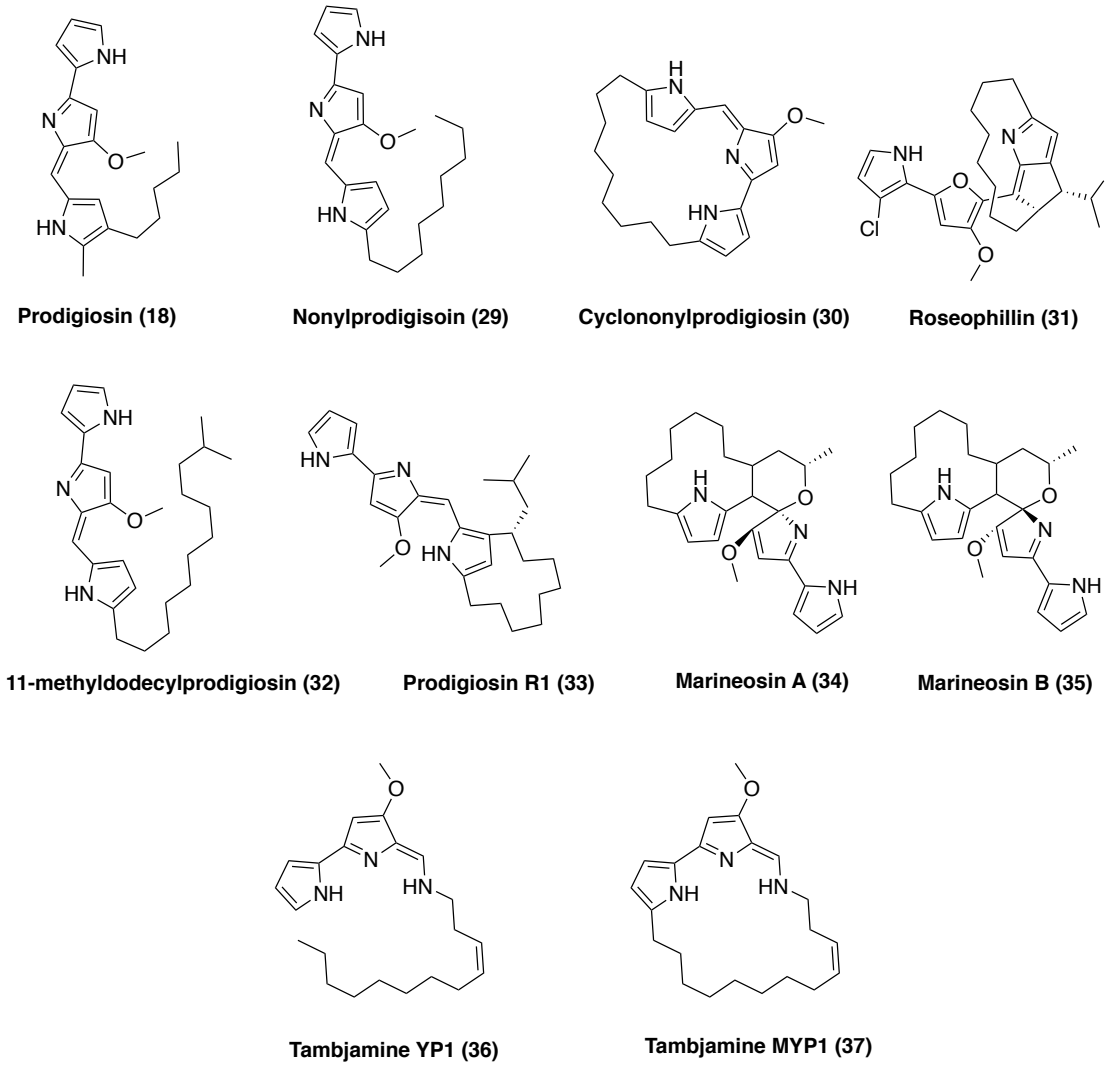
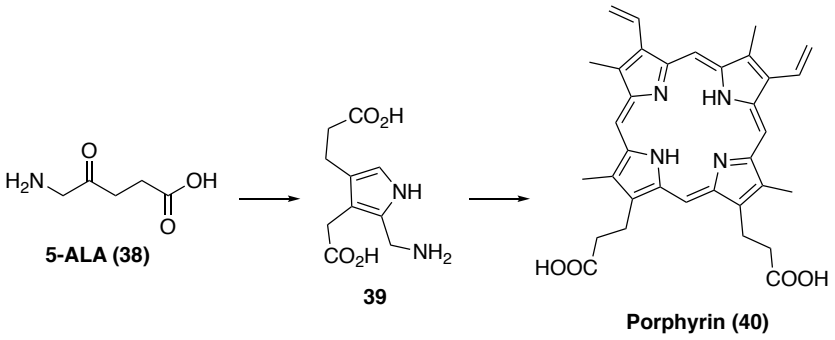



**Figure 1.8:** The structure of different linear and carbocyclic derivatives of prodiginine and prodiginine-like specialized metabolites.

### 1.3.2 Prodiginine biosynthesis

Early studies investigated whether prodiginines and porphyrin (**40**) could have a common biosynthetic pathway. However, feeding radiolabeled 5-aminolevulinic acid (5-ALA) (**38**) (a known intermediate in porphyrin (**40**) biosynthesis) to *S. marcesens* did not result in radiolabelling of prodigiosin (**18**) (scheme 1.1).[37]




**Scheme 1.1:** The biosynthesis of porphyrin (**40**).

Wasserman and co-workers carried out a series of incorporation experiments using $^{13}C$ and $^{14}C$-labelled precursors to elucidate the metabolic origins of both prodigiosin (**18**) and undecylprodigiosin (**23**) from *S. marcesens* and *S. longispororuber* respectively (figure 1.9).[38] The A-rings of both scaffolds are derived from L-proline, whilst the B-rings are derived from L-serine, acetate and the *S*-methyl group of methionine. The C-ring of prodigiosin (**18**) is derived from acetate and L-alanine whereas the C-ring of undecylprodigiosin (**23**) originated from acetate and glycine. It was therefore proposed that prodigiosin (**18**) is biosynthesized by *S. marcescens* in a similar manner to undecylprodigiosin (**23**) in *S. longispororuber via* a bifurcated pathway sharing a common intermediate, MBC (**21**). MBC (**21**) is then condensed with MAP (**22**) to furnish prodigiosin (**18**), and with 2-UP (**28**) to furnish undecylprodigiosin (**23**) (figure 1.9).[38,39]

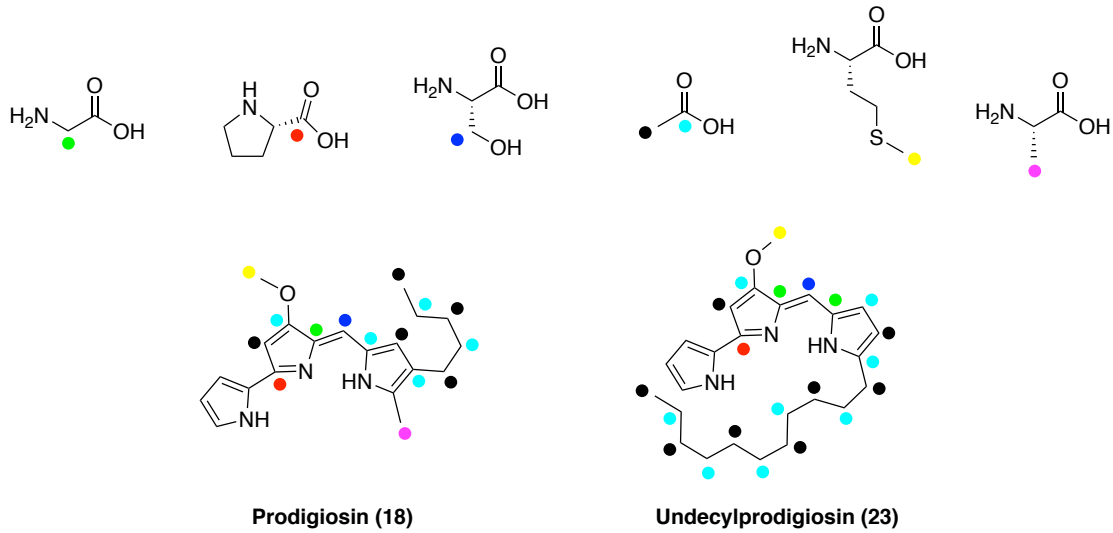



**Figure 1.9:** Incorporation experiments by Wasserman and co-workers using $^{13}C$ and $^{14}C$-labeled precursors (highlighted) to determine the origins of prodigiosin (**18**) and undecylprodigiosin (**23**).

### 1.3.3 Biosynthesis of undecylprodigiosin (23) and streptorubin B (15)

In 1980, Hopwood and co-workers used a mutant *S. coelicolor* A3(2) strain that was blocked in the biosynthesis of actinorhodin (**14**) and related metabolites, to identify the biosynthetic gene cluster for the red antibiotics later identified by Floss and co-workers as undecylprodigiosin (**23**) and a carbocyclic derivative for which a conclusive structure was not assigned.[40] The cloning and heterologous expression of this cluster in *S. lividans* was reported and later in 2001 the complete genome sequence of *S. coelicolor* became available through the *S. coelicolor* genome sequencing project.[3] The project provided the sequence of the gene cluster responsible for the biosynthesis of undecylprodigiosin (**23**) and the carbocyclic derivative.[3] This cluster contains twenty three genes arranged in four transcription units and the function of most of the encoded proteins has been confirmed using a combination of gene deletion and chemical complementation feeding experiments. In 2008, Challis and co-workers conclusively demonstrated that streptorubin B (**15**) is the carbocyclic derivative of undecylprodigiosin (**23**).[41]

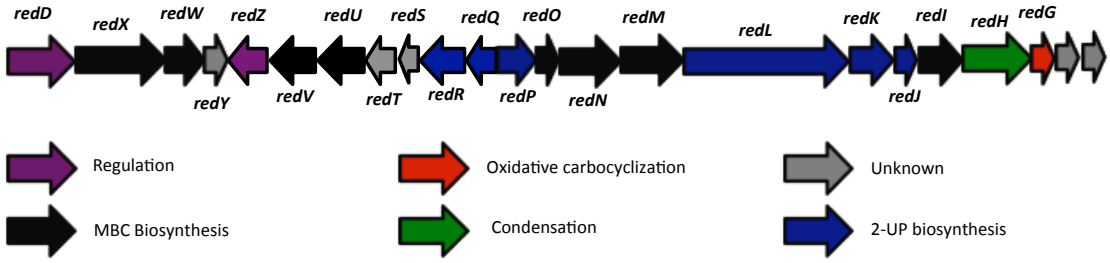



**Figure 1.10:** Organization of the biosynthetic gene cluster that directs the biosynthesis of undecylprodigiosin (**23**) and streptorubin B (**15**) in *S. coelicolor* A3(2) and the function of each encoded protein.

### 1.3.4 The biosynthesis of MBC (21)

Eight genes have been shown to direct the biosynthesis of MBC (**21**) (*redI*, *redM*, *redN*, *redO*, *redU*, *redW*, *redX* and *redV*).[42] Initial steps in the biosynthesis of MBC (**21**) involve the RedM catalyzed activation of L-proline to form an amino-acyl adenylate before phosphopantetheinylation of the peptidyl carrier protein (PCP) RedO which is catalyzed by the phosphopantetheinyl transferase RedU (figure 1.11). The resulting prolyl thioester (**41**) undergoes four electron oxidation to furnish the corresponding pyrrole-2-carboxyl thioester (**42**). This reaction is catalyzed by RedW, a flavin dependent oxidase. RedX and RedN are involved in the subsequent elongation of the pyrrole-2-carboxyl intermediate (**42**) by decarboxylative condensation with a malonyl-CoA-derived thioester to furnish the β-keto thioester (**43**). The α-oxoamine synthase (OAS) domain of RedN is proposed to catalyze the decarboxylative Claisen-condensation of L-serine with the thioester (**43**) to furnish an α,γ-dioxoamine which subsequently undergoes cyclization, dehydration and tautomerisation to yield 4-hydroxyl-2,2'-bipyrrole-5-methanol (HBM) (**44**). Unpublished work from the Challis group indicates the RedV is involved in the oxidation of the primary alcohol of HBM (**44**) to form the corresponding aldehyde HBC (**45**), prior to RedI-catalyzed methylation of the pyrrole hydroxyl group with S-adenosyl methionine (SAM) to furnish MBC (**21**).

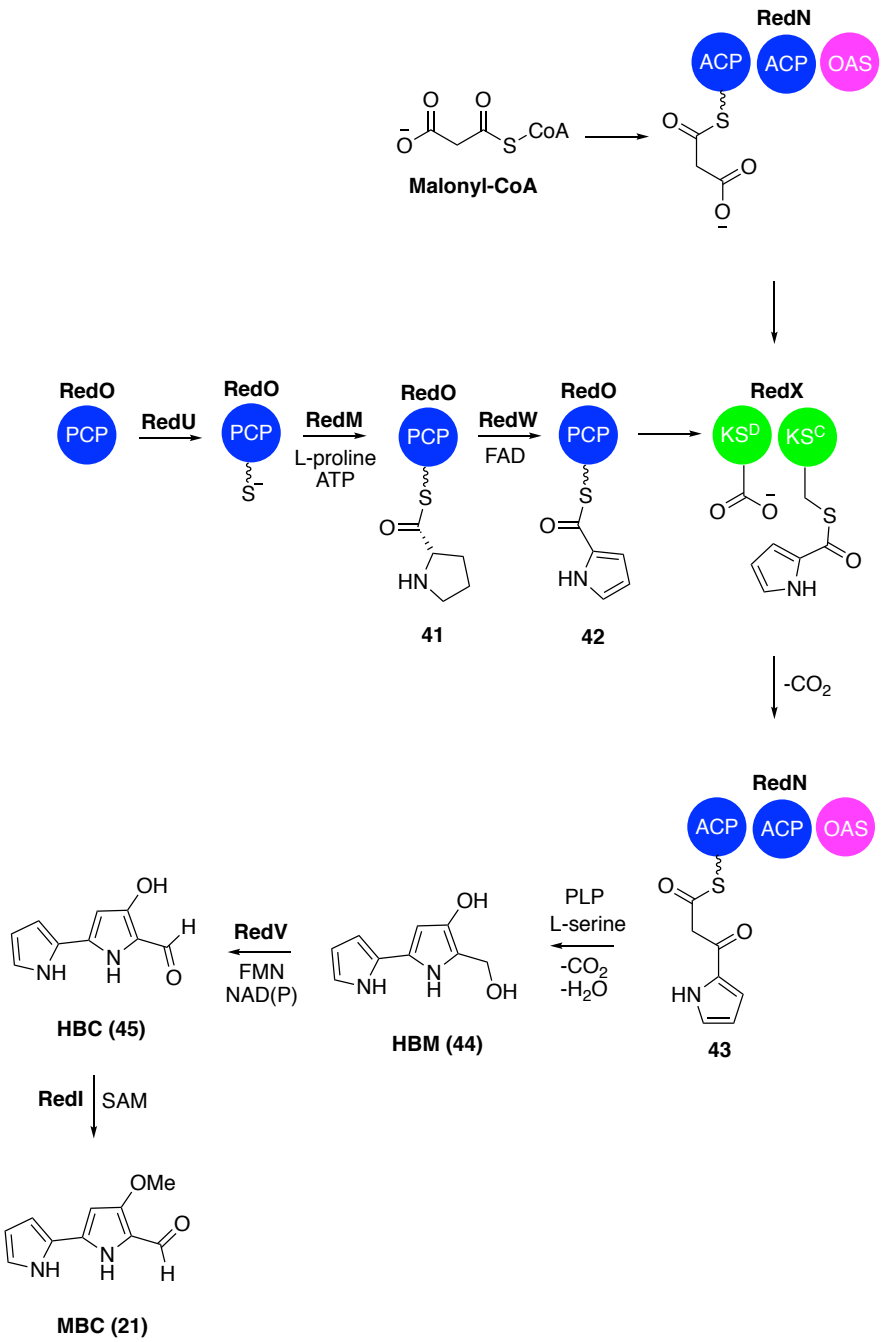



**Figure 1.11:** The proposed pathway for the assembly of MBC (**21**), in *S. coelicolor* A3(2).[42]

### 1.3.5 The biosynthesis of 2-UP (28)

It has been previously demonstrated that 2-UP (**28**) is produced from seven acetate units and one glycine unit.[41] In *S. coelicolor*, dodecanoic acid (**48)** is assembled by RedP, RedR and RedQ which are homologues of FabH, FabF and FabC enzymes responsible for fatty acid biosynthesis. RedP (FabH homologue) initiates the biosynthesis of dodecanoic acid (**48**) *via* a decarboxylative condensation of acetyl-CoA (**46**) with a malonyl extender unit bound to the ACP of RedQ (figure 1.12). The resulting β-ketothioester undergoes ketoreduction, dehydration and enoyl reduction to yield a RedQ-bound butanoylthioester. These reactions are proposed to be catalyzed by components of the primary metabolic fatty acid synthase (FAS). RedR catalyzes several further rounds of chain elongation using malonyl-RedQ with reduction of the β-keto group to a methylene catalyzed by the FAS enzymes after each chain elongation reaction, forming dodecanoyl-RedQ (**47**). The thioester is then proposed to be hydrolyzed by RedJ to form dodecanoic acid (**48**) which is transferred to the ACP domain of RedL after activation by adenylation.[43] A subsequent decarboxylative condensation with a malonyl unit is catalyzed by the KS domain of RedL to furnish a β-ketomyristoyl-ACP thioester (**49**). The PLP-dependent OAS domain of RedL catalyzes the condensation of glycine with (**49**). Subsequent cyclization and elimination of water furnishes 4-keto-2-undecylpyrroline (**50**). The ketone is reduced by the NAD(P)H dependent reductase RedK and elimination of water restores the aromaticity to give 2-UP (**28**).

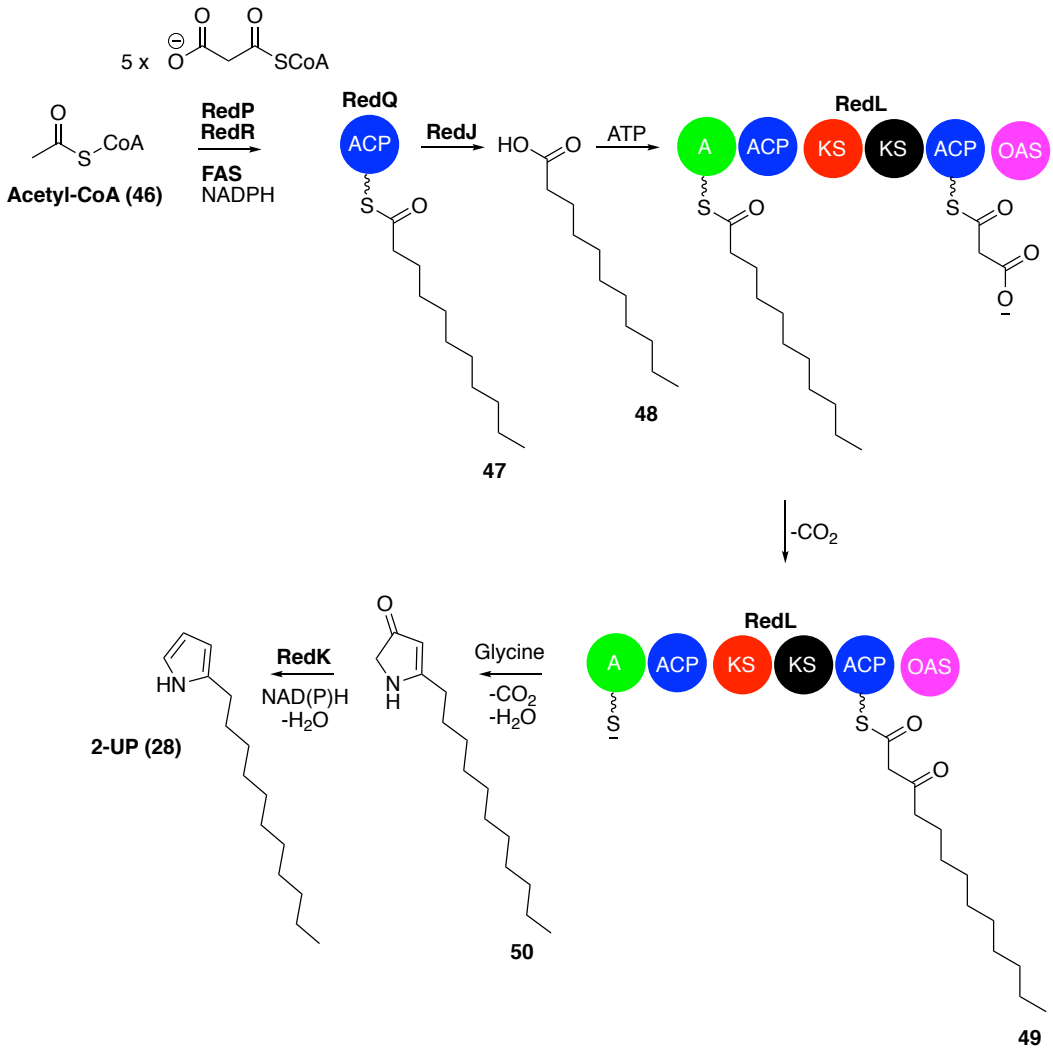



**Figure 1.12:** The biosynthesis of 2-UP (**28**).[41]

### 1.3.6 Condensation of MBC (21) and 2-UP (28)

The final step in the biosynthesis of undecylprodigiosin (**23**) and penultimate step in streptorubin B (**15**) biosynthesis is the condensation of MBC (**21**) with 2-UP (**28**).[41] Challis and co-workers have shown that this reaction is catalyzed by RedH in *S. coelicolor* A3(2).[44] A *redH* deletion mutant failed to produce prodiginines, but accumulated MBC (**21**) and 2-UP (**28**).[44] Feeding of synthetic MBC (**21**) and 2-UP (**28**) to *S. venezuelae* expressing *redH* resulted in the production of undecylprodigiosin (**23**).[44] Sequence comparison of RedH with proteins of known function showed that it contains an ATP-binding domain and a phosphotransfer domain similar to phosphoenolpyruvate synthase (PEPS) and pyruvate phosphate dikinase (PPDK). RedH contains a third domain that does not show similarity to other proteins and so is proposed to bind the MBC (**21**) substrate partner.

RedH is believed to activate the carbonyl group of MBC (**21**) *via* phosphorylation with ATP to form the iminium ion (**51**) (figure 1.13). Subsequent nucleophillic attack by 2-UP (**28**), elimination of phosphate and deprotonation, restores the aromaticity of the pyrrole and liberates undecylprodigiosin (**23**).

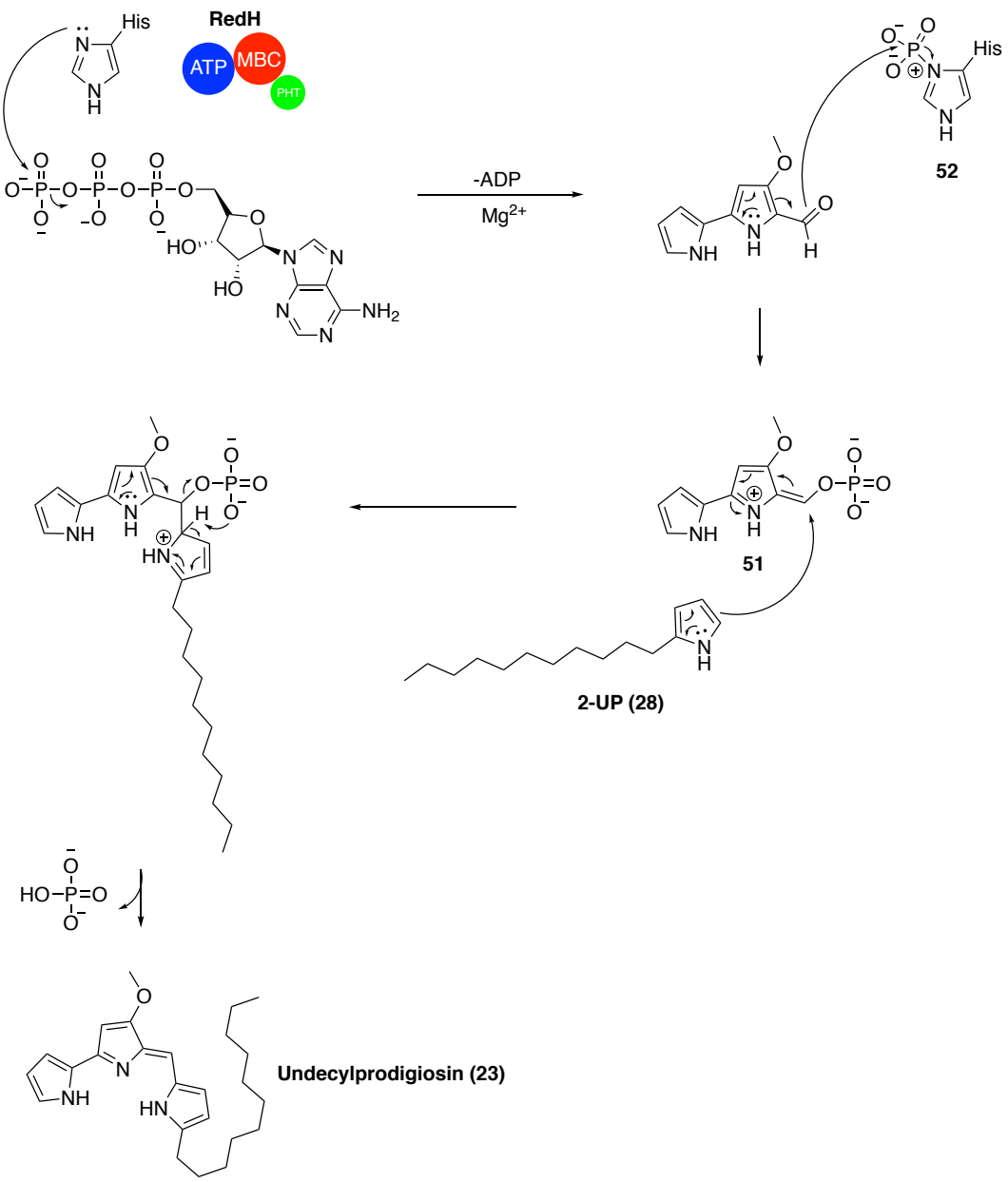
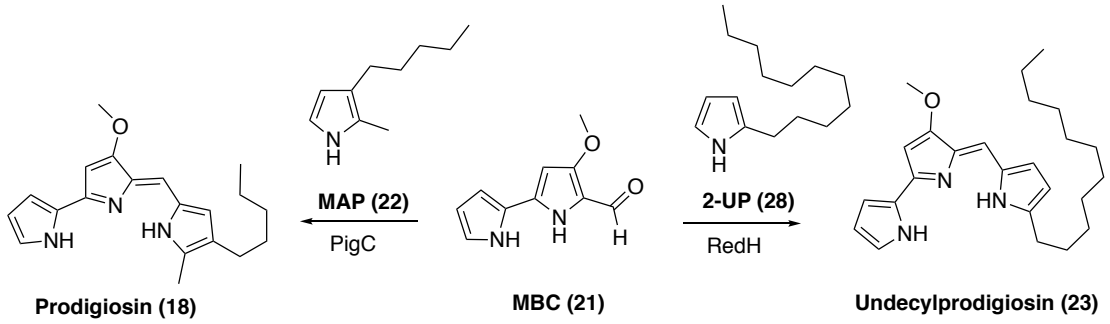



**Figure 1.13:** The final step in the biosynthesis of undecylprodigiosin (**23**), RedH catalyzed condensation of MBC (**21**) and 2-UP (**28**).[44]

Leeper and co-workers investigated PigC, a homologue of RedH that is involved in the condensation of MBC (**21**) and MAP (**22**) to furnish prodigiosin (**18**) in *S. marcesens*.[45] PigC was over-expressed and purified and the condensation of synthetic MBC (**21**) and MAP (**22**) was reconstituted *in vitro*. PigC was shown to be a membrane bound prodiginine synthetase

that utilizes ATP to condense MBC (**21**) and MAP (**22**) by phosphorylation of a histidine residue (**52**) in the phosphoryl transfer domain. Mutagenesis of the histidine in question subsequently abolished activity.

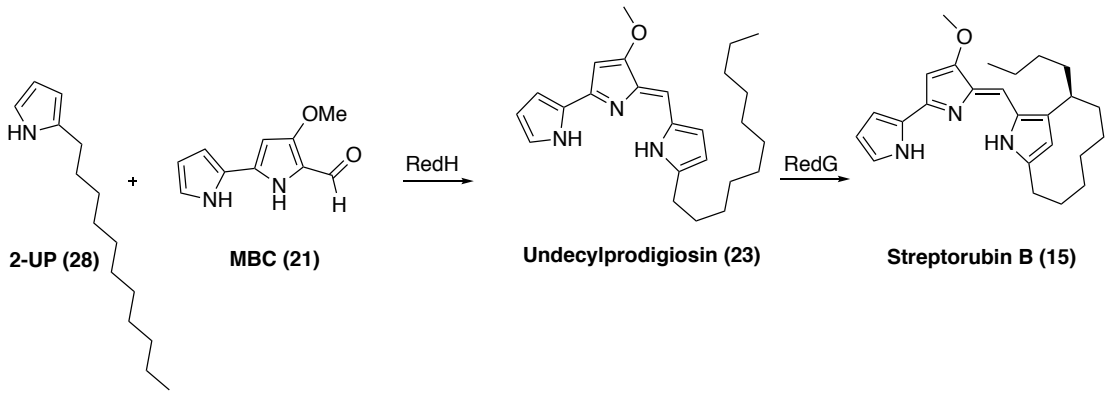



**Scheme 1.2:** The roles of PigC and RedH in the biosynthesis of prodigiosin (**18**) and undecylprodigiosin (**23**).

### 1.3.7 Oxidative carbocyclization of undecylprodigiosin (23) to streptorubin B (15)

As discussed in section 1.3.3, streptorubin B (**15**) is a prodiginine alkaloid produced by the *red* gene cluster of *S. coelicolor* A3(2).[41,46] The cluster contains a gene encoding a Rieske oxygenase, *redG*, that was identified based on sequence similarity to the Rieske oxygenase naphthalene dioxygenase (NDO).[28,46] Like NDO, the α-subunits of RedG contains an N-terminal $CXHX_{17}CX_2H$ [2Fe-2S] cluster binding motif and a C-terminal $EXHX_4H$ variant of the $DXHX_4H$ motif found in NDO.[28,46] Conserved histidine residues bind ferrous iron whilst the Asp residue of NDO is implicated in the transfer of electrons from the Rieske cluster to the non-heme iron centre. In RedG this Asp is mutated to Glu, which is believed to have no effect on the method of electron transfer.[28,46] This Asp to Glu mutation was subsequently observed by Chen and co-workers in the Rieske oxygenase CntA, which is implicated in the metabolism of carnitine in representative genomes of the human microbiota.[47]

Challis and co-workers used a *S. coelicolor* A3(2) mutant in which the *redG* gene was deleted from the chromosome and consequently the resulting strain no longer produced streptorubin B (**15**), instead accumulating the precursor to carbocyclization, undecylprodigiosin (**23**).[28] Feeding undecylprodigiosin (**23**) to *S. venezuelae* expressing *redG* produced streptorubin B (**15**), indicating that *redG* encodes the protein that is responsible for the oxidative carbocyclization reaction.[28] Feeding undecylprodigiosin (**23**) to *S. venezualae* expressing *redH* and *redG* together (*redHG*) increased the production levels of streptorubin B (**15**).[28]




**Scheme 1.3:** Late-stages in the biosynthesis of streptorubin B (**15**).

### 1.3.8 Rieske oxygenase-catalyzed oxidative carbocyclization reactions in the biosynthesis of other prodiginine alkaloids

As discussed in section 1.3.1, the biosynthetic gene clusters of several red-pigmented prodiginine alkaloids contain a gene encoding a Rieske non-heme iron-dependent oxygenase that is proposed to catalyze late stage oxidative carbocyclization reactions (figure 1.14).

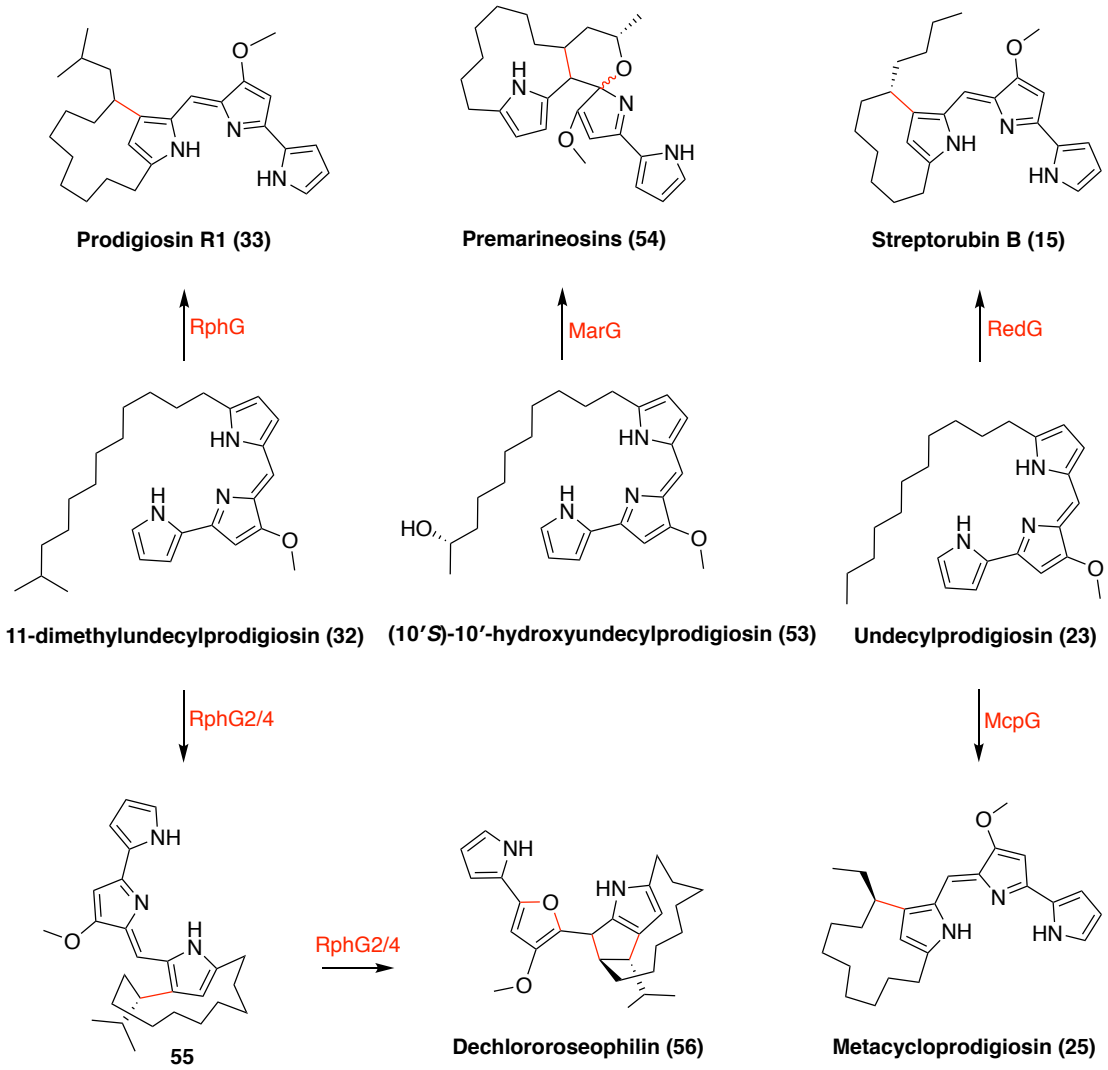
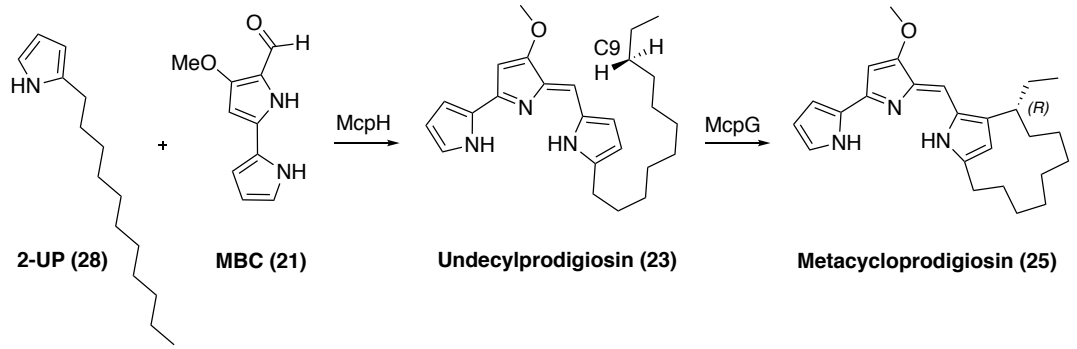



**Figure 1.14:** The bonds (red) proposed to be formed by Rieske oxygenase enzymes in the biosynthesis of several prodiginine alkaloids.

### 1.3.8.1 Role played by McpG in metacycloprodigiosin (25) biosynthesis

A homologue of RedG, McpG, catalyzes a similar oxidative carbocyclization reaction on undecylprodigiosin (**23**) in the biosynthesis of metacycloprodigiosin (**25**).[28] However, McpG catalyzes cyclization at C9 rather than C7 of the alkyl chain.[28] Metacycloprodigiosin (**25**) is structurally similar to streptorubin B (**15**), but it contains a 12-membered *ansa*-brided carbocycle rather than a 10-membered ring and is *R*-configured whereas streptorubin B (**15**) is *S*-configured. The two compounds are thus antipodal to one another (but not enantiomeric). The *mcpG* gene in *S. longispororuber* encodes a protein with 75 % sequence similarity to *redG*. A mutant of *S. coelicolor* in which the *redG* gene had been deleted and replaced with *mcpG* resulted in production of metacycloprodigiosin (**25**) (scheme 1.4).[48]

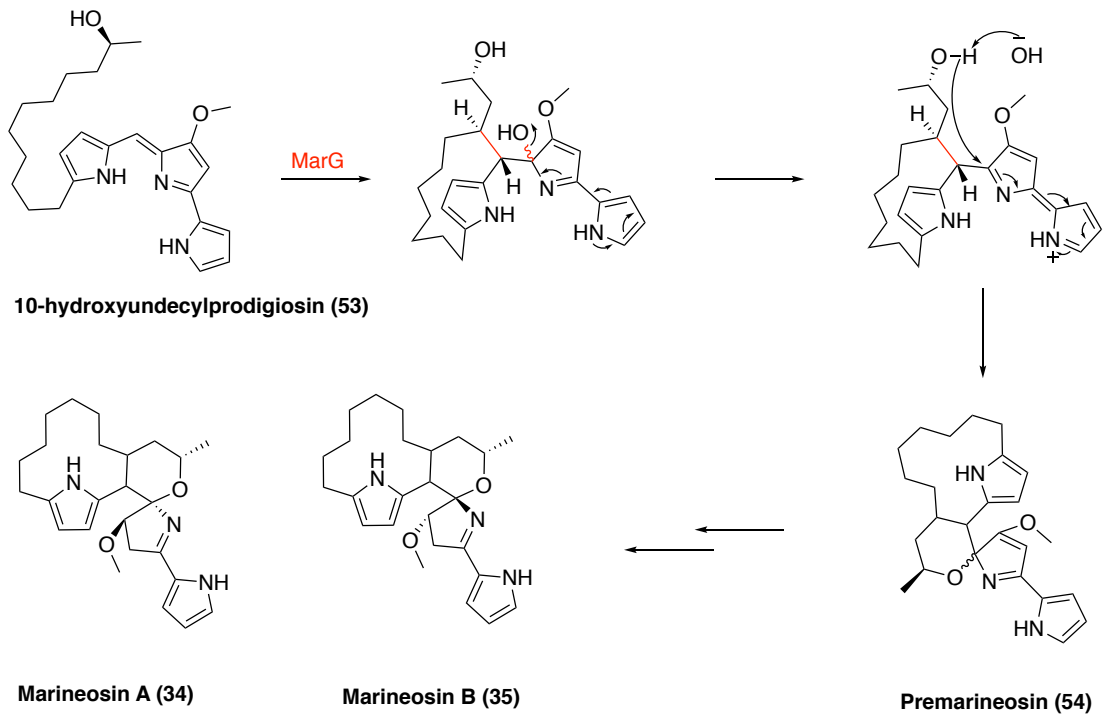



**Scheme 1.4:** The McpG-catalyzed oxidative carbocyclization of undecylprodigiosin (**23**) to metacycloprodigiosin (**25**).

### 1.3.8.2 Role played by MarG in marineosin biosynthesis

In 2008, Fenical and co-workers reported isolation and characterization of the cytotoxic spiroaminals marineosin A (**34**) and B (**35**) from a marine actinomycete.[33] The biosynthetic gene cluster for the marineosins was identified by Reynolds and co-workers and subsequently expressed in a *S. venezuelae* heterologous host.[34] A gene (*marG*) encoding a putative Rieske oxygenase was identified within the cluster which showed 64 % similarity to RedG. When the *marG* gene was deleted the strain accumulated compound (**53**), and when (**53**) was fed to *S. venezuelae* expressing *marG*, the production of premarineosin (**54**) was observed. These data imply that MarG is responsible for oxidative carbo and oxacyclization reactions that result in the generation of the premarineosin (**54**), which is subsequently reduced to marineosins A (**34**) and B (**35**) (scheme 1.5). Both synthetic enantiomers of 10-hydroxyundecylprodigiosin (**53**) were converted to premarineosin (**54**). To fully elucidate the stereochemical course of the transformation, a deuterium label was inserted at C10*'* of the 10-

hydroxyundecylprodigiosin substrate, which was retained in the premarineosin (**54**) product that resulted from feeding of the 10*S*' enantiomer.[35] The subsequent 10*R'* enantiomer is proposed to be oxidized to the ketone before reduction to the 10*S'* enantiomer prior to premarineosin (**54**) formation. Challis and co-workers have since proposed a mechanism for the cyclization of 10-hydroxyundecylprodigiosin (**53**) to form the premarineosin (**54**) and consequently marineosins A (**34**) and B (**35**) (scheme 1.5).[49]

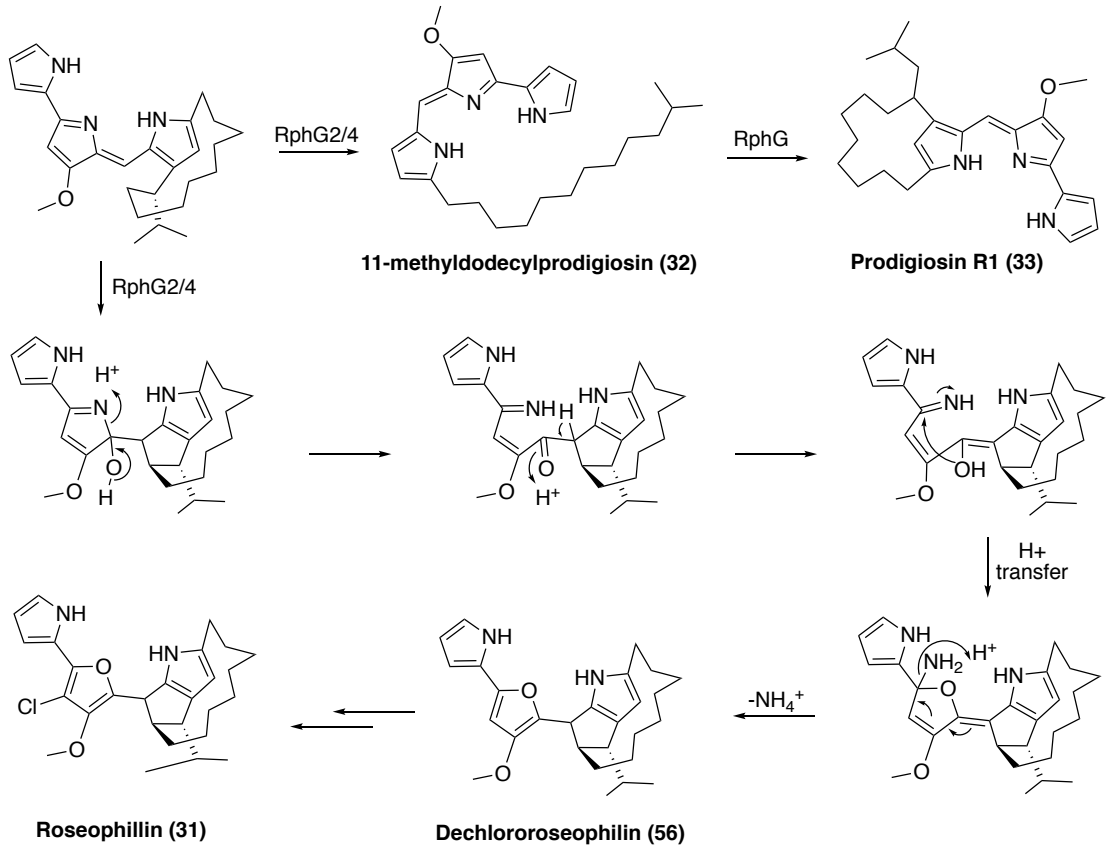



**Scheme 1.5:** The Mar-G catalyzed oxidative carbo and oxacyclization reactions to form the premarineosins (**54**) in the biosynthesis of marineosins A (**34**) and B (**35**).

### 1.3.8.3 Roles played by RphG, RphG2 and RphG4 in roseophillin (66) biosynthesis

Roseophillin (**31**), dechlororoseophillin (**56**) and prodigiosin R1 (**33**) are produced by *S. griseoviridis*.[50,51,52] Roseophillin (**31**) and dechlororoseophillin (**56**) differ from the common dipyrrole-pyrromethene scaffold of prodiginine metabolites by the central furan ring. Prodigiosin R1 (**33**) is structurally similar to metacycloprodigiosin (**25**), however, contains a branching methyl group at the terminus of the alkyl chain. Due to the make-up of the central scaffolds of (**31**), (**33**) and (**56**) it is believed that all three share a common precursor, containing a methyl group at C11. The roseophillin (**31**) biosynthetic gene cluster was reported by Kawasaki and co-workers and encodes four genes, *rphG*, *rphG2*, *rphG3* and *rphG4* encoding homologues of RedG.[50] Three of these encode proteins containing two conserved N-terminal $CXHX_7CX_2H$ Rieske cluster binding motif typical of Rieske

oxygenases. RphG is predicted to catalyze the oxidative carbocyclization reaction of 11-methyldodecylprodigiosin (**32**) to prodigiosin R1 (**33**), as expression of *rphG* in a strain of *S. coelicolor* in which *redG* had been deleted furnished a cyclic prodiginine with a C-C bond between C8 of the alkyl chain and C4 of the pyrrole C-ring. The RphG2 and RphG4 enzymes are consequently proposed to be involved in the formation of dechlororoseophillin (**56**) (scheme 1.6).[49,50]




**Scheme 1.6:** The proposed role of RedG homologues in the biosynthesis of roseophillin (**31**).

### 1.3.8.4 Rieske non-heme iron dependent oxygenases in natural product biosynthesis

Rieske oxygenase enzymes have been shown to catalyze a wide range of transformations in the biosynthesis of various specialized metabolites.[53,46] As well as oxidative carbocyclization reactions in prodiginine biosynthesis, N-oxidations, C-hydroxylations and C-C bond desaturations have also been reported (figure 1.15). The PrnD enzyme in pyrrolnitrin (**58**) biosynthesis contains a Rieske [2Fe-2S] cluster and catalyzes the N-oxidation of aminopyrrolnitrin (**57**) to form pyrrolnitrin (**58**).[54] The AuaF enzyme is reported to catalyze the N-hydroxylation in the biosynthesis of the quinolone alkaloid aurachin (**59**).[55] A C-hydroxylation reaction is reported to be catalyzed by Rieske oxygenases PpgN, VioQ and MarC in promysalin (**60**), viomycin (**61**) and methylarcyriarubin (**62**) biosynthesis

respectively.[56–58] C-C bond desaturation reactions are predicted to be catalyzed by the Rieske oxygenases AmbP and MupW in the biosynthesis of ambruticin (**63**) and mupirocin (**64**) respectively.[59,60]

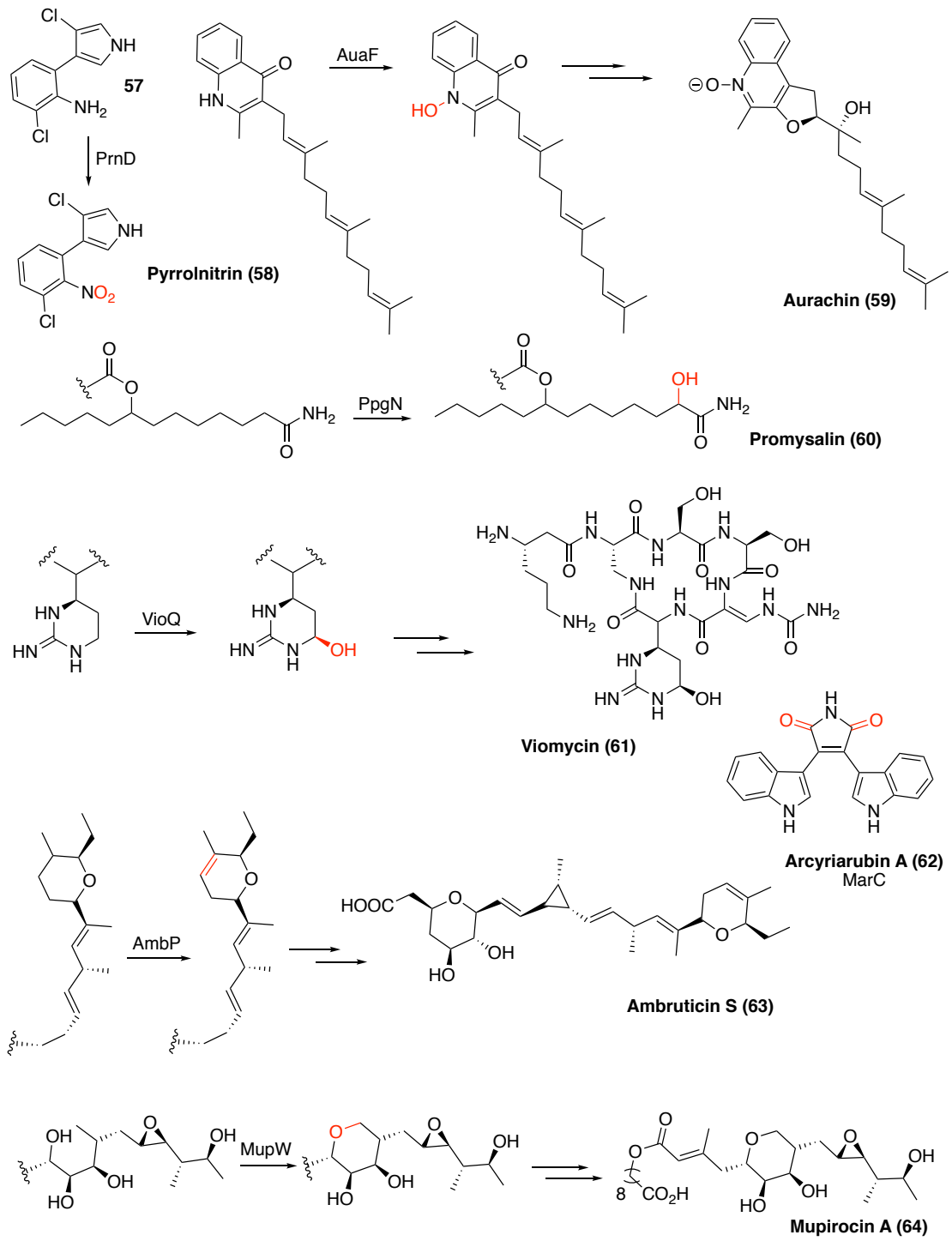



**Figure 1.15:** The predicted roles of Rieske oxygenases in the biosynthesis of specialized metabolites.

## 1.4 Rieske non-heme iron-dependent oxygenases

### 1.4.1 Iron-sulphur clusters

Iron-sulphur clusters are prevalent across nature and are an indispensible constituent of all living cells.[61–64] Proteins that contain such clusters are typically involved in fundamental biological processes such as nitrogen fixation, carbon dioxide fixation and electron transfer.[65] The presence of iron atoms in different oxidation states enabled the detection and characterization of ferredoxins, used in single electron transfer processes, by electron paramagnetic resonance (EPR) spectroscopy.[66,67] Further identification of specific classes of iron-sulphur clusters can be carried out by UV-Vis absorption spectroscopy.[68] Charge transfer between the S-$Fe^{3+}$ bonds in different classes of cluster result in distinct absorbance maxima.[54,60,69]

Typically, Fe ions are ligated by cysteine thiolate ligands and common systems include the [1Fe], [2Fe-2S], [3Fe-4S] and [4Fe-4S] clusters (figure 1.16).[70,71] The principal role of iron-sulphur clusters across nature is to facilitate the transfer of electrons by delocalization across the diffuse iron and sulphide ions. Consequently, [2Fe-2S] clusters can exist in two redox states that facilitate the transfer of one electron, the $[2Fe\text{-}2S]^{2+}$ and the $[2Fe\text{-}2S]^{1+}$ state with a highly reducing electronic potential of between -250 and -420 mV.[72] The Fe-Fe bond distance is typically 2.8 Å and both Fe ions are in the high-spin state.[73] In the all-ferric, oxidized state the two iron centres are antiferromagnetically coupled and are not detectable by EPR (S = 5/2). In the reduced $[2Fe\text{-}2S]^{1+}$ cluster ($Fe^{3+}/Fe^{2+}$ state) the two centres have fixed valencies where the ferrous (S = 2) and ferric (S = 5/2) ions are antiferromagnetically coupled to generate a cluster with net spin of 1 /2 which consequently can be detected by EPR spectroscopy.[74]

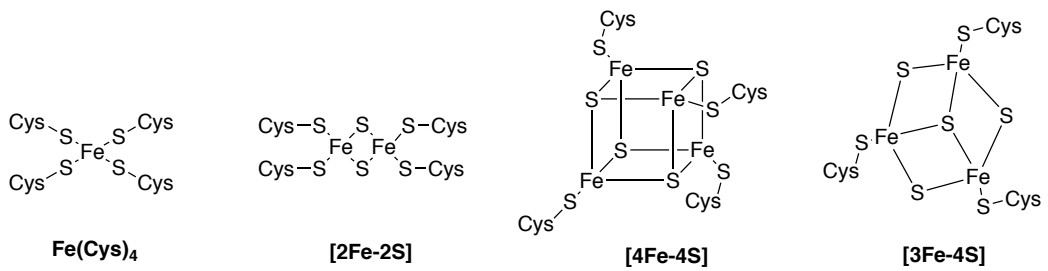



**Figure 1.16:** The different structures of characterized FeS clusters.

### 1.4.1.1 Rieske [2Fe-2S] clusters

Rieske proteins were first characterized as part of photosynthetic processes involving electron transfer, and have since been reported in bacterial oxygenases and ferredoxins.[74] As these proteins are found in all three major kingdoms they are proposed to be among the most ancient of frameworks to have developed during biological evolution. Rieske proteins contain a certain subclass of [2Fe-2S] cluster in which two of the four cysteine ligands are replaced by histidines (figure 1.17).[75,76]

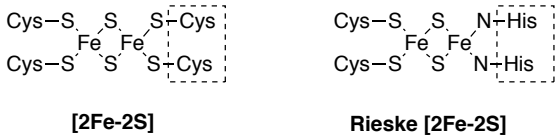


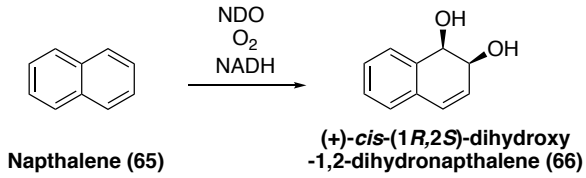

**Figure 1.17:** The structure of the [2Fe-2S] cluster and the Rieske [2Fe-2S] cluster highlighting the differences in the ligands.

The change of ligand has a significant effect on the overall electronic properties of the cluster. The electronic potential becomes positive and ranges between +100 to +400 mV.[72] As a result, Rieske clusters have a redox potential that is approximately 400-700 mV higher than the redox potentials for other [2Fe-2S] proteins. The typical scaffold for proteins containing a Rieske cluster consists of three β-sheets, with one sheet containing two 'loops' that enclose the [2Fe-2S] cluster.[77,63]

Rieske non-heme iron-dependent oxygenases (Rieske oxygenases) typically utilize a reductase that transfers electrons from NAD(P)H through a ferredoxin to the Rieske [2Fe-2S] cluster and subsequently to the non-heme iron centre.[46] Molecular oxygen is consequently activated by the non-heme iron centre and then used to oxidize the substrate of the reaction. A key component in the oxidation of substrates is the facilitation of electron transfer by the Rieske [2Fe-2S] cluster. A well characterized example of this class of enzyme is naphthalene-1,2-dioxygenase.[53,78,79]

### 1.4.2 Napthalene-1,2-Dioxygenase (NDO)

NDO is a multicomponent Rieske oxygenase from *Pseudomonas* sp. NCIB 9816-4 that has been well characterized. The enzyme is responsible for catalyzing oxidation of naphthalene (**65**) to furnish (+)-*cis*-(1*R*,2*S*)-dihydroxy-1,2-dihydronapthalene (**66**) (scheme 1.7).[78]




**Scheme 1.7:** The NDO-catalyzed dihydroxylation reaction of naphthalene (**65**) to furnish **66**.

The crystal structure of NDO shows it is an $\alpha_3\beta_3$ hexamer where the α-subunit contains a Rieske [2Fe-2S] cluster binding domain, similar to that observed in the cytochrome $bc_1$ complex.[78] The N-terminal Rieske domain is comprised of two β-hairpins forming two 'finger-like' structures, into which the Rieske cluster is embedded (Figure 1.18).[80]

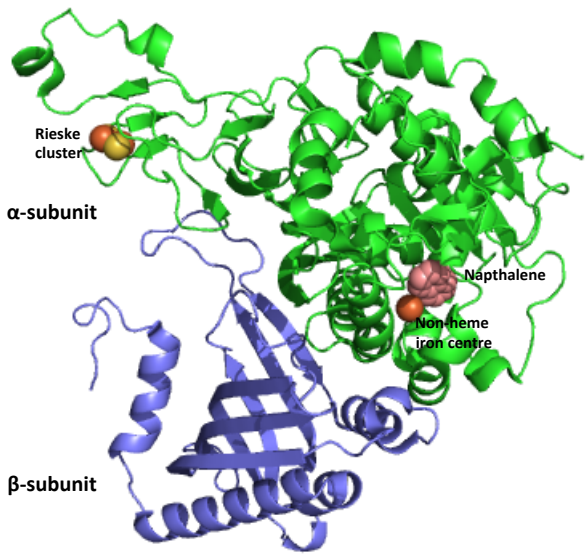



**Figure 1.18:** A crystal structure of naphthalene dioxygenase with naphthalene bound - highlighting the α and β-subunit and location of the Rieske cluster and non-heme iron centre (PDB 1O7G).

The Rieske cluster contains two Fe ions, one that is bound by His83 and His104 and another that is bound by Cys81 and Cys101. The two sulphide ions bridge the two Fe ions to form a flat rhombic structure. The C-terminal domain contains a ferric ion that is bound by residues His208 and His213. This can be reduced to the ferrous form by an external reducing agent. The two metal containing centres are separated by a distance of greater than 40 Å and are thus connected across the interface of adjacent α-subunits through Asp205 (Figure 1.19).[81]

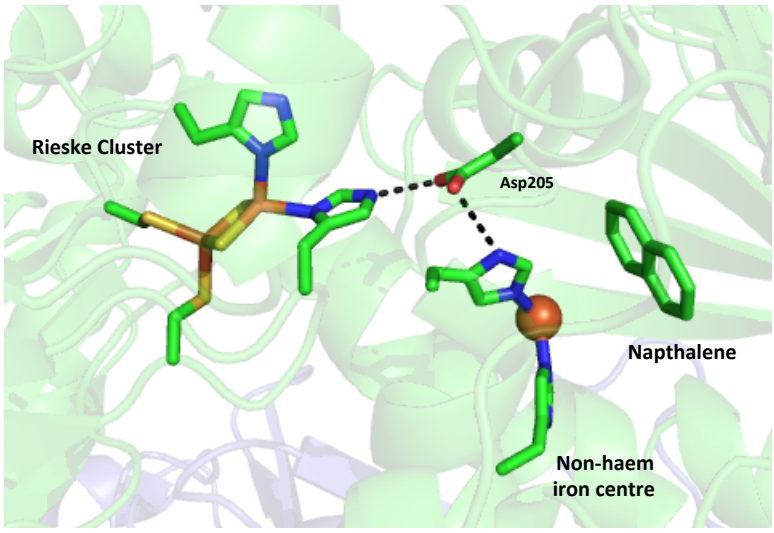
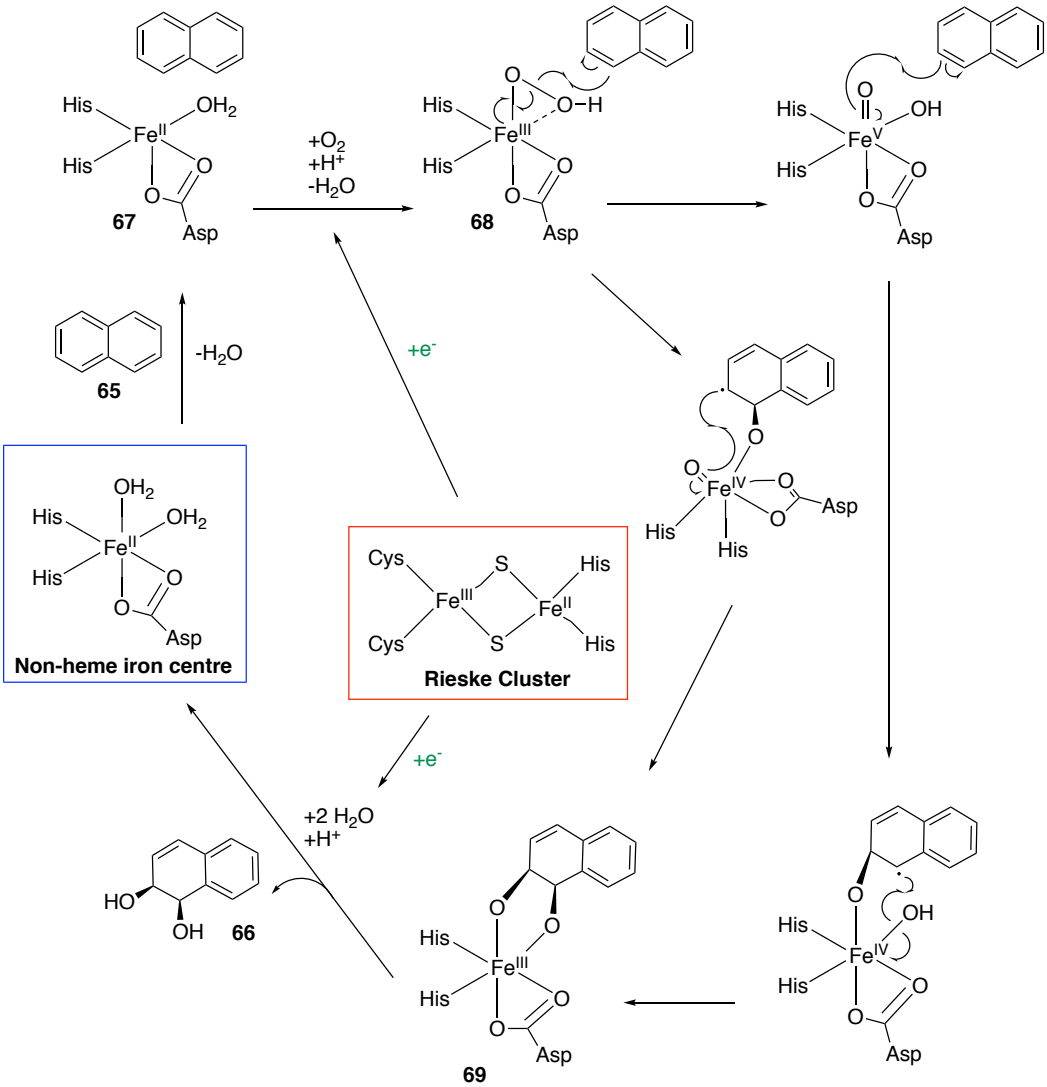



**Figure 1.19:** A crystal structure of naphthalene dioxygenase with naphthalene bound in the active site – highlighting the location of the Rieske cluster, non-heme iron centre and the bridging Asp205 residue that facilitates electron transfer between them (PDB 1O7G).

Further crystallographic studies showed that reduction of the Rieske cluster results in a conformational change that allows it to hydrogen bond to the bridging Asp205 residue.[81] Consequently, the non-heme iron centre coordination geometry changes to facilitate side-on binding of molecular oxygen resulting in formation of a hydroperoxide species.[82] The resting state of the non-heme iron centre was shown to be rigid distorted octahedral geometry and binding of naphthalene (**65**) to the active site resulted in a conformational change to form a five-coordinate square pyramid. Binding of dioxygen to the active site ferrous ion therefore appears to occur after the substrate naphthalene has bound and the geometrical change has occurred.[83]

NDO uses NADH as a co-substrate. Electrons are transferred from NADH to the Rieske cluster *via* a flavin-dependent ferredoxin reductase and a ferredoxin.[84–86] The electrons reach the non-heme iron centre after transfer from the Rieske cluster and are subsequently used for the reduction of dioxygen.[86,87] Lipscomb and co-workers used cyclopropane-containing substrates to probe the nature of the intermediates in NDO-catalyzed oxidations.[79] The results

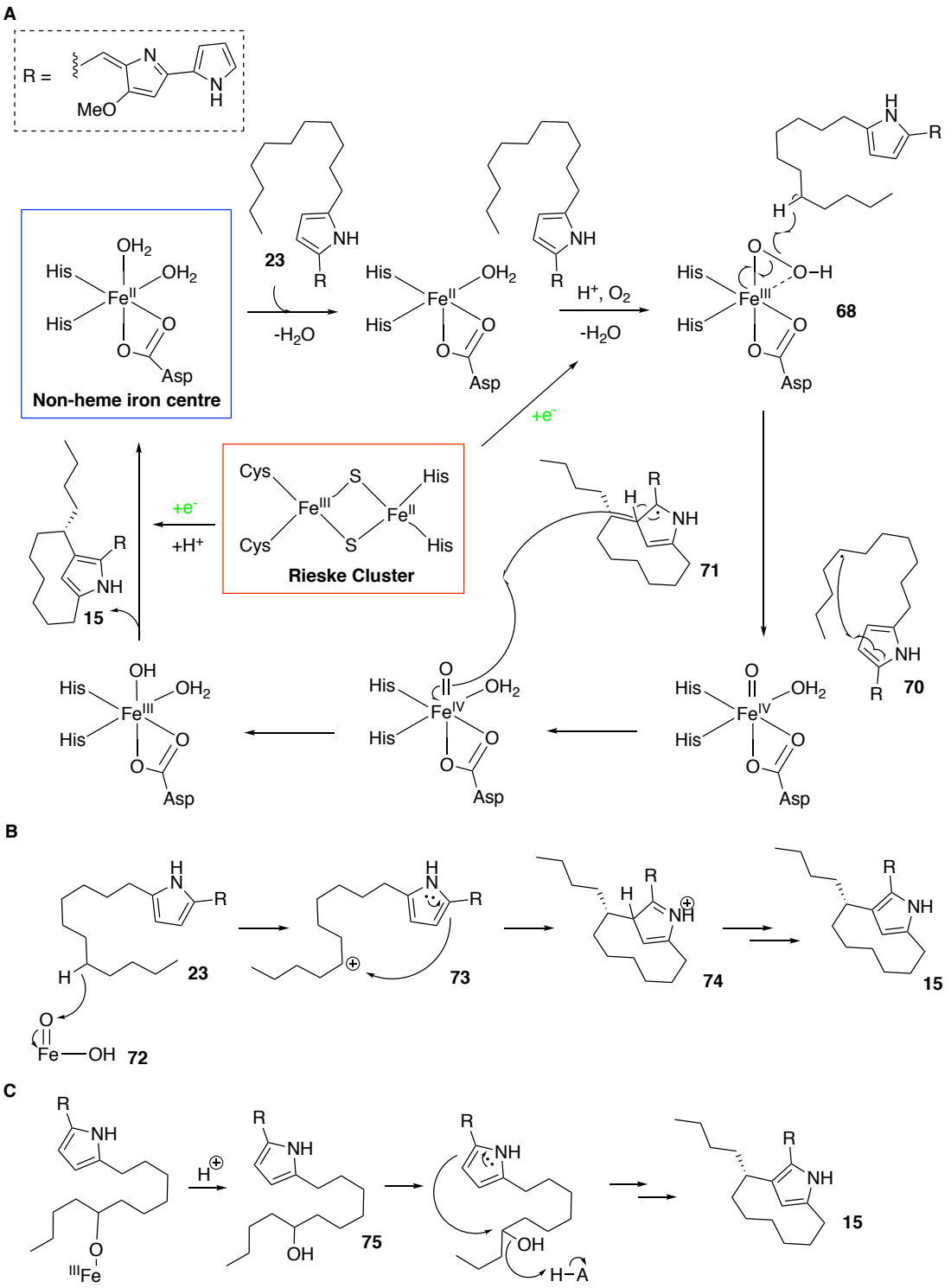
were inconsistent with the formation of cationic intermediates and suggested the involvement of carbon-centred radicals. The proposed NDO catalytic cycle involves binding of naphthalene (**65**) to the active site resulting in the loss of a water molecule to form the octahedral ferrous complex (**67**) (scheme 1.8).[84–86] Dioxygen binds to the complex and an electron is transferred from the Rieske cluster to form a peroxide complex which is protonated to form the corresponding hydroperoxide complex (**68**), of which has been experimentally observed by X-ray crystallography.[82] Homolytic O-O cleavage coupled to oxidation of **65** forms a carbon-centred radical which reacts further to form a ferric-alkoxyhydroxynapthalene complex (**69**). A second electron is transferred from the Rieske cluster to the non-heme iron centre before protonation liberates the dihydroxylated product (**66**).




**Scheme 1.8:** The proposed mechanism of the NDO catalyzed *cis*-dihydroxylation of naphthalene (**65**); The non-heme iron centre and Rieske cluster are highlighted blue and red boxes respectively and electrons donated by NADH are green.[46,84–86]

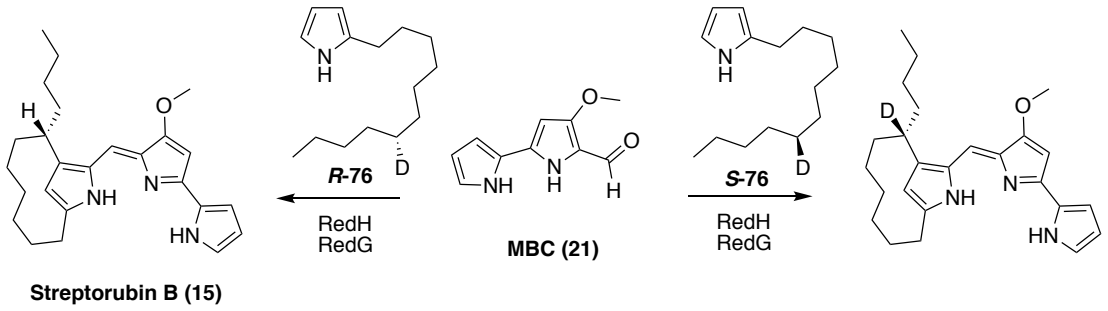
#### 1.4.2.1 The proposed mechanism of RedG-catalyzed oxidative carbocyclization of undecylprodigiosin (23) to streptorubin B (15)

A mechanism for the RedG-catalyzed oxidative carbocyclization reaction has been proposed based upon the well-characterized catalytic mechanism of NDO (**A**, scheme 1.9).[14,46,49] The binding of undecylprodigiosin (**23**) stimulates the loss of a water molecule from the non-heme iron centre. Side-on binding of dioxygen follows substrate binding with subsequent transfer of an electron from the Rieske cluster and protonation generating a Fe(III)OOH species (**68**) Homolytic cleavage of the O-O bond provides a complex that abstracts a proton from C7 of the undecylprodigiosin (**23**) alkyl chain. The resulting radical (**70**) attacks C4 of the proximal pyrrole to form the tertiary radical species (**71**), which is stabilized by the conjugated π-system of the substrate. A second hydrogen abstraction at C4 affords streptorubin B (**15**), which is released from the active site after a second electron is transferred from the Rieske cluster. Potential rearrangement of complex **68** to form an Fe(V)O(OH) species (**72**) is yet to be investigated experimentally and cations **73** and **74** are plausible alternatives to radical intermediates **70** and **71** (**B**, scheme 1.9). Also yet to be fully elucidated is the nature of the cyclization, with rebound of radical species **70** back to the iron-oxo species and formation of a hydroxylated intermediate (**75**) a potential alternative to direct carbocyclization (**C**, scheme 1.9).



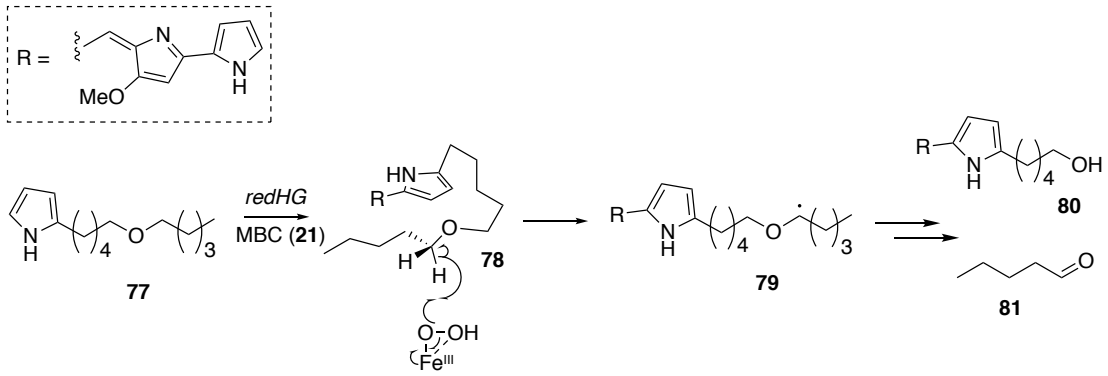

**Scheme 1.9: A.** The proposed mechanism of RedG-catalyzed oxidative carbocyclization of undecylprodigiosin (**23**) to streptorubin B (**15**) based on the well-studied Rieske oxygenase NDO; **B.** Formation of plausible alternative cationic intermediates **73** and **74** formed by a Fe=O(OH) species; **C**. The formation of hydroxylated intermediate **75** as a result of an alternative rebound mechanism.[14,46,49]

The stereochemical course of the RedG-catalyzed carbocyclization reaction that converts undecylprodigiosin (**23**) to streptorubin B (**15**) has been investigated by synthesizing analogues of 2-UP (**28**) in which a deuterium label was stereoselectively inserted into both the *R* and *S* position at C7 of the alkyl chain.[49] Both enantiomers were fed independently, alongside synthetic MBC (**21**), to *S. albus* expressing *redH* and *redG*. Feeding [7'-$^2$H](7'*S*)-2-UP (***S*-76**) resulted in a streptorubin B (**15**) where 92 % of the deuterium label was retained. The feeding of [7'-$^2$H](7'*R*)-2-UP (***R*-76**) resulted in a 77 % loss of the deuterium label in the final streptorubin B (**15**) product (scheme 1.10). The discrepancy between the amount of deuterium lost when feeding 7'*R*-2-UP (***R*-76**) compared to the amount of the deuterium that is retained when feeding 7'*S*-2-UP (***S*-76**) can be accounted for by the primary kinetic isotope effect for the abstraction of the deuterium atom from the corresponding undecylprodigiosin analogue. These results demonstrated experimentally the stereochemical course of RedG, favouring the abstraction of the *pro-R* hydrogen atom C7 to generate streptorubin B (**15**) with 7'*S* absolute stereochemistry. RedG therefore catalyzes oxidative carbocyclization with inversion of stereochemistry at C7 of the undecylprodigiosin (**23**) alkyl chain.




