

**Investigation on the Biosynthesis and Structural Diversity of  
Biarylittes Uncovers Novel Biosynthetic Paradigm**

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

**Leo Padva**

aus

Kassel

Bonn 2025

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät  
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer: Prof. Dr. Max Crüsemann

Gutachter: Prof. Dr. Martin Baunach

Tag der Promotion: 08.09.2025

Erscheinungsjahr: 2025

# Table of Contents

<b>Abstract .....</b>	<b>I</b>
<b>List of Abbreviations .....</b>	<b>II</b>
<b>List of Figures .....</b>	<b>V</b>
<b>List of Tables .....</b>	<b>VII</b>
<b>Publications included in this thesis .....</b>	<b>VIII</b>
<b>1 Introduction .....</b>	<b>1</b>
1.1 Natural Products.....	1
1.2 Cytochrome P450s: structure, mechanism, and catalytic diversity .....	1
1.2.1 Structural architecture and conservation .....	1
1.2.2 Catalytic mechanism and electron transfer.....	3
1.2.3 Functional diversity and substrate recognition.....	4
1.2.4 Spectroscopic characterisation of P450-substrate interactions.....	5
1.2.5 Roles in peptide natural product biosynthesis .....	5
1.3 Nonribosomal peptides and their biosynthesis.....	6
1.4 Biarylittides: a review of P450-modified RiPP natural products .....	7
1.4.1 Biosynthetic framework and classification of RiPPs .....	7
1.4.2 Discovery and initial characterisation of biarylittides.....	9
1.4.3 The biarylittide biosynthetic gene cluster.....	10
1.4.4 Genome mining of RiPPs and biarylittides .....	12
1.4.5 Targeted mining strategies for biarylittides .....	13
1.4.6 A myxobacterial biarylittide .....	13
1.4.7 Expansion of biarylittide diversity through homology-based genome mining of P450 enzymes involved in RiPP macrocyclisation .....	14
1.4.8 Rule-based genome mining and co-expression of <i>bytAO</i> homologues in <i>E. coli</i> .....	15
1.4.9 P450-associated RiPP genome mining through precursor peptide prediction .....	16
1.4.10 Bioactivity of biarylittides .....	19
1.4.11 Rufomycin .....	19
<b>2 Aims of the study .....</b>	<b>21</b>
2.1 Structural and functional characterisation of a biarylittide crosslinking P450 .....	21

2.2	Neofunctionalisation of biarylittide RiPP machinery in rufomycin NRPS assembly	21
2.3	Machine learning discovery and validation of novel biarylittide pathways .....	22
<b>3</b>	<b>Results and Discussion .....</b>	<b>23</b>
3.1	Structural and Functional Characterisation of P450 <sub>BLT</sub> .....	25
3.1.1	Cytochrome P450 <sub>BLT</sub> Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks.....	25
3.1.2	Reassignment of the Structure of a Tryptophan-Containing Cyclic Tripeptide Produced by the Biarylittide Crosslinking Cytochrome P450 <sub>BLT</sub> .....	28
3.1.3	Structural Insights into a Side Chain Crosslinking Biarylittide P450 from RiPP Biosynthesis.....	30
3.2	Ribosomal Pentapeptide Nitration for Nonribosomal Peptide Antibiotic Precursor Biosynthesis .....	33
3.3	Machine learning-based genome mining of biarylittide BGCs and biochemical validation.....	39
3.3.1	Bioinformatic discovery of a novel subgroup within the biarylittide family .....	39
3.3.2	Machine-learning based genome mining of biarylittides .....	40
3.3.3	Examination of biarylittide chemical diversity .....	44
3.3.4	<i>Saccharothrix variisporea</i> cluster: Discovery and characterisation of Biarylittide YVH and its modified variant .....	46
3.3.5	<i>Streptomyces gougerotii</i> cluster: A P450 with unconventional catalytic residues	52
3.3.6	<i>Streptomonospora alba</i> cluster: Integration of biarylittide and pheganomycin biosynthetic machinery.....	54
3.3.7	<i>In vitro</i> characterisation of BytO homologues .....	58
3.3.8	BytO turnover studies.....	59
3.3.9	Structural elucidation of Nle-KYWH as a widespread biarylittide.....	62
3.3.10	A second P450 in <i>byt</i> gene clusters with MKYWH motif.....	65
<b>4</b>	<b>Summary and Conclusion .....</b>	<b>66</b>
4.1	Structural and functional characterisation of P450 <sub>BLT</sub> .....	66
4.2	RiPP-NRPS integration: a novel paradigm in natural product biosynthesis.....	67
4.3	Expanding biarylittide diversity through machine learning based genome mining and biochemical validation .....	68
<b>5</b>	<b>Materials and Methods .....</b>	<b>70</b>



5.1	Materials.....	70
5.1.1	Chemicals and consumables.....	70
5.1.2	Bacterial strains .....	70
5.1.3	Plasmids.....	72
5.1.4	Media.....	72
5.1.5	Buffers .....	74
5.1.6	Primers.....	75
5.2	Microbiological Methods.....	81
5.2.1	Cultivation of bacterial strains.....	81
5.2.2	Cryopreservation .....	82
5.2.3	Transformation of chemically competent <i>E. coli</i> cells .....	82
5.2.4	Electroporation of electrocompetent <i>E. coli</i> cells.....	82
5.2.5	Intergeneric conjugation between <i>E. coli</i> and <i>Streptomyces</i> .....	83
5.3	Characterisation of biarylptide BGC workflow .....	83
5.3.1	Heterologous production of biarylptides .....	84
5.3.2	Extraction and isolation of biarylptides from bacterial culture .....	84
5.4	Molecular biological methods.....	85
5.4.1	Polymerase chain reaction.....	85
5.4.2	DNA isolation.....	87
5.4.3	Restriction digestion of DNA.....	87
5.4.4	Agarose gel electrophoresis.....	87
5.4.5	Isolation of DNA from agarose gels .....	88
5.4.6	Ligation.....	88
5.4.7	Gibson assembly.....	88
5.4.8	Sanger sequencing .....	89
5.5	Protein Methods .....	89
5.5.1	Protein expression .....	89
5.5.2	Protein purification with Ni-NTA.....	89
5.5.3	Anion exchange and size exclusion chromatography .....	90
5.5.4	SDS-Polyacrylamide gel electrophoresis (SDS-PAGE).....	90
5.5.5	<i>In vitro</i> turnover assay with P450s .....	91
5.5.6	<i>In vitro</i> aspartyl $\beta$ -hydroxylation assay .....	91
5.6	Peptide synthesis .....	92
6	References .....	93

<b>7</b>	<b>Appendix .....</b>	<b>102</b>
7.1	Appendix A .....	102
7.2	Appendix B .....	110
7.3	Appendix C .....	121
7.4	Appendix D .....	137
7.5	Appendix E .....	201
<b>8</b>	<b>Acknowledgements.....</b>	<b>219</b>

## Abstract

Ribosomally synthesised and post-translationally modified peptides (RiPPs) constitute a major group of natural products. Biarylittides, a novel class of RiPPs first isolated from the bacterial genus *Planomonospora*, are biaryl-linked tripeptides synthesised by only two genes: *bytA*, which encodes a minimal five-amino acid precursor with a conserved MxYxH motif, and *bytO*, which encodes a cytochrome P450 monooxygenase (P450) that forms a C-C biaryl crosslink between the tyrosine and histidine aromatic side chains within the peptide. This thesis presents several advances in the understanding of biarylittide biosynthesis and diversity. A detailed characterisation of the BytO homologue P450<sub>Blt</sub> from *Micromonospora* sp. MW-13 unveiled a high degree of substrate flexibility, as well as selective formation of C-N crosslinks, even when tryptophan replaces the canonical histidine in the peptide substrate. Crystal structure analysis of P450<sub>Blt</sub> in complex with its substrate MRYLH identified key residues controlling substrate coordination and reaction specificity. The central discovery of this work revealed that rufomycin biosynthesis contains an unprecedented integration of ribosomal and non-ribosomal peptide pathways, thereby solving the mystery of nitrotyrosine formation. In this pathway, a biarylittide-like precursor peptide MRYLH undergoes selective tyrosine nitration catalysed by the BytO homologue RufO. Following proteolytic cleavage of the nitrated peptide, the released nitrotyrosine is incorporated by a non-ribosomal peptide synthetase to yield the potent antibiotic rufomycins. This neofunctionalised biarylittide machinery established a new paradigm for producing non-proteinogenic amino acids through RiPP pathways using peptides as templates. To explore the broader diversity of biarylittide systems in nature, machine learning-based genome mining was utilised to identify 124 new gene clusters and expanded the known class to 277 members, including two clusters in the human oral microbiome-associated *Rothia* species. Heterologous expression facilitated the characterisation of four novel biarylittides, including a putative hydroxylated YVH tripeptide. A novel BytO homologue subgroup was biochemically validated to catalyse crosslinking of atypical peptide motifs (MRYWY). Furthermore, nuclear magnetic resonance analysis of *in vitro* enzymatic products verified C-N crosslink formation in MKYWH peptides, enabling prediction of the final structure of this widespread bacterial natural product. Collectively, these findings have illuminated the remarkable diversity and evolutionary plasticity of biarylittide biosynthetic systems, creating new opportunities for natural product discovery and engineered peptide modifications.

**List of Abbreviations**

<b>Abbreviation</b>	<b>Definition</b>
AEX	Anion exchange chromatography
Apr	Apramycin
ATP	Adenosine triphosphate
BGC	Biosynthetic gene cluster
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
<i>bytA</i>	Gene encoding biaryllytide precursor peptide
<i>bytH</i>	Gene encoding aspartyl/asparaginylyl $\beta$ -hydroxylase
<i>bytO</i>	Gene encoding biaryllytide P450 monooxygenase
<i>bytT</i>	Gene encoding amidinotransferase
<i>bytZ</i>	Gene encoding SAM-dependent methyltransferase
C	Condensation domain
CaCl <sub>2</sub>	Calcium chloride
Cm	Chloramphenicol
Da	Dalton
DCM	Dichloromethane
DEANO	Diethylamine NONOate
DIC	N,N'-Diisopropylcarbodiimide
DIEA	N,N-Diisopropylethylamine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
Fmoc	Fluorenylmethoxycarbonyl
GNPS	Global Natural Products Social Molecular Networking
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMM	Hidden Markov model
HPLC	High-performance liquid chromatography
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside

---

## List of Abbreviations

---

Kan	Kanamycin
kDa	Kilodalton
LB	Luria-Bertani
LC-MS	Liquid chromatography–mass spectrometry
$m/z$	Mass-to-charge ratio
MDa	Megadalton
MeOH	Methanol
ML	Machine learning
MS	Mass spectrometry or Mannitol soy flour medium
MS/MS	Tandem mass spectrometry
MSSN	Multilayer Sequence Similarity Network
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NO	Nitric oxide
NRPS	Nonribosomal peptide synthetase
OD600	Optical density at 600 nm
ORF	Open reading frame
P450	Cytochrome P450
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PKS	Polyketide synthase
PSI-BLAST	Position-Specific Iterative BLAST
RBS	Ribosome-binding site
RiPP	Ribosomally synthesised and post-translationally modified peptide
rpm	Revolutions per minute
<i>rufB</i>	Gene encoding serine aminopeptidase
<i>rufN</i>	Gene encoding nitric oxide synthase
<i>rufO</i>	Gene encoding nitrating P450
<i>rufT</i>	Gene encoding NRPS
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SOC	Super optimal broth with catabolite repression
SPECO	Small peptide and enzyme co-occurrence analysis

---

## List of Abbreviations

---

SSN	Sequence similarity network
TAE	Tris-acetate-EDTA
TB	Terrific broth
TFA	Trifluoroacetic acid
TIS	Triisopropylsilane
Tris	Tris(hydroxymethyl)aminomethane
TSB	Tryptic soy broth
UV	Ultraviolet
UV-Vis	Ultraviolet-visible spectroscopy
v/v	Volume per volume
w/v	Weight per volume
YEME	Yeast extract-malt extract medium
°C	Degrees Celsius

## List of Figures

<b>Figure 1</b>	Crystal structure of cytochrome P450 OxyBtei (CYP165B7) from the teicoplanin biosynthetic pathway. ....	2
<b>Figure 2</b>	Schematic and simplified representation of a generic cytochrome P450 catalytic cycle with the Fe-coordinating type-B haem depicted at its centre.....	3
<b>Figure 3</b>	A general RiPP biosynthetic gene cluster consists of a precursor peptide, modifying enzymes, and a protease.....	9
<b>Figure 4</b>	Schematic representation of biarylittide BGCs discussed in this chapter, along with the proposed biosynthetic pathway and characterised compounds. ....	11
<b>Figure 5</b>	Phylogenetic tree of detected biarylittide P450s discussed in this review, showing BytA motif diversity, bacterial strain distribution, and publication sources. ....	18
<b>Figure 6</b>	The rufomycin BGC from <i>Streptomyces atratus</i> features a hybrid RiPP-NRPS biosynthetic pathway. ....	35
<b>Figure 7</b>	<i>In vitro</i> exploration of the cytochrome P450 RufO. ....	37
<b>Figure 8</b>	The <i>Stackebrandtia nassauensis</i> <i>byt</i> BGC showing the atypical positioning and sequence of <i>bytA</i> . ....	39
<b>Figure 9</b>	Phylogenetic tree showing the diversity and distribution of putative biarylittide-modifying P450s identified through genome mining. ....	42
<b>Figure 10</b>	Overview of the <i>byt</i> gene clusters examined in this work. ....	44
<b>Figure 11</b>	Schematic representation of the <i>byt</i> BGC in <i>Saccharothrix variisporea</i> (top left); Molecular structure of Biarylittide YVH (top right); LCMS traces of two new signals (bottom left) and MS <sup>2</sup> fragmentation spectra of those compounds (bottom right). ....	47
<b>Figure 12</b>	<sup>1</sup> H-NMR spectrum of biarylittide YVH in deuterated methanol. ....	48
<b>Figure 13</b>	Schematic representation of <i>S. variisporea</i> <i>byt</i> BGC (top left) and cloned constructs to be heterologously expressed to examine enzyme function (bottom left). ....	50
<b>Figure 14</b>	SDS-Page of samples of the protein purification workflow for BytH tested at two different growth conditions, 16 °C (left) and 37 °C (right). ....	51
<b>Figure 15</b>	Schematic representation of the <i>byt</i> BGC found in the genome of <i>Streptomyces gougerotii</i> including the duplicated <i>bytA</i> gene and a section of an alignment of the	

	characterised P450 <sub>Blt</sub> and CYP <sub>S.goug.</sub> showing alterations of the crucial alcohol acid pair (ES→AP).....	53
<b>Figure 16</b>	SDS-Page analysis of BytO from <i>S. gougerotii</i> following two-step purification. ....	54
<b>Figure 17</b>	Schematic overview of pheganomycin BGC in <i>Streptomyces cirratus</i> and <i>byt</i> BGC in <i>Streptomonospora alba</i> . ....	55
<b>Figure 18</b>	Schematic representation of <i>byt</i> BGC in <i>Streptomonospora alba</i> and the genetic constructs generated based upon it. ....	57
<b>Figure 19</b>	SDS pages of six test expressions of BytO homologues. ....	59
<b>Figure 20</b>	<i>In vitro</i> turnover assay incubating Nle-RYWY (blue trace, $m/z = 800.4$ ) and BytO from <i>A. rupis</i> resulting in -2 Da product corresponding to a crosslinked Nle-RYWY mass (red trace, $m/z = 798.4$ ) (top). ....	60
<b>Figure 21</b>	<i>In vitro</i> turnover assay incubating Nle-KYWH (blue trace, $m/z = 744.8$ ) and BytO from <i>Actinomadura hibisca</i> resulting in a -2 Da product corresponding to a crosslinked Nle-KYWH mass (red trace, $m/z = 742.8$ ), and an unexpected second +14 Da product (green trace, $m/z = 758.8$ ). ....	61
<b>Figure 22</b>	NMR analysis of crosslinked Nle-KYWH peptide produced by BytO from <i>Actinomadura hibisca</i> (left) and the derived structure of the natural product biaryllytide YWH (right). ....	62
<b>Figure 23</b>	Structure of 1-phenylimidazole and NMR analysis with coloured arrows corresponding to specific NMR experiments. ....	64



**List of Tables**

<b>Table 1</b>	Summary of strains investigated, including type of analysis performed.....	45
<b>Table 2</b>	NMR assignments related to Figure 14. ....	49
<b>Table 3</b>	Information about BytO homologues investigated for <i>in vitro</i> study. ....	58
<b>Table 4</b>	Complete NMR assignments for in vitro crosslinked Nle-KYWH peptide corresponding to Figure 22.....	63
<b>Table 5</b>	NMR assignments for 1-phenylimidazole, corresponding to Figure 23. ....	64
<b>Table 6</b>	<i>E. coli</i> and <i>S. coelicolor</i> strains used for cloning and expression. ....	70
<b>Table 7</b>	Actinomycete bacterial strains examined for their ability to produce biarylittides and used to isolate genomic DNA. ....	71
<b>Table 8</b>	Plasmids used in this study. ....	72
<b>Table 9</b>	Media used in this study. ....	72
<b>Table 10</b>	Buffers used in this study.....	74
<b>Table 11</b>	Vector primers used in this study.....	75
<b>Table 12</b>	Colony PCR primer used in this study .....	76
<b>Table 13</b>	Strain specific primers used in this study. ....	76
<b>Table 14</b>	Standard PCR reaction setup when using Q5 Polymerase .....	85
<b>Table 15</b>	Thermal Cycling conditions for Q5 PCR amplification.....	86
<b>Table 16</b>	Standard PCR reaction setup using GoTaq DNA polymerase.....	86
<b>Table 17</b>	Thermal cycling conditions for GoTaq PCR amplification.....	86
<b>Table 18</b>	Composition of restriction digestion reactions. ....	87
<b>Table 19</b>	Composition of Gibson assembly master mix.....	89
<b>Table 20</b>	Biarylittide P450s identified through machine learning based genome mining. 215	

## Publications included in this thesis

**Padva, L.**, Zimmer, L., Gullick, J., Zhao, Y., Sasi, V. M., Schittenhelm, R. B., Jackson, C. J., Cryle, M., Crüsemann, M., Ribosomal pentapeptide nitration for non-ribosomal peptide antibiotic precursor biosynthesis, *Chem* 2025, *11*, 102438.

DOI: 10.1016/j.chempr.2025.102438

**Padva, L.**, Gullick, J., Coe, L. J., Hansen, M. H., De Voss, J. J., Crüsemann, M., Cryle, M. J., The biarylittides: understanding the structure and biosynthesis of a fascinating class of Cytochrome P450 modified RiPP natural products, *ChemBioChem* 2024, e202400916.

DOI: 10.1002/cbic.202400916

Coe, L. J., Zhao, Y., **Padva, L.**, Keto, A., Schittenhelm, R., Tailhades, J., Pierens, G., Krenske, E. H., Crüsemann, M., De Voss, J. J., Cryle, M. J., Reassignment of the Structure of a Tryptophan-Containing Cyclic Tripeptide Produced by the Biarylittide Crosslinking Cytochrome P450<sub>blt</sub>, *Chemistry – A European Journal* 2024, *30*, e202400988.

DOI: 10.1002/chem.202400988

Hansen, M. H., Keto, A., Treisman, M., Sasi, V. M., Coe, L., Zhao, Y., **Padva, L.**, Hess, C., Leichthammer, V., Machell, D. L., Schittenhelm, R. B., Jackson, C. J., Tailhades, J., Crüsemann, M., De Voss, J. J., Krenske, E. H., Cryle, M. J., Structural Insights into a Side Chain Crosslinking Biarylittide P450 from RiPP Biosynthesis, *ACS Catalysis* 2024, *14*, 812-826.

DOI: 10.1021/acscatal.3c05417

Zhao, Y., Marschall, E., Treisman, M., McKay, A., **Padva, L.**, Crüsemann, M., Nelson, D. R., Steer, D. L., Schittenhelm, R. B., Tailhades, J., Cryle, M. J., Cytochrome P450<sub>BlT</sub> Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks, *Angewandte Chemie International Edition* 2022, *61*, e202204957.

DOI: 10.1002/anie.202204957

# **1 Introduction**

## **1.1 Natural Products**

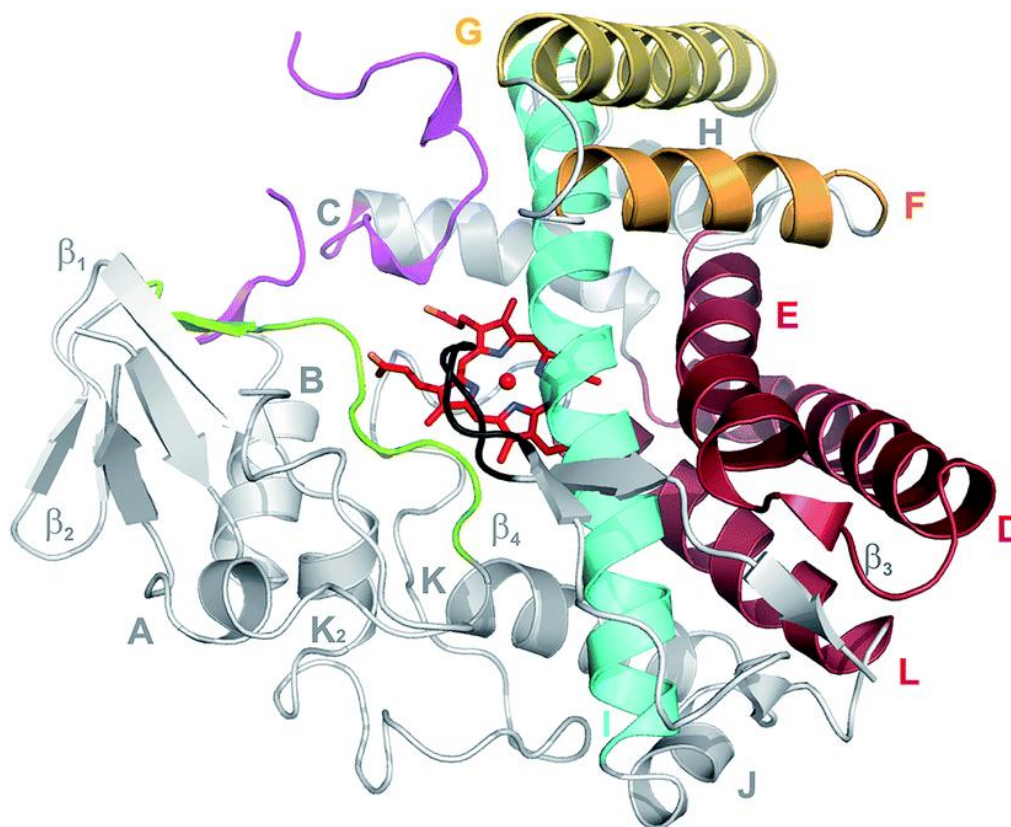
Natural products are biomolecules produced by virtually all living organisms to serve essential ecological functions. Also known as secondary metabolites, they are biosynthesised by distinct enzymatic pathways which produce diverse chemical scaffolds including polyketides, terpenes, and peptide-based natural products. Among these, peptide-based natural products have attracted particular attention due to their exceptional functional diversity, which is a reflection of their stunning structural complexity. The sophisticated architecture of peptide natural products results from an intricate interplay between dedicated biosynthetic enzymes and their substrates, based on two key elements: the assembly of diverse amino acid building blocks into peptide chains, and their subsequent modification by specialised enzymes. These post-translational modifications create rigid three-dimensional conformations that enable precise molecular targeting. Through processes such as cyclisation and functional group addition, linear peptides are transformed into complex scaffolds that have proven invaluable in drug discovery, with applications ranging from antibiotic development to pain management and food preservation.<sup>[1]</sup> Among the extensive enzymatic machinery orchestrating these transformations, cytochrome P450 monooxygenases stand out as particularly versatile catalysts. Their fundamental role in generating complex structural features is especially evident in certain peptide families such as the biarylides. The following chapter explores the structural and mechanistic features that enable these remarkable enzymes to perform such diverse chemical transformations.

## **1.2 Cytochrome P450s: structure, mechanism, and catalytic diversity**

### **1.2.1 Structural architecture and conservation**

Cytochrome P450 monooxygenases (P450s) constitute a remarkable enzyme superfamily that catalyses diverse oxidative transformations across all kingdoms of life. These haem-containing monooxygenases serve essential functions in both primary and secondary metabolism, with their name derived from a characteristic spectroscopic feature: a distinctive absorption maximum at 450 nm when the reduced enzyme is complexed with carbon monoxide. The P450s are typically 40-55 kilodalton (kDa) proteins that share a conserved structural architecture

dominated by  $\alpha$ -helical elements (conventionally labelled A-L) and select  $\beta$ -sheet regions. At their core, these enzymes feature a four-helix bundle (helices D, E, I, and L) that anchors the central haem prosthetic group (Figure 1).<sup>[2]</sup>

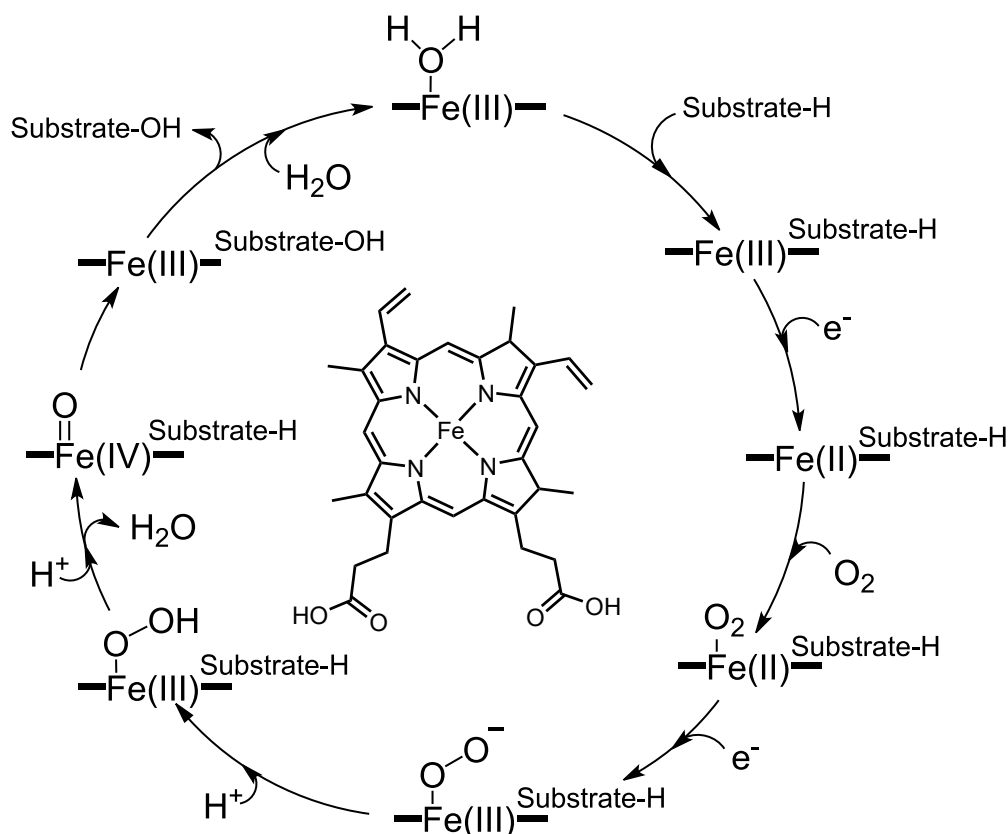


**Figure 1** Crystal structure of cytochrome P450 OxyBtei (CYP165B7) from the teicoplanin biosynthetic pathway. The structure highlights the characteristic P450 fold with its conserved  $\alpha$ -helical architecture (helices A-L) and  $\beta$ -sheet regions ( $\beta$ -1/2). The catalytically essential four-helix bundle (helices D, E, I, and L) is depicted in dark red, with the I-helix shown in turquoise. Key active site regions include the substrate-binding BC loop (pink), substrate recognition F-helix (orange) and G-helix (gold), catalytic I-helix (turquoise), structural post-K-helix  $\beta$ -strand (green), and C-terminal loop (black). The haem prosthetic group is shown in stick representation: Figure adapted from Greule et al.<sup>[2]</sup>

A highly conserved cysteine residue provides the characteristic thiolate coordination to the haem iron, enabling the unique catalytic capabilities of P450s. The enzyme's distinctive architecture, particularly the conserved I-helix, orchestrates critical proton delivery networks that, in concert with this thiolate coordination, facilitate the activation of molecular oxygen to generate reactive iron-oxo species. The remarkable functional diversity of P450s arises from variations in this common structural scaffold. The substrate binding pocket, shaped by the B-C loop, F-G helices, and I-helix, demonstrates considerable plasticity across the enzyme family, accommodating substrates ranging from small molecules to complex peptides.<sup>[3]</sup>

### 1.2.2 Catalytic mechanism and electron transfer

The catalytic mechanism (Figure 2) relies on specialised electron transfer partners - such as ferredoxins and ferredoxin reductases in bacterial systems - to supply the electrons necessary for oxygen activation. Through this sophisticated electron transport chain and precisely arranged proton delivery, P450s generate a highly reactive compound I species capable of performing diverse oxidative chemistry.<sup>[4]</sup>



**Figure 2** Schematic and simplified representation of a generic cytochrome P450 catalytic cycle with the Fe-coordinating type-B haem depicted at its centre. Two electrons and two protons are delivered to the haem-iron, changing its oxidation states.

The P450 catalytic cycle involves a series of well-defined intermediates. The cycle begins with the enzyme in its resting state, where a water molecule serves as the sixth ligand to the ferric ( $\text{Fe}^{\text{III}}$ ) haem iron.<sup>[5]</sup> When a substrate enters the active site, it displaces this water molecule, converting the iron to a pentacoordinate high-spin ferric state. This transition produces a characteristic shift in the UV-visible absorption spectrum, known as a Type I binding spectrum. The spin-state change increases the haem iron's reduction potential, making it more favourable to accept the first electron from redox partner proteins. Upon reduction, the ferrous ( $\text{Fe}^{\text{II}}$ ) haem rapidly binds molecular oxygen, forming an oxy-ferrous complex similar to that found in

oxyhemoglobin. A second electron transfer to this complex generates a ferric-peroxo species that undergoes protonation to form a ferric-hydroperoxo intermediate (Compound 0). A second protonation event triggers heterolytic cleavage of the O-O bond, releasing water and producing the highly reactive iron-oxo species called Compound I.<sup>[6]</sup> Compound I consists of an  $\text{Fe}^{\text{IV}}=\text{O}$  centre with an oxidised porphyrin radical (ferryl-oxo porphyrin  $\pi$ -cation radical) and serves as the primary oxidising species that abstracts hydrogen from the substrate, initiating P450's characteristic oxidation reactions.<sup>[7]</sup>

P450 enzymes require external electrons to complete their catalytic cycle, and these electrons are supplied by diverse redox partner systems that vary across biological domains. In the endoplasmic reticulum of mammals, microsomal P450s receive electrons from NADPH-cytochrome P450 reductase, a membrane-bound enzyme containing both FAD and FMN cofactors that transfer electrons sequentially from NADPH to the P450.<sup>[4]</sup> In contrast, mitochondrial P450s in eukaryotes and most bacterial P450s utilise a three-component electron transport chain consisting of: (1) an FAD-containing NADPH reductase that accepts electrons from NADPH; (2) a soluble iron-sulfur ferredoxin that shuttles electrons from the reductase; and (3) the P450 enzyme itself.<sup>[8]</sup> Some P450 enzymes have evolved as self-sufficient systems where the reductase components are fused to the P450 domain within a single polypeptide chain, as exemplified by P450BM3 from *Bacillus megaterium*.<sup>[9]</sup> This remarkable diversity in electron transfer systems highlights the evolutionary adaptability of P450s across different biological niches and metabolic contexts.

### 1.2.3 Functional diversity and substrate recognition

The catalytic machinery of P450s enables an extensive range of chemical transformations. In bacterial biosynthetic pathways, these enzymes exhibit precise substrate selectivity, whether processing small molecules, carrier protein-bound intermediates, or peptides. The foundational P450-catalysed reaction is selective hydroxylation of aliphatic and aromatic centres, demonstrated by P450<sub>sky</sub> in skyllamycin biosynthesis.<sup>[10]</sup> Beyond this core reaction, P450s orchestrate diverse oxidative processes, including epoxidation - as shown by PimD and EpoK in pimaricin and epothilone biosynthesis respectively.<sup>[11,12]</sup> More sophisticated transformations showcase the enzyme family's catalysed chemical diversity. Oxidative phenolic coupling reactions are catalysed by OxyA and OxyB during glycopeptide antibiotic assembly,<sup>[13]</sup> while CYP158A2 mediates flaviolin dimerisation.<sup>[14]</sup> Some P450s execute complex multistep oxidative cascades, exemplified by TamI in tirandamycin biosynthesis.<sup>[15]</sup> The evolutionary adaptability of these enzymes has produced mechanisms beyond oxygen insertion, including

OleT's oxidative decarboxylation of long-chain fatty acids <sup>[16]</sup> and TxtE's remarkable direct nitration of tryptophan in thaxtomin biosynthesis.<sup>[17]</sup> The positioning of catalytic residues within the active site determines regioselectivity and the nature of the chemical transformation. At the same time, residues in the substrate access channel control substrate specificity and orientation.<sup>[18]</sup> Molecular recognition often involves a delicate balance between specific anchoring interactions and more general hydrophobic interactions that properly position the target site for oxidation.<sup>[19]</sup>

#### **1.2.4 Spectroscopic characterisation of P450-substrate interactions**

Spectroscopic approaches serve as fundamental tools for investigating P450 enzymes and their substrate interactions.<sup>[5]</sup> The haem group within these enzymes functions as an excellent spectroscopic indicator, exhibiting distinctive absorption patterns that transform in predictable ways during substrate binding and catalytic processes.<sup>[5,20]</sup> Two distinct types of binding spectra are typically observed: Type I and Type II. Type I binding spectra result from displacement of the water molecule that serves as the sixth ligand to the haem iron, causing a shift of the Soret band from ~418 nm to ~390 nm. This spectral change is observed with productive substrates and corresponds to a transition from low-spin to high-spin state of the haem iron. In contrast, Type II binding spectra, characterised by a red shift of the Soret band to approximately 430-455 nm, occur when a ligand (often containing a nitrogen atom) directly coordinates to the haem iron. Type II interactions are typically associated with inhibitory compounds rather than productive substrates. Beyond UV-visible spectroscopy, more sophisticated techniques offer deeper insights into P450 catalysis. Resonance Raman spectroscopy provides detailed information about the iron-oxygen species formed during catalysis,<sup>[21]</sup> while electron paramagnetic resonance (EPR) spectroscopy can characterise changes in the haem electronic environment at cryogenic temperatures.<sup>[22]</sup> Stopped-flow kinetic studies enable the capture of short-lived catalytic intermediates, providing temporal resolution of the catalytic cycle under high pressures.<sup>[23]</sup> Recent advances in time-resolved crystallography and cryo-electron microscopy promise to further enhance our understanding of P450 catalytic mechanisms.<sup>[20]</sup>

#### **1.2.5 Roles in peptide natural product biosynthesis**

The catalytic diversity of P450s plays an important role in peptide natural product biosynthesis, which proceeds through two major pathways: nonribosomal peptide assembly by large enzymatic complexes, and extensively modified ribosomal peptide synthesis. In nonribosomal

peptide biosynthesis, P450s often catalyse critical oxidative modifications of amino acid building blocks or final products, introducing structural complexity and bioactivity-determining features.<sup>[24]</sup> While most P450s in natural product biosynthesis perform hydroxylation, such as the well-studied PikC from *Streptomyces*,<sup>[25,26]</sup> there are fascinating deviations from this pattern. One notable example is peptide crosslinking, such as the characteristic biaryl or biaryl ether bridges featured in glycopeptide antibiotics like vancomycin.<sup>[27]</sup> For ribosomal peptides, P450s participate in extensive post-translational modifications that transform simple peptide precursors into complex natural products with diverse bioactivities. Recently described examples include the atropopeptides, where a single P450 enzyme (TrpB) facilitates atropospecific introduction of one carbon-carbon and two carbon-nitrogen bonds,<sup>[28]</sup> and the antibiotic cihunamides characterised by P450-mediated oxidative Trp-Trp cross-linking in their biosynthesis. These are both members of the newly defined class of P450-mediated peptides called cyptides that encompass a total of 22 families RiPPs featuring biaryl macrocyclic linkages with distinct aromatic residue patterns.<sup>[29]</sup>

### 1.3 Nonribosomal peptides and their biosynthesis

Nonribosomal peptides comprise a major class of bacterial secondary metabolites synthesised independently of the ribosome by sophisticated enzymatic complexes known as nonribosomal peptide synthetases (NRPS). These remarkable megaenzymes, which range in size from 100 kDa to 150 megadalton (MDa), orchestrate the production of peptides with extraordinary structural diversity. In contrast to ribosomal peptide synthesis, which is confined to the 20 proteinogenic amino acids, NRPS machinery can incorporate hundreds of distinct building blocks, including non-proteinogenic amino acids, fatty acids, and modified monomers.<sup>[30]</sup> This expanded chemical repertoire enables the generation of compounds with a wide range of biological activities, as exemplified by clinically significant antibiotics such as vancomycin and teicoplanin, powerful immunosuppressants as ciclosporin, and anticancer agents like bleomycin.<sup>[31]</sup>

The biosynthesis of nonribosomal peptides proceeds through an elegant assembly-line mechanism, where individual NRPS modules each govern the incorporation of a specific building block.<sup>[32]</sup> Following the principle of co-linearity, the number and arrangement of modules typically correspond directly to the sequence of amino acids in the final peptide product. Each minimal NRPS module encompasses three essential domains: an adenylation (A) domain that discriminately selects and activates specific building blocks through ATP-



dependent adenylation, a peptidyl carrier protein (PCP, also termed T) domain that serves as a way station for the growing peptide chain via a phosphopantetheine prosthetic group, and a condensation (C) domain that catalyses peptide bond formation. The structural complexity of the final product can be further increased through the action of additional tailoring domains and enzymes, which perform various modifications including epimerisation, methylation, and cyclisation.

## **1.4 Biarylittides: a review of P450-modified RiPP natural products**

While nonribosomal peptide synthesis represents one major strategy for peptide natural product biosynthesis, ribosomally synthesised and post-translationally modified peptides (RiPPs) constitute an equally significant biosynthetic paradigm. This thesis focuses specifically on the latter, with particular emphasis on the emerging class of biarylittides, which exemplify the remarkable catalytic capabilities of cytochrome P450s in peptide modification. The following chapters draw substantially from a comprehensive review article co-authored by the author of this thesis, entitled "The biarylittides: understanding the structure and biosynthesis of a fascinating class of cytochrome P450 modified RiPP natural products".<sup>[33]</sup>

### **1.4.1 Biosynthetic framework and classification of RiPPs**

Ribosomally synthesised and post-translationally modified peptides (RiPPs) represent a diverse and rapidly expanding class of natural products crucial to microbial ecology, human health, and biotechnology.<sup>[34,35]</sup> Their biosynthesis begins with ribosomal synthesis of a precursor peptide, followed by enzymatic modifications, proteolytic cleavage and export.<sup>[36]</sup> The precursor peptide typically consists of an *N*-terminal leader sequence, which directs post-translational modifications, and a core peptide that becomes the final bioactive molecule (Figure 3).<sup>[37]</sup> The wide array of post-translational modifications, including cyclisation, heterocyclisation, and various side-chain alterations gives rise to complex structures with diverse bioactivities, ranging from antibiotic and anticancer agents to signalling molecules (Figure 3).<sup>[38]</sup>

The leader peptide plays a crucial role in RiPP biosynthesis, serving multiple functions that distinguish this pathway from nonribosomal peptide synthesis. It can act as a recognition element for the post-translational modification enzymes, positioning the core peptide appropriately for specific transformations.<sup>[37]</sup> The leader peptide can also prevent premature bioactivity of the developing RiPP, serving as a prodrug-like feature, and may aid in the export of the mature product.<sup>[39,40]</sup> In some RiPP systems, the leader peptide contains specific

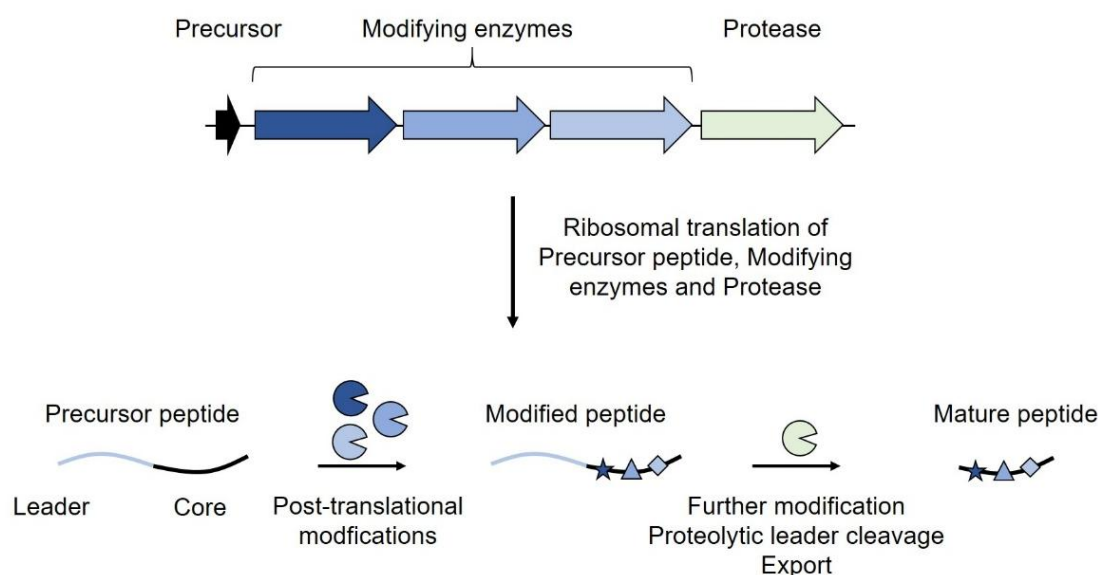
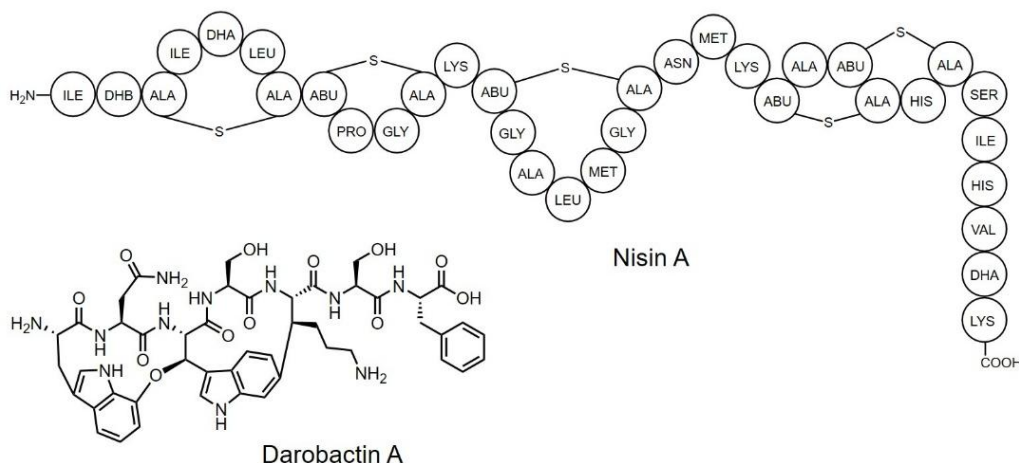
recognition sequences directing proteolytic cleavage, such as the double glycine motif found in lantibiotics.<sup>[41]</sup> The length and sequence of leader peptides vary considerably across different RiPP families, ranging from 20-100 amino acids in most systems.<sup>[35]</sup>

RiPPs are classified into numerous families based on their distinctive structural features and biosynthetic origins. Lantibiotics, exemplified by nisin, contain characteristic lanthionine and methyllanthionine thioether bridges formed between dehydrated serine/threonine residues and cysteine thiol groups.<sup>[42]</sup> Lasso peptides feature a unique threaded structure, where the C-terminal tail is threaded through an *N*-terminal macrolactam ring, creating a mechanically interlocked topology.<sup>[43]</sup> Cyanobactins, such as patellamides, are macrocyclic peptides containing heterocyclic modifications of cysteine, serine, and threonine residues.<sup>[44]</sup> These diverse structural motifs extend to linearazole-containing peptides (LAPs) like microcin B17, which contain multiple thiazole and oxazole heterocycles derived from cysteine, serine, and threonine residues.<sup>[45]</sup> Similarly complex are thiopeptides, which combine heterocycle formation with dehydrations and macrocyclisations to create complex polycyclic structures featuring a central pyridine or dehydropiperidine ring.<sup>[46]</sup> Sactipeptides contain unique thioether crosslinks between cysteine sulfur atoms and the  $\alpha$ -carbon of acceptor amino acids, formed through radical SAM enzyme activity.<sup>[47]</sup> The bottromycins feature a unique macrocyclic amidine and multiple unusual modifications, including unusual  $\beta$ -methylation of amino acid residues.<sup>[48]</sup>

Beyond their structural and functional diversity, what makes RiPPs particularly attractive for natural product engineering is their unique biosynthetic logic. The direct genetic encoding of precursor peptides facilitates easy manipulation through targeted mutagenesis, enabling the creation of diverse structural variants. This feature makes RiPPs valuable targets for drug discovery and biotechnological applications, as it allows for the rapid generation of potentially bioactive molecules.<sup>[49]</sup>

**A**

General RiPP biosynthetic gene cluster

**B**

**Figure 3** (A) A general RiPP biosynthetic gene cluster consists of a precursor peptide, modifying enzymes, and a protease. Following ribosomal translation, the core of the precursor peptide is modified by specific enzymes before the leader sequence is cleaved off, resulting in the release of the mature, bioactive peptide. (B) The chemical structures of RiPPs nisin,<sup>[50]</sup> a lantibiotic, and darobactin,<sup>[51]</sup> a recently discovered antibacterial RiPP targeting Gram-negative bacteria. ABU=aminobutyric acid; DHA=dehydroalanine; DHB=dehydrobutyryne; ALA-S-ALA=lanthionine; ABU-S-ALA=β-methyllanthionine. Figure taken from <sup>[33]</sup>.