**Scheme 1.10:** The major products from feeding deuterium-labeled 2-UP stereochemical probes to *S. albus* co-expressing *redH* and *redG*.

The site of oxidation in RedG was also investigated.[49] It is possible that the π-system of the 4-methoxypyrrolyldipyrromethene core would be oxidized rather than the alkyl chain. Challis and co-workers investigated the oxidation step by feeding a 2-UP ether analogue (**77**) to *S. albus/redHG* in which an oxygen atom is located geminal to the site of C-H abstraction. The product of this assay was 5-hydroxypentylprodigiosin (**78**), which is consistent with the proposed mechanism of abstraction at C7 of the alkyl chain and not the pyrrole moiety (scheme 1.11). The presence of the oxygen lone pair stabilizes the geminal carbon-centered radical and so reduces the rate of the cyclization of intermediate (**79**). The radical therefore rebounds onto the iron-oxo species to form a hemiketal that collapses on protonation, furnishing (**80**) and pentanal (**81**).




**Scheme 1.11:** Cleavage of the ether bond of **78** by RedG following feeding 2-UP analogue **77** to a *S. albus* heterologous host expressing *redHG*.[49]

Haynes and Withall have extensively studied the substrate tolerance of RedG by feeding 2-UP analogues to a *S. coelicolor* mutant blocked in the biosynthesis of 2-UP (**28**) (figure 1.20).[88,89] It was shown that RedG required a substrate where the alkyl chain was longer than eight carbons in length in order to catalyze cyclization. When the length of the 2-UP hydrocarbon alkyl chain was increased to twelve carbons a larger carbocycle was observed (**82**) and so presumably the increase in chain length hinders the degree of regiocontrol. RedG was also shown to accept 2-UP analogues with a methyl substituent at C10 of the alkyl chain, however moving the methyl group closer to the site of C-H abstraction (**83**) resulted in the formation of a mixture of unsaturated products (**84**) along side the corresponding carbocycle. Inserting an oxygen atom at C3 of the alkyl chain successfully generated the corresponding streptorubin B analogue (**85**), however situating the oxygen adjacent to the site of C-H abstraction (**98**) failed to produce the resulting carbocycle, presumably due to the additional stabilization of the oxygen lone pair lowering the rate of the cyclization.

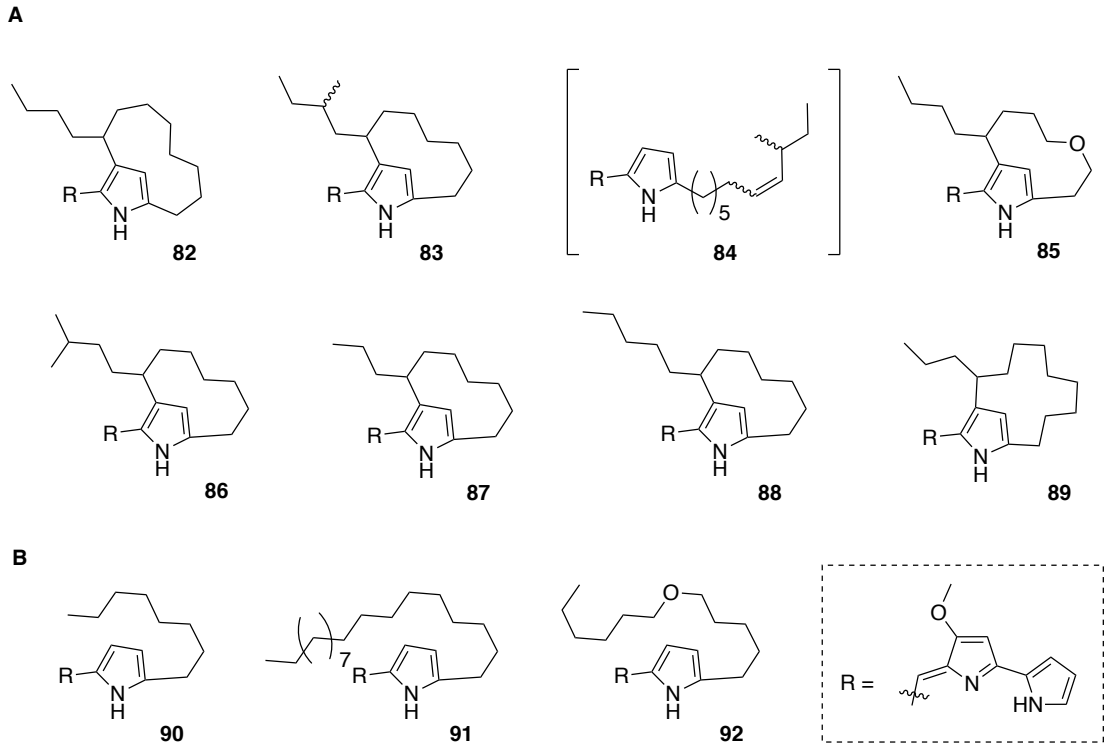



**Figure 1.20: A.** 2-UP analogues accepted by RedG; **B.** 2-UP analogues not cyclized by RedG.

The difference in absolute stereochemistry of streptorubin (**15**) and metacycloprodigiosin (**25**) poses an interesting question into the stereochemical course of cyclizations catalyzed by Rieske oxygenases in prodiginine alkaloid biosynthesis. The mechanism of cyclization of undecylprodigiosin (**23**) to metacycloprodigiosin (**25**) is predicted to proceed *via* the mechanism outlined for RedG catalysis.[46,49] To date, the stereochemical course and mechanism of oxidative carbocyclization of McpG is yet to be investigated by methods *in vivo* or *in vitro*.

## 1.5 Aims of the study

The first aim of this study was to probe the stereochemical course of the McpG-catalyzed oxidative carbocyclization of undecylprodigiosin (**23**) to metacycloprodigiosin (**25**). To do this, we envisaged creating a deletion mutant of a metacycloprodigiosin (**25**) producing strain in which the biosynthesis of 2-UP (**28**) has been abrogated, or expressing *mcpHG* in a heterologous host. We then planned to synthesize analogues of 2-UP (**28**) in which a deuterium is stereoselectively inserted into the *pro-R* and *pro-S* positions at C9 of the alkyl chain. Finally we envisaged feeding these *pro-R* and *pro-S* stereochemical probes to the *in*

*vivo* system and subsequently analyzing deuterium incorporation into the metacycloprodigiosin produced.

The second aim of this study was to further probe the mechanism of oxidative carbocyclization reactions catalyzed by Rieske oxygenases. To achieve this, we envisaged synthesizing several analogues of 2-UP (**28**) as probes to investigate the mechanism of cyclization and the involvement of radical intermediates. We aimed to feed the probes to the afore-mentioned *in vivo* system and analyze the resulting products by high resolution LC-MS. We also aimed to investigate the substrate tolerance of McpH and McpG by synthesizing 2-UP analogues that have been decorated with different functional groups and/ or altered in length of the alkyl chain, and feeding them to the *in vivo* system.

The last aim of this study was to reconstitute the activity of RedG and McpG *in vitro*. To do this, we envisaged cloning both *redG* and *mcpG* into several protein expression vectors. After overproducing the proteins in soluble form, we aimed to investigate whether the [2Fe-2S] cluster was intact using UV-Vis spectroscopy, and if necessary reconstitute the cluster under aerobic or anaerobic conditions. Finally, we aimed to demonstrate that the purified proteins can catalyze the conversion of undecylprodigiosin (**23**) to streptorubin B (**15**)/ metacycloprodigiosin (**25**).

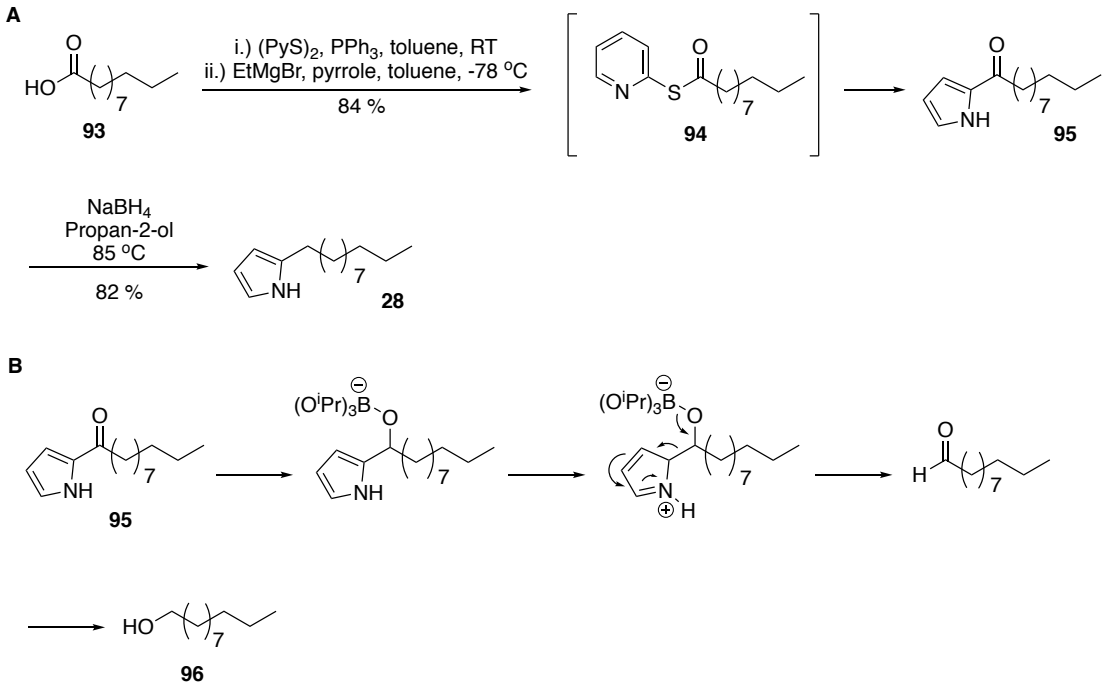
# Chapter 2: Investigating the stereochemical course of oxidative carbocyclization of undecylprodigiosin to metacycloprodigiosin by Rieske oxygenase McpG

## 2.1 Synthesis of 2-UP (28) & MBC (21)

The synthesis of 2-UP (**28**) and MBC (**21**) was undertaken based on a route previously outlined by Dr. Stuart Haynes and Dr. David Withall.[90,88,49]

### 2.1.1 Synthesis of 2-undecylpyrrole (2-UP) (28)

2-Undcylpyrrole (**28**) was synthesized from undecanoic acid (**93**), by converting to the corresponding 2-pyridyl thioester (**94**) and reacting with pyrryl magnesium bromide to afford the 2-acylpyrrole (**95**), which was reduced to the alkyl pyrrole with $NaBH_4$ (scheme 2.1).

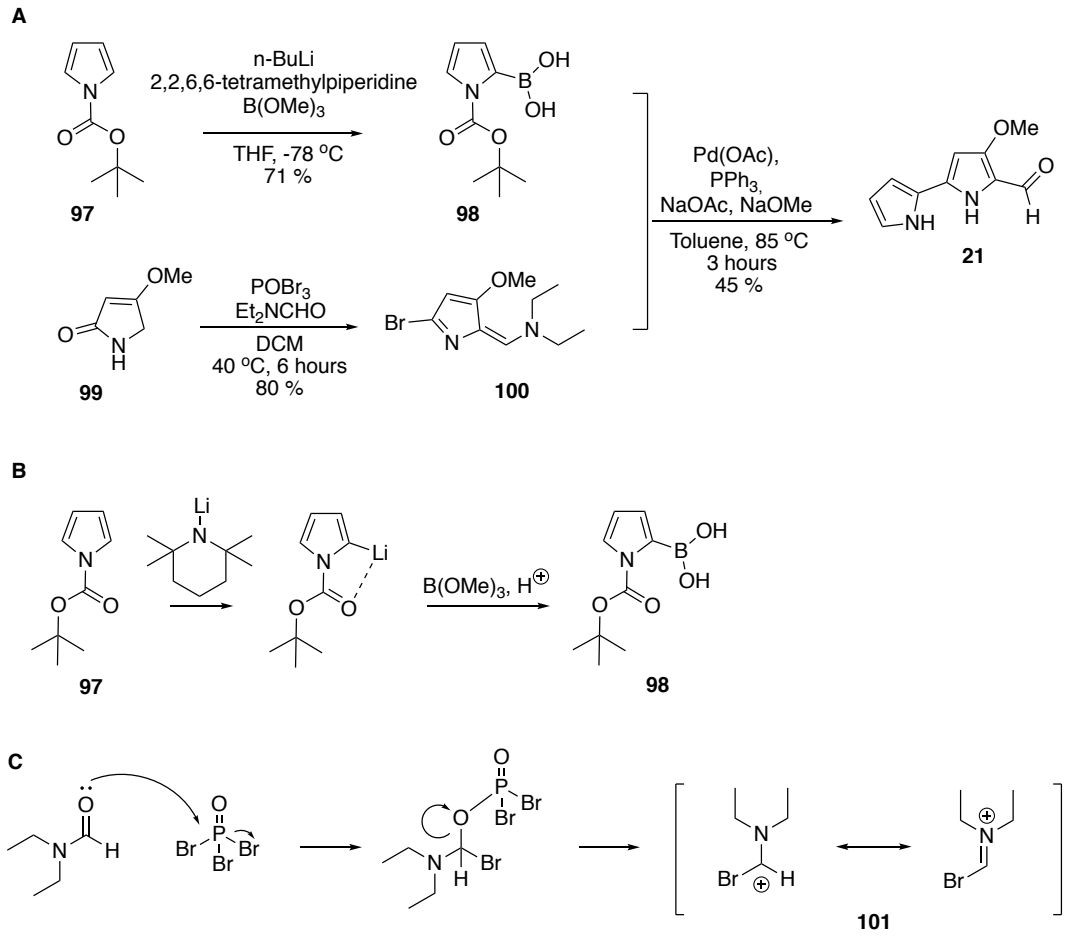



**Scheme 2.1: A.** Synthesis of 2-UP (**28**); **B.** Alternative retro-aldol pathway for sodium borohydride reduction undertaken at temperatures below reflux, affording undecanol (**96**).

### 2.1.2 Synthesis of 4-methoxy-2,2'-bipyrrole-5-carboxaldehyde (MBC) (21)

The boronic acid (**98**) was prepared from the commercially available *N*-Boc-pyrrole (**97**) by a regiospecific lithiation reaction using 2,2,6,6-tetramethylpiperidide and subsequent treatment with trimethyl borate (scheme 2.2). Electron deficient pyrroles are prone to lithiation at the 2-position due to the decrease in $pK_a$ of this proton as a result of the electron-withdrawing nature of the *tert*-butyloxycarbonyl protecting group. The carbonyl group also coordinates to

the lithium ion, further promotoing lithiation at the 2-position (scheme 2.2).[91] Commercially available 4-methoxy-3-pyrrolin-2-one (**99**) was converted to bromo-enamine (**100**) by reaction with two equivalents of the Vilsmeier-Haack salt (**101**), generated *in situ* through the reaction of phosphorus (V) oxybromide and diethylformamide. A Suzuki cross-coupling reaction between boronic acid (**98**) and bromo-enamine (**100**) generated the final MBC (**21**) product.




**Scheme 2.2: A.** The synthesis of MBC (**21**); **B** Directed lithiation of (**97**) to furnish boronic acid (**98**); **C.** Mechanism for *in situ* formation of the Vilsmeier-Haack salt (**101**).

## 2.2 Construction of an *in vivo* system to probe McpG activity – *Streptomyces albus* co-expressing *mcpH* and *mcpG*

With the synthetic late-stage intermediates 2-UP (**28**) and MBC (**21**) in hand, attention was turned to the development of an *in vivo* system for investigating the McpG-catalyzed reaction. Previous work in the Challis group utilized a *S. albus* heterologous host to co-express *redH* and *redG* to investigate the mechanism and stereochemical course of oxidative carbocyclization catalyzed by RedG.[48,89,92] It was observed that growing the strain on solid media overlaid with a sterile semi-permeable membrane generated the highest production levels of streptorubin B (**15**).[89] Synthetic material was found to be incorporated competently when a methanolic solution of 2-UP (**28**) and a DMSO solution of MBC (**21**) was pre-mixed before feeding directly onto the strain after three days of growth.[89] It was therefore envisaged that a similar method could be used to probe the stereocontrol of McpG.

### 2.2.1 *Streptomyces albus* as a heterologous host for pOSV556t*mcpHG*

*S. albus* J1074 has been used frequently in the literature as a host for the expression of single genes or entire biosynthetic gene clusters.[93–96,49] A pOSV556t/*mcpHG* vector had previously been prepared in the Challis group by Dr. Rebin Salih.[48] This is an integrative plasmid that drives the expression of cloned genes using the constitutive promoter *ermE*$^*$ and has resistance to ampicillin (for *E. coli* selection) and hygromycin (for *Streptomyces* selection). The plasmid also contains an *oriT* for conjugal transfer from *E. coli* to the *Streptomyces species*. The construct was used to transform *E. coli* ET12567 and the subsequent transconjugants were inoculated and used to transfer the plasmid DNA to *S. albus* by triparental conjugation aided by *E. coli* ET12567/pUB307.

### 2.2.2 Feeding synthetic 2-UP (28) and MBC (21) to *Streptomyces albus* expressing *mcpHG*

Solid R5 medium was overlaid with a sterilized semi-permeable membrane. Freshly prepared transconjugants of *S. albus*, expressing *mcpHG*, were plated directly onto the membrane and incubated for three days at 30 $^{o}$C until visible mycelia growth was observed. Synthetic 2-UP (**28**) (1 mg) and MBC (**21**) (1 mg) were fed (using the procedure outlined in section 6.5.6.3).

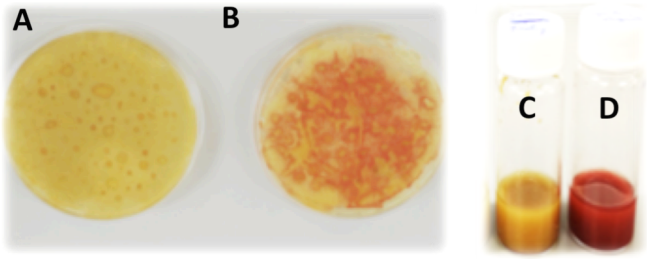



**Figure 2.1: A.** *S. albus* WT fed with synthetic 2-UP (**28**) and MBC (**21**); **B.** *S. albus* co-expressing *mcpHG* fed with synthetic 2-UP (**28**) and MBC (**21**); **C.** Mycelial extract of A; **D.** Mycelial extract of B.

After two days of incubation, red pigmentation was observed, which is indicative of prodiginine production (figure 2.1). The mycelium was collected and metabolites were extracted from it using acidified methanol, following the procedure outlined in section 6.5.6.4. The resulting mycelial extracts were analysed by high resolution LC-MS.

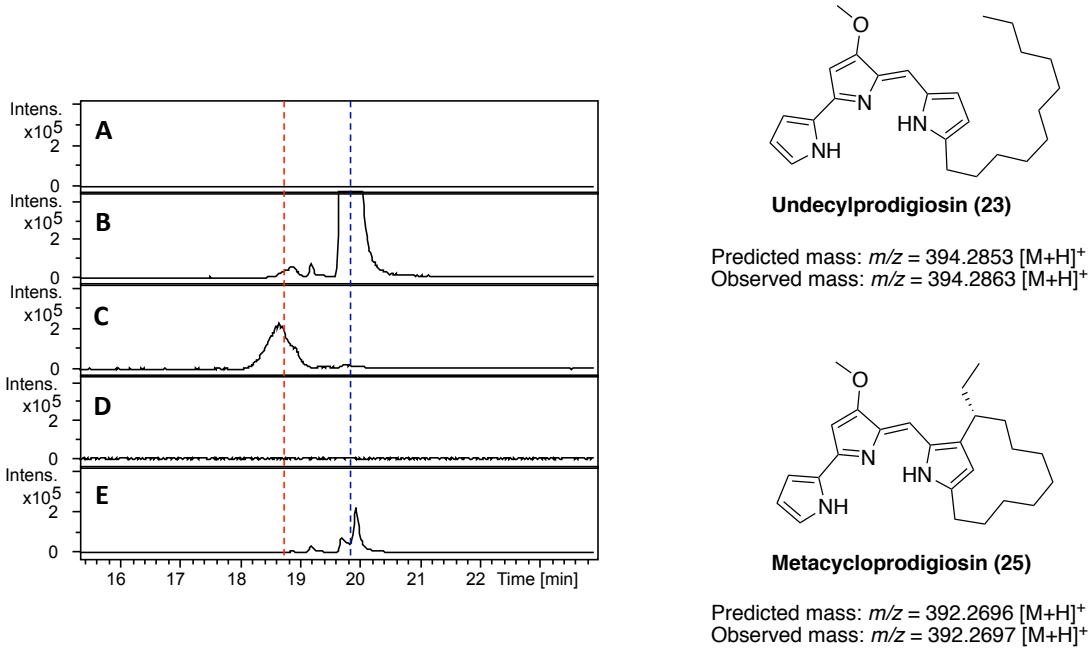



**Figure 2.2:** Extracted ion chromatograms (EICs) at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) (blue) and metacycloprodigisoin (**25**) (red) respectively from LC-MS analysis of the mycelia extracts of **A.** *S. albus*/pOSV556t/*mcpHG*; **B.** *S. albus*/pOSV556t/*mcpHG* + 2-UP (**28**) + MBC (**21**); **C.** *S. longispororuber* WT; **D.** *S. albus* WT; **E.** *S. albus* + 2-UP (**28**) + MBC (**21**).

The LC-MS chromatograms illustrate that feeding synthetic 2-UP (**28**) and MBC (**21**) to *S. albus* expressing *mcpHG* produced mainly undecylprodigiosin (**23**) along with a trace of metacycloprodigiosin (**25**) (figure 2.2, trace **B**). Small amounts of undecylprodigiosin (**23**)

were observed in the negative control (figure 2.2), resulting from spontaneous condensation of 2-UP (**28**) and MBC (**21**). This condensation occurs due to the acidic extraction conditions. The peak intensity of the carbocyclic product **25** relative to undecylprodigiosin (**23**) is significantly lower than was observed in an analogous assay with *redHG*.[92] Attempts were made to increase the amount of **25** produced by altering the concentration of 2-UP (**28**) and MBC (**21**) precursors, incubation times and production media. However, these failed to increase the amount of metacycloprodigiosin (**25**) produced relative to undecylprodigiosin (**23**). The poor conversion of (**23**) to (**25**) has been observed previously in the Challis group, where the activity of RedG has been shown to decrease over time.[89] A potential explanation for these observations is a self-inactivation or 'suicide inactivation' mechanism, which has been reported for several non-heme iron dependent oxygenases.[97,98] Because RedG is constitutively expressed from the *ermE*$^*$ promoter, in absence of substrates the enzyme may generate superoxide or hydroperoxide, which is damaging to the cell. Thus spontaneous mutants in which *redG* expression levels are lower would be at a competitive advantage. Due to the difficulties with the optimization of the feeding experiment involving *S. albus* expressing *mcpHG*, an alternative expression system was sought.

To probe streptorubin B (**15**) biosynthesis both Withall and Haynes utilized a mutant *S. coelicolor* strain in which *redL* was replaced with an apramycin resistance cassette.[88] RedL is a polyketide synthase (PKS) that plays a key role in the biosynthesis of 2-UP (**28**).[41] Deletion of this gene abrogates the production of 2-UP (**28**) and consequently streptorubin B (**15**). The use of a deletion mutant was envisaged to be a suitable alternative to *S. albus* expressing *mcpHG* as it would result in more controlled expression of *mcpG* relative to when the gene is under the control of the *ermE*$^*$ promoter.

Comparison of the gene clusters that direct the biosynthesis of streptorubin B (**15**) and metacycloprodigiosin (**25**) show that the *redL* and *redK* homologues in the *S. longispororuber mcp* cluster are fused into a single *mcpLK* gene (figure 2.3). Deletion of *mcpLK* from the genome of *S. longispororuber* was not possible as the strain was found to be genetically intractable. It was hypothesized that cloning of the entire *mcp* cluster from *S. longispororuber* into a single vector and expression in a genetically tractable host would provide an alternative approach to furnishing an *mcpLK* deletion mutant. Cloning of the *mcp* cluster using Transformation-associated recombination (TAR) was therefore undertaken.

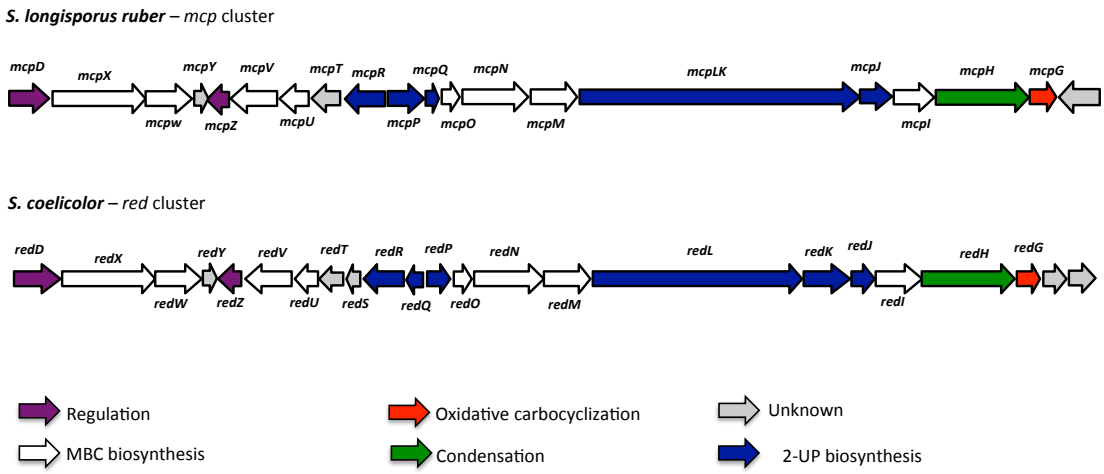



**Figure 2.3:** Comparison of the *red* cluster responsible for streptorubin B (**15**) biosynthesis and the *mcp* cluster responsible for metacycloprodigiosin (**25**) biosynthesis.

## 2.3 Direct cloning of the metacycloprodigiosin biosynthetic gene cluster by transformation-associated recombination (TAR) using *Saccharomyces cerevisiae* VL6-48N

### 2.3.1 The TAR cloning procedure

TAR is a PCR-independent tool used to selectively capture fragments of genomic DNA of size up to 250/300 kb from complex genomes onto a yeast-bacterial shuttle vector.[99] *Saccharomyces cerevisiae*, or 'baker's yeast,' is a unicellular fungus and TAR exploits the high level of natural homologous recombination.[100,101] This phenomenon was first demonstrated by Botstein and co-workers and subsequently used as a method for *in vivo* plasmid construction, as an alternative to the traditional method of *in vitro* DNA ligation.[100] A capture vector is constructed using a yeast centromere selective marker and capture arms or 'hooks' of sizes between 50-500 bp, derived from flanking sequences of the genomic DNA fragment of interest.[102] The constructed capture vector is linearized and co-transformed with fragmented native genomic DNA; and the segment of interest is selectively cloned by recombination between the recombinogenic ends in the capture vector and complementary sequences flanking the segment of interest in the genomic DNA (figure 2.4).[100]

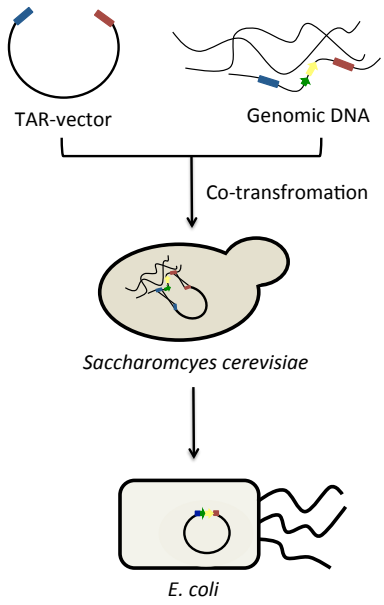



**Figure 2.4:** Schematic of the TAR cloning procedure for capturing large DNA fragments from genomic DNA onto a single vector.

#### 2.3.1.1 TAR cloning applications

TAR cloning has previously been used for the isolation of specific regions and genes from the genomes of humans, primates, mice and other organisms.[99] Through transfection of mammalian cells the function of the BRCA1 (breast cancer 1),[103] HPRT (hypoxanthine phosphoribosyl transferase),[104] KAI1 tumour-suppressor and TEY1 metastasis-suppressor genes were confirmed after capturing them using TAR cloning in yeast.[104,105] The KAI1 genes and MUC2 (mucus/gel-forming protein coding gene associated with gastric cancer) genes are both toxic to bacterial cells, so TAR cloning in yeast also facilitated subsequent sequencing, highlighting errors in the draft genome sequence that resulted from inefficient cloning in *E. coli*.[100] TAR cloning was also used in the final phases of the human genome sequencing project, contributing to closing gaps in DNA sequences by using sequence information of the flanking contigs.

#### 2.3.1.2 TAR cloning for heterologous expression of natural product biosynthetic gene clusters (BGCs)

Moore and co-workers reported the design of a TAR cloning vector called pCAP01 and used it to capture a silent 67 kb BGC that resulted in the identification of the novel antibiotic taromycin A (**102**) as it's metabolic product (figure 2.5).[106] The vector is maintained as a single copy in yeast cells and multiple copies in *E. coli*, due to the presence of a pUC *ori* element.[107] The vector confers neomycin/kanamycin resistance through an *aph(3)II* gene and is able to integrate the cloned cluster into the chromosome of heterologous Actinobacteria by intergeneric conjugal DNA transfer from *E. coli* using the φC31 integration element. This plasmid has since been used in the Moore group for the capture of the BGC responsible for the production a group of lipopeptides based on the alterochromide structure (**103**) from *Pseudoalteromonas* and in the capture and heterologous expression of a type II PKS BGC from the marine acinomycete *Salinispora pacifica*, resulting in the production of enterocin (**104**) in *S. lividans* and *S. coelicolor* M1146 (figure 2.5).[108] Thus, efforts were focused on the construction of a TAR capture vector containing two 'capture arms' corresponding to regions flanking the *mcp* gene cluster in *S. longispororuber*.

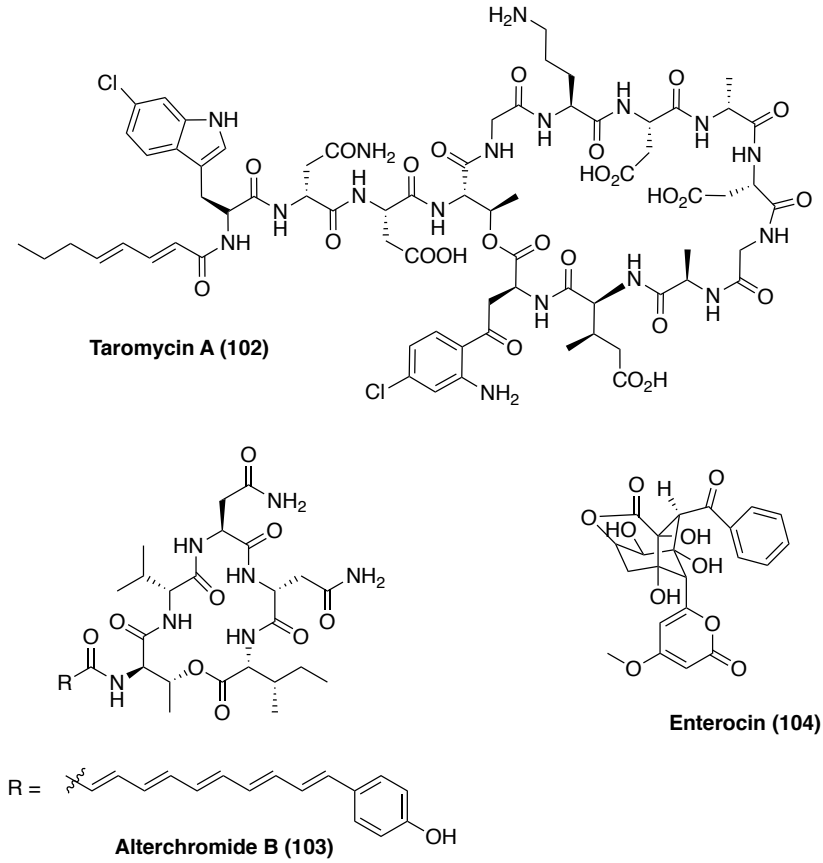



**Figure 2.5:** Metabolites produced by heterologous expression of their corresponding biosynthetic gene clusters (captured by TAR) in a heterologous host.

### 2.3.2 Construction of a TAR vector for capturing the metacycloprodigiosin biosynthetic gene cluster from *Streptomyces longispororuber*

Dr. Dani Zabala and Dr. Chuan Huang (Challis group post-doctoral researchers) further optimized the TAR vector and capture method developed by Moore and co-workers (figure 2.6). The plasmid created in the Moore lab, PCAP01 originally used capture arms that were 1 kb in size. A high number of false-positives drove the development of a second-generation plasmid, PCAP03, which used smaller 50 bp sized capture arms and also a counter-selective marker, *URA3*, which encodes for orotidine-5'-phopsphate decarboxylase (ODCase). The marker converts 5-fluoroorotic acid (5-FOA) into the toxic 5-fluorouracil, only when 5-FOA is added to the growth media. Strains that have excised the *URA3* marker *in vivo* by homologous recombination are subsequently selected for on 5-FOA supplemented media.[109,110] When using PCAP03, the number of false-positives decreases but the cloning efficiency is also lowered. Challis and co-workers relocated the *URA3* gene (under the control of the alcohol dehydrogenase promoter pADH) between the cloning locations of the two capture arms, and also used larger 1 kb sized flanking regions, for more efficient cluster capture. For clarity, the plasmid used in this work will be referred to as pCAP1000 (figure 2.6).

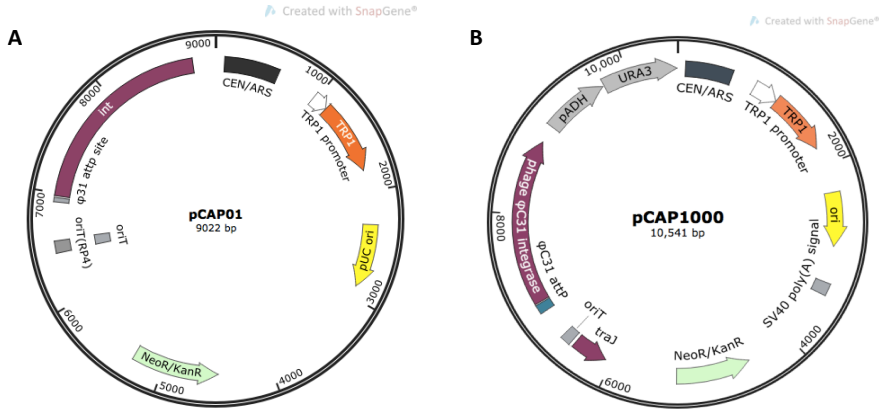



**Figure 2.6: A.** A SnapGene® created plasmid map for the TAR vector developed by the Moore group; **B.** A SnapGene® created plasmid map for the TAR vector developed by the Challis group and subsequently used in this work.

To begin constructing the TAR capture vector, 1 kb sections of DNA flanking the metacycloprodigiosin (**25**) biosynthetic gene cluster were chosen for amplification by PCR (conditions outlined in section 6.5.4.5). Flanking locations were chosen based on desirable GC content and also to ensure that all genes of the *mcp* cluster are present.

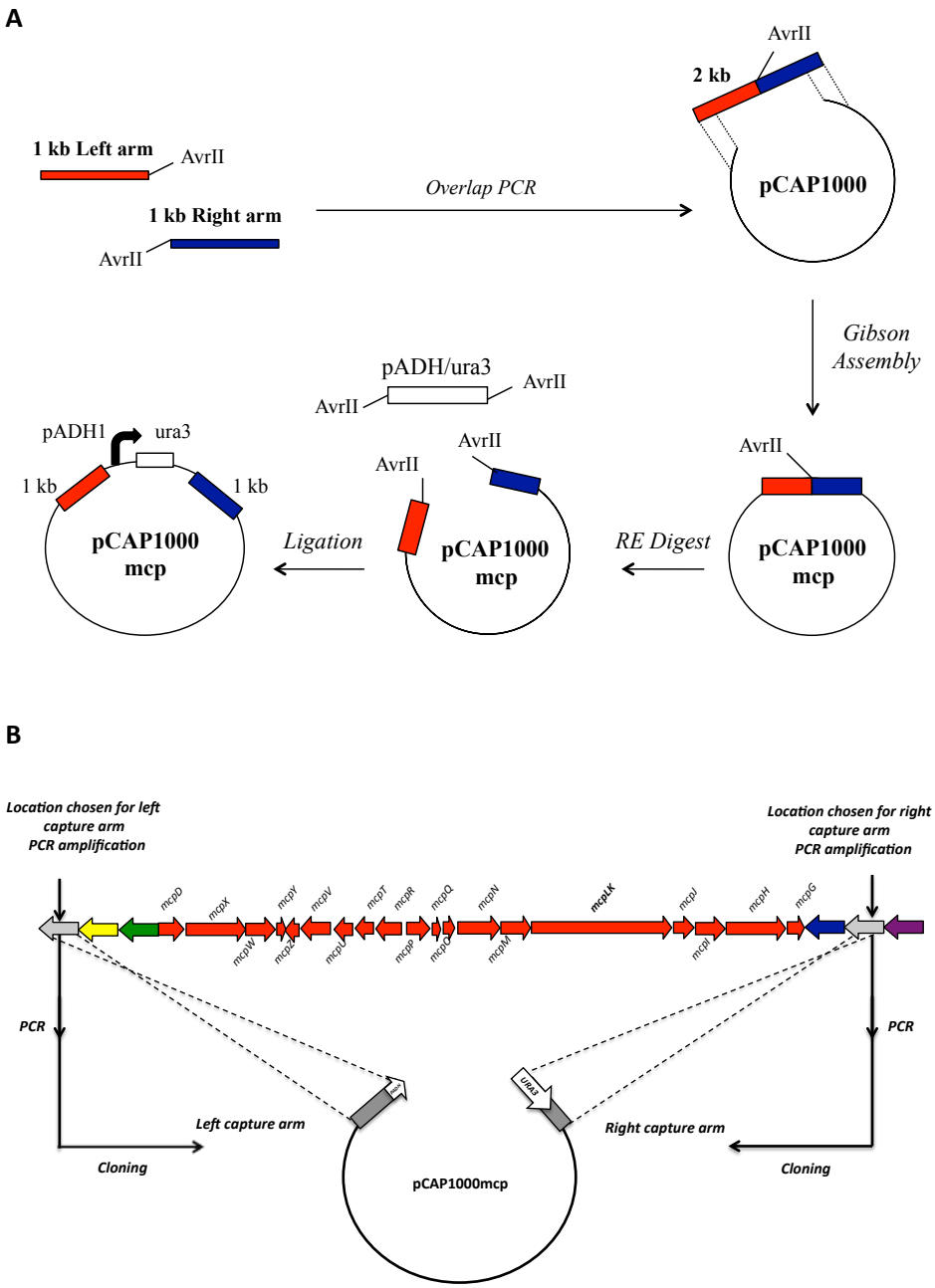



**Figure 2.7: A.** The strategy for the cloning and construction of the target capture vector; **B.** The strategy for capturing the *mcp* cluster (red) from the genome of *S. longispororuber*; additional captured genes are located either side of the cluster; flanking regions that have been chosen for PCR amplification indicated by arrows.

The left capture arm was amplified by PCR such that the 3' end contained an *Avr*II restriction site and the right capture arm were amplified by PCR such that the 5' end of the fragment contained an *Avr*II restriction site to allow for ligation of the *URA3* insert at a later stage (figure 2.7). The resulting PCR products of correct molecular size were purified by agarose gel electrophoresis (section 6.5.5.6) and subsequently used in an over-lap PCR to generate a

single 2 kb product. The overlap primers included 25 bp regions homologous to pCAP1000 to facilitate Gibson assembly.

The resulting purified PCR product was cloned into the pCAP1000 vector using Gibson assembly and positive clones were confirmed by DNA sequence analysis. Dr. Chuan Huang had previously isolated the *pADH/URA3* insert and sub-cloned it into the pJET1.2 vector. The insert was digested (*Avr*II) and directly ligated into pCAP1000 (*Avr*II). The target TAR capture vector containing the capture arms flanking the *pADH/URA3* marker was confirmed by restriction digestion and agarose gel electrophoresis (figure 2.8) (and also DNA sequencing).

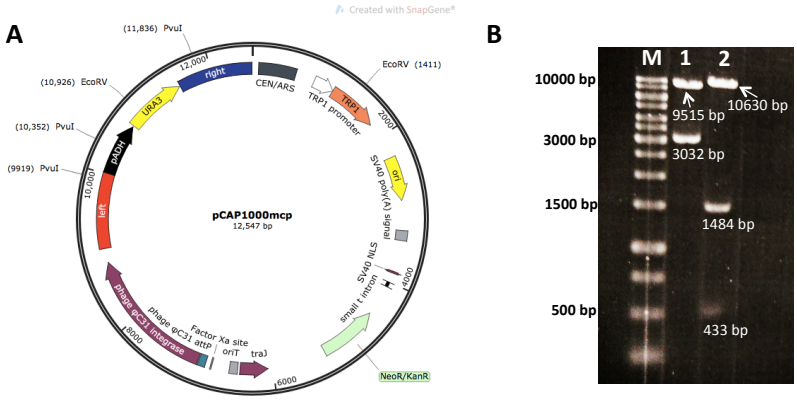



**Figure 2.8: A.** SnapGene® plasmid map of the TAR vector constructed for the capture of the metacycloprodigiosin biosynthetic gene cluster (pCAP1000mcp); **B.** An agarose gel of pCAP1000mcp vector digested using restriction enzymes where M = molecular size marker; lane 1 = *Eco*RV (9515, 3032 bp); lane 2 = *Pvu*I (10630, 1484, 433 bp).

### 2.3.3 Capture of the metacycloprodigiosin biosynthetic gene cluster and heterologous expression

#### 2.3.3.1 Co-transformation of pCAP1000-mcp capture vector and *Streptomyces longispororuber* genomic DNA in yeast

With the TAR pathway-specific vector (containing the desired homology arms that correspond to flanking regions of the *mcp* cluster) in hand, attention was turned to the yeast spheroplasting and transformation procedure. For successful capture, high quality genomic DNA and freshly prepared yeast spheroplasts are required.

The pCAP1000-mcp vector contains a single cutting site (*Pme*I) situated within the *pADH/URA3* genes and this was used to linearize the vector. *Streptomyces longisporus ruber* genomic DNA was extracted and was digested by two restriction enzymes, *Hin*dIII and *Stu*I. This was performed to create smaller DNA fragments to facilitate capture of the *mcp* cluster.

Yeast spheroplasting was undertaken following the procedure outlined in section 6.5.7.1. An overnight yeast preculture was used to inoculate 50 mL of fresh YPD media and grown for 6 hours to an $OD_{600}$ = 0.76. Leaving the cells in sorbitol overnight at 4 $^{o}$C allowed for osmotic stabilization, thus increasing the susceptibility to spheroplasting. Cells were harvested and treated with Zymolase-20T solution (an enzyme from *Arthrobacter luteus*) at 30 $^{o}$C for 40 minutes to furnish spheroplasts with an $OD_{600}$ = 0.42; relative to an $OD_{600}$ = 0.02 of the reference solution (section 6.5.7.2).[111]

The transformation procedure was undertaken following the methods outlined in section 6.5.7.3 to provide 200 µL yeast cell suspensions that were primed for the addition of linearized capture vector and fragmented genomic DNA. Linearized vector and digested gDNA were co-transformed and the resulting cell suspension was added directly to melted top agar media at 60 $^{o}$C; before pouring onto solid bottom agar media supplemented with 5-FOA and incubated at 30 $^{o}$C for one week. Yeast colonies were visible after approximately 72 hours and selected on fresh plates containing bottom agar media. Primers that targeted the central region of the metacycloprodigiosin (**25**) biosynthetic gene cluster were used to screen for successful homologous recombination with *S. longispororuber* genomic DNA.

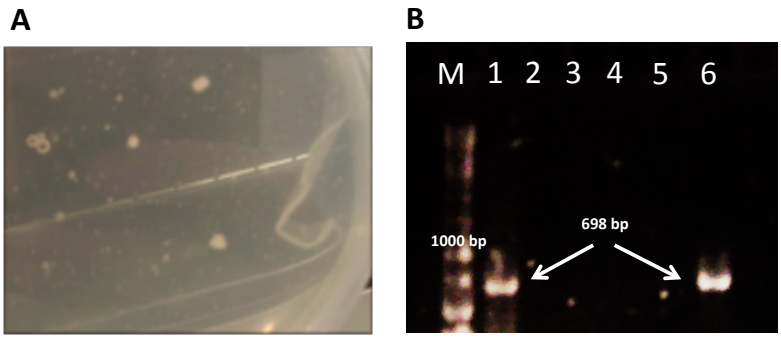



**Figure 2.9: A.** Yeast colonies present at 72 hours post transformation; **B.** Agarose gel following colony PCR of resulting transformants to screen for the presence of the mcp cluster; M = molecular size marker, lane 1 = PCR from using *S. longispororuber* gDNA; lanes 2-6 = PCR from yeast transconjugants.

The transformation procedure provided five yeast transformants to be screened by colony PCR; with one colony generating the expected band (figure 2.9). Negative colonies most likely arose from the mis-capture of other DNA regions or unlinearized capture vector. The plasmid containing the *mcp* cluster will be referred to as pCAP1000-*mcp*BGC. The yeast colony that contained pCAP1000-*mcp*BGC was cultured and the plasmid DNA extracted from the subsequent yeast cells using the procedure summarized in section 6.5.7.6. The resulting plasmid DNA was transformed in electrocompetent *E. coli* TOP10 cells for subsequent culturing and replication of the plasmid.

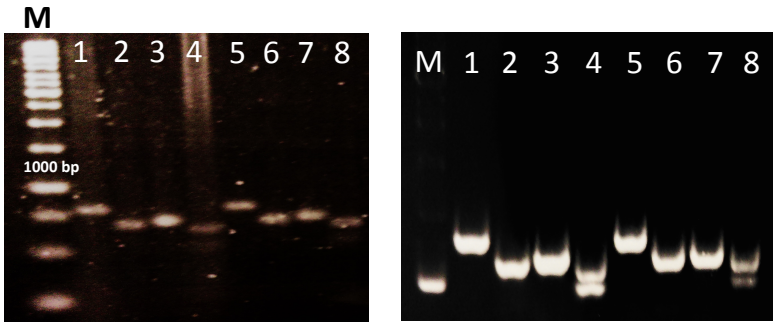



**Figure 2.10**: (*left*) Agarose gels highlighting DNA bands resulting from screening pCAP1000-*mcp*BGC for the presence of the entire gene cluster by PCR; M = molecular size marker; lanes 1-4 = DNA bands resulting from using *S. longispororuber* gDNA; lanes 5-8 = DNA bands from using pCAP1000-*mcp*BGC plasmid DNA; (*right*) Agarose gel highlighting PCR products arising from primers designed to screen for four sections of the *mcp* cluster from DNA extracted from *E. coli* ET12567 transconjugant; M = molecular size marker; lanes 1-4 = PCR products from using *S. longispororuber* gDNA as template DNA; lanes 5-8 = PCR products using pCAP1000-*mcp*BGC as template DNA.

The presence of the metacycloprodigisoin (**25**) biosynthetic gene cluster on the vector was confirmed by PCR and by restriction digest (figure 2.10). Four sets of screening primers were designed to amplify regions of DNA that span the gene cluster with *S. longispororuber*

genomic DNA used as a positive control. The sizes of the PCR products are identical from pCAP1000-*mcp*BGC and the genomic DNA positive control. The presence of the entire metacycloprodigiosin (**25**) biosynthetic gene cluster captured by TAR cloning onto a single vector was consequently confirmed.

With the metacycloprodigiosin (**25**) biosynthetic gene cluster successfully captured, attention was turned to establishing heterologous expression of the cluster. The plasmid was shuttled through *E. coli* ET12567 to *S. albus* and *S. coelicolor* M145 hosts by triparental conjugation aided by the helper strain *E. coli* ET12567/pUB307 (section 6.5.6.2). The plasmid DNA was used to transform electrocompetent *E. coli* ET12567 by electroporation before triparental conjugal transfer to *Streptomyces* strains. The presence of pCAP1000-*mcp*BGC in subsequent transformants was again confirmed by PCR.

#### 2.3.3.2 Production of the metacycloprodigiosin metabolite by expression of the *mcp* gene cluster in a *Streptomyces* heterologous host

Multiple *Streptomyces* host strains and different media were screened to find optimal metacycloprodigiosin (**25**) production conditions. The pCAP1000-*mcp*BGC plasmid was transferred to both *S. coelicolor* M1154 and *S. albus* by triparental conjugation. *S. coelicolor* M1154 was chosen because the strain is a Δ*red* mutant of *S. coelicolor* M145, a native producer of streptorubin B (**15**) containing the *red* cluster as previously mentioned. The strain has previously had the ability to produce prodiginine compounds and so incorporation of a prodiginine biosynthetic gene cluster will likely result in production of metacycloprodigiosin (**25**). Successful production of the prodiginine metabolite was easily observed due to the red-pigmentation of the subsequent mycelium (figure 2.11).

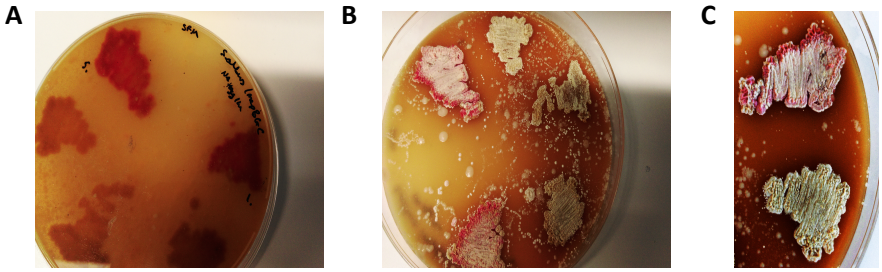



**Figure 2.11:** Images of *S. albus* transconjugants patched onto solid SFM media and incubated at 30 °C for 2 weeks; **A.** Under-plate view of prodiginine producing *S. albus* transconjugant; **B.** Front view showing two transconjugants producing a prodiginine metabolite (pink) and three non-producing (white); **C.** An image of a prodiginine producing *S. albus* transconjugant (pink) and a non-producing transconjugant (white).

*Streptomyces* transconjugants of each strain were cultured on solid SFM and incubated at 30 °C for two weeks. None of the *S. coelicolor* M1154 transconjugants showed evidence of red-pigmented metabolite production on either media. This may be a result of inefficient incorporation of the *mcp* gene cluster into the genome during conjugal transfer. In contrast, *S. albus* transconjugants displayed deep red pigmentation after two weeks of incubation on solid SFM media on both media (figure 2.12). A red colony was selected and grown on SFM plates and the spores collected. The strain will be called *S. albus*/pCAP1000-*mcp*BGC.

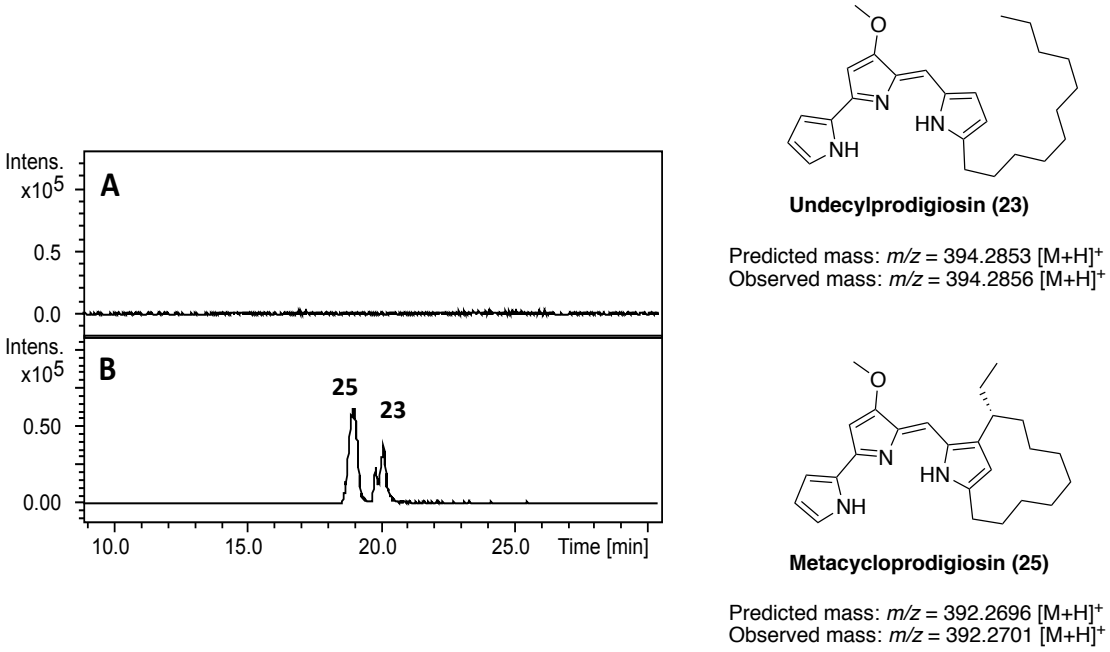



**Figure 2.12:** EICs at *m/z* = 392.27 ±0.01; 394.29 ±0.01 corresponding to the calculated $[M+H]^+$ of metacycloprodigisoin (**25**) and undecylprodigiosin (**23**) respectively from LC-MS analysis of the mycelial extract of; **A.** *S. albus* WT; **B.** *S. albus*/pCAP1000-*mcp*BGC, incubated at 30 °C on solid SFM media for two weeks.

HR LC-MS analysis of the mycelial extract of the *S. albus*/pCAP1000-*mcp*BGC grown on solid SFM media for two weeks at 30 °C confirmed the production of prodiginine metabolites undecylprodigiosin (**23**) (~20 minutes) and metacycloprodigiosin (**25**) (~19 minutes) as a result of heterologous expression of the *mcp* cluster in *S. albus*. With the production of the target metabolite **25** confirmed, attentions were focused on further optimizing the production levels.

### 2.3.3.3 Optimizing the production of metacycloprodigiosin by *S. albus*/pCAP1000-*mcp*BGC

It has been previously demonstrated in the Challis group that the use of a sterile semi-permeable membrane overlaid onto solid media prior to plating the strain increases metabolite production. The membrane allows for complete recovery of mycelia and prevents metabolites from diffusing into the agar. Therefore autoclaved membranes were overlaid onto solid SFM and R5 media and *S. albus*/pCAP1000-*mcp*BGC was grown directly on the surface for two weeks at 30 $^{o}$C. The mycelia was extracted and analysed by HR LC-MS.

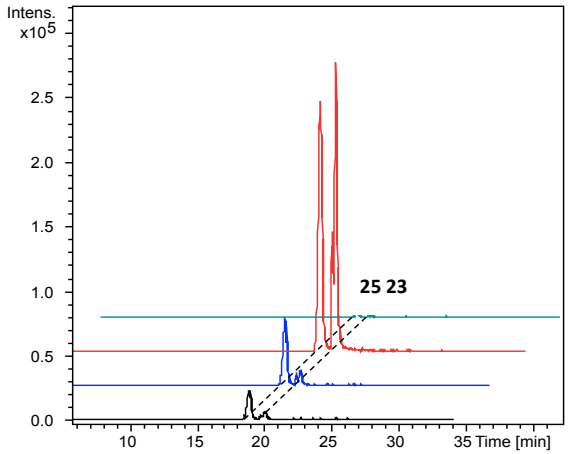
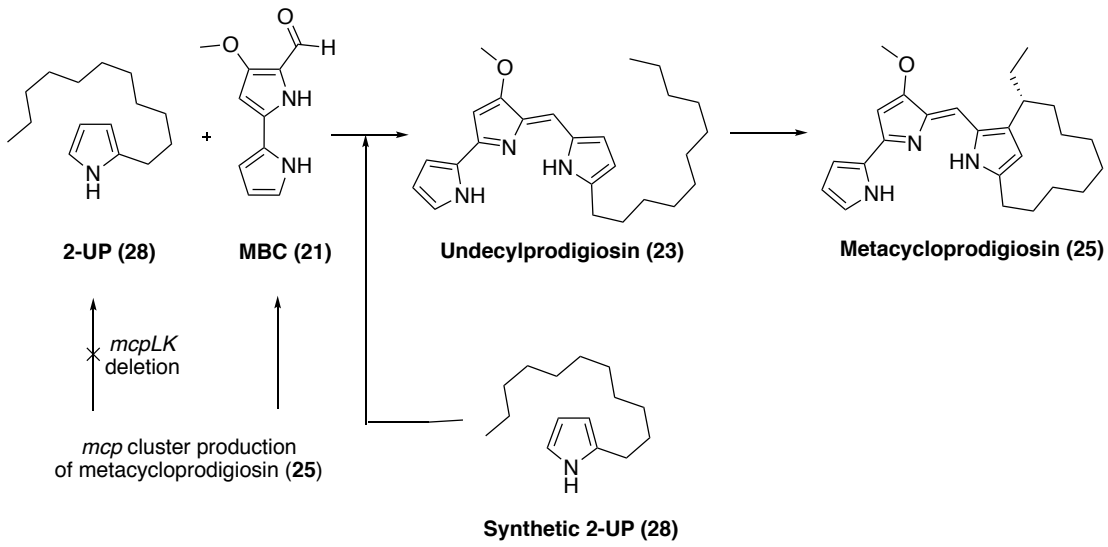



**Figure 2.13:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigisoin (**25**) respectively from LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-*mcp*BGC grown on solid SFM membrane (black trace); *S. albus*/pCAP1000-*mcp*BGC grown on solid R5 media overlaid with a permeable membrane (blue trace:); *S. albus*/pCAP1000-*mcp*BGC grown on solid SFM media overlaid with a semi-permeable membrane (red trace); *S. albus*/pCAP1000-*mcp*BGC grown on solid R5 media (green trace).

The use of a semi-permeable membrane improved production of both undecylprodigiosin (**23**) and metacycloprodigiosin (**25**), however the amount of metacycloprodigiosin (**25**) relative to undecylprodigiosin **(23)** appears lower (figure 2.13). As increasing the overall production levels of metacycloprodigiosin (**25**) was most important to isolate the resulting product for NMR characterization, the membrane was used hereafter. With the production of metacycloprodigiosin (**25**) confirmed and production levels improved, attentions were turned to constructing an *mcpLK* deletion mutant of the pCAP1000-*mcp*BGC vector, using yeast-mediated TAR as a platform.

## 2.4 Construction of an *mcpLK* deletion mutant

The natural homologous recombination that occurs in yeast can be used to repair fragmented DNA in a predictable manner when co-transforming with a pre-designed PCR product. It was envisaged that yeast could be used in the construction of an in-frame deletion of the *mcpLK* gene from pCAP1000-*mcp*BGC using a similar procedure (scheme 2.3). The method can be used to generate quick, in-frame deletions that can be easily screened for using targeted PCR.




**Scheme 2.3:** Schematic representation of the deletion of *mcpLK* in 2-UP (**28**) biosynthesis to abrogate production of 2-UP (**28**). Feeding synthetic 2-UP (**28**) to the mutant therefore restores **23** and cyclic **25** production.

### 2.4.1 Preparation of DNA for co-transformation of yeast and the selection of positive transformants

The presence of a unique cutting site in the vicinity of the *mcpLK* gene is required in order to construct an in-frame deletion. The vector can then be linearized to generate a section of broken stranded DNA primed for repair *via* homologous recombination by yeast cells. Amplification of the regions of DNA that flank the gene to be deleted will provide a PCR product that can undergo homologous recombination with the broken strands of DNA. The recombination step will subsequently replace the targeted gene (*mcpLK*) for the PCR product (figure 2.14).

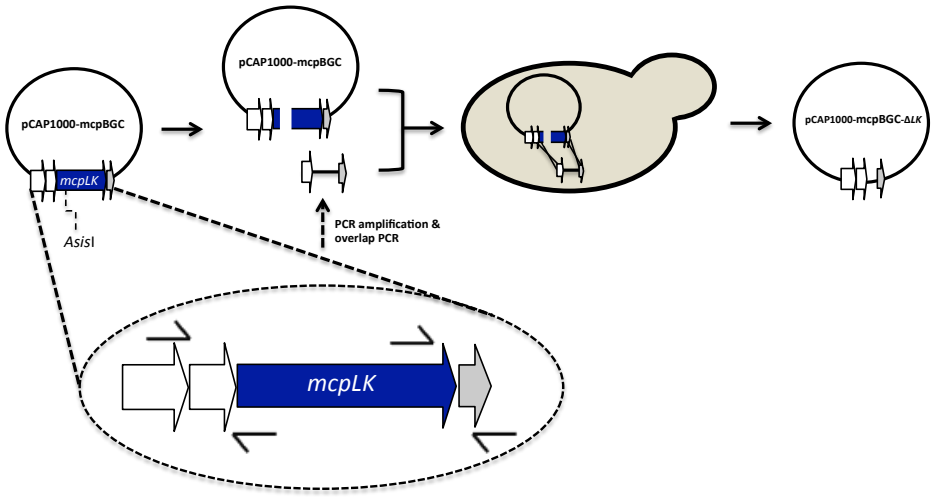



**Figure 2.14:** Schematic outlining the procedure for generating an in-frame deletion using TAR cloning in yeast; highlighting the digestion of the vector and co-transformation of yeast with a PCR product consisting of DNA that is homologous to sequences that flank the gene to be deleted.

An *Asis*I restriction site was identified that only cuts within the *mcpLK* gene and pCAP1000-*mcp*BGC was digested. Two 500 bp regions of DNA, that flank the *mcpLK* gene, were amplified by PCR and purified by agarose gel electrophoresis. The two PCR products were combined into a single 1 kb PCR product by overlap PCR (figure 2.15).

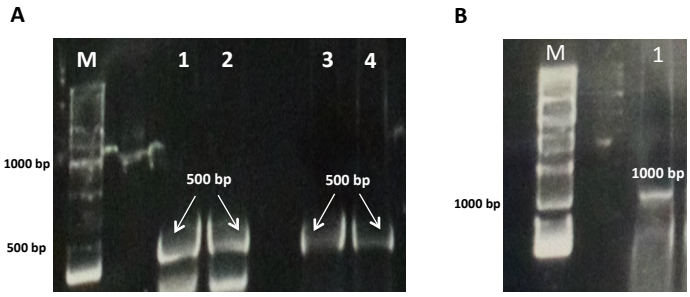



**Figure 2.15: A.** Agarose gel highlighting PCR products corresponding to the segments of *S. longispororuber* gDNA that flank the *mcpLK* gene; M = molecular size marker; lane 1 = lane 2 = left flanking region; lane 3 = lane 4 = right flanking region; **B.** Agarose gel highlighting overlap PCR product resulting from combining both PCR products from gel **A**.

With the linearized pCAP1000-*mcp*BGC vector and PCR product homologous to the segments of DNA flanking the *mcpLK* gene in hand, attentions were turned to the preparation of yeast spheroplasts for co-transformation (section 6.5.7.3). A ratio of 1:2 (vector : PCR product) was used to co-transform the primed spheroplasts and the resulting 200 µL cell-suspension was added directly to top agar media melted at 60 $^{o}$C before pouring onto solid bottom agar for incubation at 30 $^{o}$C. After four days of incubation a total of twelve yeast transformants were visible and six were selected. The DNA from six selected yeast

transconjugants was screened by colony PCR. Primers were designed that targeted the central region of the DNA fragment that had undergone homologous recombination to replace the *mcpLK* gene. Successful mutants will provide a 603 bp PCR product. Figure 2.16 demonstrates that four out of the six transconjugants generated a PCR product of the expected size (lanes 1, 3, 5 and 6). The smaller region of DNA targeted for recombination results in a higher recombination success rate, relative to capturing the entire *mcp* cluster. The negative results by PCR are likely to be colonies that contain the 'wild-type' cluster, where the in-frame deletion had not been successful. Due to the large 8 kb size of the *mcpLK* gene the resulting PCR product for these negative screens (lanes 2 and 4, figure 2.16) would not be visible. Unsuccessful deletions are likely to be a result of re-ligation/ re-circulation of the *Asis*I digested pCAP1000-*mcp*BGC plasmid before homologous recombination with the co-transformed PCR product.

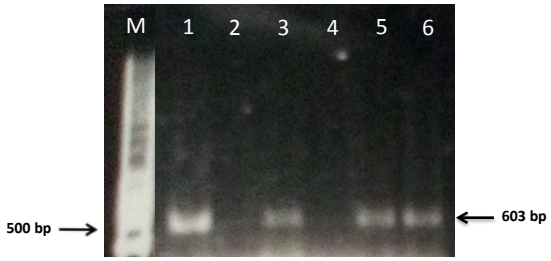



**Figure 2.16:** Agarose gel highlighting 603 bp PCR products from reactions using screening primers and the extracted DNA from six different yeast transconjugants.

A transconjugant that generated a positive PCR result was used to inoculate YPD media. DNA was extracted using the yeast MiniPrep procedure outlined in section 6.5.4. The resulting plasmid DNA was used to transform electrocompetent *E. coli* TOP10 cells, cultured and extracted for analysis by restriction enzyme digest. The plasmid DNA must be transferred to *E. coli* for replication, as yeast cells only produce a single copy of plasmid per cell compared to multiple copies per cell in the *E. coli* TOP10 strain. The sequence was confirmed by digesting with *Sph*I. The resulting deletion mutant will be called *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*, and was incubated on solid SFM media at 30 $^{o}$C for two weeks. The resulting mycelia failed to produce the red-pigmented metabolites produced by *S. albus*/pCAP1000-*mcp*BGC (figure 2.17). This suggests successful deletion, which was confirmed by analysis of the corresponding mycelial extract by HR LC-MS (figure 2.17).

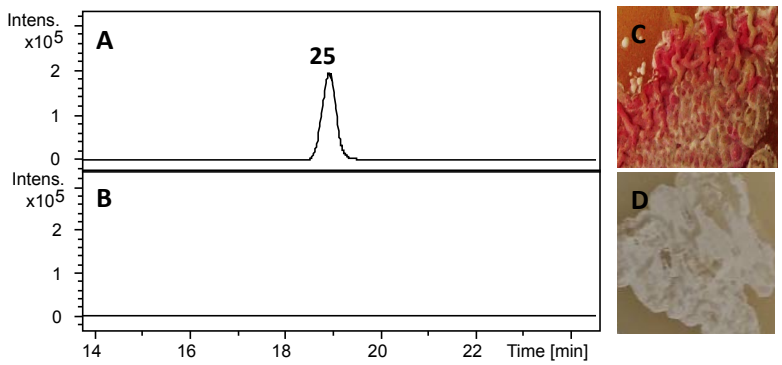



**Figure 2.17:** (*left*) EICs at *m/z* = 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of metacycloprodigisoin (**25**) from the LC-MS of the mycelial extract of **A.** *S. albus*/pCAP1000-*mcp*BGC; and **B.** *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* (*right*) Images of the mycelia of *S. albus*/pCAP1000-*mcp*BGC **C.** and *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* **D**; grown on solid SFM media at 30 °C for two weeks.

### 2.4.2 Restoring metacycloprodigiosin (25) production in *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* mutant

Attempts were made to restore the production of metacycloprodigiosin (**25**) by feeding synthetic 2-UP (**28**) to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*. *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* was grown on a sterile permeable membrane overlaid onto solid SFM media and incubated at 30 °C for two days. A methanolic solution of synthetic 2-UP (**28**) (1 mg in 50 µL) was added in small droplets onto the mycelia and the strain incubated for a further 7-10 days. Mycelium that turned red were selected and the biomass was collected. The metabolites were extracted using acidified methanol and analysed by HR LC-MS, which confirmed the production of both undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) (figure 2.18). As the peak intensities for both (**23**) and cyclic (**25**) prodiginines were low, efforts were then focused on optimizing the feeding experiment to increase the yield of metacycloprodigiosin (**25**).

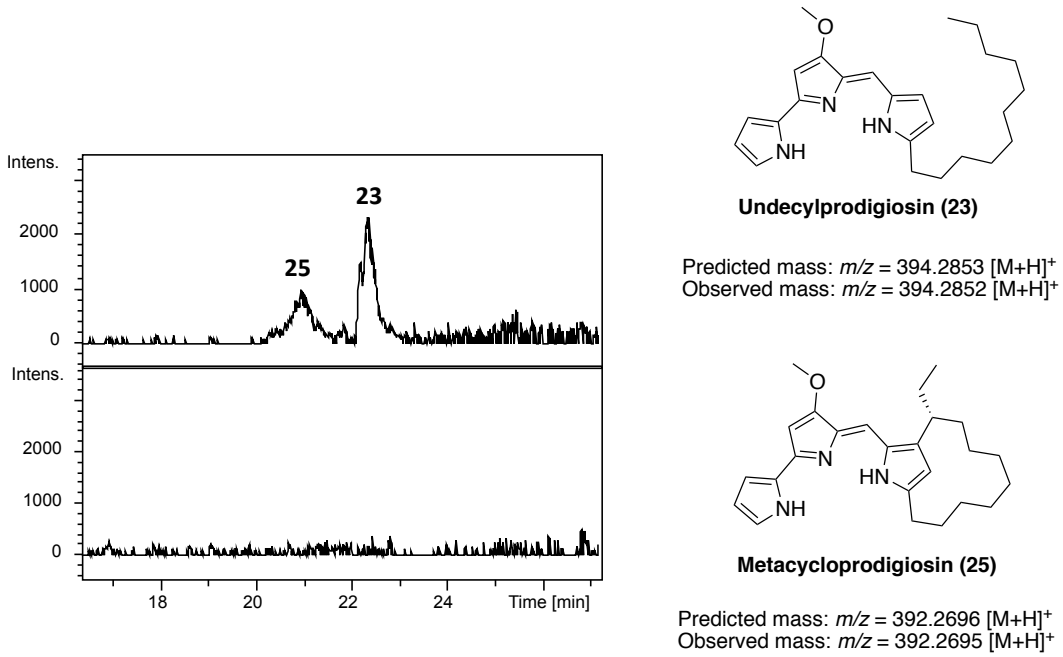



**Figure 2.18:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (blue) (**23**) and metacycloprodigisoin (**25**) (red) respectively from the LC-MS analysis of the mycelial extract of **A.** *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* fed with synthetic 2-UP (**28**); **B.** *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*.

### 2.4.3 Optimizing production of metacycloprodigiosin in *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*

*S. albus*/pCAP1000-*mcp*BGC-Δ*LK* was grown on three separate plates and fed with increasing amounts of synthetic 2-UP (**28**) (1 mg, 2 mg and 4 mg, dissolved in 50 µL methanol). HR LC-MS analysis of the mycelial extracts showed no improvement for metacycloprodigiosin (**25**) production (figure 2.19). To test whether the stability of metacycloprodigisoin (**25**) or the precursor was a problem, synthetic 2-UP (**28**) was fed to the deletion mutant and incubated for one day and two days respectively to compare the peak intensity for metacycloprodigiosin (**25**) at different time points post-feeding.

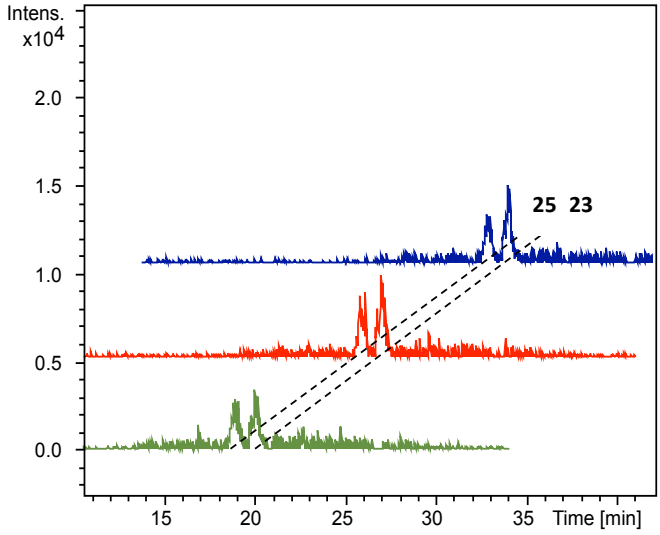



**Figure 2.19:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) respectively from LC-MS analysis of the mycelial extracts of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* fed with 4 mg synthetic 2-UP (**28**) (green); 1 mg synthetic 2-UP (**28**) (red) and 2 mg synthetic 2-UP (**28**) (blue).

Mycelial extracts at both one and two-day time points, post-feeding, were analyzed by HR LC-MS (figure 2.20). No improvement in production levels were observed, implying that the metabolites are not degraded and that the production of metacycloprodigiosin (**25**) is not limited by the 2-UP (**28**) substrate.

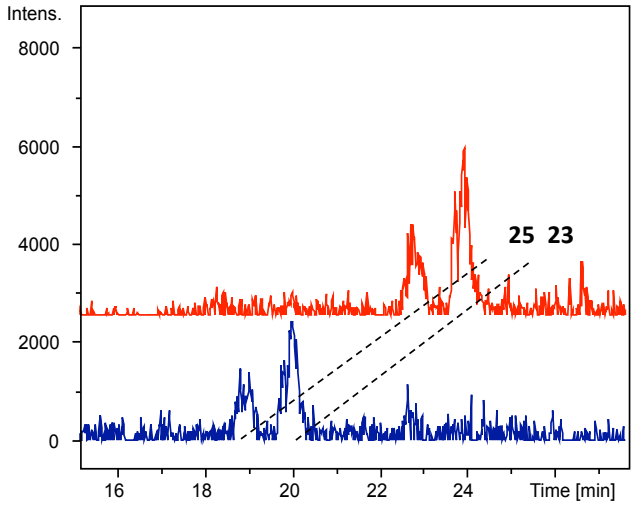
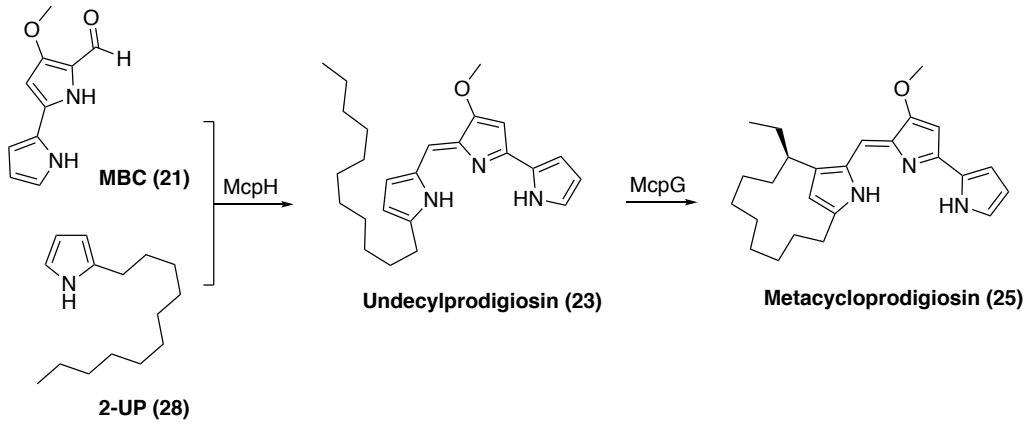



**Figure 2.20:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigisoin (**25**) respectively from the LC-MS analysis of the mycelial extract after; Incubation for two days, post-feeding with synthetic 2-UP (**28**) (red); Incubation for one day, post-feeding with synthetic 2-UP (**28**) (blue).

An alternative hypothesis for why metacycloprodigiosin (**25**) production is low is that construction of the *mcpLK* mutant has resulted in the down-regulation of MBC (**21**), the other precursor (scheme 2.4).




**Scheme 2.4:** Late stages of metacycloprodigiosin (**25**) biosynthesis, highlighting the condensation and consequent oxidative carbocyclization of 2-UP (**28**) and MBC (**21**).

Feeding a mixture of synthetic MBC (**21**) and synthetic 2-UP (**28**) did result in an increase in the production levels of metacycloprodigiosin (**25**) (figure 2.21). However, the production of linear undecylprodigiosin (**23**) is also considerably improved resulting in an increased amount of **23** to cyclic (**25**), and the duration of the feeding experiment was still eight to ten days.

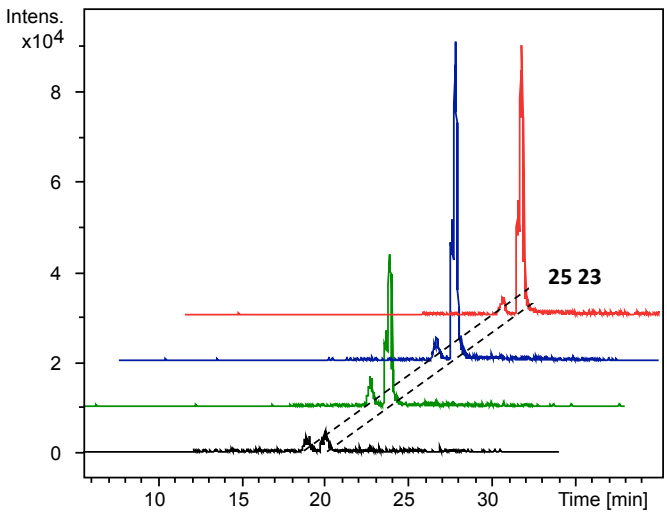



**Figure 2.21:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) respectively from the LC-MS analysis of the mycelial extract of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) (black); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) + 0.25 mg MBC (**21**) (green); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) + 0.50 mg MBC (**21**) (blue); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* + 1 mg 2-UP (**28**) + 1 mg MBC (**21**) (red).

Constitutively expressing upstream *mcpG* was predicted to increase the rate of metacycloprodigiosin (**25**) production further. An ampicillin selective *mcpHG* expression vector (where *mcpG* is expressed under the control of the constitutive promoter *ermE**) was available in the Challis group and previously used in this study (section 2.2.3). This plasmid was inserted into *S. albus*/pCAP1000-*mcp*BGC-Δ*LK* deletion mutant by conjugal transfer. For clarity, the deletion mutant that over-expresses *mcpHG* will be referred to as *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG*.

The white-coloured *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* strain developed red pigmentation after only two hours of incubation, post feeding of synthetic 2-UP (**28**). LC-MS analysis of the mycelia extracts prepared after two days of incubation showed an increase in the peak intensity in the EIC for **23** and cyclic metacycloprodigiosin (**25**) (figure 2.22, blue trace). Feeding the strain with synthetic 2-UP (**28**) and MBC (**21**) further increased production of both metabolites further (figure 2.22, red trace). The system constitutively expressing *mcpG* and feeding synthetic MBC (**21**) alongside synthetic 2-UP (**28**) was therefore used going forward.

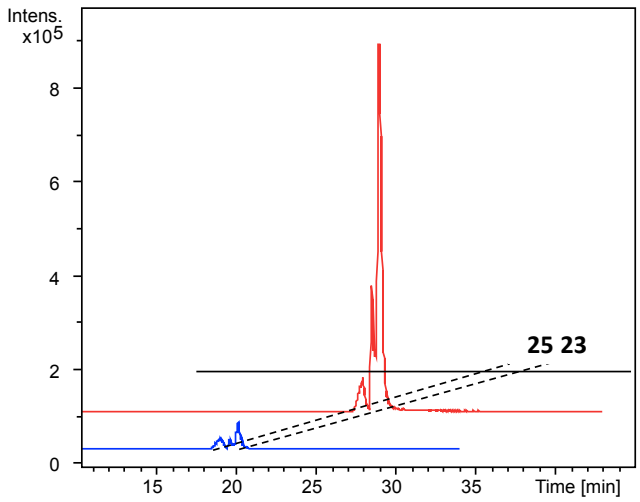



**Figure 2.22:** EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) respectively from the LC-MS analysis of the mycelial extract of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) (blue); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) + 1 mg MBC (**21**) (red); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* no feeding (black).

To confirm the identity of metacycloprodigiosin (**25**) produced under these conditions the products were directly compared with metacycloprodigiosin (**25**) produced by *S. longispororuber*. LC-MS and MS/MS analysis showed that they have identical retention times and fragmentation patterns (figure 2.23).

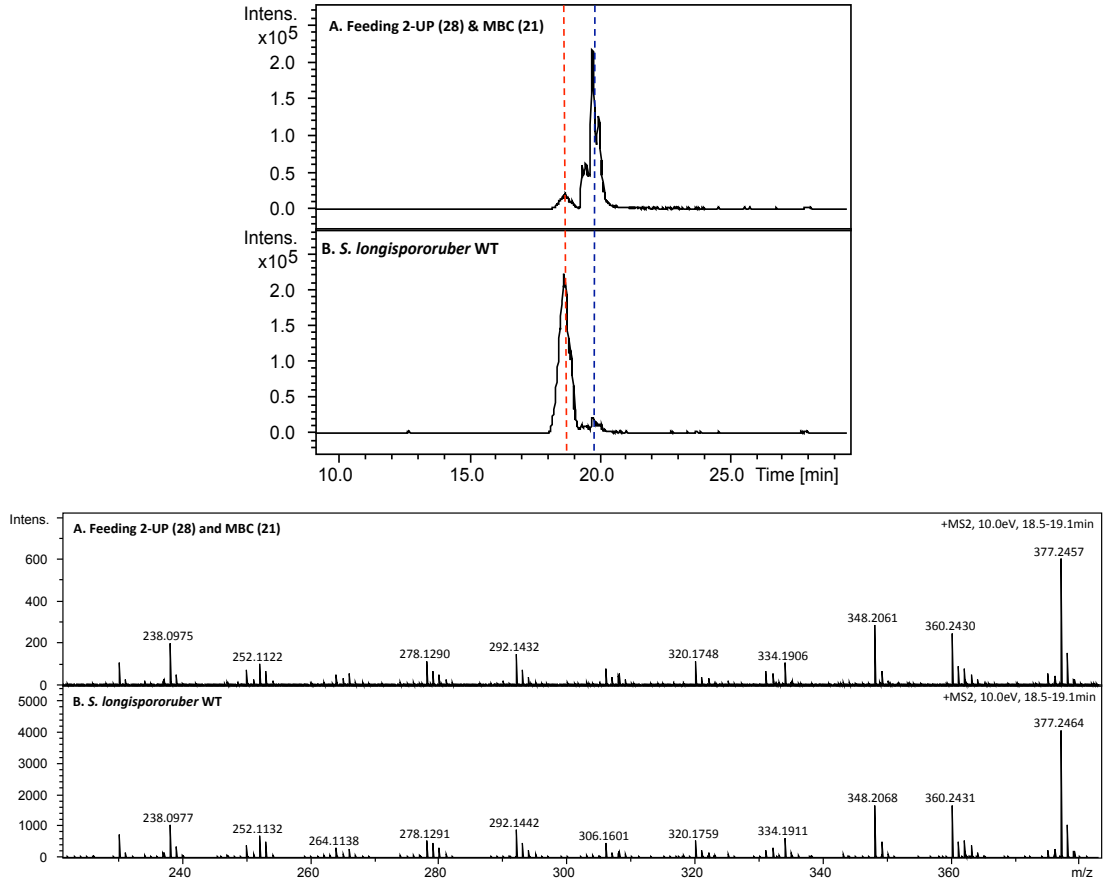



**Figure 2.23:** (*Top*) EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) (blue) and metacycloprodigiosin (**25**) (red) respectively from the LC-MS analysis of the mycelial extract of; *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) + 1mg MBC (**21**) **A**; *S. longispororuber* WT **B**; (*bottom*) MS/MS spectra of the $[M+H]^+$ metabolite at *m/z* = 392 corresponding to metacycloprodigiosin (**25**); *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* + 1 mg 2-UP (**28**) + 1 mg MBC (**21**) **A**; *S. longispororuber* WT **B**.

The fragmentation pattern of prodiginine compounds in MS/MS experiments has been studied in detail for undecylprodigiosin (**23**) and streptorubin B (**15**).[112] The study highlighted a number of atypical odd-electron ions ($OE^{+\bullet}$) formed through the fragmentation of even-electron ions ($EE^+$), where radical eliminations lead to odd-electron fragment ions that are highly stable.[112] Computational calculations were used to show that ring-C of metacycloprodigiosin (**25**) is likely to be protonated and form a hydrogen-bond with ring-B, forming a pseudo-seven-membered ring that constitutes the most stable ground state $[M+H]^+$ structure (figure 2.24).[112] The protonated prodiginine undergoes homoloytic dissociation to lose a methyl radical leading to an $OE^{+\bullet}$ or heterolytic bond cleavage to lose methanol and form an $EE^+$, the former being the most facile and therefore favoured. Subsequent fragmentation of the alkyl chain through loss of *m/z* = 14 can also be observed.

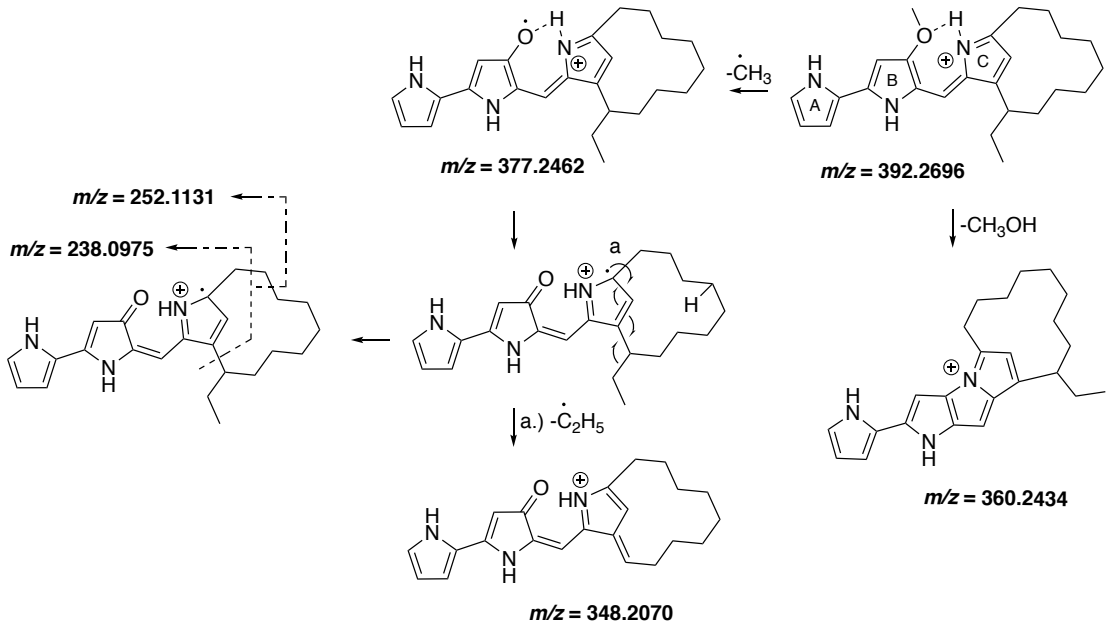
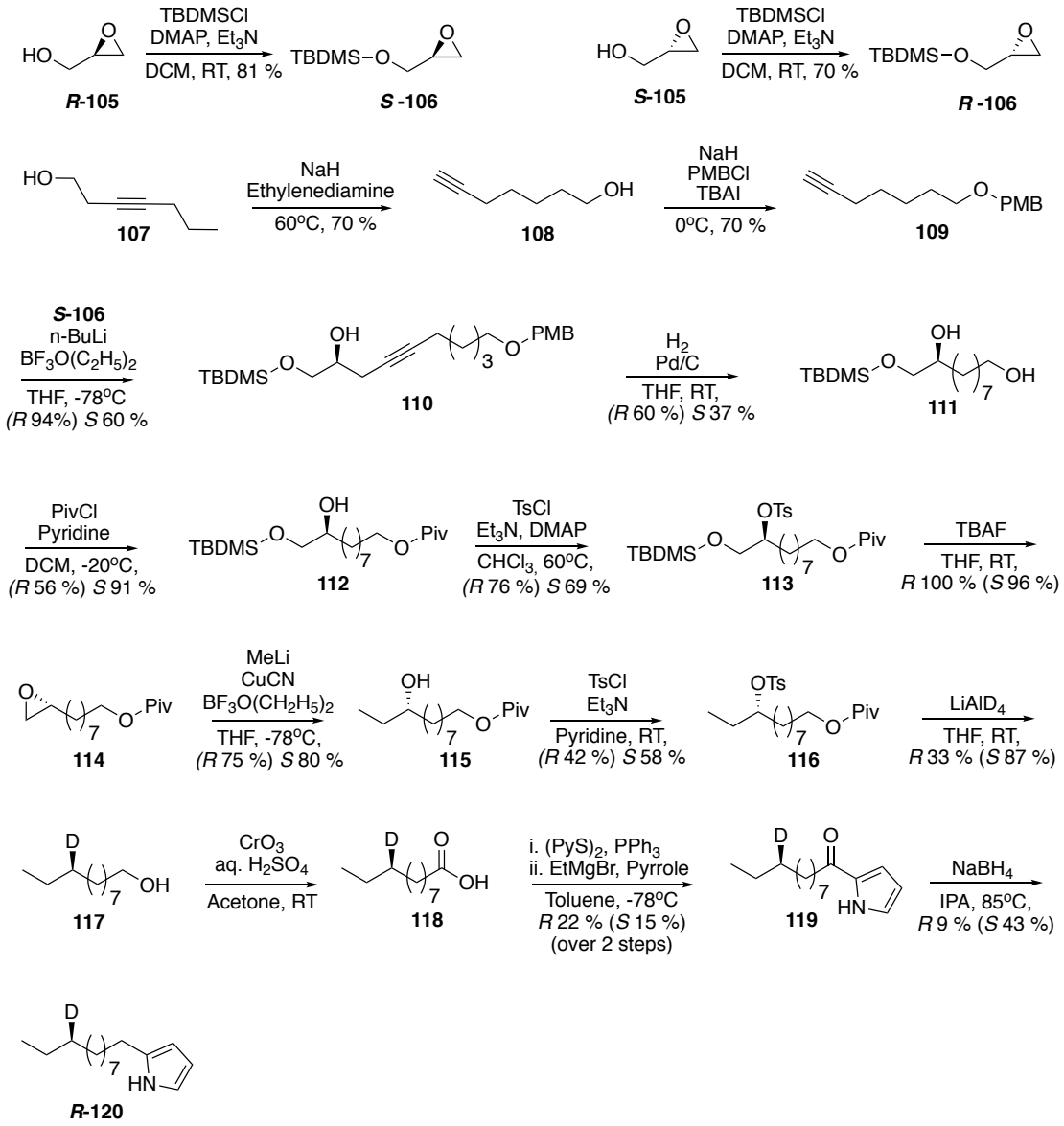



**Figure 2.24:** The proposed fragmentation pattern of metacycloprodigiosin (**25**).

Unfortunately, insufficient material was isolated for full characterization by NMR. Since the stereocontrol of McpG will by analysed by HR-MS this is not a problem. With a robust *in vivo* system that can readily accept synthetic MBC (**21**) and 2-UP (**28**) to furnish metacycloprodigiosin (**25**) in hand, attentions were turned to the synthesis of 2-UP analogues stereoselectively labelled with a deuterium atom in both the *S* and *R* configuration at C9 of the alkyl chain. Feeding both of the enantiomers separately to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* and analysis of deuterium incorporation in the resulting metacycloprodigiosin product will provide insight into the stereochemical course of the oxidative carbocyclization reaction catalyzed by McpG.

## 2.5 Synthesis of stereoselectively C9-deuterium-labeled 2-UP analogues

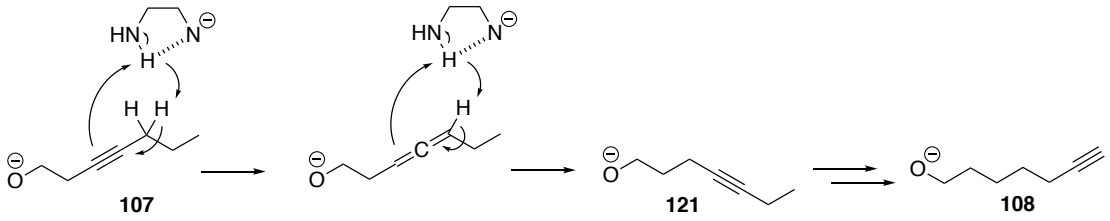
A stereoselective synthetic route was designed based upon the pathway used to synthesize the *R* and *S* C7-deuterium labelled 2-UP analogues by Withall and co-workers when investigating the stereochemical course of Rieske oxygenase RedG.[49]




**Scheme 2.5:** The synthetic route and conditions for the stereoselective synthesis of deuterium-labeled 2-UP analogues ***R***- and ***S*-120**. Chemical structures show intermediates in the route to ***R*-120**, but yields for the other enantiomer are given in brackets.

Both ***R***- and ***S***- glycidol (**105**) were protected using *tert*-butyldimethylsilyl chloride to generate ***R***- and ***S*-106** (scheme 2.5). Commercially available 3-heptyn-1-ol **107** was isomerized to the terminal alkyne **108** under basic conditions using neat sodium hydride and

ethylene diamine. The 'alkyne zipper' reaction deprotonates the methylene adjacent to the central alkyne of compound **107** (scheme 2.6). A second deprotonation then forms the alkyne moeity that is now one carbon further along the alkyl chain, compound **121**. This migration repeats until reaching the terminal position. [113]




**Scheme 2.6:** The mechanism of the alkyne zipper reaction to form terminal alkyne **108**.

Initial attempts to protect the alcohol moiety of terminal alkyne **108** using sodium hydride and 4-methoxybenzyl chloride resulted in internal migration of the alkyne moiety by one carbon atom (figure 2.25). There is precedence for a one carbon migration, noted by Brandsma and co-workers, where base-catalyzed isomerizations of 1-alkynes to 2-alkynes are outlined under various conditions.[114] These migrations are promoted by base and temperatures above 20 $^{o}$C to furnish the thermodynamic product. Protections were therefore performed at 0 $^{o}$C and the addition of *tetra*-butylammonium iodide ensured the rate of chloride nucleophillic substitution was not lowered.

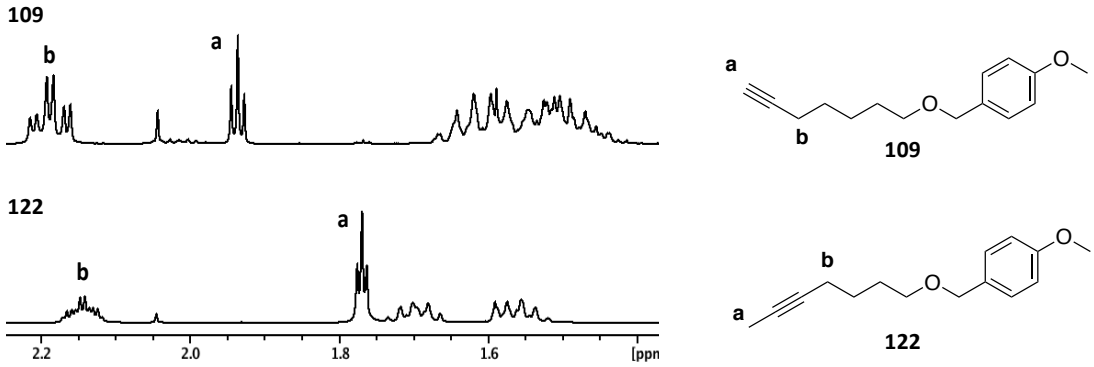



**Figure 2.25:** A section of the $^{1}$H-NMR spectra of compound **121** and the undesired migratory product **134**.

Boron trifluoride promoted regioselective ring opening occurred at the least hindered side of the protected glycidol **106** by the lithium acetylide formed *via* treatment of compound **109** with *n*-butyl lithium. Hydrogenation of the alkyne moiety and removal of the 4-methoxybenzyl protecting group was undertaken in one step using hydrogen gas and a palladium on carbon catalyst at a pressure of 50 bar. The resulting diols **115** were treated with pyridine and pivaloyl chloride to selectively protect the primary alcohol and the secondary

alcohol moiety was tosylated to provide both enantiomers of **113**. Removal of the *tert*-butyldimethylsilyl protecting group using *tetra*-butyl ammonium fluoride (TBAF) and subsequent intramolecular nucleophillic substitution of the tosyl moiety furnished epoxides ***R*-114** and ***S*-114**. Opening of the epoxides using commercial methyl lithium and copper cyanide gave the desired eleven-carbon chain and generated alcohols ***S*-115** and ***R*-115**. Copper cyanide and methyl lithium form an organocuprate $Me_2Cu(CN)Li_2$ which favours opening of the epoxide moiety rather than undesired displacement of the pivaloyl ester.[115,116] The stereochemistry of the resulting *R* and *S* enantiomers of alcohols **115** were then confirmed by $^1$H-NMR analysis of their corresponding Mosher's ester derivatives (see section 2.5.1). Tosylation of the secondary alcohols generated ***R*-** and ***S*-116** and lithium aluminum deuteride was used to stereoselectively install a deuterium atom at C9 of the alkyl chain with inversion of stereochemistry by nucleophillic substitution of the tosyl moeity, whilst simultaneously removing the pivaloyl protecting group to furnish alcohols **117**. The incorporation of deuterium was above 95 % as seen by $^1$H-NMR and confirmed by comparison the $^{13}$C-NMR of ***R*-** and ***S*-117** with that of undecanol (figure 2.26). The observed additional splitting at 29.9 ppm occurs due to the presence of an adjacent deuterium atom. The loss of peak at 63.2 ppm and the emergence of a triplet peak at 31.6 ppm is due to the presence of a deuterium atom on C9 rather than two hydrogen atoms.

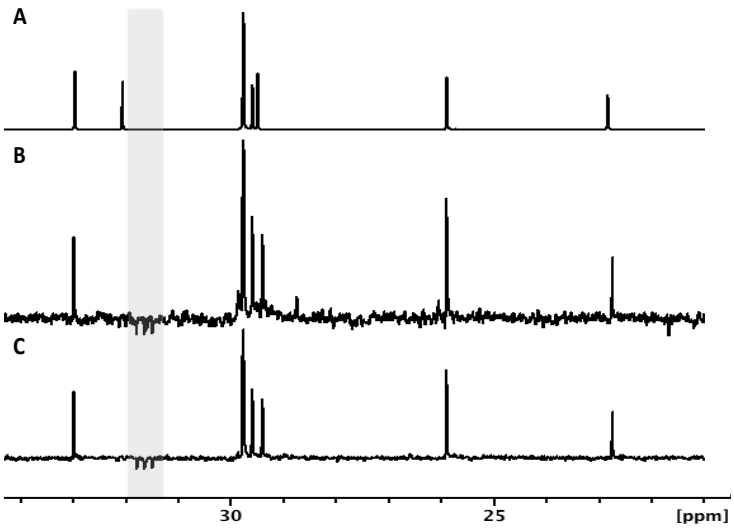
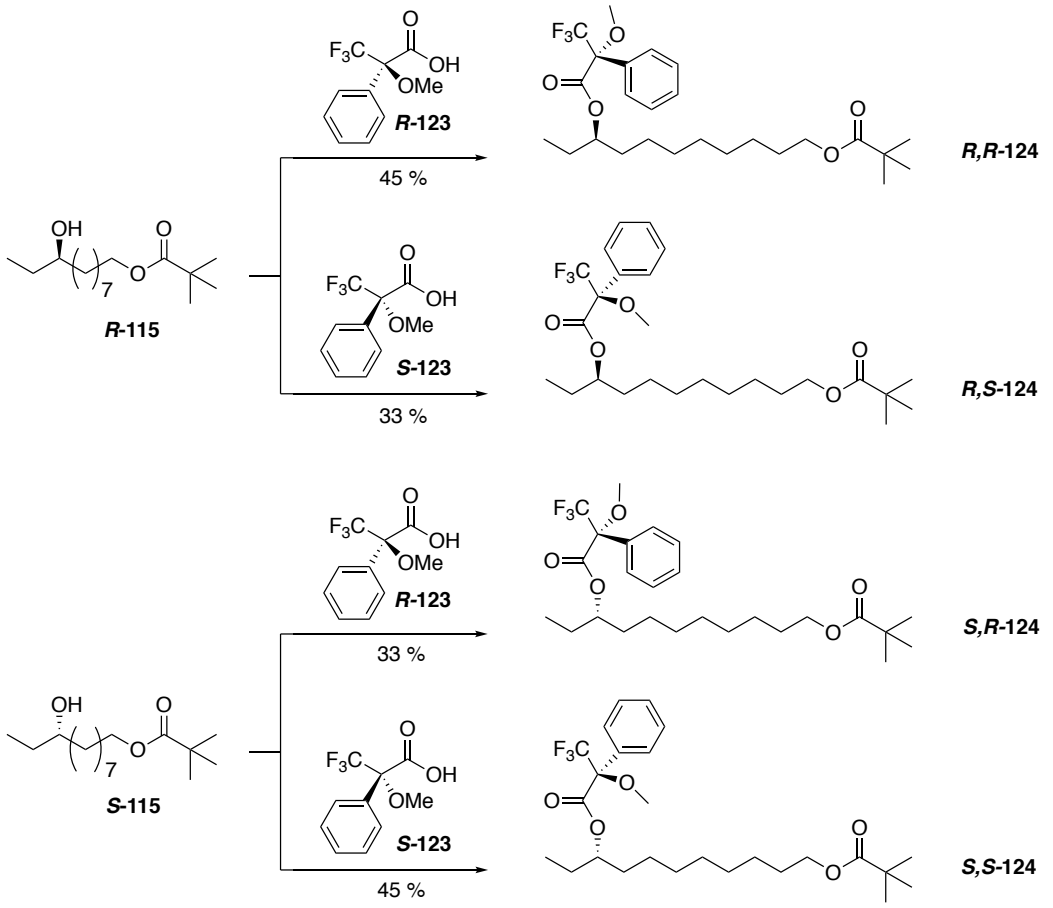



**Figure 2.26:** Comparison of the $^{13}$C-NMR spectra of undecanol **(A)**, ***R*-117 (B)** and ***S*-117 (C)**.

Both enantiomers of alcohol **117** were oxidized to their corresponding carboxylic acids using Jones' reagent generated *in situ* followed by coupling to freshly distilled pyrrole utilizing the procedure outlined in section 6.2.4. Reduction of the *R* and *S* deuterium labeled 2-ketopyyroles **119** using sodium borohydride in refluxing propan-2-ol provided the target ***R*-** and ***S*** deuterium labeled 2-UP analogues **120**.

### 2.5.1 Confirmation of the stereochemistry of alcohols *R*-115 and *S*-115 by spectroscopic analysis of their corresponding Mosher's ester derivatives

Both alcohols ***R*-115** and ***S*-115** were derivatised using both enantiomers of Mosher's acid (**123**) to generate all four possible diastereomers (***R,S*-124**), (***R,R*-124**), (***S,S*-124**) and (***S,R*-124**) (scheme 2.7).[117,118] The $^{1}$H-NMR spectra of each of the resulting diastereomers were compared.




**Scheme 2.7**: The synthesis of Mosher's ester derivatives of *R* and *S* secondary alcohol **115**. The reaction was undertaken using conditions outlined in section 6.3.3. the terminal methyl protons were observed to obtain the greatest shift in ppm as a result of the configuration of the Mosher's ester derivative and was used to confirm the predicted absolute stereochemistry.

As can be seen in figure 2.27 the two pairs of enantiomers (***R,S*-124** and ***S,R*-124**) and (***R,R*-124** and ***S,S*-124**) exhibit identical $^{1}$H-NMR spectra whilst the pairs of diastereomers (***R,S*-124** and ***R,R*-124**) and (***S,S*-124** and ***S,R*-124**) differ. The most significant difference can be observed from the terminal methyl group (figure 2.27) at 0.92 ppm for the ***R,S*-124** and ***S,R*-124** enantiomers and at 0.81 ppm for the ***R,R*-124** and ***S,S*-124** enantiomers. Comparison of the directionality of these chemical shifts with those proposed by Mosher and co-workers

allowed for the confirmation of the expected stereochemistry at C9 for ***R-*** and ***S*-115**. An optical purity of 95 % ee was assumed based on the limit of resolution by $^{1}$H-NMR.

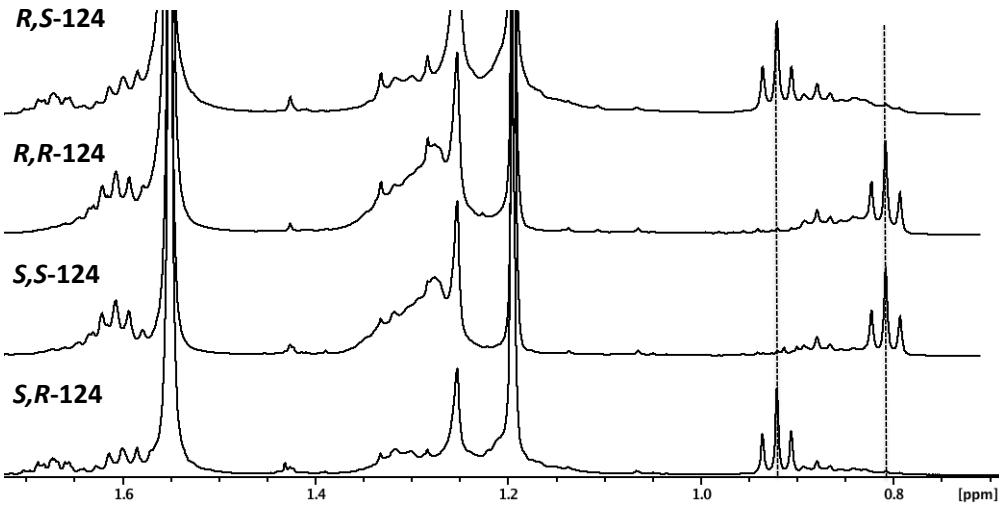



**Figure 2.27**: Comparison of the $^{1}$H-NMR spectra ($CDCl_3$, 500 MHz) of the four diastereomers of compound **124**, highlighting a key upfield shift of the terminal methyl triplet in the ***R,R* -124** and ***S,S*-124** diastereomers.

## 2.6 Feeding stereoselectively labeled 2-UP analogues *R*-120 and *S*-120 to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG*

Both ***R*-120** or ***S*-120** were fed alongside synthetic MBC **(21)** to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* grown on a sterile semi-permeable membrane for two days, then incubated for two days. Synthetic 2-UP **(28)** was also fed to the strain for use as a positive control. The metabolites were extracted using acidified methanol and the resulting extracts were analysed by HR LC-MS.

The amount of cyclic prodiginine relative to linear prodiginine following feeding ***R*- 120** (figure 2.28, chromatogram **A**) is comparable to that when feeding synthetic 2-UP **(28)** (figure 2.28, chromatogram **C**). The ratio of cyclic prodiginine relative to linear prodiginine following feeding ***R*-120** is comparable to that when feeding synthetic 2-UP **(28)** (figure 2.28). Contrastingly, the relative conversion of cyclic prodiginine following feeding ***S*-120** from linear prodiginine is decreased since deuterium abstraction would be influenced by a primary kinetic isotope effect, resulting in a slower rate of the reaction. This suggests deuterium abstraction rather than hydrogen abstraction of ***S*-120**. These data suggest that McpG abstracts the *pro-S* hydrogen during cyclization.

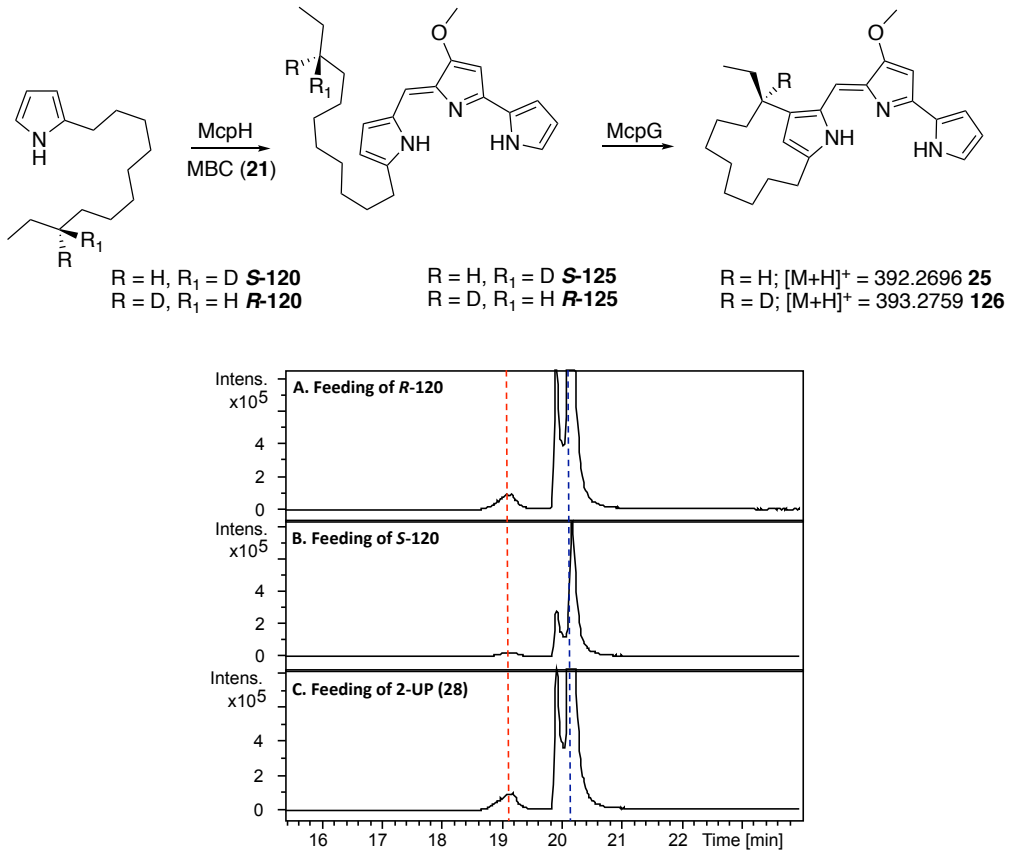



**Figure 2.28:** (*Top*) The feeding of ***R*-120** and ***S*-120** to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* resulting in the corresponding undecylprodigiosin analogues ***R*-** and ***S*-125**, and the two possible outcomes for the product of McpG catalysis; (*bottom*) EICs at *m/z* = 395.29 and 393.28 ±0.01 (*top and middle*) corresponding to the calculated $[M+H]^+$ of deuterium-labeled undecylprodigiosin (**125**) (blue) and deuterium-labeled metacycloprodigisoin (**126**) (red) respectively from the LC-MS analysis of the mycelial extract resulting from feeding **A.** *pro-R* labelled 2-UP (***R*-120**); **B.** *pro-S* labelled 2-UP (***S*-120**); (*bottom*) EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) (blue) and metacycloprodigisoin (**25**) (red) respectively of the mycelial extract from feeding **C.** 2-UP (**28**) (positive control).

The major cyclic compound observed as a result from the feeding of ***R*-120** has an *m/z* = 393.2756 $[M+H]^+$ corresponding to deuterium-labeled metacycloprodigiosin (**126**) (figure 2.29, mass spectrum **A**). The intensity of the peak at *m/z* = 393.2747 in the chromatograms is a result of a combination of deuterium-labeled metacycloprogiosin and the metacycloprodigiosin $^{13}C$-isotopomer. To quantify the contribution of metacycloprodigiosin $^{13}C$-isotopomer a comparison was made to the ratio of peak intensities of the positive control, calculated to be 22 % of the $^{12}C$-isotopomer. Analysis of the peak intensities confirms that 96 % of the deuterium label is retained indicating that McpG catalyzes oxidative carbocyclization by abstraction of the *pro-S* proton at C9 of the undecylprodigiosin alkyl chain. The peak intensity for metacycloprodigiosin (**25**) (*m/z* = 392.2696 $[M+H]^+$) following feeding of ***S*-120** is significantly higher in comparison to mass spectrum **A**, suggesting a

greater degree of deuterium abstraction by McpG. Using this data it was calculated that 53 % of the deuterium label has been lost during the McpG catalyzed oxidative carbocyclisation when ***S*-120** was fed to the deletion mutant. This experiment was repeated to generate identical data (not shown).

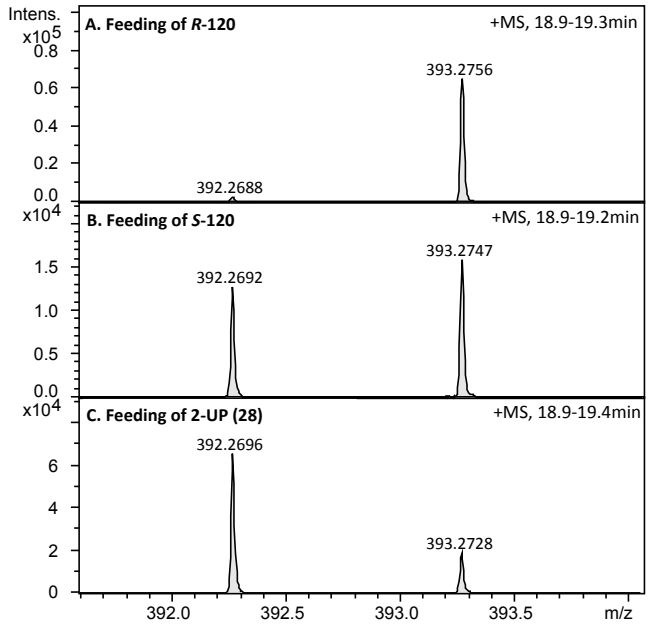



**Figure 2.29:** A comparison of the HR-MS spectrum of the cyclic prodiginine compounds eluted at ~19.0 minutes from the LC-MS analysis of the mycelial extracts resulting from feeding *pro-R* labeled 2-UP (***R*-120) (A)**; *pro-S* labeled 2-UP (***S*-120) (B)**; and 2-UP (**28**) **(C)** to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG*.

As was observed for the analogous experiments with RedG, complete abstraction of the deuterium label was not observed. This is proposed to be because of the primary kinetic isotope effect associated as removal of a deuterium instead of a hydrogen atom slows down the reaction and permits abstraction of an alternative hydrogen. Interestingly, the degree of deuterium abstracted by McpG and RedG differs (53 % compared to 77 %).[49] This may imply that McpG has the potential to abstract a hydrogen from an adjacent carbon atom along the alkyl chain of undecylprodigiosin (**23**).

These data indicate that McpG abstracts the *pro-S* proton at C9 of the undecylprodigiosin alkyl chain during oxidative carbocyclization to furnish metacycloprodigiosin (**25**). As metacycloprodigisoin (**25**) exhibits *R* absolute stereochemistry, McpG must catalyze oxidative carbocyclisation with inversion of stereochemistry at C9, in an analogous manner to RedG in streptorubin B (**15**) biosynthesis.[49]

## 2.7 Conclusions and perspectives - Stereocontrol of McpG catalyzed oxidative carbocyclization of undecylprodigiosin to metacycloprodigiosin

The metacycloprodigiosin biosynthetic gene cluster was captured by TAR cloning and heterologously expressed in *S. albus*, resulting in production of metacycloprodigiosin (**25**). An in-frame deletion was then made in the *mcpLK* gene, also using TAR, which abolished production. Metacycloprodigiosin (**25**) production could be restored by feeding synthetic 2-UP (**28**) and MBC (**21**), thus establishing a system to probe the mechanism of McpG catalysis. This study establishes the use of TAR cloning as a powerful platform for the manipulation of genetically intractable gene clusters through heterologous expression. This work also highlights TAR as a new *in vivo* method to refactor large fragments of DNA, such as gene clusters, in a quick and efficient manner to investigate targeted steps within biosynthetic pathways.

Analogues of 2-UP (**28**) with a deuterium atom in both the *pro-R* (***R*-120**) and *pro-S* (***S*-120**) position of C9 were synthesized. Feeding ***R*-120** to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-*mcpHG* furnished metacycloprodigiosin where the deuterium was almost completely retained suggesting that only the *pro-S* proton that is abstracted by McpG during oxidative carbocyclization.

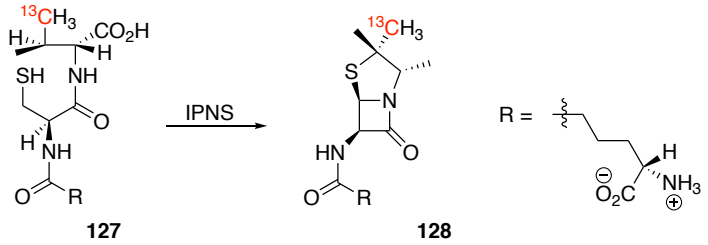
Feeding ***S*-120** generated metacycloprodigiosin where McpG had predominantly abstracted the deuterium label (53 %). The definitive reason behind the amount of remaining compound with *m/z* = 393.2747 cannot be resolved without further experiments, it is likely that McpG has the capacity to abstract a proton from an adjacent carbon atom or the other hydrogen atom at C9. The alternative C-H abstraction only occurs because the rate of the normal reaction is reduced due to the primary kinetic isotope effect occurring from breaking the C-D bond. The hypothesis can be confirmed by scale-up of the feeding experiment in a bid to isolate the resulting deuterium incorporated metacycloprodigiosin metabolite **126** to solve the structure by $^{1}$H-NMR. Circular dichroism could then be used to determine that the metacycloprodigiosin (**25**) species remains completely *R*-configured.

McpG must therefore proceed with inversion of stereochemistry at C9 of the undecylprodigiosin (**23**) alkyl chain, favouring abstraction of the *pro-S* proton to furnish metacycloprodigisoin with *R*-configuration. Both characterized Rieske oxygenases (RedG and McpG) involved in prodiginine biosynthesis therefore catalyze oxidative carbocyclization reactions with inversion of stereochemistry. This result is in contrast to other oxygenases

undertaking similar reactions that catalyze cyclization with retention of stereochemistry, such as isopenicillin-*N*-synthase and clavaminate synthase.[119–123]

### 2.7.1 Stereocontrol of isopenicillin-*N*-synthase (IPNS)

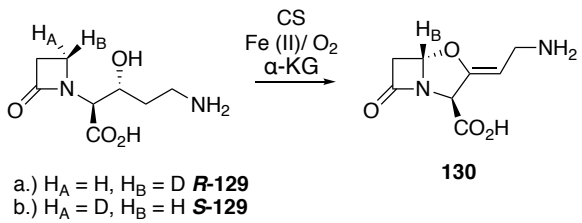
In 1986, Baldwin and co-workers investigated the stereospecificity of the carbon-sulphur bond forming reaction catalyzed by IPNS in penicillin biosynthesis.[123] A stereospecifically labelled α-aminobutyrate peptide (**127**) was synthesized from (2*S,* 3*S*)-[4-$^{13}$C]-valine through a series of racemization, protection, coupling and deprotection reactions and incubated with purified homogeneous IPNS. The resulting isotopically labelled penicillin (**128**), arising from incubating the stereospecifically prepared substrate with IPNS, was analysed by $^{13}$C-NMR and indicated that C-S bond formation occurred with complete retention of stereochemistry at the valinyl C-3 position (scheme 2.8).




**Scheme 2.8:** The stereocontrol of IPNS analysed by incubation of a stereospecifically labeled probe.

### 2.7.2 Stereocontrol of clavaminate synthase (CS)

Townsend and co-workers synthesized proclavaminates (**129**) that were stereospecifically deuterated at both the *R* and *S* configuration at position C-4' of the β-lactam ring.[119] The proclavaminates were incubated with CS, that had been isolated *in vitro*, and the results showed that the 4'-*S*-deuterium atom ($H_A$, scheme 2.9), was lost and a significant kinetic isotope effect was observed, indicating the breakage of a C-D bond rather than a C-H bond. The 4'-*R*-hydrogen atom ($H_B$, scheme 2.9) was retained (**130**). These data indicates that CS catalyzes abstraction of a hydrogen atom with retention of stereochemistry at the carbon centre.




**Scheme 2.9:** Incubation of the stereoselectively deuterium labelled proclavaminates used to probe the stereocontrol of the purified CS enzyme.

One explanation for the difference in the stereochemical course of McpG catalysis is that the carbon-centred radical formed during the catalytic cycle is not bound to the non-heme iron centre and does not react with a bound ligand, which is the case for both IPNS and CS. The pyrrole is instead orientated in such a way that it is shielded from potential reactive oxygen species by the alkyl chain. Without this orientation the non-heme iron centre may instead oxidize the pyrrole moeity. This orientation is similar to that proposed class B radical SAM methylases such as GenD1 and BmbF/ BmbB/ BmbJ for example; enzymes that also proceed with inversion of stereochemistry.[124]

### 2.7.3 Class B radical-*S*-adenosylmethionine (SAM) methylases

Class B radical SAM methylases catalyze methylation reactions across a wide range of natural products. In most biological systems methylations occur by the use of SAM as the methyl donor and proceed *via* an $S_N2$ displacement reaction.[125,126] The proposed mechanism of class B radical SAM involves electron transfer from a [4Fe-4S] cluster that initiates reductive cleavage of SAM to form L-methionine and a 5'-deoxyadenosine radical (figure 2.30). The radical species then abstracts a hydrogen atom from the unactivated carbon centre that is undergoing functionalization, to form a radical species that attacks the methyl group of a methylcobalamin cofactor. The methylated product is formed *via* homolytic cleavage of the Me-Co bond.

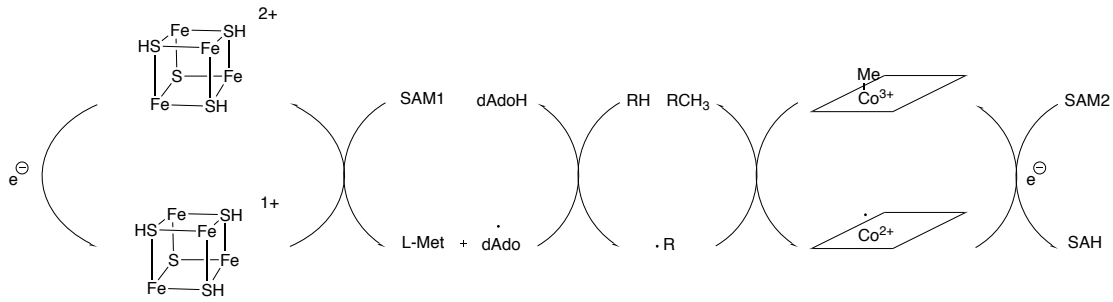
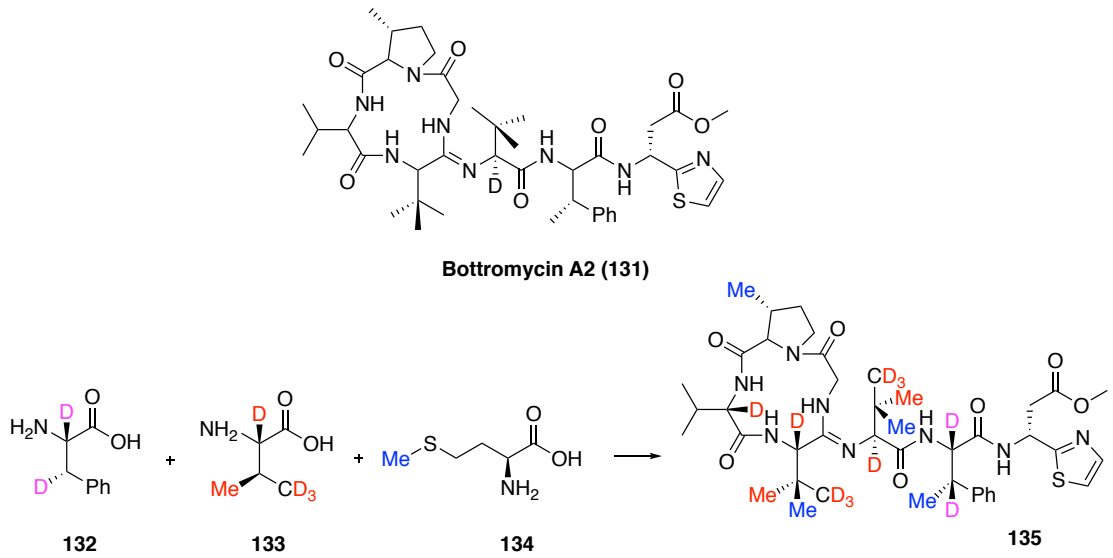



**Figure 2.30:** The mechanism of class B radical SAM methylases.

### 2.7.4 Stereocontrol of class B radical SAM methylases in the biosynthesis of bottromycin (131)

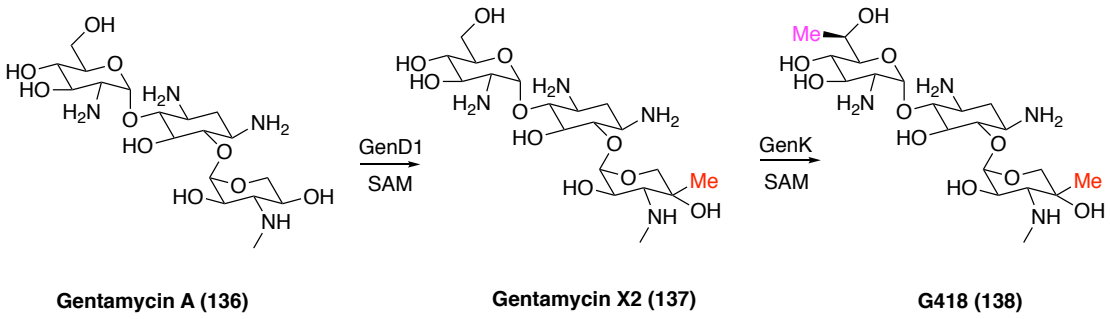
The stereochemical course of the β-methylation of phenylalanine, proline and valine by Bmb/BotRMT1, BmbF/BotRMT2 and BmbJ/BotRMT3 in the biosynthesis of bottromycin A2 (**131**) respectively, has been investigated using stereospecifically labelled synthetic mimics (scheme 2.10).[127–130] L-phenylalanine (**132**) stereospecifically labeled with deuterium at the β-position, and valine containing a deuterium label and $^{13}$C-labeled methyl group (**133**) are both incorporated into bottormycin A2 (**143**) with opposite final stereochemistry, indicating that the methylation reaction occurs with inversion.




**Scheme 2.10:** Incorporation experiments utilizing stereospecifically labeled substrates to furnish labeled-bottromycin A2 (**135**) and probe the stereochemical course of bottromycin A2 (**131**) biosynthesis.

### 2.7.5 Stereocontrol of class B methylases in gentamycin biosynthesis

The biosynthetic pathway towards the antibiotic gentamycin contains a class B radical SAM methylase, GenD1, that converts gentamycin A (**136**) to gentamycin X2 (**137**) with inversion of stereochemistry (scheme 2.11).[131] GenK catalyzes a second methylation of **137** to furnish **138**. As the carbon atom to be functionalized in **137** is achiral, the stereochemical course of the GenK-catalyzed methylation is not known.




**Scheme 2.11:** The biosynthesis of gentamycin X2 (**137**) and G418 (**138**) from radical SAM methylases GenD1 and GenK.

Therefore, these data provide further evidence for the proposed orientation of the undecylprodigiosin (**23**) alkyl chain and the reactive iron-oxo species as highlighted in figure 2.31. The orientation is predicted to be similar to that proposed for class B radical SAM methylases. In both cases, the resulting transformation proceeds with inversion of stereochemistry.

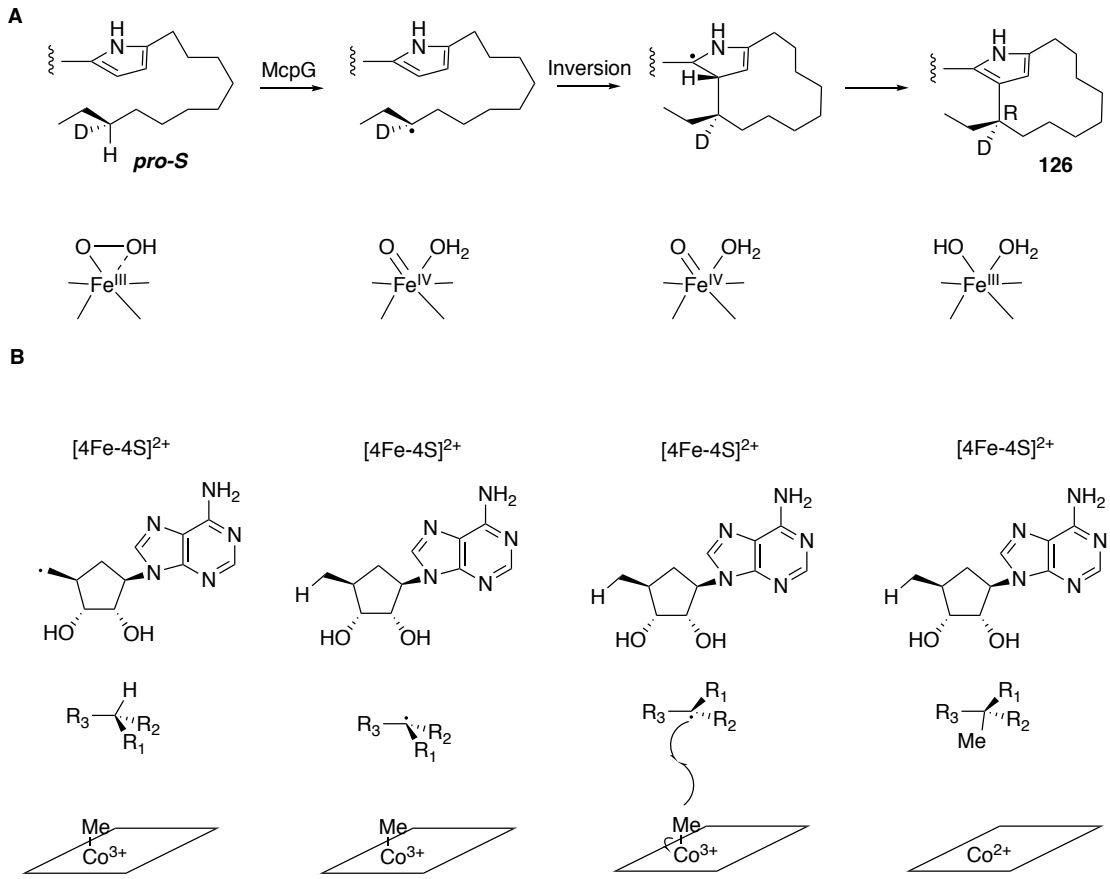


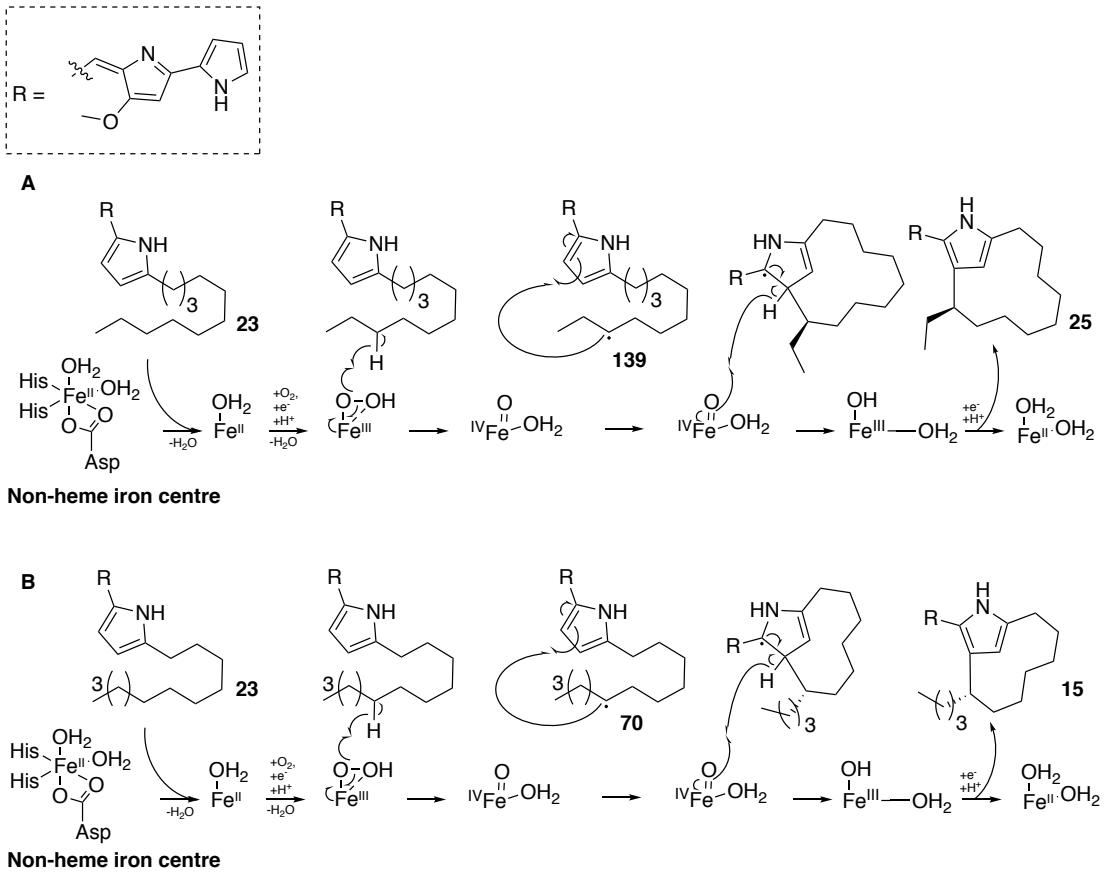

**Figure 2.31: A.** The oxidative carbocyclization reaction catalyzed by McpG, highlighting the proposed origin of the inversion of stereochemistry; **B.** The proposed reaction orientation of class B radical SAM methylases, highlighting the inversion of stereochemistry.

# Chapter 3: Mechanism and substrate tolerance of McpG

## 3.1 Introduction

### 3.1.1 Proposed mechanism for McpG-catalyzed carbocyclization

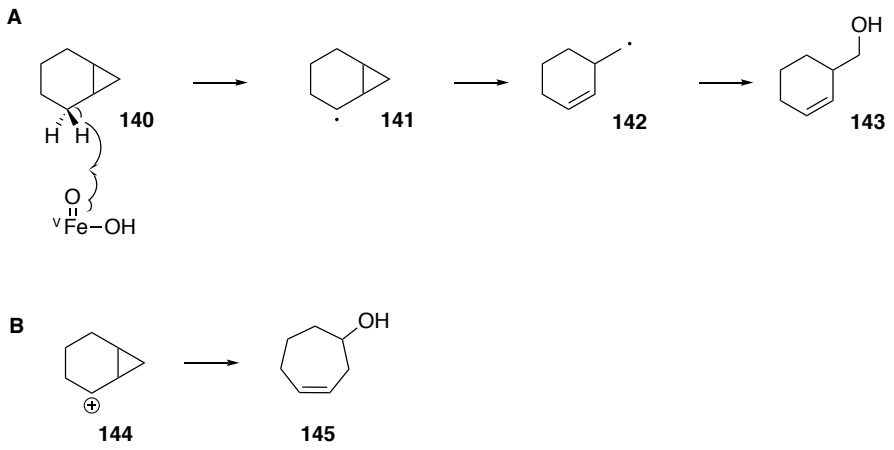
The proposed mechanism of RedG cyclization is based on the catalytic mechanism of NDO in the dihydroxylation of naphthalene (**65**) (scheme 3.1).[46,49,86,87] Based on work in the previous chapter, an analogous mechanism can also be proposed for McpG catalysis in which an iron-oxo species is responsible for the abstraction of a proton from the undecylprodigiosin (**23**) alkyl chain, to form radical intermediate **139** that undergoes cyclization with the proximal pyrrole moeity (scheme 3.1).




**Scheme 3.1:** The proposed mechanism of oxidative carbocyclization utilizing molecular oxygen, a non-heme iron centre and a Rieske cluster; (**A**) McpG-catalyzed to furnish metacycloprodigiosin (**25**); (**B**) RedG-catalyzed to furnish streptorubin B (**15**).

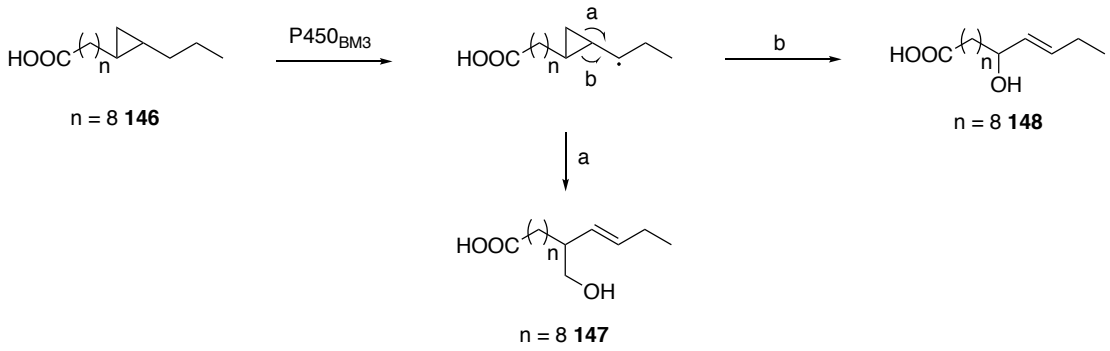
### 3.1.2 Cyclopropanes as traps for radical intermediates

Cyclopropane groups are often used as mechanistic probes to trap radical intermediates. Lipscomb and co-workers used the cyclopropane-containing norcarane (**140**) as a mechanistic probe to investigate the potential radical intermediates involved in NDO catalysis (scheme 3.2).[79] The assay furnished products that were consistent with rearrangement of the cyclopropylcarbinyl radical (**141**) and inconsistent with cationic intermediate (**144**). Ring opening of the ring-strained cyclopropane by the cyclopropylcarbinyl radical gives a primary radical (**142**) that is subsequently hydroxylated to furnish (**143**), potentially from rebound back onto an iron-oxo complex (scheme 3.2).




**Scheme 3.2: A.** Norcarane (**140**) mechanistic probe used to investigate NDO catalysis and the resulting detected product **143**; **B.** The alternative cationic intermediate (**144**) and rearrangement product (**145**) that were not observed experimentally.

The Cryle group used cyclopropyl probes to investigate the hydroxylase enzyme $P450_{BM3}$, responsible for catalyzing hydroxylation of branched fatty acids.[132] Similar to Lipscomb and co-workers the cyclopropane was located geminal to the proposed carbon centred radical. The resulting products were also consistent with radical ring opening followed by hydroxylation (scheme 3.3).




**Scheme 3.3:** Cyclopropane-containing probe used to investigate potential radical intermediates involved in $P450_{BM3}$ enzyme catalysis and the resulting products.

Although it is proposed that the RedG-catalyzed C-H abstraction generates a carbon-centred radical (**70**) and an Fe(IV)O($OH_2$) complex, it is equally plausible that the iron-oxo species rearranges to an Fe(V)O(OH) complex and that a hydride is abstracted to form a C7 carbocation (**73**). Withall attempted to probe this question by designing a cyclopropane-containing 2-UP analogue (**149**) to trap the proposed carbon-centred radical.[89] Feeding **149** to *S. albus/redHG* generated the corresponding undecylprodigiosin analogue **150** but did not furnish any RedG derived compounds (scheme 3.4). It is hypothesized that the increase in steric bulk in the central region of the hydrocarbon chain prevents **150** from adopting the correcting orientation in the RedG active site and it is therefore not a substrate.

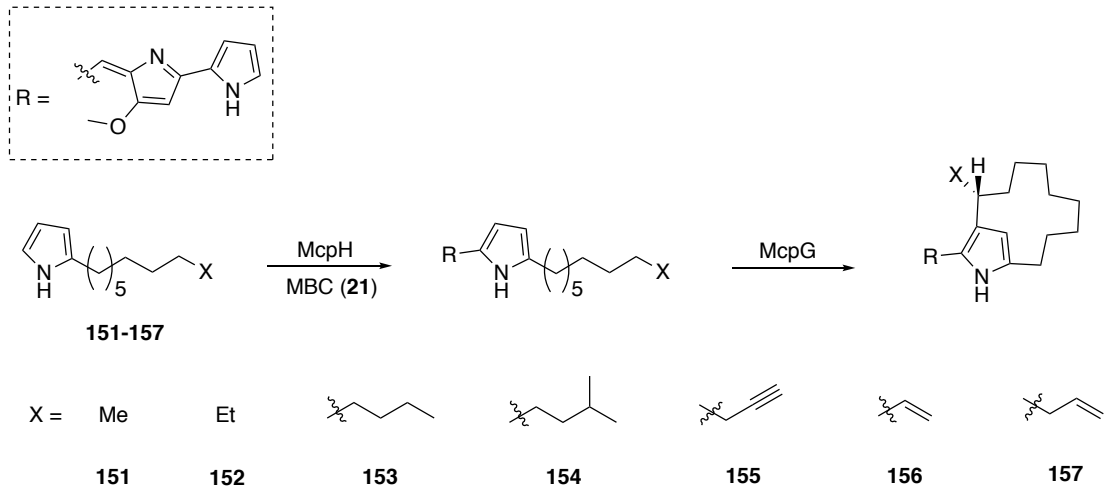

**Scheme 3.4:** The results from feeding cyclopropane-containing 2-UP analogue **160** to *S. albus* expressing *redHG*.

### 3.1.3 Aims & objectives

RedG has been shown to inefficiently accept additional steric bulk in the vicinity of C-H abstraction, and so further probing of the mechanism of carbocyclization was not possible as this limits the range of 2-UP analogues that can be chemically synthesized. McpG catalyzes cyclization by abstracting a proton from C9 rather than C7. As C9 is more terminal than C7 the additional steric bulk arising from added functionality on the mechanistic probes is less likely to affect active site binding. Various 2-UP analogues were therefore synthesized and fed to *S. albus*/pCAP1000-*mcp*BGC-Δ*LK*-pOSV556t-*mcpHG* to further understand the mechanism of Rieske oxygenase catalyzed oxidative carbocyclization, using McpG.

## 3.2 Investigating the substrate tolerance of McpG-catalyzed oxidative carbocyclization

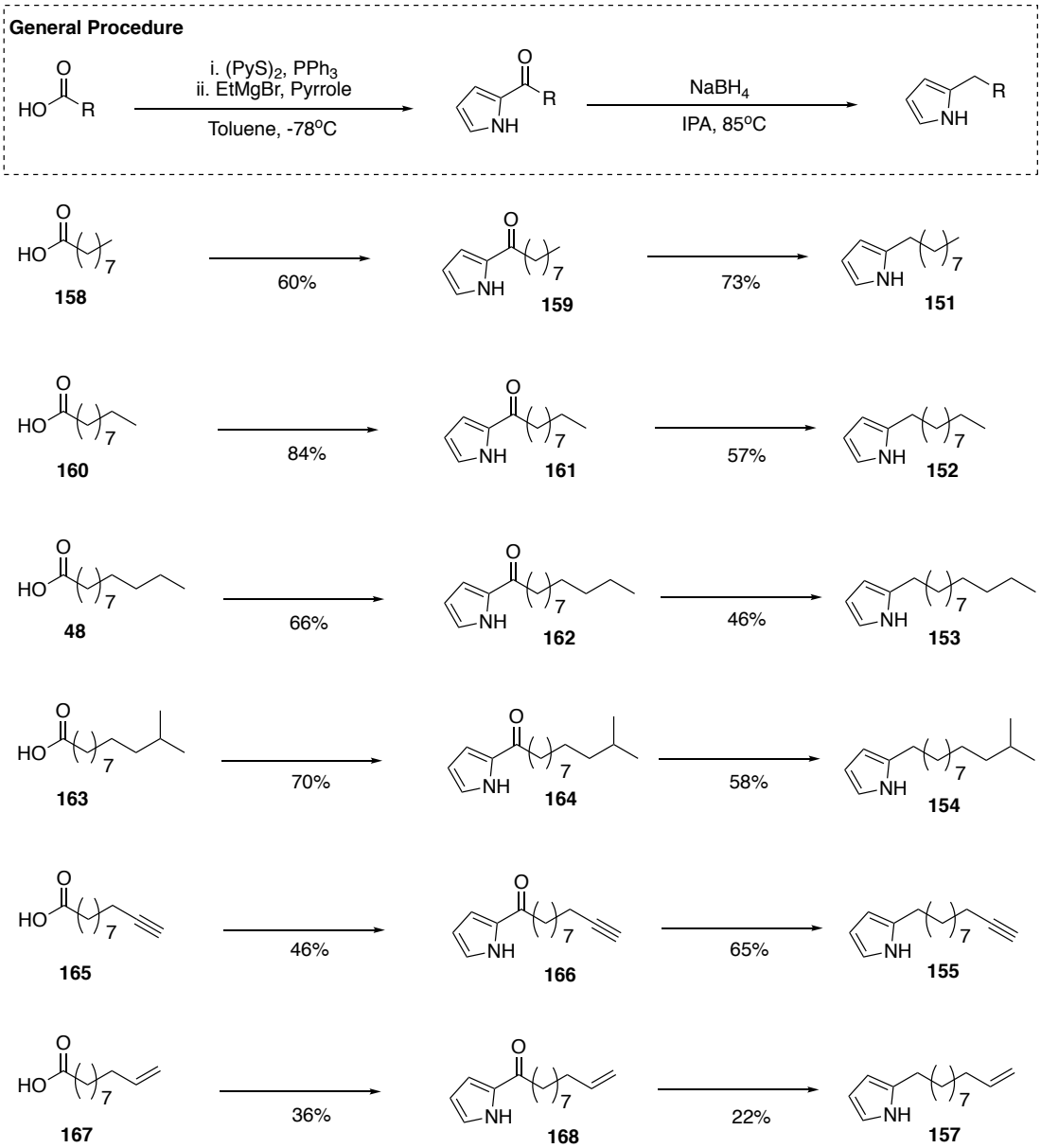
To investigate the substrate tolerance of the McpG-catalyzed oxidative carbocyclization a number of 2-UP analogues were synthesized with increased and decreased chain lengths, additional branching and with increased desaturation. It was envisaged that the analogues would be condensed with MBC (**21**) to furnish the corresponding undecylprodigiosin analogue prior to McpG-catalyzed cyclization (scheme 3.5).




**Scheme 3.5:** Strategy for investigating substrate tolerance of McpG through the synthesis of a number of 2-UP substrate analogues (**151-157**).

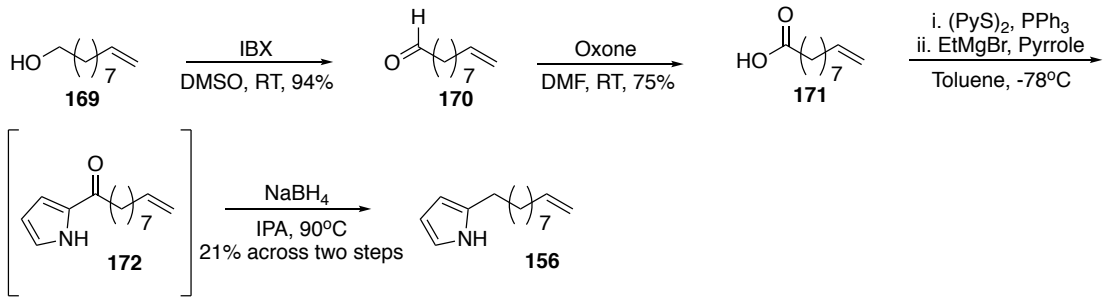
### 3.2.1 Synthesis of 2-UP analogues

The 2-UP analogues used to investigate the substrate tolerance of McpG were synthesized from their corresponding commercially available carboxylic acids (scheme 3.6). The carboxylic acid was coupled to freshly distilled pyrrole using the procedure outlined in section 6.2.6, followed by reduction of the resulting 2-ketopyrrole species using sodium borohydride in refluxing propan-2-ol.




**Scheme 3.6:** The synthesis of 2-UP analogues from their corresponding commercially available carboxylic acid.

A 2-UP analogue (**156**) in which the site of C-H abstraction is $sp^2$-hybridized was synthesized using an alternative route, as the corresponding carboxylic acid was not available for purchase. The synthesis was therefore undertaken starting with the commercially available dec-9-en-1-ol (**169**) (scheme 3.7). The alcohol **169** was oxidized to the corresponding aldehyde **170** by treatment with neat IBX. Carboxylic acid **171** was furnished by reaction of aldehyde **170** with oxone. Compound **171** was coupled to freshly distilled pyrrole and subsequently reduced using borohydride in refluxing 2-propanol to generate the target compound **156**, with a terminal C=C bond.