### 1.4.2 Discovery and initial characterisation of biarylptides

In 2021, Zdouc et al. reported the discovery and characterisation of a novel *N*-acylated tripeptide from a bacterial species of the genus *Planomonospora* using a combination of mass spectrometry-based metabolomics and molecular network analysis.<sup>[52–54]</sup> This tripeptide with the sequence Tyr-Tyr-His (YYH) features a unique carbon-carbon (C-C) biaryl linkage between

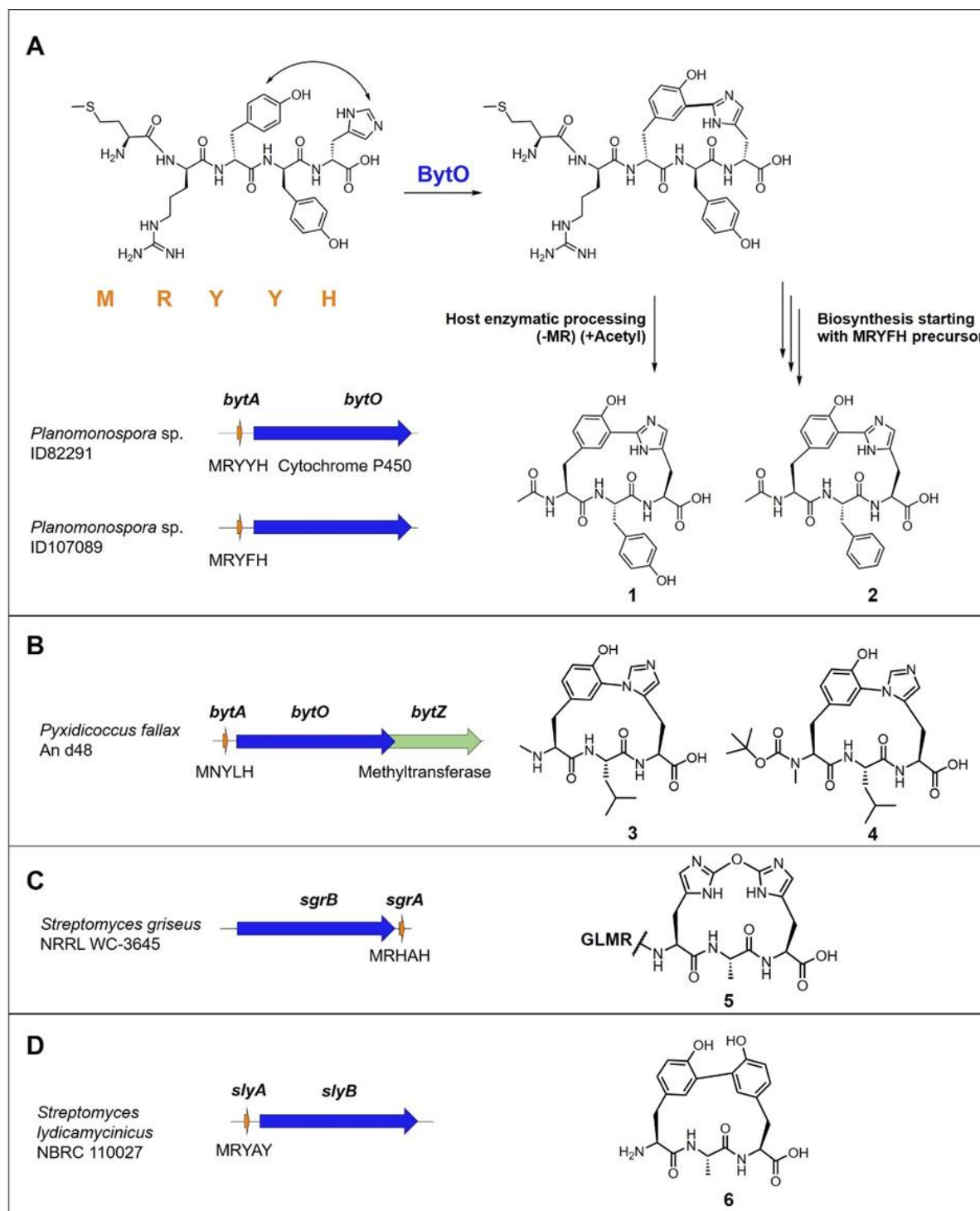
the aromatic residues of Tyr and His. Discovery, bioinformatic analysis and heterologous expression of its biosynthetic gene cluster (BGC) defined it as the first representative of a widespread subgroup of RiPPs, termed the biarylittides. Biarylittide YYH (**1**) was isolated from *Planomonospora* sp. ID82991 and its structure elucidated via nuclear magnetic resonance (NMR) spectroscopy. During this study, a closely related analogue, biarylittide YFH (**2**), was discovered through MS/MS molecular network analysis, where it was grouped with biarylittide YYH (Figure 4). **2** was shown to be produced by three additional *Planomonospora* strains and was subsequently isolated from *Planomonospora* sp. ID107089. Both isolated compounds showed characteristic UV absorption maxima at 270 and 315 nm, the latter indicating an unusual chromophore, specifically a biaryl moiety.

### 1.4.3 The biarylittide biosynthetic gene cluster

To obtain insights into the biosynthesis of the biarylittides, the 7.58-Mbp genome of *Planomonospora* sp. ID82991 was sequenced and analysed using the bioinformatics tool antiSMASH, with a focus on NRPS or RiPP BGCs. Given the known sequence of biarylittide YYH, a search was conducted for trimodular NRPS systems with specificity for aromatic building blocks, but no candidate NRPS BGC was identified. Manual analysis of all potential open reading frames (ORFs) in the genome that yield short proteins after six-frame translations revealed a very short ORF (18 base pairs) encoding the pentapeptide MRYYH, with an upstream Shine-Dalgarno sequence functioning as a ribosomal binding site. The gene immediately downstream of this ORF, PLM4\_2056, encodes a cytochrome P450 monooxygenase.

To confirm the role of the *bytAO* gene cluster in biarylittide biosynthesis, a 1303 bp sequence containing both genes was amplified, cloned into the expression plasmid pSET152 and transferred into *Streptomyces coelicolor* M1152. Cultivation of this strain, followed by high-performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS) analysis, confirmed the production of biarylittide YYH and verified the functional expression of *bytAO*. Notably, *bytA* is the smallest coding gene characterised to date, surpassing the previously smallest known gene *mccA* (21 bp), which is involved in the biosynthesis of the RiPP microcin.<sup>[55]</sup> The size of the *bytA* gene – and hence the BytA peptide substrate – is a major differentiating feature when compared to other RiPP pathways involving crosslinking by cytochrome P450s (such as cittilin and tryptorubin, *vide infra*), with the biarylittides only requiring a 2 amino acid leader sequence in comparison to the 20+ residues required in other

pathways. The size of the *bytA* gene also evokes challenges when attempting to identify biaryllytite gene clusters (*vide infra*).



**Figure 4** Schematic representation of biaryllytite BGCs discussed in this chapter, along with the proposed biosynthetic pathway and characterised compounds. (A) The first described biaryllytite BGC from *Planomonospora* sp. ID82291 contains the 18 bp *bytA* (MRYYYH) and *bytO* (P450). After ribosomal translation, the precursor peptide undergoes cyclisation, catalysed by BytO, to form a C-C biaryl crosslink between Tyr3 and His5. Host enzymes then process the peptide by cleaving the *N*-terminal leader sequence (-MR) and acetylating the *N*-terminus (+acetyl), producing the mature

biarylittide. (B) In *Pyxidicoccus fallax* An d48, the biarylittide BGC includes *bytA* (MNYLH), *bytO* (P450), and *bytZ* (SAM-dependent methyltransferase); shown with structures of the natural myxobacterial biarylittide, myxarylin (**3**), and its semi-synthetic derivative, myxarylin-Boc (**4**). Myxarylin has an *N*-terminal methylation, while myxarylin-Boc contains a Boc-protecting group at the *N*-terminus. (C) The *Streptomyces griseus* NRRL WC-3645 BGC includes *sgrA* (precursor peptide MRHAH) and *sgrB* (P450). This pathway produces Gristide 834 (**5**), which contains a cyclised HAH core with a reported ether crosslink between two histidine residues. The full peptide sequence is GLMRHAH, where GL is an artifact from cloning, MR is the leader sequence, and HAH is the core peptide. (D) The *Streptomyces lydicamycinicus* NBRC 110027 BGC contains *slyA* (MRYAY) and *slyB* (P450), leading to the formation of **6**, the first Tyr-Tyr crosslinked biarylittide. Figure taken from <sup>[33]</sup>.

#### 1.4.4 Genome mining of RiPPs and biarylittides

As our understanding of RiPP biosynthesis has grown, so too has our ability to identify and characterise these compounds encoded in nature. Genome mining, a bioinformatic approach that analyses genomic data to predict the presence of biosynthetic gene clusters (BGCs), has become an invaluable tool in this endeavour.<sup>[56]</sup> This computational approach has rapidly evolved in recent years, integrating various algorithms to detect, annotate, and classify BGCs across diverse organisms. User-friendly tools like antiSMASH and PRISM <sup>[57]</sup> have simplified BGC identification, making it accessible to non-specialists. The accuracy of predictions has improved through the integration of machine learning algorithms, as exemplified by DeepBGC,<sup>[58]</sup> particularly for novel or unusual cluster types. The development of tools like BiG-SCAPE <sup>[59]</sup> has facilitated the integration of genomic, metabolomic, and phylogenetic data, providing a more comprehensive view of BGCs and their products. Expanded databases such as MIBiG <sup>[60]</sup> have catalogued thousands of experimentally characterised BGCs, enhancing reference-based predictions. Additionally, specialised tools like RiPPMiner <sup>[61]</sup> have been developed for particular classes of natural products, improving the detection of specific BGC types. The development of sophisticated tools such as antiSMASH for secondary metabolite gene cluster prediction and RODEO for RiPP precursor peptide identification has significantly enhanced our ability to identify novel biosynthetic pathways and their products.<sup>[62,63]</sup>

However, genome mining for RiPP pathways presents more challenges compared to other natural product classes, such as polyketides and nonribosomal peptides. These are typically synthesised by large multi-modular enzyme complexes and have conserved sequence features across their families, facilitating their identification and analysis. In contrast, RiPPs are defined by family-specific modifying enzymes, which vary significantly across RiPP families.<sup>[64]</sup> This makes it difficult to identify a universally defining characteristic for the class. Additionally, RiPP precursor peptides are notably short – usually between 20 and 100 amino acids – and are often not annotated in genomic sequence data.

### 1.4.5 Targeted mining strategies for biarylittides

These challenges are further amplified in the case of biarylittides, where the precursor peptide BytA comprises only five amino acids and the only constantly associated modifying enzyme is the P450 BytO. This minimal composition, particularly the extremely short precursor peptide, explains why bioinformatic genome mining tools like antiSMASH previously failed to detect them, as the precursor was too small for automatic annotation. With the structure and biosynthetic pathway of biarylittide YYH now elucidated, a targeted genome mining strategy for biarylittides became viable. Zdouc et al. utilised a detailed bioinformatic approach, starting with a phylogenomic analysis of BytO-related P450 enzymes. These enzymes were hypothesised to catalyse biaryl crosslinks on short peptides, with the genes encoding the P450s likely colocalised with those encoding the cognate pentapeptides. Using complete genomes from the antiSMASH database and top BLAST<sup>[65]</sup> hits for BytO from *Planomonospora* sp. ID82291, they compiled an initial pool of over 3,300 sequences.

A hidden Markov model (HMM) search, using the Pfam P450 model (PF00067), was applied with strict cutoffs to ensure high-confidence hits with model and query alignment coverage over 70%. Sequences sharing more than 95% amino acid identity were clustered using an all-vs-all BLAST, and a representative sequence was selected from each cluster for further analysis. Next, the authors performed a peptide motif search in the genomic regions flanking the identified P450 sequences (+/- 500 bp). They utilised a custom Python script to scan for specific pentapeptide motifs (MxYx[Y/H]\*) and screened for the Shine-Dalgarno sequence (AGGAGG) upstream of these motifs to indicate functional peptides. This process identified homologous *bytAO* clusters in over 200 bacterial genomes, primarily within Actinobacteria, with a single cluster in Myxobacteria, confirming the widespread presence of biarylittide BGCs across bacterial lineages. The analysis revealed conserved sequence features, along with some variability, such as lysine replacing arginine at position 2 or leucine and tryptophan substituting tyrosine at position 4 of the peptide. Several homologous BGCs also included genes encoding putative tailoring enzymes, such as methyltransferases,  $\beta$ -hydroxylases, and sulfotransferases, suggesting a broad range of potential biosynthetic modifications within the biarylittide molecular family.<sup>[52]</sup>

### 1.4.6 A myxobacterial biarylittide

Building on this knowledge, in 2021 Hug et al. reported the first myxobacterial biarylittide, named myxarylin (**3**).<sup>[66]</sup> Using a genome-guided approach, they identified a unique *N*-

methyated tripeptide with a C-N biaryl crosslink in the myxobacterial strain *Pyxidicoccus fallax* An d48 (Figure 4). The producing strain contains a *bytA* homologue encoding the sequence MNYLH and a BytO homologue sharing 40.7% identity with BytO from *Planomonospora* sp. ID82291. The gene cluster in *P. fallax* An d48 includes an S-adenosyl-methionine (SAM)-dependent methyltransferase gene, designated *bytZ*. (Figure 4). This additional genetic component distinguishes the myxobacterial biarylite biosynthetic pathway from its previously characterised actinobacterial counterparts. The production of myxarylin by *P. fallax* An d48 showed dependence on specific cultivation conditions, requiring a medium containing autoclaved, inactivated microorganisms. To address the poor solubility of myxarylin, the researchers used a semi-synthetic approach by introducing a Boc protecting group at the *N*-terminal secondary amine, creating myxarylin-Boc (**4**). The compact genetic operon (*bytA*, *bytO*, and *bytZ*) was cloned and expressed in the heterologous host *Myxococcus xanthus* DK1622, resulting in low-level myxarylin production. This was attributed to potentially inefficient cleavage of the *N*-terminal leader peptide by proteases in *M. xanthus*. BytZ is postulated to catalyse *N*-terminal methylation after leader peptide cleavage, a post-translational modification that differs from the *N*-terminal acetylation observed in actinobacterial biarylites. Notably, when *bytAO* from *Planomonospora* was heterologously expressed in *Streptomyces coelicolor*, *N*-acetylation was still observed, suggesting an actinomyces-specific acetylation mechanism. *N*-Acetylation is common in various actinobacterial secondary metabolites, including K-13<sup>[67]</sup> and pseudosporamide,<sup>[68]</sup> indicating a propensity for this modification in this phylum. The observed variations in *N*-terminal modification – acetylation in actinobacterial metabolites versus methylation or unmodified *N*-termini in myxobacterial and fungal metabolites<sup>[69]</sup> – highlight the influence of phylogenetic origin on these biosynthetic processes.

#### **1.4.7 Expansion of biarylite diversity through homology-based genome mining of P450 enzymes involved in RiPP macrocyclisation**

In late 2023, three studies significantly expanded our understanding of RiPPs modified by P450 enzymes, including biarylites.<sup>[29,70,71]</sup> These studies employed diverse bioinformatic strategies to identify novel subgroups of biarylites, with two of these studies producing novel biarylites through *in vivo* expression. Nam et al. utilised a homology-based genome mining approach to identify P450 enzymes associated with peptide macrocycle biosynthesis.<sup>[29]</sup> Their methodology involved the Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST)<sup>[72]</sup> to detect homologous P450 sequences, the Rapid ORF Description and Evaluation



Online (RODEO) <sup>[63]</sup> tool to locate putative precursor peptides near these P450s, and the EFI-EST tool <sup>[73]</sup> to identify conserved aromatic sequence motifs among the identified precursors. Focusing on the P450 TrpB, known for catalysing three crosslinking reactions in tryptorubin A biosynthesis, <sup>[28]</sup> Nam et al. identified 19 novel groups of homologous BGCs. Each group was characterised by distinct aromatic residue patterns in their precursor peptides. When targeting the biarylittide P450 BytO, they uncovered a new subgroup of precursor peptides with members characterised by an MxHxH motif in *bytA*-related sequences, differing from the previously known MxYxH motifs (Figure 4).

Based on these findings, Nam et al. proposed the term ‘cyptides’ as a new name for the RiPP superfamily featuring diverse biaryl crosslinks installed by P450s. They classified cyptides into two types: type 1 for the 19 newly identified groups, and type 2 for the two groups of biarylittides. Interestingly, their analysis did not reveal direct evolutionary relationships between the P450 enzymes of groups 1–20 (type 1 cyptides) and those associated with biarylittides (type 2 cyptides). The researchers proposed that these enzymes may share a distant common P450 ancestor or may have evolved into peptide macrocyclases through two distinct evolutionary pathways. This hypothesis is supported by differences in the average genomic distances between genes encoding precursor peptides and P450s – 104 bp for type 1 cyptides, compared to 174 bp for type 2 cyptides – suggesting distinct evolutionary paths and functional adaptations. This classification and evolutionary analysis further underline the unique nature of biarylittides within the broader context of P450-mediated RiPP biosynthesis.

#### **1.4.8 Rule-based genome mining and co-expression of *bytAO* homologues in *E. coli***

He et al. developed an innovative rule-based genome mining strategy to identify novel P450 enzymes involved in RiPP macrocyclisation, addressing the limitations of traditional homology-based methods. <sup>[70]</sup> Their approach incorporated a Multilayer Sequence Similarity Network (MSSN) analysis, systematically examining small peptide-P450 pairs across 20,399 actinobacterial genomes. To enhance specificity and minimise unrelated noise, the researchers further refined their search using the Small Peptide and Enzyme Co-Occurrence Analysis (SPECOC) workflow, <sup>[74]</sup> focusing on peptides with aromatic residues. Additionally, they employed structural predictions using AlphaFold Multimer <sup>[75,76]</sup> to identify potential interactions between small peptides and P450s, distinguishing post-translational modification (PTM) enzymes from non-PTM enzymes.

From the identified enzyme families, four (KstB, ScnB, MciB, and SgrB) were prioritised for functional characterisation. Notably, SgrB encoded by *Streptomyces griseus* NRRL WC-3645

was selected due to its association with a precursor peptide (SgrA) containing a conserved HxH motif, suggesting potential histidine-histidine cross-linking. However, despite the close relationship to the biarylites, the authors did not classify SgrB as a biarylite P450 or its characteristically small precursor peptide, MRHAH, as belonging to a novel subgroup within that family (Figure 4). To elucidate the function of SgrB, the team expressed SgrA and SgrB in *E. coli*. Mass spectrometric analysis revealed a new signal with a mass increase of 14 Da compared to the unmodified precursor SgrA, suggesting the formation of a new bond and incorporation of an oxygen. Further tandem mass spectrometry localised this modification to the C-terminal HAH region, supporting histidine-histidine cross-linking. NMR spectroscopic analysis was used to suggest the presence of an unexpected ether bond between the histidine residues, representing a very unusual type of cross-linking in bacterial RiPPs. This finding would suggest an unprecedented new catalytic capability within biarylite biosynthesis, although in this publication, the peptide was termed gristide 834 (**5**) and was not attributed to the biarylite family.<sup>[71]</sup> Attempts to produce this compound through heterologous expression in *Streptomyces coelicolor* and *Streptomyces albus* were unsuccessful. This suggests that additional non-clustered genes or specific culture or expression conditions may be required for successful (heterologous) biosynthesis, highlighting the need for further research to elucidate the factors influencing the production of biarylites.

#### **1.4.9 P450-associated RiPP genome mining through precursor peptide prediction**

Hu et al. developed an innovative strategy to identify P450-modified RiPPs in genomes, addressing the challenges posed by the simplicity of these gene clusters.<sup>[70]</sup> Their approach utilised conserved RBSs upstream of precursor peptide genes to predict short peptides of 5 to 60 amino acids. This method, validated against known RiPP BGCs, proved both accurate and reliable. To manage computational efficiency, the researchers employed the UniRef90 database, focusing on genomic regions flanking P450 genes to identify potential precursor peptides.

The initial analysis yielded 12,847 peptide sequences, which was refined to 2,821 sequences linked to 2,648 P450 enzymes by focusing on peptides containing aromatic residues often involved in cross-linking reactions. Subsequent Sequence Similarity Network (SSN) analysis clustered these peptides and P450 enzymes, revealing 16 different classes of P450-modified RiPPs, including three known (cittilin, tryptorubin, and biarylite) and 13 novel compound classes. Biarylites were classified under Class C, with most precursor peptides displaying the conserved MxYxH motif. Interestingly, some exceptions included peptides with 7 or 8 residues,

typically due to additional nonaromatic amino acids at the C-terminus. A subset featuring the MxYxY motif was also identified, suggesting the potential formation of a cyclic tripeptide through a cross-link between two tyrosine residues. The study examined various gene clusters, including the *sly* BGC from *Streptomyces lydicamycinicus* NBRC 110027, encoding the precursor peptide SlyA with the sequence MRYAY (Figure 4). Initial attempts to express this cluster in *Streptomyces albus* J1074 were unsuccessful. However, expression in *E. coli* ArcticExpress (DE3) cells, supplemented with cyanobacterial ferredoxin and ferredoxin reductase, facilitated P450 functionality. Following expression and purification, LC-MS analysis of the cleaved SlyA peptide revealed a new peak with a 2 Da decrease in molecular weight, suggesting a cross-linking reaction. MS/MS analysis confirmed cross-linking between the tyrosine residues in the YAY region. NMR spectroscopy further elucidated the structure, revealing a C-C bond linkage between the tyrosine residues, resulting in a cyclic tripeptide (**6**). Notably, this compound showed significant structural similarities to the core structure of arylomycin A2, an antimicrobial lipopeptide produced via a nonribosomal peptide synthetase (NRPS) pathway.<sup>[77]</sup> A phylogenetic tree summarising the known variety of biarylittide precursor sequences from the discussed studies is shown in Figure 5.



#### 1.4.10 Bioactivity of biarylittides

The discovery of biarylittides has also led to attempts to characterise their potential bioactivity. So far, however, no biological activities have been reported for any of the discovered biarylittides. **1** was shown to be inactive against different proteases,<sup>[52]</sup> and the myxarylins **5** and **6** were inactive against all tested bacterial strains in antibiotic assays.<sup>[66]</sup> However, the vast distribution of biarylittide pathways suggests an important function of these natural products for the producing organisms, which may be related to the soil environment in which the producer strains are largely found. Recently, the structurally related YYW cyclic peptide neopetromin has been reported to possess vacuole fragmentation activity,<sup>[78]</sup> suggesting possible novel functions for such cyclic tripeptides. Future studies need to focus therefore on diverse bioactivity assays as well as investigating the consequences of biarylittide BGC deletions for the wildtype producer strains.

#### 1.4.11 Rufomycin

The rufomycins, also known as the ilamycins, are a family of cyclic heptapeptides that have garnered significant attention due to their potent and selective activity against Mycobacteria, including *Mycobacterium tuberculosis*, the causative agent of tuberculosis.<sup>[79]</sup> With over 50 characterised congeners, rufomycins exhibit exceptional bioactivity compared to existing anti-tuberculosis agents, demonstrating improved efficacy against drug-resistant strains. This positions them as compelling candidates for drug development efforts, especially considering the growing global challenge of multidrug-resistant tuberculosis. These compounds exert their antimicrobial effect by targeting cellular proteostasis via the caseinolytic protein ClpC1 in *M. tuberculosis*.<sup>[80]</sup> By inhibiting ClpC1, rufomycins disrupt protein homeostasis and stress responses in the bacterium, leading to its death and highlighting a novel mechanism of action compared to conventional antibiotics.

A key structural feature of rufomycins is the presence of three non-proteinogenic amino acids, including a polyketide synthase (PKS)-derived L-2-amino-4-hexenoic acid building block, an N-dimethylallyltryptophan, and an enigmatic 3-nitrotyrosine (3-NO<sub>2</sub>-Tyr) residue. Among these, the 3-NO<sub>2</sub>-Tyr residue has been identified as crucial for the potent antibacterial activity of rufomycins.<sup>[81]</sup> The unusual nature of these building blocks, particularly 3-NO<sub>2</sub>-Tyr, has made the biosynthesis of rufomycins a subject of intense investigation.<sup>[82]</sup>

The rufomycin BGC is a complex system that integrates various enzymatic pathways.<sup>[82]</sup> At its core is a heptamodular NRPS RufT, responsible for assembling the peptide backbone. This NRPS works in concert with a PKS system and additional tailoring enzymes to produce the final rufomycin structure.<sup>[83]</sup> Extensive research has been conducted on rufomycin biosynthesis, including the characterisation of unusual building blocks and efforts to engineer enhanced anti-tuberculosis agents.<sup>[84,85]</sup> A significant focus of these studies has been the P450 RufO, initially hypothesised to catalyse the nitration of tyrosine at C3.<sup>[82,84]</sup> However, despite its putative role, the exact substrate and reaction catalysed by RufO remained elusive. Recent structural studies by two independent groups have provided valuable insights into the potential biological role and substrate recognition of RufO,<sup>[86,87]</sup> yet the precise mechanism of 3-NO<sub>2</sub>-Tyr incorporation continued to be a subject of debate until very recent investigations by our groups.<sup>[88]</sup>

## 2 Aims of the study

Biarylittes represent a fascinating class of natural products at the intersection of ribosomal peptide biosynthesis and P450-mediated oxidative enzymatic transformations. The discovery of these minimal biosynthetic systems has raised fundamental questions about peptide modification, enzyme-substrate recognition, and natural product pathway evolution. This thesis aims to address three interconnected research questions through a combination of biochemical, structural, and computational approaches.

### 2.1 Structural and functional characterisation of a biarylittide crosslinking P450

The BytO homologue P450<sub>Bit</sub> from *Micromonospora* sp. MW-13 is planned to serve as a model system for understanding peptide modification by P450 enzymes *in vitro*. This project will investigate its mechanistic diversity through comprehensive biochemical characterisation of substrate specificity and crosslinking activity. Crystallographic analysis shall provide structural insights into substrate binding, while systematic mutagenesis of conserved amino acids in the related *bytAO* system from *Planomonospora*, expressed heterologously, shall validate these findings *in vivo*.

### 2.2 Neofunctionalisation of biarylittide RiPP machinery in rufomycin NRPS assembly

The rufomycin BGC from *Streptomyces atratus* contains machinery typically associated with biarylittide RiPP biosynthesis, including a P450 enzyme (RufO) and a short precursor peptide, suggesting an evolutionary repurposing of RiPP pathways within nonribosomal peptide synthesis. This project will investigate this unique pathway integration using genetic and biochemical approaches to understand how the RiPP biosynthetic machinery has been neofunctionalised for NRPS assembly. Producer strain manipulation, *in vitro* biochemistry, and comparative genomics will be used to elucidate the molecular basis of this pathway adaptation and to explore whether similar evolutionary innovations exist throughout bacterial natural product biosynthesis.

### **2.3 Machine learning discovery and validation of novel biarylittide pathways**

This project will adapt AtropoFinder<sup>[89]</sup> to identify novel BytO homologues and CoreFinder's<sup>[89]</sup> flexible peptide sequence recognition parameters based on recent insights into biarylittides, enabling the discovery of clusters that elude traditional searches. Selected biarylittide BGCs will be validated by heterologous expression and product characterisation, with *in vitro* biochemical analysis of BytO homologues revealing the functional diversity within this enzyme family.



### 3 Results and Discussion

This study presents three interconnected research directions that collectively advance our understanding of biarylptide biosynthesis and its broader implications for natural product chemistry. The results are organised into three main chapters, which are briefly summarised below.

First, a comprehensive mechanistic and structural investigation of the biarylptide-modifying cytochrome P450<sub>BLT</sub> from *Micromonospora* sp. MW-13 was conducted. Initial biochemical characterisation revealed the enzyme's high substrate affinity and catalytic efficiency, along with remarkable flexibility in accepting modifications at specific positions (Chapter 3.1.1). Subsequent structural analysis led to an important revision of P450<sub>BLT</sub>'s crosslinking chemistry, demonstrating the formation of C-N bonds between tryptophan and tyrosine residues rather than the previously reported ether links (Chapter 3.1.2.). This discovery was complemented by solving the first crystal structure of a peptide-modifying P450 in complex with its substrate (P450<sub>BLT</sub> with MRYLH), revealing crucial determinants of reaction specificity through key active site residues His-234 and Ser-239 (Chapter 3.1.3).

Second, an unexpected case of pathway neofunctionalisation was elucidated where biarylptide biosynthetic machinery has been repurposed within the rufomycin BGC found in *Streptomyces atratus* (Chapter 3.2). This investigation revealed how RufO, a P450 enzyme sharing high sequence identity with biarylptide P450s, evolved to perform selective tyrosine nitration instead of the typical crosslinking reaction. Through detailed characterisation of the complete pathway, it was demonstrated how the resulting nitrated pentapeptide is proteolytically processed to release 3-nitrotyrosine – an essential building block for rufomycin assembly. This discovery establishes a novel paradigm where RiPP biosynthetic machinery generates nonproteinogenic amino acids for nonribosomal peptide assembly lines, showcasing nature's ability to integrate fundamentally different peptide biosynthesis logics.

Third, the known diversity of biarylptide biosynthetic systems was substantially expanded through the development and application of a machine learning-based genome mining approach (Chapter 3.3). This comprehensive analysis led to the identification of 277 putative members, including 124 novel variants, revealing both unprecedented precursor sequence patterns and new co-occurrence patterns with putative modifying enzymes. Notably, the algorithm detected biarylptide BGCs in *Rothia* species from the human oral microbiome, significantly expanding the previously known environmental distribution of biarylptides. The bioinformatic predictions were experimentally validated through the characterisation of four new compounds (YVH,

hydroxylated YVH, YLH and YDH) from diverse bacterial sources, including *Saccharothrix variisporea*, *Streptomyces gougerotii*, and *Streptomonospora alba*. Biochemical analyses revealed that a newly discovered subgroup of BytO homologues forms crosslinks in MRYWY peptides, representing a rare sequence variant in biarylittides. Moreover, extensive structural and functional characterisation of the widely distributed MKYWH motif was achieved through NMR analysis, following its isolation from large-scale *in vitro* assays.

### 3.1 Structural and Functional Characterisation of P450<sub>Blt</sub>

#### 3.1.1 Cytochrome P450<sub>Blt</sub> Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks

Yongwei Zhao, Edward Marschall, Maxine Treisman, **Leo Padva**, Max Crüsemann, David R Nelson, David L Steer, Ralf B Schittenhelm, Julien Tailhades, Max J Cryle

*Angewandte Chemie International Edition* 2022, 61, e202204957

Please refer to Appendix A to view the full text of this publication.

#### PUBLICATION SUMMARY

The crosslinking of peptides represents a crucial modification in many important natural products, including streptide, glycopeptide antibiotics, and arylomycins, where it serves to rigidify and stabilise peptide structures while potentially introducing restricted rotation and planar chirality.<sup>[13,77,90,91]</sup> Among the enzymes responsible for installing these side chain crosslinks, cytochrome P450 monooxygenases (P450s) have emerged as particularly versatile catalysts, capable of performing diverse oxidative transformations through their ability to activate molecular oxygen into highly reactive intermediates.<sup>[2]</sup> While many P450-mediated peptide modifications occur in non-ribosomal peptide synthetase (NRPS) pathways, which often involve challenging enzyme-bound substrates, the biarylptides represent an attractive alternative as their biosynthesis involves remarkably simple ribosomally synthesised peptide precursors.<sup>[52]</sup> These precursor peptides are particularly notable for their minimal five-amino acid sequence, including a leader sequence of just two residues, making them promising candidates for biocatalysis. This research investigated a P450 from biarylptide biosynthesis as a potential biocatalyst for generating structurally diverse cyclic tripeptides.

Through comprehensive analysis of nine candidate P450s sharing minimum 47% sequence identity, P450<sub>Blt</sub> from *Micromonospora* sp. MW-13 was selected for detailed characterisation. The enzyme displayed remarkably high affinity for its native MRYLH substrate with a  $K_d$  of 2.1  $\mu$ M, binding 10-fold tighter than P450s with related substrates, and achieved >80% conversion in turnover assays.<sup>[92,93]</sup> Liberation of the cyclic tripeptide from the crosslinked product was readily accomplished via trypsin digestion, providing a straightforward route to isolate cyclic tripeptide building blocks. Through systematic comparison with the previously characterised YYH peptide produced by P450 BytO from *Planomonospora* (39% identity), the

researchers established that P450<sub>BLt</sub> installs a distinct C-N crosslink, rather than the C-C crosslink initially reported for biarylites, demonstrating the diverse chemistry possible within this enzyme family.

A systematic analysis of substrate tolerance revealed key structural requirements for P450<sub>BLt</sub> activity. The *N*-terminal leader sequence proved crucial, as removing either methionine alone or the entire Met-Arg sequence impaired catalysis by preventing the essential spin state shift of the P450<sub>BLt</sub> haem iron. This insight led to the successful replacement of methionine with norleucine (Nle), which maintained high activity (85% conversion) while eliminating unwanted sulfoxidation caused by reactive oxygen species during the P450 active cycle. Position-specific effects were also observed, with Arg to Ser/Asn substitutions at position 2 significantly reducing activity (to 15% and 39% respectively), and *C*-terminal amidation substantially decreasing binding affinity ( $K_d$  from 2.1  $\mu$ M to 179  $\mu$ M). The enzyme showed notable flexibility at position 4, accepting various side chain sizes and functionalities including alkyne, alcohol, thiol, and amine groups, though larger residues like tryptophan abolished activity, suggesting a size-limited hydrophobic binding pocket.

While modifications at the crosslinking positions generally resulted in poor acceptance and substantially reduced binding affinities ( $K_d$  values increasing from 2.1  $\mu$ M to 93-530  $\mu$ M), a notable exception emerged when replacing His5 with tryptophan. This substitution maintained high activity (>60% conversion) despite the increased size of the crosslinking residue. Further investigation through comparative NMR analysis and synthesis of a reference C-C crosslinked YLW tripeptide standard revealed that P450<sub>BLt</sub> installs an unexpected A-*O*-B linkage between the tyrosine phenol oxygen and tryptophan indole ring, confirmed by retention of all four tyrosine aromatic protons and hydrogen-deuterium exchange analysis.

This work demonstrates P450<sub>BLt</sub>'s remarkable ability to install different types of crosslinks (C-N in native substrate, A-*O*-B with tryptophan-containing peptides) while maintaining broad tolerance for substrate modifications at specific positions. The combination of this catalytic flexibility with facile leader sequence removal through proteolysis suggests significant potential for generating diverse cyclic tripeptide building blocks for pharmaceutical and biotechnological applications.

## AUTHOR CONTRIBUTIONS

For this study, L.P. analysed several homologues of biarylidade-modifying P450s and suggested a list of proteins for test expression. Of these, P450<sub>Blt</sub> from *Micromonospora* sp. MW-13 was selected. L.P. developed protocols for the heterologous expression of *bytAO* in *Streptomyces coelicolor* to produce Biarylidade YYH as an authentic standard known to have a C-C biaryl bond, to elucidate the alternate C-N crosslink formed by P450<sub>Blt</sub> *in vitro*. L.P. developed the extraction and purification procedures, confirming the compound's purity and structure through high-resolution LC-MS and NMR analysis and providing a sample for comparative analysis, and authored the corresponding methodology chapters.

### 3.1.2 Reassignment of the Structure of a Tryptophan-Containing Cyclic Tripeptide Produced by the Biarylityde Crosslinking Cytochrome P450<sub>Blt</sub>

Laura J Coe, Yongwei Zhao <sup>Ψ</sup>, **Leo Padva** <sup>Ψ</sup>, Angus Keto, Ralf Schittenhelm, Julien Tailhades, Greg Pierens, Elizabeth H Krenske, Max Crüsemann, James J De Voss, Max J Cryle

*Chemistry – A European Journal* **2024** vol. 30, 38, e202400988

<sup>Ψ</sup> Y.Z. and L.P. contributed equally to this work.

Please refer to Appendix B to view the full text of this publication.

#### PUBLICATION SUMMARY

Peptide natural products often gain their biological activity and stability through various crosslinking mechanisms, with side chain crosslinking emerging as an increasingly significant modification.<sup>[94]</sup> While many peptides utilise conventional cyclisation methods such as amide, ester, or disulfide bonds, the discovery of side chain crosslinks, particularly in ribosomally synthesised and post-translationally modified peptide (RiPP) biosynthesis pathways, has expanded our understanding of peptide modification.<sup>[95]</sup> Since the initial discovery of citrilins, biarylitydes, and tryptorubins, bioinformatic analyses have revealed a total of twenty distinct classes of P450-crosslinked peptides.<sup>[29]</sup> These cytochrome P450s have proven particularly adept at creating crosslinks between aromatic amino acid side chains,<sup>[2]</sup> with tyrosine and tryptophan being the most prevalent residues involved in crosslinks, while histidine appears in only three of the currently identified P450-crosslinked RiPP groups including the biarylitydes. This study presents a significant revision to our understanding of peptide crosslinking mechanisms by reassessing the structure of a specific crosslinked peptide produced by the biarylityde crosslinking cytochrome P450<sub>Blt</sub> from *Micromonospora* sp. MW-13.<sup>[96]</sup> Through comprehensive analysis, the research demonstrates that the crosslink in the peptide Met-Arg-Tyr-Leu-His involves a C-N bond between the meta position of tyrosine and the tryptophan indole nitrogen, rather than the previously proposed ether bond. The initial structural misassignment was reconsidered following recent investigations into reversed His-X-Tyr peptides by an engineered variant of P450<sub>Blt</sub>, which highlighted the importance of exchangeable proton analysis in NMR studies of these compounds. The investigation employed several complementary approaches. High-field NMR spectroscopy in DMSO-d<sub>6</sub> solvent proved crucial in simplifying the spectra and preserving important exchangeable OH and NH proton signals.

The change to this polar aprotic solvent markedly simplified the spectra compared to the previously used methanol-d<sub>4</sub>, particularly in the crucial aromatic region. Computational analysis supported the structural reassignment through NMR prediction and structure verification,<sup>[97]</sup> while isotope labelling experiments using deuterated tyrosine provided additional confirmation of the crosslinking position. The structural reassignment was further validated through complementary *in vivo* experiments.

These studies demonstrated that BytO, the archetypal biaryllytase P450<sup>[53]</sup> known for catalysing a C-C crosslink, could process a modified peptide sequence where the native histidine was replaced with tryptophan. The experiments involved site-directed mutagenesis of the *bytA* gene in the plasmid pSET\_ermE\*\_*bytAO* to change its native sequence from MRYYH to MRYLW, followed by conjugation to *Streptomyces coelicolor* M1152. Comparison of the *in vivo* produced peptide with a standard derived from the P450<sub>Blt</sub> generated product showed identical chromatographic behaviour and mass spectrometric fragmentation patterns, confirming that both enzymes install the same C-N crosslink. The significance of this work extends beyond the structural reassignment itself. It reveals the broader substrate tolerance of biaryllytase P450 enzymes and suggests new possibilities for engineered peptide synthesis. The ability of these enzymes to accommodate tryptophan residues, despite their larger size, opens up new avenues for creating diverse crosslinked peptides. This flexibility in substrate recognition challenges previous assumptions about the limitations of these enzymes and suggests they might be capable of processing an even wider range of amino acid combinations.

#### AUTHOR CONTRIBUTIONS

L.P. performed *in vivo* experimental validation of the structural findings, designed and executed the targeted mutagenesis of the pSET\_ermE\*\_*bytAO* plasmid, followed by successful cloning and conjugation of the modified construct into *Streptomyces coelicolor* M1152. The experimental work included cultivation of the resulting strain, extraction procedures, and comprehensive LC-MS analysis, which led to the detection of the *N*-acetylated crosslinked YLW peptide in culture extracts. L.P. conducted further purification of the extract and performed comparative analyses with the P450<sub>Blt</sub> generated standard through both individual and co-injection experiments. Additionally, L.P. was responsible for respective data analysis, figure design, and contributed to writing the corresponding manuscript sections, including the relevant methodological descriptions.

### 3.1.3 Structural Insights into a Side Chain Crosslinking Biarylittide P450 from RiPP Biosynthesis

Mathias H. Hansen, Angus Keto, Maxine Treisman, Vishnu Mini Sasi, Laura Coe, Yongwei Zhao, **Leo Padva**, Caroline Hess, Victor Leichthammer, Daniel L Machell, Ralf B Schittenhelm, Colin J Jackson, Julien Tailhades, Max Crüsemann, James J De Voss, Elizabeth H. Krenske, Max J Cryle

*ACS Catalysis* 2024, 14, 812-826

Please refer to Appendix C to view the full text of this publication.