**Scheme 3.7:** The synthesis of a 2-UP **156** analogue where the site of C-H abstraction is $sp^2$-hybridized.

### 3.2.2 Changing the length of the alkyl chain

The McpG-catalyzed oxidative carbocyclization was investigated using undecylprodigiosin analogues with varying alkyl chain lengths. The synthesis of 2-UP analogues with nine (**151**), ten (**152**) and twelve (**153**) length carbon chain was undertaken and fed to the *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* deletion mutant alongside synthetic MBC (**21**), subsequently furnishing the corresponding undecylprodigiosin analogue to be cyclized by McpG.

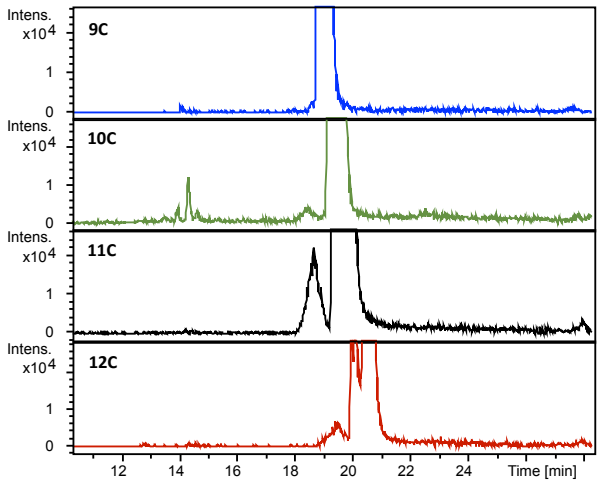
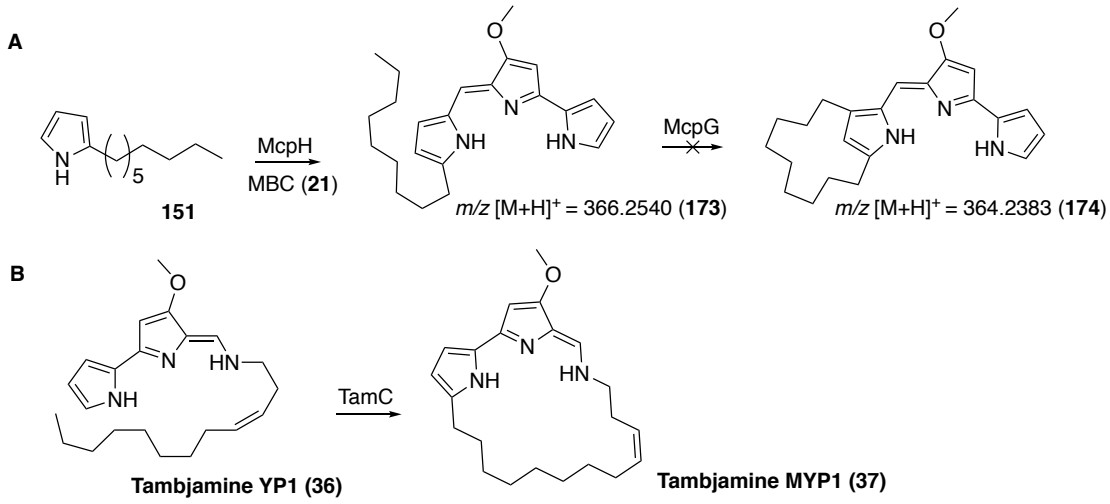



**Figure 3.1:** EICs from the LC-MS analysis of the mycelial extract resulting from feeding 2-UP (**28**) (or analogue) alongside MBC (**21**) to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **Blue trace:** Feeding nine-carbon chain 2-UP analogue (**151**): EICs at *m/z* = 366.25 and 364.24 ±0.01; **Green trace**: Feeding ten-carbon chain 2-UP analogue (**152**): EICs at *m/z* = 380.27 and 378.25 ±0.01; **Black trace:** Feeding 2-UP (**28**): EICs at *m/z* = 394.29 and 392.27 ±0.01 corresponding to the calculated $[M+H]^+$ of undecylprodigiosin (**23**) and metacycloprodigisoin (**25**) respectively; **Red trace:** Feeding twelve-carbon chain 2-UP analogue (**153**): EICs at *m/z* = 408.30 and 406.29 ±0.01 (**LHS (minor peak)** = product of McpG-catalyzed cyclization; **RHS (major peak)** = product of McpH-catalyzed condensation).

The results from feeding 2-UP analogue with a nine-carbon alkyl chain (**151**) to the *mcpLK* deletion mutant highlighted the successful McpH-catalyzed condensation with MBC (**21**), (figure 3.1). Subsequent McpG-catalyzed carbocyclization only furnished trace amounts of the cyclic analogue **174**, as seen by the low peak intensity.

These data can be rationalized as McpG-catalyzed C-H abstraction occurs at C9 of the alkyl chain. In the case of undecylprodigiosin analogue **173**, McpG must catalyze a much less energetically favourable C-H abstraction that furnishes as primary radical rather than a secondary radical, as in the case for the natural substrate **23** (scheme 3.8). The less favourable C-H abstraction therefore results in lower levels of the subsequent cyclization reaction. Contrastingly, as discussed in section 1.3.1, Ross and co-workers have reported that Rieske oxygenase TamC is proposed to catalyze a C-H abstraction and furnish a primary radical, in the biosynthesis of tambjamine MYP1 (**37**) (scheme 3.8).[36] TamC may therefore catalyze C-H abstraction *via* a different mechanism to RedG/McpG, which offers an alternative lower energy route to generate the reactive species, prior to cyclization.

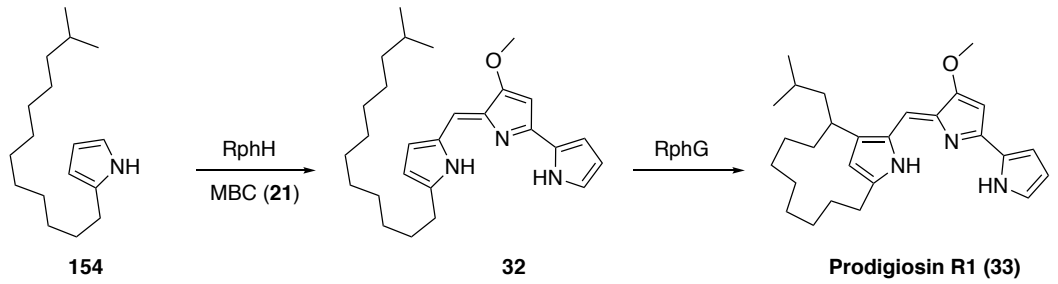



**Scheme 3.8: A.** The results from feeding nine-carbon chain 2-UP analogue (**151**) to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **B.** The TamC-catalyzed cyclization to furnish tambjamine MYP1 (**37**).

The results from feeding 2-UP analogues varying in alkyl chain length show that the synthetic (natural) precursor 2-UP (**28**) produces metacycloprodigiosin (**25**) with the highest peak intensity (figure 3.1). Feeding a nine-carbon chain 2-UP analogue (**151**) resulted in only trace amounts of the corresponding cyclic product **174**. Feeding ten-carbon chain analogue **152** and twelve-carbon chain analogue **153** generated similar peak intensities for their final cyclic products, however both lower than the natural eleven-carbon chain precursor 2-UP (**28**).

### 3.2.3 Addition of branching

Prodigiosin R1 (**33**) is a prodiginine alkaloid produced by *S. griseoviridis* and the biosynthetic gene cluster contains four RedG homologous, RphG, RphG2, RphG3 and RphG4. Expression of *rphG* in a *redG* mutant of *S. coelicolor* resulted in the production of metacycloprodigiosin (**25**) and propyl-meta-cyclooctylprodiginine, demonstrating that RphG catalyzes the conversion of 11-methyldodecylprodigiosin (**32**) to prodigiosin R1 (**33**).[32,52] Both MBC (**21**) and 11-methyldodecylprodigiosin (**32**) are late stage intermediates in the biosynthetic pathway for prodigiosin R1 (**33**) (scheme 3.9). As RphG and McpG are similar in sequence, the ability of McpG to accept **32** was investigated.




**Scheme 3.9:** The cyclization of 11-methyldodecylprodigiosin (**32**) in the biosynthesis of prodigiosin R1 (**33**).

The mutasynthesis of undecylprodigiosin analogue **32**, resulting from feeding synthetic 11-dimethyl-2-UP analogue (**154)** alongside MBC (**21**) to the *mcpLK* deletion mutant would provide the direct substrate for oxidative carbocyclization to prodigiosin R1 (**33**). Therefore the synthesis of target compound **154** was undertaken using 11-methyldodecanoic acid (**163**) (provided by Dr. Douglas Roberts) which was coupled to freshly distilled pyrrole to furnish compound (**164**) with subsequent borohydride reduction using a identical method as described for previous analogues to provide the final analogue **154** (scheme 3.6).

The 2-UP analogue **154** was fed along side MBC (**21**) to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* grown for two days on a permeable membrane on solid SFM media. The plates were incubated for a further two days before the biomass was collected and the metabolites were extracted using acidified methanol. Compound **32** displayed in chromatogram **C** (Figure 3.2) confirms the formation of the 11-methyldodecylprodigiosin (**32**) as a result of McpH-catalyzed condensation of synthetic analogue **154** and MBC (**21**). Chromatogram **D** (figure 3.2) indicates that prodigiosin R1 (**33**) was not produced as a result of McpG-catalyzed oxidative carbocyclization of **32**.

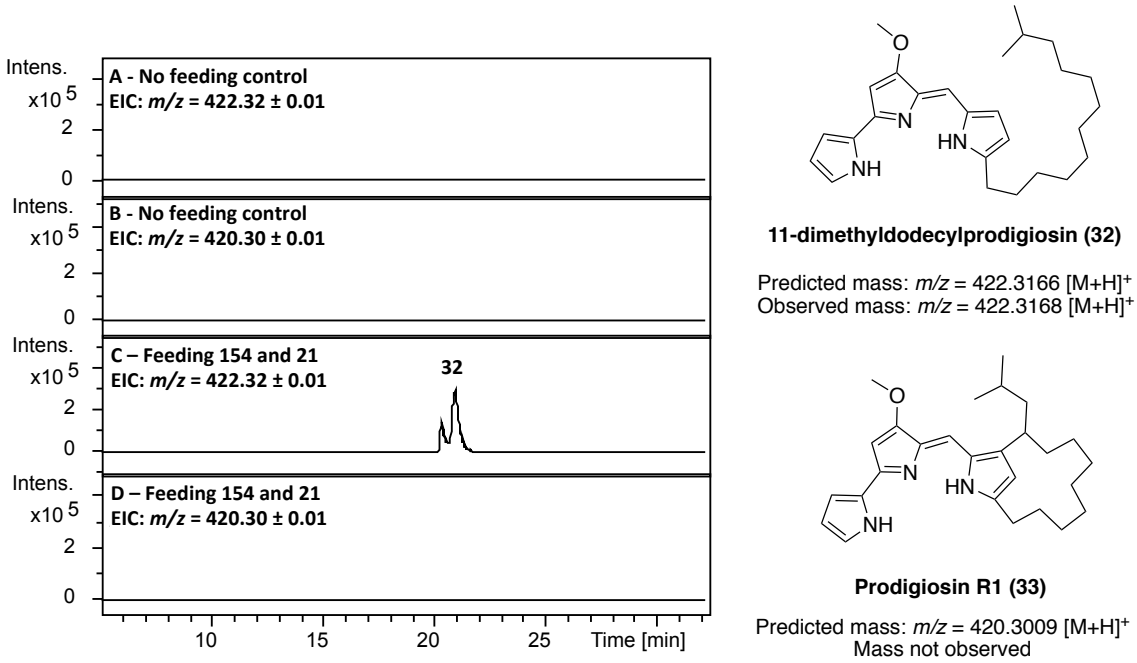
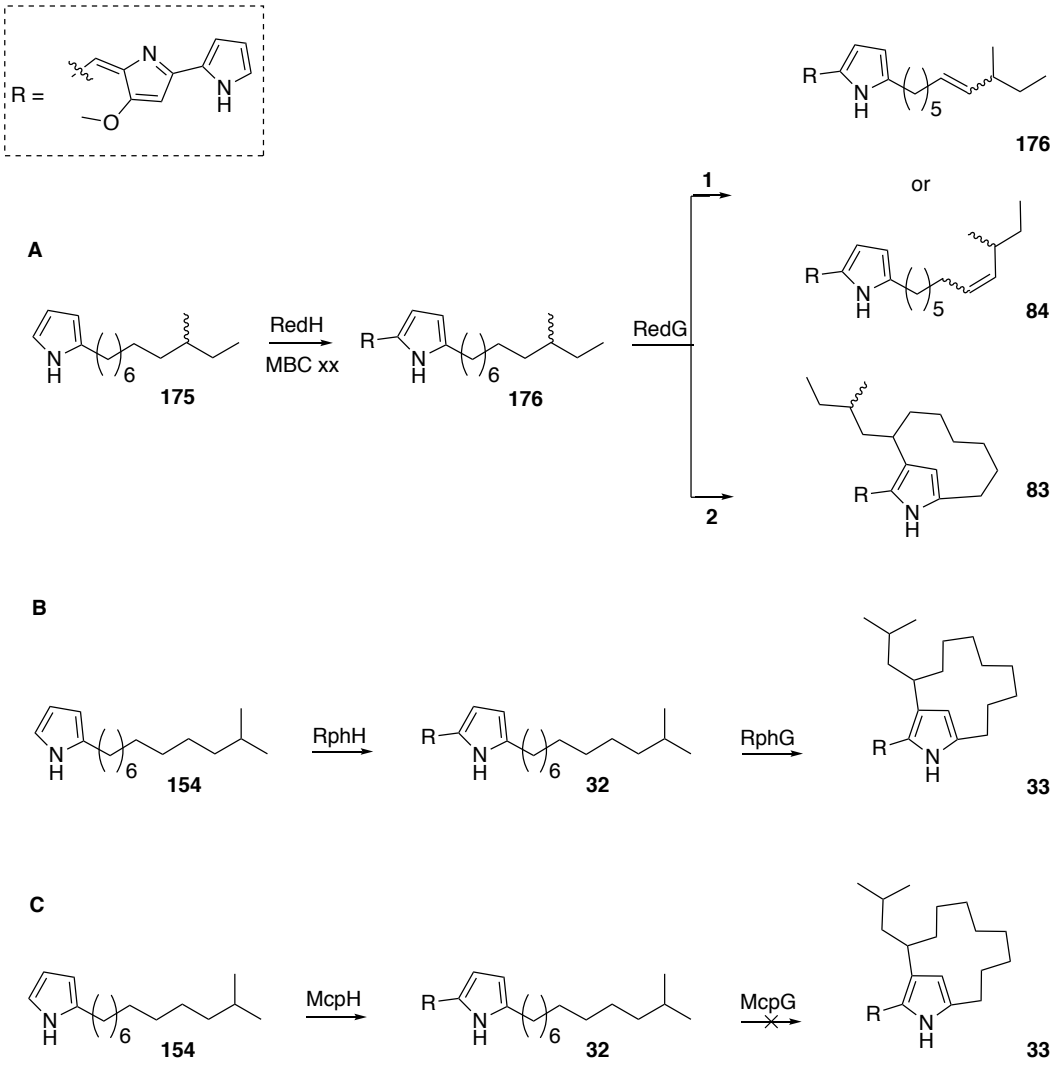



**Figure 3.2:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 422.32 ±0.01 corresponding to the $[M+H]^+$ ion of **32**: No feeding control; **B.** EIC at *m/z* = 420.3009 ±0.01 corresponding to the $[M+H]^+$ ion of **33**: No feeding control; **C.** EIC at *m/z* = 422.32 ±0.01 corresponding to the $[M+H]^+$ ion of **32**: Feeding 2-UP analogue **154** and MBC **(21)**; **D.** EIC at *m/z* = 420.30 ±0.01 corresponding to the $[M+H]^+$ ion of **33**: Feeding 2-UP analogue **154** and MBC **(21)**.

It is likely that the additional steric bulk arising from the additional methyl groups on carbon eleven of the alkyl chain of **32** hinders the McpG-catalyzed oxidative carbocyclization. These data underlines the difference in substrate tolerance of RphG and McpG and also contrasts with the activity previously observed using RedG. Feeding methyl-branched 2-UP analogue **175** to *S. albus* expressing *redHG* produced the corresponding undecylprodigiosin analogue **176** and the RedG-derived cyclic analogue **83** along with a desaturation product **176**.[89] This implied that the RedG-catalyzed cyclization reaction occurred at a much slower rate, allowing for a second hydrogen abstraction on the adjacent carbon atom to furnish the desaturated products (scheme 3.10). McpG, however, does not accept an analogue of undecylprodigiosin (**23**) where a methyl group had been installed at C-11 of the alkyl chain.




**Scheme 3.10:** Comparison of the results from feeding methyl branched 2-UP analogues **175** to probe RedG (**A**) and **154** to probe RphG **(B)** and McpG **(**C**)** respectively.

### 3.2.4 Desaturation

The ability of McpG to accept 2-UP analogues with different degrees of saturation was investigated by adding a double bond and a triple bond adjacent to the site of C-H abstraction (C10-C11). This will also probe the effect of adding π-bonds adjacent to the proposed radical species.

*S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* was incubated on a permeable membrane on solid SFM media for two days before synthetic 2-UP analogue **155**, containing a terminal acetyline moeity, and synthetic MBC (**21**) were added directly onto the mycelia. The plates were incubated for a further two days before the mycelium was extracted and metabolites extracted for HR LC-MS analysis.

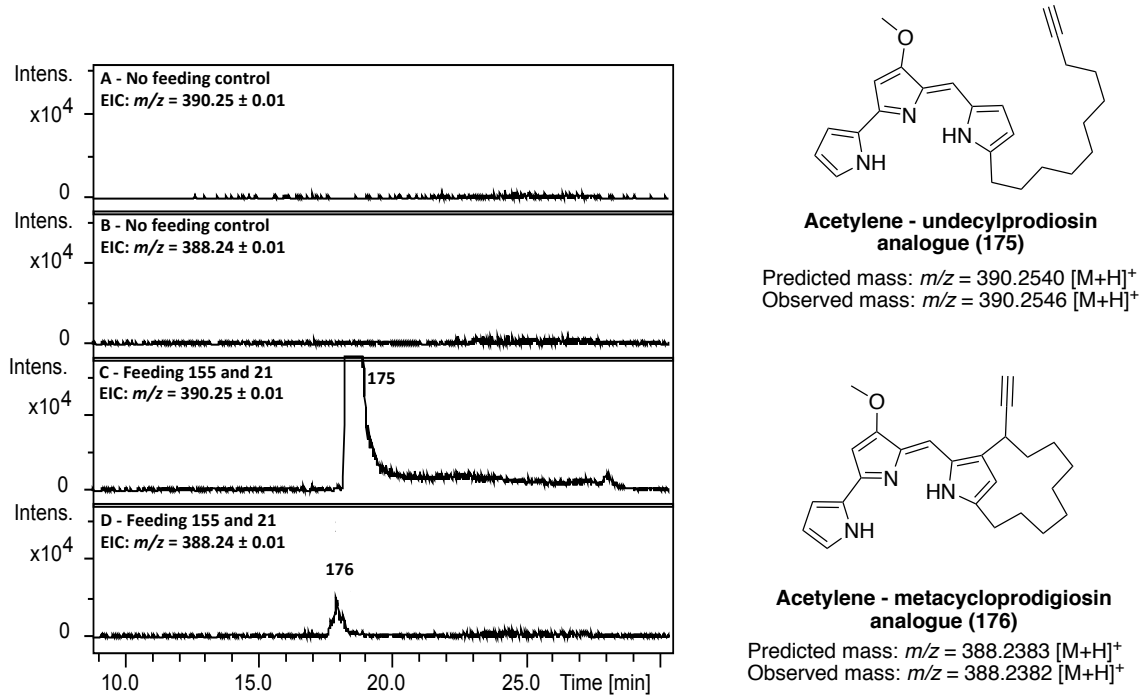
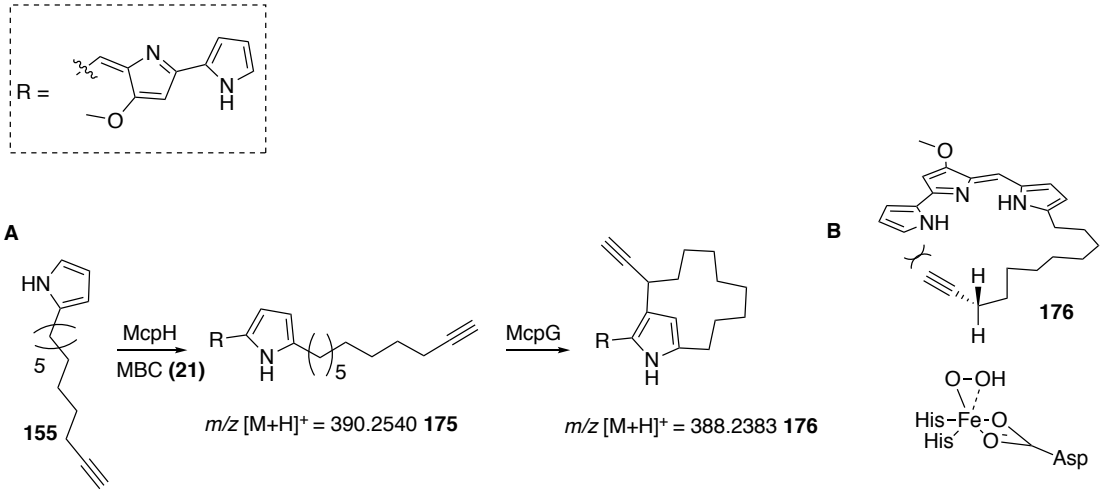



**Figure 3.3:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **175**: No feeding control; **B.** EIC at *m/z* = 388.24 ±0.01 corresponding to the $[M+H]^+$ ion of **176** No feeding control; **C.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **175**: Feeding acetylene 2-UP **155** and MBC (**21**); **D.** EIC at *m/z* = 388.24 ±0.01 corresponding to the $[M+H]^+$ ion of **176**: Feeding acetylene 2-UP **155** and MBC (**21**).

The data displayed in chromatogram **C** (figure 3.3) indicates that 2-UP analogue **155** is successfully condensed with synthetic MBC (**21**) in a reaction catalyzed by McpH. The compound **176** observed in chromatogram **D** (figure 3.3) shows that McpG successfully catalyzes the conversion of the acetyline undecylprodigiosin analogue **175** into the cyclic

derivative, although in low yield (scheme 3.3). The low peak intensity of the cyclized analogue **176** may be due to the electronic repulsion between the electron-rich acetylene and the electron-rich pyrrole core scaffold (scheme 3.11). The low pKa of the acetyline C-H may also contribute towards degradation of the Rieske cluster and/ or non-heme iron centre of McpG.




**Scheme 3.11: A.** Summary of the results from observed from feeding 2-UP analogue **155** and MBC **(21)** to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **B.** The proposed orientation of the undecylprodigiosin analogue **175** during cyclization, indicating potential electronic repulsion / steric hindrance.

The 2-UP analogue **157** containing a terminal C=C bond was fed directly to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* mycelium that had been grown for two days prior. The plates was grown for a further two days post-feeding and the mycelial extract was analyzed by HR LC-MS. Chromatogram **C** (figure 3.4) highlights the successful formation of undecylprodigiosin analogue **177** as a result of McpH catalysis. The cyclic metacycloprodigiosin analogue **178** was not produced. These data suggests that McpG-catalyzed oxidative carbocyclization reaction occurred at a lower rate for **177** relative to natural undecylprodigiosin (**23**).

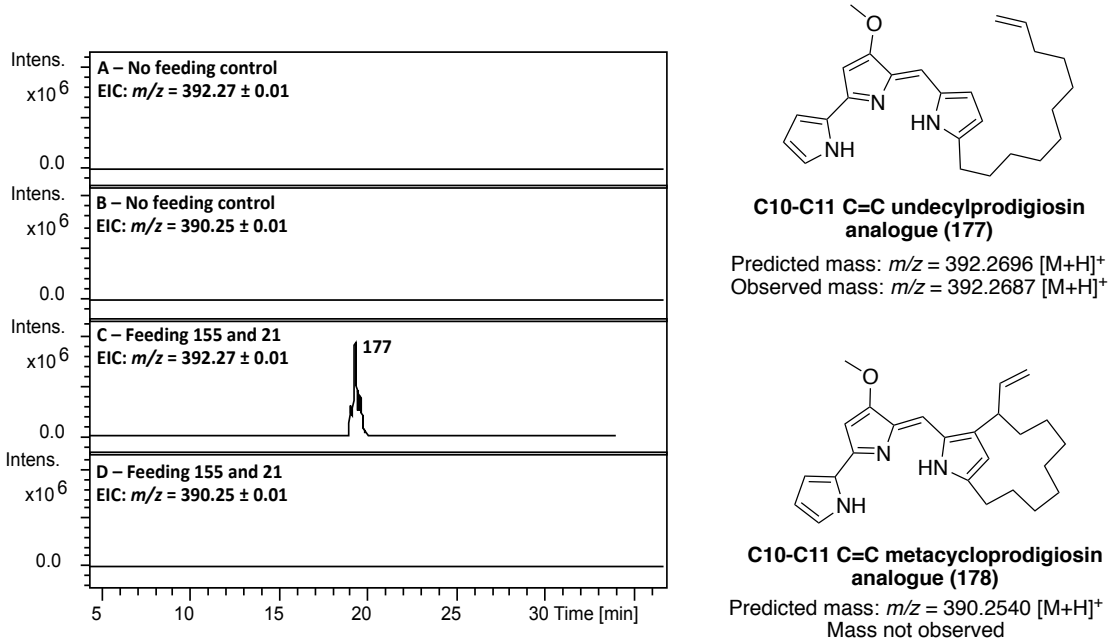
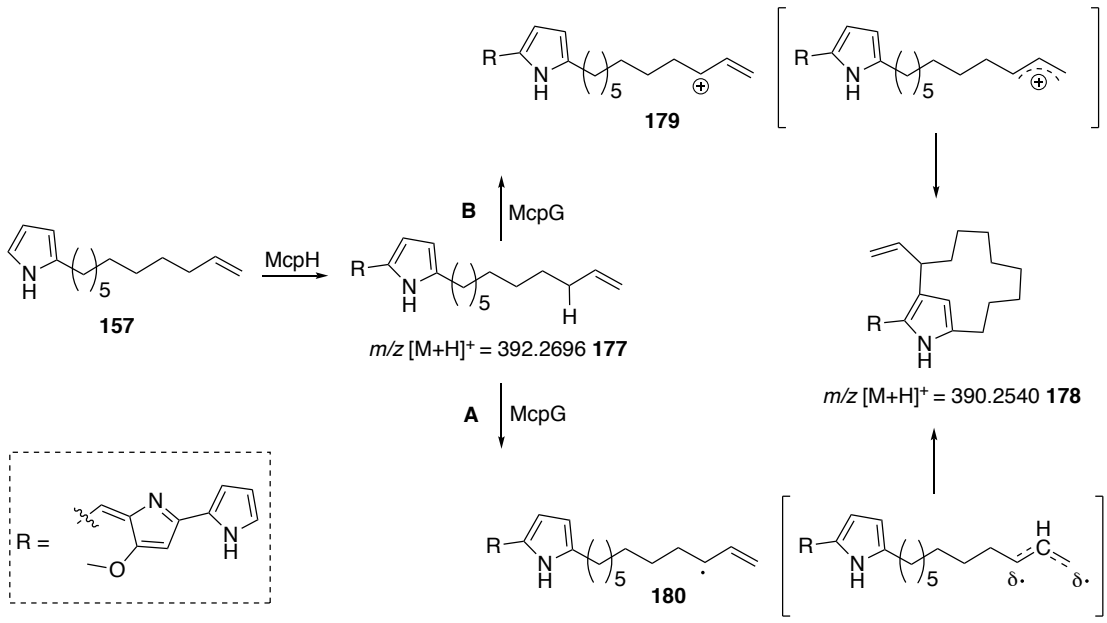



**Figure 3.4:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 392.27 ±0.01 corresponding to the $[M+H]^+$ ion of **177**: No feeding control; **B.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **178**: No feeding control; **C.** EIC at *m/z* = 392.27 ±0.01 corresponding to the $[M+H]^+$ ion of **177**: Feeding 2-UP analogue **157** and MBC (**21**); **D.** EIC at *m/z* = 390.25 ±0.01 corresponding to the $[M+H]^+$ ion of **178**: Feeding 2-UP analogue **157** and MBC (**21**).

It is plausible that electron-deficient radical **180** (or cation **179**) intermediate formed as a result of McpG-catalyzed C-H abstraction at C9 of the undecylprodigiosin **177** alkyl chain is stabilized by the resonance effects of the adjacent π-system of the C=C bond. The additional stabilization would subsequently reduce the rate of cyclization (scheme 3.12). It is plausible that similar stabilization effects are observed for the acetyline undecylprodigiosin analogue (**175**), and so may be a common trait for 2-UP analogues with adjacent π-systems.




**Scheme 3.12:** The results observed from feeding 2-UP analogue **157** to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** The proposed radical pathway leading to a stabilized radical intermediate; **B.** The alternative cationic pathway, leading to a stabilized cationic intermediate.

The 2-UP analogue **156** in which the site of C-H abstraction is $sp^2$-hybridized was fed to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*, alongside MBC (**21**), and incubated for two days, before collection of the resulting biomass, acidifying and methanolic extraction of the metabolites. The results showed successful McpH-catalyzed condensation by observation of compound **181** shown in chromatogram **C** (figure 3.5). However, as can be seen by chromatogram **D** (figure 3.5), the formation of cyclic analogue **182** was not observed. This observation can be explained, as C-H abstraction from an $sp^2$-hybridized carbon is energetically less favourable.

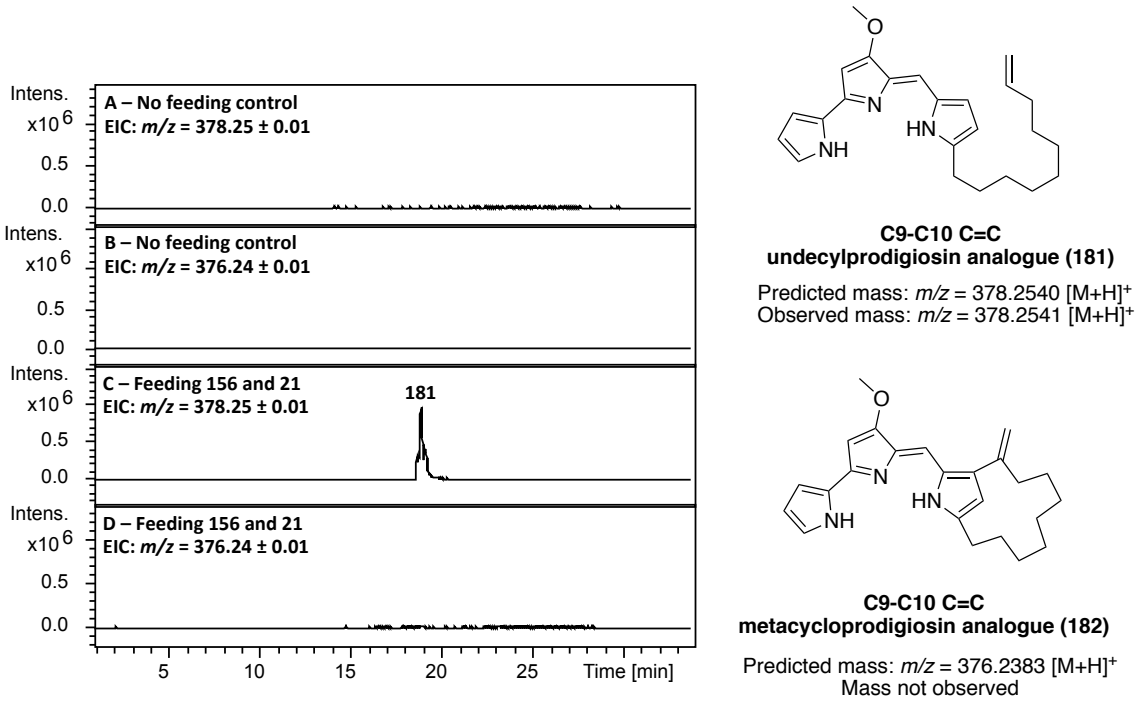



**Figure 3.5:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** EIC at *m/z* = 378.25 ±0.01 corresponding to the $[M+H]^+$ ion of **181**: No feeding control; **B.** EIC at *m/z* = 376.24 ±0. 01 corresponding to the $[M+H]^+$ ion of **182**: No feeding control; **C.** EIC at *m/z* = 378.25 ±0.01 corresponding to the $[M+H]^+$ ion of **181**: Feeding 2-UP analogue **156** and MBC **(21)**; **D.** EIC at *m/z* = 376.24 ±0.01 corresponding to the $[M+H]^+$ ion of **182**: Feeding 2-UP analogue **156** and MBC **(21)**.

Interestingly, Rieske oxygenase enzymes have been shown to catalyze a wide range of chemistries, including *C*-hydroxylation, *C*-dihydroxylation and epoxidation (figure 3.6). The highlighted moieties in figure 3.6 underlines functionalities that are believed to be installed by Rieske oxygenases, as predicted by comparative bioinformatics.[53] Such predicted Rieske oxygenase-catalyzed oxidation reactions, however, are yet to be confirmed experimentally. As McpG-catalyzed oxidative carbocyclization was not observed when feeding 2-UP analogue **156**, attentions were turned to the potential for McpG-catalyzed epoxide formation.

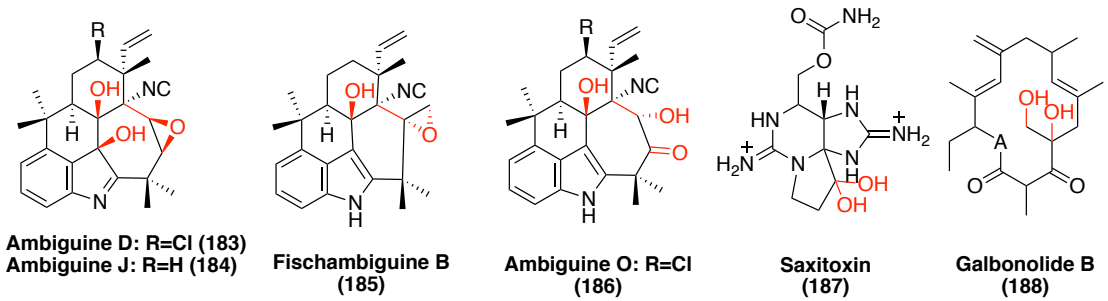



**Figure 3.6:** Natural products whose biosynthesis involves oxidation reactions catalyzed by Rieske oxygenases.

To investigate the ability of McpG to catalyze epoxidation, 2-UP analogue **156** where the site of C-H abstraction is $sp^2$-hybridized was fed to the Δ*LK* deletion mutant, alongside MBC **(21)**. The presence of a peak in chromatogram **B** (figure 3.7) corresponds to compound **189** with correct mass and chemical formulae to suggest the presence of an epoxide-containing undecylprodigiosin analogue (figure 3.7).

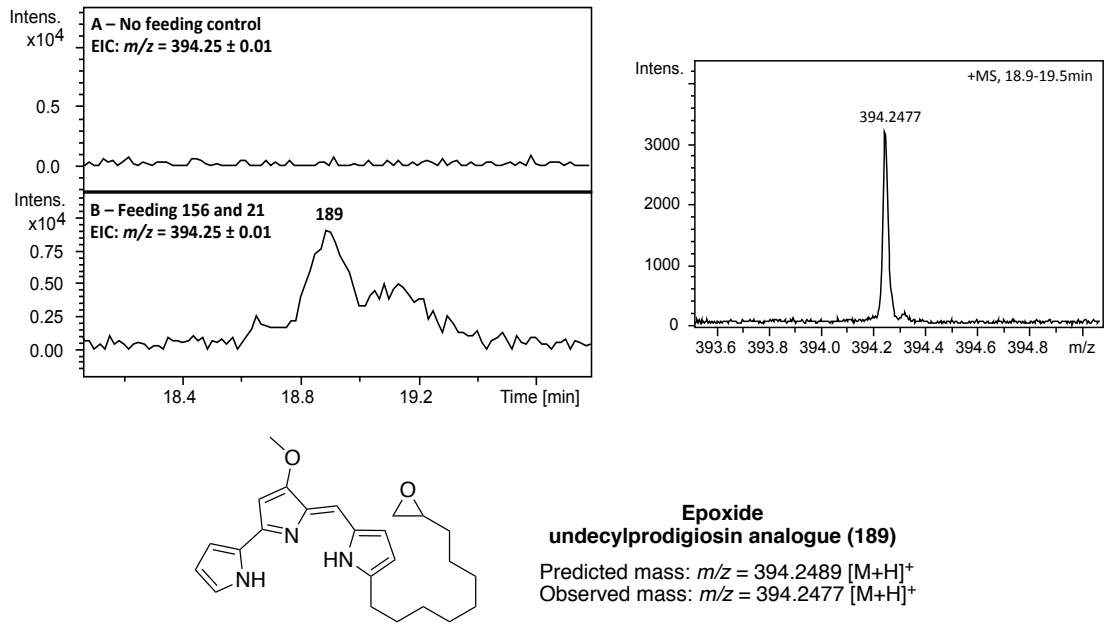



**Figure 3.7:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; EIC at *m/z* = 394.25 ±0.01 corresponding to the $[M+H]^+$ ion of **189**; **A.** No feeding control; **B.** Feeding MBC (**21**) and 2-UP analogue **156**.

As the epoxide **189** is formed in low amounts, purification/ isolation was not possible. The synthesis of an authentic standard or trapping of the epoxide using *N*-acetylcysteine (NAC), for example, would be required to confirm the structure of the compound.

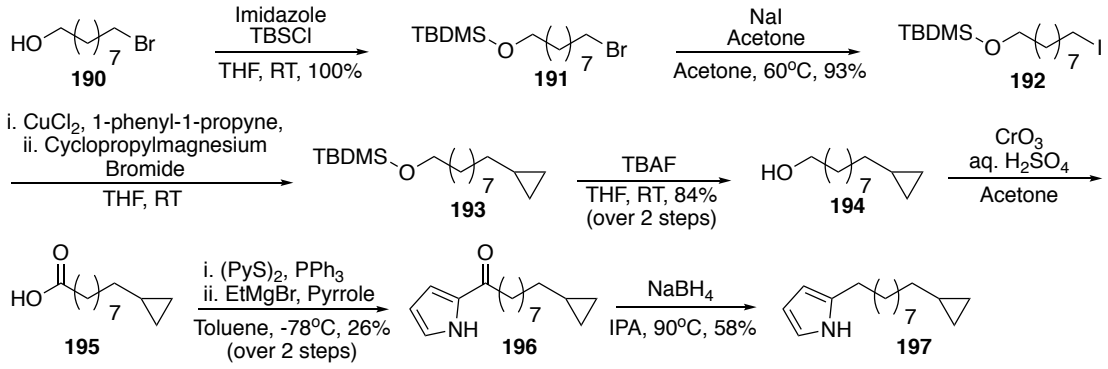
## 3.3 Investigating the product of McpG-catalyzed C-H abstraction

It was envisaged that a cyclopropane-containing 2-UP analogue could be fed to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* in order to 'trap' reactive intermediates involved in McpG catalysis. Extraction of metabolites and observation of the resulting ring-opened/ re-arranged products would provide insights into the nature of C-H abstraction. Feeding this 2-UP analogue to the deletion mutant should furnish the corresponding undecylprodigiosin analogue, following McpH-catalyzed condensation with MBC (**21**). If this compound is a substrate for McpG catalysis then the C-H abstraction will

likely result in ring-opening of the adjacent cyclopropane to form a homoallylic radical (or cationic species) that could undergo hydroxylation, desaturation or carbocyclization.

### 3.3.1 Synthesis of cyclopropane-containing 2-UP analogue 197 and subsequent feeding with MBC (21) to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*

A seven-step synthesis starting from the commercially available 9-bromononanol (**190**) was undertaken to provide the target cyclopropane-containing 2-UP analogue **197** (scheme 3.13). Silyl protection of the primary alcohol furnished bromide **191**. A subsequent Finkelstein reaction formed iodide **192**, before the cyclopropane moeity was installed through a copper-mediated Grignard reaction using cyclopropylmagnesium bromide. Fluoride promoted silyl deprotection generated the primary alcohol **194**, which was oxidized to the correspondingly carboxylic acid **195** using Jones' reagent generated *in situ*. The acid **195** was coupled to freshly distilled pyrrole to generate the 2-ketopyrrole **196** and a final borohydride reduction furnished the target product **197**.




**Scheme 3.13:** The synthesis of cyclopropane-containing 2-UP analogue (**190**).

The 2-UP analogue **197** was fed along side MBC (**21**) to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*. The resulting mycelium was collected and metabolites were extracted using acidified methanol. The extract was analyzed by HR LC-MS. Chromatogram **B** (figure 3.8) confirms the presence of the cyclopropane-containing undecylprodigiosin analogue **198** as a result of the McpH catalyzed condensation of 2-UP analogue **197** and MBC **(21)**. No compounds relating to cyclization of **198** were evident.

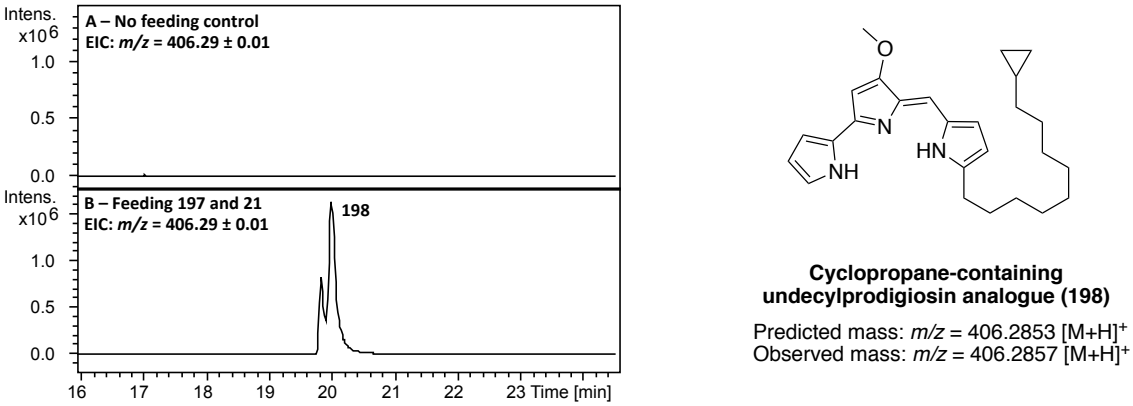



**Figure 3.8:** EICs from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; EIC at *m/z* = 406.29 ±0.01 corresponding to the $[M+H]^+$ ion of **198**; **A.** No feeding control; **B.** Feeding 2-UP analogue **197** and MBC **(21)**.

Comparison of the base peak chromatograms (figure 3.9) of the extracts arising from feeding cyclopropane-containing 2-UP analogue **197** and MBC (**21**), with the negative control highlighted the formation of a compound with an *m/z* = 364.2385 and molecular formula of $C_{23}H_{30}N_3O$. These data suggest the loss of the cyclopropane moeity with subsequent desaturation or cyclization. As cyclization of an undecylprodigiosin analogue with a nine-carbon chain has been shown to be unfavorable (section 3.2.2), it is likely desaturation follows loss of the cyclopropane moeity, to form compound **199**.

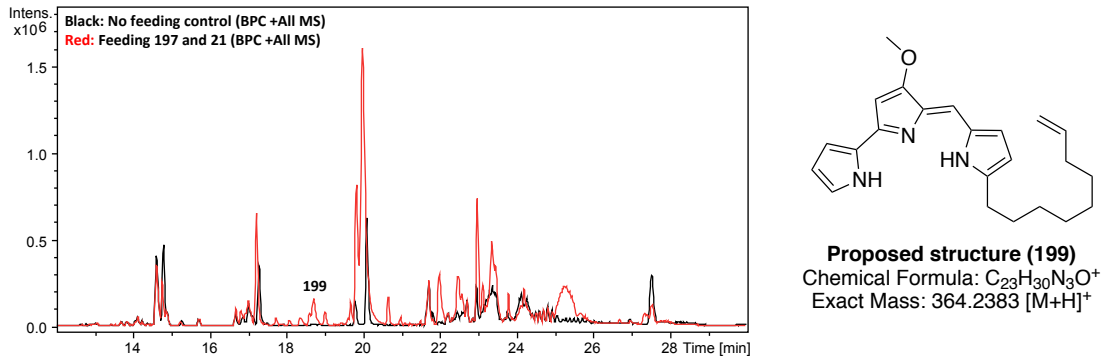
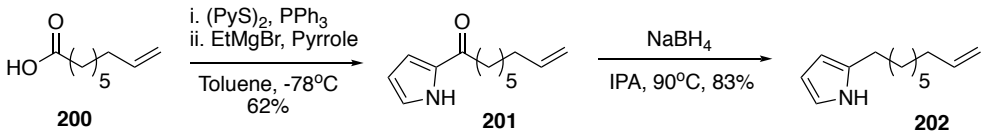



**Figure 3.9:** Base peak chromatogram from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* arising from feeding cyclopropane 2-UP analogue **190** and MBC (**21**) (*red*) and the no feeding negative control (*black*).

The proposed product **199** was synthesized *via* mutasynthesis, by feeding the corresponding synthetic 2-UP analogue (**202**) alongside MBC (**21**) to the *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* strain. The 2-UP analogue **202** was synthesized from the corresponding carboxylic acid **200** using the previously established procedure (scheme 3.14).




**Scheme 3.14**: The synthesis of 2-UP analogue **202**.

Comparison of the extracted ion chromatograms at *m/z* = 364.24 ± 0.01 for both extracts revealed that both compounds had the same retention time (figure 3.10). Scale-up and purification of the compound for analysis by NMR would be required to conclusively confirm the structure. The mechanism by which this compound is formed is unclear.

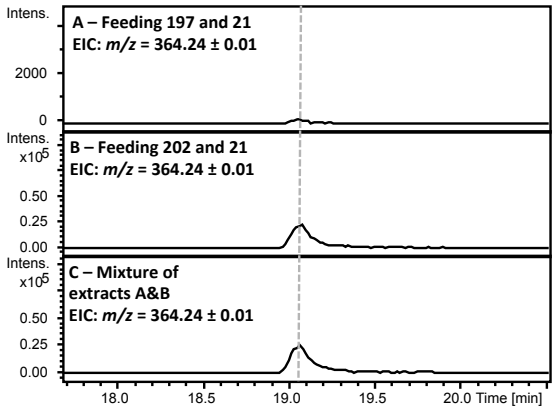



**Figure 3.10:** EIC at *m/z* = 364.24 ±0.01 corresponding to the $[M+H]^+$ ion of **199** from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** Feeding 2-UP analogue **197** and MBC (**21**); **B.** Feeding 2-UP analogue **202** and MBC (**21**); **C.)** Mixture of both extracts.

Numerous reports in the literature describe the ring-opening and hydroxylation of cyclopropane probes that trap reactive intermediates. By searching for the *m/z* of the possible 'trapped' intermediates, a peak was found with the formulae for the alkene **203**, that could be formed *via* ring-opening of the cyclopropane and subsequent hydroxylation (figure 3.11). Attempts to purify the compound were unsuccessful and so an authentic standard of **203** was synthesized.

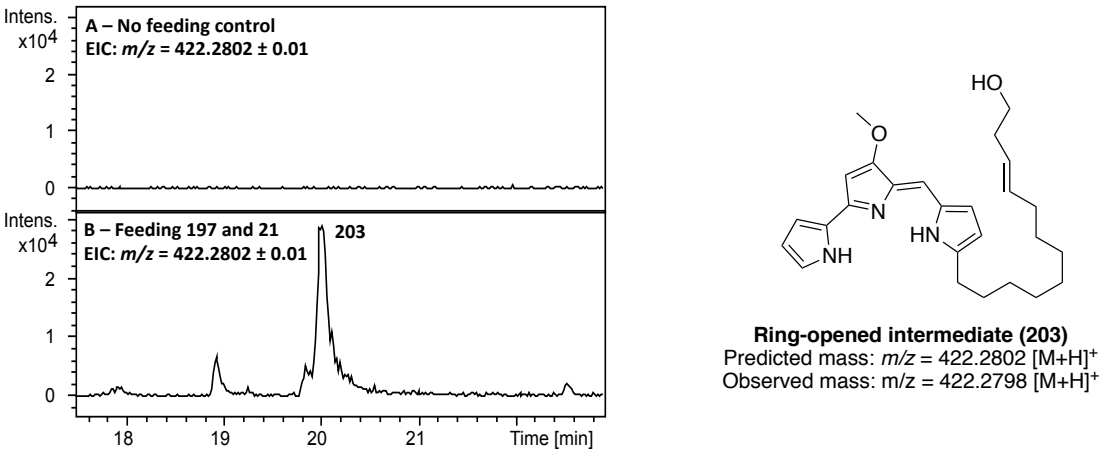
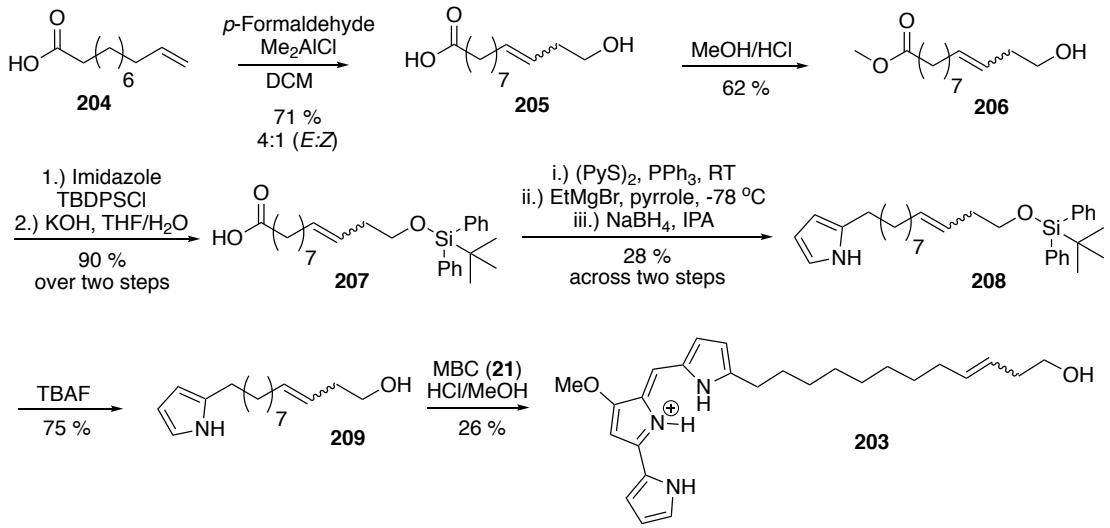



**Figure 3.11:** EIC at *m/z* = 422.28 ±0.01 corresponding to the $[M+H]^+$ ion of ring opened intermediate **203** from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** No feeding control; **B.** Feeding cyclopropane-containing 2-UP analogue **197** and MBC **(21)**.

### 3.3.2 Synthesis of ring-opened authentic standard 203 and subsequent feeding with MBC (21) to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*

An eight-step synthesis of alkene 203 was undertaken starting from undecenoic acid (**204)** (scheme 3.15). Compound **205** was furnished using formaldeheyde and dimethylaluminum chloride.[133] The corresponding methyl ester **206** was formed before silyl protection of the primary alcohol and hydrolysis to give acid **207**. Compound **207** was coupled to freshly distilled pyrrole, reduced, and the primary alcohol deprotected using fluoride. The resulting pyrrole **209** was condensed with synthetic MBC (**21**) under acidic conditions to provide authentic standard **203**.




**Scheme 3.15:** The synthesis of ring-opened authentic standard **203**.

Comparison of the extracted ion chromatograms at *m/z* = 422.2802 for the extract arising from feeding cyclopropane-containing 2-UP analogue **197** alongside MBC **(21)** with the synthetic authentic standard **203** showed that the two compounds have different retention times (figure 3.12) and are therefore not the same. A possible alternative structure is a hydroxylated cyclobutane. Scale-up and purification would be required to further confirm the structure by NMR.

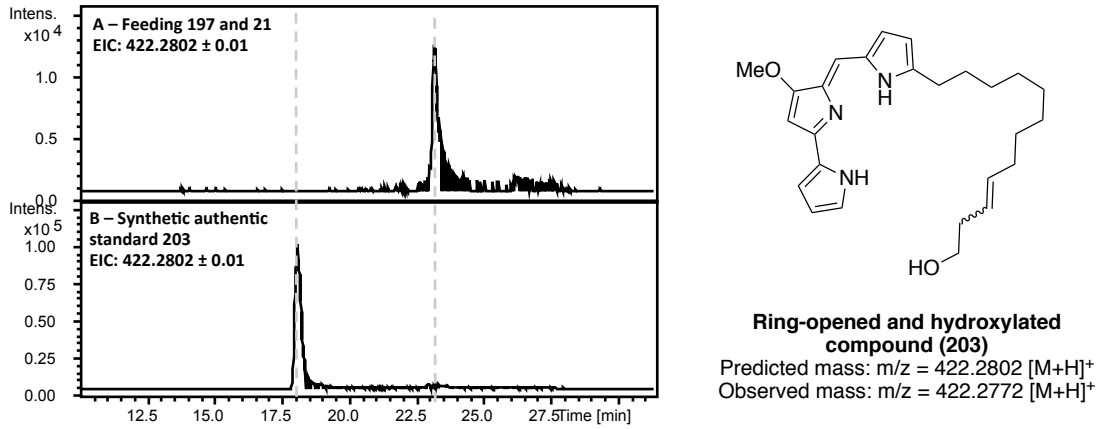
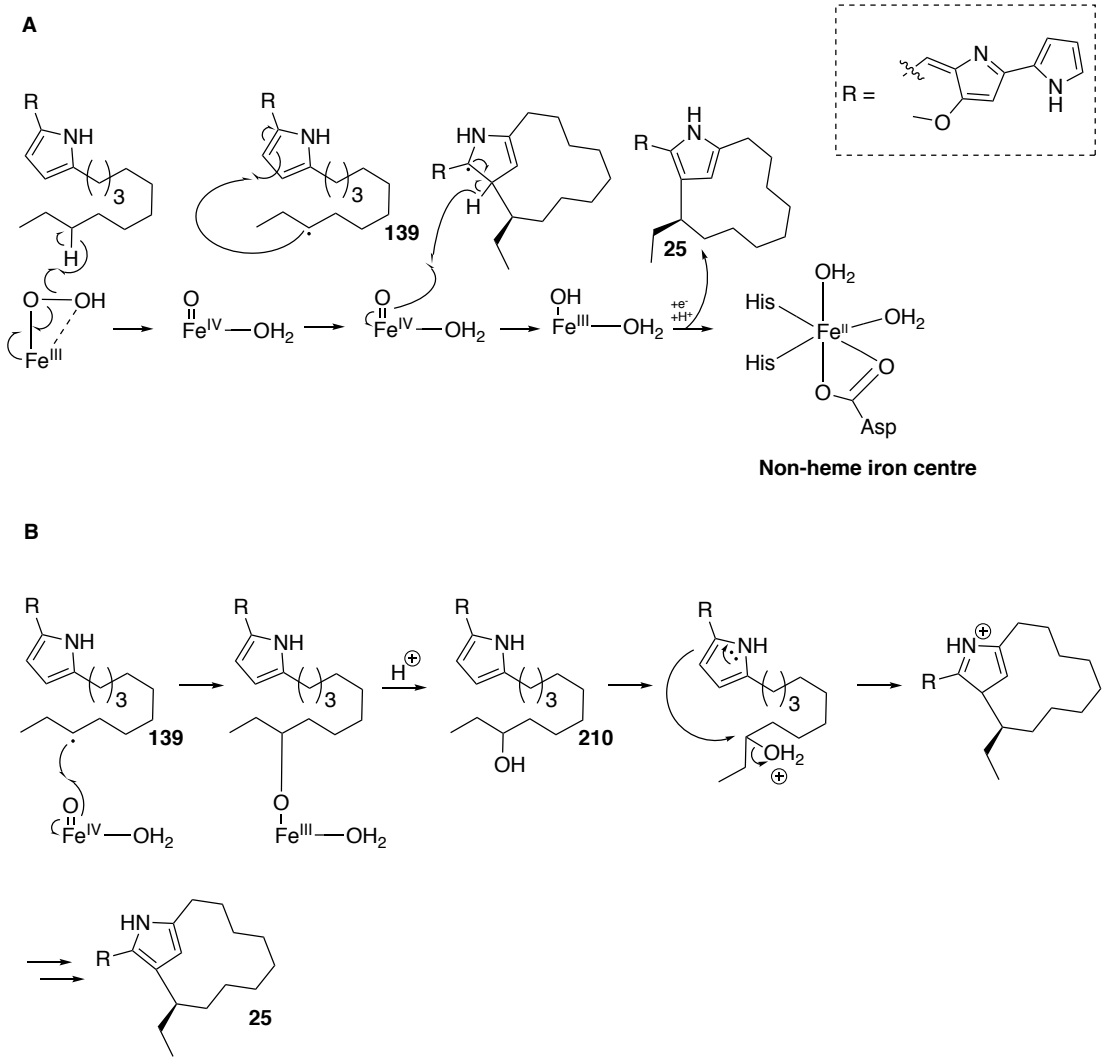



**Figure 3.12:** EIC at *m/z* = 422.28 ±0.01 corresponding to the $[M+H]^+$ ion of ring opened intermediate **203** from the LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** Feeding cyclopropane-containing 2-UP analogue **197** and MBC **(21)**; **B.** Synthetic authentic standard of intermediate **203**.

### 3.3.3 Investigation of a 'rebound' cyclization mechanism through the formation of a hydroxylated undecylprodigiosin intermediate

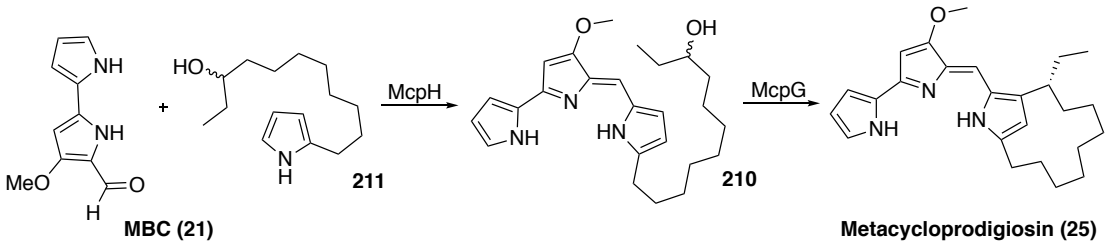
Another question yet to be investigated is the presence of C9-hydroxylated undecylprodigiosin (**210**) formed as an intermediate prior to cyclization (scheme 3.16). The proposed mechanism of oxidative carbocyclization by McpG involves C-H abstraction at C9 of the undecylprodigiosin (**23**) alkyl chain to form a carbon-centred radical (**139**) (or cation). The radical is subsequently cyclized at position-4 of the proximal pyrrole to furnish the more stable tertiary radical (scheme 3.16). It is also plausible that the carbinyl radical rebounds back to the iron-oxo species to give hydroxylated product **210** and that cyclization occurs by nucleophillic attack of the protonated C9 hydroxyl group by the proximal pyrrole. This would furnish compound **210**, which undergoes cyclization *via* protonation promoted nucleophillic substitution. The presence of the proposed hydroxylated intermediate **210** can be investigated by synthesizing 2-UP analogue **140** with a hydroxyl group located at C9 of the alkyl chain (scheme 3.16). When fed to the deletion mutant, this analogue will be condensed with

MBC (**21**) by McpH and furnish the proposed intermediate **210**. If intermediate **210** is converted to metacycloprodigiosin (**25**) then this will confirm rebound onto the iron-oxo species and that mechanism **B** (scheme 3.16) is more likely. If metacycloprodigiosin (**25**) is not observed then the hydroxylated undecylprodigiosin analogue **210** is unlikely to be an intermediate in the pathway and mechanism.




**Scheme 3.16:** Two potential mechanisms for the McpG-catalyzed cyclization reaction; **A.** Direct cyclization of the carbinyl C9 radical **139**; **B.** Rebound of the carbinyl radical **139** to the non-heme iron centre which is protonated to promote the formation of a C9 hydroxylated undecylprodigiosin analogue **210**.

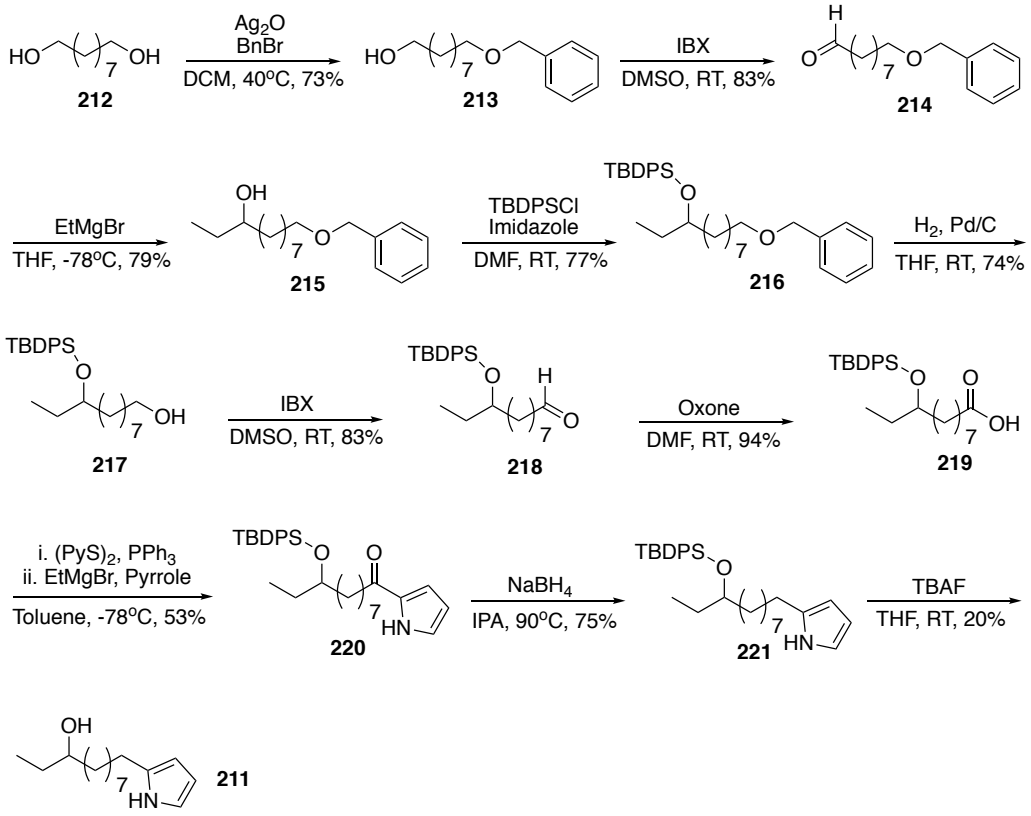
If C9-hydroxylated undecylprodigiosin **210** is an intermediate in the McpG catalytic cycle then feeding C9-hydroxyl 2-UP analogue **211** to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* will produce metacycloprodigiosin (**25**). If **210** is not a substrate then a mechanism of direct cyclization (**A**, scheme 3.16) of the carbinyl radical **139** with the proximal pyrrole is more likely.




**Scheme 3.17:** A schematic representing the feeding experiment utilizing synthetic 2-UP analogue **211** to probe whether **210** is an intermediate during cyclization to metacycloprodigisoin **(25)**.

### 3.3.4 Synthesis of C9-hydroxy-2-UP analogue 211

The first step to compound **211** was a mono-protection of 1,9-nonandiol **212** using a method optimized by Sauvé and co-workers (scheme 3.18).[134] Silver (I) oxide chelates the alcohol moieties of symmetrical diols and consequently acts as a Lewis acid, leading to selective deprotonation of the proton not involved in hydrogen bonding. Therefore, only mono-benzylation is observed.




**Scheme 3.18:** The synthetic route and reaction conditions to furnish C9-hydroxy-2UP analogue **211**.

Treatment of 1,9-nonanediol **212** with silver (I) oxide and benzyl bromide led to the formation of the mono-benzylated alcohol **213** with a yield of 73 %. Oxidation of alcohol **213**

to the corresponding aldehyde **214** was performed using neat IBX. A Grignard reaction using commercially available ethyl magnesium bromide and aldehyde **214** generates **215** with a C9 hydroxyl moeity. The secondary alcohol **215** was protected using *tert*-butyldiphenylsilyl (TBDPS) chloride to furnish **216**. The TBDPS group is more stable towards acidic conditions than the corresponding trimethylsilyl protecting group, and so was used to ensure stability throughout the oxidation reactions, which are performed under acidic conditions. Selective deprotection of the benzyl protecting group by hydrogenolysis afforded the primary alcohol **217** which was oxidized to the corresponding aldehyde **218**. Treatment with oxone furnished the carboxylic acid **219**, which was coupled to freshly distilled pyrrole to generate the 2-ketopyrrole **220**. Reduction of the carbonyl group using sodium borohydride in refluxing 2-propanol afforded the protected alcohol **221**. Treatment of compound **221** with TBAF generated the target compound **211**. The final hydroxylated 2-UP product **211** was synthesized in a racemic mixture.

### 3.3.5 Feeding C9-hydroxy-2UP analogue 211 to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*

The C9-OH-2-UP analogue **211** was fed to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* with MBC (**21**) and incubated for two days. The results showed that McpH was successful in catalyzing the condensation of **211** and MBC **(21)** to form the undecylprodigiosin analogue (**210**); confirmed by HR-MS and chemical formulae (figure 3.13). No metacycloprodigiosin (**25**) was identified in the extract.

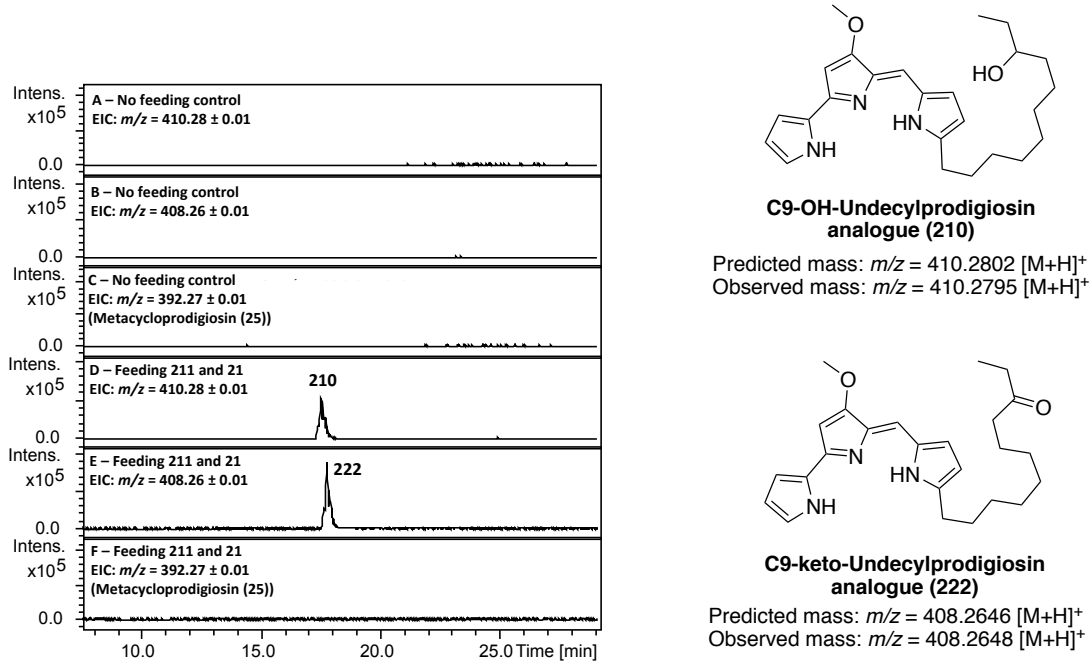
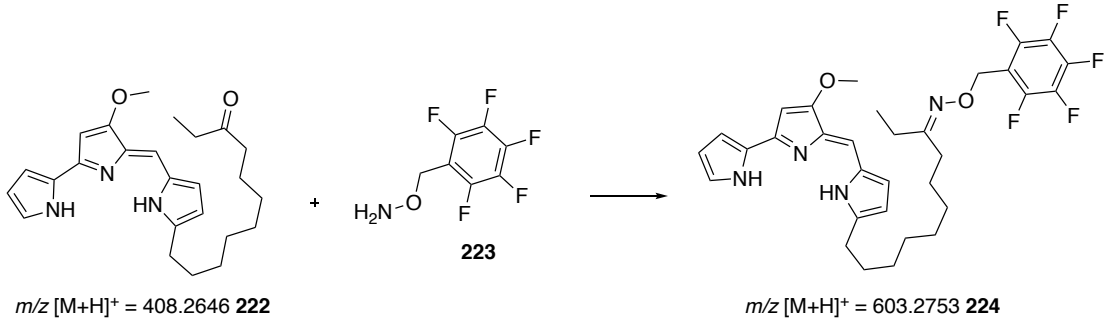



**Figure 3.13:** EICs from LC-MS analysis of the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* after feeding 2-UP analogue **211** and MBC **(21)** to; **A.** EIC at *m/z* = 410.28 ±0.01 corresponding to the $[M+H]^+$ ion of **210**: No feeding control; **B.** EIC at *m/z* = 408.26 ±0.01 corresponding to the $[M+H]^+$ ion of **222**: No feeding control; **C.** EIC at *m/z* = 392.27 ±0.01 corresponding to the $[M+H]^+$ ion of **25**: No feeding control; **D.** EIC at *m/z* = 410.28 ±0.01 corresponding to the $[M+H]^+$ ion of **210**: Feeding synthetic 2-UP analogue **211** and MBC **(21)**; **E.** EIC at *m/z* = 408.26 ±0.01 corresponding to the $[M+H]^+$ ion of **22**; **F.** EIC at *m/z* = 392.27 ±0.01 corresponding to the $[M+H]^+$ ion of **25**: Feeding synthetic 2-UP analogue **211** and MBC **(21)**.

Closer inspection showed that a second compound was also formed that has lost two mass units relative to **210**. It was proposed that intermediate **210** was oxidized to form the corresponding C9-ketone **222** (scheme 3.19). Isolation of this compound for subsequent confirmation of the structure by NMR was not attempted because it has been reported that *O*-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine.HCl (*O*-PFBHA.HCl) **223** can be used to confirm the presence of carbonyl moieties in a number of compounds (scheme 3.19).[135]




**Scheme 3.19:** The reaction of the proposed ketone **222** with hydroxylamine **223** to furnish the corresponding imine **224**, in order to confirm the presence of the ketone moiety.

An excess of *O*-PFBHA.HCl **223** was added to the mycelial extract resulting from feeding 2-UP analogue **211** with MBC **(21)** to *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG* and the reaction was incubated overnight at room temperature. The metabolites were extracted with ethyl acetate, concentrated *in vacuo* and analyzed by HR LC-MS. The LC-MS analysis showed that keto-undecylprodigiosin compound **222** had been entirely consumed and a new peak corresponding to imine **224** was present (figure 3.14).

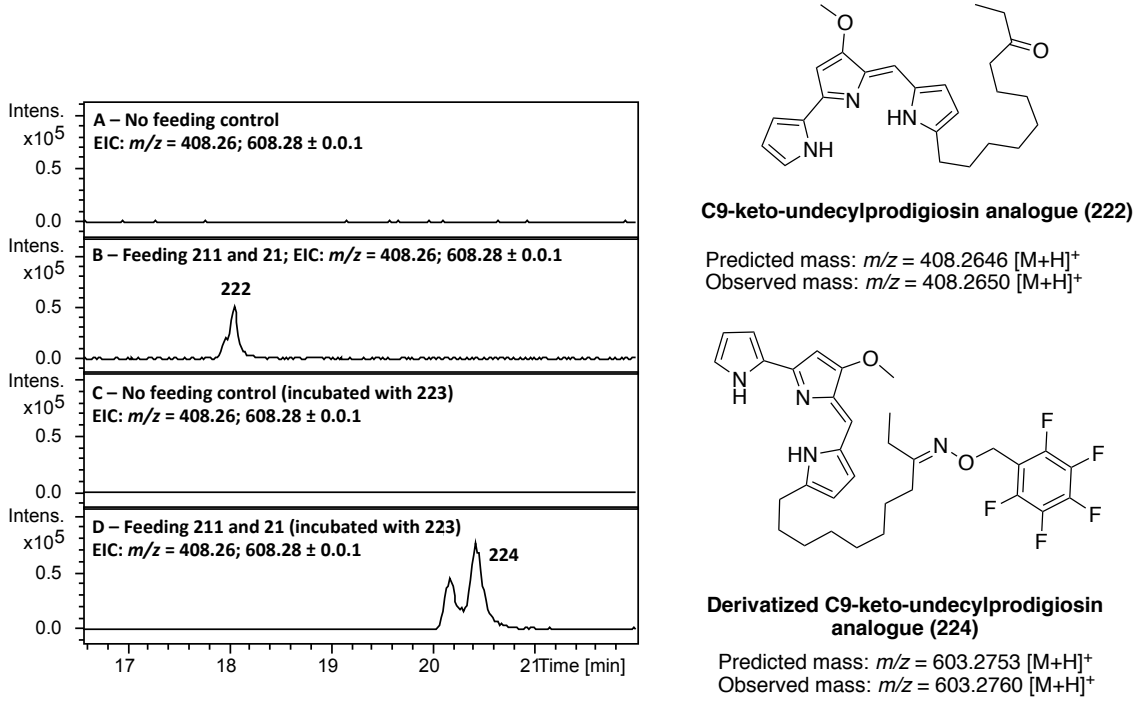



**Figure 3.14:** EIC at *m/z* = 408.26 ±0.01 and 603.29 ±0.01 corresponding to the $[M+H]^+$ ion of **222** and **224** respectively from the LC-MS analysis for the mycelial extract of *S. albus*/pCAP1000-mcpBGC-Δ*LK*-pOSV556t-mcp*HG*; **A.** No feeding control; **B.** Feeding C9-OH-2UP **211** + MBC (**21**); **C.** No feeding control + incubation of **223**; **D.** Feeding C9-OH-2-UP **211** and MBC (**21**) + incubation of **223**.

These data suggests that hydroxylated undecylprodigiosin analogue **210** is not an intermediate in the McpG catalytic cycle and that rebound back with the iron-oxo species does not occur. Formation of the corresponding ketone **222** shows that upon entering the active site, compound **210** undergoes an McpG-catalyzed oxidation rather than forming metacycloprodigiosin **(25)**. The oxidation may occur by formation of the corresponding radical intermediate that rebounds back to the $Fe(IV)O(OH_2)$ complex to furnish a hydrate that collapses to form the resulting ketone **222** (or *via* a cationic intermediate formed *via* hydride abstraction).

## 3.4 Conclusions and perspectives – Mechanism and substrate tolerance of McpG

Two mechanistic probes were designed to investigate the mechanism of McpG. Hydroxylated analogue **211** was not cyclized by McpG but was instead oxidized to a ketone **222**. This observation provides strong evidence to suggest that the proposed radical intermediate **139** does not form hydroxylated undecylprodigiosin analogue **210** by rebound onto the iron-oxo species, but instead a direct cyclization occurs. Cyclopropane analogue **197** was also not cyclized by McpG, and HR LC-MS analysis revealed the formation of a compound where the cyclopropane moeity had been cleaved (**199**). As yet, a mechanism for this cleavage is unknown. Confirmation of the structure could not be achieved by NMR characterization and so further work is required. Incubating the cyclopropane containing 2-UP analogue **197** with McpG reconstituted *in vitro* would provide further insight into the reactive intermediate formed through C-H abstraction. The *in vitro* studies would allow for investigation into reaction kinetics and for crystallographic studies, to distinguish between the possible radical or cationic intermediate.

Through the synthesis of many 2-UP mechanistic probes the mechanism of Rieske oxygenase catalysis has been extensively investigated (for RedG and McpG), providing evidence for the presence of radical intermediates rather than cationic intermediates. As the data is yet to unambiguously rule out the presence of cationic intermediates, the optimization of an *in vitro* system to investigate the over-expressed and purified RedG and McpG enzymes will be undertaken to provide further mechanistic data through kinetic and crystallographic studies.

### 3.4.1 Substrate tolerance of McpG

As RedG has been shown to have a fairly relaxed substrate tolerance, for this class of enzyme to be considered useful as a biocatalyst it is important to know whether other family members are equally as tolerant. A number of 2-UP analogues were synthesized from their corresponding carboxylic acids to investigate the substrate tolerance of McpG. Synthesis of a 2-UP analogue with a terminal alkene moeity (**157**) and feeding to the deletion mutant alongside synthetic MBC **(21)** failed to yield the corresponding metacycloprodigiosin analogue. Stabilization of the intermediate by the adjacent π-system through resonance effects drastically reduces the rate of cyclization to furnish the corresponding cyclic analogue. It is likely that the same effect is observed when feeding a 2-UP analogue with a terminal acetyline moeity (**155**), however the low pKa of the acetyline C-H may also have degradative

effects on the Rieske cluster and/ or non-heme iron centre. This functionality may also invoke electron repulsion between the electron rich π-system of the acetyline and the electron rich scaffold of the dipyrrole-pyrromethene scaffold.

A 2-UP analogue where the site of C-H abstraction is $sp^2$-hybridized (**156**) was synthesized and the compound was fed with MBC **(21)** to the *mcpLK* deletion mutant. It was envisaged that C-H abstraction at the $sp^2$ carbon would be unfavourable, and McpG would catalyze an alternative epoxidation reaction. A peak with the correct molecular formulae corresponding to epoxide **189** was observed when the resulting extract was analyzed by HR LC-MS. An authentic standard is required to confirm the structure of this product.

Feeding 2-UP analogues with a nine-carbon alkyl chain (**151**), ten-carbon alkyl chain (**152**) and twelve-carbon chain (**153**) to the deletion mutant showed that McpG accepts extensions of the alkyl chain relative to the eleven-carbon natural precursor 2-UP (**28**). A similar observation was recorded for RedG.[89]

Feeding a methyl branched 2-UP analogue (**154**) furnished the corresponding undecylprodigiosin analogue (**32**), which was not a substrate for McpG. No desaturation product was observed in the resulting culture, which differed from the results observed when a similar analogue was used to investigate RedG.[89]

Each corresponding undecylprodigiosin analogue was observed when feeding the previously discussed synthetic 2-UP analogues. These data suggests that McpH has excellent substrate tolerance and can catalyze the condensation of MBC (**21**) with a wide range of synthetic 2-UP analogues.

# Chapter 4: The overproduction and purification of Rieske oxygenases RedG and McpG

## 4.1 Introduction

### 4.1.1 The *in vitro* reconstitution of radical-forming enzymes

Radical-forming enzymes generate substrate-derived radicals through homolysis of a covalent bond or by one-electron transfer.[136] For the case of Rieske enzymes, this external electron is typically accepted by an iron-sulphur cluster and then shuttled to the non-heme iron centre for activation of $O_2$.[53,136] Radical intermediates are typically highly reactive species and the dimerization of enzyme-bound radicals or other intramolecular side-reactions are either sterically or electronically prevented by the architecture of the protein.[136] The protein-induced stabilization of radical species provides an advantage over producing similar radical species by *in vitro* chemical reactions. For this reason, there has been much interest in reconstituting radical-forming enzymes to catalyze complex chemistries with a high degree of regiocontrol and/ or specificity. However, the instability of Rieske clusters and non-heme iron centres towards oxidative degradation (by atmospheric oxygen) is a significant limitation in studying enzymes harboring a [2Fe-2S] cluster. Producing a soluble protein, correctly folded and that contains an intact [2Fe-2S] cluster is often challenging.

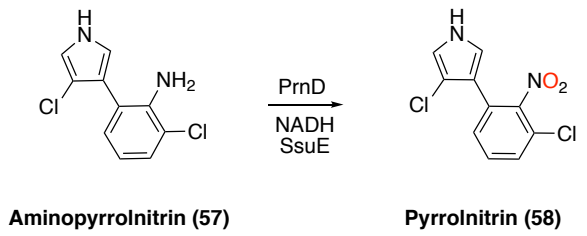
### 4.1.2 U.V-Visible absorption spectra of Rieske enzymes

Rieske [2Fe-2S] proteins are typically brown in colour, which is due to the presence of Fe ions within the core structure. The S–Fe(III) charge transfer results in the observation of a broad absorption range in the visible and near-UV region, and consequently UV-Vis absorption spectroscopy is the primary method for characterizing the architecture of the iron-sulphur cluster.[68,74,80] Rieske proteins typically show absorption maxima between 325-340 nm and 458-465 nm as well as a shoulder between 560 and 580 nm in the oxidized state.[137] In the reduced state the maxima shift to between 380-405 nm, 420-432 nm and 505 nm and 550 nm.[137]

### 4.1.3 The *in vitro* reconstitution of radical-forming Rieske oxygenase enzymes and Rieske [2Fe-2S] clusters

#### 4.1.3.1 PrnD Rieske [2Fe-2S] cluster

A well-characterized Rieske oxygenase involved in natural product biosynthesis that has been heterologously overproduced and purified is PrnD; involved in the final step of pyrrolnitrin **(58)** biosynthesis.[54,53] Pyrrolnitrin (**58**) is an antifungal compound produced by *Psuedomonas fluorscens* BL915 and was first isolated in 1997. The final step in this biosynthetic pathway is the conversion of aminopyrrolnitrin (**57**) to pyrrolnitrin (**58**) *via* the PrnD-catalyzed oxidation of an amino group to a nitro group (scheme 4.1).




**Scheme 4.1:** The PrnD-catalyzed oxidation of the amino group of aminopyrrolnitrin (**57**) to furnish pyrrolnitrin **(58)**.

Zhao and co-workers constructed a $His_6$-MBP-PrnD fusion protein that was overproduced and purified from *E. coli.* The MBP-cleaved protein was incubated with ferrous iron and sulphide under anaerobic conditions and the reconstituted protein was analyzed by UV-Vis spectroscopy. The enzyme displayed absorption maxima at 278 nm, 339 nm, 440 nm and a small peak at 570 nm (figure 4.1).[54] The dithionite-reduced Rieske cluster displayed absorption maxima that had shifted to 405 nm and 521 nm. The oxidation reaction was reconstituted *in vitro* using refactored PrnD, aminopyrrolnitrin (**57**), NADH and the flavin reductase enzyme SsuE, to furnish pyrrolnitrin (**58**).

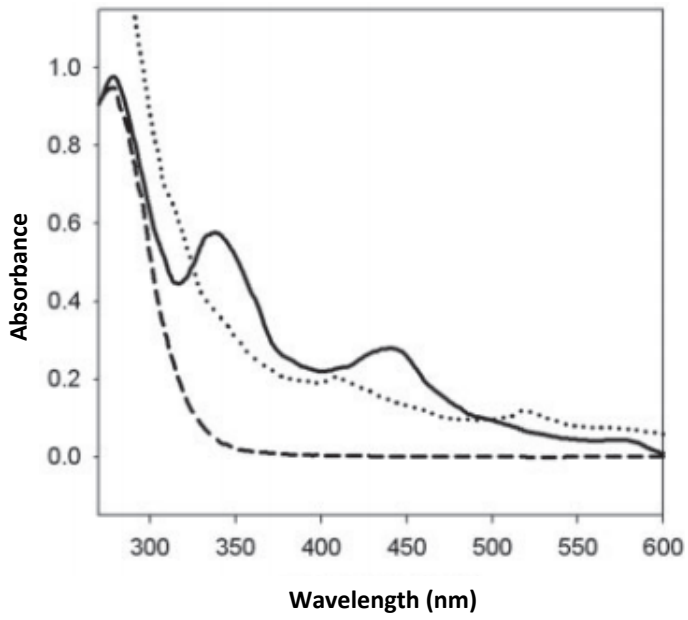
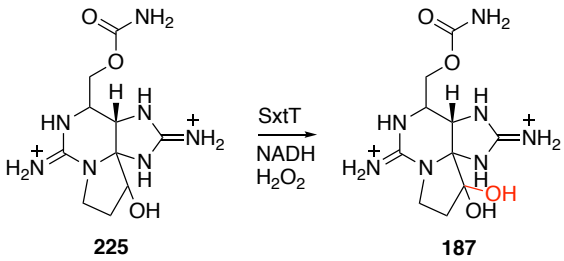



**Figure 4.1**: The UV-Vis absorption spectra of PrnD obtained by Zhao and co-workers.[54] Solid line: Reconstituted PrnD; Dashed line: Reconstituted and oxidized PrnD; Dotted line: Reconstituted and reduced PrnD.

### 4.1.3.2 SxtT Rieske [2Fe-2S] cluster

Narayan and co-workers reported that SxtT is a Rieske oxygenase enzyme that is responsible for the selective hydroxylation of a tricyclic precursor, β-saxitoxinol (**225**), to produce saxitoxin **(187)** (scheme 4.2).[69] Saxitoxin (**187**) is a paralytic shellfish toxin produced by cyanobacteria that has shown high-affinity for voltage-gated sodium channels (VGSC). The compound has therefore been used in the study of VGSC and for medicinal agents used to target VCSG isoforms that are associated with pain and disease.




**Scheme 4.2:** The SxtT-catalyzed hydroxylation of β-saxitoxinol (**225**) to furnish saxitoxin (**187**).

SxtT was overproduced as a maltose binding fusion protein (MBP) in *E. coli*, purified and the Rieske cluster characterized by UV Vis absorption spectroscopy. Key absorption maxima that indicated the presence of an oxidized Rieske cluster were observed at 325 nm, 460 nm and 580 nm (figure 4.2). The SxtT-catalyzed reaction was reconstituted *in vitro* using NADH and

stoichiometric hydrogen peroxide to generate the active oxidant, and synthetic **225**. A range of saxitoxin (**187**) related substrates were also used and subsequently hydroxylated, to highlight the substrate promiscuity of the enzyme.

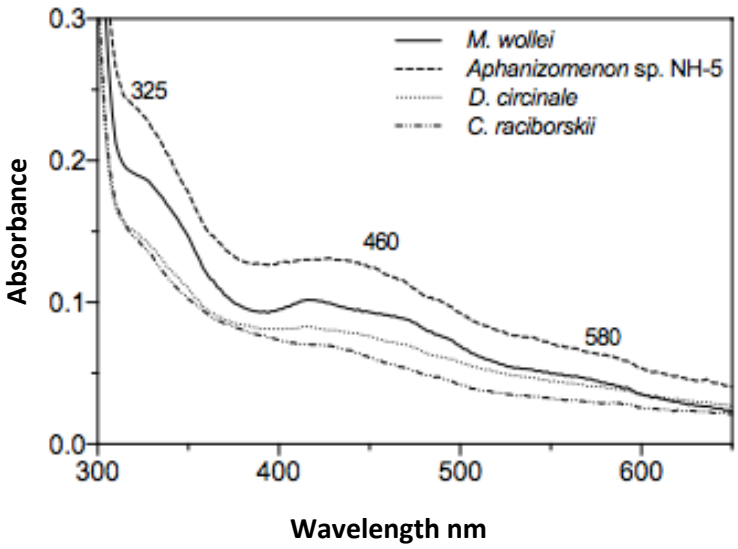
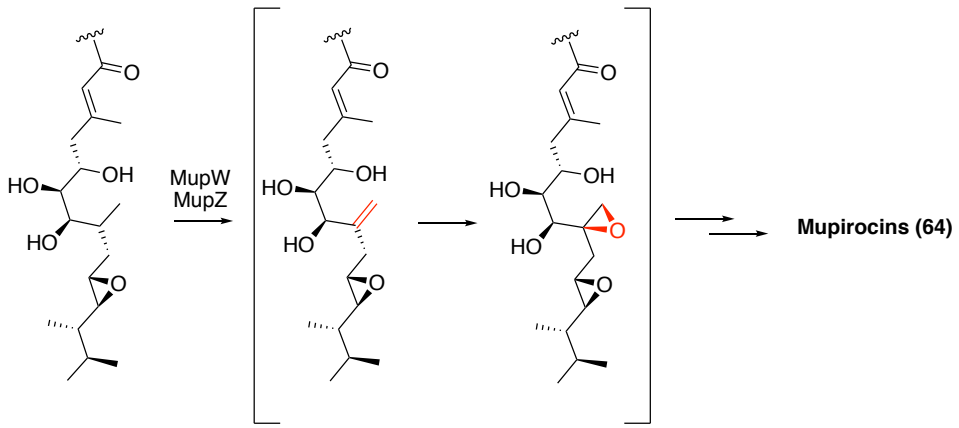



**Figure 4.2:** The UV-Vis absorption spectra for Rieske oxygenase SxtT from different cyanobacteria (highlighted by the different lines), obtained by Narayan and co-workers.[69] Absorption maxima at 325nm, 460 nm and 580 nm correspond to an oxidized Rieske cluster.

#### 4.1.3.3 MupW Rieske [2Fe-2S] cluster

Willis and co-workers have reported that the 6,7-dihydroxy-tetrahydropyran (THP) ring in Mupirocin (**64**) is biosynthesized by a tandem Rieske oxygenase/epoxide hydrolase (MupW/MupZ) cascade (scheme 4.3).[60]




**Scheme 4.3**: The MupW/MupZ-catalyzed cascade in the biosynthesis of the THP ring of the mupirocins (**64**).

MupW is a Rieske non-heme iron dependent oxygenase that catalyzes the oxidation of a C-C single bond, and MupZ catalyzes the subsequent ring formation to furnish the THP moeity of the mupirocins (**64**). The presence of an epoxide intermediate was predicted through the use of model studies. MupW was bioinformatically shown to contain conserved CXH and CXXH motifs relating to the two Cys and His residues that ligate iron atoms in a Rieske [2Fe-2S] cluster. The enzyme also contains a conserved $DX_2HX_4H$ motif that ligate a non-heme iron centre and a bridging glutamate residue. Willis and co-workers heterologously overproduced an N-terminal $His_6$-MupW fusion in *E. coli* and purified the protein aerobically. The red-brown enzyme was analyzed by UV-Vis absorption spectroscopy and displayed peaks at around 330, 460 and 560 nm. The reduced enzyme displayed absorption maxima at around 520 nm (figure 4.3).

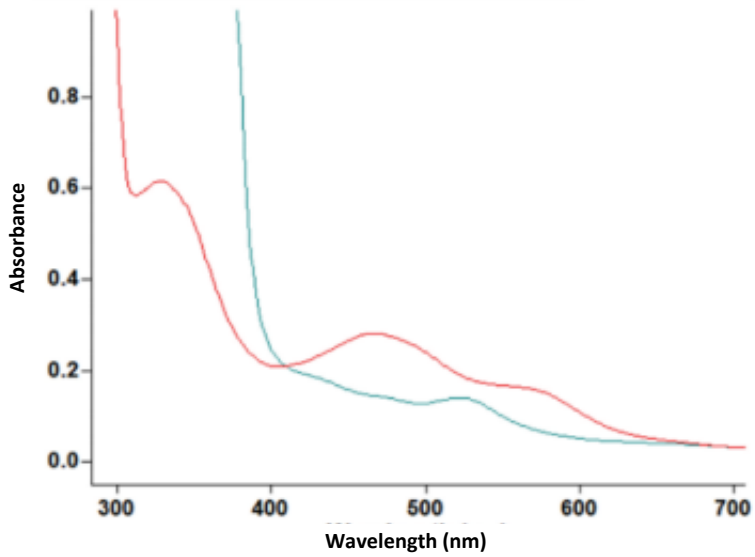



**Figure 4.3:** The UV-Vis absorption spectrum of purified MupW. Red: Oxidized MupW; Blue: Reduced MupW.[60]

#### 4.1.3.4 Rieske oxygenases NDO, RedG and McpG

Sequence alignment of RedG and McpG with NDO confirmed similarities in the protein architecture that are consistent with Rieske proteins. RedG and McpG both possess an N-terminal domain containing the characteristic $CXHX_{17}CX_2H$ [2Fe-2S] cluster-binding motif and a C-terminal domain containing an $EXHX_4H$ variant of the $DXHX_4H$ motif, found in the NDO α-subunit (section 1.4.2). The two His residues bind the active site ferrous ion and a glutamate residue (for RedG and McpG) which is associated with electron transfer from the Rieske cluster to the non-heme iron centre. Until now, neither RedG nor McpG has been studied *in vitro* nor has the bioinformatically predicted Rieske [2Fe-2S] cluster been experimentally confirmed. Dr. Paulina Sydor (former Challis group member) mutated the 2 × Cys and His residues that were predicted to bind the Rieske [2Fe-2S] cluster in RedG. This mutation subsequently abolished RedG activity *in vivo*. Development of a robust system for producing soluble recombinant RedG and McpG will facilitate the *in vitro* reconstitution of the prodiginine oxidative carbocyclization reaction.

## 4.2 Aims & objectives

The aim of this investigation was to develop methods for the production of soluble RedG and McpG containing an intact iron-sulphur cluster, as the first step towards reconstituting oxidative carbocyclization activity using synthetic undecylprodigiosin (**23**). We planned to use UV-Vis spectroscopy to characterize the Rieske cluster. As Rieske proteins are prone to oxidative degradation we envisaged exploiting both aerobic and anaerobic purification, as well as post-purification methods for the *in vitro* reconstitution of the Rieske [2Fe-2S] cluster. Reconstituting the RedG and McpG activity *in vitro* will allow for further mechanistic investigations, such as analysis of the reaction kinetics and the isolation/ detection of unstable intermediates.

## 4.3 Results & Discussion – The overproduction and purification of RedG and McpG containing an intact Rieske [2Fe-2S] cluster

### 4.3.1 Construction of RedG and McpG expression vectors

In addition to a $His_6$-*redG*/pET151 expression vector previously constructed by Paulina Sydor (former Challis group member) several other expression vectors were investigated. Their preparation is described below.

#### 4.3.1.1 Construction of $His_6$-*redG*/pET28a(+) & $His_6$-*mcpG*/pET28a(+) expression vectors

A pET28a(+) expression vector was used to avoid cross-resistance with pD1282 (see section 4.3.4.2). The *redG* and *mcpG* genes were amplified by PCR using the $His_6$-*redG*/pET151 construct and *S. longispororuber* genomic DNA, respectively, as template DNA. The resulting PCR products were ligated directly into pET28a(+) which had been linearized by restriction enzyme digest. Transformants were selected and the vectors contained within them were analyzed by restriction digests (*Xho*I and *Kpn*I, **A**, figure 4.4) or (*Eco*RV, *Not*I and *Xho*I, **B**, figure 4.4) and by sequence analysis.

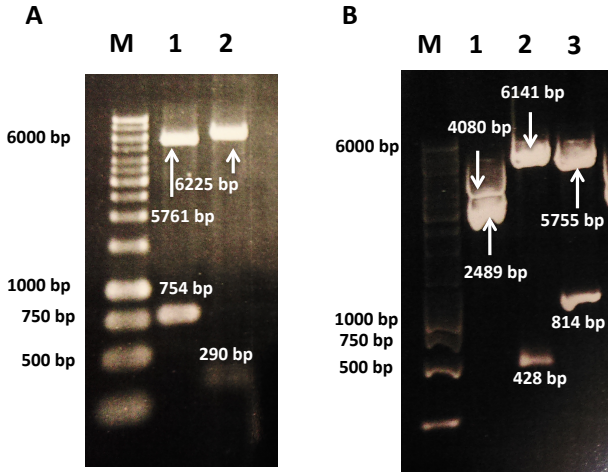



**Figure 4.4: A.** Agarose gel electrophoresis analysis of restriction enzyme digests of *redG*/pET28a(+); Lane 1 = *Xho*I digest; expected bands 6225, 754 bp; Lane 2 = *Kpn*I digested expected bands 6225, 290 bp; **B.** Agarose gel electrophoresis analysis of restriction enzyme digests of *mcpG*/pET28a(+); Lane 1 = *Eco*RV digest; expected bands 4080 2489 bp; Lane 2 = *Not*I digest; expected bands 6141, 428 bp; Lane 3 = *Xho*I digest; expected bands 5755, 814 bp.

#### 4.3.1.2 Construction of $His_6$-*redG*/pETSUMO and $His_6$-*mcpG*/pETSUMO expression vectors

A pETSUMO expression vector was used to try and improve solubility (see section 4.3.4.3). TA cloning was used to clone both *redG* and *mcpG* into the pETSUMO expression vector. The TA cloning method involves the use of Taq polymerase, that lacks the 5'-3' proof reading activity and so is capable of adding ATP to the 3' ends of double stranded PCR products that contains 3' T overhangs. Therefore *redG* and *mcpG* were amplified by PCR to contain 3' T overhangs for insertion directly into the linearized pETSUMO vector using *Taq* polymerase. The resulting transformants were screened by restriction enzyme digest and the correctly ligated plasmids were confirmed by both restriction enzyme digest and DNA sequence analysis (figure 4.5).

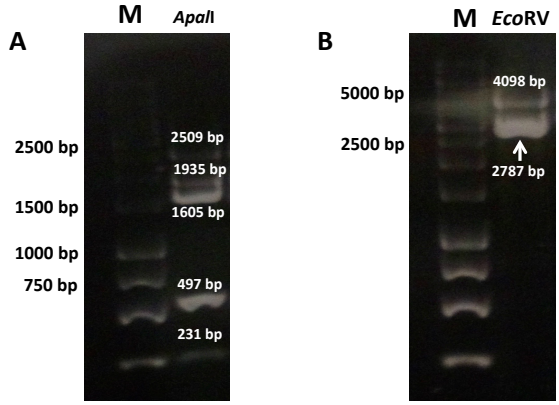



**Figure 4.5: A.** Agarose gel electrophoresis analysis of the restriction digest of $His_6$-SUMO-*redG*/pETSUMO vector; *Apal*I digest; expected bands 2509, 1935, 1605, 497, 231 bp; **B.** Agarose gel electrophoresis analysis of the restriction digest of $His_6$-SUMO-*mcpG*/pETSUMO vector; *Eco*RV digest; expected bands 4098, 2787 bp.

### 4.3.2 Initial screens of Rieske oxygenase RedG & McpG protein production

#### 4.3.2.1 Production of $His_6$-RedG using $His_6$-*redG*/pET151

The *redG*/pET151 expression vector was used to transform *E. coli* BL21 (DE3). A single transformant was used to innoculate an overnight LB pre-culture and protein overproduction was undertaken using the conditions outlined in section 6.5.9.2. The recombinant protein was purified under aerobic conditions using nickel affinity chromatography. Unbound proteins were eluted using increasing concentrations of imidazole elution buffers. The eluents were analyzed by SDS-PAGE (figure 4.6). This showed soluble recombinant $His_6$-RedG was produced under aerobic conditions. However, solutions of the purified protein did not display the typical brown colouration that is indicative of an iron-sulphur cluster.

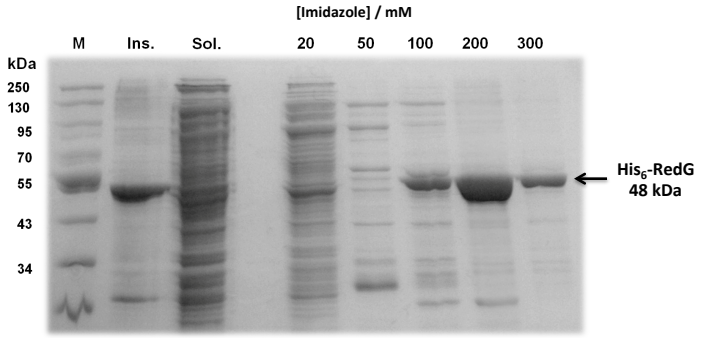



**Figure 4.6:** SDS-PAGE analysis of the purification of recombinant $His_6$-RedG using a pET151 expression vector, *E. coli* BL21 (DE3) cells and LB medium, under aerobic conditions.

### 4.3.3 Expression of recombinant Rieske oxygenases RedG and McpG using $His_6$-*redG*/pET28a(+) and $His_6$-*mcpG*/pET28a(+)

The *redG*/pET28a(+) and *mcpG*/pET28a(+) vectors were used to transform *E. coli* BL21 (DE3) cells respectively and a single transformant for each was used to inoculate an LB overnight pre-culture. Cells were grown by incubating at 37 $^{o}$C with shaking until reaching an $OD_{600}$ of 0.60-1.00. Isopropyl-β-thiogalactopyranoside (IPTG) was added to the culture before incubation at 15 $^{\circ}$C with shaking for 20 hours. The cells were harvested and protein purification was undertaken as described in section 6.5.9.2. The imidazole eluents were analyzed by SDS-PAGE, which showed that RedG **A** (figure 4.7) and McpG **B** (figure 4.7) were both produced.

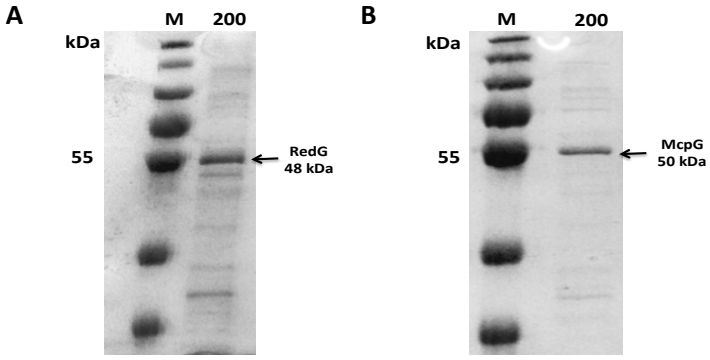



**Figure 4.7: A.** SDS-PAGE analysis of aerobically purified recombinant $His_6$-RedG resulting from the pET28a(+) expression vector, *E. coli* BL21 (DE3) cells and LB medium; **B.** SDS-PAGE analysis of aerobically purified recombinant $His_6$-McpG using a pET28a(+) expression vector, *E. coli* BL21 (DE3) cells and LB medium.

Similar to the results of the experiments utilizing the pET151 expression vector, neither RedG nor McpG displayed the brown colour that is indicative of a Rieske iron-sulphur cluster.

### 4.3.4 Attempts to optimize production conditions

To optimize protein production and promote incorporation of the Rieske [2Fe-2S] cluster various avenues were investigated, as outlined below (using $His_6$-*redG*/pET151).

#### 4.3.4.1 Anaerobic handling after cell lysis

The lysed cells were inserted into an oxygen-free glove box that had been flushed and equilibrated with nitrogen gas for 24 hours. The protein was purified inside the glove box by nickel affinity chromatography using the conditions outlined in section 6.5.9.4, where all buffers and equipment used had been purged with nitrogen gas and equilibrated inside the glove box for 24 hours.

#### 4.3.4.2 Supplementing the production medium with [2Fe-2S] cluster precursors during induction of gene expression

At the time of IPTG induction, ferrous sulphate, ferric citrate, ferric ammonium citrate and cysteine were added to the production medium, prior to overnight incubation at 15 °C with shaking.[138]

#### 4.3.4.3 Supplementing the production medium with [2Fe-2S] cluster precursors before induction of gene expression

Sodium sulphide (200 mg/L) and ferric ammonium citrate (200 mg/L) were added to the growth medium at the same time as the overnight $His_6$-*redG*/*E. coli* BL21 (DE3) pre-culture. The resulting culture was incubated at 15 °C overnight, with shaking.

#### 4.3.4.4 M9 production medium

A $His_6$-*redG*/pET151/*E. coli* BL21 (DE3) overnight pre-culture in LB medium was used to innoculate one-litre of autoclaved M9 medium (components outline in section 6.4.3). The medium was supplemented with ferrous ions. The cells exhibit a longer lag phase of between 4-5 hours than when grown in LB medium, and so longer incubation times were necessary to reach an $O.D_{600}$ = 0.60. The IPTG concentration was also lowered ten-fold.

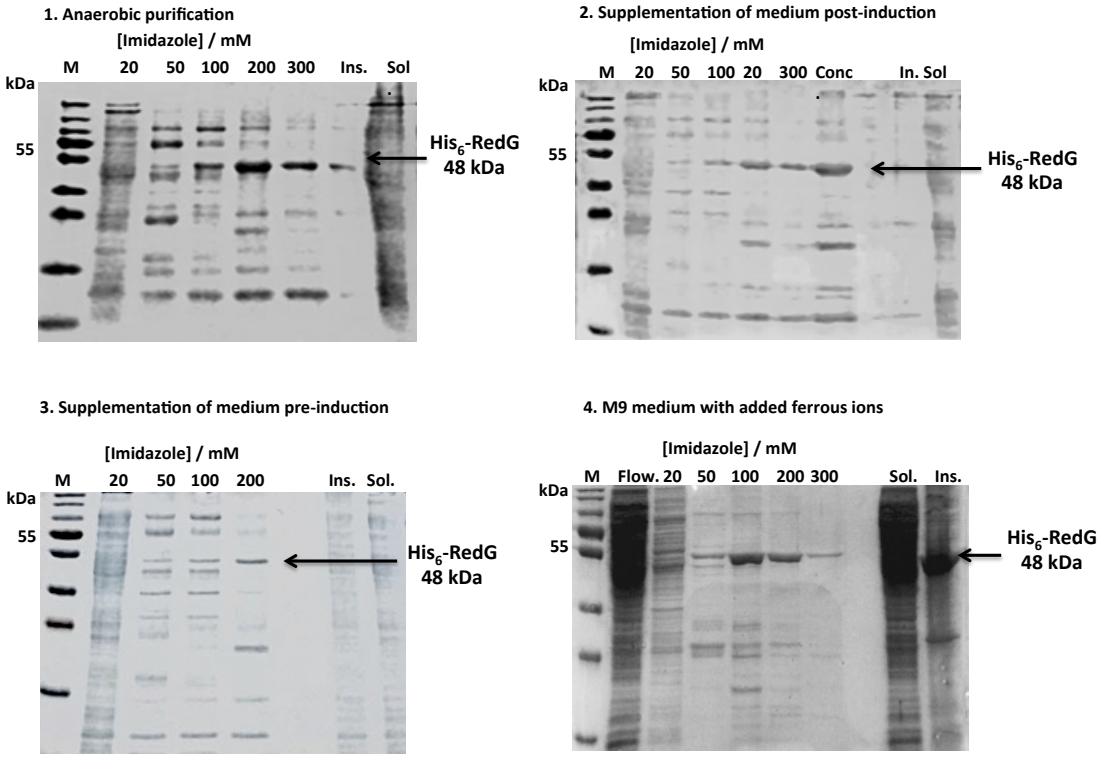



**Figure 4.8:** SDS-PAGE analysis of the purification of recombinant $His_6$-RedG using a pET151 expression vector, *E. coli* BL21 (DE3) cells and various modifications to growth and/or purification conditions.

By changing the medium from LB to M9 and lowering the concentration of IPTG ten-fold the overall expression level significantly increased. The resulting purified protein displayed a red/ brown colouration, indicating the presence of an iron-sulphur cluster. This is likely because the reduced rate of protein production facilitates Fe-S cluster incorporation. A 200 µL sample of the purified recombinant protein was analyzed by UV-Vis spectroscopy in a sealed argon-flushed quartz cuvette (figure 4.9). The spectrum displayed absorption maxima at 327 nm and 421 nm, broadly consistent with what has been observed previously for reduced Rieske iron-sulphur clusters in the Rieske protein from *Thermus thermophilus*.[139] This indicates that RedG contains the Rieske [2Fe-2S] cluster predicted based on sequence comparison with NDO. As the overall level of protein production appeared to be significantly higher than the iron-sulphur cluster incorporation level it was not possible to calculate the level of cluster incorporation by comparison of the absorbance levels at 421 nm and 280 nm (typical maxima for Trp, Tyr and Cys residues in the protein structure). To increase the incorporation levels of the Rieske [2Fe-2S] cluster, co-transformation with arabinose-inducible pDB1282 plasmid, which encodes enzymes that catalyze iron-sulphur cluster biosynthesis, was investigated.

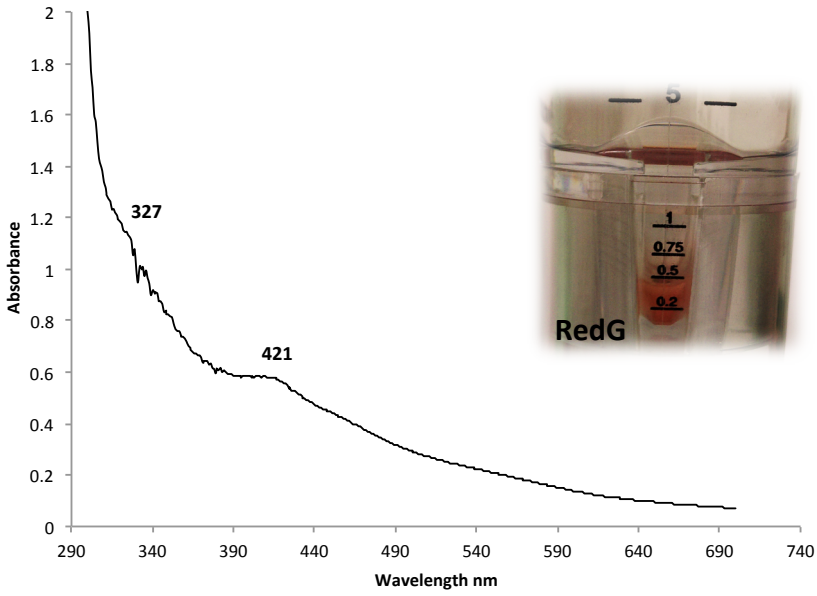



**Figure 4.9:** UV-Vis absorption spectrum of aerobically purified recombinant $His_6$-RedG from *E. coli* BL21 (DE3) cells containing $His_6$-*redG*pET151 cultured in ferrous iron supplemented M9 medium and an image of the resulting protein solution.

### 4.3.4.5 Co-transformation with an iron-sulphur cluster biosynthesis plasmid - pDB1282

A method reported by Sun and co-workers for the expression and purification of GenD1 (containing a [4Fe-4S] cluster) uses a second plasmid (pDB1282) containing genes responsible for the biosynthesis of iron-sulphur clusters.[131] This plasmid confers ampicillin resistance and contains the *isc*S-*isc*U-*isc*A-*hsc*B-*hsc*A-*fdx* portion of the *isc* operon from *Azotobacter vinelandii* and production of the corresponding proteins is induced by the addition of arabinose to the culture medium (figure 4.10).[140]

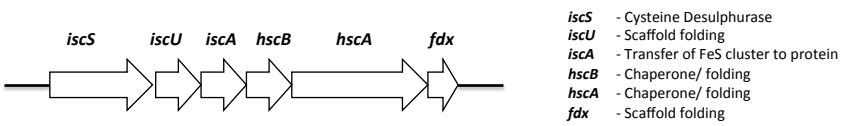



**Figure 4.10:** The genes responsible for iron-sulphur cluster biosynthesis under the control of the arabinose-inducible $P_{BAD}$ promoter in the pDB1282 plasmid.

As the yield of purified recombinant McpG was greater than the corresponding yield of RedG when using the pET28a(+) expression vector, *E. coli* BL21 (DE3) cells in LB medium containing the $His_6$-*mcpG*/pET28a(+) plasmid was used for investigation of co-transformation with pDB1282. An *E. coli* BL21 (DE3) containing $His_6$-*mcpG*/pET28a(+)/pDB1282 pre-

culture was used to inoculate one-litre of M9 medium supplemented with iron (II) sulphate heptahydrate. Following the procedure outlined by Sun and co-workers, the culture was incubated at 37 °C with shaking (220 rpm) until an $O.D_{600}$ of 1.0 was reached. Arabinose was added to induce the expression of the iron-sulphur cluster biosynthesis genes in pDB1282. The resulting culture was incubated with shaking at 90 rpm, resulting in less aeration of the culture. Reducing the oxygen concentration is proposed to reduce oxidative degradation of iron-sulphur clusters. The cells were harvested and resuspended in wash buffer. This displayed a visibly dark brown colour indicating a high level of iron-sulphur cluster incorporation (figure 4.11). After lysis by sonication, the cell lysate was subjected to nickel affinity chromatography and recombinant McpG was purified following the conditions outlined in section 6.5.9.2. The resulting imidazole eluents were analyzed by SDS-PAGE (figure 4.11).

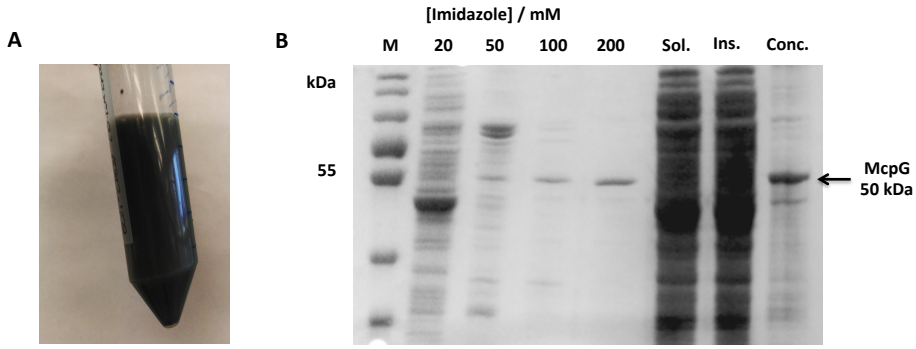



**Figure 4.11: A.** Harvested cell culture resuspended in cell lysis buffer; **B.** SDS-PAGE analysis of the purification of recombinant $His_6$-McpG using a pET28a(+) expression vector + pDB1282, *E. coli* BL21 (DE3) cells and M9 medium.

Unfortunately, although the colour of the cell suspension indicated high-level production of an FeS cluster, the purified protein was colourless. To try to increase levels of FeS cluster incorporation, the concentrations of both IPTG and arabinose were changed to modulate the levels of gene expression. Altering the ratios of IPTG and arabinose alters the rates of McpG and FeS cluster production respectively. Ten small scale (50 mL) inductions were undertaken using the $His_6$-*mcpG*/pET28a(+) expression vector and *E. coli* BL21 (DE3) cells in M9 medium. Varying amounts of IPTG and arabinose were added to cultures which were grown at 20 °C, with shaking, for 20 hours. The resulting cells were harvested and lysed and the lysate was loaded directly onto an SDS-PAGE gel for analysis by electrophoresis, without purification by nickel affinity chromatography (figure 4.12). All cultures produced a significantly larger band in the insoluble fraction compared to the soluble fraction, and none of the cultures were brown.

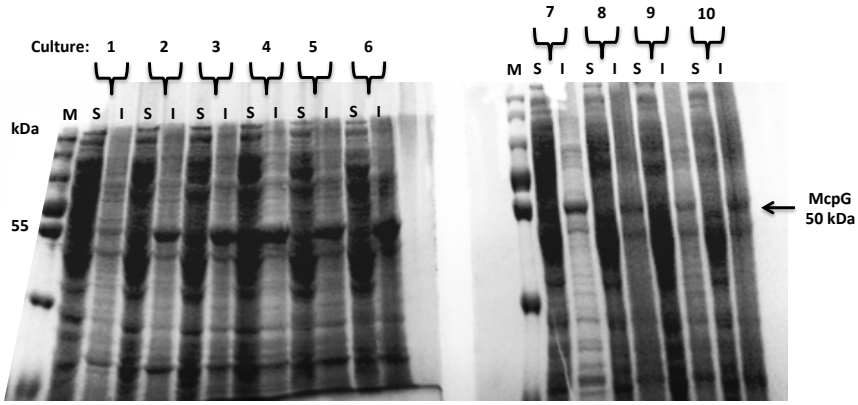



**Figure 4.12:** SDS-PAGE analysis of recombinant $His_6$-McpG produced using a pET28a(+) expression vector + pDB1282, *E. coli* BL21 (DE3) cells and M9 medium (**S** = soluble, **I** = insoluble); **Culture 1** = 0 IPTG, 3.0 mM arabinose; **2** = 1.0 mM IPTG, 3.0 mM arabinose; **3** = 1.0 mM IPTG, 1.5 mM arabinose; **4** = 1.0 mM IPTG, arabinose 0.3 mM, **5** = 0.5 mM IPTG, 3.0 mM arabinose; **6** = 0.5 mM IPTG, 1.5 mM arabinose; **7** = 0.5 mM IPTG, 0.3 mM arabinose; **8** = 0.1 mM IPTG, 3.0 mM arabinose, **9** = 0.1 mM IPTG, 1.5 mM arabinose; **10** = 0.1 mM IPTG, 0.3 mM arabinose.

These data suggest that co-expression of the FeS cluster biosynthesis plasmid pDB1282 with pET28a(+)*mcpG* failed to increase the levels of iron-sulphur cluster incorporation into recombinant McpG, instead generating high levels of insoluble protein. To try and overcome this problem the use of an N-terminal small ubiquitin-related modifier (SUMO) tag was investigated.

#### 4.3.4.6 $His_6$-SUMO-*redG*/pETSUMO + pDB1282 & $His_6$-SUMO-*mcpG*/pETSUMO + pDB1282

It has been reported that covalent conjugation of proteins with ubiquitin regulates both their function and solubility.[141,142] The small ubiquitin-related modifier (SUMO) protein has homology with ubiquitin and covalently linking the 11 kDa protein to the N-terminus of other proteins has been shown to increase their solubility. Like ubiquitin, the SUMO modifier is involved in the stabilization and localization of proteins *in vivo*. The SUMO modifier can be conjugated to both RedG and McpG *via* the use of the commercially available pETSUMO expression system. The pETSUMO vector contains the T7 promoter and the *lacO* operon for overproduction of the proteins in *E. coli*, and the SUMO protein is inserted downstream of an N-terminal $His_6$-tag which is used for protein purification.

Both *redG* and *mcpG*/pETSUMO expression vectors were constructed (section 4.3.1.2) and the plasmids were used with pDB1282 to co-transform *E. coli* BL21 (DE3) cells. The culture

growth and protein expression conditions used were based on the report of Sun and co-workers for the production of GenD1.[131] The cells were grown to an $O.D_{600}$ of 0.8. At this point, biosynthesis of the iron-sulphur cluster was induced using arabinose, and the culture was incubated with shaking for a further 45 minutes. The culture was then cooled to 4 °C before production of the Rieske proteins were induced using IPTG (0.1 mM) along with addition of 2.0 mM iron sulphate heptahydrate. The culture was incubated at 15 °C for 20 hours, with shaking at 100 rpm to reduce aeration of the culture. The cells were harvested and lysed by sonication and protein purification was undertaken by nickel affinity chromatography. All buffers and Falcon tubes were purged with argon gas for 15 minutes prior to purification.

The eluted fractions were analyzed by SDS-PAGE (figure 4.13). The gels display protein bands with the expected molecular weight for the recombinant $His_6$-SUMO-McpG and $His_6$-SUMO-RedG proteins. A smaller band can be observed beneath the expected band for $His_6$-SUMO-McpG, which is likely to be recombinant McpG from which the $His_6$-SUMO appendage has been cleaved during the purification. The 500 mM imidazole elution fractions were concentrated to 500 µL and the concentrations of the proteins were calculated to be 27 µM and 140 µM for McpG and RedG, respectively. Both proteins displayed a brown colouration indicating they contain an iron-sulphur cluster.

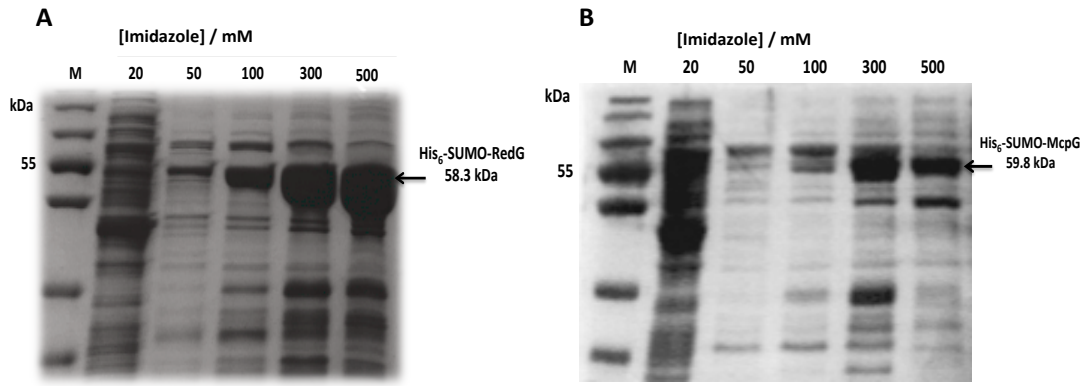



**Figure 4.13: A.** SDS-PAGE gel analysis of the purification of $His_6$-SUMO-RedG using a pETSUMO expression vector + pDB1282, *E. coli* BL21 (DE3) cells and LB medium; **B.** SDS-PAGE gel analysis of the purification of $His_6$-SUMO-McpG using a pETSUMO expression vector + pDB1282, *E. coli* BL21 (DE3) cells and LB medium.

The protein solutions were transferred to an argon-flushed quartz cuvette for UV-Vis spectroscopic analysis (figure 4.14). The UV-Vis spectrum for $His_6$-SUMO-McpG displayed an absorption maximum at 420 nm, similar to that observed for the RedG FeS cluster in section 4.3.4.4, implying the presence of a reduced Rieske [2Fe-2S] cluster.[139]

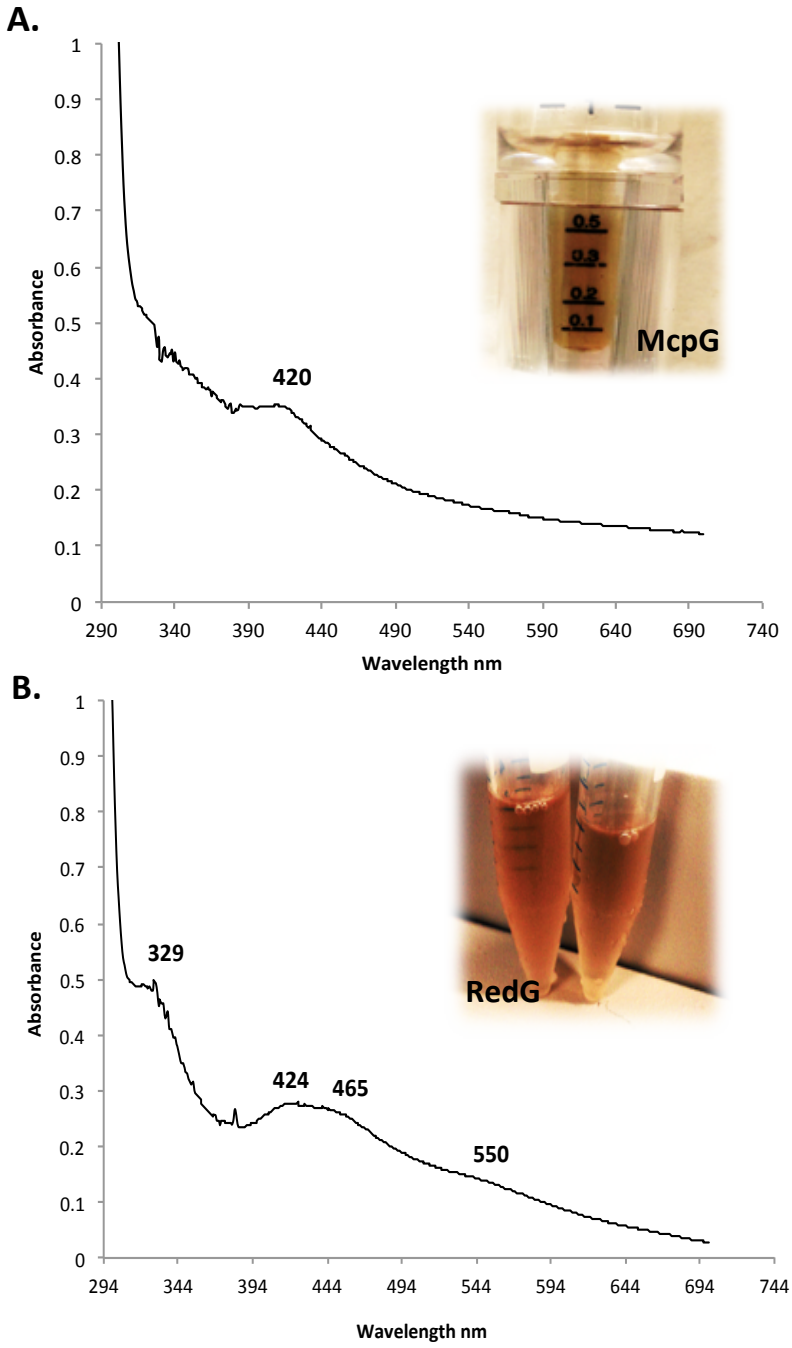



**Figure 4.14:** UV-Vis absorption spectra of aerobically purified recombinant protein co-expressed with pDB1282 using *E. coli* BL21 (DE3) cells cultured in ferrous ion supplemented LB medium; **A.** $His_6$-SUMO-McpG, **B.** $His_6$-SUMO-RedG.

The spectrum of $His_6$-SUMO-RedG displays absorption maxima at 329 nm and 424 nm, along with peaks at 465 nm and 550 nm (figure 4.14). These values are similar to that observed for the oxidized Rieske clusters of SxtT and MupW.[60,69] The spectrum for RedG also implies the presence of a mixture of reduced and oxidized FeS clusters. Since the production levels of $His_6$-SUMO-RedG were considerably higher then $His_6$-SUMO-McpG this construct was used in subsequent experiments.

With the purified recombinant $His_6$-SUMO-RedG protein containing a Rieske cluster in hand, a catalytic activity assay was undertaken using synthetic undecylprodigiosin (**23**), NADH, spinach ferredoxin and spinach ferredoxin reductase. The mixture was incubated at 37 °C overnight before the organics were extracted and analyzed by UHPLC-ESI-Q-TOF-MS. No streptorubin B (**15**) was observed in the extract. It is possible that the Rieske cluster/ non-heme iron centre degraded during the preparation of the assay mixture as a result of exposure to atmospheric oxygen. To test this hypothesis, the brown $His_6$-SUMO-RedG protein solution was exposed to atmospheric oxygen for 30 minutes and then examined by UV-Vis spectroscopy. During this time the solution became colourless, the protein precipitated and the corresponding UV-Vis spectrum no longer displayed the absorption maxima indicating the presence of a Rieske cluster. This suggested the FeS machinery undergoes rapid oxidative degradation (figure 4.15).

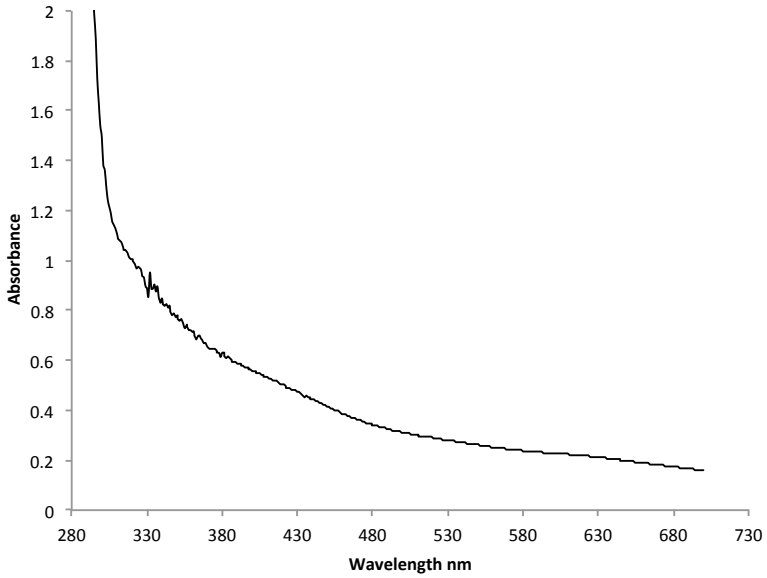



**Figure 4.15:** UV-Vis absorption spectrum of the aerobically purified recombinant $His_6$-SUMO-RedG from *E. coli* BL21 (DE3) containing PDB1282 cultured in ferrous ion supplemented LB medium after exposure to atmospheric oxygen for 30 minutes.

In a bid to reduce the exposure of the Rieske cluster to atmospheric oxygen, the purification of $His_6$-SUMO-RedG was repeated under anaerobic conditions in a glove-bag using buffers that had been purged with nitrogen gas (section 6.5.9.4). Cell growth conditions and protein production was carried out as described in section 6.5.9.2 and the cell-free extract was placed in the glove-bag. The protein was purified on a nickel affinity column, and the fraction eluted with 300 μM imidazole was concentrated to 500 μL, giving a protein concentration of 19 μM. The solution displayed the brown colour indicative of Rieske iron-sulphur cluster incorporation.

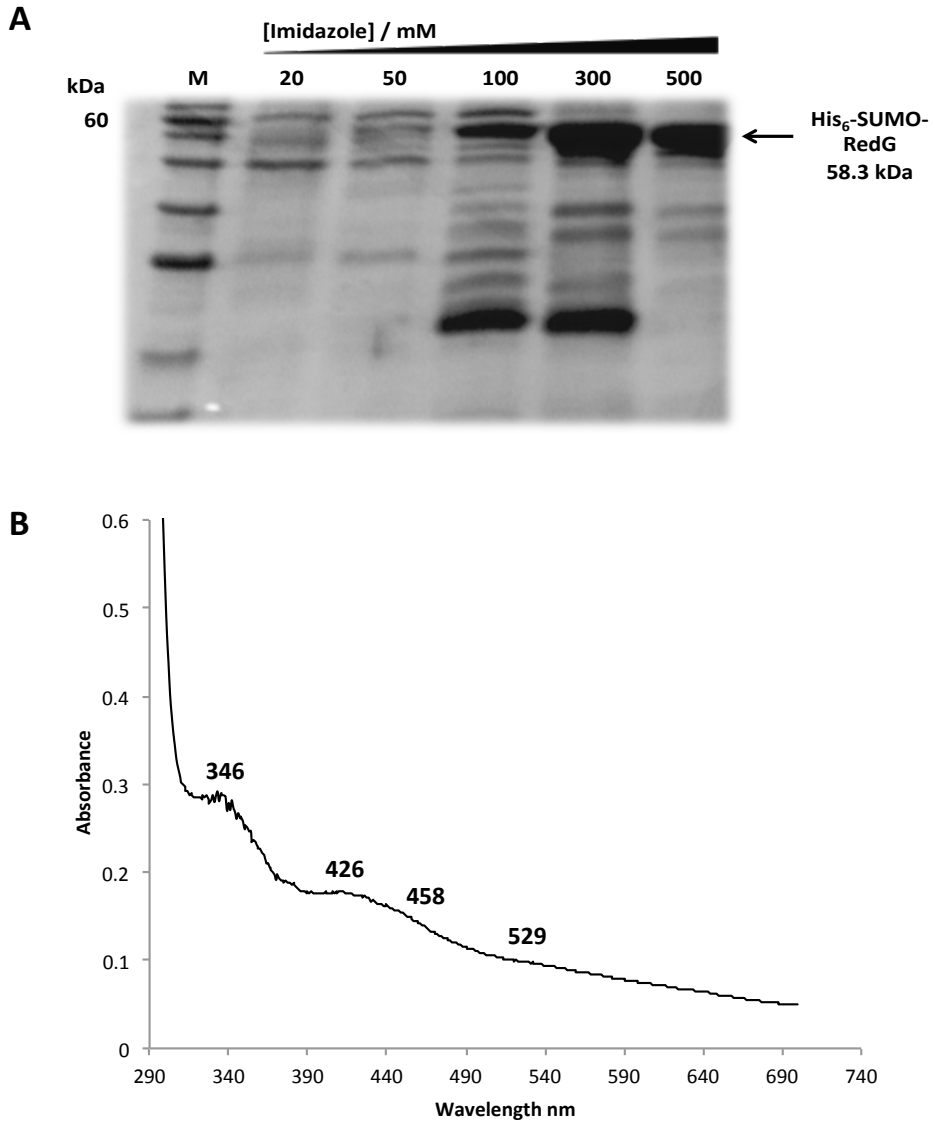



**Figure 4.16: A.** SDS-PAGE analysis of the purification of recombinant $His_6$-SUMO-RedG using a pETSUMO expression vector + pDB1282, *E. coli* BL21 (DE3) cells, LB medium and anaerobic conditions; **B.** UV-Vis spectrum of the purified protein.

The concentrated protein solution was transferred to a nitrogen flushed quartz cuvette for analysis by UV-Vis spectroscopy (figure 4.16). Similar absorption maxima were observed as for the protein purified under aerobic conditions. The protein solution still lost the brown colouration and subsequently precipitated when left open to the nitrogen atmosphere in the glove-bag. The loss of the brown colour suggests that the low levels of residual oxygen contained within the bag are significant enough to degrade the Rieske iron-sulphur cluster in 30 minutes. The instability of the FeS cluster significantly hinders efforts to reconstitute the oxidative carbocyclization activity of the protein. Production of Rieske proteins in *E. coli* typically results in the oxidized form of the Rieske cluster. Therefore, *in vitro* assays will require electron transfer proteins, such as a ferredoxin and a ferredoxin reductase. Producing the protein with a reduced Rieske cluster would allow single turnover assays to be undertaken. Chemical reduction of the oxidized Rieske cluster may increase exposure of the protein to atmospheric oxygen leading to its degradation. Therefore, attention was turned to reconstituting the Rieske cluster of RedG under reducing conditions *in vitro*. This would furnish the protein with a reduced iron-sulphur cluster that is primed for direct use in oxidative carbocyclization assays.

### 4.3.5 *In vitro* reconstitution of the $His_6$-SUMO-RedG Rieske cluster

The [4Fe-4S] protein GenD1 was reconstituted *in vitro* using dithiothreitol, sulphide ions and ferrous ions under anaerobic conditions.[131] Spontaneous formation of the iron-sulphur cluster was confirmed by UV-Vis spectroscopic analysis of the reaction mixture after purification on a PD-10 column. It was envisaged that a similar protocol could be used to reconstitute the iron-sulphur cluster of RedG *in vitro*.

Continuing to work in a glove-bag with nitrogen purged buffers, anaerobically purified $His_6$-RedG-SUMO was buffer exchanged using a PD-10 desalting column to remove imidazole. At this point, the brown colouration was no longer visible. Dithiothreitol was added (the sample was not concentrated prior to the reaction) and the solution was incubated for 15 minutes at room temperature. Ferrous ions and sulphide were added and the sample was incubated for a further two hours at room temperature. The sample displayed a black colouration after the addition of sulphide and ferrous ions to the reduced protein sample (**B**, figure 4.17). The sample was purified/buffer exchanged again using a PD-10 desalting column.

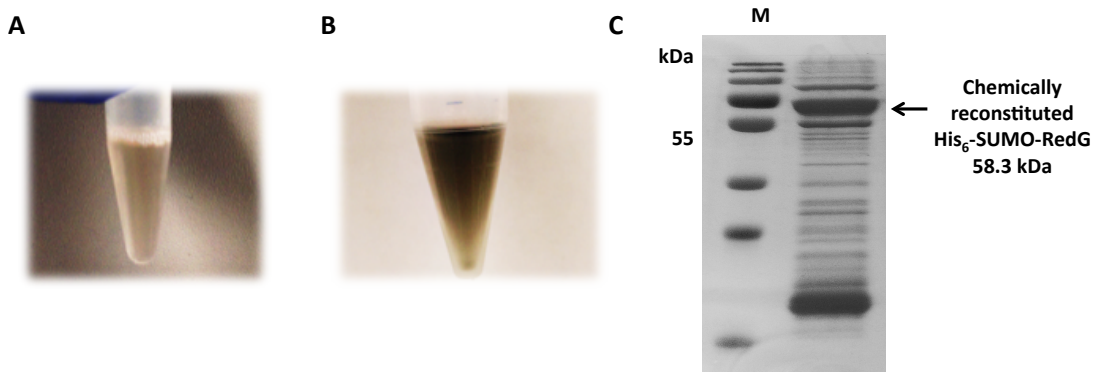



**Figure 4.17: A.** Purified $His_6$-SUMO-RedG post PD-10 column; **B.** Purified $His_6$-SUMO-RedG post PD-10 column + DTT + Fe (II) + $S^{2-}$ under anaerobic conditions; **C.** SDS-PAGE analysis of purified and reconstituted recombinant $His_6$-SUMO-RedG.

The reconstituted protein was analyzed by SDS-PAGE (figure 4.17). The gel displayed a band of the expected molecular weight for reconstituted recombinant $His_6$-SUMO-RedG. The protein solution was transferred to a quartz cuvette, under anaerobic conditions, for analysis by UV-Vis spectroscopy. The resulting spectrum displayed absorption maxima at 423 nm and a small absorption maxima was also observed at 533 nm (figure 4.18), which is more closely related to the observed spectra for the reduced Rieske cluster of NDO, MupW and the Rieske protein from *T. thermophiles*.[60,137,139] The reconstituted Rieske cluster therefore differs to that in the purified $His_6$-SUMO-RedG protein (figure 4.14 and 4.16). The absorption maximum at 458 nm has shifted to 423 nm, which is more similar to a reduced Rieske cluster.

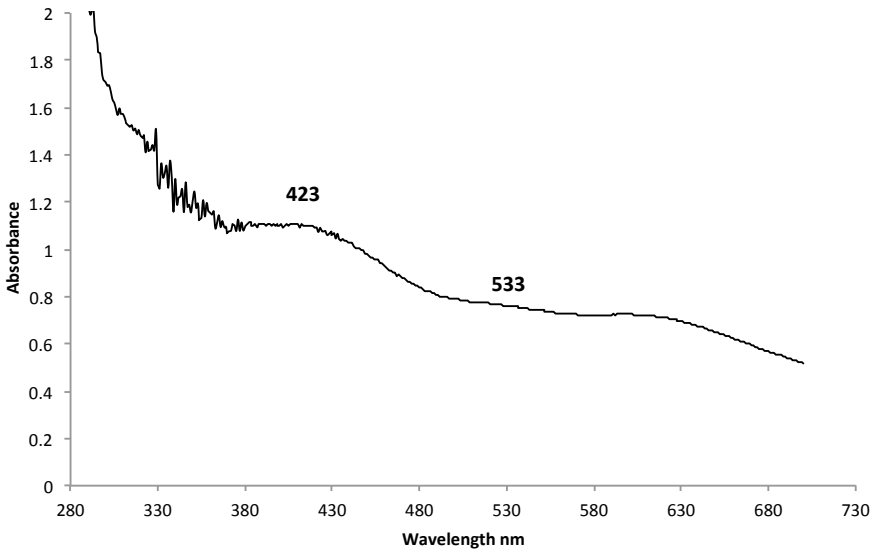



**Figure 4.18:** UV-Vis spectrum of anaerobically purified recombinant $His_6$-SUMO-RedG in which the FeS cluster has been reconstituted.

As the UV-Vis spectroscopic analysis indicated the presence of a reduced Rieske cluster, a single turnover experiment was attempted using UHPLC-ESI-Q-TOF-MS. No streptorubin B (**15**) could be detected. The reconstituted RedG protein was exposed to atmospheric oxygen for 2 hours and subsequently analyzed by UV-Vis spectroscopy (figure 4.19). The resulting UV-Vis spectrum no longer displayed the absorption maxima corresponding to either an oxidized or reduced Rieske cluster, suggesting that complete degradation had occurred.

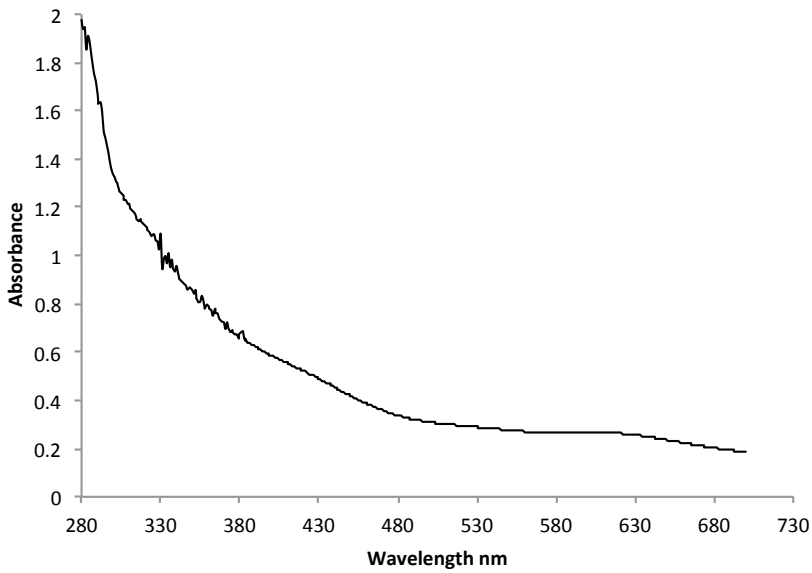



**Figure 4.19:** UV-Vis spectrum of anaerobically purified recombinant $His_6$-SUMO-RedG in which the FeS cluster has been reconstituted and exposed to atmospheric oxygen for 2 hours.

## 4.4 Conclusions and perspectives

The production and purification of soluble RedG and McpG with an intact Rieske cluster described in this work are the first examples of recombinant Rieske proteins that are involved in prodiginine biosynthesis to be isolated and purified. The use of a $His_6$-SUMO appendage on the N-terminus of either protein increased both the production and solubility levels. The use of a second plasmid encoding arabinose-induceable biosynthesis of iron-sulphur clusters and optimized production conditions increased Rieske cluster incorporation into both RedG and McpG. Analysis of the Rieske cluster by UV-Vis spectroscopy provided the first experimental evidence that the RedG and McpG iron-sulphur cluster predicted by bioinformatics is in fact a Rieske [2Fe-2S] cluster. A method for the chemical reconstitution of the RedG Rieske cluster using dithiothreitol, iron (II) sulphate heptahydrate and sodium sulphide under anaerobic conditions spontaneously re-assembled the [2Fe-2S] machinery in reduced form, as confirmed by UV-Vis spectroscopy. The established reconstitution method will aid in the future *in vitro* reconstitution of the oxidative carbocyclization reaction of undecylprodigiosin (**23**) to streptorubin B (**15**) and metacycloprodigiosin (**25**) respectively, whereby the Rieske protein can be purified and regenerated when desired to avoid degradation by atmospheric oxygen. Further characterization of the RedG and McpG Rieske cluster by electron paramagnetic resonance (EPR) spectroscopy can be used to provide additional evidence to confirm the Rieske [2Fe-2S] nature of the bioinformatically predicted iron-sulphur cluster. Air-oxidized Rieske clusters are EPR silent and so a method to fully reduce the cluster will be required or the experiment must be conducted on the reconstituted proteins. Rieske clusters display a particularly low $g_{av}$ value in the X-Band EPR spectrum and so this would provide further definitive characterization along side the distinct UV-Vis spectra.

Attempts were undertaken to demonstrate the RedG and McpG oxidative carbocyclization reactions *in vitro* using both the recombinantly produced and reconstituted Rieske proteins and neither were successful. Using the established and optimized production and purification conditions and the FeS cluster reconstitution protocol, further attempts can be made to reconstitute the RedG and McpG catalyzed oxidative carbocyclization reactions of undecylprodigiosin (**23**) to streptorubin B (**15**) or metacycloprodigiosin (**25**), respectively. The production of soluble Rieske proteins with confirmed intact Rieske clusters is a significant step forward in reconstituting the oxidative carbocyclization reactions *in vitro*, although to achieve this it may be necessary to undertake protein purifications and turnover assays in a strictly anaerobic chamber. Using the established protocols in the preliminary

work described previously it will now be possible to repeat procedures in anaerobic chambers to also further improve protein purification and FeS cluster incorporation levels. It may also be necessary to screen all potential combinations of ferredoxins and ferredoxin reductases in *Streptomyces coelicolor* M511 to identify viable electron transfer partners. The use of stoichiometric hydrogen peroxide to generate the active oxidant, along with NADH, is a potential alternative to this.

# Chapter 5: Conclusions & future work

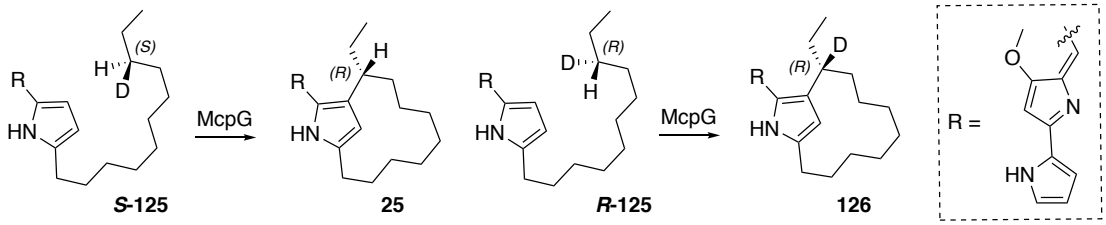
## 5.1 Construction of an *in vivo* system using yeast as a platform to probe McpG catalysis

A robust *in vivo* system to investigate the non-heme iron-dependent Rieske oxygenase McpG was constructed *via* heterologous expression of the entire metacycloprodigiosin (**25**) biosynthetic gene cluster. The cluster was captured from the genome of *S. longispororuber* using TAR cloning and production of undecylprodigiosin (**23**) and metacycloprodigiosin (**25**) was confirmed and optimized using *S. albus* as a heterologous host. TAR was used to create an in-frame deletion of the *mcpLK* gene, abrogating the biosynthesis of the late-staged intermediate 2-UP (**28**).

A procedure for feeding synthetic 2-UP (**28**) alongside MBC (**21**) to the deletion mutant was developed and the restoration of metacycloprodigiosin (**25**) production by the strain was subsequently confirmed by HR-MS and MS/MS analysis. The feeding procedure was optimized to maximize the levels of metacycloprodigiosin (**25**) production. Several 2-UP analogues were fed to the optimized system as probes of the mechanism and stereochemical course of the McpG-catalyzed reaction and to investigate the substrate tolerance of the McpH and McpG enzymes.

## 5.2 Stereochemical course of McpG-catalyzed oxidative carbocyclization

The synthesis of 2-UP analogues ***R*-120** and ***S*-120** in which each of the pro-chiral protons at C9' of the alkyl chain were replaced by deuterium labels was developed and completed. Feeding of ***R*-120** to the deletion mutant resulted in 96 % retention of the deuterium label, whereas feeding of ***S*-120** resulted in 53 % abstraction of the deuterium label from the metacycloprodigiosin (**25**) produced. These data show that McpG preferentially abstracts the pro-*S* proton at C9' of the undecylprodigiosin (**23**) alkyl chain to furnish metacycloprodigiosin (**25**), which exhibits *R*-absolute stereochemistry (scheme 5.1). McpG therefore catalyzes oxidative carbocyclization with inversion of stereochemistry at the C9' carbon centre undergoing functionalization.




**Scheme 5.1:** The major products observed from feeding deuterium-labeled 2-UP analogues **120** alongside MBC (**21**) to a 2-UP (**28**) deficient mutant.

The reason that complete loss of the deuterium label is not observed is because abstraction of a deuterium atom rather than a hydrogen atom is energetically less favourable and so generates a primary kinetic isotope effect. It is possible that the kinetic isotope effect observed when ***S*-120** is fed to the deletion mutant causes McpG to instead abstract the *pro-R* hydrogen. This hypothesis can be investigated by analyzing the product of feeding *S*-**120** using circular dichroism spectroscopy or by chiral HPLC to determine its enantiopurity. A less likely explanation is that the kinetic isotope effect causes McpG-catalyzed C-H abstraction at an alternative carbon atom.

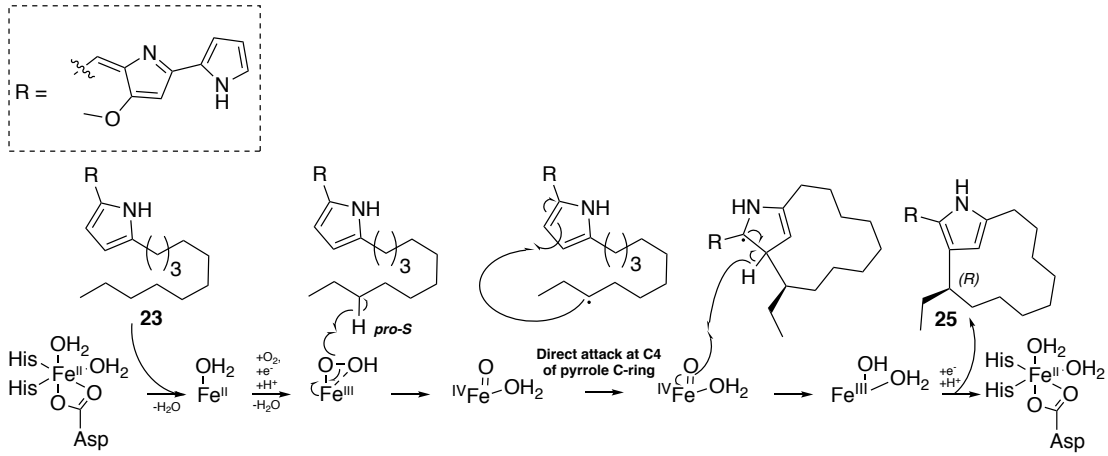
The stereochemical course of the McpG-catalyzed reaction is analogous to that observed for oxidative carbocyclization catalyzed by the homologous Rieske oxygenase RedG, in the biosynthesis of the closely-related prodiginine metabolite, streptorubin B (**15**).[49] It therefore seems likely that all Rieske oxygenase-catalyzed oxidative carbocyclizations in the biosynthesis of prodiginine natural products follows a similar stereochemical course. In the cases of non-heme iron-dependent oxygenases that catalyze oxidative cyclization reactions with retention of stereochemistry, often a ligand of the non-heme iron centre directly traps the radical intermediate and so the cyclization occurs on the same face of the carbon centre as the proton was abstracted from. The orientation of the cyclization is such that the alkyl chain shields the pyrrole moieties of the undecylprodigiosin skeleton to prevent unwanted oxidation reactions.