#### PUBLICATION SUMMARY

Cytochrome P450 monooxygenases (P450s) represent a superfamily of oxidative enzymes that play crucial roles across nature.<sup>[98]</sup> These haem-containing enzymes generate highly reactive intermediates through carefully arranged electron transfers and proton delivery to haem-bound molecular oxygen. While primarily known for hydroxylating unactivated C-H bonds, P450s can perform more complex transformations, particularly the crosslinking of aromatic groups in natural product biosynthesis.<sup>[2]</sup> This crosslinking activity is especially important in several clinically relevant antibiotics, including glycopeptides like vancomycin, teicoplanin and arylomycin.<sup>[99,100]</sup> A significant challenge in understanding these enzymes has been the lack of substrate-bound structures of P450s involved in peptide modification. While previous research has characterised many P450s involved in nonribosomal peptide biosynthesis,<sup>[101,102]</sup> including some in complex with recruitment domains,<sup>[13]</sup> the precise mechanisms of substrate recognition and specificity have remained elusive.<sup>[103]</sup> Recent discoveries of P450s in ribosomal peptide biosynthesis pathways have revealed remarkable diversity in both reactions and substrates,<sup>[104]</sup> with the biarylittides emerging as particularly intriguing examples due to their efficient use of minimal peptide sequences.<sup>[53]</sup>

This work solves the first structure of a peptide modifying P450 in complex with its substrate. A total of three crystal structures was obtained: the apo form of P450<sub>BLT</sub> (2.15 Å), its complex with the pentapeptide substrate MRYLH (1.79 Å), and an E238A mutant (1.95 Å). The substrate bound structure revealed a distinctive binding mode where the peptide lies perpendicular to the I-helix, with its C-terminal carboxylate nestled between the I- and G-helices. This orientation differs markedly from previous predictions and provides crucial insights into substrate recognition. Molecular dynamics simulations over 200 ns demonstrated that substrate binding



involves both direct interactions and an intricate network of water-mediated contacts, explaining the enzyme's ability to accommodate various substrates while maintaining specificity. Density functional theory calculations revealed that the observed C-N crosslinking is not inherently energetically favoured, highlighting the enzyme's crucial role in controlling reaction outcomes. Detailed NMR analysis provided definitive confirmation of the C-N cross-link structure between the tyrosine and histidine residues. Extensive mutational studies were performed to probe the reaction mechanism. A particularly striking finding came from the H234L mutant, which showed dramatically reduced activity (>9% vs. 85% for wild-type), demonstrating the essential role of histidine in substrate coordination. The E238A mutant retained partial activity through a fascinating compensatory mechanism, revealed structurally through the rotation of Arg-242 to maintain the proton transfer pathway.

To validate the structural insights, the biaryllytide P450 BytO from *Planomonospora* sp. was examined, as it naturally forms C-C rather than C-N cross-links, putatively due to its lack of the crucial His-234 residue.<sup>[53]</sup> Introduction of the corresponding histidine through a V219H mutation of BytO in the heterologous expression system *Streptomyces coelicolor* *bytAO* nearly abolished activity (0.6% product formation), while mutation of another conserved residue (Q216A) reduced activity to 7.6%. These results demonstrated how subtle active site differences control reaction specificity in these enzymes.

This work represents a significant advance in understanding P450-catalysed peptide modifications. As the first structure of a P450 bound to a linear peptide substrate, it provides crucial insights into substrate recognition and reaction control. Comparative analysis with other P450s revealed unique features of P450<sub>Blt</sub>, particularly the role of a conserved glutamine residue (Gln-84) as a substrate-binding sensor. These findings lay the groundwork for engineering new P450 catalysts and understanding related enzymes in both ribosomal and nonribosomal peptide biosynthesis pathways.

## AUTHOR CONTRIBUTIONS

L.P. conducted comprehensive mutagenesis studies on P450 BytO from *Planomonospora* sp. to elucidate structure-function relationships through *in vivo* experiments. This included the design and development of a targeted mutagenesis strategy to generate single, double, and triple mutations to the P450, followed by the construction of expression vectors containing these mutated variants. After conjugation and heterologous expression in *Streptomyces coelicolor*, extensive LC-MS analysis was performed to evaluate product formation, including quantitative assessment of enzymatic activities. L.P. was responsible for analysing and interpreting the

resulting data, generating Figures to visualise the findings, and authoring the corresponding experimental sections in the manuscript, including detailed methodological descriptions.

### 3.2 Ribosomal Pentapeptide Nitration for Nonribosomal Peptide Antibiotic Precursor Biosynthesis

**Leo Padva**, Lukas Zimmer, Jemma Gullick, Yongwei Zhao, Vishnu Mini Sasi, Ralf B. Schittenhelm, Colin J. Jackson, Max Cryle, Max Crüsemann

*Chem* 2025, 11, 102438

Please refer to the Appendix D to view the full text of this publication along with the supplemental information.

The following section draws substantially from a comprehensive review article co-authored by the author of this thesis, entitled "The biarylites: understanding the structure and biosynthesis of a fascinating class of cytochrome P450 modified RiPP natural products".<sup>[33]</sup>

#### PUBLICATION SUMMARY

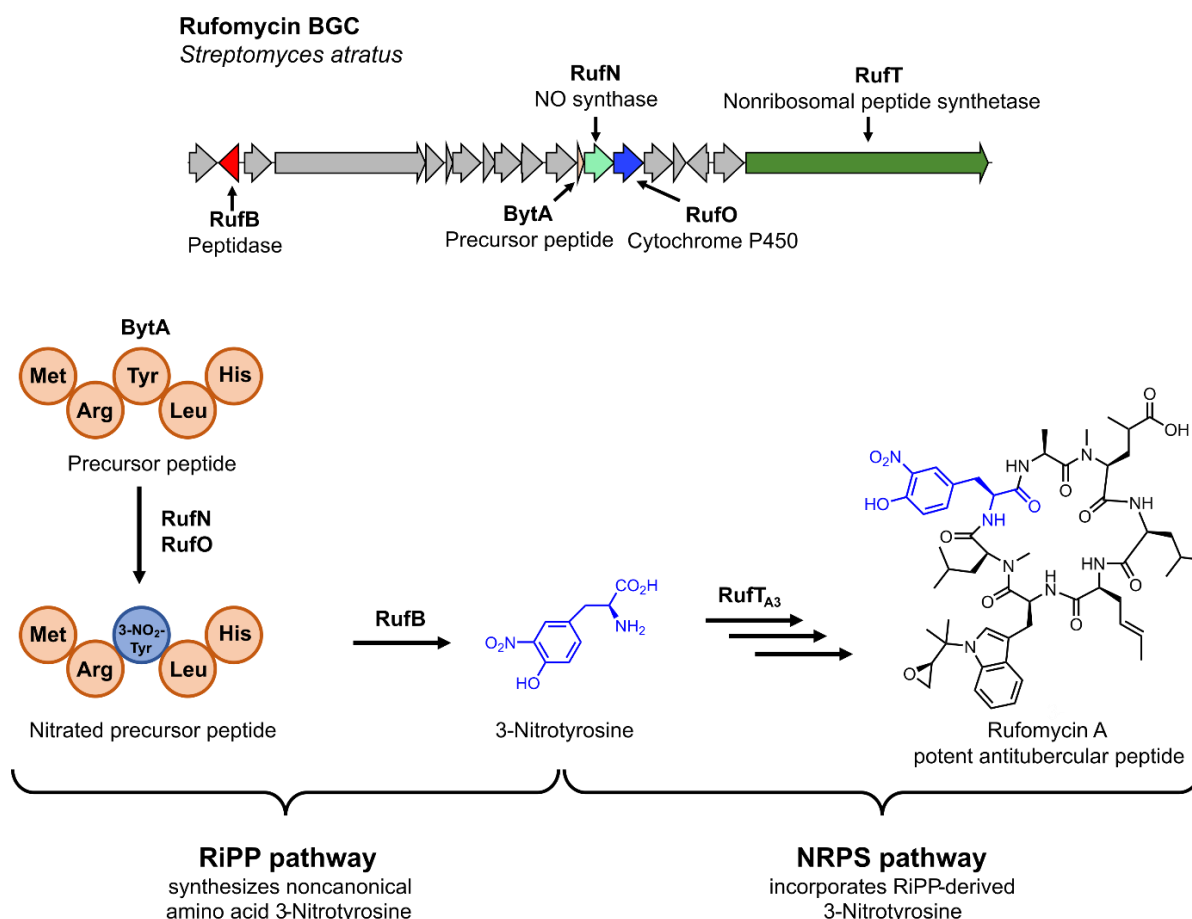
Specialised metabolism produces structurally complex molecules with diverse biological functions, with peptide natural products representing a particularly important class due to their therapeutic applications, especially as antibiotics.<sup>[105]</sup> These peptides are biosynthesised through two major routes: ribosome-dependent pathways producing ribosomally synthesised and post-translationally modified peptides (RiPPs), and ribosome-independent pathways including nonribosomal peptide synthetases (NRPSs).<sup>[30]</sup> The latter encompass both modular thiotemplated megaenzymes and alternative mechanisms of amide bond formation involving transglutaminases, ligases, and tRNA-dependent synthases.<sup>[106]</sup> While RiPPs were traditionally considered more limited in structural diversity due to their reliance on proteinogenic amino acids, recent studies have revealed that their post-translational modifications can generate structural complexity comparable to that of nonribosomal peptides, challenging previous assumptions about the biosynthetic scope of these pathways.<sup>[107–114]</sup>

The biosynthesis of rufomycin, an exceptionally potent antitubercular cyclic heptapeptide that targets the ClpC1 protease in *Mycobacterium tuberculosis*, has remained incompletely understood despite significant research efforts.<sup>[79,84,115]</sup> With over 50 characterised congeners showing improved efficacy against drug-resistant strains, rufomycins contain three non-proteinogenic amino acids including the unique 3-nitrotyrosine. Structure-activity relationships

have shown the nitro group to be crucial for the potent biological activity of rufomycins, yet its biosynthetic origin remained unclear.<sup>[81,86,87]</sup>

The elucidation of 3-NO<sub>2</sub>-Tyr biosynthesis in rufomycin emerged from an unexpected source. Rather than traditional genome mining approaches, a serendipitous observation during manual analysis led to an unconventional discovery.<sup>[88]</sup> In the strain *Streptomyces atratus* S3\_m208\_1, isolated from a belgian soil sample,<sup>[116]</sup> a P450 enzyme with 97.97% sequence identity to RufO was identified. This P450 was found adjacent to *rufN*, which encodes a nitric oxide synthase putatively generating nitric oxide from arginine. Remarkably, upstream of this *rufNO* pair, a *bytA* homologue encoding a MRYLH pentapeptide was discovered (Figure 6). This unexpected arrangement revealed a biarylptide-like BGC embedded within the rufomycin BGC, not only in *S. atratus* but also in all other known rufomycin BGCs. The discovery of a full RiPP pathway nested within an NRPS BGC was unprecedented and highlighted the importance of careful manual analysis in identifying overlooked biosynthetic elements, even within extensively characterised systems.

To elucidate the role of the newly discovered RiPP pathway in rufomycin biosynthesis, *in vivo* experiments were conducted using *S. atratus* S3\_m208\_1, focusing on the key genes *bytA* and *rufO*. A knockout mutant of *bytA* was created in *S. atratus* S3\_m208\_1 using homologous recombination. This  $\Delta$ *bytA* mutant lost the ability to produce rufomycins, demonstrating the essential nature of *bytA*. Importantly, rufomycin production could be restored in this mutant by supplementing the culture with either 3-NO<sub>2</sub>-Tyr or the nitrated pentapeptide, confirming the critical role of the MRYLH peptide in rufomycin biosynthesis. A CRISPR-Cas9-cBEST approach was then used to generate a  $\Delta$ *rufO* mutant.<sup>[117]</sup> This mutant also showed a loss of rufomycin production, mirroring the effect seen in the  $\Delta$ *bytA* mutant. Again, supplementation with either 3-NO<sub>2</sub>-Tyr or nitrated pentapeptide restored rufomycin production in the  $\Delta$ *rufO* mutant, further supporting the involvement of the RiPP pathway in generating the key 3-NO<sub>2</sub>-Tyr building block. While these mutants provided evidence for the use of the nitrated pentapeptide in rufomycin biosynthesis, the question of a release mechanism for 3-NO<sub>2</sub>-Tyr remained. The deletion of another gene, *rufB*, caused a half-fold reduction of rufomycin production.<sup>[84]</sup> RufB contains a serine-aminopeptidase domain, and *in vitro* analysis demonstrated its ability to completely cleave the nitrated peptide, thereby releasing 3-NO<sub>2</sub>-Tyr.



**Figure 6** The rufomycin BGC from *Streptomyces atratus* features a hybrid RiPP-NRPS biosynthetic pathway. The BGC contains genes encoding the precursor peptide BytA (MRYLH), the P450 enzyme RufO, NO synthase RufN, peptidase RufB, and nonribosomal peptide synthetase RufT. The biosynthetic pathway proceeds in two distinct phases: First, in the RiPP pathway, the precursor peptide BytA undergoes selective nitration at its tyrosine residue through the combined action of RufN and RufO, followed by RufB-mediated peptide hydrolysis to release 3-nitrotyrosine. Subsequently, in the NRPS pathway, the RufT synthetase incorporates this modified 3-nitrotyrosine building block into the final rufomycin A scaffold, yielding a potent antitubercular cyclic peptide. This new biosynthetic logic exemplifies how RiPP machinery can be leveraged to generate noncanonical amino acids that serve as building blocks for nonribosomal assembly lines. Figure taken from [33].

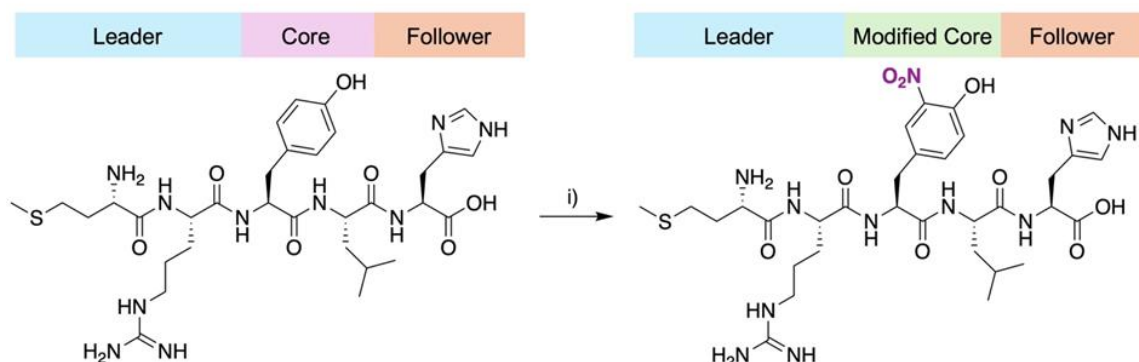
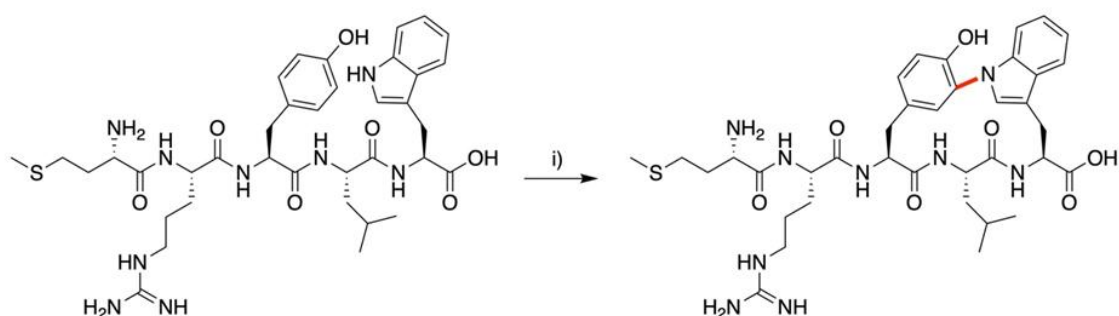
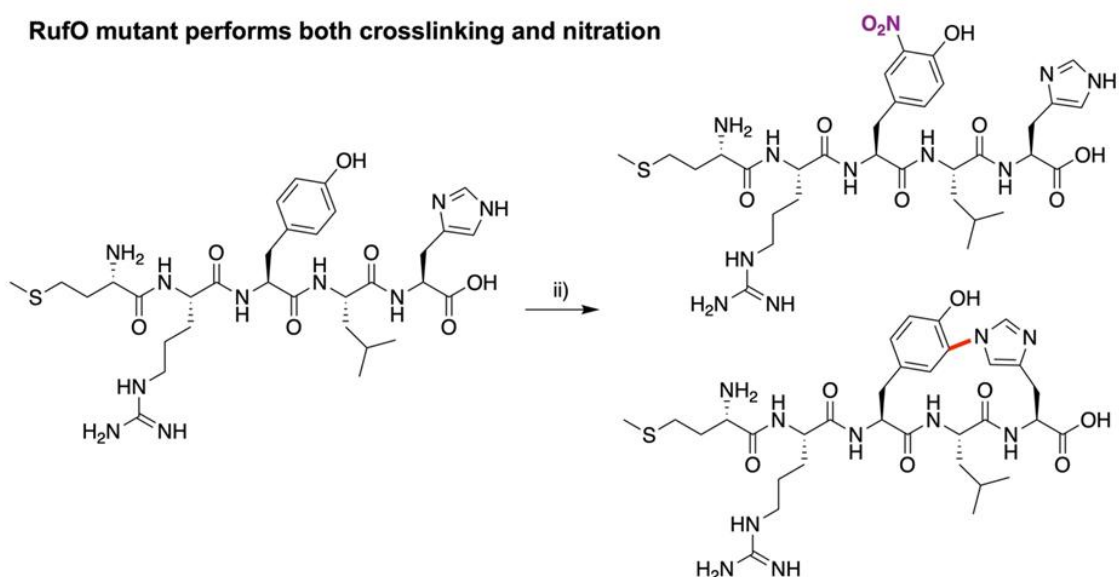
Further *in vitro* characterisation of the RufT A<sub>3</sub> domain revealed how 3-NO<sub>2</sub>-Tyr is activated for incorporation into the rufomycin peptide. Computational modelling and comparative analysis with typical tyrosine-activating domains identified unique features in RufT A<sub>3</sub>, with a critical amino acid substitution in the substrate recognition site expanding the binding pocket, potentially allowing the accommodation of the bulky nitro group of 3-NO<sub>2</sub>-Tyr.

After observing structural similarities between the crosslinking enzyme P450<sub>Blt</sub>, and RufO, and given their shared substrate (MRYLH), an exploration of the binding and activity of this cytochrome P450 was performed *in vitro*.<sup>[88]</sup> Assay protocols were adapted from work on a previously characterised nitrating cytochrome P450, TxtE, that acts to nitrate tryptophan using

diethylamine NONOate as a nitric oxide source.<sup>[17]</sup> These experiments determined that RufO binds the MRYLH peptide with comparable affinity to that seen for the crosslinking enzyme P450<sub>Blt</sub>, although with slight changes in the observed spectra. Activity experiments demonstrated that instead of installing the intramolecular crosslink between the tyrosine and histidine residues that P450<sub>Blt</sub> facilitates, RufO installs a nitro functional group, generating a 3-NO<sub>2</sub>-Tyr residue in the central position of this peptide (Figure 7).<sup>[88]</sup>

Noting the structural similarities between the I-helix of P450<sub>Blt</sub> and RufO, with the exception of two key residues (Ser239 and Val240 in P450<sub>Blt</sub>; Val240 and Pro241 in RufO), subsequent experiments exchanged these residues in RufO (RufO<sub>SV</sub>) to match those found in P450<sub>Blt</sub>.<sup>[88]</sup> The resulting RufO<sub>SV</sub> mutant possessed both crosslinking and nitrating ability in a 5:1 ratio, albeit with reduced total overall conversion (20%) (Figure 7).<sup>[88]</sup> Molecular dynamics simulations suggested that the effects of the RufO I-helix mutations on substrate binding primarily affected the positioning of the histidine-5 residue. This positioning may prevent the peptide from achieving a conformation conducive to crosslinking, which would be advantageous given that the nitration reaction is likely generated by a peroxyne species formed from reactive intermediates preceding the formation of Compound I.<sup>[17]</sup> This hypothesis was supported by the observation that RufO demonstrates crosslinking activity towards a modified MRYLW substrate (Figure 7).<sup>[88]</sup>

Finally, a genome mining approach revealed 19 biosynthetic gene clusters (BGCs) containing homologues of the complete nitrotyrosine biosynthetic system (*bytA* consistently encoding MRYLH, *rufN* and *rufO*) alongside NRPS and polyketide synthase (PKS) genes, as well as additional biosynthetic genes. The clusters were classified by the genomic arrangement of *rufN* and *rufO*, either overlapping (all *ruf* BGCs) or physically separated. The clusters encode diverse serine aminopeptidases or lack peptidases entirely, while tyrosine is predicted as a substrate for their NRPS modules, suggesting various strategies for nitrotyrosine incorporation and peptide processing.

**A) *In vitro* nitration of biaryllyte substrate****B) Mechanistic insights****RufO crosslinks MRYLW substrate****RufO mutant performs both crosslinking and nitration**

**Figure 7** *In vitro* exploration of the cytochrome P450 RufO. (A) *In vitro* nitration assay conditions: i) Peptide (50  $\mu$ M) reacted with RufO<sub>WT</sub> (0.5  $\mu$ M), PuR (0.5  $\mu$ M), PuxB (A105V mutant, 2.5  $\mu$ M),<sup>[118]</sup> DEANO (0.5 mM), NADH (2 mM), glucose (0.33%) and glucose dehydrogenase (0.033 mg/mL) in HEPES (50 mM) and NaCl (50 mM) buffer at pH 7. (B) Mechanistic insights from this investigation including that RufO<sub>WT</sub> crosslinks a MRYLW substrate and the RufO<sub>SV</sub> mutant both crosslinks and nitrates a MRYLH substrate, where conditions (ii) are the same as (i) with the RufO<sub>SV</sub> mutant. Figure taken from <sup>[33]</sup>.

These findings reveal a previously unknown connection between ribosomal and nonribosomal peptide biosynthesis, two pathways long viewed as fundamentally distinct systems. The study represents the first documented case of complete RiPP pathway integration within nonribosomal peptide biosynthesis, establishing a new paradigm for generating nonproteinogenic amino acids. The work demonstrates how a RiPP pathway can be repurposed through selective P450-mediated nitration to produce building blocks for NRPS assembly lines, showcasing an unexpected bridge between these two major biosynthetic routes. The widespread presence of similar gene clusters suggests this strategy of pathway integration might be a more common biosynthetic logic than previously recognised.

## AUTHOR CONTRIBUTIONS

L.P. discovered the biarylittide RiPP cluster within the rufomycin BGC and performed comprehensive biochemical characterisation of the pathway. This included initial characterisation of RufO through protein expression, purification, and enzymatic assays. L.P. established the essential role of the RiPP pathway through genetic manipulation of *Streptomyces atratus*, generating and characterising deletion mutants and performing complementation studies that restored rufomycin production. L.P. designed and validated constructs for the RufT A<sub>3</sub> domain and RufB aminopeptidase, conducting *in vitro* assays that revealed their roles in 3-nitrotyrosine release and incorporation. L.P. assisted in genome mining, optimised protein expression and purification protocols, performed LC-MS analyses, and prepared corresponding figures for the main text and supplementary information. L.P. contributed to manuscript writing and revision.

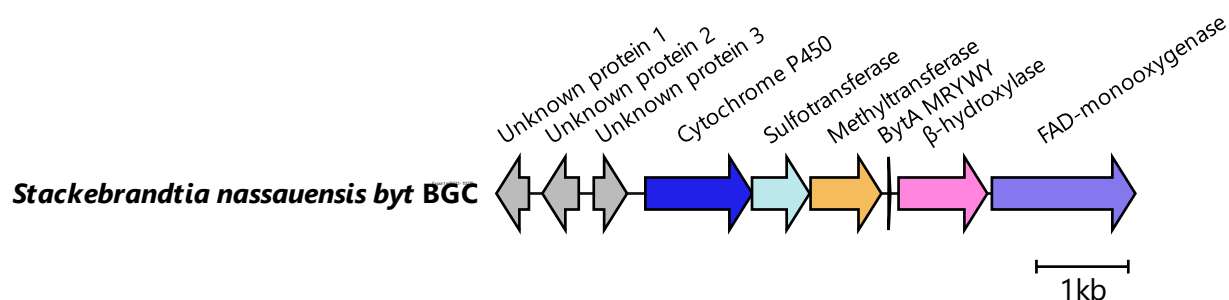


### 3.3 Machine learning-based genome mining of biarylittide BGCs and biochemical validation

The reported biosynthesis of biarylittides, first described by Zdouc et al.,<sup>[52]</sup> was based on two key characteristics: a conserved five-amino acid precursor peptide containing specific aromatic residues (MxYxY/H), and tight genomic coupling between this precursor gene (*bytA*) and its cognate crosslinking enzyme (*bytO*). This idea remained consistent as subsequent studies identified variants with alternative aromatic residue patterns, all maintaining the characteristic proximity between *bytA* and *bytO*.<sup>[29,66,70,71]</sup> However, an unexpected observation in the genome of *Stackebrandtia nassauensis* challenged this understanding.

#### 3.3.1 Bioinformatic discovery of a novel subgroup within the biarylittide family

During systematic analysis of biarylittide BGCs, a *bytO* homologue was identified in *S. nassauensis* that lacked its expected precursor peptide partner within the conventional 500-base pair window. A detailed reinvestigation of the surrounding genomic region revealed an unexpected finding: a *bytA* candidate encoding an unusual MRYWY motif was located approximately 1,400 bp downstream of *bytO*, separated by genes encoding a sulfotransferase and a methyltransferase (Figure 8). This discovery represented both a significant deviation from the established spatial arrangement of biarylittide clusters and introduced a novel precursor-sequence pattern.



**Figure 8** The *Stackebrandtia nassauensis* *byt* BGC showing the atypical positioning and sequence of *bytA*. The cluster includes genes encoding a sulfotransferase, methyltransferase,  $\beta$ -hydroxylase, and FAD monooxygenase.

The identification of this unusual cluster architecture, combined with mounting evidence of greater biarylittide diversity than previously anticipated,<sup>[29,70,88,119]</sup> prompted the development of a machine learning-based genome mining strategy to capture the full biosynthetic diversity of this natural product family.

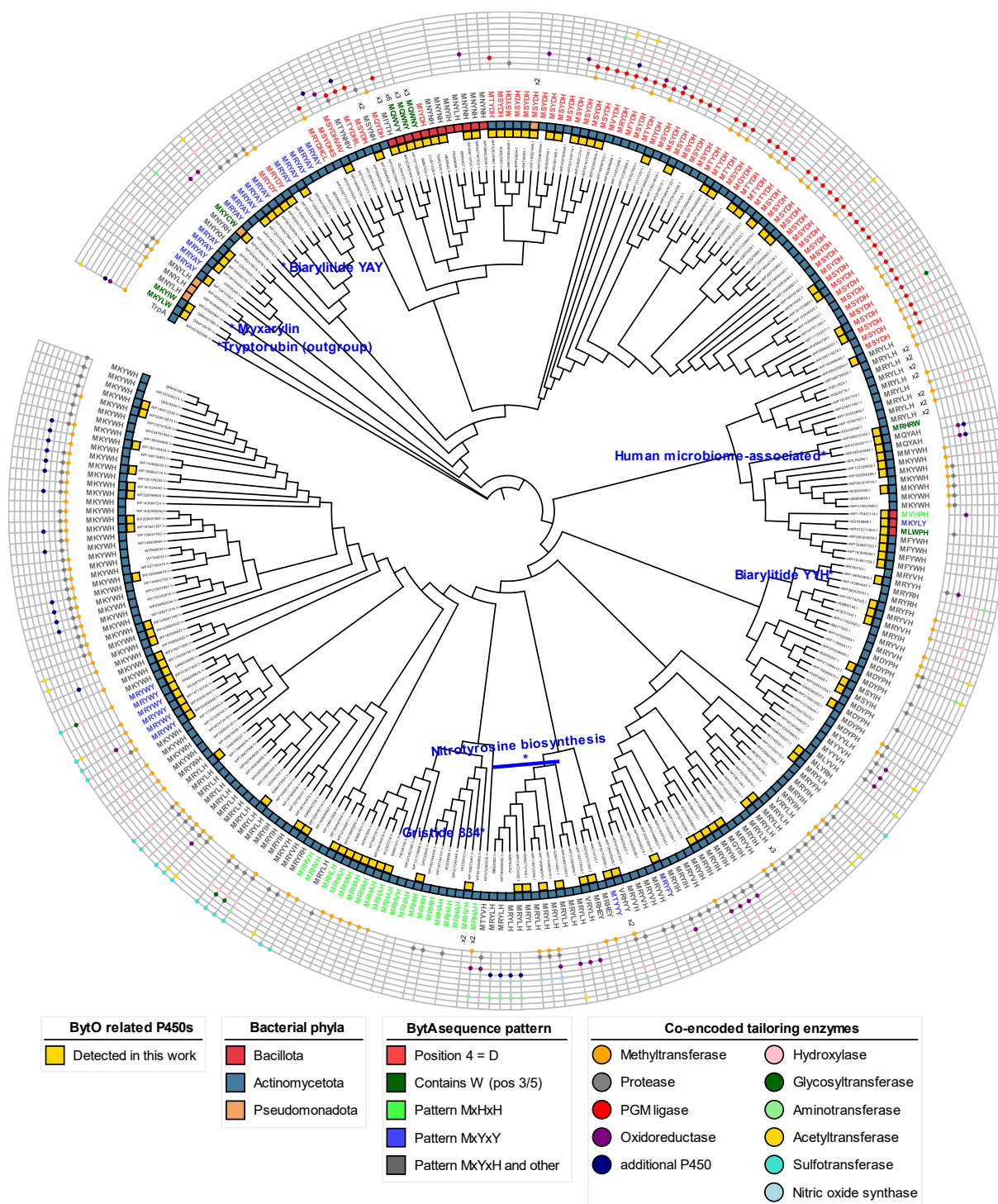
### 3.3.2 Machine-learning based genome mining of biarylittides

Traditional sequence-based approaches for identifying biarylittide biosynthetic genes have significant limitations. While BLAST searches excel at identifying closely related homologues, they prove less effective for detecting biarylittide-modifying P450s due to their characteristically low sequence conservation.<sup>[65]</sup> PSI-BLAST, despite offering greater flexibility through its iterative search approach, often yields numerous false positives by identifying P450s unrelated to biarylittide biosynthesis.<sup>[72]</sup> Furthermore, the detection of *bytA* homologues is hindered by their exceptionally small size (18 bp), which prevents proper annotation and consequently excludes them from current detection pipelines.

The AtropoFinder genome mining pipeline,<sup>[89]</sup> originally developed for the identification of atropopeptide BGCs, was thus adapted for biarylittide detection. Essentially, supervised machine learning (ML) involves training algorithms on labeled datasets to classify previously unseen data.<sup>[120]</sup> In this approach, a Random Forest classifier – a machine learning algorithm that constructs multiple decision trees and merges their outputs<sup>[121]</sup> – was trained using 37 verified atropopeptide-modifying P450s against over 9,000 unrelated P450s, with amino acids grouped by physicochemical properties to enhance generalisation. This strategy enabled the algorithm to identify key sequence features that distinguish atropopeptide-modifying P450s, resulting in the discovery of 683 putative atropopeptide BGCs – substantially outperforming traditional sequence similarity searches.<sup>[89]</sup> The training of the ML classifier and precursor peptide detection was performed by Friederike Biermann (University of Frankfurt, Germany). The ML classifier was trained on carefully curated positive and negative datasets. The positive training set initially combined 5 characterised biarylittide P450 sequences with 181 predicted putative biarylittide P450s from previous studies (55 from Zdouc et al. and 126 from Nam et al.)<sup>[29,52]</sup>. The negative dataset was assembled from diverse P450 sources: the MIBiG database (n=277)<sup>[122]</sup>, verified BGCs not in MIBiG (n=12), predicted atropopeptides from a BLAST search on WP\_007820080.1 of tryptorubin A (n=49)<sup>[28]</sup>, other predicted cyptide classes by Nam et al. (n=787)<sup>[29]</sup>, and P450s from the antiSMASH database annotated as 'cytochrome p450' (n=14,665).<sup>[62]</sup> Following dereplication at 95% sequence identity using Cdhit V4.8.1<sup>[123]</sup> with a word size of 5, the final training sets used for the classifier contained 106 positive and 9,761 negative P450 sequences. This reduction in the positive set occurred because the dereplication process removed highly similar sequences to prevent redundancy in the training data.

For precursor peptide detection, CoreFinder, also originally developed for atropopeptide mining,<sup>[89]</sup> was employed, as conventional RiPP genome mining tools are unable to detect the

unusually short five-amino-acid biarylptide precursor peptides. CoreFinder screens for open reading frames in the genomic vicinity ( $\pm 3000$  bp) of identified P450s, specifically searching for sequences between 5-8 amino acids containing aromatic amino acids at positions 3 and 5. Application of this integrated approach to the NCBI genpept database (18.03.2024)<sup>[124]</sup> identified 1,128 putative biarylptide-modifying P450s. After removal of duplicates and manual curation, 277 unique BGCs were identified, each containing both a P450 and corresponding precursor peptide. The genomic context of these clusters was further analysed using RODEO2 to detect co-encoded tailoring enzymes through Pfam domain identification.<sup>[125]</sup> For phylogenetic analysis, protein sequences were aligned using MUSCLE<sup>[126]</sup> and the alignment was submitted to the IQ-tree webserver using standard settings.<sup>[107–109]</sup> The resulting treefile was visualised and annotated using the iTOL webserver (Figure 9).<sup>[127,128]</sup>



**Figure 9** Phylogenetic tree showing the diversity and distribution of putative biarylittide-modifying P450s identified through genome mining. The integrated visualisation highlights the taxonomic distribution across bacterial phyla, characteristic BytA sequence patterns, and the variety of co-encoded tailoring enzymes, revealing distinct clades of potential biarylittide biosynthetic systems.

The ML-based genome mining approach substantially expanded the known diversity of biarylittide biosynthesis by identifying 124 novel P450s and bringing the total to 277 putative members. A complete list with all identified P450 IDs, species names and co-occurring proteins can be found in Appendix E. This comprehensive analysis confirmed the previously observed

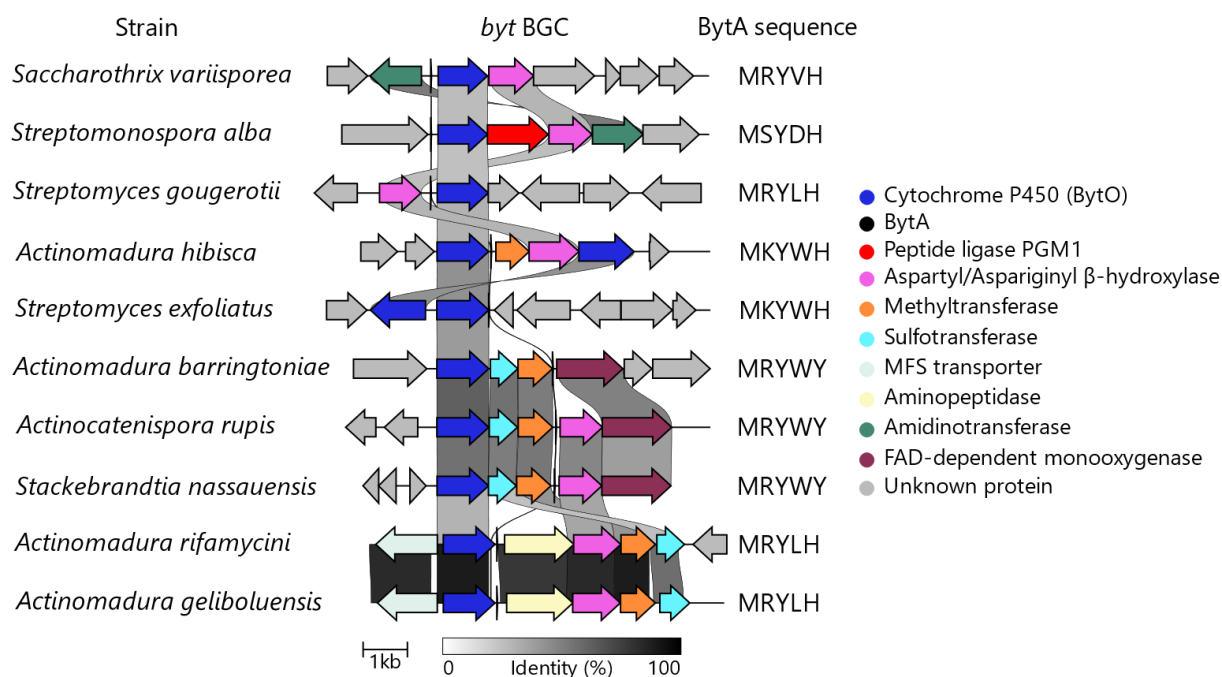
phylogenetic distribution of biarylittide BGCs across bacterial phyla, with predominant representation in Actinomycetota and additional presence in Bacillota and Pseudomonadota. Notably, the analysis revealed potential biarylittide biosynthetic capability in human microbiome-associated bacteria, specifically identifying *bytO* homologues and MQYAH precursor peptides in *Rothia dentocariosa* and *Rothia* sp. HMSC058E10. Both these Gram-positive bacteria are naturally present in the oral and respiratory microbiota,<sup>[129]</sup> representing the first identification of potential biarylittide production in human-associated microorganisms. The phylogenetic analysis further unveiled several distinct precursor sequence patterns and motif groups. The previously identified *Stackebrandtia nassauensis* *byt* BGC, notable for both its MRYWY motif and its unusual 1,400 base pair separation between *bytA* and *bytO* (see chapter 3.3.1), emerged as the founding member of a conserved clade comprising six BGCs sharing identical BytA motifs and similar tailoring enzyme compositions. Furthermore, the analysis revealed nine novel precursor peptide motifs containing tryptophan (W) at crosslinking positions, including MKYLW and MKYIW variants in *Actinomadura* strains. These findings are particularly significant given that tryptophan at position 5 has been experimentally validated to lie within the catalytic scope of both P450<sub>Blt</sub> from *Micromonospora* MW-13 *in vitro* and BytO of *Planomonospora* sp. *in vivo* (see chapters 3.1.1 and 3.1.2.). A particularly noteworthy discovery was made in four *Lysinibacillus* strains (phylum Bacillota), which harbour the MQWNY motif with tryptophan at position 3. These strains exhibited remarkable *bytA* copy number variation, with one strain containing six copies. Multiple precursor peptides have been reported for various RiPP classes such as cyanobactins, with reports of up to 10 precursor peptides leading to increased production of the compounds.<sup>[130]</sup>

The analysis also revealed distinct patterns in the co-occurrence of tailoring enzymes and specific precursor sequences. A striking example is the frequent association between Pgm1 ligase homologues and precursor motifs containing aspartic acid (D) at position 4, observed in 47 gene clusters. The Pgm1 ligase, classified as an ATP-dependent amide bond-forming enzyme of the ATP-grasp superfamily, was previously characterised in guanidinotide biosynthesis.<sup>[131]</sup> This enzyme catalyses the conjugation of ribosomal peptides with non-proteinogenic amino acids suggesting another case of biarylittide pathway integration with distinct biosynthetic logic. This integration of different biosynthetic pathways is further exemplified by gene clusters involved in nitrotyrosine biosynthesis, which contain co-encoded nitric oxide synthases and form a distinct phylogenetic clade. These clusters represent the recently characterised rufomycin-associated biarylittide pathways (see chapter 3.2). These findings substantially expand our understanding of biarylittide biosynthetic diversity and reveal

systematic patterns in the distribution of precursor sequences and tailoring enzymes across bacterial phyla. The discovery of novel motifs, particularly those containing tryptophan at crosslinking positions, combined with the identification of conserved relationships between specific precursor sequences and tailoring enzymes, provides new insights into the evolutionary diversity and biosynthetic potential of this natural product family. This approach demonstrates how machine learning (ML) tools developed for one P450-modified RiPP class <sup>[89]</sup> can be successfully repurposed for another.

### 3.3.3 Examination of biarylittide chemical diversity

The comprehensive bioinformatic analysis provided the foundation for experimental validation and exploration of biarylittide chemical diversity. Ten representative strains were selected for detailed investigation, chosen to span diverse bacterial genera and distinct phylogenetic clades identified in the analysis (Figure 10). Eight strains were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ), with two additional strains provided by Naicons Srl., Milan, Italy.



**Figure 10** Overview of the *byt* gene clusters examined in this work. Links between genes indicate sequence similarity.

Metabolomic analysis was conducted through both liquid media and solid agar plate cultivation using standard actinobacterial growth conditions (chapter 5.2.1). Secondary metabolite profiles were analysed through Diaion® HP-20 resin extraction followed by LCMS analysis (chapter

5.3.2). The resulting mass spectra were processed using molecular networking on the Global Natural Products Social Molecular Networking (GNPS) web platform to facilitate identification of related compounds and structural analogs.<sup>[54]</sup>

The experimental investigation followed three complementary strategies: (i) Analysis of wildtype producers (ii) Heterologous expression of partial and putatively complete gene clusters and (iii) Characterisation of selected BytO homologues through *in vitro* biochemical studies. The investigated strains and analyses performed are summarised in Table 1.

**Table 1** Summary of strains investigated, including type of analysis performed. The order of strains is matching Figure 10.

Strain and source	BytO ID	Experiment*
<i>Saccharothrix variisporea</i> DSM 40234	RKT69193.1	W, H, C, P
<i>Actinomadura barringtoniae</i> DSM 108647	WP_208263840.1	W, H, P
<i>Actinomadura rifamycini</i> DSM 43936	WP_034520669.1	W, H
<i>Actinocatenispora rupis</i> DSM 45178	GID12472.1	W, H, P
<i>Actinomadura hibisca</i> 83825 (Naicons)	WP_067473018.1	W, H, P
<i>Streptomyces exfoliatus</i> (Naicons)	WP_024761848.1	W, P
<i>Actinomadura geliboluensis</i> DSM 45508	WP_138637966.1	W, H
<i>Streptomyces gougerotii</i> DSM 40234	WP_191867971.1	W, H, C, P
<i>Stackebrandtia nassauensis</i> DSM 44728	WP_013017934.1	W, H, P
<i>Streptomonospora alba</i> DSM 44588	WP_052809996.1	W, H, C

\*Experiment: W = Wildtype cultivated, H = Heterologous expression of *byt* BGC, C = Compound detected, P = Protein expressed

Comprehensive analysis revealed that only *Saccharothrix variisporea* produced detectable levels of its predicted biarylittide YVH and traces of a putatively hydroxylated variant under laboratory conditions among the wildtype strains. This limited success with native producers led to the pursuit of a heterologous expression strategy. At first it was tried to establish the heterologous expression of the *byt* BGC in *Escherichia coli* using the pCDFDuet-1 vector (Table 8), to simplify the workflow and reduce cultivation times. However, after successful cloning, transformation and cultivation of the resulting strain *E. coli* BL21 pCDFDuet-1-*bytAO* (see chapters 5.2.1 and 5.4), no new signals were detected in LC-MS analysis in the cultivation extracts, in contrast to the expression of the same genes in *Streptomyces coelicolor*, which was henceforth used as the main expression host.