## 5.3 Mechanism of McpG catalysis

The oxidative carbocyclization catalyzed by Rieske oxygenase enzymes in prodiginine biosynthesis has been proposed to proceed *via* the formation of a radical intermediate (**139**) at the carbon-centre undergoing functionalization (scheme 5.2).[49] The radical-based mechanism was proposed based on the mechanism of the well-studied Rieske oxygenase NDO, and this pathway has previously been investigated by Withall, using RedG in streptorubin B (**15**) biosynthesis.[49]

Results from this work are consistent with a radical intermediate. McpG was shown to be able to catalyze both epoxidation of an alkene and oxidation of alcohol **210** to ketone **222**. Both of these reactions are commonly carried out cytochrome P450s and are well established to proceed *via* radical intermediates. The synthesis of 2-UP analogue **197** bearing a cyclopropane moiety adjacent to the site of C-H abstraction was developed and used in a bid to 'trap' the proposed C9 alkyl radical intermediate. Feeding the synthetic analogue **197** to the *mcpLK* deletion generated undecylprodigiosin analogue **199** in which the cyclopropane appears to have been cleaved. The nature of the reactive intermediate formed *via* McpG catalyzed C-H abstraction however cannot yet be fully confirmed and the corresponding carbon centred cationic species remains a potential alternative. To further study the mechanism then the above mentioned 2-UP analogues could be fed to McpG that had been reconstituted *in vitro* to analyze reaction kinetics and for investigation using crystallography.

Following formation of the alkyl radical (or cationic) intermediate one of two potential events can occur. Either the radical intermediate is hydroxylated by rebound onto the active iron-oxo species, followed by nucleophilic displacement of the hydroxyl group by C4 of the proximal pyrrole ring, or the radical intermediate directly attacks the C4 position of the pyrrole C-ring and a second hydrogen abstraction follows to regenerate a pyrrole. This work has established that hydroxylated analogue **210** is not an intermediate in the catalytic cycle, since it was not converted to metacycloprodigiosin (**25**). Thus, this work provides strong evidence to support that the generation of a hydroxylated undecylprodigiosin analogue **210** as a product of radical rebound does not occur. Questions still remain about the nature of the active iron-oxo species responsible for C-H abstraction, which could be either an Fe(III)O(OH) species or an Fe(IV)=O(OH) species.




**Scheme 5.2**: Proposed mechanism of McpG catalysis. Note that alternative cationic species are also a possibility.

## 5.4 Substrate tolerance of McpH & McpG

A variety of 2-UP analogues were synthesized in which the alkyl chain was decorated with various functionality. When 2-UP analogues with nine, ten and twelve carbons in their alkyl chains (relative to the eleven carbon chain in the natural substrate) where fed alongside MBC (**21**) to the deletion mutants, the corresponding undecylprodigiosin analogues were formed in each case. McpG accepted both the ten and twelve carbon-chained analogues. The nine carbon-chained analogue (**173**) was not cyclized by McpG. The formation of a primary radical at C-9' by McpG would be energetically unfavorable; explaining why the nine-carbon chained metacycloprodigiosin analogue **174** is not formed. This result contrasts with the proposed reaction catalyzed by TamC, a Rieske oxygenase that is proposed to catalyze the cyclization of tambjamine YP1 (**36**) to tambjamine MYP1 (**37**).[36] TamC is predicted to catalyze cyclization *via* the formation of a primary radical, suggesting an alternative mechanism for the TamC-catalyzed reaction.

Feeding of a 2-UP analogue with a methyl branch at C11' (**154**) resulted in the production of the corresponding undecylprodigiosin analogue (**32**), but this was not accepted as a substrate by McpG. The undecylprodigiosin analogue **32** is proposed to be the substrate for Rieske oxygenase RphG1 in the biosynthesis of prodigiosin R1 (**33**), and so highlighting potential differences in the substrate preference of RedG homologues.

Feeding 2-UP analogues **157** and **155** where terminal alkene and alkyne moieties had been inserted (adjacent to the site of C-H abstraction) both resulted in the production of their corresponding undecylprodigiosin analogues, however neither were cyclized by McpG.

This work has highlighted the broad substrate tolerance of McpH. All 2-UP analogues that were fed alongside MBC (**21**) to the *mcpLK* deletion mutant were accepted by McpH, resulting in production of their corresponding undecylprodigiosin analogue, with similar peak intensities in all cases.

## 5.5 Expression and purification of RedG and McpG and Rieske [2Fe-2S] cluster characterization

Prior to this study, no Rieske oxygenase enzymes involved in the biosynthesis of prodiginine alkaloids had been studied *in vitro*. The *redG* and *mcpG* genes were therefore cloned into a range of protein expression vectors and various conditions for the production of soluble protein were screened. The resulting proteins were analyzed for the presence of an intact Rieske [2Fe-2S] cluster, initially *via* the observation of a brown colour. The use of a small modifying protein appendage attached through the use of a pETSUMO vector was found to generate soluble recombinant RedG and McpG. Co-expression of the *redG*/pETSUMO and *mcpG*/pETSUMO vector with the pDB1282 plasmid (an arabinose inducible vector that is responsible for the biosynthesis of iron-sulphur clusters) and subsequent aerobic or anaerobic purification furnished protein solutions that displayed a brown colour, indicative of an iron-sulphur cluster. The cluster present in both oxygenases was confirmed to be Rieske-type by UV-Vis absorbance spectroscopy and comparison of the absorption maxima with literature values for known Rieske oxygenase enzymes. The absorbances due to the Rieske cluster of RedG were seen to disappear after exposure of the protein to atmospheric oxygen for thirty minutes, suggesting oxidative degradation of the iron-sulphur cluster. The Rieske cluster of $His_6$-SUMO-RedG was reconstituted using dithiothreitol, ferrous ion and sulphide under anaerobic conditions, and confirmed by UV-Vis absorption spectroscopy. Comparison of the absorption maxima with literature values for known Rieske oxygenases confirmed that reconstitution resulted in formation of the reduced form of the Rieske cluster.

Attempts to reconstitute the activity of RedG using synthetic undecylprodigiosin (**23**), NADH, spinach ferredoxin and ferredoxin reductase were unsuccessful, using both the oxidized and reconstituted reduced form of the Rieske enzyme. Although the overproduction, purification and characterization of RedG and McpG with intact [2Fe-2S] clusters are major steps towards reconstituting the oxidative carbocyclization activity of these enzymes, significant hurdles remain. The inherent instability towards atmospheric oxygen results in almost immediate inactivation of these proteins. Rigorous anaerobic environments must be used when handling these proteins in the future. The Rieske cluster may also be sensitive to

the pH of the reaction buffer, which will require further investigation. There are also several ferredoxins and ferredoxin reductases present in *S. coelicolor*, which are likely to be involved in the electron transfer cascade for the oxidative carbocyclization reactions catalyzed by these enzymes. Reconstituting this reaction using all possible combinations of these proteins will be necessary to find the optimum reductase partners required for successful reaction. Hydrogen peroxide could also be used to form the reactive iron-oxo species and by-pass the need for the electron transport proteins.

## 5.6 Future Work

This work, combined with the investigation into the mechanism of the RedG-catalyzed carbocyclization of undecylprodigiosin (**23**) to form streptorubin B (**15**) by both Haynes and Withall, provides data consistent for the formation of an alky radical as a result of C-H abstraction by Rieske oxygenases in prodiginine biosynthesis, however the formation of the alternative carbocation can not be ruled out.[49,88,89] To further study the mechanism of these enzymes it would be highly beneficial to develop a robust, functional *in vitro* system/assay. Synthetic mechanistic probes can be toxic to the cells and also key products can be produced in quantities that are significantly lower than other metabolites produced by the strain, rendering themselves either non-observable or not possible to isolate and characterize further by NMR spectroscopy.

This work has made substantial progress towards the reconstitution of the RedG and McpG-catalyzed oxidative carbocyclization reactions. X-ray crystallography could provide structures of both the RedG and McpG proteins. Solving the structure of these enzymes will reveal key features of their active sites, which would provide further mechanistic insights. Co-crystallizing along with the substrate would reveal why the protein catalyzes cyclization with inversion of stereochemistry at the carbon centre that undergoes functionalization. The *in vitro* system may also facilitate the characterization of low abundance products and kinetics studies. Functional enzymes will be a key step towards generating a detailed mechanistic understanding of this remarkable cyclization reaction.

# Chapter 6: Experimental

## 6.1 General chemistry experimental

Solvents and reagents were purchased from Sigma Aldrich, Fischer Scientific, Acros Organic or VWR and used without further purification unless stated otherwise. Anhydrous solvents were obtained by distillation under argon using calcium hydride as a drying reagent and stored over 4 Å molecular sieves. All water used was deionized unless otherwise stated. Thin-layer chromatography (TLC) was performed using aluminum-backed plates pre-coated with Merck silica gel 60 F254 and compounds were visualized using potassium permanganate and vanillin stains and UV radiation. Flash chromatography was performed using silica gel (Aldrich, 40-63 μM, 40-60 Å). Liquid chromatography-high resolution mas spectrometry (LC-HRMS) was performed using a Thermo Scientific Dionex Ultimate 3000 RS UHLPC coupled to a Bruker MaXis impact ESI-TOF mass spectrometer using an Agilent ZORBAX eclipse XDB-C18 analytical column (2.1 × 5mm, 1.8 μM) for small molecules and sodium formate (10 mM) was used for internal calibration. Solvents were evaporated using a BUCHI rotavapor R-200 or R210 connected to BUCHI vacuum pump V-700. All NMR spectra were measured using a Bruker DPX-300 ($^{1}$H: 300 MHz; $^{13}$C: 75 MHz) or DPX-400 ($^{1}$H: 400 MHz; $^{13}$C: 100 MHz) instrument containing a Proton Carbon Dual Probe. High field NMR spectra were measured on a Bruker AV-500 ($^{1}$H: 500 MHz; $^{13}$C: 125 MHz) and obtained by the University of Warwick NMR service using deuterated solvents purchased from Sigma Aldrich. $^{1}$H-NMR data are recorded as follows; chemical shift (δ, ppm); multiplicity, integration, proton assignment and coupling constant (Hz). The multiplicities are stated as; singlet (s); doublet (d); triplet (t); quartet (q); quintet (quint); doublet of doublets (dd); doublet of doublet of triplets (ddt); multiplet (m); broad singlet (bs). All spectra were calibrated relative to a standard ($^{1}$H: $CDCl_3$, 7.26 ppm; $^{13}$C: $CDCl_3$, 77.16 ppm) or ($^{1}$H: $D_6$-DMSO, 2.50 ppm; $^{13}$C: $D_6$-DMSO, 39.52 ppm) and the assignments were made with the aid of COSY, HMBC and HSQC spectra. Low-resolution mass spectra were measured using an Agilent 6130B Single Quad ESI spectrometer and high-resolution mass spectra were recorded by the University of Warwick Mass Spectrometry service on a Bruker MaXis ESI spectrometer. IR spectra were recorded on a Perkin Elmer Avatar 320 Fourier Transformation spectrometer and the selected absorptions stated are given in units of wave numbers.

## 6.2 General chemical procedures

### 6.2.1 Pyrrole distillation

Pyrrole was heated to 130-150 °C with vigorous stirring. Pure colourless pyrrole was collected from a brown waste residue using distillation equipment and the freshly prepared pyrrole was used directly in reactions.

### 6.2.2 *n*-Butyl lithium titration

To a solution of diphenyl acetic acid (100 mg, 0.50 mmol) in dry THF (5 mL) under argon was added *n*-butyl lithium solution until the formation of a yellow colour. The average titre of three attempts was used as the initial *n*-butyl lithium concentration.

### 6.2.3 IBX synthesis

To a solution of oxone (12.9 g, 2.09 mmol) in water (52 mL) was added 2-iodobenzoic acid (4.0 g, 16.1 mmol) and the resulting solution was stirred for 3 hours at 75 °C. The reaction mixture was put on ice and stirred for a further 2 hours. The resulting white precipitate was collected by Buchner filtration and the solid was washed with water and acetone to give the final product as a white solid (3.95 g, 87%) with identical spectroscopic data to that in the literature.

### 6.2.4 Oxidation of alcohols to carboxylic acids using Jones' reagent prepared *in situ*

Sulphuric acid (18 M, 0.65 mL) was slowly added to a solution of chromium trioxide (0.80 g) in acetone (10 mL) at 0 °C. Water (3 mL) was slowly added and the resulting red solution was added drop-wise to the alcohol in question pre-dissolved in acetone (5 mL) and at 0 °C, until the solution changed colour from red to green and then back to red. The reaction mixture was quenched with ice-cold 2-propanol and concentrated *in vacuo*. The resulting residue was re-dissolved in water and the organics extracted using chloroform, dried ($MgSO_4$) and concentrated *in vacuo*. The crude product was purified by flash chromatography.

### 6.2.5 Coupling of acids to alcohols

To a solution of acid (5.0 eqv) and EDC (1 eqv) in chloroform (1 mL) was added 4-dimethylaminopyridine (0.1 eqv) and alcohol (1 eqv) and stirred for 24 hours at RT. The reaction mixture was filtered through cotton wool and concentrated *in vacuo.* The crude product was purified by flash chromatography.

### 6.2.6 Coupling of carboxylic acids to pyrrole

To a solution of carboxylic acid (1 eqv) in toluene and under argon was added 2,2'-dipyridyl disulphide (2 eqv) and triphenyl phosphine (2 eqv) forming a bright green solution. The reaction mixture was stirred for 20 hours. This solution was cooled to -78 $^{\circ}$C and pyrryl magnesium bromide (3 eqv) was added drop-wise over 5 minutes (pyrryl magnesium bromide was prepared freshly by the addition of ethyl magnesium bromide to freshly distilled pyrrole in toluene, under argon and at -40 $^{\circ}$C. The reaction mixture was stirred for 1 hour before quenching with saturated $NH_4Cl$. The organics were extracted with EtOAc (50 mL) and the combined organics were washed with 5% $K_2CO_3$, HCl (2 M), brine and dried ($MgSO_4$) and concentrated *in vacuo.* The crude product was purified by flash chromatography.

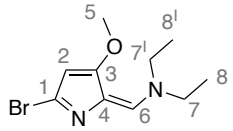
### 6.2.7 Reduction of acyl ketones

To propan-2-ol was added sodium borohydride (10 eqv) and the resulting suspension stirred at 95 $^{\circ}$C for 1 hour. The acyl ketone, dissolved in propan-2-ol, was added to the suspension and stirred for 20 hours. Water was added and the organics extracted using EtOAc before drying ($MgSO_4$) and concentration *in vacuo.* The crude product was purified by flash chromatography.

## 6.3 Chemical Synthesis

### 6.3.1 Synthesis of control substrates 2-UP (28) & MBC (21)

***N*-(5-bromo-3-methoxy-pyrrol-2-ylidenemethyl)-diethylamine 100**



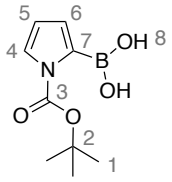

To a solution of diethyl formamide (1.63 mL, 14.6 mmol) in $CH_2Cl_2$ (10 mL) under argon and cooled to 0 °C was added phosphorous oxybromide (4.18 g, 14.6 mmol) in DCM (10 mL). The resulting solution was stirred for 25 minutes until a white precipitate formed and 4-methoxy-3-pyrolli-2-one (750 mg, 6.63 mmol) in $CH_2Cl_2$ (10 ml) was added and the solution heated to 40 °C for 6 hours. Water (80 mL) was added and the pH of the solution adjusted to 8 by addition of NaOH (2.35 M) before EtOAc (100 mL) was added and the solution filtered through a celite pad. The aqueous layer was separated and extracted using EtOAc (2 × 50 mL) and the combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated *in vacuo* to give a brown oil (1.52 g, 89 %). The compound did not require further purification. Spectroscopic data identical to literature compound.[49]

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 6.99 (s, 1H, H2), 5.59 (s, 1H, H6), 4.15 (q, 2H, H7 or H7', *J* = 7.0 Hz), 3.76 (s, 3H, H5), 3.42 (q, 2H, H7 or H7', *J* = 7.0 Hz), 1.31 (m, 6H, H8 and H8')

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 166.9 (C3), 138.7 (C2), 133.9 (C4), 121.0 (C1), 96.6 (C6), 58.1 (C5), 51.2 (C7 or C7'), 44.7 (C7 or C7'), 14.7 (C8 or C8'), 12.6 (C8 or C8')

LR-MS (ESI, +ve, MeOH) *m/z* (%): 261.0 (100, $[M+H]^+$)

***N*-Boc-pyrrole-2-boronic acid 98**




To a solution of 2,2,6,6-tetramethyl piperidine (3.37 mL, 23.3 mmol) in dry THF (20 mL) under argon, cooled to -78 °C, was added *n*-butyl lithium (14.9 mL, 23.32 mmol) and the resulting solution was left to warm to room temperature for 30 minutes. The solution was cooled back to -78 °C and *N*-Boc pyrrole (2.0 mL, 12.0 mmol) was added. The solution was left to warm to room temperature overnight. HCl (10 mL, 0.5 M) was added and the THF removed under vacuum. Water (30 mL) was added and the organics extracted using $Et_2O$ (4 ×

20 mL), dried over $MgSO_4$ and then concentrated *in vacuo* to form a brown precipitate. The precipitate was then cooled to -78 °C and the precipitate collected by vacuum filtration (2.23 g, 96 %). Spectroscopic data identical to literature compound.[49]

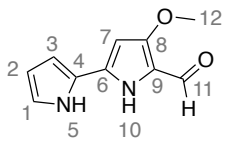
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.45 (dd, 1H, H6, *J* = 3.0 Hz, *J* = 2.0 Hz), 7.09 (dd, 1H, H4, *J* = 3.5 Hz, *J* = 1.5 Hz), 6.84 (s, 2H, H8), 6.26 (t, 1H, H5, *J* = 3.5 Hz), 1.62 (s, 9H, H1)

$^{13}$C-NMR (100 MHz, $CDCl_3$) δ: 155.9 (C3), 128.8 (C4), 127.2 (C6), 112.2 (C5), 85.6 (C2), 28.1 (C1), (C7 missing)

IR ($V_{max}$/cm$^{-1}$): 3179 (broad) (O-H stretch), 2985 (C-H stretch), 1704 (C=O stretch), 1336 (C-N stretch), 1079 (C-O stretch), 1031.91 (B-C stretch)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 234.1 (100, $[M+Na]^+$)

**4-Methoxy-2,2'-bipyrrole-carboxaldehyde (MBC) 21**




A suspension of palladium (II) acetate (143 mg, 5.87 mmol) and triphenyl phosphine (693 mg, 2.64 mmol) in degassed toluene (5 mL) and under argon was heated to 70 °C for 20 minutes. To this was added *N*-(5-bromo-3-methoxy-pyrrol-2-ylidenemethyl)-diethyl-amine **100** (1.52 g, 5.87 mmol) and *N*-Boc-pyrrole-2-boronic acid **98** (1.36 g, 6.45 mmol) in degassed (10 %) water : 1,4-dioxane (10 ml) and sodium carbonate (1.86 g, 17.6 mmol). The solution was heated to 100 °C for 3 hours. Two portions of sodium methoxide (349 mg, 1.10 mmol) were added and the reaction left to stir for 15 minutes. The reaction mixture was poured into water, neutralized to pH 7 with HCl (2 M) and stirred for 30 minutes. The resulting brown precipitate was collected by vacuum filtration and washed with water (2 × 10 mL) and acetone (2 × 5 mL) to give the final product as a yellow solid (481 mg, 43 %). Spectroscopic data identical to literature compound.[49]

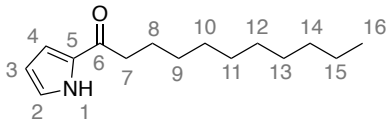
$^1$H-NMR (500 MHz, $D_6$-DMSO) δ: 11.50 (bs, 1H, H10), 11.32 (bs, 1H, H5), 9.39 (s, 1H, H11), 7.00 (m, 1H, H1), 6.84 (m, 1H, H3), 6.36 (m, 1H, H7), 6.21 (m, 1H, H2), 3.92 (s, 3H, H12)

$^{13}$C-NMR (125 MHz, $D_6$-DMSO) δ: 171.6 (C11), 159.5 (C8), 133.8 (C6), 123.4 (C4), 120.4 (C1), 117.4 (C9), 109.3 (C2), 108.3 (C3), 90.9 (C7), 57.8 (C12)

LR-MS (ESI, -ve, MeOH) *m/z* (%): 189.1 (100, $[M-H]^-$)

HR-MS (ESI, -ve, MeOH) *m/z* (%): calculated *m/z* for $C_{10}H_9N_2O_2$: 189.0670 $[M-H]^-$; observed 189.0673 $[M-H]^-$

**1-(1-H-pyrrol-2-yl)nonan-1-one 95**



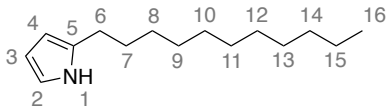

The compound was synthesized from undecanoic acid (300 mg) using the general procedure for the coupling of carboxylic acids to pyrrole. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a white powder (450 mg, 84 %). Spectroscopic data identical to literature compound.[49]

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.50 (bs, 1H, H1), 7.02 (m, 1H, H2), 6.91 (m, 1H, H4), 6.28 (m, 1H, H3), 2.75 (t, 2H, H7, *J* = 7.5 Hz), 1.74-1.68 (m, 2H, H8), 1.37-1.26 (m, 14H, H9-H15), 0.88 (t, 3H, H16, *J* = 7.0 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.3 (C6), 132.2 (C5), 124.4 (C2), 116.0 (C4), 110.7 (C3), 38.2 (C7), 32.0, 29.7, 29.6, 29.6, 29.6, 29.45, 25.4, 22.8 (C8-C15), 14.3 (C16)

HR-MS (ESI, -ve, MeOH) *m/z* (%): calculated *m/z* for $C_{15}H_{26}NO$: 236.2009 $[M+H]^+$; observed 236.2008 $[M+H]^+$

**2-Undecylpyrrole (2-UP) 28**




The compound was synthesized using the general procedure for the reduction of acyl ketones to give the crude product as a brown oil which was purified by flash chromatography ($SiO_2$; 1:10; EtOAC : Petroleum ether) to give the final product as a white solid (50 mg, 82 %). Spectroscopic data identical to literature compound.[49]

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.90 (bs, 1H, H1), 6.67 (m, 1H, H2), 6.13 (m, 1H, H3), 5.91 (m, 1H, H4), 2.59 (t, 2H, H6, *J* = 7.5 Hz), 1.64-1.58 (m, 2H, H7), 1.39-1.26 (m, 16H, H8-H15), 0.88 (t, 3H, H16, *J* = 6.5 Hz)

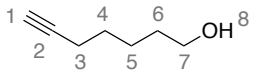
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.1 (C5), 116.1 (C2), 108.4 (C3), 104.8 (C4), 32.1, 29.8, 29.8, 29.8, 29.7, 29.6 (C7), 29.5, 29.5, 27.9 (C6), 22.8, 14.3 (C16)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 220.2 (60, $[M-H]^-$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{26}N$: 220.2071 $[M-H]^-$; observed 220.2071 $[M-H]^-$

### 6.3.2 Synthesis of stereochemical probes

**6-heptyn-1-ol 108**




To ethylenediamine (80 ml) under argon, cooled to 0 °C, was added sodium hydride powder (5.18 g, 0.22 mol). The resulting solution was vigorously stirred for 2 hours until a purple colour was seen and fizzing had ceased. The reaction mixture was then heated to 60 °C for 1 hour until the solution became blue. The temperature was then lowered to 45 °C before 3-heptyn-1-ol (8.0 ml, 7.0 mmol) was added drop-wise and then heated again to 60 °C once more, for a further 1 hour. The solution was then cooled to 0 °C before quenching with HCl (30 mL, 2 M) and the aqueous solution extracted using $Et_2O$ (6 × 100ml). The combined organics were dried over $MgSO_4$ and concentrated *in vacuo* to give an orange oil. The crude material was purified by flash chromatography ($SiO_2$; 1:1; EtOAc : Petroleum ether) to give the final product as an orange oil (5.66 g, 70 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 3.65 (t, 2H, H7, $J$ = 6.5 Hz), 2.22-2.19 (td, 2H, H3, $J_1$ =7.0 Hz $J_2$ = 2.5 Hz), 1.94 (t, 1H, H1, $J$ = 2.5 Hz), 1.62-1.54 (m, 4H, H5 & H6), 1.52-1.45 (m, 2H, H4), 1.32 (bs, 1H, H8)

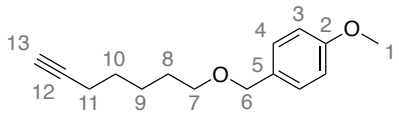
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 84.6 (C2), 68.5 (C1), 63.0 (C7), 32.4 (C6), 28.4 (C5), 25.0 (C4), 18.5 (C3)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 135.1 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_7H_{12}ONa$: 137.0780 $[M+Na]^+$;

Observed: 135.0780 $[M+Na]^+$

**1-(4-methoxybenzyloxy)-hept-6-yne 109**




To the previously synthesized 6-heptyn-1-ol **108** (10.0 g, 0.09 mol) in DMF (50 mL) under argon at 0 °C was added sodium hydride powder (2.35 g, 0.10 mol). The resulting solution was stirred for 30 minutes before 4-methoxybenzyl chloride (12.12 mL, 0.10 mol) was added drop-wise, followed by tetrabutylammonium iodide (3.29 g, 8.90 mmol). The reaction mixture was stirred at 0 °C for 16 hours. The reaction mixture was poured into water and extracted with EtOAc (30 mL). The combined organics were washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude material was purified by flash chromatography

($SiO_2$; 0.7:10; $Et_2O$ in petroleum ether) to give the final product as a colourless oil (8.85 g, 70 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.27 (d, 2H, H3, *J* = 8.5 Hz), 6.89 (d, 2H, H4, *J* = 8.5 Hz), 4.43 (s, 2H, H6), 3.80 (s, 3H, H1), 3.44 (t, 2H, H7, *J* = 6.5 Hz), 2.21-2.18 (td, 2H, H11, $J_1$ = 7.0 Hz, $J_2$ = 2.5 Hz), 1.94 (t, 1H, H13, *J* = 2.5 Hz), 1.67-1.44 (m, 6H, H8, H9 & H10)

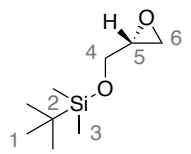
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 159.3 (C2), 130.9 (C5), 129.4 (C4), 113.9 (C3), 84.7 (C13), 72.7 (C6), 70.1 (C7), 68.3 (C12), 55.4 (C1), 29.4 (C8), 28.5 (C9), 25.6 (C10), 18.4 (C11)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 255.1 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{20}O_2Na$: 225.1356 $[M+Na]^+$;

Observed: 225.1357 $[M+Na]^+$

***(R)*-tert-butyldimehtyl(oxiran-2-ylmethoxy)silane 106**




To *(S)*-glycidol **105** (2.0 g, 27.0 mmol), 4-dimethylamino pyridine (4.95 g, 40.5 mmol), and triethylamine (7.51 mL, 53.9 mmol) in $CH_2Cl_2$ (50 ml) at 0 °C, under argon, was added *tert*-butyldimethylsilyl chloride (6.10 g, 40.52 mmol) drop-wise over 1 hour. The reaction mixture was then stirred at RT for 16 hours. The solution was then diluted with $CH_2Cl_2$ and the combined organics were washed with water, HCl (2 M), and sat. $NaHCO_3$, dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was then purified by flash chromatography ($SiO_2$; 1:20; EtOAc : Petroleum ether) to give the product as a colourless oil (3.58 g, 70 %).

The same procedure was repeated using *(R)*-glycidol starting material (2.0 g, 27.0 mmol) to generate the *(S)*-enantiomer product **106** (4.16 g, 81 %). Spectroscopic data identical to literature compounds.[49]

$^1$H-NMR (400 MHz, $CDCl_3$) δ: 3.87 (dd, 1H, $H4_a$, *J* = 12.0 Hz, *J* = 3.0 Hz), 3.68 (dd, 1H, $H4_b$, *J* = 12.0 Hz, *J* = 5.0 Hz), 3.09-3.07 (m, 1H, H5), 2.78-2.76 (m, 1H, $H6_a$), 2.64-2.63 (m, 1H, $H6_b$), 0.90 (s, 9H, H1), 0.08 (s, 6H, H3)

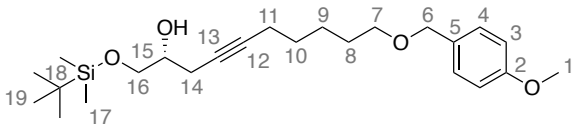
$^{13}$C-NMR (100 MHz, $CDCl_3$) δ: 63.9 (C4), 52.65 (C5), 44.6 (C6), 32.4 (C2), 26.0 (C1), -5.22 (C3)

IR ($V_{max}/cm^{-1}$): 1136 (C-O stretching)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 399.20 (100, $[2M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_9H_{20}O_2SiNa$: 211.1125 $[M+Na]^+$; observed 211.1125 $[M+Na]^+$

**(*R*)-1-((*tert*-butyldimethylsilyl)oxy)-10-((4-methoxybenzyl)oxy)dec-4-yn-2-ol 110**




To a solution of the previously made 1-(4-methoxybenzyloxy)-hept-6-yne **109** (4.26 g, 18.0 mmol) in dry THF (30 mL), under argon, at -78 °C was added *n*-butyl lithium (12.2 mL, 18.0 mmol) and the resulting solution stirred for 20 minutes. Boron trifluoride diethyl etherate (2.41 mL, 20.0 mmol) was added and the solution stirred for a further 20 minutes. *(R)*-tert-butyldimehtyl(oxiran-2-ylmethoxy)silane **106** (2.30 g, 12.0 mmol) in dry THF (20 mL) was then added and the reaction mixture left stirring for 4 hours. The solution was quenched using saturated $NH_4Cl$ and the aqueous was extracted using EtOAc (3 × 100 mL). The combined organics were dried over $MgSO_4$ and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (4.52 g, 94 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.27 (d, 2H, H4, *J* = 8.5 Hz), 6.88 (d, 2H, H3, *J* = 8.5 Hz), 4.43 (s, 2H, H6), 3.80 (s, 3H, H1), 3.74 (m, 1H, H15), 3.70-3.68 (dd, 1H, $H16_a$, *J* = 10.0 Hz, *J* = 4.0 Hz), 3.61 (dd, 1H, $H16_b$, *J* = 10.0 Hz, *J* = 6.0 Hz), 3.43 (t, 2H, H7, *J* = 6.5 Hz), 2.40-2.36 (m, 1H, $H14_a$), 2.20-2.19 (m, 1H, $H14_b$), 2.17-2.14 (m, 2H, H11), 1.64-1.41 (m, 6H, H8, H9 & H10), 0.90 (s, 9H, H19), 0.08 (s, 6H, H17)

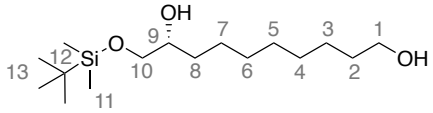
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 159.3 (C2), 130.9 (C5), 129.4 (C4), 113.9 (C3), 82.6 (C12), 76.0 (C13), 72.7 (C6), 70.6 (C15), 70.1 (C7), 65.8 (C16), 55.4 (C1), 29.4 (C10), 29.0 (C9), 28.5 (C8), 26.0 (C19), 23.6 (C14), 18.9 (C18), 18.5 (C11), -5.3 (C17)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 443.3 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{24}H_{40}O_4SiNa$: 443.2588 $[M+Na]^+$; Observed: 443.2591 $[M+Na]^+$

The same procedure was followed using *(S)*-tert-butyldimehtyl(oxiran-2-ylmethoxy)silane **106** (2.32 g, 12.0 mmol) to generate (*S*)-1-((*tert*-butyldimethylsilyl)oxy)-10-((4-methoxybenzyl)oxy)dec-4-yn-2-ol **110** final product (3.0 g, 60 %) with identical spectroscopic data.

***(R)*-10-((tert-butyldimethylsilyl)oxy)decane-1,9-diol 111**




To a solution of *(R)*-((tert-butyldimethylsilyl)oxy)-10-((4-methoxybenzyl)oxy)dec-4-yn-2-ol **110** (4.52 g, 10.0 mmol) in THF (10 ml) under argon at RT, was added 10% palladium on activated carbon (343 mg). The vessel was then flushed with hydrogen gas and left to stir at RT for 48 hours under a hydrogen atmosphere of 50 bar. The reaction mixture was then filtered through a pad of celite and concentrated *in vacuo*. The crude material was then purified by flash chromatography ($SiO_2$; 1:1; EtOAc : Petroleum ether) to give the product as a colourless oil (1.82 g, 60 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 3.65-3.60 (m, 4H, H1, H9 & $H10_a$), 3.38 (dd, 1H, $H10_b$, $J$ = 10.5 Hz, $J$ = 8.5 Hz), 1.80 (bs, 1H, C-OH), 1.59-1.53 (m, 2H, H2), 1.45-1.31 (m, 12H, H3-H8), 0.90 (s, 9H, H13), 0.07 (s, 6H, H11)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 72.0 (C9), 67.4 (C10), 63.2 (C1), 32.9 (C8), 32.9 (C2), 29.8, 29.6, 29.5, 26.0 (C13), 25.8, 25.7, 18.4 (C12), -5.2 (C11)

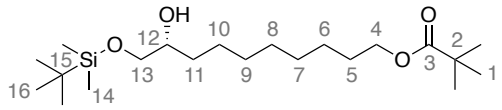
IR ($V_{max}$/cm$^{-1}$): 3333 (O-H stretch), 1252 (C-O stretch)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 327.20 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{36}O_3SiNa$: 327.2326 $[M+Na]^+$; Observed: 327.2327 $[M+Na]^+$

The same procedure was repeated using *(S)*-((tert-butyldimethylsilyl)oxy)-10-((4-methoxybenzyl)oxy)dec-4-yn-2-ol **110** (3.0 g, 7.14 mmol) to generate the *(S)*-10-((tert-butyldimethylsilyl)oxy)decane-1,9-diol **111** (810 mg, 37 %) product with identical spectroscopic data.

***(R)*-10-((tert-butyldimethylsilyl)oxy)-9-hydroxydecyl pivalate 112**




To a solution of *(R)*-10-((tert-butyldimethylsilyl)oxy)decane-1,9-diol **111** (1.82 g, 5.98 mmol) in pyridine (15 mL) under argon at -20 °C was added trimethylacetyl chloride (0.96 mL, 7.77 mmol) in $CH_2Cl_2$ (4.0 ml) and left to stir for 4 hours. The reaction mixture was poured into water and extracted using $CH_2Cl_2$. The combined organics were washed with water (10 mL), HCl (2 M, 5 mL), saturated $NaHCO_3$ (10 mL), dried over $MgSO_4$ and concentrated

*in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (1.31 g, 56 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 4.04 (t, 2H, H4, *J* = 6.5 Hz), 3.63-3.60 (m, 2H, H12 & $13_a$), 3.38 (dd, 1H, $H13_b$, *J* = 10.5 Hz, *J* = 8.5 Hz), 2.40 (bs, 1H, C-OH), 1.64-1.58 (m, 2H, H5), 1.45-1.30 (m, 12H, H6-H11), 1.19 (s, 9H, H1), 0.90 (s, 9H, H16), 0.07 (s, 6H, H14).

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 72.0 (C12), 67.4 (C13), 64.6 (C4), 38.9 (C2), 30.6, 29.8, 29.6, 29.2, 28.7 (C5), 27.4 (C1), 26.5, 26.0 (C16), 25.7, 18.4 (C15), -5.2 (C14).

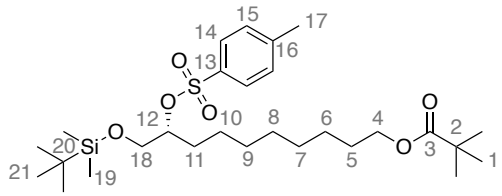
IR ($V_{max}$/$cm^{-1}$): 3503 (O-H stretch), 1253 (C-O stretch)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 411.30 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{21}H_{44}O_4SiNa$: 411.2901 $[M+Na]^+$; Observed: 411.2901 $[M+Na]^+$

The same procedure was repeated using *(S)*-10-((tert-butyldimethylsilyl)oxy)decane-1,9-diol (810 mg, 2.66 mmol) to generate the *(S)*-enantiomer product **112** (940 mg, 91 %) with identical spectroscopic data.

***(R)*-10-((tert-butyldimethylsilyl)oxy)-9-(tosyloxy)decyl pivalate 113**




To a stirred solution of *(R)*-10-((tert-butyldimethylsilyl)oxy)-9-hydroxydecyl pivalate **112** (1.31 g, 3.37 mmol), 4-dimethylamino pyridine (113.0 mg, 1.01 mmol), triethylamine (1.17 mL, 8.43 mmol) in chloroform (20 ml) under argon, at RT, was added *p*-toluenesulphonyl chloride (836 mg, 4.38 mmol). The reaction mixture was heated to 60 °C for 20 hours. The resulting solution was poured into water and extracted using $CH_2Cl_2$. The combined organics were washed with water, HCl (2 M), saturated $NaHCO_3$ (10 mL), dried over $MgSO_4$ and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (1.39 g, 76 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.80 (d, 2H, H14, *J* = 8.5 Hz), 7.33 (d, 2H, H15, *J* = 8.0 Hz), 4.92-4.45 (m, 1H, H12), 4.03 (t, 2H, H4, *J* = 6.5 Hz), 3.67-3.59 (m, 2H, H18), 2.44 (s, 3H, H17), 1.68-1.60 (m, 4H, H5 & H11), 1.34-1.21 (m, 10H, H6-H10), 1.19 (s, 9H, H1), 0.84 (s, 9H, H21), -0.01 (s, 6H, H19)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 144.6 (C16), 134.6 (C13), 129.8 (C15), 128.0 (C14), 83.4 (C12), 64.6 (C4), 64.1 (C18), 38.9 (C2), 31.2 (C11), 29.4 (C5), 29.3, 29.2, 28.6, 27.4 (C1), 26.0, 25.9 (C21), 24.7, 21.7 (C17), 18.4, -5.4 (C19)

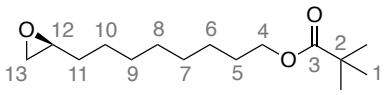
IR ($V_{max}/cm^{-1}$): 1726 (C=O stretching), 1154 (C-O stretching), 1097 (S=O stretching)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 565.30 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* $C_{28}H_{50}O_6SSiNa$: 565.2990 $[M+Na]^+$; observed 565.2988 $[M+Na]^+$

The same procedure was repeated using *(S)*-10-((tert-butyldimethylsilyl)oxy)-9-hydroxydecyl pivalate **112** (940 mg, 2.42 mmol) to generate the *(S)*-enantiomer product **113** (900 mg, 69 %) with identical spectroscopic data.

***(S)*-8-(oxiran-2-yl)octyl pivalate 114**




To a stirring solution of *(R)*-10-((tert-butyldimethylsilyl)oxy)-9-(tosyloxy)decyl pivalate **113** (1.39 g, 2.56 mmol) in THF (10 mL) under argon at RT was added tetra-n-butylammonium fluoride (3.08 mL, 3.08 mmol). The resulting solution was stirred at RT for 3 hours. The reaction mixture was concentrated *in vacuo* and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (630 mg, 96 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 4.04 (t, 2H, H4, $J$ = 6.5 Hz), 2.92-2.88 (m, 1H, H12), 2.74 (m, 1H, $H13_a$), 2.47 (dd, 1H, $H13_b$, $J$ = 5.0 Hz, $J$ = 2.5 Hz), 1.64-1.58 (m, 2H, H5), 1.55-1.32 (m, 12H, H6-H11), 1.19 (s, 9H, H1)

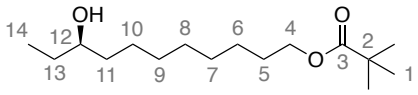
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 64.6 (C4), 52.5 (C12), 47.3 (C13), 38.9 (C2), 32.6, 29.6, 29.5, 29.3, (C7-C10), 28.73 (C5), 27.4 (C1), 26.1, 26.0

LR-MS (ESI, +ve, MeOH) *m/z* (%): 279.20 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{28}O_3Na$: 279.1931 $[M+Na]^+$; observed 279.1934 $[M+Na]^+$

The same procedure was repeated using *(S)*-10-((tert-butyldimethylsilyl)oxy)-9-(tosyloxy)decyl pivalate **113** (900 mg, 1.66 mmol) to generate the *(R)*-enantiomer product **114** (425 mg, 100 %) with identical spectroscopic data.

***(R)*-9-hydroxyundecyl pivalate 115**




To copper cyanide (440 mg, 4.92 mmol) in a dry flask purged with argon at RT was added THF (5 ml). The solution was cooled to -78 °C and methyl lithium (6.15 mL, 9.84 mmol) was added drop-wise and stirred slowly. The reaction mixture was allowed to warm to 0 °C until complete dissolution to a homogenous solution. The reaction was re-cooled to -78 °C and *(S)*-8-(oxiran-2-yl)octyl pivalate **114** (630 mg, 2.46 mmol) in dry THF (2 mL) was added drop-wise followed by $BF_3OEt_2$ (0.36 mL, 2.95 mmol) at -78 °C for 30 minutes left to warm to RT for 16 hours. Water was then added to the reaction mixture and the aqueous extracted using EtOAc (3 × 50 ml), dried over $MgSO_4$ and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (500 mg, 75 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 4.04 (t, 2H, H4, $J$ = 6.5 Hz), 3.54-3.49 (m, 1H, H12), 1.64-1.58 (m, 2H, H5), 1.55-1.30 (m, 14H, H6-H11 & H13), 1.19 (s, 9H, H1), 0.94 (t, 3H, H14, $J$ = 7.4 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 73.5 (C12), 64.6 (C4), 38.9 (C2), 37.1, 30.3, 29.7, 29.6, 29.3, 28.7, 27.4 (C1), 26.0, 25.8, 10.0 (C14)

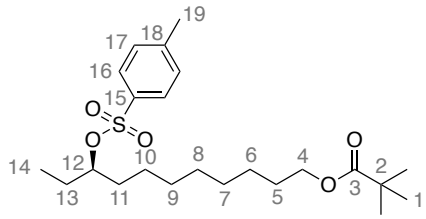
IR ($V_{max}$/cm$^{-1}$): 2959 (O-H stretching), 1728 (C=O stretching), 1459 (C-H bending), 1155 (C-O stretching)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 295.20 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated for $C_{16}H_{32}O_3Na$: 295.2244 $[M+Na]^+$; observed 295.2244 $[M+Na]^+$

The same procedure was repeated using *(R)*-8-(oxiran-2-yl)octyl pivalate **114** (425 mg, 1.66 mmol) to generate the *(S)*-enantiomer product **115** (360 mg, 80 %) with identical spectroscopic data.

***(R)*-9-(tosyloxy)undecyl pivalate 116**




To a solution of *(R)*-9-hydroxyundecyl pivalate **115** (470 mg, 1.73 mmol) in pyridine (30 mL) under argon at RT was added triethylamine (0.31 mL, 2.34 mmol). The reaction mixture was cooled to 0 °C and *p*-toluenesulphonyl chloride (3.30 g, 0.02 mol) was added in small portions. The resulting solution was left to stir at RT overnight. The reaction mixture was diluted with EtOAc and the organics washed with water (10 mL), HCl (2 M), saturated $NaHCO_3$ (10 mL), dried over $MgSO_4$ and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (310 mg, 42 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.81 (d, 2H, H16, *J* = 8.5 Hz), 7.34 (d, 2H, H17, *J* = 8.5 Hz), 4.55-4.47 (m, 1H, H12), 4.04 (t, 2H, H4, 6.6 Hz), 2.44 (s, 3H, H19), 1.64-1.57 (m, 6H, H5, H11 & H13), 1.41-1.22 (m, 10H, H6-H10), 1.19 (s, 9H, H1), 0.82 (t, 3H, H14, *J* = 7.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 144.5 (C18), 134.9 (C15), 129.8 (C17), 127.8 (C16), 85.7 (C12), 64.6 (C4), 38.9 (C2), 33.7, 29.4, 29.4, 29.2, 29.0, 27.4, (C5-C10), 27.2 (C1), 26.0, 24.8, 21.8 (C19), 14.2 (C14)

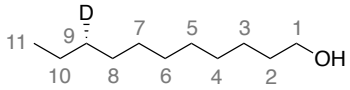
IR ($V_{max}$/cm$^{-1}$): 1724 (C=O stretching), 1460 (C-H bending), 1361 (S=O stretching), 1284 (C-O stretching)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 449.20 (100, M+Na]$^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{23}H_{38}O_5SNa$: 449.2332 [M+Na]$^+$; observed 449.2332 [M+Na]$^+$

The procedure was repeated using *(S)*-9-(tosyloxy)undecyl pivalate **115** (330 mg, 1.21 mmol) to generate the *(S)*-enantiomer product **116** (300 mg, 58 %) with identical spectroscopic data.

***(S)*-undecan-9-d-1-ol 117**




To *(R)*-9-(tosyloxy)undecyl pivalate **116** (510 mg, 1.20 mmol) in THF (10 mL) at 0 °C under argon was added lithium aluminium deuteride (504 mg, 0.01 mol). The reaction mixture was left to stir at RT for 24 hours. The solution was cooled to 0 °C and quenched with water. HCl (2 M) was added and the aqueous layer was extracted using EtOAc (3 × 20 mL), dried over $MgSO_4$ and concentrated *in vacuo* to give the product as a waxy oil (180 mg, 87 %).

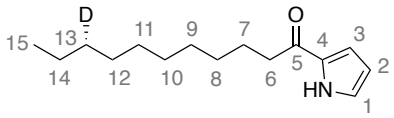
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 3.64 (t, 2H, H1, *J* = 7.0 Hz), 1.58-1.56 (m, 2H, H2), 1.36-1.26 (m, 15H, H3-H10), 0.88 (t, 3H, H11, *J* = 7.0 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 63.3 (C1), 33.0 (C2), 31.6 (t, C9, *J* = 18.5 Hz), 29.9, 29.8, 29.7, 29.6, 29.4, 25.9, 22.7, (C3-C8, C10) 14.2 (C11)

IR ($V_{max}$/cm$^{-1}$): 3030 (O-H stretching), 765 (C-H bending)

The same procedure was repeated using *(S)*-9-(tosyloxy)undecyl pivalate **116** (300 mg, 0.70 mmol) to generate the *(R)*-enantiomer product **117** (40 mg, 33 %) with identical spectroscopic data.

***(S)*-1-(1H-pyrrol-2-yl)undecan-1-one-9-d 119**




Oxidation of alcohol *(S)*-undecan-9-d-1-ol **117** (150 mg, 0.87 mmol) was undertaken following the general procedure for Jones' oxidation of alcohols to generate a crude product (135 mg) that was carried forward without need for further purification.

The crude product (135 mg, 0.72 mmol) was the subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole to generate the final product as a thick oil (30 mg, 15 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.27 (bs, 1H, N-H), 7.01-7.00 (m, 1H, H1), 6.90 (m, 1H, H3), 6.28-6.27 (m, 1H, H2), 2.75 (t, 2H, H6, *J* = 7.5 Hz), 1.74-1.68 (m, 2H, H7), 1.35-1.25 (m, 13H, H8-H14), 0.87 (t, 3H, H15, *J* = 7.0 Hz)

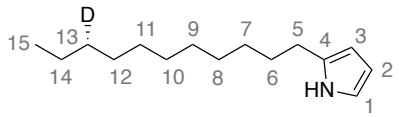
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.3 (C5), 132.3 (C4), 124.2 (C1), 115.9 (C3), 110.7 (C2), 38.2 (C6), 31.6 (t, C13, *J* = 19.0 Hz), 29.7, 29.7, 29.6, 29.6, 29.1, 25.4, 22.7, (C7-C12, C14) 14.2 (C15)

IR ($V_{max}$/cm$^{-1}$): 3281 (N-H stretching), 1641 (N-H bending), 735 (C-H bending)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{24}DNONa$: 259.1891 $[M+Na]^+$; observed; 259.1891 $[M+Na]^+$

The same two-step procedure was repeated using *(R)*-undecan-9-d-1-ol **117** (30 mg, 0.17 mmol) to generate the *(R)*-enantiomer product **119** (12 mg, 22 %) as a colourless oil with identical spectroscopic data.

**(*S*)-2-(undecyl-9-*d*)-1*H*-pyrrole 120**




*(S)*-1-(1H-pyrrol-2-yl)undecan-1-one-9-d **119** (30 mg, 0.13 mmol) was subjected to the general reaction conditions for the reduction of acyl ketones to generate the final product as a white solid (12 mg, 43 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.90 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13-6.12 (m, 1H, H2), 5.91 (m, 1H, H3), 2.59 (t, 2H, H5, *J* = 7.5 Hz), 1.64-1.59 (m, 2H, H6), 1.34-1.21 (m, 15H, H7-H14), 0.88 (t, 3H, H15, *J* = 7.0 Hz)

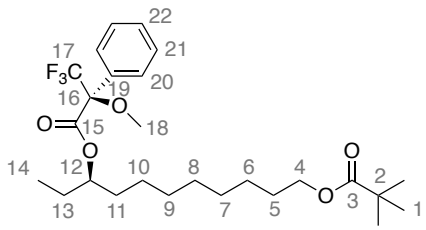
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.1 (C4), 116.1 (C1), 108.4 (C2), 105.0 (C3), 31.6 (t, C13, *J* = 21.0 Hz), 29.8, 29.8, 29.8, 29.7, 29.6, 29.6, 29.4, 27.9, 22.7, (C5-C12, C14) 14.2 (C15)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{27}DN$: 223.2279 $[M+H]^+$; observed; 223.2275 $[M+H]^+$

The same procedure was repeated using the *(R)*-1-(1H-pyrrol-2-yl)undecan-1-one-9-d **110** starting material to generate the *(R)*-2-(undecyl-9-*d*)-1*H*-pyrrole **120** final product (1 mg, 9 %) as a white solid with identical spectroscopic data.

### 6.3.3 Synthesis of Mosher's ester derivatives of alcohols *R* & *S*-115

**(*R*)-11-(pivaloyloxy)undecan-3-yl-(*S*)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoate (*R,S*) 124**




*(R)*-9-hydroxyundecyl pivalate **115** (5.0 mg, 0.02 mmol) and (*S*)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoic acid (5.0 mg, 0.02 mmol) were subjected to the general procedure for the coupling of alcohols to acids. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (3.0 mg, 33 %).

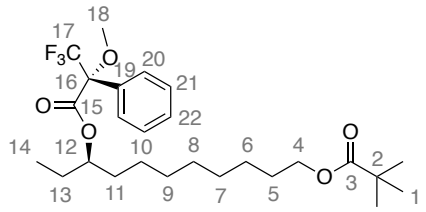
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.56-7.39 (m, 5H, H20, H21 & H22), 5.06–5.01 (m, 1H, H12), 4.04 (t, 2H, H4, *J* = 6.5 Hz), 3.57 (s, 3H, H18), 1.69-1.58 (m, 6H, H5, H11 & H13), 1.33-1.20 (m, 19H, H1 & H6-H10), 0.92 (t, 3H, H14, *J* = 7.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 166.5 (C15), 132.7 (C19), 129.6, 128.5, 127.5, 78.9 (C12), 64.6 (C4), 55.6 (C18), 33.1 (C2), 29.9, 29.4, 29.4, 29.2, 28.7, 27.4 (C1), 26.8, 26.0, 25.0, 9.7 (C14) (2 signals missing)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 511.20 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated m/z $C_{26}H_{39}F_3O_5Na$: 511.2642 $[M+Na]^+$; observed 511.2654 $[M+Na]^+$

**(*R*)-11-(pivaloyloxy)undecan-3-yl-(*R*)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoate (*R,R*) 124**




*(R)*-9-hydroxyundecyl pivalate **115** (5.0 mg, 0.018 mmol) and (*R*)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoic (5.0 mg, 0.02 mmol) were subjected to the general procedure for the coupling of alcohols to acids. The crude material was then purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (4.0 mg, 45 %).

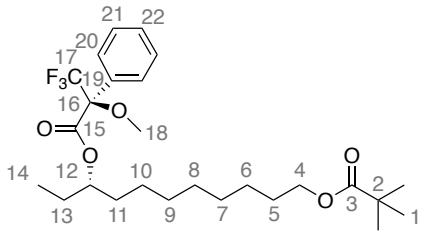
$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 7.56-7.39 (m, 5H, H20, H21 & H22), 5.06-5.01 (m, 1H, H12), 4.04 (t, 2H, H4, *J* = 6.5 Hz), 3.55 (s, 3H, H18), 1.64-1.58 (m, 6H, H5, H11 & H13), 1.28-1.19 (m, 19H, H1 & H6-H10), 0.81 (t, 3H, H14, *J* = 7.5 Hz).

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.9 (C3), 168.3 C15), 132.6 (C19), 129.6, 128.5, 127.5, 78.9 (C12), 64.6 (C4), 55.5 (C18), 33.3, 29.9, 29.5, 29.3, 28.7, 27.4 (C1), 26.5, 26.0, 25.3, 9.4 (C14) (3 signals missing)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 511.20 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{26}H_{39}F_3O_5Na$: 511.2642 $[M+Na]^+$; observed 511.2647

**(*S*)-11-(pivaloyloxy)undecan-3-yl-(*S*)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoate (*S,S*) 124**




*(S)*-9-hydroxyundecyl pivalate **115** (5.0 mg, 0.02 mmol) and (*R*)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoic (5.0 mg, 0.02 mmol) were subjected to the general procedure for the coupling of alcohols to acids. The crude material was then purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (4.0 m g, 45 %).

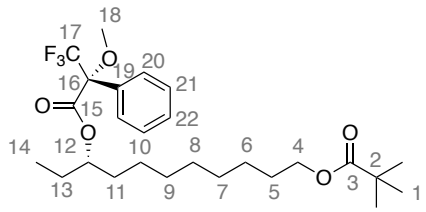
$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 7.56-7.39 (m, 5H, H20, H21 & H22), 5.06-5.01 (m, 1H, H12), 4.04 (t, 2H, H4, *J* = 6.5 Hz), 3.55 (s, 3H, H18), 1.63-1.58 (m, 6H, H5, H11 & H13), 1.28-1.19 (m, 19H, H1 & H6-H10), 0.81 (t, 3H, H14, *J* = 7.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 166.5 (C15), 132.6 (C19), 129.7, 128.5, 127.5, 78.9 (C12), 64.6 (C4), 55.5 (C18), 33.3, 29.9, 29.5, 29.3, 28.7, 27.4 (C1), 26.5, 26.0, 25.3, 9.4 (C14) (2 signals missing)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 511.20 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{26}H_{40}F_3O_5$: 489.2822 $[M+H]^+$; observed 489.2824 $[M+H]^+$

***(S)*-11-(pivaloyloxy)undecan-3-yl-*(R)*-3,3,3-trifluoro-2-methoxy-2-phenylpropanoate *(S,R)* 124**



*(S)*-9-hydroxyundecyl pivalate **115** (5.0 mg, 0.02 mmol) and (*R*)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoic (5.0 mg, 0.02 mmol) were subjected to the general procedure for the coupling of alcohols to acids. The crude material was then purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the product as a colourless oil (3.0 mg, 33 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.56-7.38 (m, 5H, H20, H21 & H22), 5.06-5.01 (m, 1H, H12), 4.04 (t, 2H, H4, $J$ = 6.5 Hz), 3.56 (s, 3H, H18), 1.69-1.58 (m, 6H, H5, H11 & H13), 1.30-1.20 (m, 19H, H1 & H6-H10), 0.92 (t, 3H, H14, $J$ = 7.5 Hz)

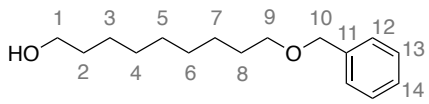
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 178.8 (C3), 166.5 (C15), 132.7 (C19), 129.6, 128.5, 127.5, 78.9 (C12), 64.6 (C4), 55.6 (C18), 33.1, 29.9, 29.4, 29.4, 29.2, 28.7, 27.4 (C1), 26.8, 26.0, 25.0, 9.7 (3 signals missing)

LR-MS (ESI, +ve, MeOH) *m/z* (%): 511.20 (100, $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{26}H_{40}F_3O_5$: 489.2822 $[M+H]^+$; observed 489.2823 $[M+H]^+$

### 6.3.4 Synthesis of mechanistic probes

**9-(benzyloxy)nonan-1-ol 213**



To 1,9-nonanediol **212** (4.0 g, 25 mmol) in $CH_2Cl_2$ (50 mL) was added silver (I) oxide (8.68 g, 37 mmol) and benzyl bromide (3.26 mL, 27.0 mmol) under an argon atmosphere. The reaction mixture was heated to reflux and left to stir for 16 hours. After cooling to RT, the solution was filtered through a pad of celite and concentrated *in vacuo.* The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (4.58 g, 73 %). Spectroscopic data identical to literature compound.[134]

$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 7.38-7.22 (m, 5H, H12-H14), 4.50 (s, 2H, H10), 3.65-3.62 (m, 2H, H1), 3.46 (t, 2H, H9, *J* = 6.5 Hz), 1.68-1.54 (m, 4H, H2 & H8), 1.42-1.22 (m, 10H, H3-H7)

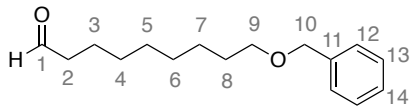
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 138.9 (C11), 128.5 (C13), 127.8 (C12), 127.6 (C14), 73.0 (C10), 70.6 (C9), 63.2 (C1), 32.9 (C8), 29.9 (C2), 29.7, 29.5, 29.5, 26.3, 25.9 (C3-C7)

IR ($V_{max}$/$cm^{-1}$): 3364 (O-H stretching), 2923 (C-H stretching)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 273.3 $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{26}O_2Na$: 273.1825 $[M+Na]^+$; observed 273.1820 $[M+Na]^+$

**9-(benzyloxy)nonanal 214**



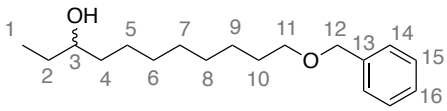

To 9-(benzyloxy)nonan-1-ol **213** (2.0 g, 7.99 mmol) in DMSO (6 mL) at RT was added neat IBX (2.46 g, 8.79 mmol). The reaction mixture was left to stir for 16 hours. EtOAc and water was added and the solution filtered through a pad of celite. The organic layer was extracted using EtOAc (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo.* The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (1.65 g, 83 %). Spectroscopic data identical to literature compound.[143]

$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 9.76 (s, 1H, H1), 7.36-7.26 (m, 5H, H12-H14), 4.50 (s, 2H, H10), 3.46 (t, 2H, H9, *J* = 7.0 Hz), 2.41 (dt, 2H, H2, *J* = 1.5 Hz), 1.64-1.53 (m, 4H, H3 & H8), 1.39-1.25 (m, 8H, H4-H7)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 203.1 (C1), 138.8 (C11), 128.5 (C13), 127.8 (C12), 127.6 (C14), 73.0 (C10), 70.6 (C9), 44.0 (C2), 29.9 (C8), 29.4, 29.4, 29.2, 26.3 (C4-C7), 22.2 (C3)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{24}O_2Na$: 271.1669 $[M+Na]^+$; observed 271.1665 $[M+Na]^+$

**11-(benzyloxy)undecan-3-ol 215**




To 9-(benzyloxy)nonanal **214** (1.65 g, 6.65 mmol) in dry THF (30 mL) under an argon atmosphere and at -78 °C was added ethylmagnesium bromide (1 M, 9.97 mL, 9.97 mmol)

drop-wise over 5 minutes. The reaction mixture was allowed to warm to RT and left to stir for 16 hours. The solution was cooled to 0 °C before saturated $NH_4Cl$ solution was added. The organics were extracted using EtOAc (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:20; EtOAc : Petroleum ether) to give the final product as a colourless oil (1.46 g, 79 %).

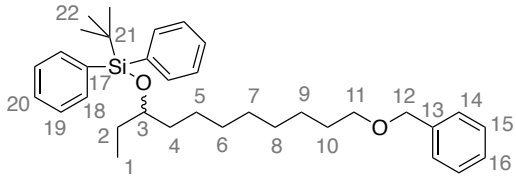
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.36-7.26 (m, 5H, H14-H16), 4.50 (s, 2H, H12), 3.55-3.50 (m, 1H, H3), 3.46 (t, 2H, H11, $J$ = 7.0 Hz), 1.64-1.58 (m, 2H, H10), 1.53-1.28 (m, 14H, H2-H9), 0.94 (t, 3H, H1, $J$ = 7.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 138.9 (C13), 128.5 (C15), 127.8 (C16), 127.6 (C14), 73.5 (C3), 73.0 (C12), 70.1 (C11), 37.1 (C4), 30.3 (C2), 29.9, 29.8, 29.7, 29.6, 26.3, 25.8 (C5-C10), 10.0 (C1)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 301.3 $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{18}H_{30}O_2Na$: 301.2138 $[M+Na]^+$; observed 301.2136

**((11-(benzyloxy)undecan-3-yl)oxy)(*tert*-butyl)diphenylsilane((11-(benzyloxy)undecan-3-yl)oxy)(*tert*-butyl)diphenylsilane 216**




To 11-(benzyloxy)undecan-3-ol **215** (1.46 g, 5.35 mmol) in DMF (20 mL) was added imidazole (786 mg, 12.0 mmol) and *tert*-butyldiphenylsilyl chloride (1.53 mL, 5.77 mmol) and the reaction mixture stirred at RT for 20 hours. Water was added and the organics extracted using EtOAc (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (2.0 g, 77 %)

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.72-7.66 (m, 6H, H18 & H14), 7.42-7.34 (m, 9H, H15, H16, H19 & H20), 4.50 (s, 2H, H12), 3.67-3.64 (m, 1H, H3), 3.45 (t, 2H, H11, $J$ = 7.0 Hz), 1.62-1.55 (m, 2H, H10), 1.48-1.11 (m, 14H, H2-H9), 1.05 (s, 9H, H22), 0.78 (t, 3H, H1, $J$ = 7.5 Hz)

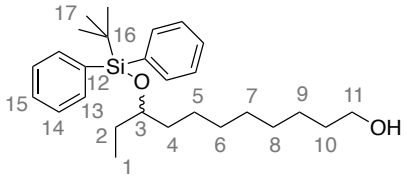
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 138.9 (C13), 136.1 (C18), 135.3 (C17), 135.0 (C14), 129.8 (C20), 128.5 (C19), 127.9 (C15), 127.8 (C16), 74.4 (C3), 73.0 (C12), 70.7 (C11), 35.8, 29.9, 29.8, 29.6, 29.5, 29.1, 26.3, 25.0 (C2 & C4-C10), 27.2 (C22), 19.2 (C21), 9.3 (C1)

IR ($V_{max}$/cm$^{-1}$): 2929 (C-H stretching), 700 (benzene C-H)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 539.4 $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{34}H_{48}O_2SiNa$: 539.3316 $[M+Na]^+$; observed 539.3314

**9-((tert-butyldiphenylsilyl)oxy)undecan-1-ol 217**




To ((11-(benzyloxy)undecan-3-yl)oxy)(*tert*-butyl)diphenylsilane((11-(benzyloxy)undecan-3-yl)oxy)(*tert*-butyl)diphenylsilane **216** (2.0 g, 3.91 mmol) in THF (20 mL) was added palladium on activated carbon (300 mg) and the reaction vessel was flushed with hydrogen gas with vigorous stirring. The reaction mixture was left under a hydrogen atmosphere at atmospheric pressure and at RT for 20 hours. The solution was filtered through a pad of celite and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : PET) to give the final product as a colourless oil (1.22 g, 74 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.68-7.67 (m, 4H, H13), 7.42-7.34 (m, 6H, H14 & H15), 3.68-3.62 (m, 3H, H3 & H11), 1.49-1.11 (m, 16H, H2-H10), 1.05 (s, 9H, H17), 0.78 (t, 3H, H1, $J$ = 7.5 Hz)

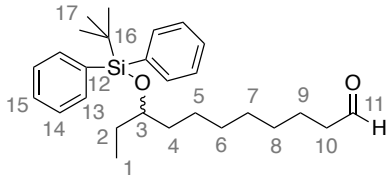
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 136.1 (C13), 135.0 (C12), 129.5 (C15), 127.5 (C14), 74.4 (C3), 63.3 (C11), 35.8, 33.0, 29.7, 29.6, 29.5, 29.1, 25.8, 25.0 (C2 & C4-C10), 27.2 (C17), 19.6 (C16), 9.3 (C1)

IR ($V_{max}/cm^{-1}$): 3352 (O-H stretching), 2927 (C-H stretching)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 449.4) $[M+Na]^+$

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{27}H_{42}O_2SiNa$: 449.2846 $[M+Na]^+$; observed 449.2843 $[M+Na]^+$

**9-((*tert*-butyldiphenylsilyl)oxy)undecanal 218**




To 9-((*tert*-butyldiphenylsilyl)oxy)undecan-1-ol **217** (1.23 g, 2.89 mmol) in DMSO (10 mL) was added neat IBX (889 mg, 3.17 mmol) under argon and at RT. The reaction mixture was left to stir for 16 hours. Water (40 mL) and EtOAc (40 mL) were added and the solution

filtered through a celite pad and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (720 mg, 59 %).

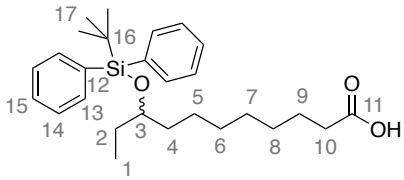
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.75 (s, 1H, H11), 7.68-7.66 (m, 4H, H13), 7.42-7.34 (m, 6H, H14 & H15), 3.68-3.64 (m, 1H, H3), 2.39 (dt, 2H, H10, *J* = 1.5 Hz), 1.61-1.54 (m, 2H, H9), 1.48-1.42 (m, 2H, H2), 1.41-.136 (m, 2H, H4), 1.27-1.211 (m, 8H, H5-H8), 0.78 (t, 3H, H1, *J* = 7.0 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 203.1 (C11), 136.1 (C13), 135.0 (C12), 129.5 (C15), 127.5 (C14), 74.4 (C3), 44.1 (C10), 35.7 (C4), 29.2 (C2), 29.6, 29.4, 29.1, 24.9, (C5-C8), 27.2 (C17), 22.2 (C9), 19.6 (C16), 9.3 (C1)

IR ($V_{max}$/cm$^{-1}$): 2920 (C-H stretching), 1736 (aldehyde C=O stretching)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{27}H_{40}O_2SiNa$: 447.2690 $[M+Na]^+$; observed 447.2689 $[M+Na]^+$

**9-((*tert*-butyldiphenylsilyl)oxy)undecanoic acid 219**




To 9-((*tert*-butyldiphenylsilyl)oxy)undecanal **218** (720 mg, 1.70 mmol) in DMF (10 mL) and under an argon atmosphere was added oxone (1.25 g, 2.04 mmol). The reaction mixture was left to stir for 3 hours at RT. The reaction was poured into water and extracted using EtOAc (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo* to give the final product as a colourless oil (700 mg, 94 %), which did not require further purification.

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.68-7.66 (m, 4H, H13), 7.42-7.34 (m, 6H, H14 & H15), 3.68-3.63 (m, 1H, H3), 2.32 (t, 2H, H10, *J* = 7.5 Hz), 1.62-1.10 (m, 14H, H2-H9), 1.05 (s, 9H, H17), 0.78 (t, 3H, H1, *J* = 7.5 Hz)

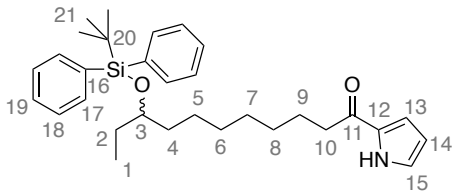
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 176.9 (C11), 136.1 (C13), 135.0 (C12), 129.5 (C15), 127.5 (C14), 74.4 (C3), 35.7, 33.6, 29.6, 29.3, 29.1, 29.1, 24.9, 24.8, (C2 & C4-C9), 27.2 (C17) 19.6 (C16), 9.3 (C1)

IR ($V_{max}$/cm$^{-1}$): 2927 (C-H stretching), 1427 (C-O-H bend), 1709 (C=O stretch)

LR-MS (ESI, +ve, MeOH) *m/*z (%): (100, 463.4) $[M+Na]^+$

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{27}H_{40}O_3SiNa$: 463.2639 $[M+Na]^+$; observed 463.2644 $[M+Na]^+$

**9-((*tert*-butyldiphenylsilyl)oxy)-1-(1*H*-pyrrol-2-yl)undecan-1-one 220**




Compound 9-((*tert*-butyldiphenylsilyl)oxy)undecanoic acid **219** (700 mg, 1.59 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (410 mg, 53 %).

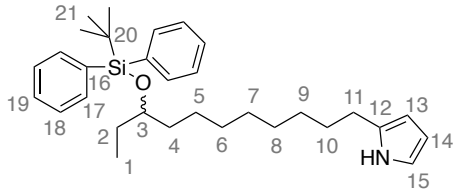
$^1$H-NMR (400 MHz, $CDCl_3$) δ: 9.29 (bs, 1H, N-H), 7.68-7.62 (m, 4H, H17), 7.40-7.35 (m, 6H, H18 & H19), 7.01 (m, 1H, H15), 6.89 (m, 1H, H13), 6.28 (m, 1H, H14), 3.65-3.64 (m, 1H, H3), 2.72 (t, 2H, H10, $J$ = 7.5 Hz), 1.68-1.65 (m, 2H, H9) 1.47-1.11 (m, 12H, H2-H8), 1.05 (s, 9H, H21), 0.78 (t, 3H, H1, $J$ = 7.0 Hz)

$^{13}$C-NMR (100 MHz, $CDCl_3$) δ: 194.7 (C11), 136.1 (C17), 129.5 (C19), 127.5 (C18), 124.2 (C15), 115.9 (C13), 110.7 (C14), 74.4 (C3), 38.2, 35.7, 29.7, 29.5, 29.5, 29.1, 25.0 (C2 & C4-C8), 27.2 (C21), 25.4 (C9), 19.6 (C20), 9.3 (C1) (C12 & C16 missing)

IR ($V_{max}$/cm$^{-1}$): 2929 (C-H stretching), 1637 (C=O stretching)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{31}H_{43}O_2SiNNa$: 512.2955 $[M+Na]^+$; observed 512.2957 $[M+Na]^+$

**2-(9-((*tert*-butyldiphenylsilyl)oxy)undecyl)-1*H*-pyrrole 221**




Compound 9-((*tert*-butyldiphenylsilyl)oxy)-1-(1*H*-pyrrol-2-yl)undecan-1-one **220** (410 mg, 0.84 mmol) was subjected to the general reaction conditions for the reduction of acyl ketones and purified by flash chromatography ($SiO_2$; 1.2:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (300 mg, 75 %).

$^1$H-NMR (400 MHz, $CDCl_3$) δ: 7.90 (bs, 1H, N-H), 7.68-7.67 (m, 4H, H17), 7.42-7.34 (m, 6H, H18 & H19), 6.66 (m, 1H, H15), 6.13 (m, 1H, H13), 5.91 (m, 1H, H14), 3.67-3.65 (m, 1H, H3), 2.58 (t, 2H, H11, $J$ = 7.5 Hz), 1.63-1.55 (m, 2H, H10), 1.47-1.10 (m, 14H, H2-H9), 1.05 (s, 9H, H21), 0.78 (t, 3H, H1, $J$ = 7.5 Hz)

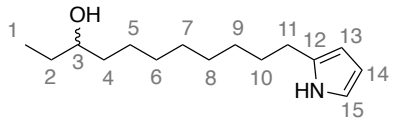
$^{13}$C-NMR (100 MHz, $CDCl_3$) δ: 136.1 (C17), 135.0 (C16), 129.5 (C19), 127.5 (C18), 116.1 (C15), 108.4 (C13), 105.0 (C14), 74.5 (C3), 35.8, 29.8 (C10), 29.6, 29.5, 29.1, 27.9 (C11), 27.2 (C21), 25.0, 9.3 (C1) 4 signals missing.

IR ($V_{max}$/$cm^{-1}$): 2948 (C-H stretching), 1111 (C-O stretching)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 498.5) $[M+Na]^+$

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{31}H_{45}OSiNNa$: 498.3163 $[M+Na]^+$; observed 498.3167 $[M+Na]^+$

**11-(1*H*-pyrrol-2-yl)undecan-3-ol 211**



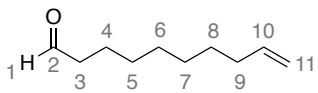

To compound 2-(9-((*tert*-butyldiphenylsilyl)oxy)undecyl)-1*H*-pyrrole **221** (50 mg, 0.11 mmol) in dry THF (5 mL) under an argon atmosphere was added TBAF (1 M, 0.16 mL, 0.16 mmol) and left to stir at RT for 20 hours. The reaction mixture was concentrated *in vacuo* and the crude material was purified by flash chromatography ($SiO_2$; 1:2; EtOAc : Petroleum ether) to give the final product as a waxy solid (5 mg, 20 %).

$^1$H-NMR (400 MHz, $CDCl_3$) δ: 7.91 (bs, 1H, N-H), 6.66 (m, 1H, H15), 6.13-6.12 (m, 1H, H14), 5.91 (m, 1H, H13), 3.52-3.49 (m, 1H, H3), 2.59, (t, 2H, H11, $J$ = 8.0 Hz), 1.63-1.56 (m, 2H, H10), 1.45-1.24 (m, 14H, H2-H9), 0.94 (t, 3H, H1, $J$ = 7.5 Hz)

$^{13}$C-NMR (100 MHz, $CDCl_3$) δ: 133.0 (C12), 116.1 (C15), 108.4 (C14), 105.0 (C13), 73.5 (C3), 37.1, 30.3, 29.9 (C10), 29.8, 29.7, 29.5, 29.5, 27.9 (C11), 25.8, 10.0 (C1)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{26}ON$: 236.2020 $[M-H]^-$; observed 236.2014 $[M-H]^-$

**Dec-9-enal 170**




To 9-decen-1-ol **169** (324 mg, 2.07 mmol), in DMSO (5 mL) under argon and at RT was added neat IBX (68 mg, 2.28 mmol) and left to stir for 20 hours. Water (30 mL) and EtOAc (30 mL) were added and the solution filtered through a celite pad and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (300 mg, 94 %).

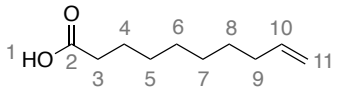
$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 9.76 (s, 1H, H1), 5.84-5.76 (ddt, 1H, H10, *J* = 17.0 Hz, *J* = 10.0 Hz, *J* = 6.5 Hz), 5.01 (dd, 2H, $H11_a$, *J* = 17.0 Hz, *J* = 1.5 Hz), 4.94 (dd, 1H, $H11_b$, *J* = 10.0 Hz, *J* = 1.0 Hz), 2.42 (dt, 2H, H3, *J* = 1.5 Hz), 2.06-2.02 (m, 2H, H9), 1.66-1.60 (m, 2H, H4), 1.39-1.31 (m, 8H, H5-H8)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 203.1 (C2), 139.3 (C10), 114.4 (C11), 44.1 (C3), 33.9 (C9), 29.3, 29.2, 29.0, 29.0, (C5-C8), 22.2 (C4)

IR ($V_{max}$/$cm^{-1}$): 2925 (C-H stretching), 1725 (aldehyde C=O stretching), 1640 (C=C stretching)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{10}H_{17}O$: 153.1285 $[M-H]^-$; observed 153.1285 $[M-H]^-$

**Dec-9-enoic acid 171**




To dec-9-enal **170** (300 mg, 1.95 mmol) in DMF (5 mL) under argon and at RT was added oxone (1.68 g, 2.73 mmol) and left to stir for 1 hour. Water was added and the organics extracted using EtOAc (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (250 mg, 75 %).

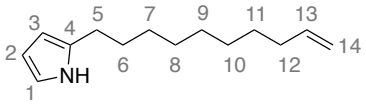
$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 10.56 (bs, 1H, H1), 5.85-5.76 (ddt, 1H, H10, *J* = 17.0 Hz, *J* = 10.0 Hz, *J* = 6.5 Hz), 5.01 (dd, H, $H11_a$, *J* = 17.0 Hz, *J* = 1.0 Hz), 4.94 (dd, 1H, $H11_b$, *J* = 10.0 Hz, *J* = 1.0 Hz), 2.35 (t, 2H, H3, *J* = 7.5 Hz), 2.06-2.02 (m, 2H, H9), 1.66-1.60 (m, 2H, H4), 1.39-1.26 (m, 8H, H5-H8)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 179.0 (C2), 139.1 (C10), 114.2 (C11), 33.8 (C3), 33.8 (C9), 29.1, 29.0, 28.9, 28.8 (C5-C8), 24.7 (C4)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 169.2 $[M-H]^-$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{10}H_{17}O_2$: 169.1234 $[M-H]^-$; observed 169.1231 $[M-H]^-$

**2-(dec-9-en-1-yl)-1*H*-pyrrole 156**



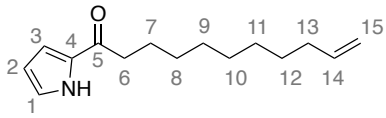

Dec-9-enoic acid **171** (300 mg, 1.47 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material was flushed through a silica pad using a 1:10; EtOAc : Petroleum ether solvent system. The resulting material was subjected to the general reaction conditions for the reduction of acyl ketones and the resulting crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (62 mg, 21 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.89 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13 (m, 1H, H2), 5.91 (m, 1H, H3), 5.81 (ddt, 1H, H13, *J* = 17.0 Hz, *J* = 10.0 Hz, *J* = 6.5 Hz), 5.01 (dd, 1H, $H14_a$, *J* = 17.0 Hz, *J* = 1.5 Hz), 4.94 (dd, 1H, $H14_b$, *J* = 10.0 Hz, *J* = 1.0 Hz), 2.59 (t, 2H, H5, J = 8.0 Hz), 2.06-2.02 (m, 2H, H12), 1.65-1.59 (m, 2H, H6), 1.40-1.26 (m, 10H, H7-H12)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 139.4 (C13), 133.0 (C4), 116.1 (C1), 114.3 (C14), 108.4 (C2), 105.0 (C3), 33.9 (C12), 29.8, 29.5, 29.5, 29.5, 29.3, 29.1 (C6-C11), 27.9 (C5)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* calculated for $C_{14}H_{24}N$: 206.1903 $[M+H]^+$; observed 206.1903 $[M+H]^+$

**1-(1*H*-pyrrol-2-yl)undec-10-en-1-one 168**




Undecylenoic acid **167** (500 mg, 2.71 mmol) was subjected to the general conditions for the coupling of carboxylic acids to pyrrole. The crude material was purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a white solid (230 mg, 31 %).

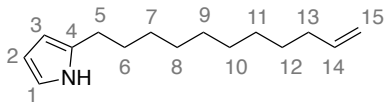
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.36 (bs, 1H, N-H), 7.01 (m, 1H, H1), 6.90 (m, 1H, H3), 6.28 (m, 1H, H2), 5.81 (dtt, 1H, H14, *J* = 17.0 Hz, *J* = 10.0 Hz, *J* = 6.5 Hz), 5.01 (dd, 1H, $H15_a$, *J* = 17.5 Hz, *J* = 1.0 Hz), 4.94 (dd, 1H, $H15_b$, *J* = 10.5 Hz, *J* = 0.5 Hz), 2.75 (t, 2H, H6, *J* = 7.5 Hz), 2.05-2.01 (m, 2H, H13), 1.39-1.29 (m, 12H, H7-H12)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.3 (C5), 139.4 (C14), 132.2 (C4), 124.3 (C1), 116.0, (C3), 114.3 (C15), 110.7 (C2), 38.2 (C6), 33.9, 29.6, 29.5, 29.5, 29.2, 29.0, (C7-C13), 25.4 (C13)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 256.2 $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{23}NONa$: 256.1672 $[M+Na]^+$; observed 256.1672 $[M+Na]^+$

**2-(undec-10-en-1-yl)-1*H*-pyrrole 157**



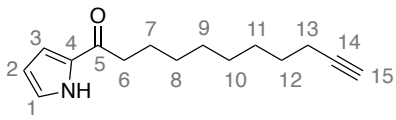

1-(1*H*-pyrrol-2-yl)undec-10-en-1-one **168** (230 mg, 0.99 mmol) was subjected to the general procedure for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a thick brown oil (47 mg, 22 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.89 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13 (m, 1H, H2), 5.91 (m, 1H, H3), 5.82 (dtt, 1H, H14, J = 7.0 Hz, J = 10.2 Hz, J = 16.9 Hz), 5.01 (dd, 1H, $H15_a$ *J* = 17.0 Hz, *J* = 1.5 Hz), 4.94 (dd, 1H, $H15_a$, *J* = 10.0 Hz, *J* = 0.90 Hz), 2.59 (t, 2H, H5, *J* = 7.5 Hz), 2.06-2.02 (m, 2H, H13), 1.65-1.58 (m , 2H, H6), 1.37-1.28 (m, 12H, H7-H12)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 139.4 (C14), 133.0 (C4), 116.1 (C1), 114.3 (C15), 108.4 (C2), 105.0 (C3), 34.0 (C13), 29.8, 29.7 (C6), 29.6, 29.6, 29.5, 29.7, 29.1, 27.9 (C5)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{26}N$: 220.2060 $[M+H]^+$; Observed 220.2056 $[M+H]^+$

**1-(1*H*-pyrrol-2-yl)undec-10-yn-1-one 166**




10-undecynoic acid **165** (500 mg, 2.74 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) giving as a white solid product (290 mg, 46 %).

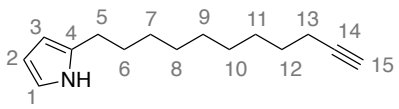
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.01 (m, 1H, H1), 6.90 (m, 1H, H3), 6.28 (m, 1H, H2), 2.75 (t, 2H, H6, *J* = 7.5 Hz), 2.18 (dt, 2H, H13, *J* = 7.0 Hz, *J* = 2.5 Hz), 1.94-1.93 (m, 1H, H15), 1.74-1.68 (m, 2H, H7), 1.54-1.49 (m, 2H, H12), 1.42-1.27 (m, 8H, H8-H11)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.1 (C5), 133.7 (C4), 124.2 (C1), 115.8 (C3), 110.6 (C2), 84.8 (C14), 68.1 (C15), 38.0 (C6), 29.5, 29.4, 29.1, 28.8 (C12), 28.6, 25.4 (C7), 18.5 (C13)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 254.3 $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{21}NONa$: 254.1515 $[M+Na]^+$; observed 254.1515 $[M+Na]^+$

**2-(undec-10-yn-1-yl)-1*H*-pyrrole 155**



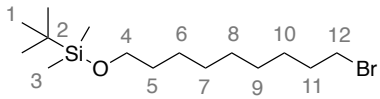

1-(1*H*-pyrrol-2-yl)undec-10-yn-1-one **166** (290 mg, 1.25 mmol) was subjected to the general procedure for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a light yellow oil (189 mg, 65 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.90 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13 (m, 1H, H2), 5.91 (m, 1H, H3), 2.59 (t, 2H, H5), *J* = 8.0 Hz), 2.19 (dt, 2H, H13, *J* = 7.0 Hz, *J* = 2.5 Hz), 1.94 (m, 1H, H15), 1.65-1.59 (m, 2H, H6), 1.54-1.49 (m, 2H, H12), 1.40-1.29 (m, 10H, H7-H11)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.0 (C4), 116.1 (C1), 108.4 (C2), 104.5 (C3), 84.9 (C14), 68.2 (C15), 29.8, 29.5 C6), 29.5, 29.2, 29.1, 28.9, 28.6 (C12), 27.9 (C5), 18.5 (C13)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{15}H_{24}N$: 218.1903 $[M+H]^+$; observed 218.1900 $[M+H]^+$

**((9-bromononyl)oxy)(*tert*-butyl)dimethylsilane 191**



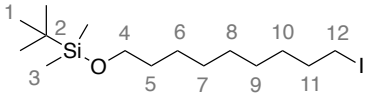

To 9-bromononan-1-ol **190** (2.0 g, 8.96 mmol) in THF (30 mL) at RT was added *tert*-butyldimethylsilyl chloride (1.62g, 11.0 mmol) and imidazole (911 mg, 13.0 mmol) and left to stir for 16 hours. The reaction mixture was diluted with EtOAc and washed with water and brine before drying ($MgSO_4$) and concentration *in vacuo.* The resulting colourless oil did not require further purification (2.92 g, 100 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 3.60 (t, 2H, H4, *J* = 6.5 Hz), 3.41 (t, 2H, H12, *J* = 7.0 Hz), 1.90-1.80 (m, 2H, H5), 1.58-1.38 (m, 4H, H6 & H11), 1.33-1.26 (m, 8H, H7-H10), 0.89 (s, 9H, H1), 0.05 (s, 6H, H3)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 63.4 (C4), 34.2 (C12), 33.0 (C5), 33.0 (C11), 29.6, 29.5, 28.9, 28.3, 26.1 (C1), 25.9, 18.5 (C2), -5.0 (C3)

IR ($V_{max}$/$cm^{-1}$): 2927 (C-H stretch), 1096 (C-O stretching), 565 (C-Br stretching)

**_tert_-butyl((9-iodononyl)oxy)dimethylsilane 192**



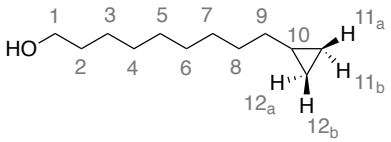

To ((9-bromononyl)oxy)(*tert*-butyl)dimethylsilane **191** (2.92 g, 8.69 mmol) in acetone (30 mL) was added sodium iodide (2.73 g, 18.0 mmol) and heated to reflux for 4 hours. The reaction mixture was cooled and concentrated *in vacuo*. The resulting residue was diluted with EtOAc and washed with water and brine before drying ($MgSO_4$) and concentration *in vacuo.* The resulting yellow oil did not require further purification (3.12 g, 93 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 3.60 (t, 2H, H4, $J$ = 6.5 Hz), 3.19 (t, 2H, H12, $J$ = 7.0 Hz), 1.85-1.79 (m, 2H, H11), 1.51-1.49 (m, 2H, H5), 1.41-1.36 (m, 2H, H6), 1.32-1.29 (m, 8H, H7-H10)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 63.4 (C4), 33.7 (C11), 33.0 (C4), 30.6 (C10), 29.5, 29.5, 28.6, 26.1 (C1), 25.9, 18.5 (C2), 7.5 (C12), -5.1 (C3)

IR ($V_{max}$/cm$^{-1}$): 2926 (C-H stretching), 1096 (C-O stretching), 505 (C-I stretching)

**9-cyclopropylnonan-1-ol 194**




To *tert*-butyl((9-iodononyl)oxy)dimethylsilane **192** (3.12 g, 8.12 mmol) in THF (20 mL) under argon and at 0 °C was added copper chloride (33 mg, 0.24 mmol) and 1-phenyl-1-propyne (152 µL, 1.22 mmol) and stirred for 5 minutes. Cyclopropyl magnesium bromide (0.5 M, 32.48 mL, 16.0 mmol) was added drop-wise over 5 minutes and the reaction mixture stirred at RT for 2 hours. The resulting brown solution was cooled to 0 °C and quenched with saturated $NH_4Cl$ and the organics extracted with $Et_2O$ (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo.* The resulting residue was dissolved in THF (20 mL) and cooled to 0 °C before TBAF (1.0 M, 16.0 mL, 16.0 mmol) was added and the solution left to stir at RT for 3 hours. The reaction mixture was quenched with ice-cold water, extracted with $Et_2O$ (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo.* The crude material was purified by flash chromatography ($SiO_2$; 1:5; EtOAc : Petroleum ether) to give the final product as a colourless oil (1.29 g, 84 %).

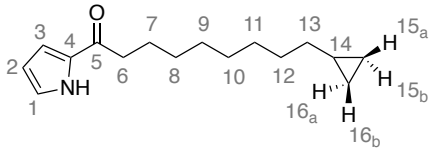
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 3.64 (t, 2H, H2, $J$ = 6.5 Hz), 1.61-1.52 (m, 2H, H2), 1.44-1.26 (m, 12H, H3-H8), 1.26-1.14 (m, 2H, H9), 0.71-0.58 (m, 1H, H10), 0.41-0.35 (m, 2H, H$11_a$ & $12_b$ or H$11_b$ & $12_a$), 0.01-0.04 (m, 2H, H$11_a$ & $12_a$ or H$11_b$ & $12_b$)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 63.3 (C1), 34.9 (C9), 33.0, 29.8, 29.8, 29.8, 29.7, 29.6, 25.9 (C2-C8), 11.1 (C10), 4.5 (C11 & C12)

IR ($V_{max}$/cm$^{-1}$): 3348 (OH stretch), 2922 (C-H stretch),

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{12}H_{25}O$: 185.1900 $[M+H]^+$; observed 185.899 $[M+H]^+$

**9-cyclopropyl-1-(1*H*-pyrrol-2-yl)nonan-1-one 196**




The previously synthesized 9-cyclopropylnonan-1-ol **194** (1.26 g, 6.84 mmol) was subjected to the general reaction conditions for the oxidation of alcohols to carboxylic acids using Jones' reagent generated *in situ*. The resulting material (1.16 g, 5.85 mmol) was then subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material purified by flash chromatography ($SiO_2$; 1:5; EtOAc : Petroleum ether) to give the final product as a brown oil (440 mg, 26 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.30 (bs, 1H, N-H), 7.01 (m, 1H, H1), 6.90 (m, 1H, H3), 6.28 (m, 1H, H2), 2.75 (t, 2H, H6, *J* = 7.5 Hz), 1.74-1.68 (m, 2H, H7), 1.37-1.27 (m, 10H, H8-H12), 1.19-1.15 (m, 2H, H13), 0.68-0.60 (m, 1H, H14), 0.39-0.36 (m, 2H, $H15_a$ & $16_b$ or $H15_b$ & $H16_a$), 0.01-0.03 (m, 2H, $H15_a$ & $16_b$ or $H15_b$ & $H16_a$)

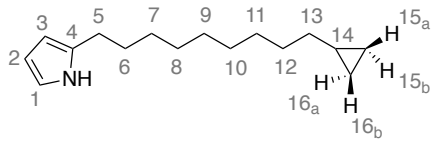
$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.3 (C5), 132.3 (C4), 124.2 (C1), 115.9 (C3), 110.7 (C2), 38.2 (C6), 34.9 (C13), 29.8, 29.7, 29.6, 29.6, 25.4 (C7-C11), 110.0 (C14), 4.5 (C15 & C16) (1 signal missing)

IR ($V_{max}$/cm$^{-1}$): 2923 (C-H stretching), 1638 (C=O stretching)

LR-MS (ESI, +ve, MeOH) m/z (%): (100, 270.3 $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{25}NONa$: 270.1828 $[M+Na]^+$; observed 270.1830 $[M+Na]^+$

**2-(9-cyclopropylnonyl)-1*H*-pyrrole 197**



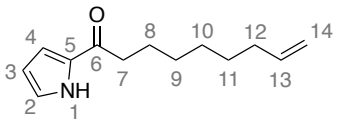

Compound 9-cyclopropyl-1-(1*H*-pyrrol-2-yl)nonan-1-one **196** (40 mg, 0.16 mmol) was subjected to the general reaction conditions for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a brown oil (22 mg, 58 %).

$^1$H-NMR (300 MHz, $CDCl_3$) δ: 7.88 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13 (m, 1H, H2), 5.91 (m, 1H, H3), 2.59 (t, 2H, H5, *J* = 7.5 Hz), 1.67-1.59 (m, 2H, H6), 1.37-1.28 (14H, H6-H12), 1.21-1.14 (m, 2H, H13), 0.71-0.59 (m, 1H, H14), 0.39 (m, 2H, $H15_a$ & $H16_b$ or $H15_b$ & $H16_a$)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.1 (C4), 116.1 (C1), 108.4 (C2), 105.0 (C3), 34.9, 29.8, 29.8, 29.7, 29.7 (C6), 29.6, 29.5, 27.9 (C5), 110.1 (C14), 4.5 (C16)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{26}N$: 232.2071 $[M-H]^-$; observed 232.2072 $[M-H]^-$

**1-(1*H*-pyrrol-2-yl)non-8-en-1-one 201**



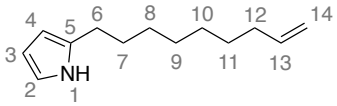

8-nonenoic acid **200** (500 mg, 3.20 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material purified by flash chromatography ($SiO_2$; 1:5; EtOAc : Petroleum ether) to give the final product as a brown oil (410 mg, 62 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.38 (1H, H1, NH), 7.02 (m, 1H, H2), 6.90 (m, 1H, H4), 6.28-6.27 (m, 1H, H3), 5.84-5.76 (m, 1H, H13), 5.01-4.92 (m, 2H, H14), 2.76 (t, 2H, H7, *J* = 7.5 Hz), 2.06-2.01 (m, 2H, H12), 1.75-1.69 (m, 2H, H8), 1.41-1.34 (m, 6H, H9-H11)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.7 (C6), 139.4 (C13), 132.8 (C5), 124.8 (C2), 116.5 (C4), 115.0 (C14), 111.7 (C3), 38.2 (C7), 34.2 (C12), 29.9, 29.5, 29.3 (C9-C11), 26.2 (C8)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{13}H_{19}NO$: 204.1391 $[M-H]^-$; observed 204.1394 $[M-H]^-$

**2-(non-8-en-1-yl)-1*H*-pyrrole 202**



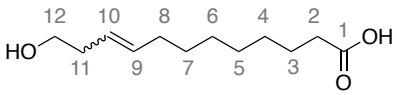

Compound **201** (200 mg, 0.98 mmol) was subjected to the general reaction conditions for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (160 mg, 83 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.90 (1H, H1, NH), 6.66 (m, 1H, H2), 6.13-6.12 (m, 1H, H3), 5.91 (m, 1H, H4), 5.85-5.77 (m, 1H, H13), 5.01-4.92 (m, 2H, H14), 2.59 (t, 2H, H6, *J* = 7.5 Hz), 2.06-2.01 (m, 2H, H12), 1.65-1.59 (m, 2H, H7), 1.38-1.31 (m, 8H, H8-H11)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 139.2 (C13), 132.8 (C5), 116.0 (C2), 114.2 (C14), 108.3 (C3), 104.9 (C4), 33.8 (C12), 29.7, 29.4, 29.3, 29.0, 28.9 (C7-C11), 27.5, (C6)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{13}H_{21}N$: 190.1601 $[M-H]^-$; observed 190.1601 $[M-H]^-$

**12-hydroxydodec-9-enoic acid 205**



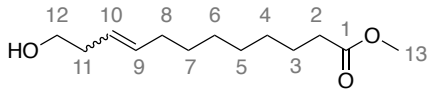

To undecenoic acid **204** (0.92 g, 5.0 mmol) in DCM (25 mL) was added paraformaldehyde (0.30 g, 10.0 mmol) under argon, at -15 °C. Dimethyl aluminum chloride (10 mL, 10.0 mmol) was added drop-wise and the solution was warmed to RT and stirred for 2 hours. The reaction mixture was quenched with diethyl ether (10 mL), water (40 mL) and HCl was added until all visible aluminum salts had dissolved. The organics were extracted using diethyl ether (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo.* The crude material was purified by flash chromatography ($SiO_2$; 9:1; EtOAc : Petroleum ether) to give the final product as a white solid (760 mg, 71 %). Spectroscopic data identical to that reported in the literature for the E-isomer (with the same ratio of geometric isomers (4:1 E:Z)).[133]

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 5.55 (1H, H10, dtt, *J* = 7.0, 15.0, 1.0 Hz), 5.37 (1H, H9, dtt, *J* = 7.0 Hz, 15.0 Hz, 1.0 Hz), 3.63 (t, 2H, H12, *J* = 6.5 Hz), 2.34 (t, 2H, H2, *J* = 7.5 Hz), 2.26 (ddt, 2H, H11, *J* = 6.5 Hz, 6.3 Hz, 1.2 Hz), 2.01 (dt, 2H, H8, *J* = 6.0 Hz, 6.5 Hz), 1.63 (tt, 2H, H3, *J* = 7.5 Hz, *J* = 7.0 Hz), 1.27-1.40 (m, 8H, H4-H7)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 179.1 (C1), 134.4 (C10), 125.9 (C9), 62.2 (C12), 36.1 (C11), 34.0 (C2), 32.7(C8), 29.4, 29.1, 29.1, 29.0 (C4-C7), 24.8 (C3)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{12}H_{22}O_3$: 213.1498 $[M-H]^-$; observed 213.1496 $[M-H]^-$

**Methyl 12-hydroxydodec-9-enoate 206**



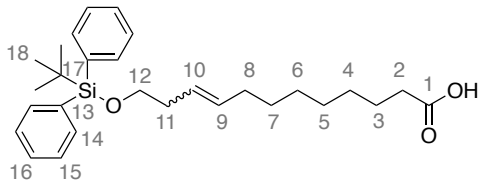

Compound **205** and 6 M HCl (1 mL) was refluxed in methanol (10 mL) for 20 hours. The reaction mixture was concentrated in vacuo and the crude material was purified by flash chromatography ($SiO_2$; 3:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (500 mg, 62 %) in a ratio of 4:1 (E:Z isomers).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 5.58-5.51 (m, 1H, H10), 5.41-5.33 (m, 1H, H9), 3.67 (s, 3H, H13), 3.62 (t, 2H, H12, *J* = 6.0 Hz), 2.3 (t, 2H, H2, *J* = 7.5 Hz), 2.26 (m, 2H, H11), 2.01 (m, 2H, H8), 1.27-1.40 (m, 8H, H4-H7)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 174.5 (C1), 134.4 (C10), 125.9 (C9), 62.2 (C12), 51.6 (C13), 36.1 (C11), 34.2 (C2), 32.7 (C8), 29.7, 29.5, 29.2, 29.0 (C4-C7), 25.0 (C3)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{12}H_{24}O_3$: 251.1618 $[M+Na]^+$; observed 251.1618 $[M+Na]^+$

**12-((*tert*-butyldiphenylsilyl)oxy)dodec-9-enoic acid 207**



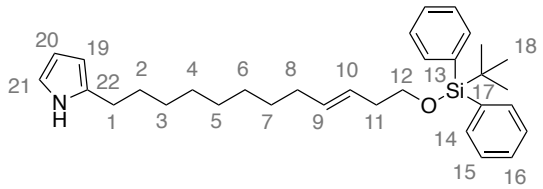

To compound **206**, in DCM (20 mL) was added triethylamine (0.61 mL, 4.33 mmol), DMAP (27.0 mg, 0.22 mmol) and t-buyldiphenylsilyl chloride (0.70 mL, 2.65 mmol) and stirred under argon at RT overnight. Water (30 mL) was added and the organics were extracted using DCM (3 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The crude material was dissolved in $H_2O$:THF (1:4) and LiOH (256 mg) was added and the reaction mixture was stirred at RT for 1 hour. Water (20 mL) was added and the organics were acidified using HCl and extracted using diethyl ether (3 × 20 mL). The combined organics were dried ($MgSO_4$), concentrated *in vacuo* and was purified by flash chromatography ($SiO_2$; 1:2; EtOAc : Petroleum ether) to give the final product (4:1 E:Z) as a colourless oil (900 mg, 90 %).

$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 7.68-7.65 (m, 2H, H16), 7.43-7.36 (m, 8H, H14&15), 5.57-5.36 (m, 2H, H9&10), 3.66 (t, 2H, H12, *J* = 6.5 Hz), 2.34 (t, 2H, H2, *J* = 7.5), 2.24 (m, 2H, H11), 1.96 (m, 2H, H8), 1.62 (m, 2H, H3), 1.36-1.25 (m, 8H, H4-H7), 1.04, (s, 9H, H18)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 177.9, 135.8 (C16), 135.3 (C13), 132.7 (C10), 129.63 (C15), 127.8 (C14), 126.6 (C9), 64.2 (C12), 36.2 (C11), 33.8 (C2), 32.8 (C8), 29.5, 29.2, 29.1, 29.1, 27.0 (C18), 24.8 (C3) (C17 missing)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{28}H_{40}O_3Si$: 475.2642 $[M+Na]^+$; observed 475.2639 $[M+Na]^+$

**(*E*)-12-hydroxy-1-(1*H*-pyrrol-2-yl)dodec-9-en-1-one 208**



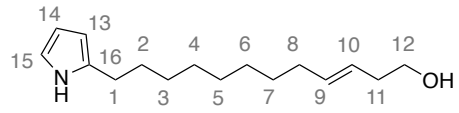

Compound **207** was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material was subsequently subjected to the general reaction conditions for the reduction of acyl ketones. The crude material purified by flash chromatography ($SiO_2$; 3:10; EtOAc : Petroleum ether) to give the final product as a colourless oil (280 mg, 29 %).

$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 7.89 (1H, N-H), 7.67-7.66 (m, 2H, H16), 7.43-7.36 (m, 8H, H14&15), 6.66 (m, 1H, H21), 6.13-6.12 (m, 1H, H19), 5.91 (m, 1H, H20), 5.48-5.35 (m, 2H, H9&10), 3.66 (t, 2H, H12, *J* = 7.0 Hz), 2.58 (t, 2H, H2, *J* = 8.5), 2.24 (m, 2H, H11), 1.96 (m, 2H, H8), 1.61 (m, 2H, H3), 1.38-1.21 (m, 8H, H4-H7), 1.04 (s, 9H, H18)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 190.4 (C1), 135.9 (C16), 134.9 (C13), 133.0 (C22), 132.8 (C10), 129.7 (C15), 127.7 (C14), 126.5 (C9), 116.1 (C21), 108.4 (C19), 104.5 (C20), 64.2 (C12), 36.2 (C11), 32.8 (C8), 29.8, 29.6, 29.6, 29.4, 29.3 (C7-C3), 27.9 (C2), 27.0 (C18) (C17 missing)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{25}NO_2$: 510.3163 $[M+H]^+$; observed 510.3163 $[M+H]^+$

**(*E*)-12-(1*H*-pyrrol-2-yl)dodec-3-en-1-ol 209**



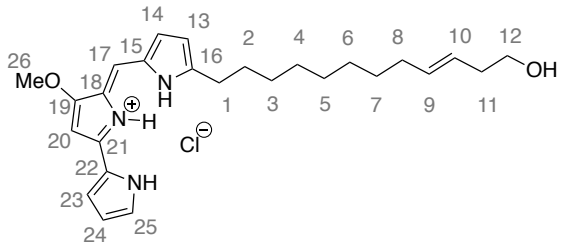

Compound **208** (260 mg, 0.53 mmol) was stirred in THF (5 mL) under argon and at RT. TBAF (1.07 mL, 1.07 mmol) was added and the reaction mixture stirred overnight. The mixture was concentrated *in vacuo* and the resulting slurry was partitioned between EtOAc (10 mL) and water (10 mL). The organics were extracted using EtOAc (3 × 20 mL) and the combined organics were dried ($MgSO_4$), concentrated *in vacuo* and purified by flash chromatography ($SiO_2$; 3:10; EtOAc : Petroleum ether) to give the final product as a white solid (100 mg, 75 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.95 (1H, N-H), 6.66 (m, 1H, H15), 6.13-6.12 (m, 1H, H15), 5.91 (m, 1H, H13), 5.58-5.52 (m, 1H, H9), 5.40-5.34 (m, 1H, 10), 3.63 (t, 2H, H12, *J* = 6.5 Hz), 2.59 (t, 2H, H1, *J* = 7.5), 2.26 (m, 2H, H11), 2.01 (m, 2H, H8), 1.61 (m, 2H, H2), 1.40-1.22 (m, 10H, H3-H7)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 134.6 (C9), 133.0 (C16), 125.9 (C10), 116.1 (C15), 108.4 (C14), 105.0 (C13), 62.2 (C12), 36.1 (C11), 32.8 (C8), 29.8, 29.8, 29.5, 29.5, 29.5, 19.2 (C2-C7), 27.9 (C1)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{27}NO$: 272.1985 $[M+Na]^+$; observed 272.1985 $[M+Na]^+$

**(*E*)-12-(5-((*Z*)-(4'-methoxy-1*H*,5'*H*-[2,2'-bipyrrol]-5'-ylidene)methyl)-1*H*-pyrrol-2-yl)dodec-3-en-1-ol hydrochloride salt 203**




Compound **209** (30.0 mg, 0.12 mmol) and MBC 21 (21.0 mg, 0.11 mmol) were stirred in methanol (3 mL) at RT and under argon and 1 M HCl (0.20 mL) was added. The reaction mixture was stirred at RT for 15 minutes. $H_2O$ (3 mL) was added and the organics were extracted with EtOAc (3 × 5mL). The combined organics were dried ($MgSO_4$), concentrated *in vacuo* and purified by flash chromatography ($SiO_2$; 100 % EtOAc) to give the final product as a red solid/ film (12 mg, 26 %).

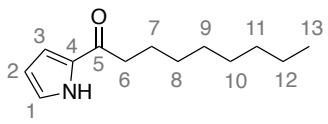
$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 12.91, 12.70, 12.68 (NH), 7.01 (m, 1H, H1), 6.95 (m, 1H, H2), 6.84 (m, 1H, H13), 6.37 (m, 1H, H3), 6.21 (m, 1H, H12), 6.09 (m, 1H, H6), 5.57-5.52 (m, 1H, H23), 5.40-5.34 (m, 1H, H24), 5.10 (m, 1H, H10), 4.02 (s, 3H, H8), 3.62 (t, 2H, H26, *J* = 6.0 Hz), 2.94 (t, 2H, H15, *J* = 6.5 Hz), 2.28-2.24 (m, 2H, H25), 2.02-1.98 (m, 2H, H22), 1.80-1.25 (m, 12H, H16-21)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 166.4 (C7), 153.4 (C14), 149.1 (C9), 134.5 (C23), 128.7 (C13), 126.2 (C11), 125.8 (C24), 123.8 (C4), 122.4 (C5), 118 (C2), 116.6 (C1), 112.7 (C12), 112.1 (C3), 93.2 (C6), 67.1 (C10), 62.2 (C26), 58.9 (C8), 36.1 (C25), 32.8 (C22), 29.8 (C15) 29.5, 29.5, 29.4, 29.4, 29.2, 28.6 (C16-C21)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{26}H_{36}N_3O_2$: 422.2802 $[M+H]^+$; observed 422.2803 $[M+Na]^+$

### 6.3.5 Synthesis of 2-UP chain length/branched analogues

**1-(1*H*-pyrrol-2-yl)nonan-1-one 159**



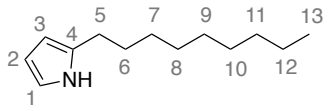

Nonanoic acid **158** (500 mg, 3.16 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material purified by flash chromatography ($SiO_2$; 1:2; EtOAc : Petroleum ether) to give the final product as a thick colourless oil ( 250 mg, 38 %)

$^{1}$H-NMR (500 MHz, $CDCl_3$) δ: 9.33 (bs, 1H, N-H), 7.01 (m, 1H, H1), 6.90 (m, 1H, H3), 6.27 (m, 1H, H2), 2.75 (m, 2H, H6, *J* = 7.5 Hz), 1.74-1.68 (m, 2H, H7), 1.38-1.26 (m, 10H, H8-H12), 0.88 (t, 3H, H13, *J* = 6.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.2 (C5), 132.1 (C4), 124.1 (C1), 115.8 (C3), 110.6 (C2), 38.1 (C6), 31.9, 29.5, 29.4, 29.2, 25.3 (C7), 22.7, 14.1 (C13)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{13}H_{21}ONNa$: 230.1515 $[M+Na]^+$; observed 230.1516 $[M+Na]^+$

**2-nonyl-1*H*-pyrrole 151**



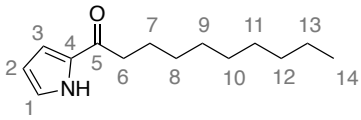

1-(1*H*-pyrrol-2-yl)nonan-1-one **159** (250 mg, 1.21 mmol) was subjected to the general reactions conditions for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:20; EtOAc : Petroleum ether) to give the final product as a light yellow oil (170 mg, 73 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.90 (bs, 1H, N-H), 6.66 (m, 1H H1), 6.13 (m, 1H, H2), 5.91 (m, 1H, H3), 2.59 (t, 2H, H5, *J* = 7.5 Hz), 1.65-1.59 (m, 2H, H6), 1.35-1.22 (m, 12H, H7-H12), 0.88 (t, 3H, H13, *J* = 6.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.0 (C4), 116.1 (C1), 109.5 (C2), 105.0 (C3), 32.0, 29.8, 29.7, 29.6 (C6), 29.5, 29.5, 27.9 (C5), 22.8, 14.3 (C13)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{13}H_{24}N$: 194.1903 $[M+H]^+$; observed 194.1902 $[M+H]^+$

**1-(1*H*-pyrrol-2-yl)decan-1-one 161**




Decanoic acid **160** (500 mg, 2.90 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material purified by flash chromatography ($SiO_2$; 1:20; EtOAc : Petroleum ether) to give the final product as a white solid (540 mg, 84 %).

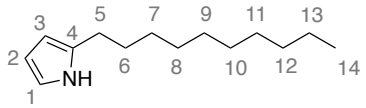
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.28 (bs, 1H, N-H), 7.01 (m, 1H, H1), 6.90 (m, 1H, H3), 6.27 (m, 1H, H2), 2.75 (t, 2H, H6, J = 7.7 Hz), 1.74-1.68 (m, 2H, H7), 1.36-1.25 (m, 12H, H8-H13), 0.88 (t, 3H, H14, *J* = 7.0 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.3 (C5), 132.3 (C4), 124.2 (C1), 115.9 (C3), 110.7 (C2), 38.2 (C6), 32.0, 29.6, 29.6, 29.4, 25.4 (C7), 22.8, 14.3 (C14)

IR ($V_{max}/cm^{-1}$): 2920 (C-H stretching), 1638 (C=O)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{14}H_{22}ON$: 220.1707 $[M-H]^-$; observed 220.1706 $[M-H]^-$

**2-decyl-1*H*-pyrrole 152**



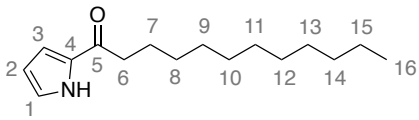

1-(1*H*-pyrrol-2-yl)decan-1-one **161** (540 mg, 2.44 mmol) was subjected to the general reaction conditions for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:20; EtOAc : Petroleum ether) to give the final product as a white thick oil (290 mg, 57 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.90 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13 (m, 1H, H2), 5.91 (m, 1H, H3), 2.59 (t, 2H, H5, *J* = 7.5 Hz), 1.64-1.58 (m, 2H, H6), 1.34-1.26 (m, 14H, H7-H13), 0.88 (t, 3H, H14, *J* = 6.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.1 (C4), 116.1 (C1), 108.4 (C2), 105.0 (C3), 32.1, 29.8, 29.8, 29.7, 29.6 (C6), 29.6, 29.5, 27.9 (C5), 22.8, 14.3 (C14)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{14}H_{26}N$: 208.2060 $[M+H]^+$; observed 208.2057 $[M+H]^+$

**1-(1*H*-pyrrol-2-yl)dodecan-1-one 162**




Dodecanoic acid **48** (500 mg, 2.50 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material purified by flash chromatography ($SiO_2$; 1:20; EtOAc : Petroleum ether) to give the final solid product (413 mg, 66 %).

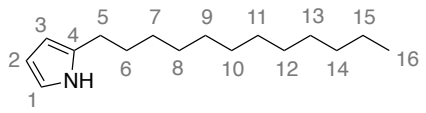
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.28 (bs, 1H, N-H), 7.01 (m, 1H, H1), 6.90 (m, 1H, H3), 6.28 (m, 1H, H2), 2.75 (t, 2H, H6, *J* = 7.5 Hz), 1.74-1.65 (m, 2H, H7), 1.36-1.25 (m, 16H, H8-H15), 0.88 (t, 3H, H16, *J* = 7.0 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.3 (C5), 132.3 (C4), 124.2 (C1), 115.9 (C3), 110.7 (C2), 38.2 (C6), 32.1 (C7), 29.8, 29.6, 29.6, 29.6, 29.5, 25.4 (C7), 22.8, 14.3 (C16)

IR ($V_{max}/cm^{-1}$): 2920 (C-H stretching), 1642 (C=O stretching)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{26}NO$: 248.2020 $[M-H]^-$; observed 248.2017 $[M-H]^-$

**2-dodecyl-1*H*-pyrrole 153**



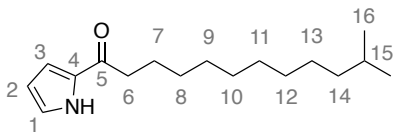

1-(1*H*-pyrrol-2-yl)dodecan-1-one **162** (412 mg, 1.65 mmol) was subjected to the general reaction conditions for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:20; EtOAc : Petroleum ether) to give the final product as a white solid (180 mg, 46 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.90 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13 (m, 1H, H3), 5.91 (m, 1H, H2), 2.59 (t, 2H, H5, *J* = 7.5 Hz), 1.65-1.59 (m, 2H, H6), 1.36-1.26 (m, 18H, H7-H15), 0.88 (t, 3H, H16, *J* = 6.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.1 (C4), 116.1 (C1), 108.4 (C2), 105.0 (C3), 32.1, 29.8, 29.8, 29.8, 29.7, 29.7, 29.6 (C6), 29.6, 29.5, 27.9 (C5), 22.8, 14.3 (C16)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{16}H_{30}N$: 236.2373 $[M+H]^+$; observed 236.2368 $[M+H]^+$

**11-methyl-1-(1*H*-pyrrol-2-yl)dodecan-1-one 164**




11-methyldodecanoic acid **163** (345 mg, 1.61 mmol) was subjected to the general reaction conditions for the coupling of carboxylic acids to pyrrole and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the final product as a white solid (299 mg, 70 %).

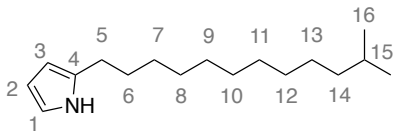
$^1$H-NMR (500 MHz, $CDCl_3$) δ: 9.28 (bs, 1H, N-H), 7.01 (m, 1H, H1), 6.90 (m, 1H, H3), 6.28 (m, 1H, H2), 2.75 (t, 2H, H6, *J* = 7.0 Hz), 1.74-1.68 (m, 2H, H7), 1.55-1.47 (m, 1H, H15), 1.36-1.24 (m, 12H, H8-H13), 1.16-1,12 (m, 2H, H14), 0.87 (d, 6H, H16, *J* = 6.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 191.3 (C5), 132.3 (C4), 124.2 (C1), 115.9 (C3), 110.7 (C2), 39.2 (C6), 38.2 (C7), 30.1, 29.8, 29.7, 29.6, 29.6, 29.4, 28.1 (C15), 27.6, 22.8 (C16)

LR-MS (ESI, +ve, MeOH) *m/z* (%): (100, 286.2 $[M+Na]^+$)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{17}H_{29}NONa$: 286.2141 $[M+Na]^+$; observed 286.2143 $[M+Na]^+$

**2-(11-methyldodecyl)-1*H*-pyrrole 154**




11-methyl-1-(1*H*-pyrrol-2-yl)dodecan-1-one **164** (200 mg, 0.76 mmol) was subjected to the general reaction conditions for the reduction of acyl ketones and the crude material purified by flash chromatography ($SiO_2$; 1:10; EtOAc : Petroleum ether) to give the oil product (110 mg, 58 %).

$^1$H-NMR (500 MHz, $CDCl_3$) δ: 7.89 (bs, 1H, N-H), 6.66 (m, 1H, H1), 6.13 (m, 1H, H2), 5.91 (m, 1H, H3), 2.59 (t, 2H, H5, $J$ = 7.5 Hz), 1.65-1.59 (m, 2H, H6), 1.55-1.49 (m, 1H, H15), 1.36-1.26 (m, 14H, H7-H13), 1.16-1.13 (m, 2H, H14), 0.87 (d, 6H, H16, $J$ = 6.5 Hz)

$^{13}$C-NMR (125 MHz, $CDCl_3$) δ: 133.1 (C4), 116.1 (C1), 108.4 (C2), 105.0 (C3), 39.2, 30.2, 29.9, 29.8, 29.8, 29.7, 29.6, 29.5, 28.1 (C15), 27.9 (C5), 27.6, 22.8 (C16)

HR-MS (ESI, +ve, MeOH) *m/z*: calculated *m/z* for $C_{17}H_{30}N$: 248.2384 $[M-H]^-$; observed 248.2381 $[M-H]^-$

## 6.4 Biology experimental

### 6.4.1 General

*Streptomyces* strains were handled in a Bassaire Laminar flow hood. Cultures grown on liquid and solid media were incubated in a New Brunswick Scientific Innova shakers and Heraeus static incubators (Thermo Scientific). Media was autoclaved at 121 °C and 1 bar for 10 minutes. Centrifugations were performed using Eppendorf models 5415D, 5424, 580R, 5810R or a Sorvell RC 6 Plus equipped with an SS34 or SLA3000 rotor. An Eppendorf Mastercycler Gradient or Eppendorf Mastercycler nexus GX2 machine was used to perform PCR reactions, and DNA/ protein concentrations were measured using a Nanodrop Lite spectrophometer (Thermo Scientific) and visualized using agarose gel electrophoresis and acrylamide SDS-Page gel electrophoresis for proteins. This was performed in a BioRad tank and a PowerPac™ (300 or basic) machine. The DNA was visualized using a UVP BioDoc-It® II Imaging System 2UV Transilluminator. Optical densities were measured by a BioMate 3 spectophotometer (Thermo Scientific). Proteins were visualized/ analyzed by gel electrophoresis using acrylamide SDS-PAGE and a BioRad Mini-PROTEAN® Tetra cell. UV-Vis measurements were conducted using a Perkin Elmer® Lambda 35 UV/Vis Spectophotomoter of solutions added to a sub-micro quartz cuvettes of 3 mm path length. Electroporations were conducted using a BioRad Gene Pulser® (200 Ω, 25 µM, 1.8 kV) with a BioRad Pulse Controller Plus. Samples for LC-MS analysis were prepared at a volume of between 250 1 µL-1 mL. LC-HRMS analysis of metabolites was performed on a Bruker MaXis Impact (Q-TOF) mass spectrometer coupled with a Dionex UltiMate 3000 UHPLC, using Agilent Zorbax Eclipse Plus C18 column (100 × 2.1 mm, 1.8 µm) with a flow rate of 0.2 mL/min.

### 6.4.2 Materials, kits and enzymes

GelRed Nucleic Acid Gel Stain and InstaBlue stain were purchased from Cambridge Bioscience and Expedeon. Bacto yeast extract and Bacto agar were purchased from Becton Dickinson and Company. All other chemicals and reagents were purchased from Sigma Aldrich. Taq polymerase, restriction enzymes and buffers were purchased from Thermo Fisher Scientific. Phusion polymerase and Gibson assembly kits were purchased from New England Biolabs. T4 DNA ligase and buffers were purchased from Roche. The kits used for DNA purification and plasmid DNA extraction were obtained from Thermo Fischer Scientific

### 6.4.3 Buffers & media

**Nickel (II) affinity washing/ elution buffer:** 100 mM NaCl, 20 mM Tris-HCl, 20 mM imidazole (wash) 50, 100, 200, 300 and 500 mM imidazole (elution), pH 8.0.

**PD-10 elution buffer**: 100 mM NaCl, 20 mM Tris-HCl, pH 8.0.

**Storage buffer**: 100 mM NaCl, 20 mM Tris-HCl, 10% glycerol, pH 8.0.

**SDS-PAGE running buffer (10**x): 250 mM Tris-HCl, 2 M glycine, 1 % SDS, pH 8.8.

**TBE buffer (10x):** 108 g Tris-HCl, 55.0 g boric acid, 11.7 g EDTA, water to 1 L, pH 8.0-8.3.

**Luria-Bertani (LB) Medium:** LB Broth powder (Miller's) (25.0 g) was dissolved in water (1 L) and the solution sterilized by autoclaving. LB agar was made by addition of Bacto agar (15.0 g).

**Soya Flour Mannitol (SFM) Medium**: Soya flour (20.0 g), D-mannitol (20.0 g) and Bacto agar (20.0 g) were dissolved in deionized water and sterilized by autoclaving.

**R5 Medium**: $K_2SO_4$ (0.25 g/L), $MgCl_2.6H_2O$ (10.12 g/L), glucose (10 g/L) sucrose (103 g/L), Difco Casamino acids (0.1 g/L), Difco yeast extract (5 g/L), trace element solution[a] (2 mL/L) and TES (5.73 g/L) were all dissolved in water and sterilized by autoclaving (Bacto agar added beforehand for solid media). When ready to use; 0.5 % (w/w) $KH_2PO_4$ (10 mL/L), 5 M $CaCl_2.H_2O$ (4 mL/L), 20 % L-proline (15 mL/L) and 1 N NaOH (7 mL/L) were added.

**International *Streptomyces* Project Medium - 2 (ISP2) Medium:** Yeast extract (4.0 g), malt extract (10.0 g), dextrose (4.0 g), agar (20.0 g), deionized water (1 L)

**M9 Medium:** (10x) M9 salts[b] (100 mL), $MgSO_4$ (1 M, 2 mL), $NH_4Cl$ (1 M, 1 mL), glucose (20 % w/v, 20 mL), casamino acids (20 % w/v, 10 mL), $CaCl_2$ (1 M, 100 µL), (1x) Trace elements (1 mL), cysteine.HCl (1 M, 1 mL), $FeSO_4.7H_2O$ (1 M, 1 mL), micronutrients solution (1 µL)

[a]Trace element solution: $ZnCl_2$ (40 mg/L), $FeCl_3.6H_2O$ (200 mg/L), $CuCl_2.2H_2O$ (10 mg/L), $MnCl_2.4H_2O$ (10 mg/L), $Na_2Ba_4O_7.10H_2O$ (10 mg/l) and $(NH_4)_6Mo_7.4H_2O$ (10 mg/L) dissolved in water.

[b](10x) M9 salts solution: $Na_2HPO_4$ (64.0 g), $KH_2PO_4$ (15.0 g), NaCl (2.50 g) dissolved in deionized water (1 L) and sterilized by autoclaving.

### 6.4.4 TAR cloning media and solutions

Media and solutions required for the TAR cloning procedure are from the protocol 'Direct Cloning and Heterologous Expression of Biosynthetic Gene Cluster,' by the Moore Lab.[ref]

**10 × Nitrogen Bases (100 mL):** Yeast nitrogen base without amino acids and ammonium sulphate (1.70 g), yeast synthetic drop-out medium supplements without tryptophan (1.90 g), ammonium sulphate (5.0 g) were dissolved in 100 mL warmed Milli-Q water and filtered using a 0.22 µm filter to prepare a 10x stock solution to be stored at 4 $^{o}$C.

**100 × Adenine (50 mL):** Adenine (0.50 g), HCl (1 M, 3.7 mL).

**Top Selective Agar (50 mL):** Sorbitol (9.10 g), dextrose (1.10 g), agar (1.50 g). This was microwaved and 10 × nitrogen bases and 100 × adenine stock were added before use.

**Bottom Selective Agar (500 mL):** Sorbitol (91.0 g), dextrose (11.0 g), agar (10.0 g) were mixed and left to set. The solid media was microwaved and 10 × nitrogen bases solution was added followed by 100 × adenine and 0.0001 % 5-fluoroorotic acid (5-FOA), before use. The 5-FOA was added from a 100 × stock solution (100 mg/mL in DMSO).

**Liquid Yeast Extract Peptone Dextrose (YPD) Medium (500 mL)**: Yeast extract (5.0 g) and peptone (10.0 g) were dissolved in water and sterilized by autoclaving. Glucose (2 % w/v) was added before use.

**Yeast Extract Peptone Dextrose (YPD) Agar (500 mL):** Yeast extract (5.0 g), peptone (10.0 g) and agar (7.50 g) were dissolved in water. This was microwaved to melt before use and glucose added (2 % w/v).

**Yeast selective liquid medium:** 10 × nitrogen base stock and glucose (2 % w/v) were mixed and filtered through a 0.22 µm syringe filter.
**1 M Sorbitol:** Sorbitol (91.0 g) dissolved in deionized water (500 mL).

**Milli-Q water:** Sterilized before use (200 mL).

**1 M HEPES buffer (pH 7.5):** HEPES (119.5 g) added to 400 mL deionized water and stirred. The pH was adjusted using NaOH and water was added to a total volume of 500 mL. The solution was filtered through a 0.22 μm syringe-filter and stored.

**EDTA Solution:** Solid EDTA was dissolved in water to a concentration of 0.5 M and the pH adjusted to 8.0 with NaOH.

**100 mM $CaCl_2$:** $CaCl_2$ (1.10 g) dissolved in water (100 mL).

**SPE:** HEPES buffer (500 μL), EDTA solution (1 mL), sorbitol (9.10 g), water (50 mL) and sterilized by autoclaving.

**SOS:** $CaCl_2$ solution (3.25 mL), yeast extract (0.125 g), sorbitol (9.10 g), peptone (0.50 g) and water (50 mL) and sterilized by autoclave.

**Tris Buffer**: Tris base (121.14 g) added to deionized water and the pH adjusted to 7.5 using 2 M HCl.

**PEG:** Tris buffer (100 μL), $CaCl_2$ (1 mL), PEG8000 (2.0 g), water (10 mL), filtered and adjusted to pH 7.5.

**Zymolyase 20T:** 10 mg/ mL zymolyase-20T (MP Bio), glycerol (25 % w/v), 50 mM Tris-HCl (pH 7.5) stored at -20 $^{o}$C in 500 μL aliquots.

### 6.4.5 Bacterial strains

**Table 1** : A list of bacterial strains used throughout this work.

| Bacterial strain | Use/ Function | Supplier |
|---|---|---|
| ***Escherichia coli*** | | |
| One Shot ® TOP10 chemically competent cells | Transformation, replication and DNA isolation | Thermo Fischer Scientific |
| One Shot ® BL21 Star™ (DE3) chemically competent cells | High-level recombinant protein expression | |
| One Shot ® BL21 (DE3) chemically competent | High-level recombinant protein expression | |
| Rosetta (DE3) pLysS | | Lab stock |
| Origami B(DE3) | | |
| ET12567 | Conjugation with *Streptomyces* strains, used to shuttle plasmid DNA | |
| ET12567/pUB307 | Conjugation with *Streptomyces* strains, used to aid conjugation | |
| ***Streptomyces*** | | |
| *albus* (WT) | Heterologous expression host for the captured metacycloprodigiosin biosynthetic gene cluster | Lab stock |
| *coelicolor M145* | Heterologous expression host for the captured metacycloprodigiosin biosynthetic gene cluster | |
| *longisporus ruber* (WT) | Capturing the entire metacycloprodigiosin biosynthetic gene cluster | |
| *albus* + mcpBGC | Heterologous host expressing metacycloprodigiosin biosynthetic gene cluster | This work |

| | |
|---|---|
| *albus* + mcpBGC-Δ*lk* | Heterologous host expressing metacycloprodigiosin biosynthetic gene cluster with mcp*LK* gene deletion |

### 6.4.6 Antibiotics

Antibiotics were made as stock solutions using the appropriate amounts and solvents and sterilized by filtration through a 0.20 µm syringe filter.

**Table 2:** Antibiotics and the corresponding concentrations used for bacteria selection.

| Antibiotic | Stock solution (mg/mL) | Solvent |
|---|---|---|
| Ampicillin | 100 | Water |
| Hygromycin | 50 | Water |
| Naladixic acid | 25 | 0.3 M NaOH |
| Kanamycin | 50 | Water |
| Chloramphenicol | 25 | EtOH |

### 6.4.7 PCR Primers

**Table 3:** PCR primers used for specific DNA amplifications within this project.

| Primer ID | Sequence (5'→3') | Use |
|---|---|---|
| | | |
| Mcp1-L-Fwd | cgccgatggtttctacaaagatcgactagttcagaccaagtgcaggcgctcctgg | Left *mcp* BGC capture arm |
| Mcp1-L-Rev | tccatgtgacctaggccagttcgtgagcctgagcccctgggactg | Left *mcp* BGC capture arm |
| Mcp1-M-Fwd | aggctcacgaactggcctacaacaactaagaaaatggcta | Middle *mcp* BGC capture region - G.A |
| Mcp1-M-Rev | agctgttccatgtgattagttttgctggccgcatcttctc | Middle *mcp* BGC capture region – G.A |
| Mcp1-R-Fwd | aggctcacgaactggcctaggtcacatggaacagctccagcgcct | Right *mcp* |

| | | |
|---|---|---|
| | | BGC capture arm |
| Mcp1-R-Rev | acctcaagtctcgagtgaacgcatgtacagcagactaaacacccg | Right *mcp* BGC capture arm |
| **TAR - capture vector sequencing & checking primers** | | |
| Mcp1-Fwd-Seq | gagctgacagttcagggtgc | Sequencing of cloned *mcp* cluster capture arms |
| Mcp1-Rev-Seq | cgagcgcttctgcgagct | |
| CK1-Fwd | gatgctaaggtagagggtg | Confirming *mcp* capture arm cloning |
| CK2-Rev | tctttgtctttgttcttcc | |
| **TAR – primers used to confirm the entire *mcp* biosynthetic gene cluster** | | |
| mcpCK1-Fwd | agcgagaagaatgcggagg | Confirming DNA regions spanning entire *mcp* BGC |
| mcpCK1-Rev | cgtgagtgagattgacgaaagg | |
| mcpCK2-Fwd | ccgcagaatcggataccact | |
| mcpCK2-Rev | tctggctggtcaggaacgtc | |
| mcpCK3-Fwd | cgtctccgtcacttgccg | |
| mcpCK3-Rev | cttgccgaacagttgcgtc | |
| mcpCK4-Fwd | tgtcccgtgaaggggtagt | |
| mcpCK4-Rev | ctccgtcgtcaaggtggc | |
| **TAR – primers used to construct mcp*LK* deletion** | | |
| *mcpLK*-deletion-1-Fwd | tgacgcaggtcgtcgtcgccgagcccttggtcc | PCR amplification of DNA regions flanking mcp*LK* gene |
| *mcpLK*-deletion-1-Rev | cccgaccgagcgcgcacgcgtgccccgctga | |
| *mcpLK*-deletion-2-Fwd | cgcgtgcgcgctcggtcgggccttgggtcat | |
| *mcpLK*-deletion-2-Rev | gggcgactgtgtgacccccggctactggcgccg | |
| | | |
| *mcpLK*-CK-Fwd | gcggtagacggaaggggtgc | Primers used to confirm the deletion of |
| *mcpLK*-CK- | cgcaaggaccgcatggtc | |

| | | |
|---|---|---|
| Rev | | mcp*LK* gene |
| **Primers used to clone *RedG* & *McpG* into protein expression vectors** | | |
| pET28a(+) vectors | | |
| McpG-Fwd | cccgcatatgatcccgaatcagtggtatccga | Cloning McpG from *S. longisporus ruber* |
| McpG-Rev | atagaattctcagggcgcggtcttgctgagcc | |
| RedG-Fwd | cccgcatatgatccccaaccagtggtatcccatcgtcgagg | Cloning RedG from *S. coelicolor* M511 WT |
| RedG-Rev | atatgaattcctacgggcgcgcggcggcgtcctccttgg | |
| pETSUMO vectors | | |
| McpG-SUMO-Fwd | atgatcccgaatcagtggtatccga | Cloning McpG from *S. longisporus ruber* |
| McpG-SUMO-Rev | tcagggcgcggtcttgctgagcc | |
| RedG-SUMO-Fwd | atgatccccaaccagtggtatcccatcgtcgagg | Cloning RedG from *S. coelicolor* M511 WT |
| RedG-SUMO-Rev | ctacgggcgcgcggcggcgtcctccttgg | |

### 6.4.8 Vectors

**Table 4:** A list of vectors used within this project and their corresponding function.

| Vector | Resistance Marker | Use | Genotype | Supplier |
|---|---|---|---|---|
| pCAP1000 | Kanamycin | Capturing the metacycloprodigiosin biosynthetic gene cluster fro *S. longis porus ruber* | *ori, oriT, URA3, phage φC31 integrase* | Lab stock-Dr. Chuan Huang |
| pCAP1000mcpBGC | Kanamycin | Heterologous expression of metacycloprodigiosin gene cluster | *ori, oriT, URA3, phage φC31 integrase* | This work |

| | | | | |
|---|---|---|---|---|
| pCAP1000mcpBGC-Δ*LK* | Kanamycin | Heterologous expression of metacycloprodigiosin gene cluster with deletion in 2-UP biosynthesis | *ori, oriT, URA3, phage φC31 integrase* | |
| pOSV556t*mcpHG* | Ampicillin | Expression of *mcpHG* genes from *S. longisporus ruber* | *ori, oriT, lacI* | Lab stock – Dr. Rebin Salih |
| pET151*RedG* | Ampicillin | High-level protein expression | *ori, lacI* | Lab stock – Dr. Paulina Sydor |
| pET28a(+)*RedG* | Kanamycin | High-level protein expression | *ori, lacI* | This work |
| pET28a(+)*McpG* | Kanamycin | High-level protein expression | *ori, lacI* | |
| pETSUMO*RedG* | Kanamycin | High-level protein expression and increased protein solubility | *ori, lacI,* SUMO tag | |
| pETSUMO*McpG* | Kanamycin | High-level protein expression and increased protein solubility | *ori, lacI,* SUMO tag | |
| pDB1282 | Ampicillin/ carbenicillin | Expression of proteins responsible for the biosynthesis of iron-sulphur clusters | Arabinose-inducible promoter | Lab stock |

## 6.5 Biological procedures

### 6.5.1 General bacterial growth conditions

A number of strains of *E. coli* and *Streptomyces* were used throughout this work. *E. coli* was grown overnight on solid LB agar at 37 °C, or in liquid LB broth at 180 rpm and at 37 °C. Selected media contained the appropriate antibiotic for bacteria selection. *Streptomyces* was grown on solid SFM, ISP2 or R5 media at 30 °C for 5-7 days.

### 6.5.2 Preparation of chemically competent *E. coli* cells

*E. coli* cells from glycerol stocks were used to inoculate 5 mL LB media, with the addition of the appropriate antibiotic, and incubated with shaking at 37 °C overnight. The resulting overnight preculture (5 mL) was then used to inoculate fresh LB media (500 mL) and the cells were incubated with shaking at 37 °C. The cells were incubated for approximately 3-4 hours (until the $OD_{600}$ of 0.40) before being placed on ice for 20-30 minutes. The culture was centrifuged at 300 rpm for 10 minutes at 4 °C. The supernatant was removed and the cells resuspended in ice-cold $CaCl_2$ (0.1 M, 500 mL) and submerged in ice for 30 minutes. The cells were centrifuged at 300 rpm for 10 minutes at 4 °C and the supernatant removed. The pelleted cells were resuspended in cold $CaCl_2$ containing 15 % glycerol and stored as 100 µL aliquots in a micro-centrifuge tube and left on ice for 1 hour. The aliquots were flash frozen in liquid nitrogen and stored at -78 °C until use.

### 6.5.3 Preparation of electrochemically competent *E. coli* cells

The appropriate *E. coli* strain (typically *E. coli* ET12567, 100 µL) was used to inoculate LB media (5 mL) with the appropriate antibiotics and grown overnight at 37 °C with shaking. The overnight preculture (100 µL) was used to inoculate fresh LB media (5 mL) and incubated at 37 °C with shaking until reaching an $OD_{600}$ of 0.40-0.60. The cells were centrifuged at 400 rpm for 10 minutes at 4 °C and supernatant removed. The cell pellet was resuspended in ice-cold sterile glycerol (10 %) solution and centrifuged at 400 rpm for 10 minutes at 4 °C. The supernatant was removed and the glycerol wash repeated. The cell pellet was resuspended in glycerol (10 %) in aliquots of 100 µL and flash frozen in liquid nitrogen and the microcentrifuge tube stored at -78 °C until use.

### 6.5.4 Plasmid isolation, DNA manipulation and transfer

#### 6.5.4.1 Transformation of chemically competent *E. coli cells*

The required plasmid (3 μL) was added to the frozen aliquot of prepared chemically competent *E. coli* cells and left on ice for 30-45 minutes. The cells were then submerged in a water bath at 42 °C for one minute and then placed directly on ice. LB media (300 μL) was added to the cells and incubated with shaking at 37 °C for 1 hour to aid recovery. The resulting culture was spread onto the appropriate solid media containing the appropriate antibiotics and incubated at 37 °C overnight.

#### 6.5.4.2 Transformation of electrocompetent *E. coli* cells

The required plasmid DNA (typically pCAP1000 or pOSV556t variants, 3-7 μL) was added to electrocompetent *E. coli* ET12567 cells (100 μL). A 0.20 cm electroporation cuvette was put on ice until cold. The mixture was added to the ice-cold cuvette and shocked at 1.8 kV. LB media (1 mL) was immediately added to the cuvette, transferred to a sterilized Eppendorf tube and incubated at 37 °C with shaking for 1 hour for recovery. Different volumes of the transformation mixture were added to LB agar plates containing the appropriate antibiotics and incubated at 37 °C overnight.

#### 6.5.4.3 Co-transformation of chemically competent *E. coli* cells

Based upon the transformation procedure outlined in section 6.4.5.2 but using two plasmids rather than one. The two plasmids of interest (3 μL) were added to the frozen aliquot of prepared chemically competent *E. coli* cells and left on ice for 1 hour. The cells were then submerged in a water bath at 42 °C for one minute and then placed directly on ice. LB media (300 μL) was added to the cells and incubated with shaking at 37 °C for 1 hour to aid recovery. The resulting culture was spread onto the appropriate solid media containing two appropriate antibiotics, to select for colonies containing both of the desired plasmids, and incubated at 37 °C overnight.

#### 6.5.4.4 Plasmid isolation from *E. coli*

A single colony or 1 µL of *E. coli* glycerol stock solution harboring the plasmid DNA of interest was used to inoculate LB media (5 mL) containing the appropriate antibiotic and incubated with shaking at 37 °C overnight. The culture was centrifuged at 4000 rpm for 10 minutes at 4 °C and the supernatant discarded. The plasmid DNA was extracted from the resulting pellet using a GeneJET Plasmid Miniprep kit according to the manufacturers instructions.

#### 6.5.4.5 Polymerase Chain Reaction (PCR)

PCR reactions (50 µL) contained the following components; Phusion polymerase mastermix (25 µL), template DNA (1 µL, 100 ng/ µL), primer 1 (2 µL, 100 mM), primer 2 (2 µL, 100 mM), DMSO (4 µl)*, deionized $H_2O$ (16 µL).

**PCR cycle**:

1. Denaturation: 98 °C, 30 s
2. Denaturation: 98 °C, 10 s
3. Annealing: $T_m$ – 5 °C**, 30 s
4. Extension: 72 °C, X min***
5. (Repeat of 2, 3 & 4 (x30 cycles))
6. Final extension: 72 °C, 10 minutes
7. Resting: 4 °C, until end

*Higher concentrations of DMSO were used in certain PCR reactions.
**Annealing temperatures were optimized by gradient PCR trials or were approximately 5 °C below the melting temperatures of the primers.
***Extension times 'X' were based on 1 minute per 1000 bp.

#### 6.5.5.6 Agarose gel electrophoresis, gel extraction and DNA quantification

Agarose (1 g) was added to TBE buffer (1 × stock solution, 100 mL) and dissolved by microwave heating. Once cooled, GelRed (1 µL) was added and the solution poured into a mold fitted with a suitably sized comb.* The solid gel was transferred to a gel electrophoresis tank and filled until covered with TBE buffer (1 ×). DNA loading dye was added to DNA samples (5 µL dye to 20 µL sample) and the samples loaded into the wells. A DNA molecular

weight marker was loaded alongside samples to use as a reference. Electrophoresis was undertaken at 80-100 V for approximately 1 hour. Once good separation had occurred, the gel was removed and visualized using a UVP UV transilluminator. The DNA fragments were cut from the agarose gel and purified using a GeneJET Gel Extraction kit according to the manufacturers instructions. A NanoDrop spectrophotometer was used to quantify the concentration of DNA.

*In some cases, GelRed addition was omitted and the following gel electrophoresis steps were continued. After 1 hour the gel was removed from the tank and placed into water (500 mL) containing GelRed (~10 µL) and left to sit overnight. The gel/ DNA was visualized using then visualized using a UVP UV transilluminator as described above.

#### 6.5.5.7 Restriction enzyme digests

Typical restriction digests were comprised of DNA (20 µL), appropriate restriction enzyme (2 µL), deionized water (23 µL) corresponding buffer (5 µL), as prescribed by the manufacturer or by using the Thermo Fischer DoubleDigest Calculator were incubated in a 37 °C water bath for 1-3 hours. Digested DNA was then purified using the GeneJET PCR purification kit or by the agarose gel electrophoresis and gel extraction procedure described in section 6.5.5.6.

#### 6.5.5.8 Ligation

Ligations were undertaken using the Rapid DNA Ligation kit according to the manufacturers instructions. The insert : vector ratios were calculated using the *in silico* ligation calculator, using PCR products purified by the methods outlined in section 6.5.5.6 and digested using the methods in section 6.5.5.7. The ligation mixture (5-10 µL) was used directly to transform *E. coli* TOP10 cells and spread onto agar containing the appropriate antibiotics to select for transformants harboring successfully ligated vectors. The DNA was isolated from the resulting colonies using the DNA isolation methods outlined in section 6.5.5.4 and the vectors confirmed by PCR and/or DNA sequencing.

#### 6.5.5.9 Gibson Assembly

Gibson Assembly was undertaken as a method of ligation following the manufacturers instructions using the recommended amounts/ratios of vector to insert. Primers were designed in advance following the requirements recommended by the manufacturers instructions. DNA vectors were linearized in advance using the restriction digest methods described in section 6.5.5.7. A typical procedure was prepared on ice and comprises of; DNA vector (50-100 ng), insert (2-3 eqv), Gibson Assembly master mix (2 ×) and deionized water to a total volume of 20 µL. The mixture was incubated at 50 °C for 1 hour and used directly (5-8 µL) to transform *E. coli* TOP10 cells before spreading onto solid agar containing appropriate antibiotics and incubated at 37 °C overnight. Transformants were screened for successfully assembled vectors by isolating the DNA following the previously described procedure and confirmed by PCR and/ or DNA sequencing.

### 6.5.6 Molecular engineering of *Streptomyces* strains

#### 6.5.6.1 Preparation of surface grown *Streptomyces* spore stock solutions

A single colony, or 10 µL liquid culture, of *S. albus, S. coelicolor* or *S. longisporus ruber* was spread onto solid SFM media and grown at 30 °C until the formation of a confluent lawn (5-7 days). Between 1-3 mL of sterile water was added to the plates and the spores were suspended by carefully scratching the mycelia with a sterile loop. The suspension was filtered through sterile non-absorbant glass wool. Glycerol was added to a final concentration of 50 % and the spores stored at -78 °C until use.

#### 6.5.6.2 Intergeneric triparental conjugation – DNA transfer from *E. coli* to *Streptomyces* strains

Triparental conjugation was used as a method of transferring plasmid DNA to *Streptomyces* strains from *E. coli.* The plasmid DNA in question was typically pCAP100 or pOSV556t variants and can be used directly for conjugation as they each contain *oriT* on their backbone.

Electrocompetent cells of *E. coli* ET12567 were prepared using the procedure outlined in section 6.5.3. The competent cells were transformed with the desired plasmid DNA using the procedure outlined in section 6.5.4.2, using the appropriate antibiotic and chloramphenicol. A colony of *E. coli* ET12567/plasmid DNA was used to inoculate LB media (10 mL) containing

the required antibiotics and grown overnight at 37 °C with shaking. *E. coli* ET12567/pUB307 was also used to inoculate LB media (10 mL) in the same manner.

The overnight precultures (100 µL) were used to inoculate fresh LB media (10 mL) and the appropriate antibiotics and grown at 37 °C with shaking for approximately 4 hours until an $OD_{600}$ of 0.40-0.60. The cells were washed twice by centrifugation at 400 rpm for 10 minutes at 4 °C, discarding the supernatant and resuspending in fresh LB media (10-50 mL) to remove antibiotics. *E. coli* ET12567/pUB307 was resuspended in 0.5-1 mL LB media and *E. coli* ET12456/plasmid DNA in 0.20-0.40 mL LB.

*Streptomyces* spores (10-20 µL) were added to LB media (~ 200 µL) and placed in a water bath at 50 °C for 10-15 minutes then allowed to cool at RT. The heat-shocked spores were mixed with each of the prepared *E. coli* cell suspensions (100 µL) and different amounts of the mixture were spread onto solid SFM media containing 10 mM $MgCl_2$, minus antibiotics, and incubated at 30 °C overnight.

The plates were overlaid with 1.5 mL water containing 0.50 mg nalidixic acid and 2-4 mg of the appropriate antibiotic to select for the plasmid DNA in question. The plates were incubated at 30 °C for 5-7 days.

A single colony of Streptomyces was picked and spread onto a solid SFM media plate containing nalidixic acid and the appropriate antibiotic for plasmid DNA selection to further remove *E. coli* contaminations. The plate was incubated at 30 °C for 5-7 days and spore stocks were generated following the previously described procedure.

#### 6.5.6.3 Feeding synthetic substrates to *Streptomyces* strains expressing *mcpH* and *mcpG* or pCAP100mcpBGC-Δ*LK*

Feeding experiments were undertaken on solid SFM media overlaid with a sterile semi-permeable membrane. The desired mutant *Streptomyces* strain spore suspension (10-20 µL) was spread directly onto the semi-permeable membrane and incubated at 30 °C for 2-3 days. Synthetic 2-UP and analogues (1 mg) were dissolved in MeOH (50 µL). Synthetic MBC (1 mg) was dissolved in DMSO (50 µL). The two solutions were mixed together and the resulting 100 µL mixture was pipetted directly onto the *Streptomyces* mycelia as small droplets, covering as much of the mycelia surface area as possible. The plates were allowed to dry for 30-60 minutes and incubated at 30 °C for a further 3 days.

#### 6.5.6.4 Extraction of prodiginine compounds and preparation of LC-MS samples

After 3 days of incubation, post-feeding, the biomass was scraped from the membrane and placed into a 20 mL glass vial. MeOH (5 mL) was added directly to the mycelia followed by 2 M HCl (250 µL). The resulting solution was sonicated for 5 minutes. MeOH (400 µL) was added to a 0.20 µm spin-filter along with 100 µL of the sonicated mycelia extract. The content was centrifuged at 1000-1500 rpm for 3 minutes (repeated if necessary) and 500 µL of the filtrate was added to MeOH (500 µL) and analysed by LC-MS.

### 6.5.7 TAR cloning – Preparation of yeast spheroplast cells and transformation

The TAR cloning procedure is based upon a method outlined by the Moore group.[106]

#### 6.5.7.1 Preparation of yeast spheroplasts

A single colony of *Saccharomyces cerevisiae* VL6-48N was used to inoculate YPD medium with adenine (10 mL) and incubated at 30 °C overnight with shaking. The yeast preculture (1 mL) was used to inoculate YPD media (50 mL) with adenine, in a 250 mL Erlenmeyer flask and grown at 30 °C with shaking (220 rpm) for 5-8 hours until an $OD_{600}$ of 0.70-1.0. The contents was transferred to a Falcon tube and placed on ice for 10 minutes before centrifuging the cells at 1800 g for 3 minutes at 4 °C. The supernatant was removed and the cell pellet was resuspended in ice-cold sterile water. The cells were centrifuged again at 1800 g for 3 minutes at 4 °C. The supernatant was removed and the cell pellet resuspended in ice-cold sorbitol (1 M). The cells were left overnight at 4 °C to osmotically stabilize the cells before spheroplasting.

The Falcon tube containing the yeast cells was inverted (× 10) and the centrifuged at 1800 g for 3 minutes at 4 °C. The supernatant was removed, thoroughly, and the cells resuspended using SPE solution (20 mL) before the addition of 2-mercaptoethanol (40 µL) and the tube inverted (× 10) to mix. Zymolase solution (80 µL) was added and inverted (× 10) to mix and initiate spheroplast formation. The cells were incubated at 30 °C for 40 minutes and the tube inverted every 5 minutes.

#### 6.5.7.2 Determining the level of yeast spheroplasting

The level of spheroplasting was determined by measuring the optical density ($OD_{600}$) of the cell suspension. A small sample of the zymolase 20T-digested cells was split into 2 samples. One sample was diluted with 1 M sorbitol (5x dilution). The second was diluted with an SDS solution (2 % v/v in water, 5 × dilution). The optical densities of both samples were measured. The yeast cells are determined to be at a suitable level of spheroplasting when the difference between the two samples is approximately 20-fold or above. This will indicate that approximately 95 % of the cells have been converted to spheroplast. The cells can be incubated further to increase the level of spheroplasting if the $OD_{600}$ is not yet at the required level. Incubation time must not exceed 50 minutes.

#### 6.5.7.3 Transformation procedure

When the level of spheroplasting is deemed to be sufficient, ice-cold 1 M sorbitol was added to the cells suspension up to 50 mL and the tube inverted gently to mix. The cells were centrifuged at 600 g for 10 minutes at 4 °C. The supernatant was removed thoroughly and the spheroplasts resuspended by gentle pipetting using ice-cold 1 M sorbitol (20 mL). Sorbitol was added to a final volume of 50 mL and the cells were centrifuged at 600 g for 10 minutes at 4 °C. The supernatant was removed thoroughly and the spheroplasts resuspended in 2-3 mL STC solution by gentle pipetting. The spheroplasts were incubated at RT for 10 minutes. The linearized capture vector (0.2-0.5 µg) and the desired genomic DNA (2-3 µg) were mixed with 200 µL of the spheroplast solution, in a microcentrifuge tube, by gentle pipetting. The cell suspension was incubated at RT for 10 minutes before 800 µL of PEG solution (20 %) was added and the tube inverted. The mixture was incubated at RT for 20 minutes before the cells were microcentrifuged at 700 g for 10 minutes at 4 °C. The supernatant was removed using a 1 mL pipette. SOS solution (800 µL) was added to the cell pellet and gently pipetted to mix. The cell suspension was incubated at 30 °C for 40 minutes. Top agar was melted in a 50 °C water bath (8 mL). The cell suspension was added directly to the heated top agar, inverted, and immediately poured onto solid bottom agar containing 5-FOA, in a petri dish. The plates were incubated for 7 days at 30 °C.

#### 6.5.7.4 Selecting yeast transformants

Yeast transformants form colonies both on and within the top agar after approximately 1 week of incubation at 30 °C. Candidate colonies were selected using a pipette tip and small patches were spread onto fresh bottom agar plates without 5-FOA. The plates were incubated for ~ 4 days at 30 °C.