Each *byt* gene cluster was PCR-amplified from genomic DNA using primers containing overhangs compatible with the genome integrating pSET152\_ermE\* expression vector (Table 8). The amplified gene clusters were integrated into the vector through Gibson assembly (chapter 5.4.7), verified by restriction digest (chapter 5.4.3) and Sanger sequencing (chapter 5.4.8), and introduced into three *Streptomyces* hosts (*S. coelicolor*, *S. lividans*, and *S. albus*) via conjugation using *E. coli* ET12567/pUZ8002 (Table 8). Of the ten examined clusters, three successfully yielded novel compounds through heterologous expression:

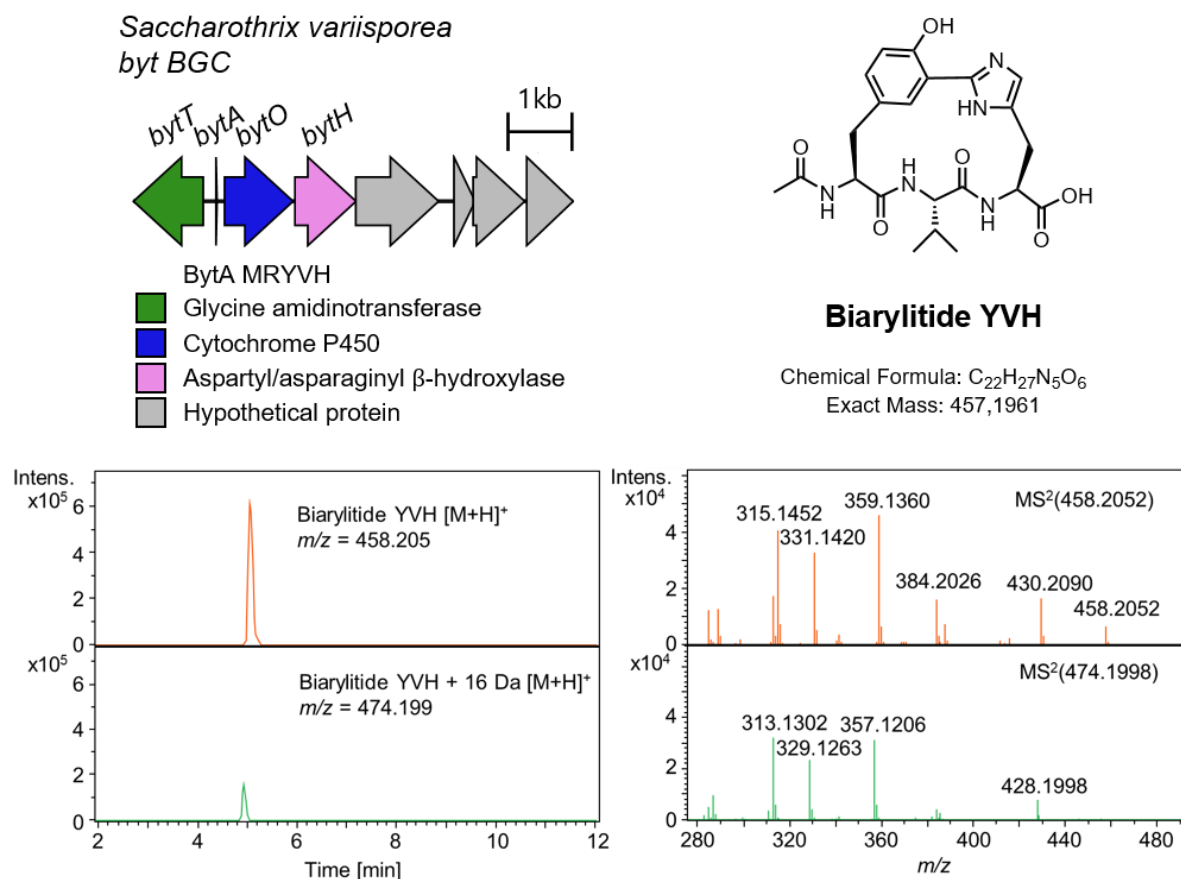
1. *Saccharothrix variisporea* - Biarylittide YVH and a putative hydroxylated variant.
2. *Streptomyces gougerotii* - Biarylittide YLH.
3. *Streptomonospora alba* - Biarylittide YDH.

These successful cases, representing distinct positions within the phylogenetic tree, showcase different aspects of biarylittide biosynthetic diversity. The following chapters detail the characterisation of each system, beginning with the extensively studied *byt* cluster from *S. variisporea*.

### **3.3.4 *Saccharothrix variisporea* cluster: Discovery and characterisation of Biarylittide YVH and its modified variant**

The *S. variisporea* *byt* BGC emerged as the most productive system in this study, yielding detectable compounds in both wildtype and heterologous expression systems. This gene cluster is particularly noteworthy for its expanded biosynthetic potential through additional tailoring enzymes. Beyond the canonical *bytA* and *bytO* genes, the cluster contains *bytT*, encoding an amidinotransferase, and *bytH*, predicted to encode an aspartyl/asparaginyl  $\beta$ -hydroxylase (Figure 11). Initial analysis of extracts from both wildtype and heterologous expression revealed two distinct compounds: the expected Biarylittide YVH and a second compound with a mass shift of +16 Da, suggesting hydroxylation.

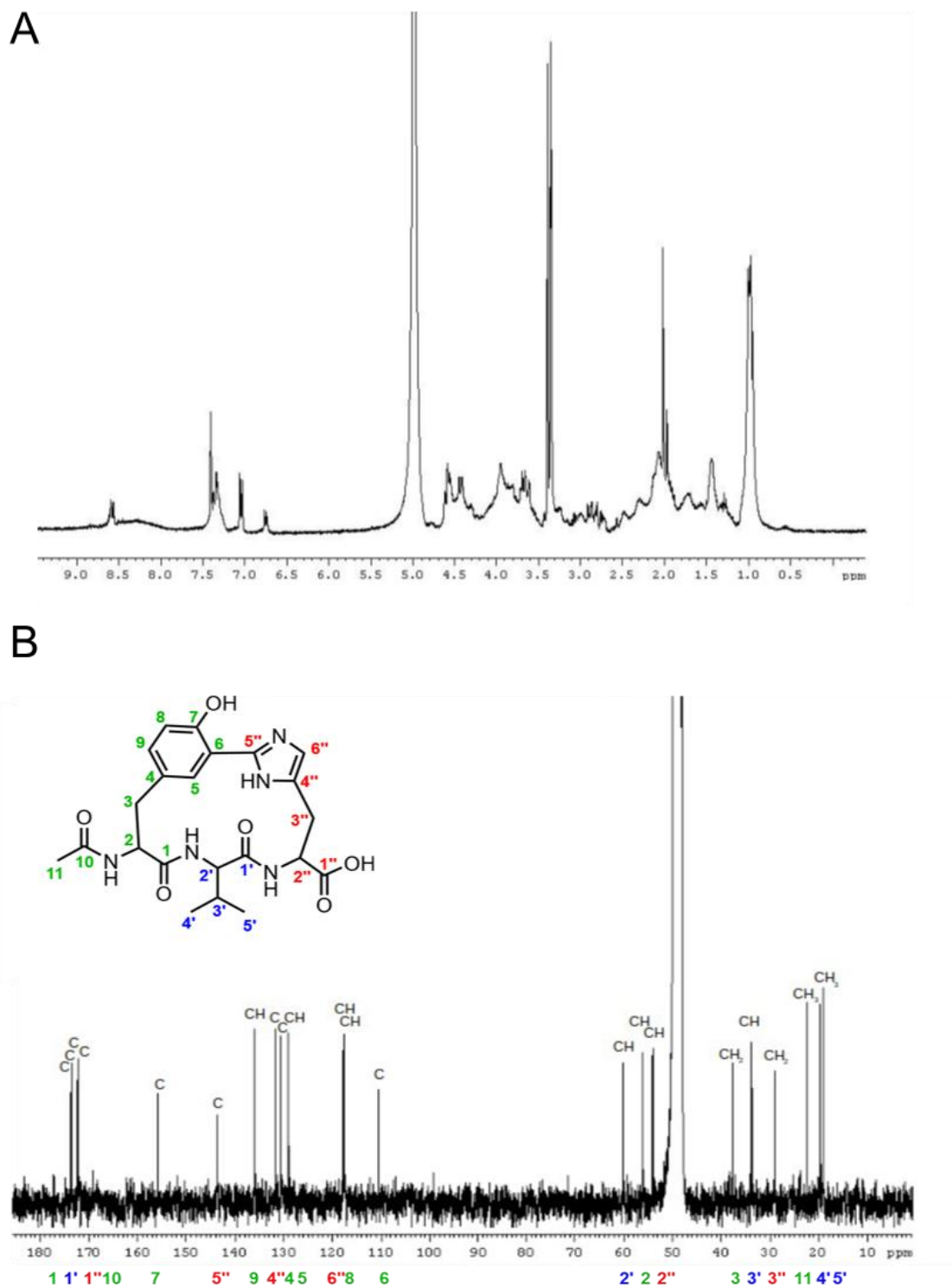




**Figure 11** Schematic representation of the *byt* BGC in *Saccharothrix variisporea* (top left); Molecular structure of Biarylittide YVH (top right); LCMS traces of two new signals (bottom left) and MS<sup>2</sup> fragmentation spectra of those compounds (bottom right).

For isolation of biarylittide YVH, 15 litres of main culture *Streptomyces coelicolor* pSET\_ermE\*-bytAOHT (Figure 13) were cultivated, extracted and purified according to the procedures outlined in chapter 5.3. After multiple rounds of semi-preparative HPLC, 8.7 mg of a white substance was isolated, which was sufficient for structural elucidation through comprehensive NMR analysis recorded in deuterated methanol (CD<sub>3</sub>OD) as a solvent (Figure 12). The <sup>13</sup>C-NMR spectral analysis provided definitive structural confirmation through direct comparison with the previously reported data for biarylittide YYH (Figure 12 and Table 2).<sup>[52]</sup> The carbon chemical shifts demonstrated precise correlation with the characterised biarylittide YYH structure throughout all conserved regions (Table 2), with predictable variations observed only at positions corresponding to the valine substitution (positions 3', 4', and 5') relative to the original tyrosine. Critical to the structural elucidation was the determination of the Tyr-His crosslink configuration. Based on the 68.78% sequence homology between BytO and its *S. variisporea* homologue, a C-C crosslink was hypothesised. This is conclusively supported by the chemical shift of tyrosine C6 at  $\delta_C$  110.3 ppm, which falls within the diagnostic range ( $\delta_C$  108-115 ppm) characteristic of C-C crosslinks

in these systems,<sup>[33]</sup> matching the signal observed in biaryllyte YYH and is distinctly upfield from the chemical shift range associated with C-N crosslinks ( $\delta_C$  125-130 ppm).<sup>[33,66,96]</sup>



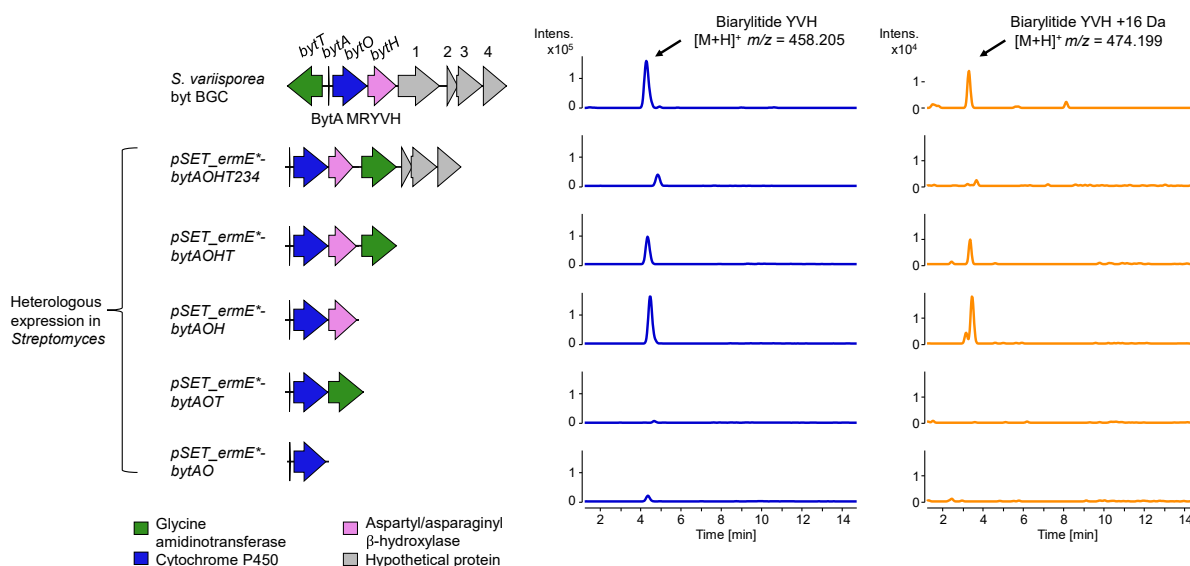
**Figure 12** (A)  $^1\text{H}$ -NMR spectrum of biaryllyte YVH in deuterated methanol. (B) Carbon assignment of signals from  $^{13}\text{C}$ -NMR spectrum of biaryllyte YVH in methanol. To elucidate the structure, the signals were compared with the measurements of biaryllyte YYH.<sup>[52]</sup>

**Table 2** NMR assignments related to Figure 14. Position 6 shows the characteristic shift for a C-C bond.

	Biarylityde YVH (this study) $\delta_c$ (type) in CD <sub>3</sub> OD	Biarylityde YYH Zdouc et al. <sup>[52]</sup> $\delta_c$ (type) in CD <sub>3</sub> OD
<b>Tyrosine</b>		
1 (C)	173.6	173.5
2 (CH)	56.0	55.5
3 (CH <sub>2</sub> )	37.8	37.2
4 (C)	130.5	130.3
5 (CH)	129	128.7
6 (C)	110.3	110
7 (C)	155.8	155
8 (CH)	117.4	117
9 (CH)	136.0	136
10 (C)	172.0	171.9
11 (CH <sub>3</sub> )	22.6	22.3
<b>Valine</b>		
1' (C)	173.5	-
2' (CH)	60.2	-
3' (CH <sub>2</sub> )	33.8	-
4' (CH <sub>3</sub> )	19.9	-
5' (CH <sub>3</sub> )	19	-
<b>Histidine</b>		
1'' (C)	172.2	172.3
2'' (CH)	54	54
3'' (CH <sub>2</sub> )	29	29
4'' (C)	131.5	131.4
5'' (C)	144	142
6'' (CH)	117.4	117

To systematically investigate the roles of the various tailoring enzymes, a series of genetic constructs were designed and heterologously expressed in various strains including *S. coelicolor*, *S. lividans* and *S. albus*. Although production levels of both biarylityde YVH and the hydroxylated variant among the strains were comparable, *S. coelicolor* was chosen due to favourable growing behaviour resulting in higher amounts of biomass. Of particular interest

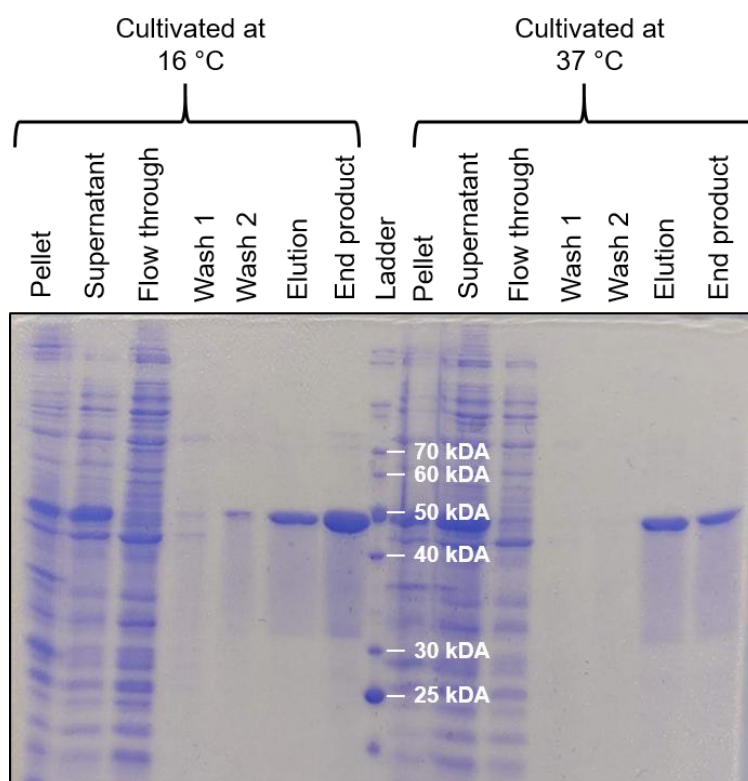
was BytH, the putative  $\beta$ -hydroxylase hypothesised to be responsible for the +16 Da modified biarylittide. Despite multiple attempts, this second compound could not be isolated in sufficient quantities for full structural characterisation. Six genetic constructs were designed to elucidate the roles of various genes in the cluster (Figure 13). Notably, while pSET\_ermE\*-bytAOHIT was successfully cloned (construct not shown), multiple attempts at conjugation to *Streptomyces* hosts failed.



**Figure 13** Schematic representation of *S. variisporea* byt BGC (top left) and cloned constructs to be heterologously expressed to examine enzyme function (bottom left). Upstream of *bytA* is the designated gene *bytT*, which encodes an amidinotransferase domain containing enzyme, while downstream of *bytO* is *bytH*, which is predicted to encode  $\alpha$ -ketoglutaric acid-dependent  $\beta$ -hydroxylase. Downstream of *bytH* are four more genes in the same orientation but no annotated function. Below are the six constructs that were designed based on this genomic region of *S. variisporea*. Due to its different orientation, *bytT* was cloned downstream of *bytH* to align all the examined genes. pSET\_ermE\*-bytAOHT was cloned successfully but conjugation to *Streptomyces* hosts failed on multiple attempts, in contrast to the other gene clusters. LC-MS chromatograms of *S. variisporea* wildtype carrying a byt BGC and genetic constructs based on this cluster expressed in *S. coelicolor* M1152. The first row of chromatograms shows signals corresponding to the [M+H]<sup>+</sup> ion of Biarylittide YVH (blue traces), while the second row shows signals corresponding to the [M+H]<sup>+</sup> ion of Biarylittide YVH +16 Da (orange traces).

To further investigate the putative hydroxylase BytH, its DNA sequence was amplified from genomic DNA of *Saccharothrix variisporea* using specifically designed forward and reverse primers (Table 13). The amplified gene was subsequently cloned into linearised pET-28a(+) expression vector (see Table 8 and Table 11) via Gibson assembly, following the protocol described in 5.4.7. Recombinant protein expression was performed in *E. coli* BL21(DE3) cells under IPTG induction (see 5.5.1), followed by Ni-NTA affinity chromatography purification (see 5.5.2). SDS-PAGE analysis of the purified protein (see Chapter 5.5.4), revealed an unusual

migration pattern, with BytH consistently appearing at approximately 50 kDa, which is significantly higher than its theoretically calculated molecular mass including the His-tag (41.7 kDa). This anomalous migration behaviour persisted across multiple expression experiments conducted at different temperatures (16 °C and 37 °C) and various induction conditions, suggesting this characteristic is intrinsic to the protein rather than an artifact of expression conditions (Figure 14).



**Figure 14** SDS-Page of samples of the protein purification workflow for BytH tested at two different growth conditions, 16 °C (left) and 37 °C (right). The purified protein appears to run at a higher size (slightly below 50 kDa) than the expected size for the purified protein including His-tag is 41.7 kDa.

*In vitro* characterisation of BytH was attempted using both the isolated biaryllytite YVH and the available linear peptide (MRYLH) as substrates according to published assay conditions (chapter 5.5.6).<sup>[132]</sup> While no new products were detected in these assays, it is important to note that the linear peptide differed from the natural substrate (MRYPVH). The absence of hydroxylation activity could be attributed to several factors. First, BytH may exhibit strict substrate specificity, requiring the exact amino acid sequence of the natural substrate. Second, the unusual migration pattern observed during SDS-PAGE analysis might indicate post-translational modifications or structural changes that could affect enzyme functionality. Third, the recombinant enzyme might lack essential cofactors or interaction partners present in the native host. Some preliminary information about the position of the hydroxy group can be

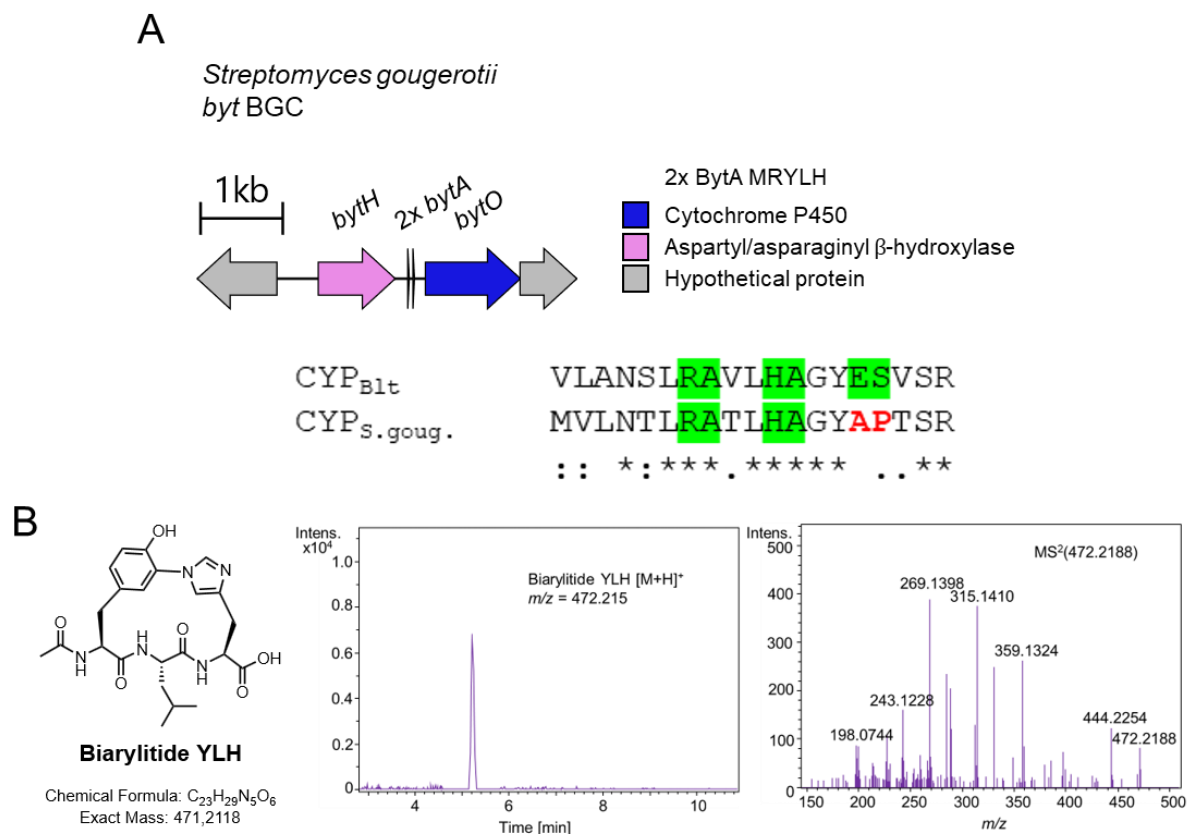
gained from the MS fragmentation patterns of the +16 Da compound (Figure 11), which revealed neutral valine loss, suggesting hydroxylation likely occurs at a position other than valine.

For future experiments, chemical synthesis of the natural substrate MRYVH would be a logical next step to determine whether the hydroxylase exhibits the expected activity with its native substrate. Additionally, alternative expression systems, cofactor supplementation, and optimisation of assay conditions could be explored to successfully reconstitute the hydroxylase activity *in vitro*. Co-expression with potential interaction partners identified in the BGC might also facilitate proper folding and activity of BytH.

### **3.3.5 *Streptomyces gougerotii* cluster: A P450 with unconventional catalytic residues**

The biarylptide BGC identified in *S. gougerotii* presents two notable features: two identical BytA precursor peptides and a BytO enzyme containing unusual residues in its I-helix region (Figure 15). P450 enzymes typically contain a conserved glutamic acid (E) and serine (S) pair that is crucial for catalytic activity, and this motif has been specifically demonstrated to be important for crosslinking reactions in biarylptide biosynthesis (chapter 3.1.3). Remarkably, in this cluster's BytO, these residues are replaced with alanine (A) and proline (P), representing a significant deviation from the established catalytic motif, posing questions on how formation of the crosslink is yet catalysed.

Heterologous expression of the *S. gougerotii* *byt* BGC successfully produced a new compound corresponding to the predicted *N*-acetylated tripeptide Biarylptide YLH (Figure 15). Although the cluster contains a *bytH* gene encoding an aspartyl/asparaginyl  $\beta$ -hydroxylase, no additional modified compounds were detected in the cultivation extracts.

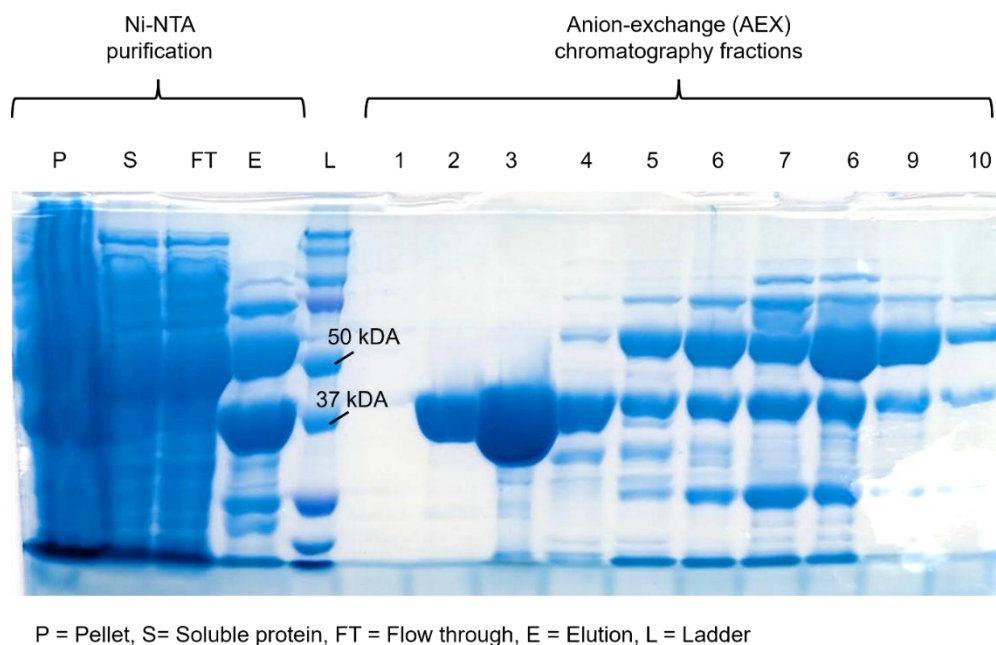


**Figure 15** (A) Schematic representation of the *byt* BGC found in the genome of *Streptomyces gougerotii* including the duplicated *bytA* gene and a section of an alignment of the characterised P450<sub>Blt</sub> and CYP<sub>S.goug.</sub> showing alterations of the crucial alcohol acid pair (ES→AP). (asterisk\*)=identical residues, (colon)=conservation between groups with strongly similar properties, (period)=conservation between groups with weakly similar properties. (B) Molecular structure of expected Biarylityde YLH LCMS traces of a new signal and fragmentation spectrum of this peak.

For closer inspection of BytO<sub>S.goug.</sub>, the DNA sequence encoding the protein was commercially synthesised with codon optimisation for *E. coli* expression from Twist Biosciences (San Francisco, USA). The synthetic gene construct was delivered pre-cloned into the pET28(+) expression vector (see 5.1.3), which incorporates an *N*-terminal His6-tag for affinity purification. The recombinant plasmid was transformed into *E. coli* ArcticExpress (DE3) cells (see 5.1.2 and 5.2.3), specifically selected for their ability to co-express cold-adapted chaperonins (Cpn10 and Cpn60) that facilitate proper protein folding at lower temperatures. [133] Protein expression was conducted under optimised conditions following the protocol detailed in section 5.5.1, with induction using IPTG at reduced temperature (16 °C) to enhance soluble protein production. The recombinant protein was subsequently purified using a two-step chromatographic approach. Initial purification was performed via Ni-NTA affinity chromatography (see 5.5.2), followed by anion-exchange chromatography (see 5.5.3) to remove contaminating proteins. The purity and apparent molecular weight of BytO from *S. gougerotii*



were assessed by SDS-PAGE analysis (see 5.5.4), which confirmed the presence of a protein band corresponding to the expected molecular weight (Figure 16).



**Figure 16** SDS-PAGE analysis of BytO from *S. gougerotii* following two-step purification. The left panel shows fractions from Ni-NTA affinity chromatography: pellet (P), soluble fraction (S), flow-through (FT), and elution (E), alongside a protein molecular weight ladder (L). The right panel displays fractions 1-10 collected during anion-exchange chromatography (AEX) purification. BytO appears as a prominent band at the expected size of 44.03 kDa. The highest purity is observed in fractions 2 and 3, which were pooled and used for further analysis.

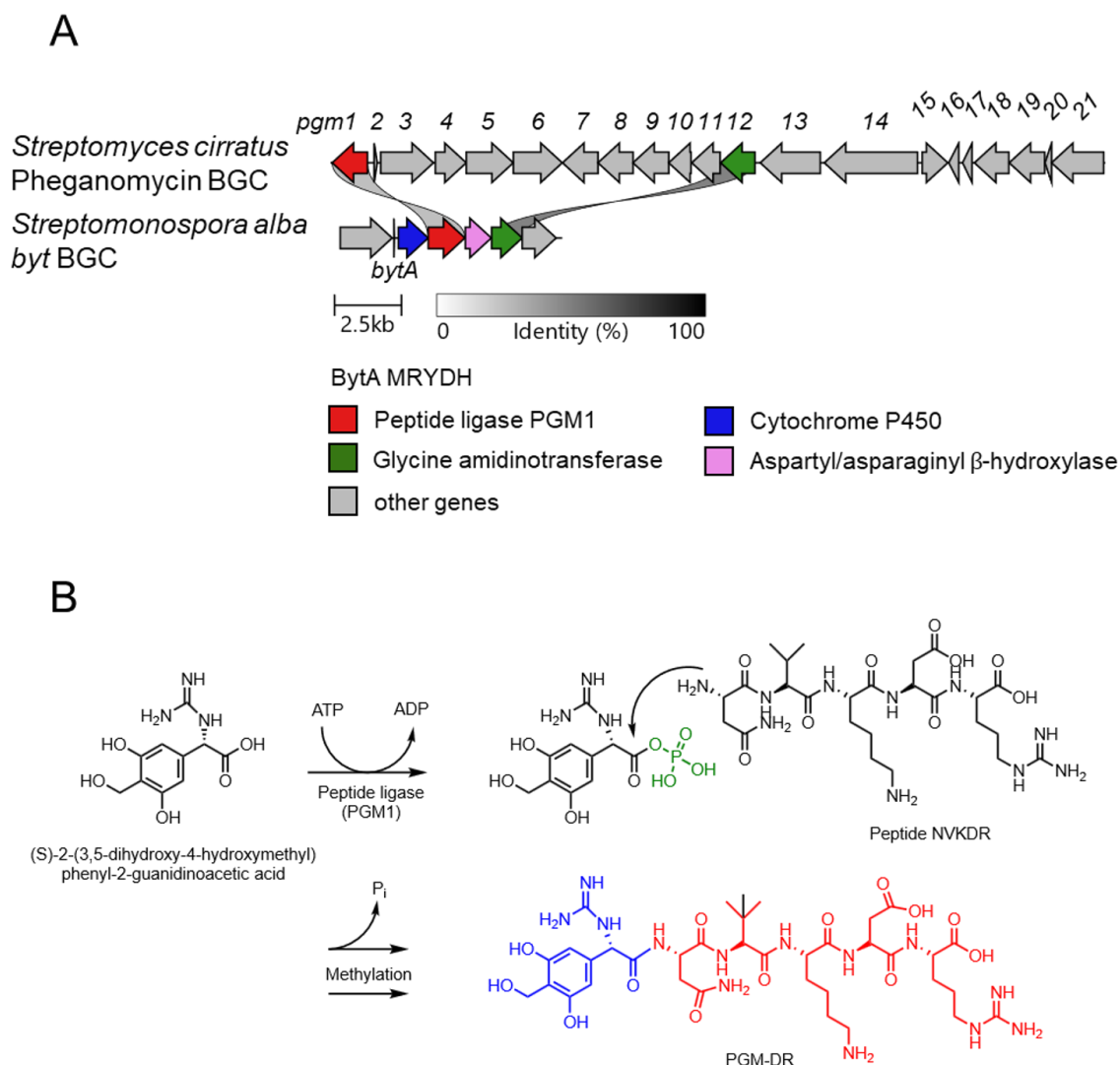
*In vitro* turnover assays (see 5.5.5) of the purified BytO from *S. gougerotii* demonstrated that the enzyme maintains activity with its cognate substrate MRYLH, despite the unconventional I-helix residues. However, the enzyme showed relatively low binding affinity and achieved only approximately 30% conversion (data not shown). Notably, introduction of a P228S mutation in BytO, designed to partially restore the canonical motif, did not significantly alter these catalytic parameters, suggesting that other structural features may compensate for the unusual I-helix residues, as seen with the E238A mutant of P450<sub>Blt</sub>.<sup>[134]</sup>

### 3.3.6 *Streptomonospora alba* cluster: Integration of biarylittide and pheganomycin biosynthetic machinery

The genome of *Streptomonospora alba* contains a remarkable *byt* BGC that combines elements from both biarylittide and pheganomycin biosynthesis. A key feature of this cluster is a Pgm1 homologue that shares 36% sequence identity with the specialised peptide ligase involved in pheganomycin biosynthesis in *Streptomyces cirratus*. This Pgm1 ligase is the same type as



previously discussed (Chapter 3.3.2), catalysing the unusual ligation of a ribosomally synthesised pentapeptide to a non-proteinogenic amino acid during pheganomycin assembly.<sup>[131]</sup> Notably, this Pgm1 homologue is positioned immediately downstream of the *bytAO* gene pair (Figure 17), suggesting its role in further modification.



**Figure 17** (A) Schematic overview of pheganomycin BGC in *Streptomyces cirratus* and *byt* BGC in *Streptomonospora alba*. Two genes of those clusters share a significant identity (Pgm1 – 36%, glycine amidinotransferase 57%). (B) Proposed biosynthesis of pheganomycin (PGM-DR).<sup>[131]</sup> (S)-2-(3,5-dihydroxy-4-hydroxymethyl)phenyl-2-guanidinoacetic acid is likely synthesised by a type III PKS synthase (PGM11) and ligated by Pgm1 to the peptide NVKDR and methylated to yield the final PGM-DR.

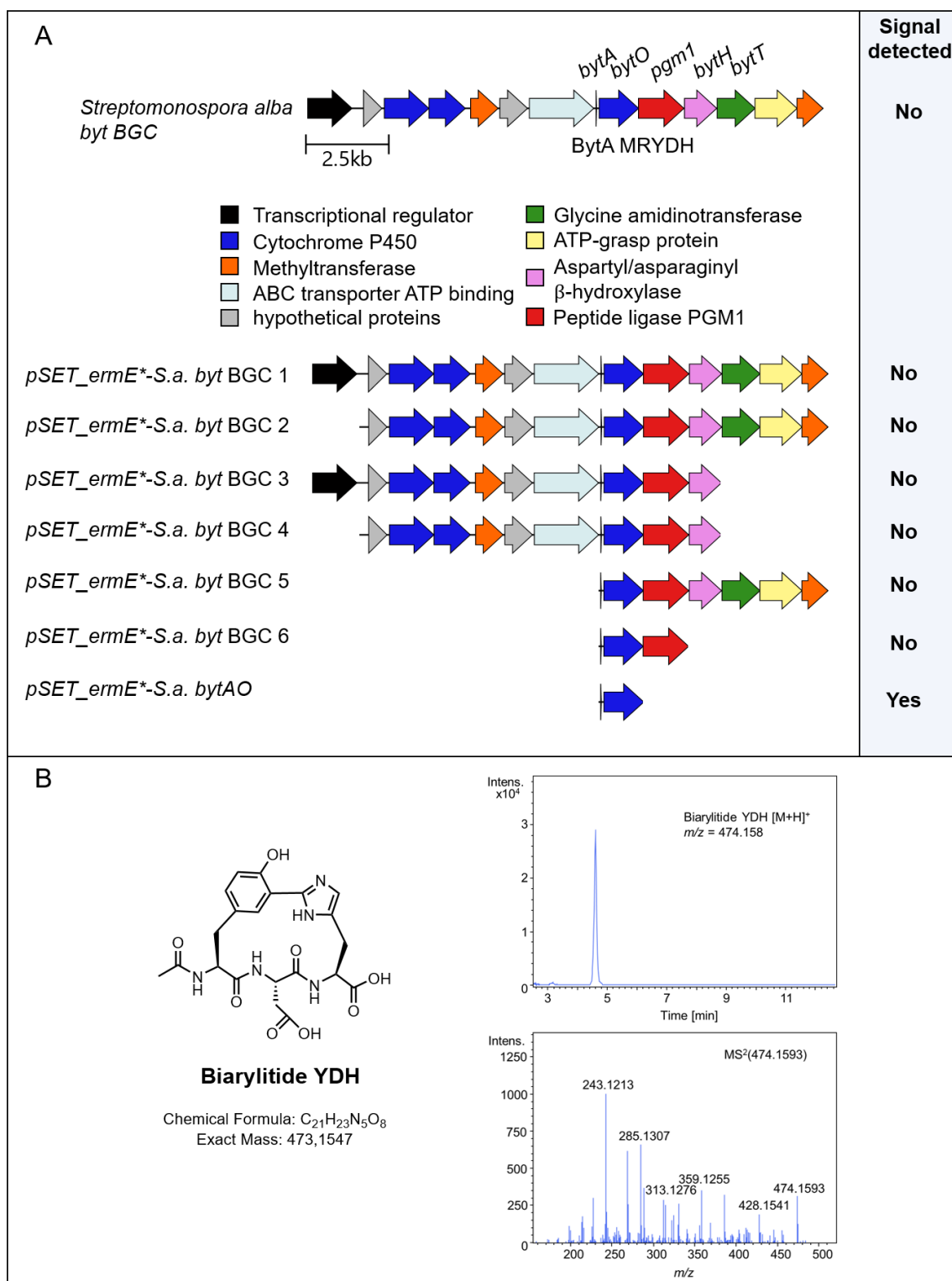
The cluster's architecture reveals parallels with the pheganomycin pathway, including the presence of a glycine amidinotransferase that maintains 57% sequence identity with its pheganomycin counterpart. The widespread occurrence of this architectural pattern, found in 46 additional biarylite BGCs across bacterial genomes (Figure 9), suggests these compounds

play significant biological roles and that the integration of pheganomycin-like biosynthetic machinery with biarylittide pathways represents an evolutionarily conserved strategy for secondary metabolite production in diverse bacterial species. These shared genetic elements between the two pathways strongly suggest that the *byt* BGC may encode a hybrid biosynthetic pathway capable of producing novel compounds that incorporate features of both biarylittides and pheganomycin-like molecules. The presence of the Pgm1 homologue in the *byt* gene cluster is particularly significant, as it could potentially catalyse the ligation of a biarylittide core to additional chemical moieties, similar to its role in pheganomycin biosynthesis.

To systematically investigate the gene cluster's biosynthetic potential, seven distinct genetic constructs with varying gene compositions were generated (Figure 18). The approximately 13 kb gene cluster was successfully cloned through a systematic four-step process. Comprehensive LC-MS analysis was conducted to detect potential products from the heterologous expression of these constructs in *S. coelicolor*. The analysis focused on identifying masses corresponding to predicted biarylittides and potential hybrid products, complemented by molecular networking to detect novel metabolites.<sup>[54]</sup> A control *S. coelicolor* strain containing an empty pSET-*ermE*\* plasmid was used for comparative analysis.

Interestingly, among all constructs, only the minimal *bytAO* gene pair successfully produced a detectable product, identified as biarylittide YDH (Figure 18). No novel masses were detected in constructs containing the Pgm1 homologue or other genes from the extended cluster. This observation suggests that when Pgm1 is present, the biosynthetic pathway may be redirected toward different, currently undetected products, or that the expression of the complete pathway requires specific conditions or precursors absent in the heterologous host.

The widespread distribution of this hybrid cluster architecture, combining elements of both biarylittide and pheganomycin biosynthesis, suggests a significant evolutionary relationship between these pathways. While the heterologous expression efforts revealed only the core biarylittide product from the minimal construct, the conservation of pheganomycin-like biosynthetic machinery indicates additional biosynthetic capabilities that may require specific conditions or regulatory factors for expression. Future investigation of this cluster type would benefit from focusing on a homologous cluster from an alternative source, as *S. alba*'s extreme growth requirements (30-day cultivation period and 10% salt concentration) present significant experimental challenges for native producer studies.



**Figure 18** (A) Schematic representation of *byt* BGC in *Streptomonospora alba* and the genetic constructs generated based upon it. (B) Putative molecular structure of Biaryllyte YDH as well as LCMS trace and MS<sup>2</sup> fragmentation spectrum.

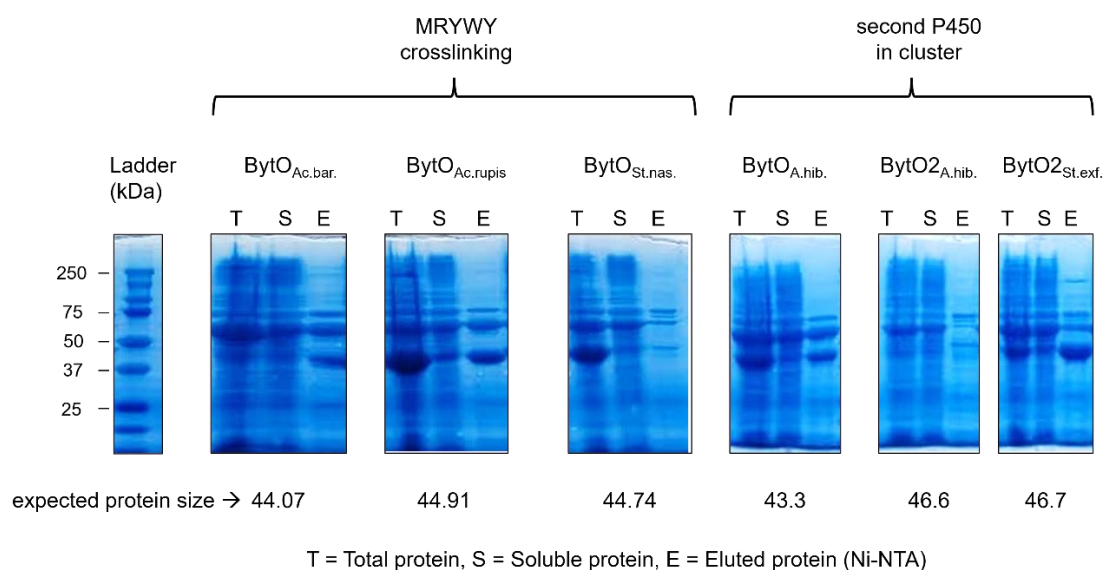
### 3.3.7 *In vitro* characterisation of BytO homologues

To expand our understanding of the functional diversity within the BytO enzyme family, several homologues were selected for biochemical characterisation (Table 3). These candidates were chosen to represent distinct features of interest: RufO from the rufomycin biosynthetic pathway (chapter 3.2), members of a newly identified subgroup containing the MRYWY sequence motif, and representatives of the second P450 enzyme that appears in some biarylite gene clusters.