#### 6.5.7.5 Screening candidate yeast colonies by colony PCR

Screening yeast colonies for plasmid DNA that has successfully captured the desired gene cluster was undertaken by colony PCR. Primers were designed in advance that targeted the central region of the biosynthetic gene cluster. Colony PCR was undertaken by extracting plasmid DNA from the candidate yeast cells to be used as template DNA. Three solutions were prepared for the extraction of DNA from yeast;

**Solution 1:** Sucrose 10 % w/v, 50 mM Tris-HCl (pH 8), 10 mM EDTA, 2-mercaptoethanol (1 µL) and Zymolase 20T (0.20 mg/mL)

**Solution 2:** NaOH (0.20 M), SDS (1% w/v)

**Solution 3:** 3 M potassium acetate, pH 8

The patched cells were collected and resuspended in solution 1 (200 µL) before incubation at 37 °C for 2 hours. The cell suspension was mixed with solution 2 (400 µL) and mixed by inversion. Solution 3 (300 µL) was added and the suspension microcentrifuged at maximum speed for 10 minutes. An equal volume of propan-2-ol was added to precipitate the DNA and the resulting solution was microcentrifuged at maximum speed for 5 minutes. The supernatant was removed carefully using a pipette and dried for 10-15 minutes in a 60 °C oven. The DNA pellet was dissolved in water or buffer and used in 50 µL PCR reactions.

#### 6.5.7.6 Extracting DNA from yeast colonies that generated a positive PCR result

The positive yeast colony was used to inoculate yeast selective medium and incubated for 30 hours at 30 °C with shaking. The cells were centrifuged at 1800 g for 5 minutes at 4 °C and the supernatant removed. The cell pellet was resuspended in solution 1 (200 µL) and incubated for 2 hours at 37 °C. The resulting cell suspension was mixed with solution 2 (400

μL) followed by solution 3 (300 μL). The cell suspension was centrifuged at maximum speed for 10 minutes. The residual proteins were removed by phenol-chloroform denaturation and the DNA was precipitated using propan-2-ol and EtOH (70 % v/v). The DNA pellet was dissolved in water or buffer before ready for use in transformation procedures described previously with the appropriate *E. coli* strain (typically *E. coli* TOP10 or *E. coli* ET12567).

### 6.5.8 LC-MS analysis of mycelium-extracted metabolites

A gradient of water and acetonitrile was used to separate the metabolites with a run time of 34 minutes:

**Table 5:** LC-MS conditions used to analyze prodiginine metabolites.

| Time (min) | Acetonitrile % (v/v) | Water % (v/v) |
|---|---|---|
| 0-5 | 5 | 95 |
| 6-17 | 5-100 | 95-0 |
| 18-22 | 100 | 0 |
| 23-25 | 100-5 | 0-95 |
| 26-34 | 5 | 95 |

### 6.5.9 Protein overexpression and purification

#### 6.5.9.1 General procedure for the overexpression of recombinant fusion proteins

The protein expression vector of interest was used to transform the appropriate *E. coli* strain using the transformation procedure previously outlined. A single transformant was used to inoculate sterile LB media (10 mL), with the appropriate antibiotic, and grown at 37 $^{\circ}$C overnight with shaking. The overnight preculture was then used to inoculate sterile LB media (1 L) with the appropriate antibiotic and incubated at 37 $^{o}$C with shaking until an $OD_{600}$ of 0.60-1.00. When the $OD_{600}$ has reached the desired value (1 M) isopropyl-β-thiogalactopyranoside (IPTG) (1 mL) was added to the culture, over a flame, before incubation at 15 $^{\circ}$C with shaking for 16-20 hours.

#### 6.5.9.2 General procedure for the purification of recombinant fusion proteins

Proteins were typically cloned into expression vectors (listed in table 4) to contain N-terminal $His_6$-recombinant fusion proteins and purified using a 1 mL HisTrap$^{TM}$ HP affinity column with binding buffer. The 1 L induced cell culture was centrifuged at 4000 rpm at 4 $^{\circ}$C for 20-30 minutes. The supernatant was discarded and the cell pellet resuspended in washing buffer (~ 20 mL). The cells were lysed using sonication (× 5) and centrifuged 17000 rpm at 4 $^{\circ}$C for 1 hour. The supernatant/ cell-free extract was filtered through a 0.45 μM syringe filter and loaded onto the HisTrap$^{TM}$ affinity column, which had been equilibrated with washing buffer (5-10 mL). Unbound proteins were eluted using low concentration imidazole elution buffers (20 mM – 20 mL, 50 mM – 15 mL and 100 mM – 10 mL imidazole. The protein of interest was eluted from the column using elution buffers with imidazole concentration typically between 300-500 mM at a volume of 1-2 mL to provide a protein sample that was typically ~ 30 μM. Fractions containing the desired fusion protein as determined by acrylamide SDS-PAGE analysis were either concentrated using Amicon® Ultra-15 centrifugal filters with a 10 kDa molecular weight cut-off membrane and washed with storage buffer, or passed through a PD-10 column and eluted using storage buffer. Glycerol was added to the protein solution to a final concentration of between 10-20 % and flash frozen in liquid nitrogen (storage -78 $^{\circ}$C).

#### 6.5.9.3 Optimized conditions for the overexpression of $His_6$-RedG with an intact FeS cluster

Both the RedGpETSUMO (3 μL) protein expression vector and the plasmid pDB1282 (3 μL) responsible for ISC biosynthesis were used to transform *E. coli* BL21 (DE3) using the procedure for cotransfromation of *E. coli* strains. A single transformant was used to inoculate LB media (10 mL) with kanamycin and ampicillin antibiotics and incubated at 37 $^{o}$C with shaking for 16-20 hours. The preculture was used to inoculate 1 L sterile LB media containing ampicillin and kanamycin and incubated at 37 $^{o}$C until an $OD_{600}$ of 0.80-1.0. Arabinose (3 g) was added to the culture to induce the biosynthesis of the iron-sulphur cluster by expression of the vector pDB1282, and incubated for 45 minutes at 37 $^{\circ}$C with shaking. The culture was cooled to 4 $^{o}$C before IPTG (0.1 mM final concentration) and $FeSO_4.7H_2O$ (27.8 mg) were added over a flame. The culture was incubated at 15 $^{\circ}$C for 20-25 hours with gentle shaking at 100 rpm to minimize oxygenation of the cell suspension.

#### 6.5.9.4 Optimized conditions for the anaerobic purification of $His_6$-RedG with an intact FeS cluster

All buffers were purged of oxygen by bubbling through argon gas for 15-20 minutes prior to the purification. Argon gas was also used to purge any required centrifuge tubes, Falcon tubes and Eppendorf tubes before use. Purifications of $His_6$-RedG were undertaken within an Atmos two-hand non-sterile glove bag with a tape-seal. The glove bag was purged of oxygen prior to purifications using nitrogen gas. All required equipment (HisTrap™ affinity column, wash/ elution buffers, syringes and Falcon tubes) were placed within the glove bag before filling with nitrogen gas and the seal secured using tape and metal clips. The 1 L induced cell culture was centrifuged at 4000 rpm at 4 °C for 20-30 minutes. The supernatant was discarded and the brown/black cell pellet resuspended in argon-purged washing buffer (~ 20 mL). The cells were lysed using sonication (× 5), generating a brown/black solution and centrifuged at 17000 rpm for 1 hour at 4 °C. The supernatant/ cell-free extract was filtered through a 0.45 μM syringe filter in to an argon-purged Falcon tube and placed into the purged glove bag. The brown cell-free extract was loaded onto the HisTrap™ affinity column, which had been equilibrated with argon-purged washing buffer (5-10 mL) prior to the purification. The presence of RedG recombinant protein bound to the column is indicated by a dark brown/ black colouration. The unbound proteins were eluted with 20 mM imidazole buffer (20 mL), 50 mM imidazole buffer (15 mL), 100 mM imidazole elution buffer (10 mL) and 200 mM imidazole elution buffer (5 mL). The brown protein was eluted with 300 mM imidazole buffer (3 mL) and 500 mM imidazole elution buffer (2 mL) and the eluent was visibly brown in colour. Due to the inherent oxygen sensitivity of the FeS cluster, further purification and/or concentration of the protein-containing fractions using Amicon® an Ultra-15 centrifugal filter was avoided, as this would inevitably involve exposure to atmospheric oxygen. Removal of imidazole from protein-containing fractions, if necessary, was undertaken using a PD-10 column within the glove bag following the manufacturers instructions/procedure outlined in section 6.4.9.4. The buffers used were purged with argon gas prior to use. In most cases the recombinant protein was used for analysis directly after elution from the HisTrap™ affinity column.

#### 6.5.9.5 Chemical *in vitro* reconstitution of $His_6$-RedG iron-sulphur cluster

To 500 µL of protein (eluted in 500 mM elution buffer from a HisTrap™ column under anaerobic conditions and, in some cases, undergone a further PD-10 column purification step in storage buffer) was added 1 M DTT solution (250 µL). This was incubated at RT for 15 minutes. To this was added 1 M $Na_2S$ solution (1 µL) and 1 M $FeSO_4(NH_4)_2.6H_2O$ (1 µL) to instantly form a black solution which was incubated at RT for 1 hour. The reaction mixture was loaded directly onto a PD-10 column and eluted using storage buffer (minus glycerol, 1.5 mL) according to the manufacturers instructions. The brown/black protein remained visible throughout, indicating a successful reconstitution. The resulting protein was analysed by acrylamide SDS-PAGE gel electrophoresis according to the guidelines outlined previously.

#### 6.5.9.6 SDS-PAGE analysis of purified recombinant proteins

An acrylamide SDS-PAGE gel (10 %) for protein sizes 35-100 kDa was typically used throughout this work and prepared as follows;

**Table 6:** Components and amounts used to prepare SDS-PAGE resolving gels for protein analysis.

| **Resolving Gel** | |
|---|---|
| **Components** | **Volume (mL)** |
| Water | 1.90 |
| 30 % acrylamide mix | 1.70 |
| 1.5 M Tris-HCl (pH 8.8) | 1.300 |
| 10 % SDS | 0.05 |
| 10 % ammonium persulphate | 0.05 |
| TEMED | 0.002 |

**Table 7:** Components and amounts used to prepare SDS-PAGE stacking gels for protein analysis.

| **Stacking Gel** | |
|---|---|
| **Components** | **Volume (mL)** |
| Water | 1.40 |
| 30 % acrylamide mix | 0.33 |
| 1 M Tris-HCl (pH 6.8) | 0.25 |
| 10 % SDS | 0.02 |

| | |
|---|---|
| 10 % ammonium persulphate | 0.02 |
| TEMED | 0.002 |

Acrylamide SDS-PAGE loading dye (1 ×) was added to 20 µL protein samples and mixed by pipetting up and down slowly. The samples were loaded onto the gel next to a protein molecular weight marker (ladder) and electrophoresis was undertaken at 180 V for 1 hour or until good separation of the protein ladder had occurred. The gel was then stained using InstaBlue solution for 1 hour and destained with deionized water and small amount of 70 % ethanol (v/v) to visualize protein samples.

#### 6.5.9.7 Calculating protein concentration

A NanoDrop spectrophotometer was used to measure the protein absorbance at 280 nm, relative to a standard/blank using the appropriate buffer that the protein was eluted or stored in. Molecular weight of the protein was determined using SnapGene Viewer *in silico* program and the extinction coefficient was obtained form ExPASy ProtParam tool. Using these data the protein concentration was calculated;

$c = A(\lambda_{max}) / \varepsilon l$ where ($l = 1$ cm)

#### 6.5.9.8 UV-Vis analysis of recombinant $His_6$-RedG FeS cluster

The iron-sulphur cluster of $His_6$-RedG was visualized/analyzed by ultraviolet-visible spectrometry using a PerkinElmer Lambda 35 UV/VIS Spectrometer. A blank was used, typically storage buffer (minus glycerol) or 500 mM imidazole elution buffer, of equal volume. Purified recombinant protein or chemically reconstituted protein (300 µL) of concentration between 20-100 µM was added to a quartz cuvette for analysis;

**Table 8:** UV-Vis conditions used to analyze FeS cluster of recombinant $His_6$-RedG.

| Start (nm) | End (nm) | Ordinate mode | Slit width (nm) | Scan speed (nm/min) | Data interval (nm) | No. of cycles | Cycle time (s) |
|---|---|---|---|---|---|---|---|
| 700 | 200 | A | 1 | 480 | 1 | 1 | 1 |

## References


(1) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces genetics*; John Innes Foundation, **2000**.

(2) Meng, L.; Mohan, R.; Kwok, B. H.; Elofsson, M.; Sin, N.; Crews, C. M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (18), 10403–10408.

(3) Bentley, S. D.; Chater, K. F.; Cerdeño-Tárraga, A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C.-H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M.-A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. *Nature* **2002**, *417* (6885), 141–147.

(4) Corre, C.; Song, L.; O'Rourke, S.; Chater, K. F.; Challis, G. L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (45), 17510–17515.

(5) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci.* **2003**, *100* (4), 1541–1546.

(6) Gomez-Escribano, J. P.; Song, L.; Fox, D. J.; Yeo, V.; Bibb, M. J.; Challis, G. L. *Chem. Sci.* **2012**, *3* (9), 2716.

(7) Aoki, Y.; Matsumoto, D.; Kawaide, H.; Natsume, M. *J. Antibiot. (Tokyo).* **2011**, *64* (9), 607–611.

(8) Williams, R. P.; Gott, C. L.; Qadri, S. M.; Scott, R. H. *J. Bacteriol.* **1971**, *106* (2), 438–443.

(9) Pandey, R.; Chander, R.; Sainis, K. B. *Indian J. Biochem. Biophys.* **2007**, *44*, 295–302.

(10) Colotta, F.; Golay Mario Ferrari, J.; Fornasiero, C.; Alessandra Mortellaro, A.; Songia, S.; Gnocchi, P.; Mortellaro, A.; Ferrari, M.; Isetta, A.; Golay, J. *J Immunol* **1999**, *162*, 7102–7109.

(11) D'Alessio. R.; Bargiotti. A.; Carlini. O.; Colotta. F.; Ferrari. M.; Gnocchi. P.; Isetta. A.; Mongelli. N.; Motta. P.; Rossi. A.; Rossi. M.; Tibolla . M.;Vanotti. E. *J. Med. Chem*. **2000**, *43*, (13), 2557-2565

(12) Pérez-Tomás, R.; Montaner, B.; Llagostera, E.; Soto-Cerrato, V. *Biochem. Pharmacol.* **2003**, *66* (8), 1447–1452.

(13) Elahian, F.; Moghimi, B.; Dinmohammadi, F.; Ghamghami, M.; Hamidi, M.; Mirzaei, S. A. *DNA Cell Biol.* **2013**, *32* (3), 90–97.

(14) Hu, D. X.; Withall, D. M.; Challis, G. L.; Thomson, R. J. *Chem. Rev.* **2016**, *116* (14), 7818–7853.

(15) Wrede, F.; Hettche, O. *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.)* **1929**, *62* (9), 2678–2685.

(16) Williams, R. P.; Hearn, W. R. In *Biosynthesis*; Springer Berlin Heidelberg: Berlin, Heidelberg. **1967**, 410–432.

(17) Santer, U. V; Vogel, H. J. *Biochim. Biophys. Acta* **1956**, *19* (3), 578–579.

(18) Wasserman, H. H.; McKeon, J. E.; Smith, L.; Forgione, P. *J. Am. Chem. Soc.* **1960**, *82* (2), 506–507.

(19) Perry, J. J. *Nature.* **1961**, *191* (4783), 77–78.

(20) Harashima, K.; Tsuchida, N.; Tanaka, T.; Nagatsu, J. *Agric. Biol. Chem.* **1967**, *31* (4), 481–489.

(21) Wasserman, H. H.; Rodgers, G. C.; Keith, D. D. *Chem. Commun.* **1966**, *22*, 825.

(22) Gerber, N. N. *J. Antibiot. (Tokyo).* **1975**, *28* (3), 194–199.

(23) Gerber, N. N.; McInnes, A. G.; Smith, D. G.; Walter, J. A.; Wright, J. L. C.; Vining, L. C. *Can. J. Chem.* **1978**, *56* (9), 1155–1163.

(24) Gerber, N. N.; Gauthier, M. J. *Appl. Environ. Microbiol.* **1979**, *37* (6), 1176–1179.

(25) Wasserman, H. H.; Rodgers, G. C.; Keith, D. D. *J. Am. Chem. Soc.* **1969**, *91* (5), 1263–1264.

(26) Hu, D. X.; Clift, M. D.; Lazarski, K. E.; Thomson, R. J. *J. Am. Chem. Soc.* **2011**, *133* (6), 1799–1804.

(27) Haynes, S. W.; Sydor, P. K.; Corre, C.; Song, L.; Challis, G. L. *J. Am. Chem. Soc.* **2011**, *133* (6), 1793–1798.

(28) Sydor, P. K.; Barry, S. M.; Odulate, O. M.; Barona-Gomez, F.; Haynes, S. W.; Corre, C.; Song, L.; Challis, G. L. *Nat. Chem.* **2011**, *3* (5), 388–392.

(29) Clift, M. D.; Thomson, R. J. *J. Am. Chem. Soc.* **2009**, *131* (40), 14579–14583.

(30) Gerber, N. N. *Appl. Environ. Microbiol.* **1969**, *18* (1).

(31) Hayakawa, Y.; Kawakami, K.; Seto, H.; Furihata, K. *Tetrahemon Lett.* **1992**, *33* (19), 2701–2704.

(32) Kawasaki, T.; Sakurai, F.; Hayakawa, Y. *J. Nat. Prod.* **2008**, *71* (7), 1265–1267.

(33) Boonlarppradab, C.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. *Org. Lett.* **2008**, *10* (24), 5505–5508.

(34) Salem, S. M.; Kancharla, P.; Florova, G.; Gupta, S.; Lu, W.; Reynolds, K. A. *J. Am. Chem. Soc.* **2014**, *136* (12), 4565–4574.

(35) Kancharla, P.; Lu, W.; Salem, S. M.; Kelly, J. X.; Reynolds, K. A. *J. Org. Chem.* **2014**, *79* (23), 11674–11689.

(36) Picott, K. J.; Deichert, J. A.; Dekemp, E. M.; Schatte, G.; Sauriol, F.; Ross, A. C. *Med. Chem. Commun.* **2019**, *10*, 478.

(37) Marks, G. S.; Bogorad, L. *Proc. Natl. Acad. Sci. U. S. A.* **1960**, *46* (1), 25–28.

(38) Wasserman, H. H.; Shaw, C. K.; Sykes, R. J.; Cushley, R. J. *Tetrahedron Lett.* **1974**, *15* (33), 2787–2790.

(39) Wasserman, H. H.; Sykes, R. J.; Peverada, P.; Shaw, C. K.; Cushley, R. J.; Lipsky, S. R. *J. Am. Chem. Soc.* **1973**, *95* (20), 6874–6875.

(40) Ruddt A N, A. N. A. M.; Hopwood, D. A. *J. Generui Microbiol.* **1980**, *119*, 333–340.

(41) Mo, S.; Sydor, P. K.; Corre, C.; Alhamadsheh, M. M.; Stanley, A. E.; Haynes, S. W.; Song, L.; Reynolds, K. A.; Challis, G. L. *Chem. Biol.* **2008**, *15* (2), 137–148.

(42) Stanley, A. E.; Walton, L. J.; Kourdi Zerikly, M.; Corre, C.; Challis, G. L.; Fürtsner, A.; Dairi, K.; Tripathy, S.; Attardo, G.; Lavallée, J.-F.; Mortellaro, A.; Songia, S.; Gnocchi, P.; Ferrari, M.; Fornasiero, C.; D'Alessio, R.; Isetta, A.; Colotta, F.; Golay, J.; Wasserman, H. H.; Rodgers, G. C.; Keith, D. D.; Wasserman, H. H.; Keith, D. D.; Rodgers, G. C.; Wasserman, H. H.; Shaw, C. K.; Sykes, R. J.; Tsao, S. W.; Rudd, B. A.; He, X. G.; Chang, C. J.; Floss, H. G.; Rudd, B. A. M.; Hopwood, D. A.; Feitelson, J. S.; Hopwood, D. A.; Feitelson, J. S.; Malpartida, F.; Hopwood, D. A.; Cerdeno, A. M.; Bibb, M. J.; Challis, G. L.; Thomas, M. G.; Burkart, M. D.; Walsh, C. T.; Mo, S. J.; Kim, B. S.; Reynolds, K. A.; Williamson, N. R.; Simonsen, H. T.; Ahmed, R. A. A.; Goldet, G.; Slater, H.; Woodley, L.; Leeper, F. J.; Salmond, G. P. C.; Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F.; Kerbarh, O.; Campopiano, D. J.; Baxter, R. L.; Luo, G.; Pieper, R.; Rosa, A.; Khosla, C.; Cane, D. E.; Ehmann, D. E.; Trauger, J.; Stachelhaus, T.; Walsh, C. T.; Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Khosla, C.; Walsh, C. T.; Cox, R. J.; Crosby, J.; Daltrop, O.; Glod, F.; Jarzabek, M. E.; Nicholson, T. P.; Reed, M.; Simpson, T. J.; Smith, L. H.; Soulas, F.; Szafranska, A. E.; Westcott, J.; Weissman, K. J.; Hong, H.; Oliynyk, M.; Siskos, A. P.; Leadlay, P. F. *Chem. Commun.* **2006**, *42* (38), 3981–3983.

(43) Whicher, J. R.; Florova, G.; Sydor, P. K.; Singh, R.; Alhamadsheh, M.; Challis, G. L.; Reynolds, K. A.; Smith, J. L. *J. Biol. Chem.* **2011**, *286*, 22558-22569.

(44) Haynes, S. W.; Sydor, P. K.; Stanley, A. E.; Song, L.; Challis, G. L.; Mortellaro, A.; Songia, S.; Gnocchi, P.; Ferrari, M.; Fornasiero, C.; D'Alessio, R.; Isetta, A.; Colotta, F.; Golay, J.; Gerber, N. N.; Isaka, M.; Jaturapat, A.; Kramyu, J.; Tanticharoen, M.; Thebtaranonth, Y.; Montaner, B.; Perez-Tomas, R.; Fürtsner, A.; Nguyen, M.; Marcellus, R. C.; Roulston, A.; Watson, M.; Serfass, L.; Madiraju, S. R. M.; Goulet, D.; Viallet, J.; Belec, L.; Billot, X.; Acoca, S.; Purisima, E.; Wiegmans, A.; Cluse, L.; Johnstone, R. W.; Beauparlant, P.; Shore, G. C.; Tsao, S. W.; Rudd, B. A.; He, X. G.;

Chang, C. J.; Floss, H. G.; Mo, S. J.; Sydor, P. K.; Corre, C.; Alhamadsheh, M. M.; Stanley, A. E.; Haynes, S. W.; Song, L.; Reynolds, K. A.; Challis, G. L.; Cerdeno, A. M.; Bibb, M. J.; Challis, G. L.; Stanley, A. E.; Walton, L. J.; Kourdi-Zerikly, M.; Corre, C.; Challis, G. L.; Thomas, M. G.; Burkart, M. D.; Walsh, C. T.; Garneau-Tsodikova, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T.; Williamson, N. R.; Simonsen, H. T.; Ahmed, R. A. A.; Goldet, G.; Slater, H.; Woodley, L.; Leeper, F. J.; Salmond, G. P. C.; Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F.; Dairi, K.; Tripathy, S.; Attardo, G.; Lavallée, J.-F.; Kayakawa, Y.; Kawakami, K.; Seto, H.; Furihata, K.; Wasserman, H. H.; Rodgers, G. C.; Keith, D. D. *Chem. Commun.* **2008**, *162* (16), 1865.

(45) Chawrai, S. R.; Williamson, N. R.; Mahendiran, T.; Salmond, G. P. C.; Leeper, F. J. *Chem. Sci.* **2012**, *3*, 447-454.

(46) Barry, S. M.; Challis, G. L. *ACS. Catal.* **2013**, *3*, (10), 2362-2370.

(47) Zhu, Y.; Jameson, E.; Crosatti, M.; Schäfer, H.; Rajakumar, K.; Bugg, T. D. H.; Chen, Y. *PNAS.* **2014**, *111*, (11), 4268-4273.

(48) Salih, R. Investigations of Carbocyclisation Reactions in Prodiginine Alkaloid Biosynthesis, *PhD Thesis*, University of Warwick, **2017**.

(49) Withall, D. M.; Haynes, S. W.; Challis, G. L. *J. Am. Chem. Soc.* **2015**, *137* (24), 7889–7897.

(50) Kawasaki, T.; Sakurai, F.; Hayakawa, Y. *J. Nat. Prod.* **2008**, *71* (7), 1265-1267.

(51) Hayakawa, Y.; Nagatsuka, S.; Kawasaki, T. *J. Antibiot. (Tokyo).* **2009**, *62* (9), 531–532.

(52) Kawasaki, T.; Sakurai, F.; Nagatsuka, S.; Hayakawa, Y. *J. Antibiot. (Tokyo).* **2009**, *62* (5), 271–276.

(53) Perry, C.; De Los Santos, E. C.; Alkhalaf, L. M.; Challis, G. L. *Nat. Prod. Rep.* **2018**, *35* (7).

(54) Lee, J.; Simurdiak, M.; Zhao, H. *J. Biol. Chem.* **2005**, *280*, (44), 36719-36727.

(55) Pistorius, D.; Li, Y.; Sandmanny, A.; Muïler, R. *Mol. BioSyst. Mol. BioSyst* **2011**, *7* (7), 3308–3315.

(56) Fei. X.; Yin. X.; Zhang. L.; Zabriskie. M. T. *J. Nat. Prod.* **2007**, *70*, (4), 618-622.

(57) Li, W.; Estrada-de los Santos, P.; Matthijs, S.; Xie, G.-L.; Busson, R.; Cornelis, P.; Rozenski, J.; De Mot, R. *Chem. Biol.* **2011**, *18* (10), 1320–1330.

(58) Chang, F.-Y.; Brady, S. F. *ChemBioChem* **2014**, *15* (6), 815–821.

(59) Julien, B.; Tian, Z.-Q.; Reid, R.; Reeves, C. D. *Chem. Biol.* **2006**, *13*, 1277–1286.

(60) Wang, L.; Parnell, A.; Williams, C.; Bakar, N. A.; Challand, M. R.; van der Kamp, M. W.; Simpson, T. J.; Race, P. R.; Crump, M. P.; Willis, C. L. *Nat. Catal.* **2018**, *1* (12), 968–976.

(61) Stiefel, E. I.; George, G. N. Ferredoxins, hydrogenases, and Notrogenases: Metal-Sulfide Proteins. *Bioinorg. Chem.* Bertini, I.; Ed; University Science Books: Mill Valley. **1994**, 365-453.

(62) Beinert, H. *J. Biol. Inorg. Chem.* **2000**, *5* (1), 2–15.

(63) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. *Annu. Rev. Biochem.* **2005**, *74* (1), 247–281.

(64) Bian, S.; Cowan, J. A. *Coord. Chem. Rev.* **1999**, *190-192*, 1049–1066.

(65) Glaser, T.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *Acc. Chem. Res.* **2000**, *33* (12), 859–868.

(66) Mortenson, L. E.; Valentine, R. C.; Carnahan, J. E. *Biochem. Biophys. Res. Commun.* **1962**, *7* (6), 448–452.

(67) Tagawa, K.; Arnon, D. I. *Nature.* **1962**, *195* (4841), 537–543.

(68) Mapolelo, D. T.; Zhang, B.; Naik, S. G.; Huynh, B. H.; Johnson, M. K. *Biochemistry* **2012**, *51* (41), 8071.

(69) Lukowski, A. L.; Ellinwood, D. C.; Hinze, M. E.; DeLuca, R. J.; Du Bois, J.; Hall, S.; Narayan, A. R. H. *J. Am. Chem. Soc.* **2018**, *140* (37), 11863–11869.

(70) Venkateswara Rao, P.; Holm, R. H. *Chem. Rev.* **2004**, *104* (2), 527–560.

(71) Beinert, H.; Holm, R. H.; Münck, E. *Science* **1997**, *277* (5326), 653–659.

(72) Colbert, C. L.; Couture, M. M.-J.; Eltis, L. D.; Bolin, J. T. *Structure* **2000**, *8* (12), 1267–1278.

(73) Thomson, A. J. *Biochem. Soc. Trans.* **1975**, *3* (4), 468.3–472.

(74) Grazina, R.; Pauleta, S. R.; Moura, J. J. G.; Moura, I. In *Comprehensive Inorganic Chemistry II. Elsevier*. **2013**, 103–148.

(75) Meyer, J. *JBIC J. Biol. Inorg. Chem.* **2008**, *13* (2), 157–170.

(76) Lebrun, E.; Santini, J. M.; Brugna, M.; Ducluzeau, A.-L.; Ouchane, S.; Schoepp-Cothenet, B.; Baymann, F.; Nitschke, W. *Mol. Biol. Evol.* **2006**, *23* (6), 1180–1191.

(77) Kim, J. H.; Füzéry, A. K.; Tonelli, M.; Ta, D. T.; Westler, W. M.; Vickery, L. E.; Markley, J. L. *Biochemistry* **2009**, *48* (26), 6062–6071.

(78) Kauppi, B.; Lee, K.; Carredano, E.; Parales, R. E.; Gibson, D. T.; Eklund, H.; Ramaswamy, S. *Structure* **1998**, *6* (5), 571–586.

(79) Chakrabarty, S.; Austin, R. N.; Deng, D.; Groves, J. T.; Lipscomb, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 3514–3515.

(80) Ferraro, D. J.; Gakhar, L.; Ramaswamy, S. Biochem. *Biophys. Res. Com*. **2005**, *338*, (1), 175-190.

(81) Parales, R. E.; Parales, J. V; Gibson, D. T. *J. Bacteriol.* **1999**, *181* (6), 1831–1837.

(82) Karlsson, A.; Parales, J. V; Parales, R. E.; Gibson, D. T.; Eklund, H.; Ramaswamy, S. *Science* **2003**, *299* (5609), 1039–1042.

(83) Ohta. T.; Chakrabarty. S.; Lipscomb. J. D.; Solomon. E. I. *J. Am. Chem. Soc*. **2008**, *130*, (5), 1601-1610.

(84) Abu-Omar, M. M.; Loaiza, A.; Hontzeas, N. *Chem. Rev*. **2005**, *105*, (6), 2227-2252.

(85) Costas. M.; Mehn. M. P.; Jensen. M. P.; Que. L. J. *Chem. Rev.* **2004**, *104*, (2), 939-986.

(86) Bugg, T. D.; Ramaswamy S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 134–140.

(87) Ferraro, D. J.; Gakhar, L.; Ramaswamy, S. *Biochem. Biophys. Res. Commun.* **2005**, *338* (1), 175–190.

(88) Haynes, S. Mutasynthesis approaches to the preparation of streptorubin B analogues, *Phd Thesis,* University of Warwick, **2010**.

(89) Withall, D. Mechanistic investigation of the RedG-catalysed oxidative carbocyclisation of undecylprodigiosin to streptorubin B, *Phd Thesis*, University of Warwick, **2014**.

(90) Papireddy, K.; Smilkstein, M.; Kelly, J. X.; Shweta; Salem, S. M.; Alhamadsheh, M.; Haynes, S. W.; Challis, G. L.; Reynolds, K. A. *J. Med. Chem.* **2011**, *54* (15), 5296–5306.

(91) Slocum, D. W.; Jennings10, C. A. *Directed Metalation Reactions. 6.la Competition of Substituents for Ortho Direction of Metalation in Substituted Anisóles*; UTC, **1976**, *41*

(92) Hopwood, D. A. *Natural product biosynthesis by microorganisms and plants. Elsevier*. **2012**, *516 (Part B).*

(93) Kallifidas, D.; Jiang, G.; Ding, Y.; Luesch, H. *Microb. Cell Fact.* **2018**, *17*, 25.

(94) Baltz, R. H. *J. Ind. Microbiol. Biotechnol.* **2010**, *37* (8), 759–772.

(95) Biggins, J. B.; Gleber, C. D.; Brady, S. F. *Org. Lett.* **2011**, *13* (6), 1536–1539.

(96) Kallifidas, D.; Kang, H.-S.; Brady, S. F. *J. Am. Chem. Soc*. **2012**, *134*, (48), 19552-19555.

(97) Liu, A.; N Ho, R. Y.; Que, L.; Ryle, M. J.; Phinney, B. S.; Hausinger, R. P. *Biol. Appl. Raman Spectrosc.* **1999**, *9* (2), 5126–5127.

(98) Que, L.; Ho, R. Y. N. *Chem. Rev.* **1996**, *96* (7), 2607–2624.

(99) Kouprina, N.; Larionov, V. *Nat. Rev. Genet.* **2006**, *7* (10), 805–812.

(100) Kouprina, N.; Larionov, V. *Chromosoma* **2016**, *125* (4), 621–632.

(101) Schneiter, R. *Genetics, Molecular and Cell Biology of Yeast*; **2004**.

(102) Bhargava, J.; Shashikant, C. S.; Carr, J. L.; Juan, H.; Bentley, K. L.; Ruddle, F. H. *Genomics* **1999**, *62* (2), 285–288.

(103) Annab, L. A.; Kouprina, N.; Solomon, G.; Cable, P. L.; Hill, D. E.; Barrett, J. C.; Larionov, V.; Afshari, C. A. *Gene* **2000**, *250* (1-2), 201–208.

(104) Kouprina, N.; Annab, L.; Graves, J.; Afshari, C.; Barrett, J. C.; Resnick, M. A.; Larionov, V. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (8), 4469–4474.

(105) Nihei, N.; Kouprina, N.; Larionov, V.; Oshima, *J. Cancer. Res*. **2002**, *62*, 367-370.

(106) Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. *PNAS*. **2014**, *111*, (5), 1957-1962.

(107) Lin-Chao, S.; Chen, W.-T.; Wong, T.-T. *Mol. Microbiol.* **1992**, *6* (22), 3385–3393.

(108) Bonet, B.; Teufel, R.; Crüsemann, M.; Ziemert, N.; Moore, B. S. *J. Nat. Prod.* **2015**, *78* (3), 539–542.

(109) Lee, T.; Bradley, M. E.; Walowitz, J. L. *Nucleic Acids Res.* **1998**, *26* (13), 3215–3220.

(110) Flynn, P. J.; Reece, R. J. *Mol. Cell. Biol.* **1999**, *19* (1), 882.

(111) Ross, A. C.; Gulland, L. E. S.; Dorrestein, P. C.; Moore, B. S. *ACS. Synth. Biol.* **2015**, *4*, (4), 414-420.

(112) Chen, K.; Rannulu, N. S.; Cai, Y.; Lane, P.; Liebl, A. L.; Rees, B. B.; Corre, C.; Challis, G. L.; Cole, R. B. *JAM.* **2008**, *19*, 1856–1866.

(113) Brown, C. A.; Yamashita, A. *J. Am. Chem. Soc.* **1975**, *97* (4), 891–892.

(114) Brandsma, L.; Heus-Kloos, Y. A.; Heiden, R. van der.; Verkruijsse, H. D. *Preparative acetylenic chemistry*. *Elsevier*. **1988**.

(115) Lipshutz, B. H.; Wilhelm, R. S.; Floyd, D. M. *J. Am. Chem. Soc.* **1981**, *103* (25), 7672–7674.

(116) Lipshutz, B. H.; Wilhelm, R. S.; Kozlowski, J. A.; Parker, D. *J. Org. Chem.* **1984**, *49* (21), 3928–3938.

(117) Dale, J. A.; Mosher, H. S. *J. Am. Chem. Soc.* **1973**, *95* (2), 512–519.

(118) Hoye, T. R.; Jeffrey, C. S.; Shao, F. *Nat. Protoc.* **2007**, *2* (10), 2451–2458.

(119) Basak, A.; Salowe, S.; Townsend. C. A. *Biochemistry*. **1990**, *29*, (27), 6499-6508.

(120) Roach, P. L.; Clifton, I. J.; Fülöp, V.; Harlos, K.; Barton, G. J.; Hajdu, J.; Andersson, I.; Schofield, C. J.; Baldwin, J. E. *Nature.* **1995**, *375* (6533), 700–704.

(121) Long, A. J.; Clifton, I. J.; Roach, P. L.; Baldwin, J. E.; Rutledge, P. J.; Schofield, C. J. *Biochemistry.* **2005**, *44* (17), 6619–6628.

(122) Ogle, J. M.; Clifton, I. J.; Rutledge, P. J.; Elkins, J. M.; Burzlaff, N. I.; M. Adlington, R.; Roach, P. L.; Baldwin, J. E. *Chem. Biol.* **2001**, *8* (12), 1231–1237.

(123) Baldwin, J. E.; Adlington, R. M.; Domayne-Hayman, B. P.; Ting, H.-H.; Turner, N. J. *J. Chem. Soc., Chem. Commun.* **1986**, No. 2, 110–113.

(124) Zhou, S.; Alkhalaf, L. M.; de los Santos, E. L.; Challis, G. L. *Curr. Opin. Chem. Biol.* **2016**, *35*, 73–79.

(125) Bauerle, M. R.; Schwalm, E. L.; Booker, S. J. *J. Biol. Chem.* **2015**, *290* (7), 3995–4002.

(126) Lin, H. *Bioorg. Chem.* **2011**, *39* (5-6), 161–170.

(127) Hou, Y.; Tianero, M. D. B.; Kwan, J. C.; Wyche, T. P.; Michel, C. R.; Ellis, G. A.; Vazquez-Rivera, E.; Braun, D. R.; Rose, W. E.; Schmidt, E. W.; Bugni, T. S. *Org.*

*Lett.* **2012**, *14* (19), 5050–5053.

(128) Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L. *Chem. Sci.* **2012**, *3* (12), 3522.

(129) Huo, L.; Rachid, S.; Stadler, M.; Wenzel, S. C.; Müller, R. *Chem. Biol.* **2012**, *19* (10), 1278–1287.

(130) Crone, W. J. K.; Leeper, F. J.; Truman, A. W. *Chem. Sci.* **2012**, *3* (12), 3516.

(131) Huang, C.; Huang, F.; Moison, E.; Guo, J.; Jian, X.; Duan, X.; Deng, Z.; Leadlay, P. F.; Sun, Y. *Chem. Biol.* **2015**, *22* (2), 251–261.

(132) Cryle, M. J.; Stuthe, J. M. U.; Ortiz de Montellano, P. R.; De Voss, J. J. *Chem. Commun. (Camb).* **2004**, No. 5, 512–513.

(133) Metzger, J. *Synthesis*. **1992**

(134) Bouzide, A.; Sauvé, G. *Tetrahedron Lett.* **1997**, *38* (34), 5945–5948.

(135) Sugaya. N.; Sakurai. K.; Nakagawa. T.; Onda. N.; Onodera. S.; Morita. M.; Tezuka. M. *Anal. Sci.* **2004**, *20* (5), 865–870.

(136) Buckel, W.; Golding, B. T. *Annua. Rev. Microbiol.* **2006**, *60*, 27-49.

(137) Karlsson, A.; Parales, J. V; Parales, R. E.; Gibson, D. T.; Eklund, H.; Ramaswamy, S. *J. Inorg. Biochem.* **2000**, *78* (1), 83–87.

(138) Jaganaman, S.; Pinto, A.; Tarasev, M.; Ballou, D. P. *Protein Expr. Purif.* **2007**, *52* (2), 273–279.

(139) Fee. J.; Findling. K.; Yoshida. T.; Hille. R.; Tarr. G. E.; Hearson. D. O.; Dunham. W. R.; Day. E. P.; Kent. T. A.; Munck. E. *J. Biol. Chem*. **1984**, *259*, 124-133.

(140) Johnson, D. C.; Unciuleac, M.-C.; Dean, D. R. *J. Bacteriol.* **2006**, *188* (21), 7551–7561.

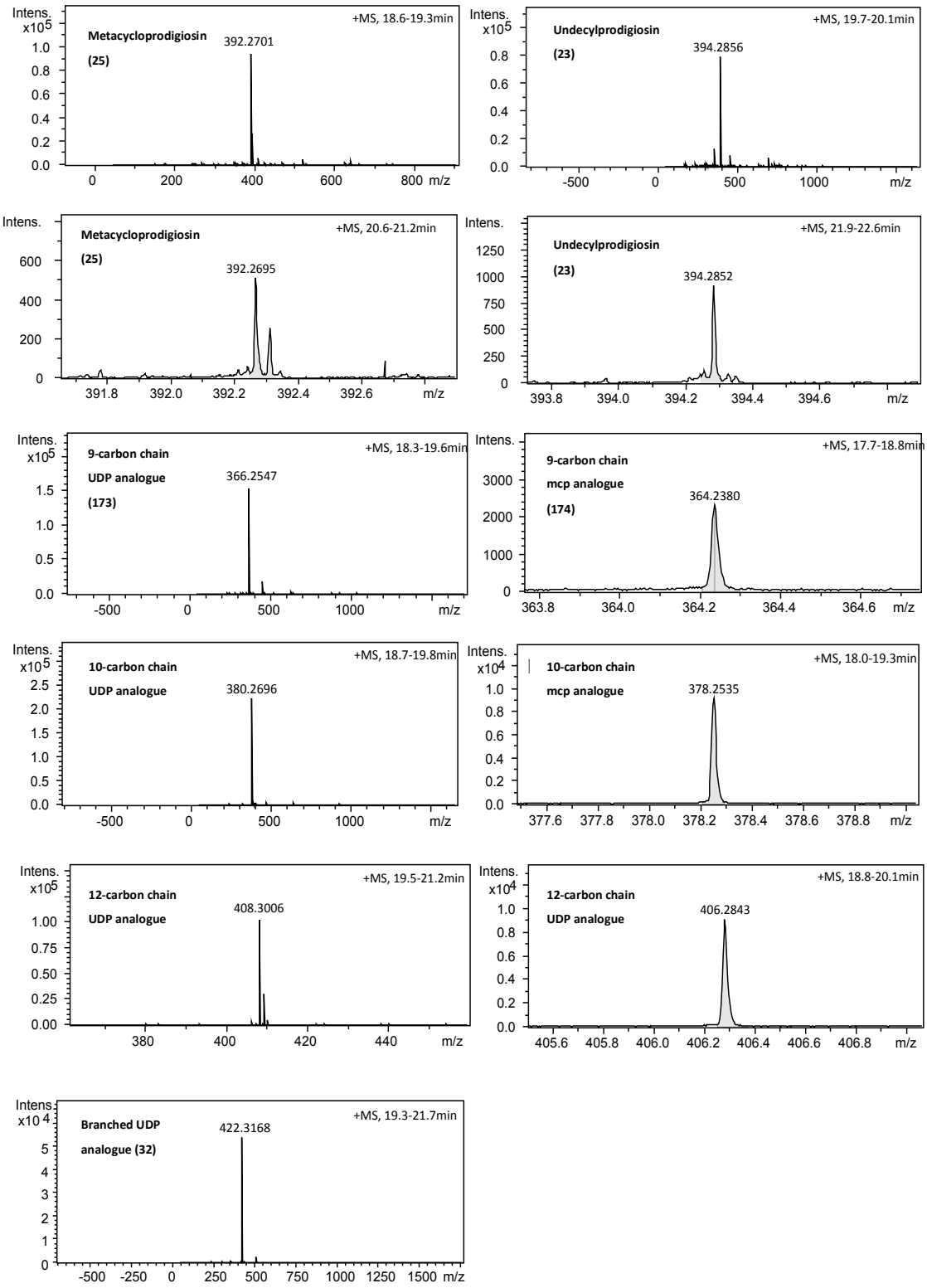
(141) Marblestone, J. G.; Edavettal, S. C.; Lim, Y.; Lim, P.; Zuo, X.; Butt, T. R. *Protein Sci.* **2006**, *15* (1), 182–189.

(142) Lee, C.-D.; Sun, H.-C.; Hu, S.-M.; Chiu, C.-F.; Homhuan, A.; Liang, S.-M.; Leng, C.-H.; Wang, T.-F. *Protein Sci.* **2008**, *17* (7), 1241–1248.

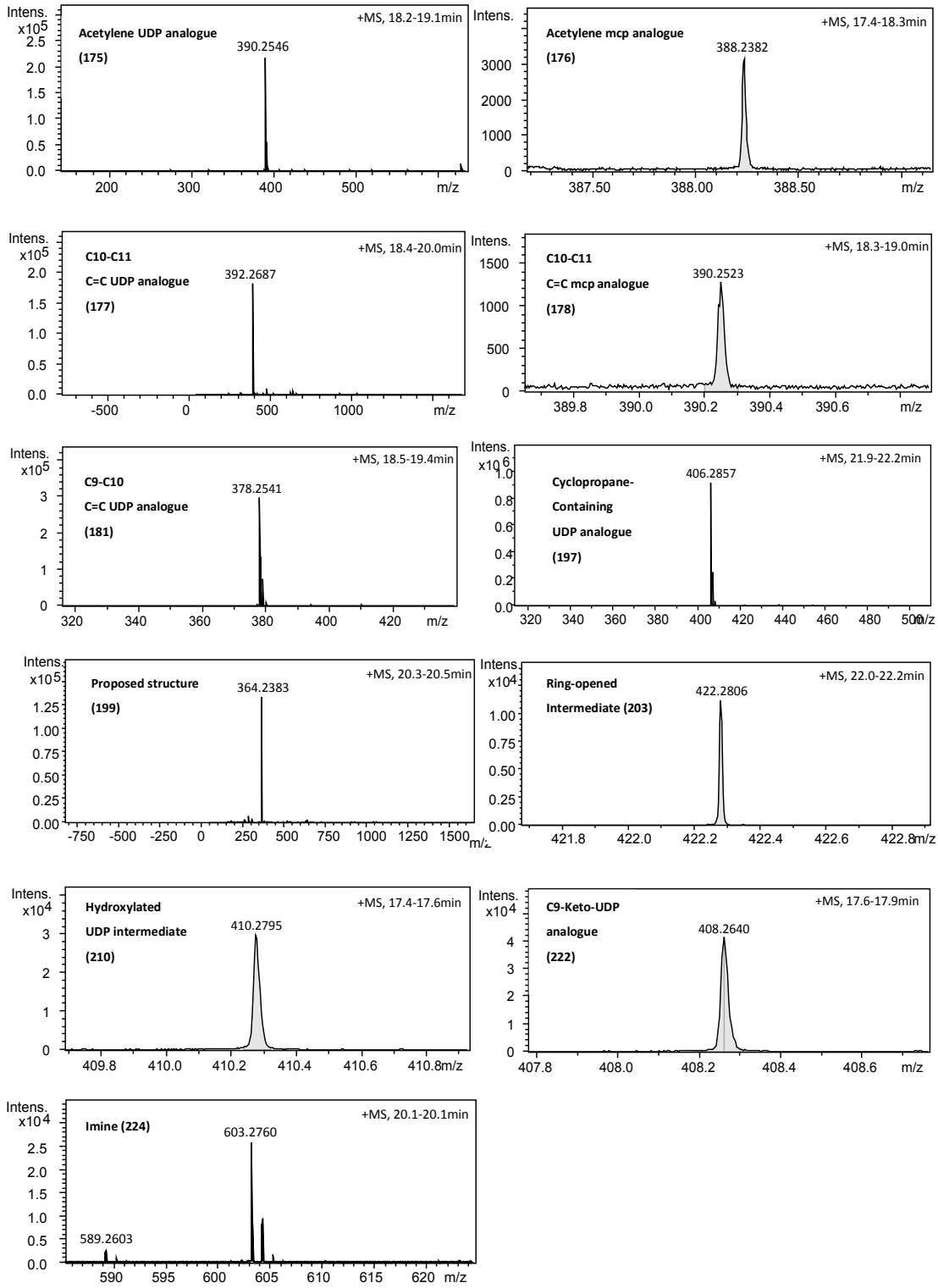
(143) Raghavan, S.; Krishnaiah, V. *J. Org. Chem.* **2010**, *75* (3), 748–761.

# Appendix

## Mass spectra

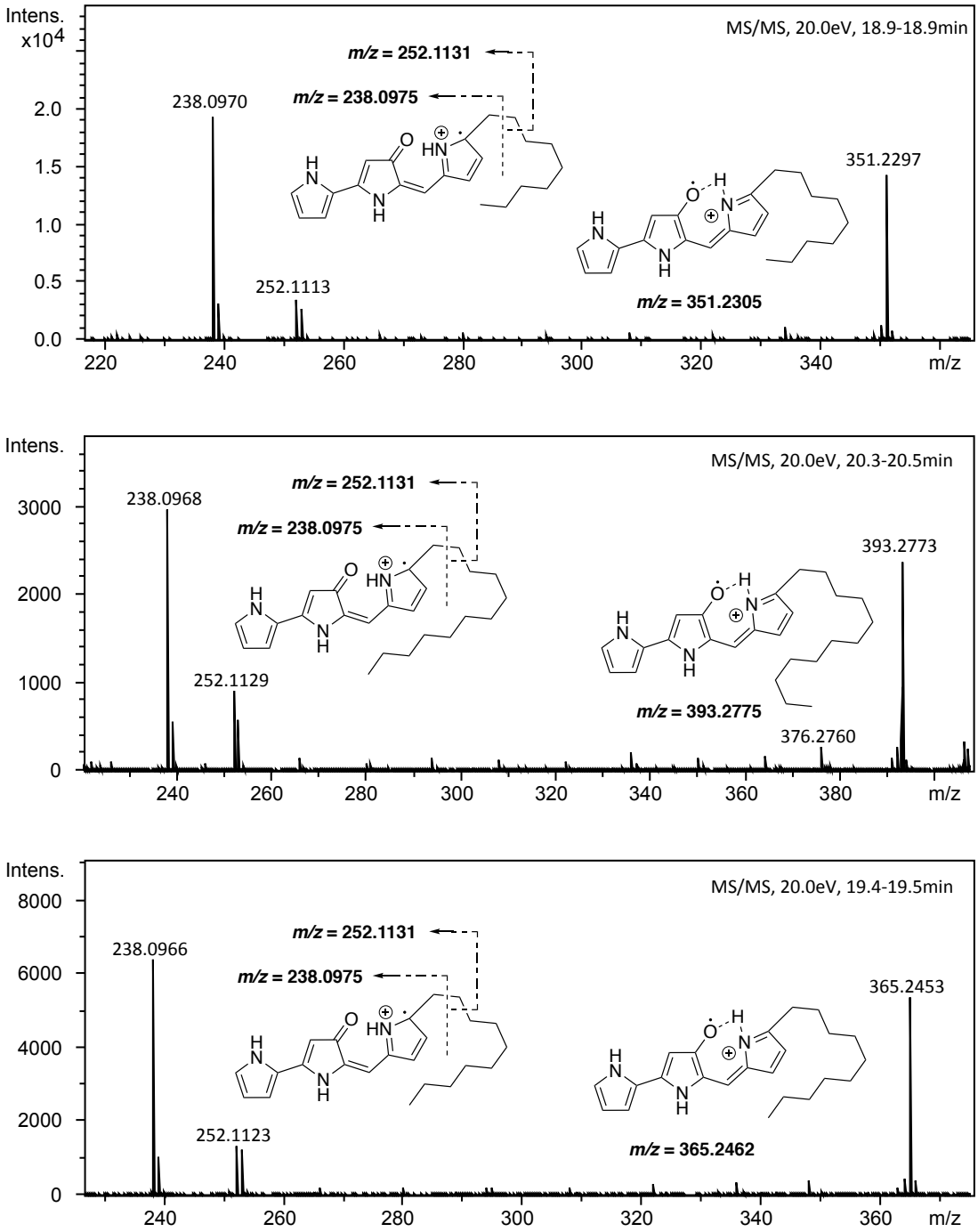



**Appendix 1:** Mass spectra of compounds observed by LC-MS in this work.

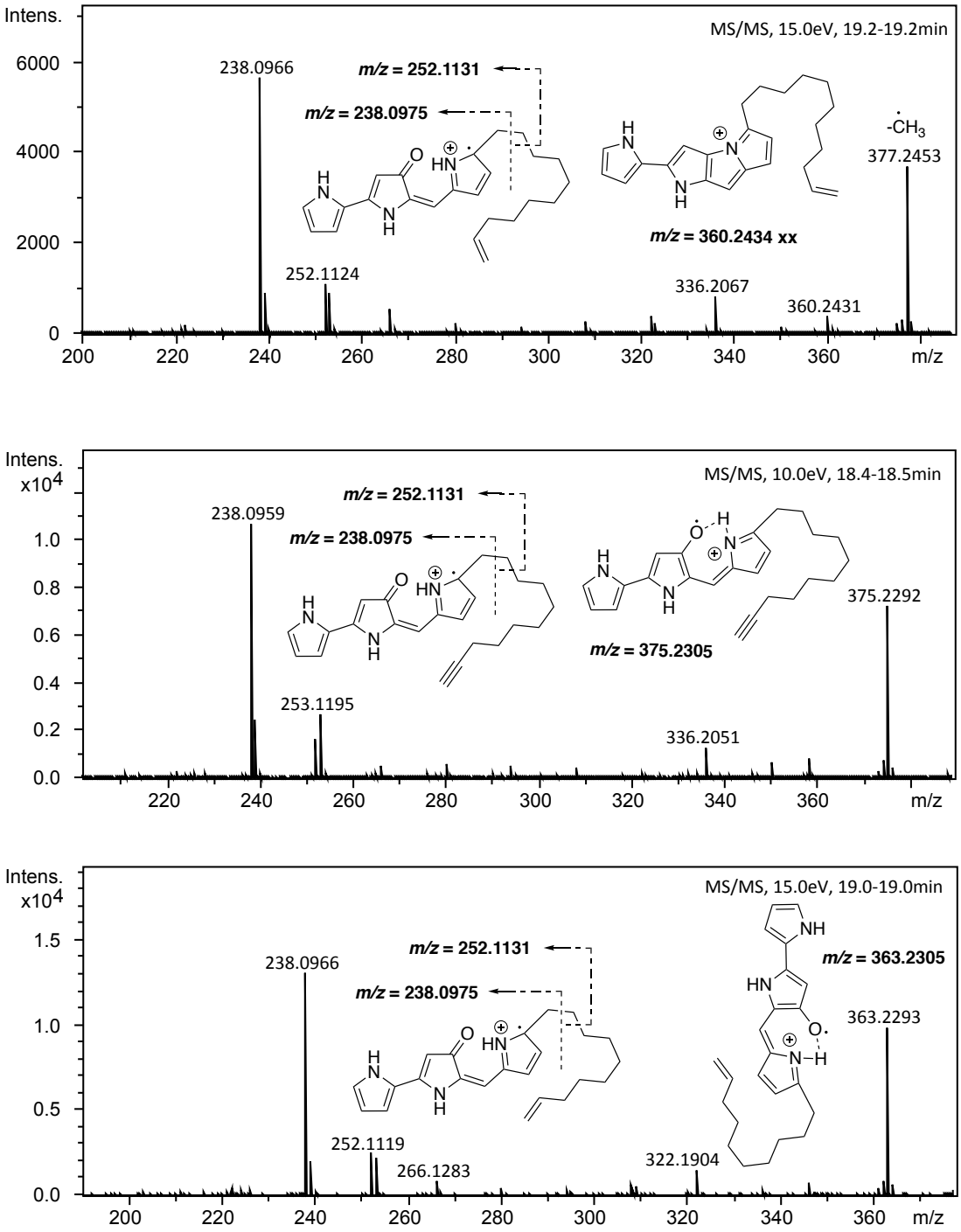



**Appendix 2:** Mass spectra of compounds observed by LC-MS in this work.

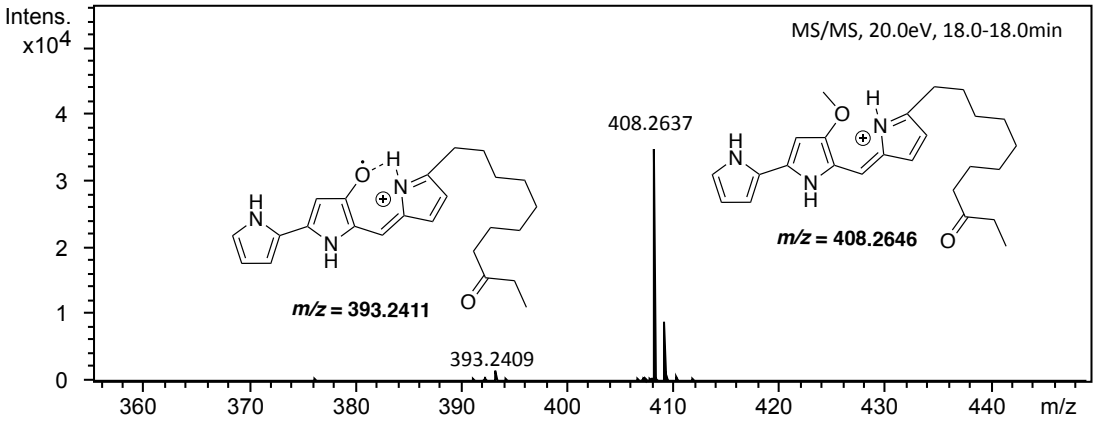
## MS/MS Spectra

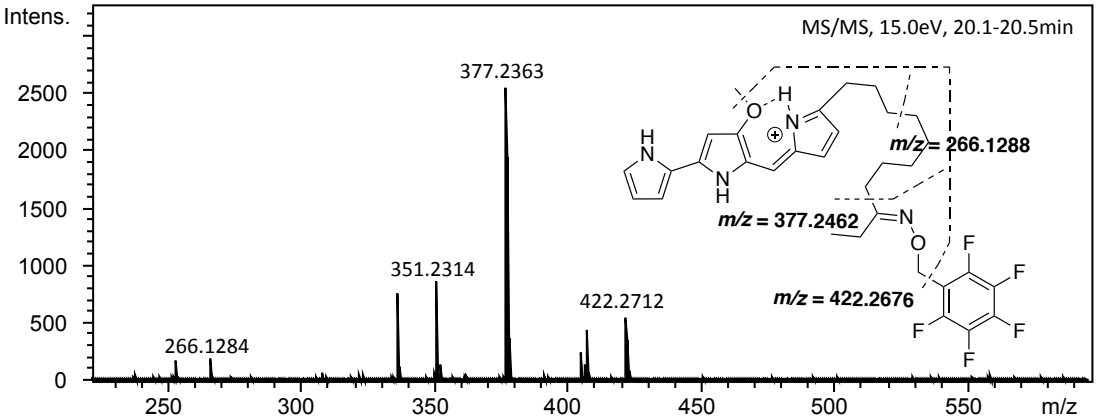



**Appendix 3:** MS/MS spectra of compounds observed by LC-MS in this work.



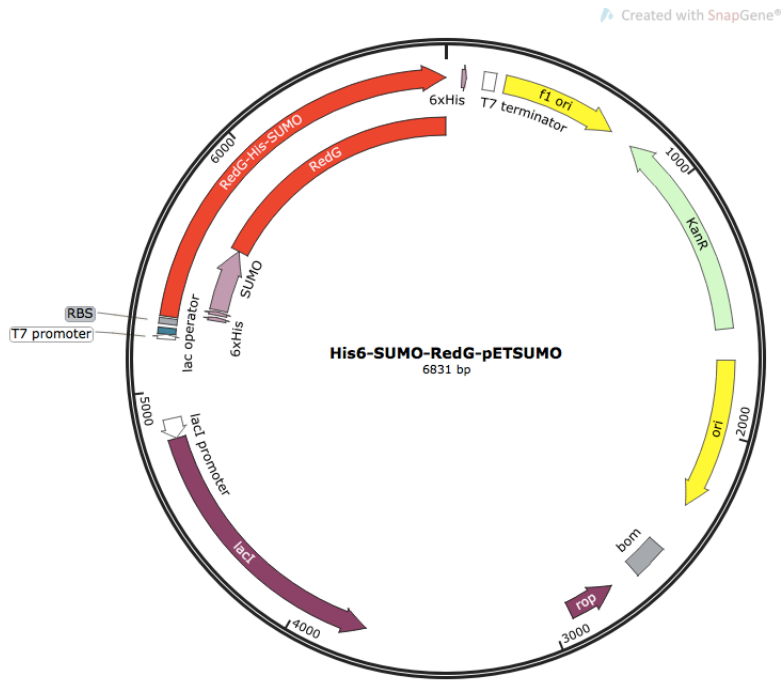

**Appendix 4:** MS/MS spectra for compounds observed by LC-MS in this work.

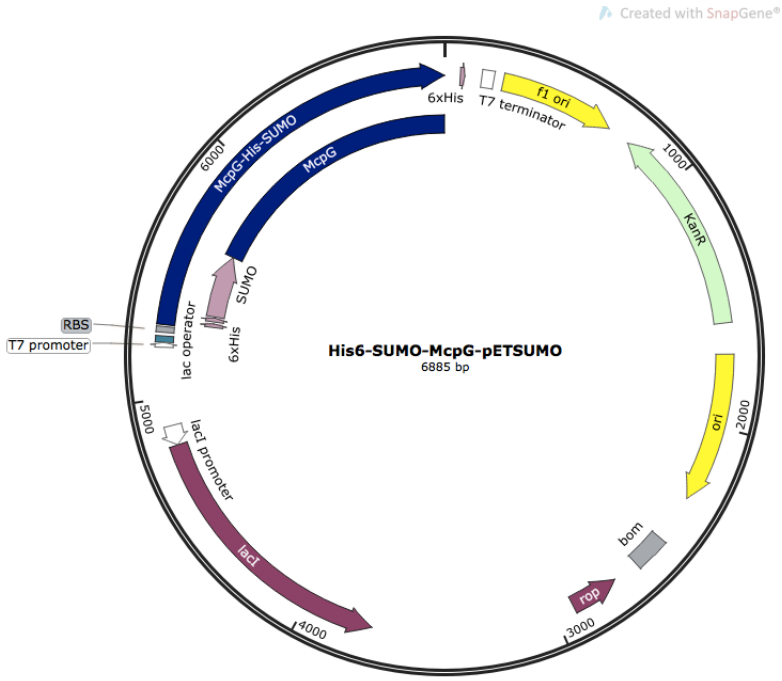


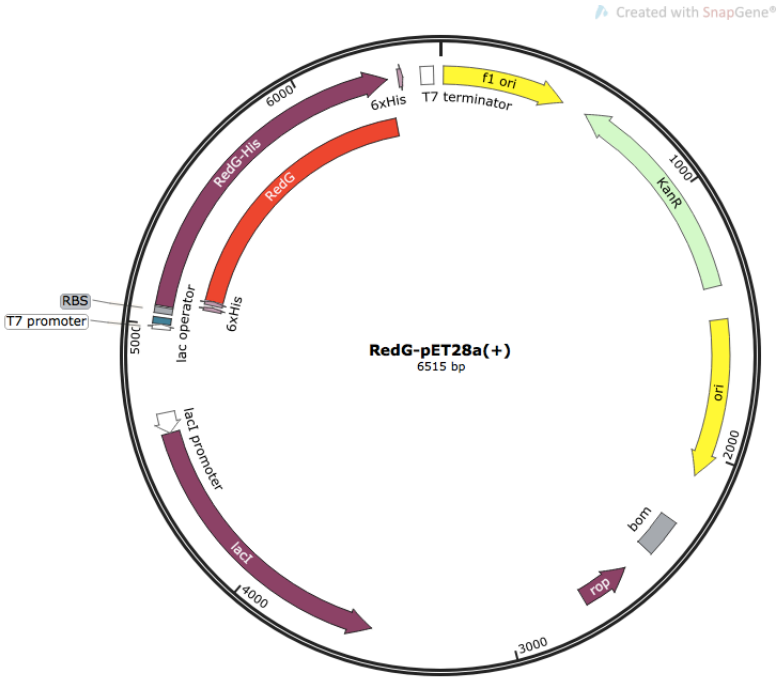

**Appendix 5:** MS/MS spectra for compounds observed by LC-MS in this work.

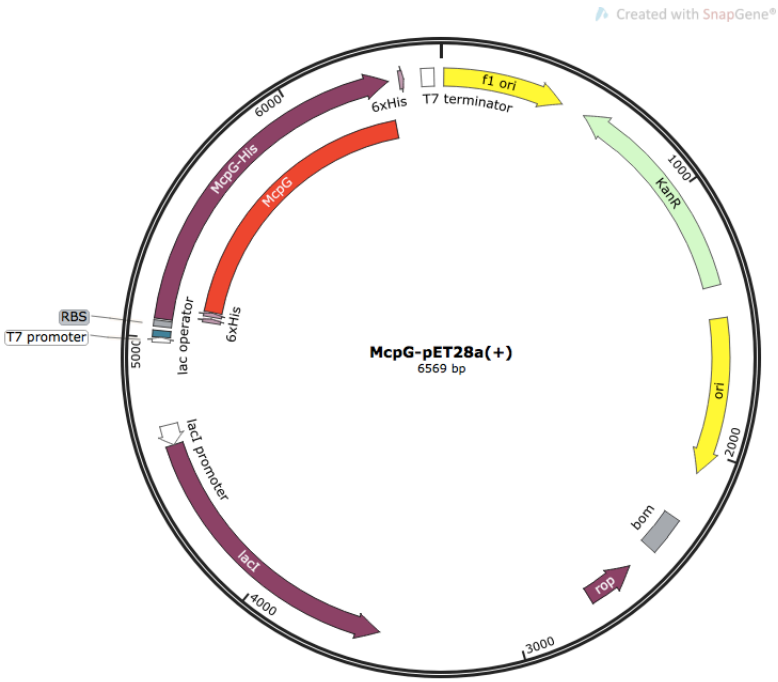
## Plasmid Maps



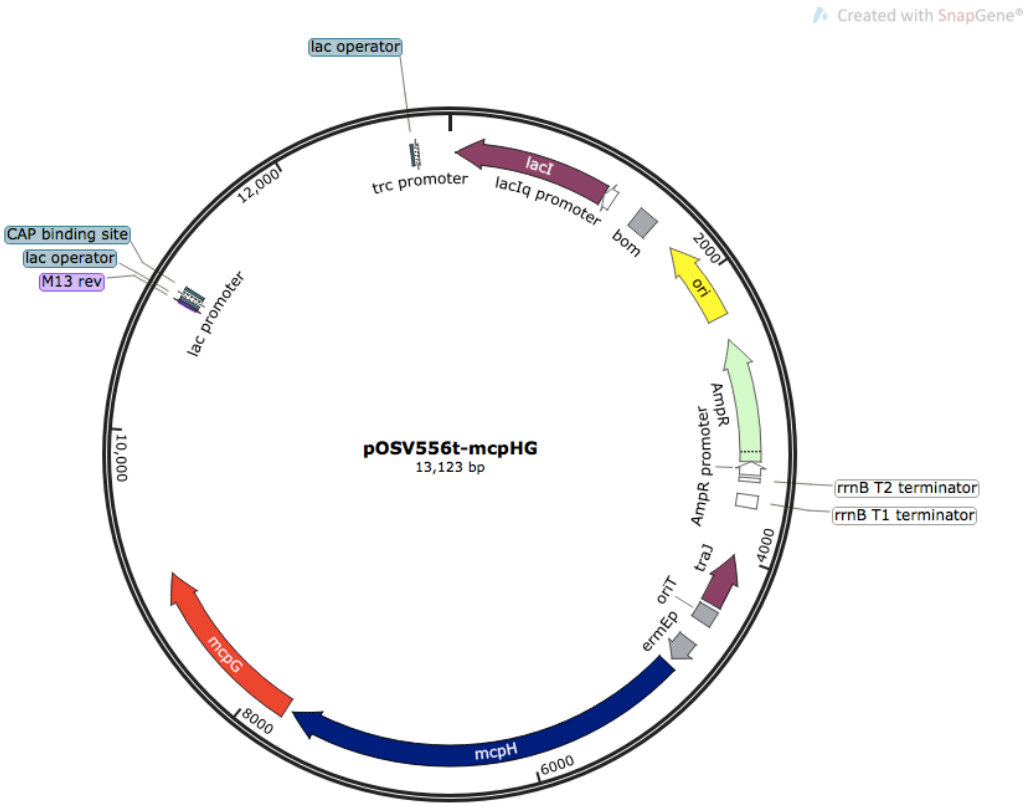




**Appendix 6:** Plasmid aps of pETSUMO/*redG* (top) and pETSUMO/*mcpG* (bottom).

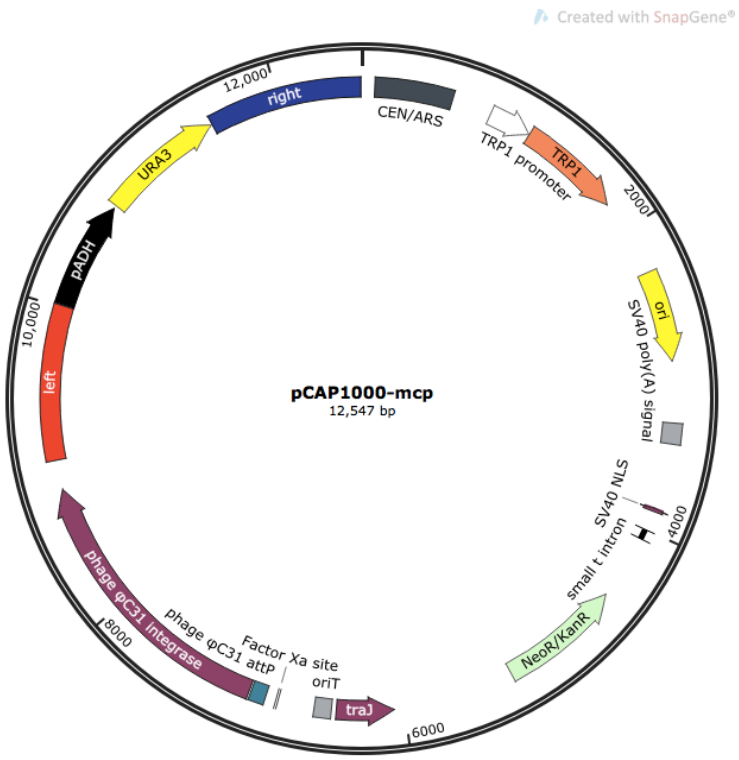


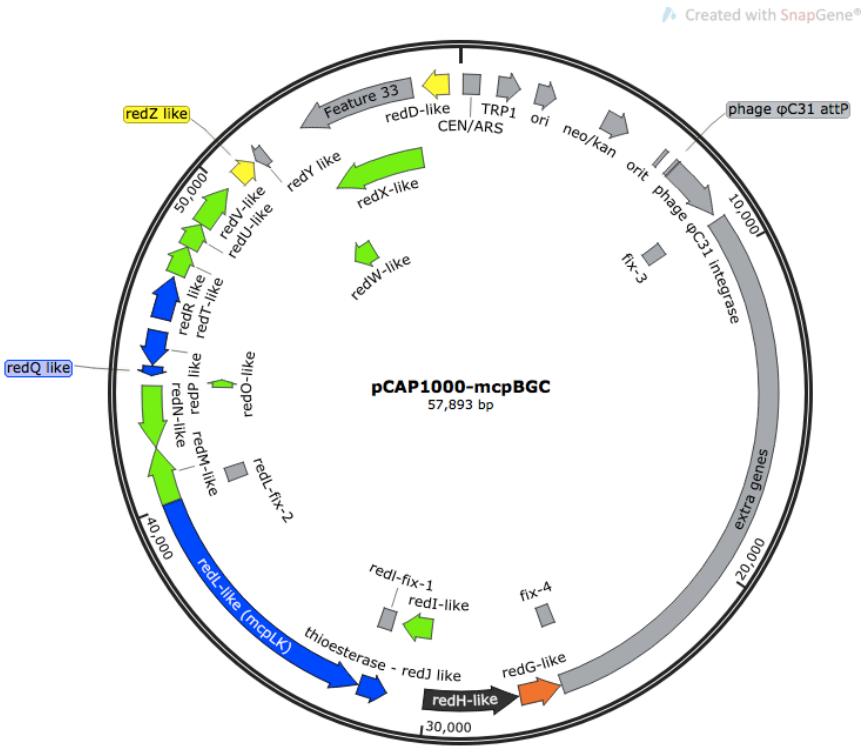

**Appendix 7:** Plasmid maps for pET28a(+)/*redG* (top) and pET28a(+)/*mcpG* (bottom).



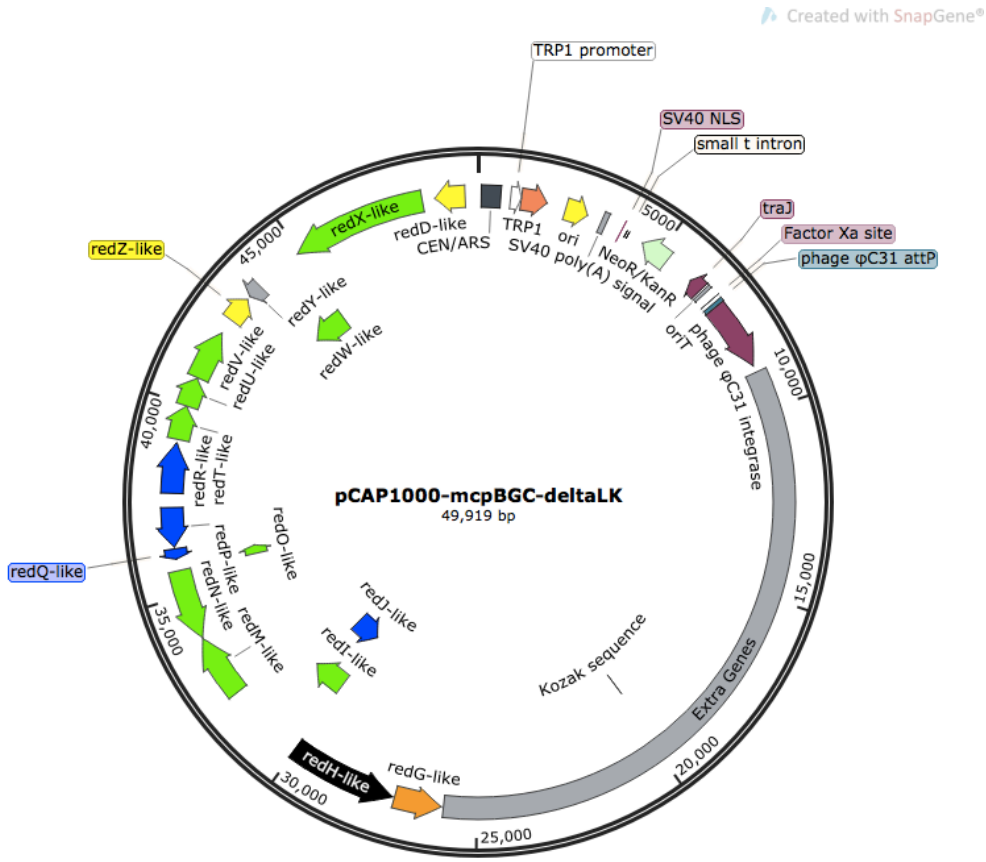

**Appendix 8:** Plasmid map of pOSV556t/*mcpH*




**Appendix 9:** Plasmid map for pCAP1000-mcp capture vector.




**Appendix 10:** Plasmid map for pCAP1000-mcpBGC (Captured cluster).




**Appendix 11:** Plasmid map for pCAP1000-mcpBGC-ΔLK deletion mutant.