**Table 3** Information about BytO homologues investigated for *in vitro* study.

Strain	BytA sequence	Characteristic	Protein expressed
<i>Actinomadura barringtoniae</i>	MRYWY	New subgroup member	Yes
<i>Actinocatenispora rupis</i>	MRYWY	New subgroup member	Yes
<i>Stackebrandtia nassauensis</i>	MRYWY	New subgroup member	No
<i>Actinomadura hibisca</i> CYP1	MKYWH	Widespread motif	Yes
<i>Actinomadura hibisca</i> CYP2	-	Second encoded P450	No
<i>Streptomyces exfoliatus</i> CYP2	-	Second encoded P450	Yes

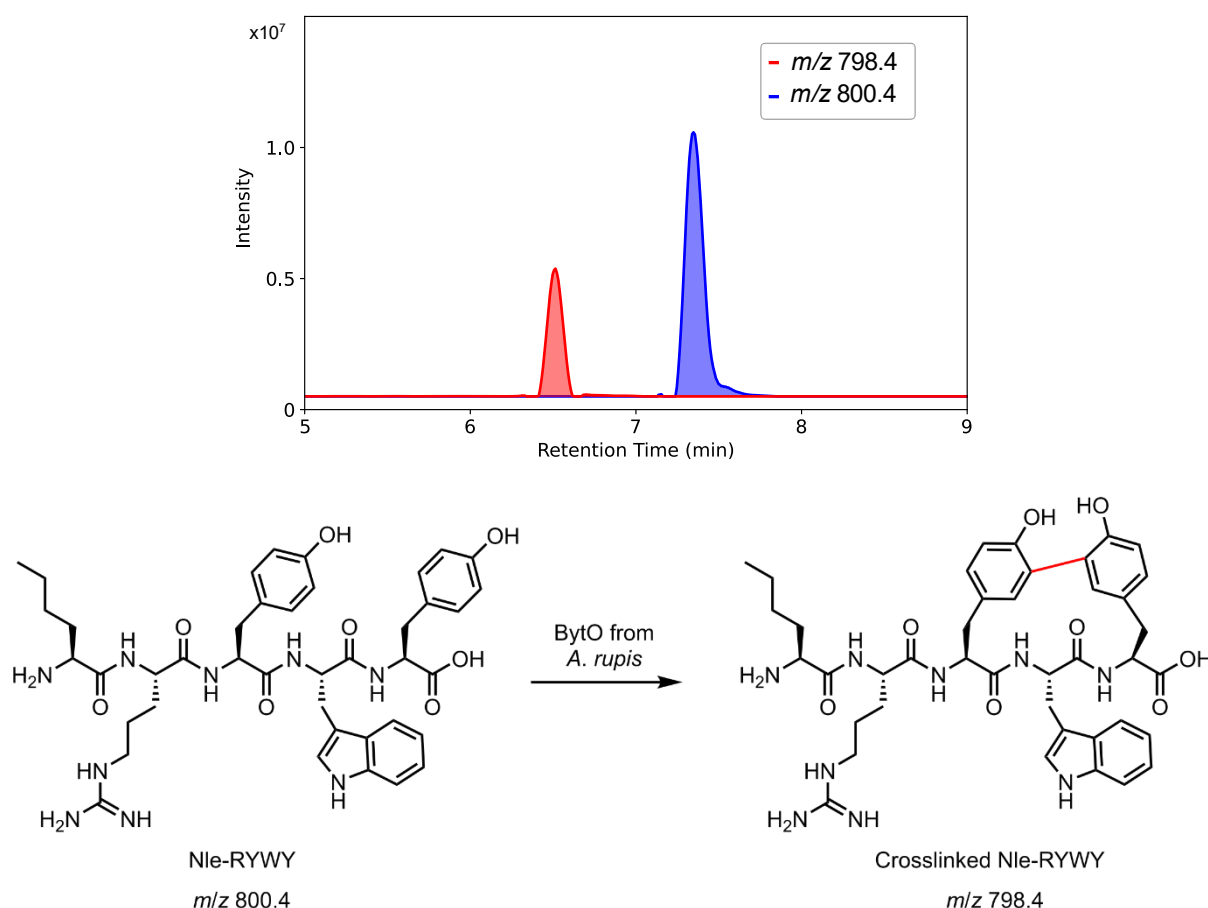
The selected BytO proteins were heterologously expressed in *E. coli* and evaluated for solubility (Figure 19). Analysis of the six BytO variants showed that four proteins (BytO<sub>Ac.bar.</sub>, BytO<sub>Ac.rupis</sub>, BytO<sub>A.hib.</sub>, and BytO<sub>2A.hib.</sub>) were successfully produced as soluble proteins, with two representatives each from the MRYWY crosslinking and second P450 cluster groups, enabling their further analysis.



**Figure 19** SDS pages of six test expressions of BytO homologues. A band at the expected size in the elution fraction indicates successful protein expression.

### 3.3.8 BytO turnover studies

After expression and purification of the respective BytO variants (methodology described in chapters 5.5.1, 5.5.2 and 5.5.3), turnover assays were performed with chemically synthesised pentapeptides matching their natural substrate structure (methodology described in chapters 5.5.5 and 5.6). The BytO homologue from *Actinocatenispora rupis* BytO<sub>Ac.rupis</sub> represented a particularly intriguing case as a member of a previously missed subgroup of biaryllytite P450s, encoding the unusual MRYWY peptide sequence (chapter 3.3.1). The presence of tyrosine at position 5, rather than the more common histidine, suggested potential novel crosslinking chemistry. Using a synthetic Nle-RYWY substrate (where methionine was replaced by norleucine to prevent sulfoxidation),<sup>[96]</sup> LC-MS analysis revealed formation of a product with a -2 Da mass shift ( $m/z = 798.4$ ) relative to the substrate ( $m/z = 800.4$ ), consistent with the expected oxidative cyclisation reaction (Figure 20).

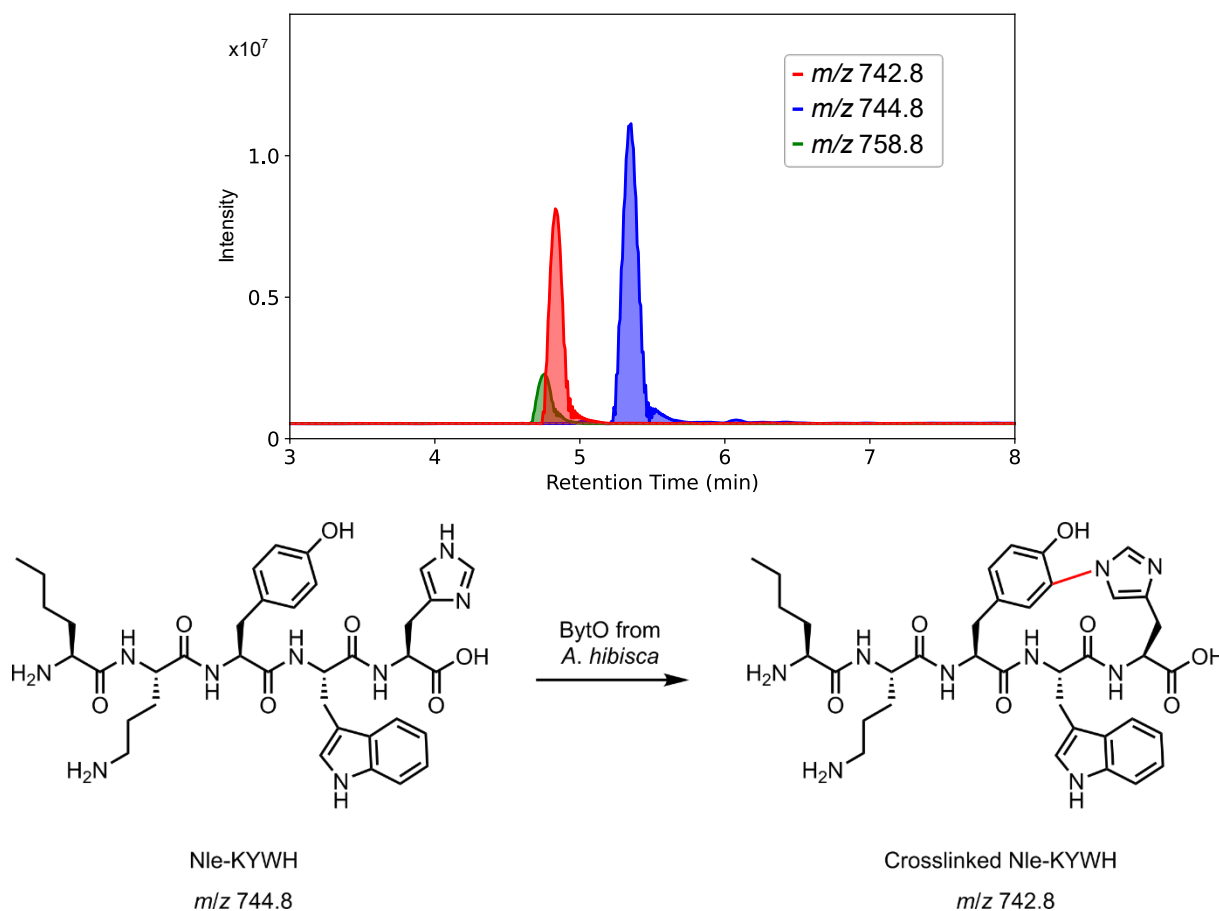


**Figure 20** *In vitro* turnover assay incubating Nle-RYWY (blue trace,  $m/z$  = 800.4) and BytO from *A. rupis* resulting in -2 Da product corresponding to a crosslinked Nle-RYWY mass (red trace,  $m/z$  = 798.4) (top). The catalysed crosslink is likely C-C, but warrants further experiments to validate (bottom).

While this initial result confirmed the enzyme's catalytic competence, the exact nature of the installed crosslink (whether C-O or C-C) remains to be determined. Several approaches could help to resolve this question: synthesis of authentic standards containing either linkage type for comparative analysis or isolation of sufficient material for comprehensive NMR analysis. The simplest strategy would be to repeat the assay with deuterium-labelled peptides to track bond formation via LC-MS.<sup>[33]</sup>

A second major focus was the characterisation of BytO homologues associated with the MKYWH motif, one of the most widely distributed sequence patterns across bacterial genomes (Figure 9). This motif's prevalence, combined with the observation that P450<sub>Blt</sub> did not tolerate large residues like tryptophan at position 4 (chapter 3.1.1), made structural characterisation and testing for crosslinking activity particularly intriguing. *In vitro* turnover studies with purified BytO from *Actinomadura hibisca* BytO<sub>A.hibisca</sub> yielded unexpected complexity: alongside the anticipated -2 Da crosslinked product ( $m/z$  = 742.8), a second species with +14 Da mass shift

( $m/z = 758.8$ ) was observed, suggesting incorporation of an oxygen in addition to the crosslink (Figure 21).



**Figure 21** *In vitro* turnover assay incubating Nle-KYWH (blue trace,  $m/z = 744.8$ ) and BytO from *Actinomadura hibisca* resulting in a -2 Da product corresponding to a crosslinked Nle-KYWH mass (red trace,  $m/z = 742.8$ ), and an unexpected second +14 Da product (green trace,  $m/z = 758.8$ ).

To enable structural characterisation, the reaction was scaled up 100-fold while maintaining relative component concentrations. Products were purified by reversed-phase HPLC using a semi-preparative C18 column with an optimised water/acetonitrile gradient. While this approach successfully yielded sufficient quantities of the primary crosslinked product for NMR analysis and structural elucidation (chapter 3.3.9), the +14 Da species proved unstable under preparative conditions.

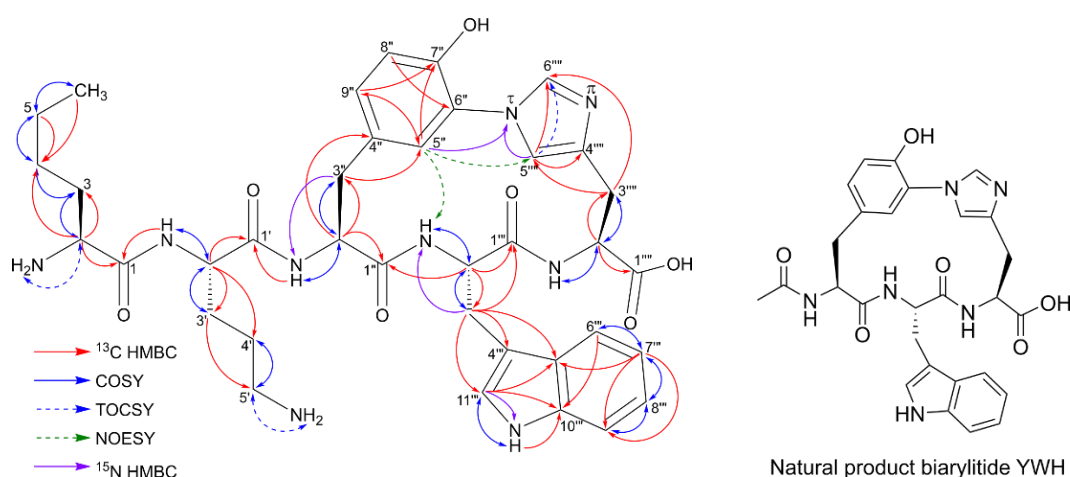
These experiments demonstrated successful reconstitution of activity with two phylogenetically distinct BytO homologues: one processing the newly identified MRYWY motif from *A. rupis*, and another acting on the widely distributed MKYWH motif from *A. hibisca*. The successful isolation and structural characterisation of the Nle-KYWH crosslinked product marks the first elucidation of a biaryllytite structure from the prevalent MKYWH family, providing crucial

insights into this widespread peptide modification system. The structural details of this novel crosslinked product are presented in the following chapter.

### 3.3.9 Structural elucidation of Nle-KYWH as a widespread biaryllytide

The measurement, analysis and Figure preparation was performed by Laura Coe and James de Voss (University of Queensland). The NMR analysis of the *in vitro* crosslinked Nle-KYWH peptide by BytO from *Actinomadura hibisca* was conducted using high-field NMR spectroscopy in DMF- $d_7$  (Figure 22 and Table 4). To aid in structural assignment, 1-phenylimidazole was analysed analogously as a model compound for the crosslinked Tyr-His moiety (Table 5 and Figure 23). Initial examination of the  $^1\text{H}$  NMR spectrum of crosslinked Nle-KYWH revealed five  $\alpha$ -protons at  $\delta_{\text{H}}$  4.18, 4.48, 4.77, 5.12, and 4.73, with their corresponding alpha carbons detected by  $^{13}\text{C}$ -HMBC at  $\delta_{\text{C}}$  54.3, 54.2, 54.8, 54.9, and 53.2 respectively. The expected five amide protons were identified at  $\delta_{\text{H}}$  8.65, 8.75, 7.65, 8.78, and 8.85, and assigned through COSY and TOCSY correlations. Five carbonyl signals were observed at  $\delta_{\text{C}}$  170.3, 171.6, 170.9, 173.1, and 173.9. Examination of the aromatic region revealed signals characteristic of a C-N biaryl crosslink such as the shift of Tyr-C6'' at  $\delta_{\text{C}}$  125.4. In contrast to this, there is a well described shift within the C-C crosslinked peptides such as biaryllytide YYH ( $\delta_{\text{C}}$  110) and aciculitin B ( $\delta_{\text{C}}$  114).<sup>[52][135]</sup>

Based on these results, BytO from *Actinomadura hibisca* installs a C-N crosslink, matching the bond facilitated by P450<sub>Blt</sub>.<sup>[96]</sup> Thus, the wide spread natural product biaryllytide YWH is expected to have the structure shown in Figure 22 (right).



**Figure 22** NMR analysis of crosslinked Nle-KYWH peptide produced by BytO from *Actinomadura hibisca* (left) and the derived structure of the natural product biaryllytide YWH (right). The coloured arrows correspond to the specific NMR experiments.



**Table 4** Complete NMR assignments for *in vitro* crosslinked Nle-KYWH peptide corresponding to Figure 22.

	Obs. DMF-d <sub>7</sub> 700 MHz $\delta_H(J)$	176 MHz $\delta_C$	70 MHz $\delta_N$
<b>Norleucine</b>			
1		170.3	
2	4.18 (t, 6.5, br)	54.3	
2-NH	8.65 (s, br)		Not Obs.
3	1.92 (dq, 14.0, 7.0, ov) <sup>1</sup> 1.89 (dq, 14.0, 7.0, ov) <sup>1</sup>	32.3	
4	1.44 (pent, 7.4, ov)	27.6	
5	1.28 (sext, 7.4)	23.1	
6	0.84 (t, 7.4)	14.4	
<b>Lysine</b>			
1'	-	171.6	
2'	4.48 (td, 8.8, 4.4)	54.2	
2'-NH	8.75 (d, 8.8)		120.1
3'	1.61-1.67 (m), 1.76-1.82 (m, ov)	32.6	
4'	1.68-1.78 (m, ov)	27.8	
5'	2.98-3.30 (m)	40.5	
5-NH	8.32 (s, br)		Not Obs.
<b>Tyrosine</b>			
1''	-	170.9	
2''	4.77 (dt, 7.5, 4.4)	54.8	
2''-NH	7.65 (d, 7.3)		114.0
3''	3.03 -3.08 (m, ov)	37.2	
4''	-	129.0	
5''	6.97 (d, 2.2)	129.7	
6''	-	125.4	
7''	-	150.0	
7''-OH	10.48 (s) <sup>2</sup>		
8''	6.99 (d, ov, 8.2)	117.3	
9''	6.93 (dd, 8.2, 2.2)	130.6	
<b>Tryptophan</b>			
1'''	-	173.1	
2'''	5.12 (td, 8.7, 4.9)	54.9	
2'''-NH	8.78 (d, 8.7)		121.1
3'''	3.22 (dd, 14.7, 4.9), 3.13 (14.7, 8.7)	31.0	
4'''	-	111.14	
5'''	-	128.8	
6'''	7.71 (d, 7.5)	119.5	
7'''	6.98 (t, 7.5)	119.4	
8'''	7.08 (t, 7.5)	121.9	
9'''	7.38 (d, 7.5)	112.4	
10'''	-	137.7	
11'''	7.27 (d, 2.0)	125.0	
11'''-NH	10.89 (d, 2.0)		129.9

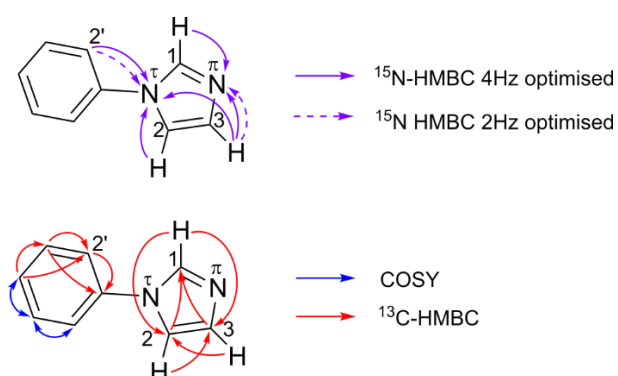
<b>Histidine</b>			
1''''	-	173.9	
2''''	4.73 (ddd, 12.2, 9.5, 2.6)	53.2	
2''''-NH	8.85 (d, 9.5)		119.1
3''''	3.26 (dd, 14.7, 2.6), 2.87 (dd, 4.7, 12.2)	32.3	
4''''	-	137.9	
5''''	7.09 (s, br, ov)	120.2	
5''''-N			177.0
6''''	7.91 (s, br)	138.2	

<sup>1</sup> Two overlapped signals. ABX2Y system appearing as a nine-peak signal.

<sup>2</sup> No correlations observed. Exchangeable proton. Assignment by chemical shift.

**Table 5** NMR assignments for 1-phenylimidazole, corresponding to Figure 23.

	Obs. DMF-d <sub>7</sub> 700 MHz $\delta_H$ (J)	176 MHz $\delta_C$	70 MHz $\delta_N$
<b>1</b>	8.31 (t, 1.2)	136.7	
<b>1-N (<math>\tau</math>)</b>			
<b>2</b>	7.79 (t, 1.2)	119.1	
<b>3</b>	7.16 – 1.17 (m)	131.0	
<b>3-N (<math>\pi</math>)</b>			265.6
<b>1'</b>	-	138.6	
<b>2'</b>	7.72 - 7.74 (m)	121.6	
<b>3'</b>	7.54 - 7.57 (m)	131.2	
<b>4'</b>	7.40 (tt, 7.5, 1.1)	128.0	
<b>5'</b>	7.54 - 7.57 (m)	131.2	
<b>6'</b>	7.72 - 7.74 (m)	121.6	



**Figure 23** Structure of 1-phenylimidazole and NMR analysis with coloured arrows corresponding to specific NMR experiments.

### 3.3.10 A second P450 in *byt* gene clusters with MKYWH motif

A notable feature among *byt* gene clusters, such as those found in *A. hibisca* and *S. exfoliatus*, is the presence of a second P450 gene (Figure 10). This additional P450 was hypothesised to perform a reaction similar to LauB2, a P450 involved in atropopeptide biosynthesis that forms an aryl ether crosslink between adjacent Trp and Tyr residues after LauB1 establishes the three canonical crosslinks characteristic of atropopeptides.<sup>[89]</sup> To test this hypothesis, turnover assays were conducted using purified P450s from both *A. hibisca* and *S. exfoliatus* (see 5.5.5). The enzymes were tested both independently and in combination with their respective cluster-associated BytO homologues. While the BytO homologues successfully catalysed the formation of the expected crosslinked products from the Nle-KYWH substrate, the second P450 showed no catalytic activity in either configuration. The absence of additional modifications suggests that this second P450 either requires specific substrate features not present in this test system or depends on additional biosynthetic partners for activity.

## 4 Summary and Conclusion

### 4.1 Structural and functional characterisation of P450<sub>Blt</sub>

Biochemical characterisation of P450<sub>Blt</sub> revealed its ability to process several pentapeptide substrates (chapter 3.1.1). Through systematic binding affinity analyses and turnover assays, the enzyme demonstrated a remarkably high affinity for its native substrate MRYLH ( $K_d$  of 2.1  $\mu$ M) and achieved >80% conversion rates. Substrate modification studies, including position-specific amino acid substitutions and leader sequence changes, revealed critical structural requirements for enzyme activity. In particular, the enzyme showed an unexpected acceptance of tryptophan as a substitute for the native histidine residue at the cross-linking position, consistent with the subsequent discovery of naturally occurring peptide motifs containing tryptophan at the cross-linking position (chapter 3.3.2). The nature of this unexpected crosslink was elucidated by comparative NMR analysis and hydrogen-deuterium exchange studies, revealing an A-O-B bond between the tyrosine phenol oxygen and the tryptophan indole ring. This finding, combined with the observed crosslinking diversity within the biarylptide family, [29,66,70,74,136] highlights the potential of these P450s as enzymatic tools for cyclic tripeptide synthesis. The practical utility of this system was enhanced by the development of a simple isolation protocol using trypsin digestion to release the cyclic tripeptide products from their cross-linked precursors.

In a subsequent study, critical revision of the reaction mechanism revealed that P450<sub>Blt</sub> forms a C-N bond between the indole nitrogen of tryptophan and the metacarbon of tyrosine, rather than the C-O bond originally proposed (chapter 3.1.2). High-field NMR spectroscopy in DMSO- $d_6$  proved crucial, providing simplified spectra and preserving essential exchangeable proton signals that were key to the structural determination. The revised structure was further validated by computational NMR predictions and isotope labelling experiments using deuterated tyrosine. Complementary *in vivo* studies with the related enzyme BytO from *Planomonospora*, which naturally catalyses C-C bonds between tyrosine and histidine,<sup>[52]</sup> involving site-directed mutagenesis and heterologous expression in *Streptomyces coelicolor* M1152, yielded identical products as confirmed by chromatographic analysis and mass spectrometric fragmentation patterns, definitively establishing the C-N cross-linking mechanism.

The first crystal structure of a peptide-modifying substrate-bound P450 (P450<sub>blt</sub> with MRYLH) revealed specific molecular interactions that control the catalytic process (chapter 3.1.3). High-resolution X-ray crystallography of three forms – apo (2.15 Å), substrate-bound (1.79 Å), and

E238A mutant (1.95 Å) – showed that I-helix residues, including a key acid/alcohol pair, interact with the C-terminus of the substrate to orient the tyrosine and histidine side chains for cross-linking. The enzyme's specificity for C-N cross-link formation is critically dependent on His-234, as demonstrated by mutational studies in which H234L showed dramatically reduced activity (>9% vs. 85% wild-type), while the conserved Asp-238/Ser-239 pair suggests standard P450 oxygen activation <sup>[137]</sup> rather than alternative pathways such as superoxide-mediated catalysis.<sup>[138]</sup> Molecular dynamics simulations over 200 ns and density functional theory calculations revealed that these residues in particular have a dual function, with Ser-239 coordinating the binding of the substrate, while Asp-238 organises a network of water molecules for the delivery of protons to the active site. These structural insights were further validated by extensive mutational studies of BytO, which naturally forms C-C cross-links. After cloning and heterologous expression in *Streptomyces coelicolor*, followed by extraction and LC-MS analysis of the products, the introduction of a histidine by the V219H mutation nearly abolished activity (0.6% product formation), while the Q216A mutation reduced activity to 7.6%, demonstrating how subtle active site differences control reaction specificity in these enzymes.

## 4.2 RiPP-NRPS integration: a novel paradigm in natural product biosynthesis

Manual genome analysis of the rufomycin BGC revealed an unexpected biarylptide pathway that bridges ribosomal and non-ribosomal peptide synthesis (chapter 3.2). This discovery, made through careful manual analysis rather than traditional genome mining approaches, revealed a P450 enzyme with 97.97% identity to RufO encoded adjacent to *rufN* in *Streptomyces atratus* S3\_m208\_1, leading to the identification of a complete RiPP pathway nested within the rufomycin NRPS BGC. This discovery not only resolved the long-standing question of 3-nitrotyrosine biosynthesis in rufomycin, a potent antitubercular cyclic heptapeptide <sup>[82,84,86,87,139]</sup> but also established a novel paradigm for the integration of ribosomal and non-ribosomal peptide synthesis.

The characterisation of RufO revealed a remarkable evolutionary repurposing of P450 enzymes through minimal active site modifications. *In vitro* analysis using DEANO as a source of nitric oxide revealed that despite high structural similarity to P450<sub>Blt</sub>, including nearly identical I-helix architecture except for two key residues (Ser239/Val240 in P450<sub>Blt</sub> versus Val240/Pro241 in RufO), RufO catalyses selective nitration rather than cross-linking. This evolutionary flexibility mirrors the rapid functional diversification seen in radical SAM-dependent

enzymes.<sup>[140]</sup> Notably, substitution of these residues in RufO (RufO<sub>SV</sub>) resulted in both crosslinking and nitrating activity in a 5:1 ratio, albeit with reduced overall conversion (20%), demonstrating how subtle active site variations can dramatically alter reaction outcomes.

*In vivo* experiments demonstrated the essential nature of this embedded RiPP pathway. Using homologous recombination and CRISPR-Cas9-cBEST approaches,<sup>[117]</sup> deletion of either *bytA* or *rufO* abolished rufomycin production, whereas supplementation with either 3-nitrotyrosine or the nitrated pentapeptide restored production. A complete mechanistic understanding of this neofunctionalised RiPP pathway emerged with the identification of RufB as the peptidase that releases 3-nitrotyrosine from the modified peptide, supported by computational modelling and characterisation of the distinctive nitro group-binding pocket of the RufT A<sub>3</sub> domain.

The widespread distribution of this biosynthetic logic, as evidenced by genome mining that revealed 19 additional gene clusters containing *rufNO* homologues and conserved MRYLH-encoding genes, suggests that this is a more general strategy in nature. These gene clusters display diverse genetic architectures and peptide processing strategies, revealing multiple evolutionary solutions for incorporating RiPP-derived building blocks into complex natural products. This discovery fundamentally changes our understanding of the relationship between ribosomal and nonribosomal peptide biosynthesis, two pathways traditionally considered to be distinct systems, and suggests that similar pathway integration events may be more common than previously recognised.

While previous examples of RiPP pathway expansion include the addition of polycyclic terpene moieties<sup>[109]</sup> and PEARL enzyme systems,<sup>[141–143]</sup> rufomycin biosynthesis represents the first complete integration of a RiPP pathway into NRPS assembly. This hybrid pathway establishes a new paradigm for non-proteinogenic amino acid biosynthesis in NRPS pathways, which differs from traditional approaches involving free amino acid synthesis or carrier protein-based modifications.<sup>[144,145]</sup>

### **4.3 Expanding biarylptide diversity through machine learning based genome mining and biochemical validation**

The systematic study of biarylptide biosynthesis using a combined computational and experimental approach revealed significant insights into both precursor sequences and enzyme functions (chapter 3.3). An adapted machine learning pipeline based on AtropoFinder<sup>[89]</sup> enabled the identification of 277 unique biarylptide biosynthetic gene clusters, including 124 novel members. This analysis extended the known distribution of biarylptide producers to

include human microbiome-associated bacteria, specifically identifying potential production capability in *Rothia* species. The computational analysis identified several novel precursor peptide motifs, particularly variants containing tryptophan at cross-linking positions. An interesting finding was the discovery of multiple precursor copy variants in *Lysinibacillus* strains, some containing up to six copies of the precursor gene. The analysis also revealed systematic patterns in the distribution of tailoring enzymes, with 47 gene clusters encoding both a Pgm1 ligase homologue and precursor motifs with aspartic acid at position 4. This widespread occurrence of hybrid clusters combining biarylptide and pheganomycin biosynthetic machineries suggests a significant evolutionary relationship between these pathways and has broader implications for understanding the evolution of peptide natural product biosynthesis.

Experimental validation focused on ten representative strains from different bacterial genera. Heterologous expression studies in *Streptomyces* hosts yielded four novel compounds: Biarylptide YVH and its hydroxylated variant from *Saccharothrix variisporea*, Biarylptide YLH from *Streptomyces gougerotii*, and Biarylptide YDH from *Streptomonospora alba*. The *S. gougerotii* cluster revealed a functional BytO enzyme despite the absence of the two residues essential for oxygen activation,<sup>[134]</sup> demonstrating unexpected flexibility in P450 function.

The biochemical characterisation of the BytO homologues yielded two important results: First, scaled-up *in vitro* assays with BytO from *Actinomadura hibisca* processing the widespread MKYWH motif enabled NMR analysis that confirmed the formation of C-N cross-links, providing structural validation of this widespread biarylptide variant. Second, the BytO from *Actinocatenispora rupis* showed enzymatic cross-linking activity with its natural substrate, the atypical MRYWY sequence, establishing the catalytic competence of a previously unrecognised biarylptide subgroup.

The systematic relationships observed between precursor sequences and tailoring enzymes, combined with the discovery of multiple pathway integration events, suggest that similar biosynthetic innovations may be more abundant in nature than previously thought.

## 5 Materials and Methods

### 5.1 Materials

#### 5.1.1 Chemicals and consumables

Unless specified otherwise, all standard reagents and consumables were sourced from the following suppliers: Carl Roth (Karlsruhe, Germany), Eppendorf (Hamburg, Germany), Merck (Darmstadt, Germany), Sarstedt (Nümbrecht, Germany), Sigma-Aldrich (St. Louis, USA), Thermo Fisher Scientific (Waltham, USA), and VWR (Darmstadt, Germany). The used enzymes were supplied by New England Biolabs (Frankfurt am Main, Germany), Promega (Madison, USA) and Takara Bio (Kusatsu, Japan).

#### 5.1.2 Bacterial strains

*Escherichia coli* strains  $\alpha$ -select silver and NEB® 10-  $\beta$  were used for cloning and maintenance of plasmids. *E. coli* BL21(DE3), Lemo21(DE3), LOBSTR and Rosetta were used for protein expression and ET12567 for intergeneric conjugation with *Streptomyces*. *Streptomyces coelicolor* M1152, *Streptomyces lividans* TK 24 and *Streptomyces albus* J1074 were used for heterologous expression of biosynthetic gene clusters. The remaining bacteria were used for analysis of their respective Biarylittide gene clusters.

**Table 6** *E. coli* and *S. coelicolor* strains used for cloning and expression.

Strain	Genotype	Origin
<i>E. coli</i> $\alpha$ -select silver	F- <i>deoR endA1 recA1 relA1 gyrA96</i> <i>hsdR17</i> (rk <sup>-</sup> , mk <sup>+</sup> ) <i>supE44 thi-1 phoA</i> $\Delta$ ( <i>lacZYA argF</i> )U169 80 <i>lacZ</i> $\Delta$ M15 $\lambda$ <sup>-</sup>	Bioline
<i>E. coli</i> NEB® 10- $\beta$	$\Delta$ ( <i>ara-leu</i> ) 7697 <i>araD139 fhuA</i> $\Delta$ <i>lacX74</i> <i>galK16 galE15 e14- <math>\phi</math>80dlacZ</i> $\Delta$ M15 <i>recA1</i> <i>relA1 endA1 nupG rpsL</i> (StrR) <i>rph spoT1</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> )	New England Biolabs
<i>E. coli</i> BL21(DE3)	<i>fhuA2</i> [lon] <i>ompT gal</i> ( $\lambda$ DE3) [ <i>dcm</i> ] $\Delta$ <i>hsdS</i> $\lambda$ DE3 = $\lambda$ <i>sBamHI</i> o $\Delta$ EcoRI-B <i>int::</i> ( <i>lacI::PlacUV5::T7 gene1</i> ) <i>i21 <math>\Delta</math>in5</i>	New England Biolabs
<i>E. coli</i> LOBSTR	Derived from <i>E. coli</i> BL21(DE3). $\Delta$ <i>arnA</i> $\Delta$ <i>slyD</i>	Andersen <i>et al.</i> , 2013 <sup>[146]</sup>



<i>E. coli</i> Lemo21(DE3)	<i>fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/</i> <i>pLemo(CamR) λ DE3 = λ sBamHI ΔEcoRI-</i> <i>B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5</i> <i>pLemo = pACYC184-PrhaBAD-lysY</i>	New England Biolabs
<i>E. coli</i> Rosetta	F- <i>ompT hsdSB(rB- mB-) gal dcm</i> (DE3) pRARE	New England Biolabs
<i>E. coli</i> ET12567	F- <i>dam-13::Tn9 dcm-6 hsdM hsdR</i> , Cmr	MacNeil <i>et al.</i> , 1992 <sup>[147]</sup>
<i>E. coli</i> ArcticExpress (DE3)	F- <i>ompT hsdS(rB- mB-) dcm+ Tetr gal</i> <i>λ(DE3) endA Hte [cpn10 cpn60 Gentr]</i>	Agilent
<i>S. coelicolor</i> M1152	<i>Δact Δred Δcpk Δcda rpoB[C1298T]</i> <i>rpsL[A262G C271T]</i>	Till F. Schäberle. Justus Liebig University Gießen
<i>S. lividans</i> TK24	-	as above
<i>S. albus</i> J1074	-	as above

**Table 7** Actinomycete bacterial strains examined for their ability to produce biarylides and used to isolate genomic DNA.

Strain	Genotype	Origin
<i>Actinomadura hibisca</i> 83825	Wildtype	Naicons
<i>Actinomadura hibisca</i> 83826	Wildtype	Naicons
<i>Streptomyces exfoliatus</i>	Wildtype	Naicons
<i>Saccharothrix variisporea</i> DSM 40234	Wildtype	DSMZ
<i>Actinomadura barringtoniae</i> DSM 108647	Wildtype	DSMZ
<i>Actinomadura rifamycini</i> DSM 43936	Wildtype	DSMZ
<i>Actinocatenispora rupis</i> DSM 45178	Wildtype	DSMZ
<i>Actinomadura geliboluensis</i> DSM 45508	Wildtype	DSMZ
<i>Streptomyces gougerotii</i> DSM 40234	Wildtype	DSMZ
<i>Stackebrandtia nassauensis</i> DSM 44728	Wildtype	DSMZ
<i>Streptomonospora alba</i> DSM 44588	Wildtype	DSMZ
<i>Micromonospora</i> MW-14	Wildtype	Harald Groß, Tübingen

### 5.1.3 Plasmids

Different plasmids were used for heterologous expression in *Streptomyces* (pSET152-*ermE*\*, pGM1202) and in *E. coli* (pET-28a(+)) and pCDFDuet-1) as well as for aiding in intergeneric conjugation of DNA (pUZ8002). The plasmids used in this thesis are presented in Table 8.

**Table 8** Plasmids used in this study.

Plasmid	Description	Origin
pSET152- <i>ermE</i> *	Apr <sup>R</sup> , <i>Streptomyces</i> integrative vector, <i>lacZα</i> , MCS, <i>rep</i> <sup>PUZ</sup> , <i>ermE</i> * promotor, <i>oriT</i>	Institute for Pharmaceutical Biology, University of Bonn
pET-28a(+)	Kan <sup>R</sup> , bacterial expression vector with T7 lac promoter, <i>N</i> -terminal His tag, MCS-1, thrombin cleavage site, internal T7 epitope tag, <i>C</i> -terminal His tag, pBR322 <i>ori</i>	Institute for Pharmaceutical Biology, University of Bonn
pCDFDuet-1	Apr <sup>R</sup> , bacterial expression vector, lac operator, RBS, <i>N</i> -terminal His tag, MCS, pCloDF12 <i>ori</i>	Institute for Pharmaceutical Biology, University of Bonn
pUZ8002	Kan <sup>R</sup> , RK2 derivative, defective <i>oriT</i> ( <i>aph</i> )	Institute for Pharmaceutical Biology, University of Bonn

### 5.1.4 Media

Bacterial cultures were grown in either liquid media or on solid agar plates, depending on their intended use. Demineralised water was used for all media preparations. After adding the necessary ingredients and adjusting the pH with NaOH or HCl, the media were sterilised by autoclaving at 121 °C for 20 minutes. When agar was included in the media, antibiotics were added, if required, after the media had cooled to 50 °C, and then the mixture was poured into petri dishes.

**Table 9** Media used in this study. Unless otherwise specified, the ingredients listed are for 1 litre of medium. For agar media, the amounts of agar added are indicated in brackets. Components marked with an asterisk (\*) were either added after autoclaving or sterile filtration.

Medium	Ingredient	Concentration
10x TB-salts pH 7.2	KH <sub>2</sub> PO <sub>4</sub>	0.17 mM
	K <sub>2</sub> HPO <sub>4</sub>	0.72 mM

## Materials and Methods

2x YT	Yeast extract	1 g/L
	Tryptone	1.6 g/L
	NaCl	0.5 g/L
AF	Dextrose	20 g/L
	Soybean meal	8 g/L
	Yeast extract	2 g/L
	CaCO <sub>3</sub>	4 g/L
	NaCl	1 g/L
ISP-2	Yeast extract	4 g/L
	Malt extract	10 g/L
	dextrose	4 g/L
ISP-4	Soluble starch	10 g
	MgSO <sub>4</sub> x 7H <sub>2</sub> O	1 g
	NaCl	1 g
	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	2 g
	CaCO <sub>3</sub>	2 g
	Trace salts solution*	1 mL
	(Agar)	20 g/L
Luria-Broth (LB) pH 7.5	Tryptone	10 g/L
	Yeast extract	5 g/L
	NaCl	10 g/L
	(Agar)	15 g/L
Mannitol soy flour (MS)	Mannitol	20 g/L
	Soy flour	20 g/L
	(Agar)	20 g/L
SOC pH 7.5	Tryptone	10 g/L
	Yeast extract	5 g/L
	NaCl	0.5 g/L
	1M KCl	2.5 mL
	1M Glucose*	20 mL
Terrific Broth (TB)	Tryptone	12 g/L
	Yeast extract	24 g/L
	Glycerol	5 g/L
	10x TB-salts*	100 mL
Trace salts solution	FeSO <sub>4</sub> x 7H <sub>2</sub> O	0.1 g
	MnCl <sub>2</sub> x 4H <sub>2</sub> O	0.1 g
	ZnSO <sub>4</sub> x 7H <sub>2</sub> O	0.1 g
	100 mL H <sub>2</sub> O	100 mL
Tryptic Soy Broth	Tryptone	17 g/L
	Peptone	3 g/L
	Glucose	2.5 g/L
	NaCl	5 g/L

YEME	Yeast extract	3 g/L
	Malt extract	3 g/L
	Peptone	5 g/L
	Glucose	10 g/L
	Sucrose	170 g/L

### 5.1.5 Buffers

**Table 10** Buffers used in this study.

Medium	Ingredient	Concentration
50x TAE pH 8.0	Tris	2 M
	Acetic acid	1 M
	EDTA	50 mM
$\gamma^{18}\text{O}_4$ -ATP exchange assay buffer pH 7.5	Tris	20 mM
	Glycerol	5% (v/v)
Lysis Buffer pH 8.0	Tris	50 mM
	NaCl	300 mM
	Imidazole	10 mM
	DTT	1 mM
Wash buffer I pH 8.0	Tris	50 mM
	NaCl	300 mM
	Imidazole	20 mM
	DTT	1 mM
Wash buffer II pH 8.0	Tris	50 mM
	NaCl	300 mM
	Imidazole	35 mM
	DTT	1 mM
Elution buffer	Tris	50 mM
	NaCl	300 mM
	Imidazole	300 mM
	DTT	1 mM
Storage buffer	Tris	50 mM
	NaCl	150 mM
	Glycerol	5% (v/v)
	DTT	1 mM
Low salt storage buffer	Tris	50 mM
	NaCl	50 mM
	Glycerol	5% (v/v)
	DTT	1 mM
SEC Buffer pH 7.4	Tris	20 mM
	NaCl	100 mM

AEX Buffer A pH 8.0	Tris	20 mM
	NaCl	10 mM
AEX Buffer B pH 8.0	Tris	20 mM
	NaCl	1 M
5x isothermal (ISO) buffer pH 7.5	Tris	500 mM
	MgCl <sub>2</sub>	50 mM
	dNTPs	1 mM each
	DTT	50 mM
	PEG-8000	25% (v/v)
	NAD	5 mM
SDS gel separating Buffer pH 8.8	Tris	1.5 M
	SDS	0.4%
SDS gel stacking Buffer pH 6.8	Tris	0.5 M
	SDS	0.4%
10x SDS gel running buffer	Tris	250 mM
	Glycine	1.92 M
	SDS	1%

### 5.1.6 Primers

Primers were designed using SnapGene 5, with typical lengths ranging from 12 to 20 base pairs for routine applications. For Gibson assembly (chapter 5.4.7), additional 20 base pair overhangs were included. Primers were obtained from Eurofins Genomics (Ebersberg, Germany).

**Table 11** Vector primers used in this study

Template	Primer Name	Sequence (5'-3')
pSET_ermE*	pSET_linearised_fwd	CCTCTCTAGAGTCGACCTGCAGC
pSET_ermE*	pSET_linearised_rev	CCTTCCGTACCTCCGTTGCT
pSET_ermE*	pSET_fragment1_rev	GCATTTGGACGGCTGATCCGGGACG GCTGATCCG
pSET_ermE*	pSET_fragment2_fwd	GCCAATCGGATCAGCCGTCCCGGAT CAGCCGTCCAAATGC
pSET_ermE*	pSET_core_for	CATATCTCATTGCCCCCGGACGAGC GTCTG
pSET_ermE*	pSET_core_rev	CAGACGCTCGTCCGGGGGCAATGA GATATG
pSET_ermE*	pSET_RBS_opti_for	CCTCTCTAGAGTCGACCTGCAGC
pSET_ermE*	pSET_RBS_opti_rev	CCTCCTCCTTCCGTACCTCC

pET28(+)	Linearised_BytH_for	GATCTTCTCCAGACCCATTTGCTGT CCACCAG
pET28(+)	Linearised_BytH_for	GGCGGTGTCGTGACACCACCACCAC CACCAC

**Table 12** Colony PCR primer used in this study

Primer Name	Sequence (5'-3')
M13_fwd	TGTAAACGACGGCCAGT
M13_rev	CAGGAAACAGCTATGAC
ColPCR_pSET_for	GGACAGGTATCCGGTAAGCGG
ColPCR_pSET_rev	GCGGCTTGCCGAACCTTG
ColPCR_pSET-S.variisporea-bytAO_rev	AACGCGGACAGCGAAGT
ColPCR_pSET-S.gougerotii_rev	AGGTGGTGGTACATGGGGT
ColPCR_pSET-St.exfoliatus_rev	TCGGCGAACTGGCCGAA
ColPCR_pSET-A.geliboluensis_rev	AGTCCGAGGCGAACGATTCC
ColPCR_Micromonospora_contig 11_rev	CGTAGGACGGCGGTGC
ColPCR_Micromonospora_contig 44_rev	ACGGAGGGGGCCAGTG
ColPCR_pSET_44588_4frags_noReg	GGCACTCGGGAGCCG
ColPCR_pSET_44588_4frags_+reg	GGTACTCGGCGTTGATGCC
ColPCR_pSET_44588_pme1_2_for	CCGAGCTGTACGGCATCC
ColPCR_pSET_44588_pme1_3_for	GGAGCGAGCAGCGCA
ColPCR_pSET_44588_pme1_4_for	CCGTGAGGGATTACGCAC
ColPCR_pSET-S.nassauensis_rev	CCGGGGTGTTCTCGCG
ColPCR_pSET-A.rupis_rev	GGTACGCGTTCGCCGAC
ColPCR_pSET-A.rifamycini_rev	GGATGGCCGTGTGCTCG
ColPCR_pSET-A.barringtoniae_rev	CGTAGCCGGTACCGAT
ColPCR_pSET-A.hibisca_rev	CCGGTCGTACTCGGGG
ColPCR_pSET-S.nassauensis_rev	AAAGCACGGGGGCGT
ColPCR_pRSF-Planomonospora-bytAO_for	GGATCTCGACGCTCTCCCT
ColPCR_pRSF-Planomonospora-bytAO_rev	GGTTCGGCGAACTCCGATACG

**Table 13** Strain specific primers used in this study.

Strain	Primer Name	Sequence (5'-3')
<i>Planomonospora</i> ID82291	pRSF-bytAO_for	GTATAAGAAGGAGATATACAAT GCGCTACTACCACTGATCCG

<i>Streptomyces gougerotii</i>	pRSF-bytAO_rev	TTTACCAGACTCGAGGGTACCT AGCGGGGGAGAAGGACG
	pSET_S.gougerotii_RBS_opti_rev	GCAGGTCGACTCTAGAGAGGTC AGCGGGGCAGCGG
	pSET_S.gougerotii_RBS_opti_for	GGAGGTACGGAAGGAGGAGGG AACACGTGAGAGCGTCACTCG
<i>Actinomadura hibisca</i>	pSET_S.gougerotii_RBS_opti_2	GGAGGTACGGAAGGAGGAGGC GGACAATGATCAAGGGTGCG
	pSET_A.hibisca_RBS_opti_for	GGAGGTACGGAAGGAGGAGGG GCAACGTGGACAACGATG
	pSET_A.hibisca_RBS_opti_for_2	GGAGGTACGGAAGGAGGAGGG CACCCATGGGAATGCGG
	pSET_A.hibisca_RBS_opti_rev	GCAGGTCGACTCTAGAGAGGGC CCGCTCACGCCGA
	pSET_A.hibisca_bytmin_for	AGGTACGGAAGGGGTGTGCGACA TTGGCGAGGA
	pSET_A.hibisca_bytmin_rev	GA CTCTAGAGAGGCCGGTTCTC ATCTGGTGCTTAGT
	pSET_A.hibisca_for	AGCAACGGAGGTACGGAAGGC AGGCGGTGCTGTTTCGAC
	pSET_A.hibisca_rev	GCAGGTCGACTCTAGAGAGGAT CTGGCCCGCTCACG
<i>Actinomadura barringtoniae</i>	pSET_A.barringtoniae_for	GGGCTGCAGGTTCGACTCTAGAG AGGCCAGGACCCCTGATCAGAG
	pSET_A.barringtoniae_rev	AGCGGAGCAACGGAGGTACGG AAGGCGAGAGTGAGGTACGTCA CATGAG
<i>Actinomadura rifamycini</i>	pSET_A.rifamycini_for	AGCGGAGCAACGGAGGTACGG AAGGCCGCTCGCCGGT
	pSET_A.rifamycini_rev	GGGCTGCAGGTTCGACTCTAGAG AGGGGCGTCGAGAAACGTGAG G
<i>Actinocatenispora rupis</i>	pSET_A.rupis_for	AGCGGAGCAACGGAGGTACGG AAGGCCTGCGGAGGAACGCATG CCC
	pSET_A.rupis_rev	GGGCTGCAGGTTCGACTCTAGAG AGGCCTGTACCCGGGGCTGAGG TCACCAGGCG
	pSET_A.rupis_version2_for	CGGAGGTACGGAAGGCCGCTCG GTGTGTTGATCAC
	pSET_A.rupis_version3_for	CGGAGGTACGGAAGGTGATCAC GTCTGGGCGATAC
	pSET_A.rupis_version2_rev	TCGACTCTAGAGAGGGCTGTTC CTCGTACCGTTGC

<i>Stackebrandtia nassauensis</i>	pSET_A.rupis_version3_rev	TCGACTCTAGAGAGGGTTCCTC GTACCGTTGCCA
	pSET_S.nassauensis.Frag1_for	AGGTACGGAAGGCCGGATACAA GGGAAGGTGAAT
	pSET_S.nassauensis.Frag1_rev	CCGCGAATCGTGAGGTTGCGCA GCT
	pSET_S.nassauensis.Frag1_version2_for	AGGTACGGAAGGCCGGATACAA GGGAAGGTGAATCC
	pSET_S.nassauensis.Frag2_for	AGCTGGCGAACCTCACGATTCG CG
<i>Streptomyces exfoliatus</i>	pSET_S.nassaeunsis.Frag2_rev	GACTCTAGAGAGGCGTGCGGGG ATGACTCAG
	pSET_St.exfoliatus_bytmin_for	GACTCTAGAGAGGCCGCGCTCT CTCAGGTTCA
	pSET_St.exfoliatus_bytmin_rev	AGGTACGGAAGGCCTTTCTGCG CAAGG
	pSET_St.exfoliatus_template_for	GGAAAAGGTGAGCGCGGTGC
	St.exfoliatus_template_rev	CGCTGCTCGCCACGTG
	St.exfoliatus.Frag_1_for	AGCAACGGAGGTACGGAAGGC CTTTCTGCGCAAGGGGAG
	St.exfoliatus.Frag_1_rev	GCCCCGCCCACCGCGGGACCG CGCTCTCTCAGGT
	St.exfoliatus.Frag_2_for	CCCGCCGGTGGGCGGGGCTGCC CGAGATCTTTCCTCTTGC
	St.exfoliatus.bytmin_version2_for	AGCAACGGAGGTACGGAAGGG GTGGGCGGGGCACC
	St.exfoliatus.bytmin_version2_rev	GCAGGTCGACTCTAGAGAGGGG CCTTTCTGCGCAAGG
	St.exfoliatus_byt_neu_for	AGCAACGGAGGTACGGAAGGA CCGCGCTCTCTCAGG
	St.exfoliatus_byt_neu_rev	GCAGGTCGACTCTAGAGAGGGA CGCCCGGTCAGCC
<i>Micromonospora</i> sp MW-13	M.sp.MW-13_contig44_long_for	AGCAACGGAGGTACGGAAGGG CTGAGTCGTTACGTCATCACC GTCGGTCGGGCCGCCACCGTCG
	M.sp.MW-13_contig44_long_ref	GCAGGTCGACTCTAGAGAGGGC TACGCACGACGCCTCCGGTCGT TGTCGGGCACGGTCAGT
	M.sp.MW-13_contig44_short_for	AGCAACGGAGGTACGGAAGGTT CACCGACGATCCGAGCG
	M.sp.MW-13_contig44_short_version2_for	AGCAACGGAGGTACGGAATTAT TGCCGGATCGCCGTCAG
	M.sp.MW-13_bytmin_for	AGCGGAGCAACGGAGGTACGG AAGGCCGAGCGGAAGGGG



<i>Saccharothrix variisporea</i>	M.sp.MW-13_bytmin_rev	GGGCTGCAGGTCGACTCTAGAG AGGGACGCCTCCGGTCGTTG
	M.sp.MW-13_bytmin_version2_rev	GCAGGTCGACTCTAGAGAGGAG CGCCCCGTTACCG
	M.sp.MW-13_contig11_for	AGCAACGGAGGTACGGAAGGA AGGCGCATGCCTACGCC
	M.sp.MW-13_contig11_rev	GCAGGTCGACTCTAGAGAGGAC GTCGCTTGACCTACCG
	pET28(+)_BytH_for	ACAGCAAATGGGTCTGGAGAAG ATCGCCATGCAAG
	pET28(+)_BytH_rev	GTGGTGGTGGTGTACGACACC GCCTCGC
	S.variisporea_for	AGCAACGGAGGTACGGAAGGTC AGCCCAGGTAGTCTCTGC
	S.variisporea_rev	GCAGGTCGACTCTAGAGAGGCA TCGCGGCCCTCCT
	S.variisporea_bytmin_for	AGCAACGGAGGTACGGAAGGG AGGTTGTCATGCGTTACGT
	S.variisporea_bytmin_rev	GCAGGTCGACTCTAGAGAGGTC AGCGGGGAAGGAGGCGCAGG
	S.variisporea_bytAOH_rev	GCTGCAGGTCGACTCTAGAGAG GTCACGACACCGCCTCGC
	S.variisporea_bytAOHAm_rev	GGTTCGCTCCCAATCCGCCGTC ACGACACCGCCT
	S.variisporea_bytAOHAm2_for	TGCGCGAGGCGGTGTCGTGACG GCGGATTGGGAGCG
	S.variisporea_bytAOHAm2_rev	GCAGGTCGACTCTAGAGAGGTC AGCCCAGGTAGTCTCTGC
	S.variisporea_wholebyt_1_rev	CGCGGCAGCGTGCTCGTGGATC GCCGCGTTCATCCGGTCG
	S.variisporea_wholebyt_2_for	CGACCGGATGAACGCGGCGATC CACGAGCACGCTGCCGCG
	S.variisporea_incl_regulators_1_for	AGCAACGGAGGTACGGAAGGG CTCGTGACCATTCAAACCGTG
	S.variisporea_incl_regulators_1_rev	ATGGTGCGCATCACATTCGCCT CCTTCCCTGGGTG
	S.variisporea_incl_regulators_2_for	GAATCACCAGGGAAGGAGGC GAATGTGATGCGC
	S.variisporea_incl_regulators_2_rev	GCAGGTCGACTCTAGAGAGGCA CTCGCGACCTGCC
	S.variisporea_bytAOH_+3genes_for	CAGAGGACTACCTGGGCTGACC TCTCTAGAGGCACCTCTACGGG CATT

	S.variisporea_bytAOH_+3genes_rev	GCCAAGCTTGGGCTGCAGGTCTG ACTCTAGACGCCCCGCGGTGTGA
<i>Streptomyces gougerotii</i>	S.gougerotii_for	AGCAACGGAGGTACGGAAGGG TGAGAGCGTCACTCGACT
	S.gougerotii_rev	GCAGGTCGACTCTAGAGAGGTC AGCGGGGCAGCGG
	S.gougerotii_bytmin_for	AGCAACGGAGGTACGGAAGGCT GGCACGAGGAGGTTACCAT
	S.gougerotii_bytmin_rev	GCAGGTCGACTCTAGAGAGGTC ATACGGGCATTCCGCTCACCGG
	S.gougerotii_bytAOM_rev	GCAGGTCGACTCTAGAGAGGTC GGAGTAGTCAGCGGGG
	S.gougerotii_bytHAOM_for	AGCAACGGAGGTACGGAAGGA ACAACCGGGAGGGGAACACGT GAG
	S.gougerotii_reannotated_for	AGCAACGGAGGTACGGAAGGC CAGCCGCCCAGCTATCAAG
<i>Actinomadura geliboluensis</i>	A.geliboluensis_for	AGCAACGGAGGTACGGAAGGA TGTCGGAGGCGCCG
	A.geliboluensis_rev	GCAGGTCGACTCTAGAGAGGCG TGGCCGCGTCGT
	A.geliboluensis_bytOA_for	AGCAACGGAGGTACGGAAGGC GTGCGAGGAGGTGATATG
	A.geliboluensis_bytOA_rev	GCAGGTCGACTCTAGAGAGGTC AGTGGAGGTAACGCATCCGTG
	A.geliboluensis_bytOAPep_rev	GCAGGTCGACTCTAGAGAGGTC ATGGTCGCTGCTCCCGGG
	A.geliboluensis_bytOAPepH_rev	GCAGGTCGACTCTAGAGAGGTC ACCTGCCTGTCTGGT
	A.geliboluensis_OAPepHM_rev	GCAGGTCGACTCTAGAGAGGTC AGGCCGTACGGCGG
<i>Streptomonospora alba</i>	A.geliboluensis_OAPepHMS_rev	GCAGGTCGACTCTAGAGAGGTC GAGTGGTCTGTGGGTCAC
	S.alba_byt_for	GCAGGTCGACTCTAGAGAGGTG GCGATGGGAGAACTCGATATGG
	S.alba_byt_rev	GCAGGTCGACTCTAGAGAGGTG GCGATGGGAGAACTCGATATGG
	S.alba_bytmin_for	AGCAACGGAGGTACGGAAGGTC CACGGAAGGGGGCA
	S.alba_bytmin_rev	GCAGGTCGACTCTAGAGAGGTC ACGGGTGCCTCCCGGAC
	S.alba_bytAOHyp_rev	GCAGGTCGACTCTAGAGAGGTC ATTCGCTCGGTGCTGGGGAC

S.alba_byt_+regulator_for	AGCAACGGAGGTACGGAAGGG AAGACAGCGCTGCCCCACACC
S.alba_byt_+regulator_rev_1	TGGACGCGGGCGGACGGATCGG AATCGGACTCACCGGCGT
S.alba_pme1_1_for	CGGAGGTACGGAAGGGGAAGA CAGCGCTGCCCCACA
S.alba__no regulator_for	AGCAACGGAGGTACGGAAGGG GCAGGCGACGATACC
S.alba__pme1_1_rev	TCGACTCTAGAGAGGGTTTAAA CCCTCGGATGCGACGGCCACAC
S.alba__pme1_2_for	GTGTGGCCGTCGCATCCGAGGG TTTGTGTGGCCGTCGCATCCGA
S.alba__pme1_2_rev	TGCAGGTCGACTCTAGAGAGGG TTTAAACCGTGCGGCTACGGGA
S.alba__pme1_3_for	AAGAAGTCCCGTAGCCGCACGG TTTCCGGTGAGTCCG
S.alba__pme1_3_rev	TGCAGGTCGACTCTAGAGAGGG TTTAAACGGCCTTTCAGTGTTTCG TTGGG
S.alba_pme1_4_for	CCCAACGAACACTGAAAGGCCG TTTTTCGGTATGCGAC
S.alba_pme1_4_rev	TGCAGGTCGACTCTAGAGAGGG TTTCAGCGGTCTTCAGCGTGC
S.alba_pme1_1_version2_noreg_for	CGGAGGTACGGAAGGCAGGCA GGTGAGTCTTGAGTAGG
S.alba_pme1_2_version2_for	CCTAACGATGGTTTGTGTGGCC GTCGCATCC
S.alba_Fragment 1.1_rev	CGCCGATGGGTCCGTTGG
S.alba_Fragment 1.2_for	CCAACGGACCCATCGGCG
S.alba_Fragment 2.1_rev	GCACGGTGGTCGCTGG
S.alba_Fragment 2.2_for	CCAGCGACCACCGTGCCGTCGC ATCCGAGGGTTTGCCTAACGAT GGTGAACGTC

---

## 5.2 Microbiological Methods

### 5.2.1 Cultivation of bacterial strains

*E. coli* bacteria for cloning were mainly cultivated in LB at 37 °C and shaking at 200 rpm, whereas *Streptomyces* were cultivated at 30 °C and 200 rpm in an optimised medium comprising equal parts TSB and YEME with high saccharose to promote dispersed mycelial growth, further enhanced by the addition of glass beads (5 g per 100 ml).

### 5.2.2 Cryopreservation

To preserve bacteria for long term storage either cells, mycelium or spores were resuspended in 25% glycerol in a suitable medium (LB for *E. coli*, TSB for *Streptomyces*) and flash frozen with liquid nitrogen and stored at -80 °C until further use.

### 5.2.3 Transformation of chemically competent *E. coli* cells

To prepare chemically competent *E. coli* cells, 1 mL of an overnight culture was inoculated into 100 mL of LB medium and incubated at 37 °C with shaking at 200 rpm until the OD<sub>600</sub> reached between 0.4 and 0.6. The cells were then collected by centrifugation at 5000 x *g* for 3 minutes at 4 °C, and resuspended in 3.5 mL of a solution containing 70 mM CaCl<sub>2</sub> and 20 mM MgSO<sub>4</sub>. After 30 minutes incubation on ice, 875 µL of sterile glycerol was added to the suspension. Aliquots of 50 µL were distributed into 1.5 mL Eppendorf tubes, flash frozen, and stored at -80 °C for future use.

For transformation, the competent cells were thawed on ice, then mixed with 5 µL of sample DNA by flicking the tube. After another 30 minutes on ice, the tubes were subjected to a heat shock at 42 °C for 60 seconds, followed by a 5-minute incubation on ice. To recover the transformed cells, 1 mL of SOC medium was added, and the mixture was incubated at 37 °C for 1 hour. Then, 100 µL of the cell suspension was plated onto one half of an LB agar plate with the appropriate antibiotic. The remaining suspension was centrifuged, the pellet was resuspended in the residual medium, and spread onto the second half of the plate and incubated overnight at 37 °C.

### 5.2.4 Electroporation of electrocompetent *E. coli* cells

To create electrocompetent cells, 200 µL of an overnight culture was inoculated into 20 mL of LB medium and incubated at 37 °C with shaking at 200 rpm until the OD<sub>600</sub> reached between 0.4 and 0.6. The cells were then collected by centrifuging at 5000 x *g* for 3 minutes at 4 °C and washed twice with 10 mL of 10% glycerol to remove medium. After washing, the cell pellet was resuspended in 1 mL of glycerol and divided into 100 µL aliquots in 1.5 mL tubes.

For electroporation, 5 µL of sample DNA was combined with an aliquot of the competent cells. This mixture was transferred to a UV-sterilised electroporation cuvette and subjected to a 2.5 kV pulse for 6.5 ms using a BioRad MicroPulser. Immediately following the pulse, 1 mL of SOC medium were added, and the cells were incubated at 37 °C for 1 hour to recover. The transformed cells were then plated on selective agar plates.

### 5.2.5 Intergeneric conjugation between *E. coli* and *Streptomyces*

To transfer successfully cloned constructs from *E. coli* to *Streptomyces*, an intergeneric conjugation protocol was employed. The donor strain used was the methylation-deficient *E. coli* ET12567 harbouring the helper plasmid pUZ8002, while *Streptomyces* spores served as the recipient. The plasmid of interest was transformed into electrocompetent *E. coli* ET12567/pUZ8002 cells and plated on LB agar containing chloramphenicol (25 µg/mL), kanamycin (50 µg/mL) and apramycin (50 µg/mL). After overnight incubation, a single colony was inoculated into 5 mL of LB medium with the appropriate antibiotics and grown overnight to serve as a pre-culture. The following day, 10 mL of LB medium was inoculated with 100 µL of the pre-culture and incubated until an OD<sub>600</sub> of 0.5 was reached. The cells were harvested by centrifugation (5000 rpm, 5 min, 4 °C) and washed twice with cooled LB containing 10 mM MgCl<sub>2</sub> to remove antibiotics. The cells were then resuspended in 500 µL of LB and kept on ice until conjugation. For the recipient preparation, 100 µL of spores were thawed and washed twice with 1 mL of 2xYT medium. The spores were resuspended in 500 µL of fresh 2x YT medium and activated by heating to 50 °C for 10 minutes. Spores were allowed to cool down to room temperature before proceeding with conjugation. For conjugation, the donor and recipient were mixed in a 1.5 mL tube and mixed by pipetting. The bacterial mixture was centrifuged (13,000 x g, 1 min) and the pellet resuspended in 100 µL of TSB. This mixture was then spread on an MS agar plate containing 10 mM MgCl<sub>2</sub> and incubated at 30 °C. After 20 hours, the plate was treated with nalidixic acid (0.5 mg) and apramycin (1 mg), and incubation was continued until ex-conjugants emerged. Single colonies were picked and transferred to new MS agar plates with apramycin (50 µg/mL) and nalidixic acid (25 µg/mL) to confirm the ex-conjugants. Following 5-7 days of incubation, six colonies were validated through colony PCR, and three positive clones cultivated to test production.

### 5.3 Characterisation of biarylittide BGC workflow

The heterologous expression of biosynthetic gene clusters is a common strategy to study the functions and products of genes. As a general workflow, the biarylittide BGCs of interest are amplified from isolated genomic DNA of the producer bacteria via polymerase chain reaction (PCR) and specific primers. Then, linear amplicons are cloned into an integrating plasmid pSET\_152-*ermE*\* which is conjugated to the well-established expression host *Streptomyces coelicolor* M1152. Cultivation of this strain is followed by solid phase extraction and LC-MS analysis of extracts, comparing with extracts from control strains without the additional genes.

The spectral data is subsequently converted to mzXML format and analysed with molecular networking. A detailed description of the heterologous expression of *bytAO* in *S. coelicolor* is described in the following chapter.

### 5.3.1 Heterologous production of biarylittides

*S. coelicolor* M1152 pSET-*bytAO* <sup>[53]</sup> was utilised for the production of Biarylittide YYH. The initial culture inoculated from spores was grown in 20 mL of production medium comprising a 1:1 mixture of yeast extract-malt extract (YEME) broth and TSB medium in a 100 mL Erlenmeyer flask. The culture conditions were maintained at 30 °C with shaking at 200 rpm for 3 days. This seed culture was then used to inoculate 80 mL of TSB/YEME (1:1) in a 300 mL Erlenmeyer flask containing 5 g of 3 mm glass beads (Merck, Darmstadt, Germany) and grown under the same conditions for another 3 days. This secondary culture was used to inoculate five 300 mL Erlenmeyer flasks, each containing 80 mL of TSB/YEME (1:1) and 5 g of glass beads. The cultures were incubated for 7 days at 30 °C, 200 rpm, with the addition of 50 µg/mL apramycin (Sigma-Aldrich).

### 5.3.2 Extraction and isolation of biarylittides from bacterial culture

For the extraction of biarylittides, the cultures were processed as follows: After seven days of incubation, cultures were centrifuged at 10,000 rpm for 20 minutes. The supernatants containing the peptides were incubated with 5% (w/v) HP-20 resin (Supelco/Bellefonte, USA), which was activated earlier by soaking in methanol for 20 minutes and subsequently washed twice with water. The total resin particles in water were transferred to the culture supernatants and incubated at room temperature for one hour with continuous shaking on a shaking platform. After incubation, the resin particles were added to appropriate columns, and the flow-through was discarded. The resin was washed with 5% acetonitrile and 0.1% formic acid and eluted with methanol. The eluate was evaporated using a Laborata 4000 rotary evaporator (Heidolph Instruments GmbH & Co. KG, Schwabach, Germany) and a CVC 24 vacuum controller (Vacubrand GmbH + Co. KG, Wertheim, Germany) with a 40 °C temperature-controlled water bath at 100 rpm until completely dry components were obtained. The samples were stored at -20 °C until further purification with flash chromatography. The dried extract was fractionated using a Reveleris C18 flash column (220 g, 40 µm). The mobile phase consisted of two phases: Phase A: 0.1% trifluoroacetic acid (TFA) in H<sub>2</sub>O, and Phase B: 0.1% TFA in MeOH. The elution gradient applied was as follows: 5% Phase B for 5 minutes, increasing to 25% Phase B over 15 minutes, followed by 100% Phase B from 25 to 30 minutes. Fractions

containing Biarylthide YYH were identified, dried, and further purified by high-performance liquid chromatography (HPLC) using a semi-preparative Nucleodur C18 column (250 × 8 mm, 5 µm, Macherey-Nagel) with a C18 pre-column. The isocratic elution was performed with 70/30 0.05% TFA in H<sub>2</sub>O/MeOH at a flow rate of 2.5 mL/min. The purification was repeated twice under the same conditions to yield 15 mg of pure Biarylthide YYH as a white powder from a total of 13 L of cultivation broth. Purity was assessed using HPLC, LC-MS and NMR spectroscopy.

## 5.4 Molecular biological methods

### 5.4.1 Polymerase chain reaction

For cloning procedures Q5 High-Fidelity DNA Polymerase (NEB) was routinely used. The standard Polymerase chain reaction (PCR) reaction setup (Table 14) and standard thermal cycling conditions (Table 15) for Q5 Polymerase were adjusted accordingly depending on template DNA, amplicon length und primers used.

**Table 14** Standard PCR reaction setup when using Q5 Polymerase

Component	25 µL reaction	Final concentration
5× Q5 reaction buffer	5 µL	1×
5× Q5 high GC enhancer	(5 µL)	1×
10 mM dNTPs	0.5 µL	200 µM
10 µM forward primer	1.25 µL	0.5 µM
10 µM reverse primer	1.25 µL	0.5 µM
Template DNA	variable	~50 ng genomic ~1 ng plasmid
Q5 DNA polymerase	0.125 µL	0.25 U
Nuclease-free water	to 25 µL	-

**Table 15** Thermal Cycling conditions for Q5 PCR amplification. Optimal annealing temperature (\*) was calculated using the NEB T<sub>m</sub> Calculator and experimentally determined by gradient PCR.

Step	Temperature	Time
Initial denaturation	98 °C	30-120 s
	98 °C	10 s
30 cycles	50-72 °C*	15 s
	72 °C	30 s/kb
Final extension	72 °C	2 min
Hold	4 °C	-

To screen for the successful integration of inserted DNA, colony PCR was employed. Individual colonies were picked with a sterile 10 µL tip and resuspended in 5 µL of water, which served as the template. This template was then added to 20 µL of GoTaq PCR master mix, prepared as outlined in Table 16, and amplified in a thermal cycler under the conditions specified in

**Table 16** Standard PCR reaction setup using GoTaq DNA polymerase.

Component	25 µL reaction	Final concentration
5× Green GoTaq reaction buffer	5 µL	1×
25 mM MgCl <sub>2</sub>	2.5 µL	2.5 mM
10 mM dNTPs	1 µL	200 µM
10 µM forward primer	1.25 µL	0.5 µM
10 µM reverse primer	1.25 µL	0.5 µM
DMSO	1 µL	4%
Template DNA	variable	~1-50 ng
GoTaq DNA polymerase (5 u/µl)	0.25 µL	1.25 U
Nuclease-free water	to 25 µL	-

**Table 17** Thermal cycling conditions for GoTaq PCR amplification. Optimal annealing temperature (\*) was calculated using the Promega T<sub>m</sub> Calculator and experimentally determined by gradient PCR.

Step	Temperature	Time
Initial denaturation	95 °C	120 s
	95 °C	30 s
25 cycles	50-72 °C*	30 s



	72 °C	1 min/kb
Final extension	72 °C	2 min
Hold	4 °C	

#### 5.4.2 DNA isolation

For PCR and cloning procedures, a high purity template, genomic or plasmid DNA, is required. DNA was routinely isolated according to manufacturer's protocols (for genomic DNA: GenElute™ Bacterial Genomic DNA Kit (Thermo Fisher Scientific) and for plasmid DNA: FastGene Plasmid Mini Kit (NIPPON Genecits EUROPE)).

#### 5.4.3 Restriction digestion of DNA

Linear DNA and circular plasmids were digested with restriction enzymes like *EcoRI* (NEB) to generate fragments with overlapping (sticky) ends, facilitating subsequent ligation, or to analyse cloning success. All digestions were performed in rCutSmart Buffer, enabling concurrent digestion with multiple restriction enzymes. Table 12 details the composition of a typical digest. The reactions were incubated at 37 °C for 60 minutes, before adding 1 µl of Shrimp alkaline phosphatase to desphosphorylate 5'-ends of the DNA-fragments by incubating for another 60 minutes. For the following cloning step, the enzymes were either heat-inactivated or purified using the Promega PCR Extraction Kit.

**Table 18** Composition of restriction digestion reactions.

Component	Volume
DNA (1 µg)	variable
10x rCutsmartBuffer	1 µL
Restriction enzyme	1 µL
Water	to 20 µL

#### 5.4.4 Agarose gel electrophoresis

Agarose gel electrophoresis was employed as an analytical technique to confirm the presence of the desired DNA products. Agarose was dissolved at 1% (w/v) in 1x TAE buffer using a microwave. After cooling for a few minutes 0.7 µL of Midori Green Direct (Nippon Genetics) was added to 25 ml of liquid agarose solution, which was then poured into a cast, and a comb was inserted to create wells for the samples. For analysing fragments with high molecular

weight, such as gDNA, 0.7% (w/v) agarose was utilised. Once the gel had solidified, the mold was filled with 1x TAE running buffer. Prior to loading, 5  $\mu$ L of the sample was mixed with 1  $\mu$ L of 6x TriTrack DNA loading dye (ThermoFisher). As a reference, 4  $\mu$ L of GeneRuler DNA Ladder Mix (ThermoFisher) was used. After all samples were loaded, an electric voltage of 120 V was applied for 30 minutes and subsequently the gel was analysed in a UV cabinet.

#### **5.4.5 Isolation of DNA from agarose gels**

If PCR products matched the predicted size, DNA bands were cut out from the gel and transferred to a 1.5 mL Eppendorf tube and purified using the Zymoclean™ Gel DNA Recovery Kit following the manufacturers' instructions. To achieve high concentrations of DNA, the final elution step was performed with 6  $\mu$ L of water. This step was particularly important when PCR product had low yield.

#### **5.4.6 Ligation**

Ligation was carried out with DNA fragments possessing compatible overhangs using T4 DNA ligase (NEB). A mixture of 8  $\mu$ L DNA (at a 1:3 molar ratio of vector to insert), 1  $\mu$ L of 10x T4 ligase buffer, and 1  $\mu$ L of T4 ligase was prepared and incubated overnight at 16 °C. This ligation mixture was then directly used for transformation into *E. coli*.

#### **5.4.7 Gibson assembly**

Gibson assembly, an isothermal DNA amplification and cloning method, was employed in this thesis to generate constructs for heterologous expression and site-directed mutagenesis of plasmid constructs. Primers were designed with a 20-base overlap, and educt fragments were generated via Q5-PCR using genomic DNA, plasmid DNA, or the plasmid to be mutated as a template. Post-PCR purification involved removing any remaining template DNA by adding 1  $\mu$ L *Dpn*I and incubating at 37 °C for 1 hour. For the assembly, 5  $\mu$ L of DNA (with insert and vector mixed at a 3:1 ratio) were combined with 15  $\mu$ L of Gibson assembly master mix (Table 6) and incubated at 50 °C for 60 minutes. The resulting mixture was purified using the DNA Clean & Concentrator™-5 Kit (Zymo Research), eluted in 5  $\mu$ L water, and subsequently used for chemical transformation into *E. coli*.

**Table 19** Composition of Gibson assembly master mix. Aliquots of 15  $\mu\text{L}$  were stored at  $-20\text{ }^{\circ}\text{C}$  until further.

Component	Volume
5x ISO buffer	320 $\mu\text{L}$ (1x)
T5 exonuclease (10 U/ $\mu\text{L}$ )	0.64 $\mu\text{L}$
Phusion polymerase (2 U/ $\mu\text{L}$ )	20 $\mu\text{L}$
Taq ligase (40 U/ $\mu\text{L}$ )	160 $\mu\text{L}$
Nuclease-free water	1.2 mL

### 5.4.8 Sanger sequencing

Plasmids and PCR amplicons were sent for Sanger sequencing using the LightRun or TubeSeq Supreme service from Eurofins Genomics with suitable primers. Resulting files containing the sequence and quality information were downloaded and ab1 files aligned and constructs confirmed in SnapGene 5.

## 5.5 Protein Methods

### 5.5.1 Protein expression

Protein expression was performed in 100-1000 mL of LB or TB medium, depending on the protein type and expression levels. Baffled DURAN® flasks with screw closures were used to ensure sufficient aeration, with the culture medium not exceeding 30% of the flask volume. For the expression of P450 enzymes,  $\delta$ -aminolevulinic acid (200  $\mu\text{g/L}$ ) was added to facilitate haem incorporation. Media were inoculated with 1% of an overnight culture and grown on a rotary shaker at 200 rpm and  $37\text{ }^{\circ}\text{C}$  until the optical density at 600 nm reached 0.4 - 0.6 for LB or 0.8 - 1.2 for TB. At this point, cultures were cold-shocked on ice for 30 minutes, and protein expression was induced by adding IPTG (0.05 - 0.1 mM), followed by incubation at  $18\text{ }^{\circ}\text{C}$  for 16 hours or under conditions optimal for each protein.

### 5.5.2 Protein purification with Ni-NTA

Cells were harvested by centrifugation (4,000  $\times g$ ,  $4\text{ }^{\circ}\text{C}$ , 10 min), and the resulting pellet was resuspended in chilled lysis buffer. To minimise protein degradation, all subsequent steps were conducted on ice. Lysozyme (1 mg/mL) was added to lyse the cells, followed by sonication (3 rounds of 1 min). The lysate was clarified by centrifugation (10,000  $\times g$ ,  $4\text{ }^{\circ}\text{C}$ , 10 min), and the supernatant was incubated with Ni-NTA resin (500  $\mu\text{L}/100\text{ mL}$  culture) for 1 hour. The resin

was washed twice with 4 mL buffers containing increasing concentrations of imidazole (wash I and wash II) to remove non-specifically bound proteins. The target protein was eluted with a high-imidazole concentration buffer (elution buffer).

The eluted protein was concentrated and buffer-exchanged into storage buffer using a 30 kDa MWCO centrifuge filter unit (Amicon). Protein concentration was determined using a Nanodrop spectrophotometer, requiring the protein's molecular weight and molar extinction coefficient (computed via ProtParam, ExPASy). The concentration was adjusted to the desired value (e.g., 100  $\mu$ M), flash-frozen with liquid nitrogen, and stored at -80 °C until further use. Samples (30  $\mu$ L) of each fraction (total protein, soluble protein, flow-through, wash I, wash II, and elution) were collected for subsequent analysis with SDS-PAGE.

### **5.5.3 Anion exchange and size exclusion chromatography**

Proteins requiring further purification were subjected to anion exchange chromatography (AEX) and size exclusion chromatography (SEC). For AEX, the protein was dialysed overnight into a low-salt buffer (AEX Buffer A) and eluted with a high-salt buffer (AEX Buffer B). Fractions containing the target enzyme were combined and concentrated for SEC, which was performed using a Sepax SRT-10 SEC-300 column on an Äkta system (Cytiva) with SEC buffer. The purified protein fractions were pooled and concentrated, and the final protein concentration was determined using a Nanodrop spectrophotometer. The purified protein was then aliquoted, flash-frozen in liquid nitrogen, and stored at -80 °C until further use.

### **5.5.4 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)**

SDS-PAGE was employed for the qualitative analysis of protein expression and purification. Each sample (30  $\mu$ L) was mixed with 10  $\mu$ L of NuPAGE LDS Sample Buffer (4x) and 4  $\mu$ L of NuPAGE Sample Reducing Agent (10x). From this mixture, 10  $\mu$ L was loaded into a well of a 12% acrylamide gel. The Color Prestained Protein Standard, Broad Range (10-250 kDa) from NEB served as the molecular weight ladder. The proteins were separated at 200 V for 1 h using a Bio-Rad PowerPac. After electrophoresis, the gel was stained with 25 mL of staining solution (0.1% w/v brilliant blue, 1% acetic acid, and 5% methanol) for 20 minutes. Following staining, the gel was washed twice with water and then immersed in a destaining solution (10% acetic acid, 20% methanol, 70% water) overnight before analysis.

### 5.5.5 *In vitro* turnover assay with P450s

Standard enzyme reactions were performed in Eppendorf tubes and contained BytO (0.5  $\mu$ M) supplemented with the redox partners PuR (0.5  $\mu$ M) and PuxB A105V variant (2.5  $\mu$ M)<sup>[118]</sup> in 50 mM HEPES buffer pH 7.0 containing 50 mM NaCl. NADH (2 mM) was used as electron donor with an additional regeneration system consisting of glucose (0.33%) and glucose dehydrogenase (0.033 mg/mL). Synthetic peptide substrates were added to a final concentration of 25  $\mu$ M. Reactions were incubated overnight at 30 °C with shaking at 300 rpm.

Products were isolated using solid phase extraction cartridges (Bond Elut C18, Agilent). Cartridges were activated with methanol containing 0.1% formic acid (500  $\mu$ L) followed by equilibration with water containing 0.1% formic acid (1 mL). Samples were diluted to 1 mL with acidified water before loading. After washing with acidified water (1 mL), products were eluted with 600  $\mu$ L methanol containing 0.1% formic acid. For preparative-scale reactions, products were eluted using an acetonitrile/water mixture (0.1% formic acid). Methanol fractions were concentrated using an Eppendorf Concentrator while acetonitrile fractions were lyophilized. Dried samples were reconstituted in 20% acetonitrile/water (0.1% formic acid) for LC-MS analysis.

### 5.5.6 *In vitro* aspartyl $\beta$ -hydroxylation assay

Standard enzyme reactions were performed in Eppendorf tubes with a final volume of 100  $\mu$ L. The reaction mixture contained 50 mM Tris-HCl (pH 7.8), 5 mM substrate, 2 mM L-ascorbic acid, bovine serum albumin (0.1%), ferrous ammonium sulfate (50  $\mu$ M), dithiothreitol (10  $\mu$ M), and  $\alpha$ -Ketoglutaric acid (125  $\mu$ M). After adding these components in a volume of 25  $\mu$ L, 35  $\mu$ L of water was added. Following a 5-minute pre-incubation at 37 °C in a water bath, reactions were initiated by adding 40  $\mu$ L of enzyme source and incubated for 3 hours with shaking. Reactions were terminated by adding 100  $\mu$ L of 0.5 M  $\text{KH}_2\text{PO}_4$ . Products were recovered using C18 Sep-Pak cartridges. Cartridges were first washed with 5 mL of 100% methanol and equilibrated with 5 mL of water. Reaction mixtures were loaded onto the cartridges and washed with 5 mL of water. Products were eluted with 2 mL of 100% methanol and evaporated to dryness. The extracts dissolved in 1 mL of 10% acetonitrile and subjected to LCMS analysis.

## 5.6 Peptide synthesis

Peptide synthesis was carried out in accordance with established protocols <sup>[148]</sup> on a 0.1 mmol scale employing N,N'-Diisopropylcarbodiimide (DIC) and Oxyma Pure as coupling agents. Initially, 100 mg of 2-chlorotriyl chloride resin was subjected to a series of washes with dichloromethane (DCM) (3 x 2 mL), followed by swelling in DCM (2 mL) for 30 minutes and subsequent washing with DCM (3 x 2 mL). The first Fmoc-protected amino acid (0.15 mmol) was conjugated to the resin overnight using N,N-Diisopropylethylamine (DIEA). To block any remaining reactive chlorotriyl sites, the resin was treated with a mixture of dimethylformamide (DMF), DIEA, and methanol (MeOH) in a 7:2:1 ratio (3 x 2 mL). The Fmoc protecting group was then removed with 20% piperidine in DMF. Subsequent coupling of Fmoc-amino acids (3 equivalents) was conducted using DIC (3 equivalents) and Oxyma Pure (6 equivalents) at 50 °C for 10 minutes. This cycle of deprotection and coupling was repeated until the desired peptide sequence was fully assembled. A final deprotection step was carried out to remove the Fmoc group from the terminal amino acid. Peptide cleavage from the resin was achieved with a solution comprising trifluoroacetic acid (TFA), triisopropylsilane (TIS), and water in a 95:2.5:2.5 ratio, with gentle agitation for 1 hour at room temperature. The resin was separated from the cleavage mixture by filtration. The resin was washed with TFA (2 x 1 mL), and the combined filtrate was concentrated under a nitrogen stream to approximately 1 mL. Peptide precipitation was accomplished using pre-cooled diethyl ether (20 mL), followed by centrifugation in a spark-free centrifuge (Spintron, GT-175) and washing with diethyl ether (3 x 10 mL). The resulting peptides were freeze dried and a portion resuspended and analysed by LC-MS.

## 6 References

- [1] A. G. Atanasov, S. B. Zotchev, V. M. Dirsch, C. T. Supuran, *Nature Reviews Drug Discovery* **2021**, 20, 200.
- [2] A. Greule, J. E. Stok, J. J. de Voss, M. J. Cryle, *Natural Product Reports* **2018**, 35, 757.
- [3] T. L. Poulos, E. F. Johnson, In: Ortiz de Montellano, P. (eds) *Cytochrome P450*. Springer, Cham. **2015**, pp. 3–32.
- [4] A. W. Munro, H. M. Girvan, K. J. McLean, *Biochimica et Biophysica Acta* **2007**, 1770, 345.
- [5] I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, *Chemical Reviews* **2005**, 105, 2253.
- [6] J. Rittle, M. T. Green, *Science* **2010**, 330, 933.
- [7] J. T. Groves, *Journal of Inorganic Biochemistry* **2006**, 100, 434.
- [8] F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochimica et Biophysica Acta* **2007**, 1770, 330.
- [9] A. W. Munro, D. G. Leys, K. J. McLean, K. R. Marshall, T. W. B. Ost, S. Daff, C. S. Miles, S. K. Chapman, D. A. Lysek, C. C. Moser et al., *Trends in Biochemical Sciences* **2002**, 27, 250.
- [10] S. Uhlmann, R. D. Süßmuth, M. J. Cryle, *ACS Chemical Biology* **2013**, 8, 2586.
- [11] M. V. Mendes, N. Antón, J. F. Martín, J. F. Aparicio, *The Biochemical Journal* **2005**, 386, 57.
- [12] H. Ogura, C. R. Nishida, U. R. Hoch, R. Perera, J. H. Dawson, P. R. Ortiz de Montellano, *Biochemistry* **2004**, 43, 14712.
- [13] K. Haslinger, M. Peschke, C. Brieke, E. Maximowitsch, M. J. Cryle, *Nature* **2015**, 521, 105.
- [14] B. Zhao, F. P. Guengerich, A. Bellamine, D. C. Lamb, M. Izumikawa, L. Lei, L. M. Podust, M. Sundaramoorthy, J. A. Kalaitzis, L. M. Reddy et al., *The Journal of Biological Chemistry* **2005**, 280, 11599.
- [15] J. C. Carlson, S. Li, S. S. Gunatilleke, Y. Anzai, D. A. Burr, L. M. Podust, D. H. Sherman, *Nature Chemistry* **2011**, 3, 628.
- [16] Y. Liu, C. Wang, J. Yan, W. Zhang, W. Guan, X. Lu, S. Li, *Biotechnology for Biofuels and Bioproducts* **2014**, 7, 28.
- [17] S. M. Barry, J. A. Kers, E. G. Johnson, L. Song, P. R. Aston, B. Patel, S. B. Krasnoff, B. R. Crane, D. M. Gibson, R. Loria et al., *Nature Chemical Biology* **2012**, 8, 814.

- [18] P. Urban, T. Lautier, D. Pompon, G. Truan, *International Journal of Molecular Sciences* **2018**, *19*.
- [19] V. Cojocaru, P. J. Winn, R. C. Wade, *Biochimica et Biophysica Acta* **2007**, *1770*, 390.
- [20] P. J. Mak, I. G. Denisov, *Biochimica et Biophysica acta (BBA) - Proteins and Proteomics* **2018**, *1866*, 178.
- [21] T. J. Deng, L. M. Proniewicz, J. R. Kincaid, H. Yeom, I. D. Macdonald, S. G. Sligar, *Biochemistry* **1999**, *38*, 13699.
- [22] N. R. Orme-Johnson, W. H. Orme-Johnson, *Methods in Enzymology* **1978**, *52*, 252.
- [23] C. Jung, *Biochimica et Biophysica Acta* **2002**, *1595*, 309.
- [24] J. D. Rudolf, C.-Y. Chang, M. Ma, B. Shen, *Natural Product Reports* **2017**, *34*, 1141.
- [25] Y. Xue, D. Wilson, L. Zhao, H. w. Liu, D. H. Sherman, *Chemistry & biology* **1998**, *5*, 661.
- [26] E. I. Graziani, D. E. Cane, M. C. Betlach, J. T. Kealey, R. McDaniel, *Bioorganic & Medicinal Chemistry Letters* **1998**, *8*, 3117.
- [27] K. Woithe, N. Geib, K. Zerbe, D. B. Li, M. Heck, S. Fournier-Rousset, O. Meyer, F. Vitali, N. Matoba, K. Abou-Hadeed et al., *Journal of American Chemical Society*. **2007**, *129*, 6887.
- [28] P. Nanudorn, S. Thiengmag, F. Biermann, P. Erkoc, S. D. Dirnberger, T. N. Phan, R. Fürst, R. Ueoka, E. J. N. Helfrich, *Angewandte Chemie International Edition* **2022**, *61*, e202208361.
- [29] H. Nam, J. S. An, J. Lee, Y. Yun, H. Lee, H. Park, Y. Jung, K.-B. Oh, D.-C. Oh, S. Kim, *Journal of the American Chemical Society* **2023**, *145*, 22047.
- [30] R. D. Süßmuth, A. Mainz, *Angewandte Chemie International Edition* **2017**, *56*, 3770.
- [31] K. J. Weissman, *Nature Chemical Biology* **2015**, *11*, 660.
- [32] M. A. Marahiel, *Natural Product Reports* **2016**, *33*, 136.
- [33] L. Padva, J. Gullick, L. Coe, M. Hansen, J. de Voss, M. Crüsemann, M. J. Cryle, *ChemBioChem* **2024**, e202400916.
- [34] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy et al., *Natural Product Reports* **2013**, *30*, 108.
- [35] M. Montalbán-López, T. A. Scott, S. Ramesh, I. R. Rahman, A. J. van Heel, J. H. Viel, V. Bandarian, E. Dittmann, O. Genilloud, Y. Goto et al., *Natural Product Reports* **2021**, *38*, 130.
- [36] M. A. Ortega, W. A. van der Donk, *Cell Chemical Biology* **2016**, *23*, 31.



- [37] T. J. Oman, W. A. van der Donk, *Nature Chemical Biology* **2010**, 6, 9.
- [38] K. J. Hetrick, W. A. van der Donk, *Current Opinion in Chemical Biology* **2017**, 38, 36.
- [39] J. R. van der Meer, H. S. Rollema, R. J. Siezen, M. M. Beerthuyzen, O. P. Kuipers, W. M. de Vos, *Journal of Biological Chemistry* **1994**, 269, 3555.
- [40] B. Li, J. P. J. Yu, J. S. Brunzelle, G. N. Moll, W. A. van der Donk, S. K. Nair, *Science* **2006**, 311, 1464.
- [41] L. S. Håvarstein, H. Holo, I. F. Nes, *Microbiology* **1994**, 140 ( Pt 9), 2383.
- [42] O. McAuliffe, R. P. Ross, C. Hill, *FEMS Microbiology Reviews* **2001**, 25, 285.
- [43] M. O. Maksimov, S. J. Pan, A. James Link, *Natural Product Reports* **2012**, 29, 996.
- [44] J. Koehnke, A. F. Bent, W. E. Houssen, G. Mann, M. Jaspars, J. H. Naismith, *Current Opinion in Structural Biology* **2014**, 29, 112.
- [45] F. Collin, A. Maxwell, *Journal of Molecular Biology* **2019**, 431, 3400.
- [46] M. C. Bagley, J. W. Dale, E. A. Merritt, X. Xiong, *Chemical Reviews* **2005**, 105, 685.
- [47] L. Flühe, M. A. Marahiel, *Current Opinion in Chemical Biology* **2013**, 17, 605.
- [48] L. Franz, U. Kazmaier, A. W. Truman, J. Koehnke, *Natural Product Reports* **2021**, 38, 1659.
- [49] G. A. Hudson, D. A. Mitchell, *Current Opinion in Microbiology* **2018**, 45, 61.
- [50] A. Hurst in *Advances in Applied Microbiology*, v.27, Academic Press, New York, **1981**, pp. 85–123.
- [51] Y. Imai, K. J. Meyer, A. Iinishi, Q. Favre-Godal, R. Green, S. Manuse, M. Caboni, M. Mori, S. Niles, M. Ghiglieri et al., *Nature* **2019**, 576, 459.
- [52] M. M. Zdouc, M. M. Alanjary, G. S. Zarazúa, S. I. Maffioli, M. Crüsemann, M. H. Medema, S. Donadio, M. Sosio, *Cell Chemical Biology* **2021**, 28, 733-739.e4.
- [53] M. M. Zdouc, M. Iorio, S. I. Maffioli, M. Crüsemann, S. Donadio, M. Sosio, *Journal of Natural Products* **2021**, 84, 204.
- [54] M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapon, T. Luzzatto-Knaan et al., *Nature Biotechnology* **2016**, 34, 828.
- [55] J. E. González-Pastor, J. L. San Millán, F. Moreno, *Nature* **1994**, 369, 281.
- [56] M. H. Medema, M. A. Fischbach, *Nature Chemical Biology* **2015**, 11, 639.
- [57] M. A. Skinnider, C. W. Johnston, M. Gunabalasingam, N. J. Merwin, A. M. Kieliszek, R. J. MacLellan, H. Li, M. R. M. Ranieri, A. L. H. Webster, M. P. T. Cao et al., *Nature Communications* **2020**, 11, 6058.

- [58] G. D. Hannigan, D. Prihoda, A. Palicka, J. Soukup, O. Klempir, L. Rampula, J. Durcak, M. Wurst, J. Kotowski, D. Chang et al., *Nucleic Acids Research* **2019**, *47*, e110.
- [59] J. C. Navarro-Muñoz, N. Selem-Mojica, M. W. Mullowney, S. A. Kautsar, J. H. Tryon, E. I. Parkinson, E. L. C. de Los Santos, M. Yeong, P. Cruz-Morales, S. Abubucker et al., *Nature Chemical Biology* **2020**, *16*, 60.
- [60] S. A. Kautsar, K. Blin, S. Shaw, J. C. Navarro-Muñoz, B. R. Terlouw, J. J. J. van der Hooft, J. A. van Santen, V. Tracanna, H. G. Suarez Duran, V. Pascal Andreu et al., *Nucleic Acids Research* **2020**, *48*, D454-D458.
- [61] P. Agrawal, S. Khater, M. Gupta, N. Sain, D. Mohanty, *Nucleic Acids Research* **2017**, *45*, W80-W88.
- [62] K. Blin, S. Shaw, H. E. Augustijn, Z. L. Reitz, F. Biermann, M. Alanjary, A. Fetter, B. R. Terlouw, W. W. Metcalf, E. J. N. Helfrich et al., *Nucleic Acids Research* **2023**, *51*, W46-W50.
- [63] J. I. Tietz, C. J. Schwalen, P. S. Patel, T. Maxson, P. M. Blair, H.-C. Tai, U. I. Zakai, D. A. Mitchell, *Nature Chemical Biology* **2017**, *13*, 470.
- [64] A. H. Russell, A. W. Truman, *Computational and Structural Biotechnology Journal* **2020**, *18*, 1838.
- [65] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *Journal of Molecular Biology* **1990**, *215*, 403.
- [66] J. J. Hug, N. A. Frank, C. Walt, P. Šenica, F. Panter, R. Müller, *Molecules* **2021**, *26*, 7483.
- [67] T. Yasuzawa, K. Shirahata, H. Sano, *The Journal of Antibiotics* **1987**, *40*, 455.
- [68] S. Saito, K. Atsumi, T. Zhou, K. Fukaya, D. Urabe, N. Oku, M. R. U. Karim, H. Komaki, Y. Igarashi, *Beilstein Journal of Organic Chemistry* **2020**, *16*, 1100.
- [69] S. Sano, K. Ikai, K. Katayama, K. Takesako, T. Nakamura, A. Obayashi, Y. Ezure, H. Enomoto, *The Journal of Antibiotics* **1986**, *39*, 1685.
- [70] Y. L. Hu, F. Z. Yin, J. Shi, S. Y. Ma, Z. R. Wang, R. X. Tan, R. H. Jiao, H. M. Ge, *Journal of the American Chemical Society* **2023**, *145*, 27325.
- [71] B.-B. He, J. Liu, Z. Cheng, R. Liu, Z. Zhong, Y. Gao, H. Liu, Z.-M. Song, Y. Tian, Y.-X. Li, *Angewandte Chemie International Edition* **2023**, *62*, e202311533.
- [72] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Research* **1997**, *25*, 3389.
- [73] R. Zallot, N. Oberg, J. A. Gerlt, *Biochemistry* **2019**, *58*, 4169.

- [74] B.-B. He, Z. Cheng, Z. Zhong, Y. Gao, H. Liu, Y.-X. Li, *Angewandte Chemie International Edition* **2022**, *61*, e202212447.
- [75] R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim et al., *BioRxiv* **2021**, 2021.10.04.463034
- [76] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, *Nature Methods* **2022**, *19*, 679.
- [77] H. Aldemir, S. Shu, F. Schaefer, H. Hong, R. Richarz, S. Harteis, M. Einsiedler, T. M. Milzarek, S. Schneider, T. A. M. Gulder, *Chemistry – A European Journal* **2022**, *28*, e202103389.
- [78] Y. Hitora, A. H. El-Desoky, Y. Sadahiro, A. Sejiyama, A. Kinoshita, Y. Ise, E. D. Angkouw, R. E. P. Mangindaan, T. Higaki, S. Tsukamoto, *Journal of Natural Products* **2024**, *87*, 1197.
- [79] B. Zhou, G. Shetye, Y. Yu, B. D. Santarsiero, L. L. Klein, C. Abad-Zapatero, N. M. Wolf, J. Cheng, Y. Jin, H. Lee et al., *Journal of Natural Products* **2020**, *83*, 657.
- [80] N. M. Wolf, H. Lee, M. P. Choules, G. F. Pauli, R. Phansalkar, J. R. Anderson, W. Gao, J. Ren, B. D. Santarsiero, H. Lee et al., *ACS Infectious Diseases* **2019**, *5*, 829.
- [81] M. P. Choules, N. M. Wolf, H. Lee, J. R. Anderson, E. M. Grzelak, Y. Wang, R. Ma, W. Gao, J. B. McAlpine, Y.-Y. Jin et al., *Antimicrobial Agents and Chemotherapy* **2019**, *63*.
- [82] H. Tomita, Y. Katsuyama, H. Minami, Y. Ohnishi, *Journal of Biological Chemistry* **2017**, *292*, 15859.
- [83] G. Perez Ortiz, J. D. Sidda, E. L. C. de Los Santos, C. B. Hubert, S. M. Barry, *Chemical Communications* **2021**, *57*, 11795.
- [84] J. Ma, H. Huang, Y. Xie, Z. Liu, J. Zhao, C. Zhang, Y. Jia, Y. Zhang, H. Zhang, T. Zhang et al., *Nature Communications* **2017**, *8*, 391.
- [85] Y. Wang, J. He, M. S. Alam, F. Wang, Z. Shang, Y. Chen, C. Sun, Z. Lu, Y. Gao, T. Zhang et al., *ACS Synthetic Biology* **2024**, *13*, 930.
- [86] S. Jordan, B. Li, E. Traore, Y. Wu, R. Usai, A. Liu, Z.-R. Xie, Y. Wang, *Journal of Biological Chemistry* **2023**, *299*, 105049.
- [87] B. D. Dratch, K. L. McWhorter, T. C. Blue, S. K. Jones, S. M. Horwitz, K. M. Davis, *ACS Chemical Biology* **2023**, *18*, 1713.
- [88] L. Padva, L. Zimmer, J. Gullick, Y. Zhao, V. M. Sasi, R. B. Schittenhelm, C. J. Jackson, M. J. Cryle, M. Crüsemann, *Chem* **2025**, *11*, 102438.

- [89] F. Biermann, B. Tan, M. Breitenbach, Y. Kakumu, P. Nanudorn, Y. Dimitrova, A. S. Walker, R. Ueoka, E. J. N. Helfrich, *Chemical Science* **2024**, *15*, 17506.
- [90] K. R. Schramma, L. B. Bushin, M. R. Seyedsayamdost, *Nature Chemistry* **2015**, *7*, 431.
- [91] E. J. Culp, N. Waglechner, W. Wang, A. A. Fiebig-Comyn, Y.-P. Hsu, K. Koteva, D. Sychantha, B. K. Coombes, M. S. van Nieuwenhze, Y. V. Brun et al., *Nature* **2020**, *578*, 582.
- [92] P. Belin, M. H. Le Du, A. Fielding, O. Lequin, M. Jacquet, J.-B. Charbonnier, A. Lecoq, R. Thai, M. Courçon, C. Masson et al., *Proceedings of the National Academy of Sciences* **2009**, *106*, 7426.
- [93] M. J. Cryle, S. G. Bell, I. Schlichting, *Biochemistry* **2010**, *49*, 7282.
- [94] A. Zorzi, K. Deyle, C. Heinis, *Current Opinion in Chemical Biology* **2017**, *38*, 24.
- [95] H. Li, W. Ding, Q. Zhang, *RSC Chemical Biology* **2024**, *5*, 90.
- [96] Y. Zhao, E. Marschall, M. Treisman, A. McKay, L. Padva, M. Crüsemann, D. R. Nelson, D. L. Steer, R. B. Schittenhelm, J. Tailhades et al., *Angewandte Chemie International Edition* **2022**, *61*, e202204957.
- [97] A. Howarth, J. M. Goodman, *Chemical Science* **2022**, *13*, 3507.
- [98] S. Ariyasu, J. K. Stanfield, Y. Aiba, O. Shoji, *Current Opinion in Chemical Biology* **2020**, *59*, 155.
- [99] M. H. Hansen, E. Stegmann, M. J. Cryle, *Current Opinion in Biotechnology* **2022**, *77*, 102767.
- [100] H. Aldemir, S. Shu, F. Schaefer, H. Hong, R. Richarz, S. Harteis, M. Einsiedler, T. M. Milzarek, S. Schneider, T. A. M. Gulder, *Chemistry – A European Journal* **2022**, *28*, e202103389.
- [101] K. Haslinger, M. J. Cryle, *FEBS letters* **2016**, *590*, 571.
- [102] M. J. Cryle, J. Staaden, I. Schlichting, *Archives of Biochemistry and Biophysics* **2011**, *507*, 163.
- [103] H. S. Ali, R. H. Henchman, S. P. de Visser, *Organic & Biomolecular Chemistry* **2020**, *18*, 4610.
- [104] S. Kunakom, H. Otani, D. W. Uduary, D. T. Doering, N. J. Mouncey, *Journal of Industrial Microbiology & Biotechnology* **2023**, *50*.
- [105] L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang, C. Fu, *Signal Transduction and Targeted Therapy* **2022**, *7*, 48.
- [106] M. Dell, K. L. Dunbar, C. Hertweck, *Natural Product Reports* **2022**, *39*, 453.

- [107] A. Bhushan, P. J. Egli, E. E. Peters, M. F. Freeman, J. Piel, *Nature Chemistry* **2019**, *11*, 931.
- [108] N. M. Bösch, M. Borsa, U. Greczmiel, B. I. Morinaka, M. Gugger, A. Oxenius, A. L. Vagstad, J. Piel, *Angewandte Chemie International Edition* **2020**, *59*, 11763.
- [109] F. Hubrich, N. M. Bösch, C. Chepkirui, B. I. Morinaka, M. Rust, M. Gugger, S. L. Robinson, A. L. Vagstad, J. Piel, *Proceedings of the National Academy of Sciences* **2022**, *119*.
- [110] R. S. Ayikpoe, L. Zhu, J. Y. Chen, C. P. Ting, W. A. van der Donk, *ACS Central Science* **2023**, *9*, 1008.
- [111] J. Z. Acedo, I. R. Bothwell, L. An, A. Trouth, C. Frazier, W. A. van der Donk, *Journal of the American Chemical Society* **2019**, *141*, 16790.
- [112] L. An, D. P. Cogan, C. D. Navo, G. Jiménez-Osés, S. K. Nair, W. A. van der Donk, *Nature Chemical Biology* **2018**, *14*, 928.
- [113] F. Hubrich, S. K. Kandy, C. Chepkirui, C. Padhi, S. Mordhorst, P. Moosmann, T. Zhu, M. Gugger, J. R. Chekan, J. Piel, *Chem* **2024**, *10*, 3224.
- [114] S. Mordhorst, F. Ruijne, A. L. Vagstad, O. P. Kuipers, J. Piel, *RSC Chemical Biology* **2023**, *4*, 7.
- [115] H. Tomita, Y. Katsuyama, H. Minami, Y. Ohnishi, *Journal of Biological Chemistry* **2017**, *292*, 15859.
- [116] S. S. Qi, A. Bogdanov, M. Cnockaert, T. Acar, S. Ranty-Roby, T. Coenye, P. Vandamme, G. M. König, M. Crüsemann, A. Carlier, *Environmental Microbiology* **2021**, *23*, 2132.
- [117] Y. Tong, C. M. Whitford, K. Blin, T. S. Jørgensen, T. Weber, S. Y. Lee, *Nature Protocols* **2020**, *15*, 2470.
- [118] S. G. Bell, F. Xu, E. O. D. Johnson, I. M. Forward, M. Bartlam, Z. Rao, L.-L. Wong, *Journal of Biological Inorganic Chemistry* **2010**, *15*, 315.
- [119] B.-B. He, Z. Cheng, J. Liu, R. Liu, Z. Zhong, Y. Gao, H. Liu, Y.-X. Li, *Angewandte Chemie International Edition* **2023**, *62*, e202311533
- [120] I. H. Sarker, *Sn Computer Science* **2021**, *2*, 160.
- [121] L. Breiman, *Machine Learning* **2001**, *45*, 5.
- [122] B. R. Terlouw, K. Blin, J. C. Navarro-Muñoz, N. E. Avalon, M. G. Chevrette, S. Egbert, S. Lee, D. Meijer, M. J. J. Recchia, Z. L. Reitz et al., *Nucleic Acids Research* **2023**, *51*, D603-D610.
- [123] W. Li, A. Godzik, *Bioinformatics* **2006**, *22*, 1658.

- [124] E. W. Sayers, E. E. Bolton, J. R. Brister, K. Canese, J. Chan, D. C. Comeau, R. Connor, K. Funk, C. Kelly, S. Kim et al., *Nucleic Acids Research* **2022**, *50*, D20-D26.
- [125] G. A. Hudson, B. J. Burkhardt, A. J. DiCaprio, C. J. Schwalen, B. Kille, T. V. Pogorelov, D. A. Mitchell, *Journal of the American Chemical Society* **2019**, *141*, 8228.
- [126] R. C. Edgar, *Nucleic Acids Research* **2004**, *32*, 1792.
- [127] I. Letunic, "iTOL: Interactive Tree Of Life", can be found under <https://itol.embl.de/>, **2025**.
- [128] I. Letunic, P. Bork, *Nucleic Acids Research* **2024**, *52*, W78-W82.
- [129] S. R. West, A. B. Suddaby, G. R. Lewin, C. B. Ibberson, *Trends in Microbiology* **2024**, *32*, 720.
- [130] D. Sardar, E. Pierce, J. A. McIntosh, E. W. Schmidt, *ACS Synthetic Biology* **2015**, *4*, 167.
- [131] M. Noike, T. Matsui, K. Ooya, I. Sasaki, S. Ohtaki, Y. Hamano, C. Maruyama, J. Ishikawa, Y. Satoh, H. Ito et al., *Nature Chemical Biology* **2015**, *11*, 71.
- [132] R. S. Gronke, W. J. VanDusen, V. M. Garsky, J. W. Jacobs, M. K. Sardana, A. M. Stern, P. A. Friedman, *Proceedings of the National Academy of Sciences of the United States of America* **1989**, *86*, 3609.
- [133] M. Ferrer, T. N. Chernikova, K. N. Timmis, P. N. Golyshin, *Applied and Environmental Microbiology* **2004**, *70*, 4499.
- [134] M. H. Hansen, A. Keto, M. Treisman, V. M. Sasi, L. Coe, Y. Zhao, L. Padva, C. Hess, V. Leichthammer, D. L. Machell et al., *ACS Catalysis* **2024**, *14*, 812.
- [135] C. A. Bewley, H. He, D. H. Williams, D. J. Faulkner, *Journal of the American Chemical Society* **1996**, *118*, 4314.
- [136] M. Treisman, L. Coe, Y. Zhao, V. M. Sasi, J. Gullick, M. H. Hansen, A. Ly, V. Leichthammer, C. Hess, D. L. Machell et al., *Organic Letters* **2024**, *26*, 1828.
- [137] T. H. Yosca, J. Rittle, C. M. Krest, E. L. Onderko, A. Silakov, J. C. Calixto, R. K. Behan, M. T. Green, *Science* **2013**, *342*, 825.
- [138] H. E. Gering, X. Li, H. Tang, P. D. Swartz, W.-C. Chang, T. M. Makris, *Journal of the American Chemical Society* **2023**, *145*, 19256.
- [139] E. Higashide, M. Shibata, H. Yamamoto, K. Nakazawa, H. Iwasaki, J. Ueyanagi, A. Miyake, *Agricultural and Biological Chemistry* **1962**, *26*, 234.
- [140] T. W. Precord, S. Ramesh, S. R. Dommaraju, L. A. Harris, B. L. Kille, D. A. Mitchell, *ACS Bio & Med Chem Au* **2023**, *3*, 240.

- [141] C. P. Ting, M. A. Funk, S. L. Halaby, Z. Zhang, T. Gonen, W. A. van der Donk, *Science* **2019**, 365, 280.
- [142] P. N. Daniels, H. Lee, R. A. Splain, C. P. Ting, L. Zhu, X. Zhao, B. S. Moore, W. A. van der Donk, *Nature Chemistry* **2022**, 14, 71.
- [143] Y. Yu, W. A. van der Donk, *BioRxiv* **2023**, 2023.12.23.573212
- [144] J. B. Hedges, K. S. Ryan, *Chemical Reviews* **2020**, 120, 3161.
- [145] F. J. Ehinger, C. Hertweck, *Current Opinion in Chemical Biology* **2024**, 81, 102494.
- [146] K. R. Andersen, N. C. Leksa, T. U. Schwartz, *Proteins* **2013**, 81, 1857.
- [147] D. J. MacNeil, K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons, T. MacNeil, *Gene* **1992**, 111, 61.
- [148] E. Marschall, R. W. Cass, K. M. Prasad, J. D. Swarbrick, A. I. McKay, J. A. E. Payne, M. J. Cryle, J. Tailhades, *Chemical Science* **2023**, 15, 195.

## 7 Appendix

### 7.1 Appendix A

The following pages feature the article “Cytochrome P450Blt Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks” as it was published in *Angewandte Chemie International Edition* by Wiley-VCH GmbH. According to the copyright policies of Wiley-VCH GmbH, the authors of this article reserve the right to reuse the article in whole or in part within a thesis or dissertation, provided that these are not commercially published. A written consent from Wiley-VCH GmbH is not necessary for this purpose.

The article is reprinted with permission from:

Yongwei Zhao, Edward Marschall, Maxine Treisman, Leo Padva, Max Crüsemann, David R Nelson, David L Steer, Ralf B Schittenhelm, Julien Tailhades, Max J Cryle. *Angewandte Chemie International Edition* 2022, 61, e202204957.

Copyright 2021 The Authors. *Angewandte Chemie International Edition* published by Wiley-VCH GmbH.

To view the publication and supplementary information online, please follow the link below.

<https://doi.org/10.1002/anie.202204957>




**Biosynthesis Hot Paper**

 How to cite: *Angew. Chem. Int. Ed.* **2022**, 61, e202204957

International Edition: doi.org/10.1002/anie.202204957

German Edition: doi.org/10.1002/ange.202204957

# Cytochrome P450<sub>BLT</sub> Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks

Yongwei Zhao, Edward Marschall, Maxine Treisman, Alasdair McKay, Leo Padva, Max Crüsemann, David R. Nelson, David L. Steer, Ralf B. Schittenhelm, Julien Tailhades,\* and Max J. Cryle\*

**Abstract:** We report our investigation of the utility of peptide crosslinking cytochrome P450 enzymes from biaryllytite biosynthesis to generate a range of cyclic tripeptides from simple synthons. The crosslinked tripeptides produced by this P450 include both tyrosine-histidine (A–N–B) and tyrosine-tryptophan (A–O–B) crosslinked tripeptides, the latter a rare example of a phenolic crosslink to an indole moiety. Tripeptides are easily isolated following proteolytic removal of the leader peptide and can incorporate a wide range of amino acids in the residue inside the crosslinked tripeptide. Given the utility of peptide crosslinks in important natural products and the synthetic challenge that these can represent, P450 enzymes have the potential to play roles as important tools in the generation of high-value cyclic tripeptides for incorporation in synthesis, which can be yet further diversified using selective chemical techniques through specific handles contained within these tripeptides.

The crosslinking of peptides is common to many important, synthetically challenging natural products (Figure 1), including streptide,<sup>[1]</sup> the glycopeptide antibiotics (GPAs)<sup>[2]</sup> and

arylomycin,<sup>[3]</sup> where it rigidifies and stabilises peptides as well as leading to potential restricted rotation and planar chirality.<sup>[4]</sup> Within the biosynthesis pathways of crosslinked peptides, one enzyme class commonly responsible for insertion of these side chain crosslinks is the cytochrome P450s (P450s).<sup>[5]</sup> P450s can perform an array of oxidative transformations in many biosynthetic pathways via activating molecular oxygen as a highly powerful intermediate (compound I),<sup>[6]</sup> making them capable of a wide range of transformations beyond C–H hydroxylation.<sup>[5]</sup> Given this synthetic utility and combination of oxidative power and regiochemical precision, P450s have been widely implicated as potential biocatalysts.<sup>[7]</sup> Beyond the use of natural P450s, many approaches have been explored to expand the scope of P450s yet further, including the use of alternate metalated enzymes as well as techniques to alter their substrate preference, such as library shuffling, ancestral variants, directed evolution and the use of decoy substrates.<sup>[8]</sup>

Whilst the power of P450s as biocatalysts is apparent, a challenge to their application in peptide crosslinking is that many pathways in which they occur are non-ribosomal and are difficult to engineer or exploit for synthesis due to their challenging (typically enzyme bound) substrates.<sup>[9]</sup> Ribosomal (RiPP) pathways offer greater potential to identify P450 (and other)<sup>[10]</sup> enzymes as biocatalysts due to their simpler substrates,<sup>[11]</sup> although these can still be long peptides due to the large leader sequences required for cyclisation (as seen in darobactin and cittilin).<sup>[12]</sup> Given this, we focussed on the biaryllytides, a class of RiPPs that contain YxH tricyclic motifs installed by P450s.<sup>[13]</sup> Strikingly, this crosslink is found in a pentapeptide substrate, implying that the leader peptide is only two amino acids, and suggesting great biocatalytic potential for these P450s. We show that this pathway can

[\*] Y. Zhao, E. Marschall, M. Treisman, J. Tailhades, M. J. Cryle  
 Department of Biochemistry and Molecular Biology, The Monash  
 Biomedicine Discovery Institute, Monash University  
 Clayton, VIC 3800 (Australia)  
 and  
 EMBL Australia, Monash University  
 Clayton, VIC 3800 (Australia)  
 and  
 ARC Centre of Excellence for Innovations in Peptide and Protein  
 Science, Clayton, VIC 3800 (Australia)  
 E-mail: julien.tailhades@monash.edu  
 max.cryle@monash.edu

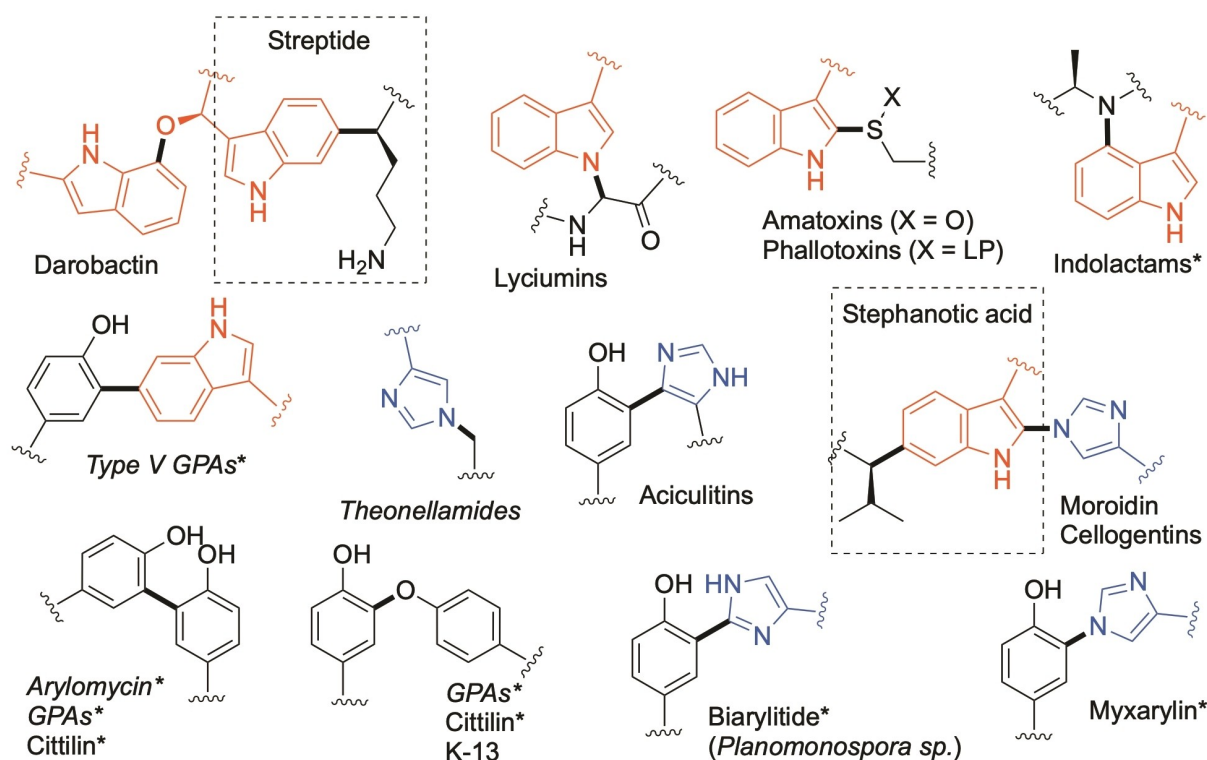
A. McKay  
 Department of Chemistry, Monash University  
 Clayton, VIC 3800 (Australia)

L. Padva, M. Crüsemann  
 Institute of Pharmaceutical Biology, University of Bonn  
 53115 Bonn (Germany)

D. R. Nelson  
 Department of Microbiology, Immunology and Biochemistry,  
 University of Tennessee  
 Memphis, TN 38163 (USA)

D. L. Steer, R. B. Schittenhelm  
 Monash Proteomics and Metabolomics Facility, Monash University  
 Clayton, VIC 3800 (Australia)

© 2022 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.



**Figure 1.** Peptide crosslinks found in RiPP and NRPS biosynthesis pathways. NRPS pathways shown in italics, crosslinks shown in bold, those installed by P450s indicated (\*). LP = lone pair.

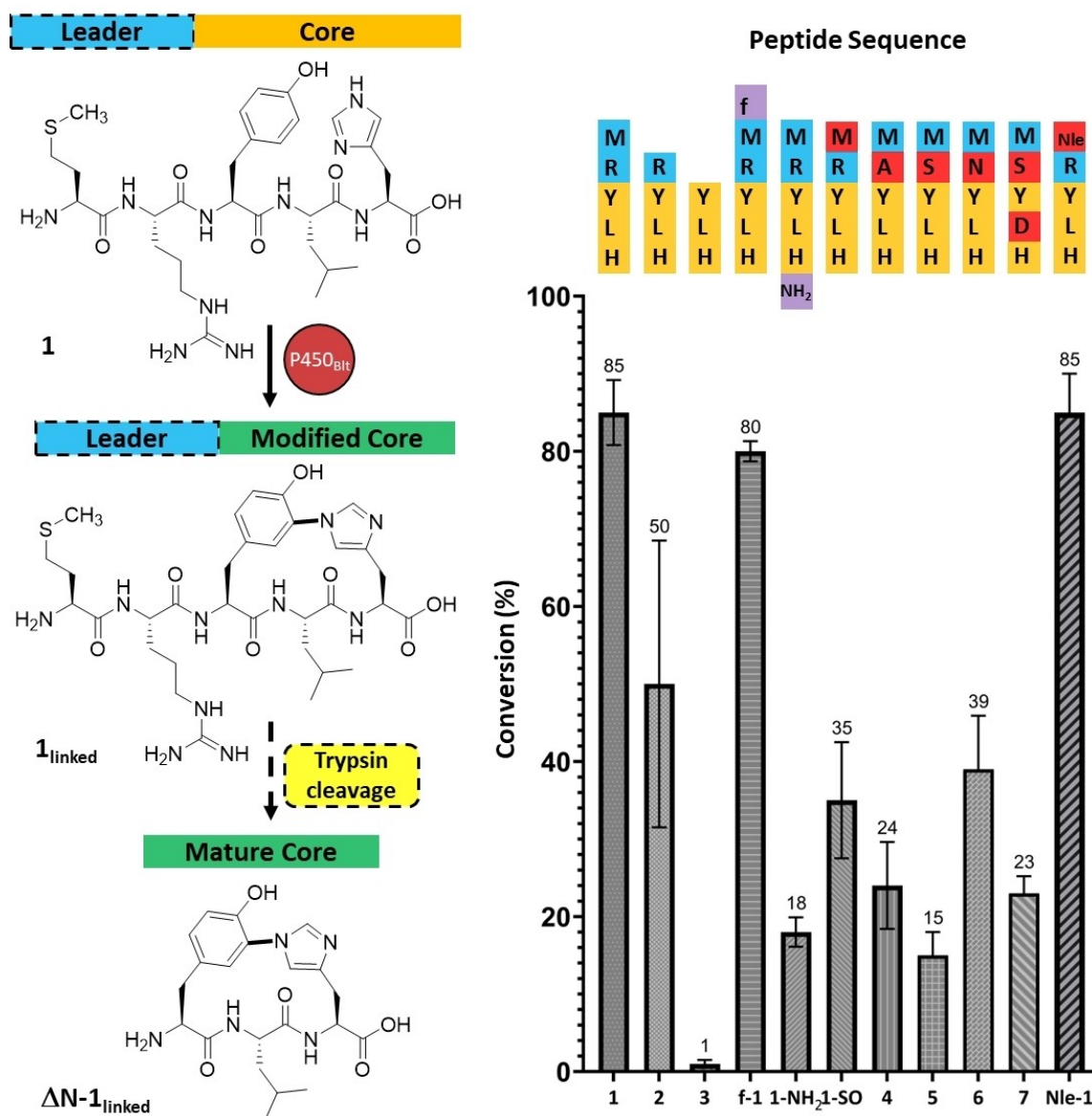
generate a range of cyclic tripeptides via P450-enabled biocatalysis and further that this allows incorporation of tryptophan in such crosslinked peptide building blocks.

To commence our study, we selected 9 P450s from a range of biarylittide pathways whose substrate sequences were MxYxH (Supporting Information Figure 1).<sup>[13]</sup> These P450s share a minimum sequence identity of 47 % and are from different families (Supporting Information Table 1, Supporting Information Figure 2). Expression of these constructs in *E. coli* demonstrated high level expression of 4, (Supporting Information Figure 3) with analysis of their catalytic competence and cyclisation activity towards

**Table 1:** Exploring the substrate tolerance of P450<sub>Blt</sub> for alterations within the directly crosslinked residues (P3/P5) of pentapeptide substrates.

Entry	N-term	Peptide sequence	C-term	Conversion [%]
17	H-	MRYLh	-OH	6 (±1)
18	H-	MRyLH	-OH	2 (±1)
19	H-	mrylh	-OH	3 (±1)
20	H-	MRHLY	-OH	2 (±1)
21	H-	MRYL-Hpg	-OH	2 (±1)
22	H-	MR-Hpg-LH	-OH	8 (±1)
23	H-	MRYL-Thz	-OH	2 (±1)
24	H-	MRYLW	-OH	55 (±5)
26	H-	MRYGW	-OH	19 (±5)
27	H-	MRYAW	-OH	21 (±3)
28	H-	MRY-Nle-W	-OH	66 (±4)

MRLYH peptide **1** leading to the selection of CYP1251 C3 (*Micromonospora* sp. MW-13) for further investigation (referred to as P450<sub>Blt</sub>). The CO complex of P450<sub>Blt</sub> showed a moderate proportion of catalytically competent enzyme (characterised by 450 nm absorption, Supporting Information Figure 4), and turnover of the **1** using the PuR/PuxB electron transport system from *Rhodospseudomonas palustris* CGA009<sup>[14]</sup> afforded >80 % conversion to the cyclic peptide under the conditions of our assay (Figure 2, Supporting Information Figure 5). We noted significant Met sulfoxidation in these assays, likely due to non-specific oxidation by side products of the P450<sub>Blt</sub> activation cycle, although this is unproblematic for generating cyclic tripeptides via tryptic digestion (see below). The binding of **1** to P450<sub>Blt</sub> displayed a relatively low spin state shift (5 %) and affinity in the low micromolar range ( $k_d = 2.1 \mu\text{M}$ , Supporting Information Figure 5), 10-fold tighter than P450 binding to related substrates (diketopiperazines, CYP121–21  $\mu\text{M}$ ;<sup>[15]</sup> CYP134A5–24  $\mu\text{M}$ )<sup>[16]</sup> and instead comparable to that seen with the peptidyl-PCP-X didomain substrates present in GPA cross-linking ( $k_d = 1.7 \mu\text{M}$ ).<sup>[17]</sup> Liberation of the cyclic tripeptide  $\Delta\text{N-1}_{\text{linked}}$  from **1**<sub>linked</sub> was performed via addition of trypsin following P450<sub>Blt</sub> assays, demonstrating a facile route to isolate cyclic tripeptide building blocks from such assays. Analysis of  $\Delta\text{N-1}_{\text{linked}}$  by NMR and the hydrogen/deuterium exchange (HDX) revealed that this crosslinked peptide contained an A–N–B crosslink (Supporting Information Figures 6–18), which is the same type of crosslink reported in myxarylin<sup>[18]</sup> but different to the A–B crosslink initially

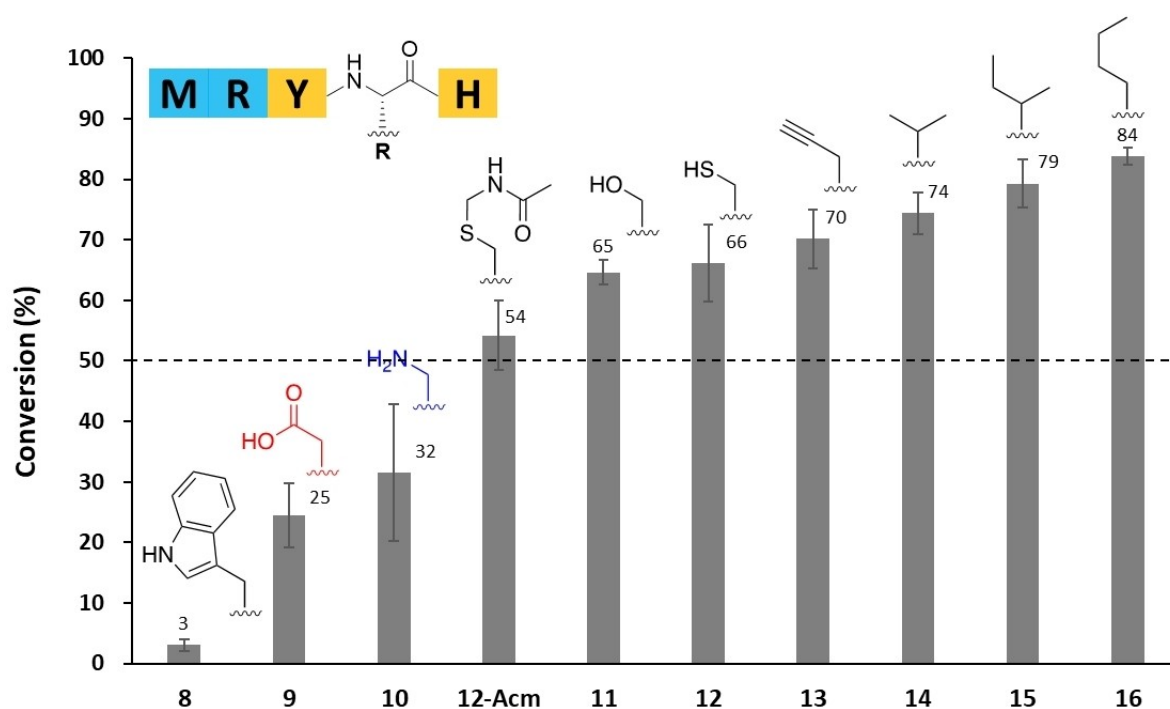


**Figure 2.** Peptide synthesis, P450-mediated cyclisation, and proteolytic cleavage to afford A–N–B linked cyclic tripeptides (left) with exploration of a range of modified biarylptide peptides (right). Reactions were conducted as described in small scale enzymatic cyclisation section (see Supporting Information). Conversions were calculated using AUC (area under curve) of the total ion current from LCMS analysis. Conversion (%) =  $(AUC_{cyclic} + AUC_{cyclic(SO)}) / (AUC_{linear} + AUC_{linear(SO)} + AUC_{cyclic} + AUC_{cyclic(SO)}) \times 100$ ; for substrates without Met at position 1,  $AUC_{linear(SO)}$  and  $AUC_{cyclic(SO)}$  equal 0. Blue: leader, yellow: core, green: cyclised core, red: altered residue, purple: modification of N/C-termini.

reported for biarylptides.<sup>[13]</sup> Given this, we re-analysed the acetylated biarylptide  $YYH_{linked}$  produced by the P450 BytO (39% sequence identity to P450<sub>Blt</sub>) in *Planomonospora*<sup>[13]</sup> and compared the spectra obtained to  $\Delta N-1_{linked}$  (Supporting Information Figure 16–17), which confirmed the difference in crosslinking pattern produced by these two related P450s and highlighting their impressive catalytic diversity.

With activity of P450<sub>Blt</sub> demonstrated towards **1**, we next analysed the effect of alterations to the peptide substrate to examine the potential synthetic utility of this P450 (Figure 2, Supporting Information Figures 19–29). First, we investigated the effect of removing the (already minimal) leader sequence, which showed that removal of Met (**2**)

significantly reduced activity and removal of Met-Arg (**3**) essentially abolished activity. This change appears due to the inability of these truncated substrates to generate the requisite spin state shift of the P450<sub>Blt</sub> heme iron (i.e., non-productive binding) rather than a loss in direct enzyme affinity. Modification of the peptide N-terminus through formylation (**f-1**) did not affect cyclisation, whilst truncation of the Met side chain (ARYLH, **4**) resulted in a major reduction of cyclisation (24%), further supporting the importance of Met for effective P450<sub>Blt</sub>-catalysed peptide cyclisation. Met sulfoxidation (**1-SO**) led to a reduction in total cyclisation by more than 50%, indicating that these interactions are likely hydrophobic in nature. As significant



**Figure 3.** Exploration of substrate tolerance around P4 of the pentapeptide. Reactions conducted as described in the Supporting Information and conversions calculated as for Figure 2. Side chains indicated for each residue.

sulfoxidation is seen in most P450<sub>Blt</sub> assays, this further suggests that peptide cyclisation is rapid, whilst sulfoxidation is slower and likely caused through oxidation by reactive oxygen species generated during the P450 active cycle. To avoid sulfoxidation, we also tested the replacement of the Met residue with Nle within the wildtype peptide sequence (**Nle-1**), which showed that this replacement is well tolerated (85 % conversion) and also simplifies purification. Replacement of the C-terminal carboxylate of **1** with an amide (**1-NH<sub>2</sub>**) revealed a 4-fold loss of P450<sub>Blt</sub> activity, attributable to a loss of two-orders of magnitude binding affinity ( $k_d = 179 \mu\text{M}$ , Supporting Information Figure 5) and suggestive of the loss of a salt bridging interaction.

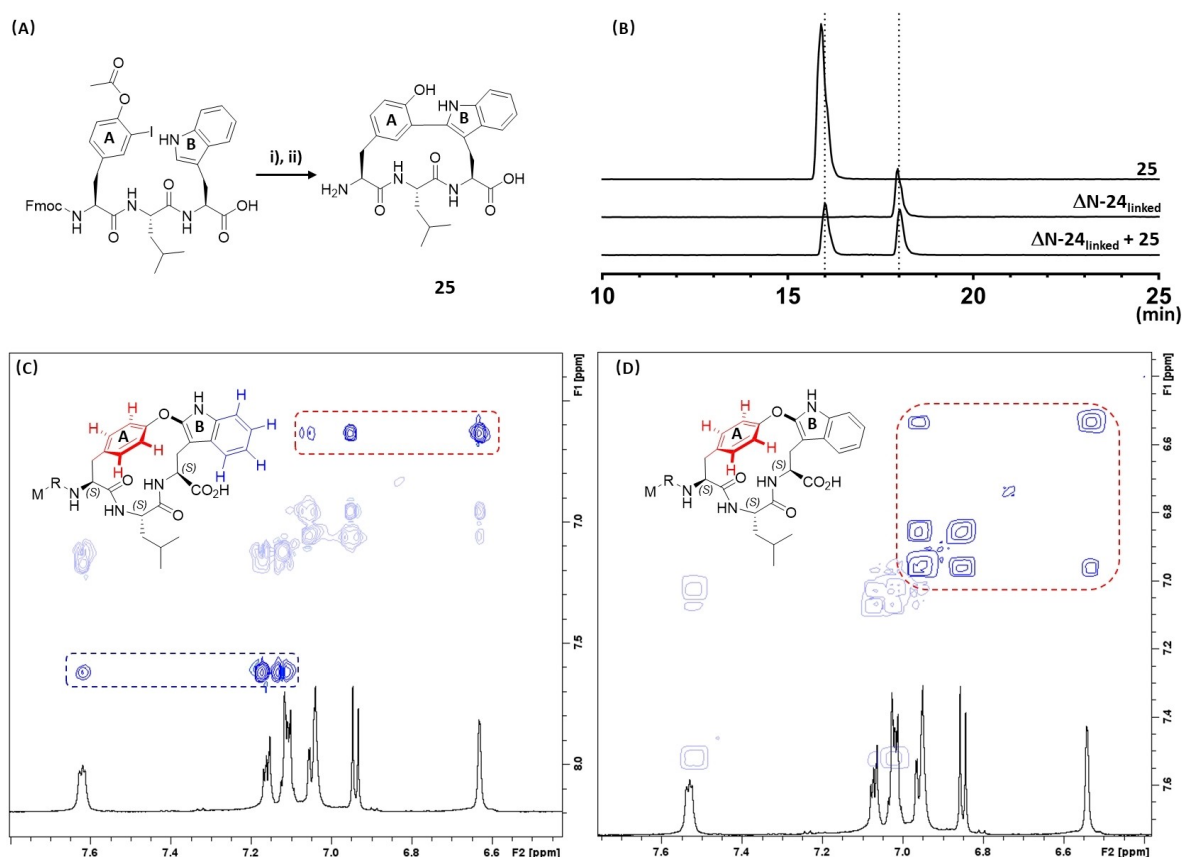
Turning to position 2 (P2), we next replaced Arg with Ser (**5**) and Asn (**6**), which both showed reduced cyclisation activity ( $\approx 6$ -fold for **5** (15 %),  $\approx 2$ -fold for **6** (39 %)). This further supports the importance of the Arg residue for peptide binding to P450<sub>Blt</sub> (Arg also being adventitious for isolation of cyclic tripeptides using trypsin), although curiously the cyclisation of the MSYDH peptide (**7**) showed no further reduction in cyclisation than for **5** alone (Figure 2, Supporting Information Figures 26–28). As a biaryllyte crosslinking P450 (43 % sequence identity to P450<sub>Blt</sub>) has recently been reported to introduce an A–N–B crosslink in a MNYLH pentapeptide,<sup>[18]</sup> we also carefully inspected the NMR data of myxarylin in comparison to that of  **$\Delta\text{N-1}_{\text{linked}}$** , with the differences observed suggesting that this altered A–N–B crosslinking pattern may be specific to each individual P450.

Given that the most interesting position outside of the crosslinking positions themselves is arguably P4 of the

pentapeptide, we next studied the acceptance of a range of peptides with altered residues at P4 by P450<sub>Blt</sub> (Figure 3, Supporting Information Figures 30–39). We tested a range of peptides exploring the effect of side chain size (Val (**14**), Ile (**15**), Nle (**16**), Trp (**8**)) and charge (Asp (**9**)) on P450<sub>Blt</sub>-catalysed cyclisation. These experiments showed that there is considerable tolerance for alterations in the size of P4, although charge appears problematic for P450<sub>Blt</sub>. Curiously, whilst cyclisation of **14–16** showed no change in cyclisation compared to Leu (**1**), incorporation of Trp (**8**) almost totally abolished activity, showing that there is a limit to the size of residue that can be included inside the crosslink of the current (biosynthetic, not engineered) enzyme. The reduced level of cyclisation seen for **9** further revealed the “pocket” for P4 is likely hydrophobic in nature, although it is important note that P450s accepting peptides containing Trp<sub>4</sub> and Asp<sub>4</sub> residues have been reported.<sup>[13]</sup> Thus, future access to a peptide bound structure of P450<sub>Blt</sub> would be highly informative when combined with sequence data and could well enable the engineering of P450<sub>Blt</sub> to support activity towards such peptides.

Given the potential utility of an enzymatic route to a range of cyclic tripeptide building blocks, we further explored modifications of P4, concentrating on residues whose side chains contained moieties of use for chemical diversification (**10–13**, Figure 3, Supporting Information Figures 32–36). Peptides tested included those with an alkyne-containing Pra residue (**13**), alcohol containing Ser (**11**), thiol-containing Cys (both unprotected **12** and Acm-protected **12-Acm**) and amine-containing Dap (**10**) at P4. Peptides **11–13** were well accepted by P450<sub>Blt</sub>, showing an impressive degree





**Figure 4.** A) Synthesis of A–B crosslinked YLW standard **25**. i) 5 % mol % Pd(OAc)<sub>2</sub>, 1.0 equiv of AgBF<sub>4</sub>, 1.5 equiv of 2-NO<sub>2</sub>BzOH in DMF: PBS (1:1), 80 °C (microwave), 12 h. ii) 0.5 M NaOH, 20 min. B) TIC analysis of enzymatically cyclised YLW peptide  $\Delta$ N-**24**<sub>linked</sub> and the synthetic A–B crosslinked YLW standard **25**. C) Superposition of <sup>1</sup>H-spectrum over a <sup>1</sup>H-<sup>1</sup>H total correlation spectra (TOCSY) for **24**<sub>linked</sub>, showing a zoomed view of the aromatic protons together with relevant interactions highlighted with red and blue; relevant cross peaks are boxed using dotted line with the corresponding colour. D) Superposition of <sup>1</sup>H-spectrum over a <sup>1</sup>H-<sup>1</sup>H correlation spectrum (COSY) for **24**<sub>linked</sub>, showing a zoomed view of the aromatic protons together with relevant interactions highlighted with red; relevant cross peaks are also boxed using dotted line.

of cyclisation despite the diversity of side chains present in these peptides. Even the Dap-containing peptide **10** was a moderate substrate; given that tryptic digestion allows for simple isolation of cyclic tripeptides, even a 30 % conversion of **10** allows this to remain a possible route to access such a range of cyclic synthons.

Next, we investigated the acceptance of peptides containing different residues within the site of the crosslink (Table 1, Supporting Information Figures 40–46). In these experiments, we explored the tolerance of P450<sub>Blt</sub> for altering stereochemistry within the peptide substrate (**17**–**19**), switching the positions of the Tyr/His residues (**20**) and replacing them with a more rigid 4-hydroxyphenylglycine (Hpg) residue (**21**–**22**), as well as ascertaining the importance of the NH moiety in these residues by replacement of the P5 imidazole side chain with a thiazoline moiety (**23**). As was perhaps expected for such a biosynthetic enzyme, peptides **17**–**22** were not well accepted—the highest conversions seen for **17** (D-His<sub>5</sub> residue) and **22** (Hpg<sub>3</sub> residue) were >10 % than that of **1**. This likely stems from the low affinity of P450<sub>Blt</sub> for these modified peptides, which is seen in the loss of affinity determined for **17** ( $k_d$  = 93 μM) and **18** ( $k_d$  = 530 μM) compared to **1** ( $k_d$  = 2.1 μM). Despite these

low yields, the natural occurrence of a biaryllyte A–(X)–B linked peptide with the sequence MRHEY shows that altering the ring substituents is possible,<sup>[13]</sup> further supporting the crucial importance of obtaining structural information of member/s of this versatile peptide crosslinking class of P450s. Turnover of **23** also supported the site of crosslinking present in the P450<sub>Blt</sub> reaction, as **23** is unable to be cyclised by P450<sub>Blt</sub> despite it being able to bind to the enzyme ( $k_d$  = 59 μM).

Finally, we tested the exchange of the His<sub>5</sub> residue for Trp, given that such crosslinks are widely reported in cyclic peptides (see Figure 1). P450<sub>Blt</sub>-catalysed cyclisation of MRYLW (**24**) was very effective, showing 2/3<sup>rd</sup>s of the level of cyclisation of **1** despite the alteration of a crosslinking residue and the large increase in size of this residue (Supporting Information Figure 47). Such differences are seen in the spectral response of P450<sub>Blt</sub> upon binding of **24**, where the spectrum no longer resembles the activation spectra seen for **1** and displays a slight reduction in affinity ( $k_d$  = 4.2 μM, Supporting Information Figure 5). We further explored the tolerance of P450<sub>Blt</sub> in accepting alterations in the MRYxW peptide (**26**–**28**) and observed a similar trend as had been seen for MRYxH peptides with higher activity

seen with larger, hydrophobic substrates (19 % MRYGW  $\approx$  21 % MRYAW < 66 % MRY-Nle-W; Supporting Information Figures 48–50). Cyclisation of **24** by P450<sub>Blt</sub> raised the important question of the nature of the crosslink in this peptide given that the type of linkage could affect the ring size of the crosslinked peptide. To address this, we synthesised a cyclic YLW tripeptide standard (**25**) containing an A–B linkage (Figure 4)<sup>[19]</sup> and compared the retention time of this standard to the product formed by P450<sub>Blt</sub>-catalysed cyclisation of **24** with subsequent tryptic digestion ( $\Delta$ N-**24**<sub>linked</sub>). Having seen a different retention time in this analysis, we next performed extensive NMR characterisation of both **24**<sub>linked</sub> and **25** to understand the nature of the crosslink in **24**<sub>linked</sub> (Supporting Information Figure 51–60). These analyses revealed that—in addition to not containing an aryl A–B crosslink—**24**<sub>linked</sub> retained all four Tyr aromatic protons, implying the crosslink is not through the Trp indole nitrogen to the Tyr ring and it does not resemble the crosslinking seen in pseudosporamide.<sup>[20]</sup> This led to the conclusion that the crosslink installed in this case is an A–O–B crosslink to the Trp indole ring through the Tyr phenol oxygen. HDX analysis of  $\Delta$ N-**24**<sub>linked</sub> and **25** further supported a heteroatom containing crosslink for  $\Delta$ N-**24**<sub>linked</sub> (Supporting Information Figure 18). The ability of P450<sub>Blt</sub> to generate **24**<sub>linked</sub> shows that there is significant plasticity in P450<sub>Blt</sub> for the generation of alternate rings linkages in tripeptides and supports the potential of this P450 class as a general tool for the generation of a range of cyclic tripeptide building blocks containing A–B,<sup>[13]</sup> A–N–B<sup>[18]</sup> and A–O–B crosslinks.

In summary, we have performed a detailed analysis of the substrate tolerance of the biaryllyte peptide crosslinking P450<sub>Blt</sub>. We have shown that the minimal leader sequence of these pentapeptides remains important for catalysis, although this can be removed through proteolysis to allow potential access to A–N–B crosslinked cyclic tripeptides. Furthermore, we have demonstrated that generating various crosslinked tripeptides is highly feasible. Perhaps most impressive is the ability of this enzyme to install a crosslink in which the natural His residue is replaced by Trp. This, combined with altered crosslinking and substrates reported for other members of this biosynthetic pathway demonstrates that biaryllyte biosynthetic P450s could well play important roles as future biocatalysts for the generation of a diverse range of cyclic tripeptide building blocks. In this endeavour, a structure of the substrate-bound P450 remains a priority for the field to allow rational engineering of these versatile enzymes and to address unanswered mechanistic questions concerning P450-mediated peptide crosslinking, an important biosynthetic process central to the generation of many crosslinked peptide natural products.

## Acknowledgements

We thank Prof. Hutton (Uni Melbourne) and Prof. De Voss (UQ) for helpful discussions. This work was supported by Monash University, EMBL Australia and was conducted by the Australian Research Council Centre of Excellence for

Innovations in Peptide and Protein Science (CE200100012) funded by the Australian Government. This study used BPA-enabled (Bioplatforms Australia)/NCRIS-enabled (National Collaborative Research Infrastructure Strategy) infrastructure located at the Monash Proteomics and Metabolomics Facility. Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available in the Supporting Information of this article.

**Keywords:** Amino Acids · Biocatalysis · Cytochrome P450 · Metalloenzymes · Peptide Crosslinking · Peptide Cyclisation · Peptides

- [1] K. R. Schramma, L. B. Bushin, M. R. Seyedsayamdost, *Nat. Chem.* **2015**, *7*, 431–437.
- [2] a) K. Haslinger, M. Peschke, C. Brieke, E. Maximowitsch, M. J. Cryle, *Nature* **2015**, *521*, 105–109; b) E. J. Culp, N. Waglechner, W. Wang, A. A. Fiebig-Comyn, Y.-P. Hsu, K. Koteva, D. Sychantha, B. K. Coombes, M. S. Van Nieuwenhze, Y. V. Brun, G. D. Wright, *Nature* **2020**, *578*, 582–587.
- [3] H. Aldemir, S. Shu, F. Schaefer, H. Hong, R. Richarz, S. Harteis, M. Einsiedler, T. M. Milzarek, S. Schneider, T. A. M. Gulder, *Chem. Eur. J.* **2022**, *28*, e202103389.
- [4] a) R. López, C. Palomo, *Angew. Chem. Int. Ed.* **2022**, *61*, e202113504; *Angew. Chem.* **2022**, *134*, e202113504; b) S. H. Reisberg, Y. Gao, A. S. Walker, E. J. N. Helfrich, J. Clardy, P. S. Baran, *Science* **2020**, *367*, 458–463.
- [5] a) A. Greule, J. E. Stok, J. J. De Voss, M. J. Cryle, *Nat. Prod. Rep.* **2018**, *35*, 757–791; b) F. P. Guengerich, A. W. Munro, *J. Biol. Chem.* **2013**, *288*, 17065–17073.
- [6] J. Rittle, M. T. Green, *Science* **2010**, *330*, 933–937.
- [7] a) G. Di Nardo, G. Gilardi, *Trends Biochem. Sci.* **2020**, *45*, 511–525; b) E. O'Reilly, V. Köhler, S. L. Flitsch, N. J. Turner, *Chem. Commun.* **2011**, *47*, 2490–2501; c) L. E. Zetzsche, J. A. Yazarians, S. Chakrabarty, M. E. Hinze, L. A. M. Murray, A. L. Lukowski, L. A. Joyce, A. R. H. Narayan, *Nature* **2022**, *603*, 79–85.
- [8] a) Y. Gumulya, J.-M. Baek, S.-J. Wun, R. E. S. Thomson, K. L. Harris, D. J. B. Hunter, J. B. Y. H. Behrendorff, J. Kulig, S. Zheng, X. Wu, B. Wu, J. E. Stok, J. J. De Voss, G. Schenk, U. Jurva, S. Andersson, E. M. Isin, M. Bodén, L. Guddat, E. M. J. Gillam, *Nat. Catal.* **2018**, *1*, 878–888; b) S. Ariyasu, J. K. Stanfield, Y. Aiba, O. Shoji, *Curr. Opin. Chem. Biol.* **2020**, *59*, 155–163; c) M. Karasawa, K. Yonemura, J. K. Stanfield, K. Suzuki, O. Shoji, *Angew. Chem. Int. Ed.* **2022**, *61*, e202111612; *Angew. Chem.* **2022**, *134*, e202111612; d) K. Zhang, B. M. Shafer, M. D. Demars, H. A. Stern, R. Fasan, *J. Am. Chem. Soc.* **2012**, *134*, 18695–18704; e) S. Li, M. R. Chaulagain, A. R. Knauff, L. M. Podust, J. Montgomery, D. H. Sherman, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18463–18468; f) H. Joo, Z. Lin,

- F. H. Arnold, *Nature* **1999**, 399, 670–673; g) A. Li, C. G. Acevedo-Rocha, L. D'Amore, J. Chen, Y. Peng, M. Garcia-Borràs, C. Gao, J. Zhu, H. Rickerby, S. Osuna, J. Zhou, M. T. Reetz, *Angew. Chem. Int. Ed.* **2020**, 59, 12499–12505; *Angew. Chem.* **2020**, 132, 12599–12605.
- [9] R. D. Süssmuth, A. Mainz, *Angew. Chem. Int. Ed.* **2017**, 56, 3770–3821; *Angew. Chem.* **2017**, 129, 3824–3878.
- [10] a) K. Sogahata, T. Ozaki, Y. Igarashi, Y. Naganuma, C. Liu, A. Minami, H. Oikawa, *Angew. Chem. Int. Ed.* **2021**, 60, 25729–25734; *Angew. Chem.* **2021**, 133, 25933–25938; b) T. Q. N. Nguyen, Y. W. Tooh, R. Sugiyama, T. P. D. Nguyen, M. Purushothaman, L. C. Leow, K. Hanif, R. H. S. Yong, I. Agatha, F. R. Winnerdy, M. Gugger, A. T. Phan, B. I. Morinaka, *Nat. Chem.* **2020**, 12, 1042–1053; c) S. Ma, H. Chen, H. Li, X. Ji, Z. Deng, W. Ding, Q. Zhang, *Angew. Chem. Int. Ed.* **2021**, 60, 19957–19964; *Angew. Chem.* **2021**, 133, 20110–20117.
- [11] C. Wu, W. A. van der Donk, *Curr. Opin. Biotechnol.* **2021**, 69, 221–231.
- [12] a) Y. Imai, K. J. Meyer, A. Iinishi, Q. Favre-Godal, R. Green, S. Manuse, M. Caboni, M. Mori, S. Niles, M. Ghiglieri, C. Honrao, X. Ma, J. J. Guo, A. Makriyannis, L. Linares-Otaya, N. Böhringer, Z. G. Wuisan, H. Kaur, R. Wu, A. Mateus, A. Typas, M. M. Savitski, J. L. Espinoza, A. O'Rourke, K. E. Nelson, S. Hiller, N. Noinaj, T. F. Schäberle, A. D'Onofrio, K. Lewis, *Nature* **2019**, 576, 459–464; b) J. J. Hug, J. Dastbaz, S. Adam, O. Revermann, J. Koehnke, D. Krug, R. Müller, *ACS Chem. Biol.* **2020**, 15, 2221–2231.
- [13] M. M. Zdouc, M. M. Alanjary, G. S. Zarazúa, S. I. Maffioli, M. Crüsemann, M. H. Medema, S. Donadio, M. Sosio, *Cell Chem. Biol.* **2021**, 28, 733–739.e734.
- [14] S. G. Bell, F. Xu, E. O. D. Johnson, I. M. Forward, M. Bartlam, Z. Rao, L.-L. Wong, *J. Biol. Inorg. Chem.* **2010**, 15, 315–328.
- [15] P. Belin, M. H. Le Du, A. Fielding, O. Lequin, M. Jacquet, J.-B. Charbonnier, A. Lecoq, R. Thai, M. Courçon, C. Masson, C. Dugave, R. Genet, J.-L. Pernodet, M. Gondry, *Proc. Natl. Acad. Sci. USA* **2009**, 106, 7426–7431.
- [16] M. J. Cryle, S. G. Bell, I. Schlichting, *Biochemistry* **2010**, 49, 7282–7296.
- [17] M. Peschke, K. Haslinger, C. Brieke, J. Reinstein, M. Cryle, *J. Am. Chem. Soc.* **2016**, 138, 6746–6753.
- [18] J. J. Hug, N. A. Frank, C. Walt, P. Šenica, F. Panter, R. Müller, *Molecules* **2021**, 26, 7483.
- [19] L. Mendive-Tapia, S. Preciado, J. García, R. Ramón, N. Kielland, F. Albericio, R. Lavilla, *Nat. Commun.* **2015**, 6, 7160.
- [20] S. Saito, K. Atsumi, T. Zhou, K. Fukaya, D. Urabe, N. Oku, M. R. U. Karim, H. Komaki, Y. Igarashi, *Beilstein J. Org. Chem.* **2020**, 16, 1100–1110.

Manuscript received: April 4, 2022

Accepted manuscript online: July 18, 2022

Version of record online: August 3, 2022

## 7.2 Appendix B

The following pages feature the article “Reassignment of the Structure of a Tryptophan-Containing Cyclic Tripeptide Produced by the Biaryllytide Crosslinking Cytochrome P450blt” as it was published *Chemistry – A European Journal* by Wiley-VCH GmbH. According to the copyright policies of Wiley-VCH GmbH, the authors of this article reserve the right to reuse the article in whole or in part within a thesis or dissertation, provided that these are not commercially published. A written consent from Wiley-VCH GmbH is not necessary for this purpose.

The article is reprinted with permission from:

Laura J Coe, Yongwei Zhao, Leo Padva, Angus Keto, Ralf Schittenhelm, Julien Tailhades, Greg Pierens, Elizabeth H Krenske, Max Crüsemann, James J De Voss, Max J Cryle. *Chemistry – A European Journal* 2024, 30, e202400988.

Copyright 2024 The Authors. *Chemistry – A European Journal* published by Wiley-VCH GmbH.

To view the publication and supplementary information online, please follow the link below.

<https://doi.org/10.1002/chem.202400988>



# Reassignment of the Structure of a Tryptophan-Containing Cyclic Tripeptide Produced by the Biaryllytite Crosslinking Cytochrome P450<sub>blt</sub>

Laura J. Coe,<sup>[a]</sup> Yongwei Zhao<sup>+, [b, c, d]</sup> Leo Padva<sup>+, [e]</sup> Angus Keto<sup>+, [a]</sup> Ralf Schittenhelm,<sup>[f]</sup> Julien Tailhades,<sup>[b, c, d]</sup> Greg Pierens,<sup>[g]</sup> Elizabeth H. Krenske,<sup>[a]</sup> Max Crüsemann,<sup>[e]</sup> James J. De Voss,<sup>\*, [a, d]</sup> and Max J. Cryle<sup>\*, [b, c, d]</sup>

The structure of the sidechain crosslinked Tyr-Leu-Trp peptide produced by the biaryllytite crosslinking cytochrome P450<sub>blt</sub> from *Micromonospora* sp. MW-13 has been reanalysed by a series of NMR, computational and isotope labelling experiments and shown to contain a C–N rather than a C–O bond. Additional *in vivo* experiments using such a modified peptide show there is a general tolerance of biaryllytite crosslinking

P450 enzymes for histidine to tryptophan mutations within their minimal peptide substrate sequences despite the lack of such residues noted in natural biaryllytite gene clusters. This work further highlights the impressive ability of P450s from biaryllytite biosynthesis pathways to act as biocatalysts for the formation of a range of sidechain crosslinked tripeptides.

## Introduction

The crosslinking of peptide natural products is an important process that is responsible both for the activity of these molecules and often their improved pharmacokinetic stability.<sup>[1]</sup> Whilst many peptides are cyclised via amide/ester bonds (through their N-termini, C-termini and/or side chain residues) or disulfide bonds (via cysteine residues), an increasingly large

number of peptide natural products have been identified that involve the crosslinking of side chain residues. Whilst well known examples of side chain crosslinked peptides are often found in the products of nonribosomal biosynthesis pathways,<sup>[2]</sup> there has been a recent and dramatic increase in the reports of such peptides from ribosomal pathways (RiPPs).<sup>[3]</sup> Such ribosomal pathways contain precursor peptides composed of both leader and core sequences, with the leader sequence responsible for recruitment of modifying enzymes that then act upon the core portion of the peptide.

Within the broader family of RiPPs, many enzyme classes are implicated in the introduction of peptide crosslinks,<sup>[3g–j]</sup> although as a family the Cytochromes P450 (P450s)<sup>[4]</sup> have been shown to be one of the most common involved in sidechain crosslinking. This is particularly the case with amino acids containing aromatic side chains.<sup>[3a–f,k,5]</sup> Recent bioinformatic studies have revealed how earlier reports of the citrilins,<sup>[6]</sup> the biaryllytides/myxarylin<sup>[7]</sup> and the tryptorubins<sup>[3k]</sup> have now expanded to twenty different classes of peptides crosslinked by P450s, of which many have only recently been characterised (Figure 1).<sup>[3a,b,f]</sup>

Within such P450-mediated crosslinking pathways, the most prevalent residues involved in crosslinks are tyrosine and tryptophan; histidine is less common and is found within only three of the currently identified P450-crosslinked RiPP groups (Figure 1). Two of these classes are biaryllytite type RiPPs,<sup>[5a,7]</sup> which themselves are fascinating examples of such peptides that only require a minimal (2 amino acid) leader sequence within a total peptide length of five amino acids. This stands in contrast to the other types of P450-crosslinked RiPPs that require significantly longer leader peptide sequences.<sup>[3a,b,f]</sup>

Characterisation of P450s from the biaryllytite/myxarylin class of RiPPs have shown that the enzymes within these minimal biosynthetic gene clusters can be versatile biocatalysts for the generation of cyclic YxH tripeptides (1), including the

[a] L. J. Coe, A. Keto, E. H. Krenske, J. J. De Voss

School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia  
E-mail: j.devoss@uq.edu.au

[b] Y. Zhao, J. Tailhades, M. J. Cryle

Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, VIC 3800, Australia  
E-mail: max.cryle@monash.edu

[c] Y. Zhao, J. Tailhades, M. J. Cryle

EMBL Australia, Monash University, Clayton, VIC 3800, Australia

[d] Y. Zhao, J. Tailhades, J. J. De Voss, M. J. Cryle

ARC Centre of Excellence for Innovations in Peptide and Protein Science, Australia

[e] L. Padva, M. Crüsemann

Institute of Pharmaceutical Biology, University of Bonn, 53115 Bonn, Germany


[f] R. Schittenhelm


Monash Proteomics and Metabolomics Platform, Monash University, Clayton, VIC 3800, Australia

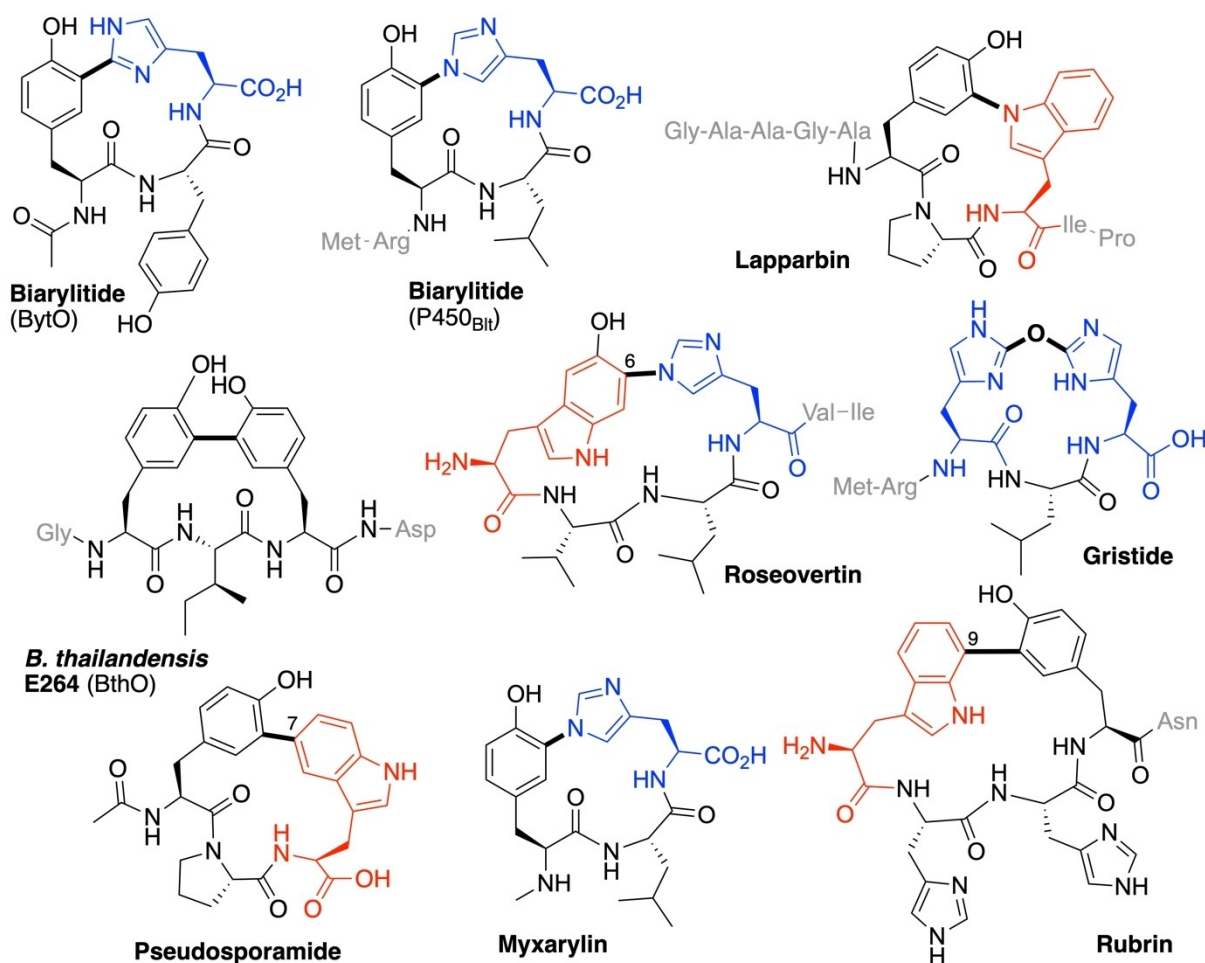
[g] G. Pierens

Centre for Advanced Imaging, The University of Queensland, Brisbane, QLD 4072, Australia

[\*] Authors contributed equally

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/chem.202400988>

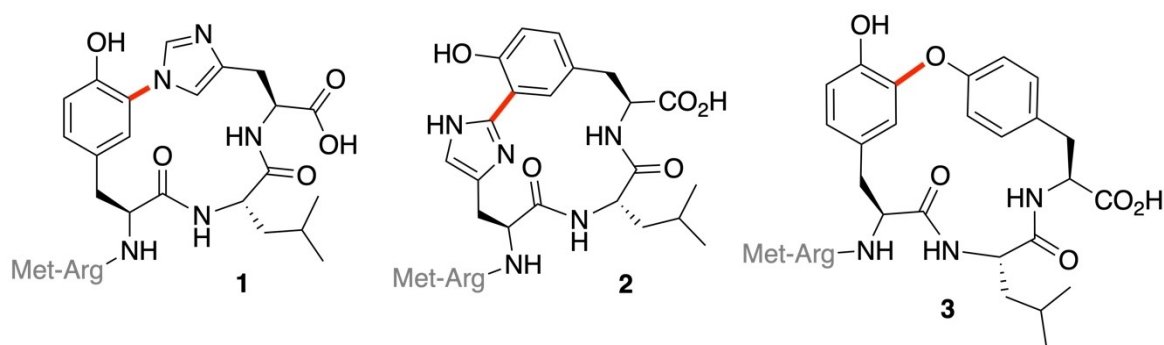
 © 2024 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.



**Figure 1.** Examples of cytochrome P450-mediated crosslinking within RiPP biosynthesis. Histidine residues are labelled in blue and tryptophan residues are labelled in red. Additional residues outside of the crosslinked core of these peptides are shown in 3-letter amino acid code and coloured grey.

generation of different crosslinks (C–C and C–N) between the Tyr and His residues by the P450 enzymes BytO from *Planomonospora* sp. and P450<sub>Blt</sub> from *Micromonospora* sp. MW-13 respectively (Figure 1).<sup>[5a,7–8]</sup> More recently, the structure-enabled engineering<sup>[8]</sup> of the active site of P450<sub>Blt</sub> has demonstrated that this enzyme is further capable of installing a C–C

crosslink within a reversed peptide sequence (2, HxY) and is also able to crosslink YxY peptides (3) in a manner that generates selective formation of one crosslinked product (Figure 2).<sup>[9]</sup> Our original investigations into P450<sub>Blt</sub>-mediated peptide crosslinking had revealed that this enzyme was capable of crosslinking MRYxW peptides,<sup>[5a]</sup> a discovery that was somewhat



**Figure 2.** Structures of crosslinked peptides produced by P450<sub>Blt</sub> (1) together with those produced by active site mutants (2–3) showing wildtype (1, YxH), reversed (2, HxY) and alternate (3, YxY) crosslinking patterns that are enabled by this versatile biaryllytide P450 enzyme. Additional residues outside of the crosslinked core of these peptides are shown in 3-letter amino acid code and coloured grey.

surprising given the size of the Trp sidechain and the observation that no biarylittides encode a Trp residue within a typical biarylittide crosslink.<sup>[3b,7b]</sup>

Here, we report our re-evaluation of the structure of the Tyr-X-Trp crosslink produced by P450<sub>Blt</sub>, which was inspired by our recent studies on the reversed His-X-Tyr peptide by an engineered variant of this P450. Our results show that this peptide is crosslinked via the indole nitrogen of the Trp residue to the *m*-position of Tyr (4, Figure 3), which agrees with high resolution NMR, computational and isotope labelling experiments and matches the results of other recent studies concerning the structure of P450-crosslinked RiPPs bearing Trp residues.<sup>[3a,b]</sup>

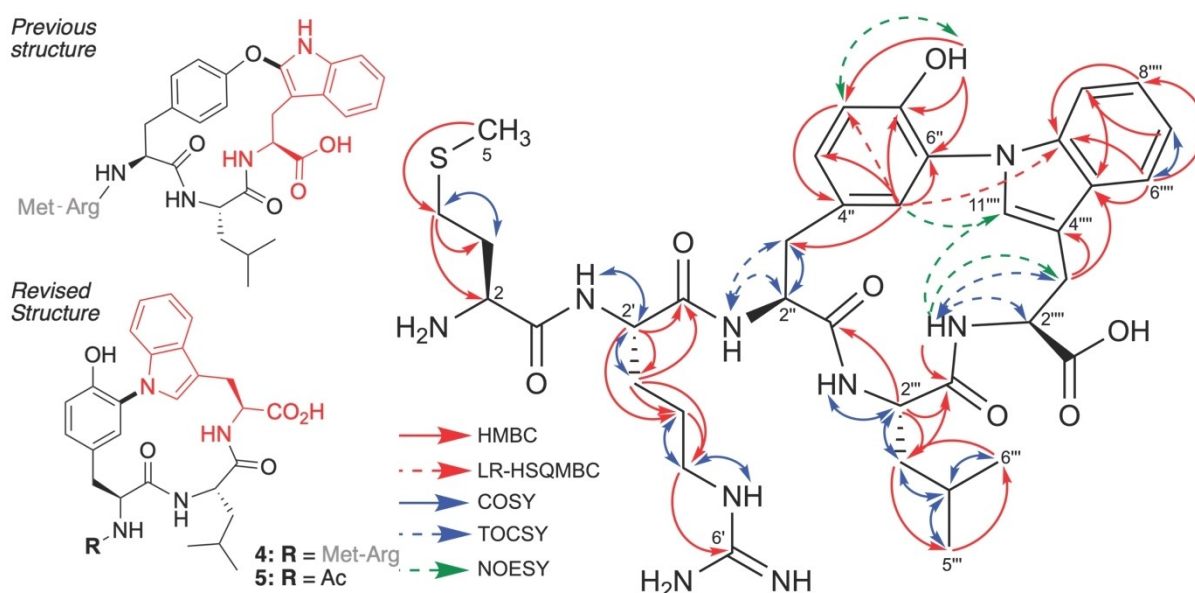
## Results and Discussion

Our ongoing investigations of the catalytic potential of P450<sub>Blt</sub> has given insight into the wide variety of aromatic crosslinks that this enzyme can install in peptides. We recently reported TyrC-TyrO, TyrC-HisN and HisC-TyrC crosslinks that required extensive structure elucidation.<sup>[8–9]</sup> (Figure 2). Insights from this work, particularly the importance of exchangeable proton analysis in the NMR, prompted re-examination of our previously reported P450<sub>Blt</sub> cyclised MRYLW structure (Figure 3).<sup>[5a]</sup> A hydrogen-deuterium exchange (HDX) analysis indicated heteroatom involvement in the crosslink formation. This, along with the challenges in determining the site of crosslinking in such heteroatom-containing systems, prompted us to undertake a re-analysis of the structure of this crosslinked peptide. The cyclised MRYLW peptide was carefully re-examined using high-field NMR spectroscopy in DMSO-*d*<sub>6</sub> solvent. The change to a polar aprotic solvent preserved the exchangeable OH and NH

proton signals and pleasingly also greatly simplified the spectra when compared to those obtained in methanol-*d*<sub>4</sub>.<sup>[5a]</sup> As was the case in our prior studies, this was particularly evident in the crucial aromatic region. Importantly, no evidence of multiple slowly-interconverting conformers was observed in the spectra of the crosslinked MRYLW (see SI for further details).<sup>[5a]</sup> This has previously been reported in such cyclised peptides and can complicate interpretation.<sup>[3b]</sup>

Initial examination of the <sup>1</sup>H NMR and TOCSY spectra revealed the presence of four  $\alpha$ -protons at  $\delta_{\text{H}}$  3.78, 4.39, 4.57 and 4.74 and their corresponding alpha carbons were detected by <sup>13</sup>C-HSQC ( $\delta_{\text{C}}$  52.0, 52.5, 53.2, 50.6 respectively). The alpha proton of the C-terminal Trp was identified by TOCSY as a very broad peak at  $\delta_{\text{H}}$  4.48 (the corresponding carbon resonance was not observed). The expected four amide protons at  $\delta_{\text{H}}$  8.57, 7.70, 8.53 and 8.86 were identified and assigned by COSY and TOCSY correlations. One Arg guanidino NH was identified ( $\delta_{\text{H}}$  7.58, br t, *J*=5). Four carbonyl signals were identified at  $\delta_{\text{C}}$  172.9, 170.3, 169.0, and 171.9 by analysis of the HMBC, HSQC and <sup>13</sup>C spectra. The carbonyl signal belonging to the C-terminal Trp residue was not observed. The expected Met-Arg-Tyr-Leu-Trp pentapeptide aliphatic side chain signals were then fully assigned using 1D <sup>1</sup>H and 2D COSY, TOCSY, <sup>13</sup>C-HSQC, HMBC, and NOESY (Table 1). <sup>15</sup>N-HSQC experiments allowed assignment of some nitrogen resonances (Table 1).

Signals associated with the aromatic residues were then carefully examined to determine the structure of the installed crosslink. The relevant region of the <sup>1</sup>H NMR spectrum was complex with multiple overlapping CH and NH resonances. Seven aromatic CH signal crosspeaks were readily identified in the <sup>13</sup>C-HSQC spectrum. An eighth crosspeak from a broad singlet proton at  $\delta_{\text{H}}$  6.96 presented more weakly and was correlated to  $\delta_{\text{C}}$  128.3 (Figure 4). This  $\delta_{\text{H}}$  6.96 proton signal



**Figure 3.** Structures of YxW crosslinked peptides produced by P450<sub>Blt</sub> as originally reported (upper left) and as have been reassigned here (4, lower left) together with the NMR correlations that support the reassignment (right). The structure of the acetylated tripeptide standard 5 prepared as a standard for *in vivo* analysis (*vide infra*) is also shown (lower left). The correlations from specific NMR experiments are colour coded based on the key within the figure. Additional residues contained in the biarylittide peptides but outside of the crosslinked core are shown in 3-letter amino acid code and coloured grey.

**Table 1.** Complete NMR assignments for crosslinked MRYLW peptide 4.

	DMSO- $d_6$ 700 MHz $\delta_H$ (J)	176 MHz $\delta_C$	70 MHz $\delta_N$ <sup>6</sup>
1	–	172.9 <sup>1</sup>	
2	3.78 (br)	52.0	
3	1.95–2.00 (m, br) 1.91–1.95 (m, br)	31.8	
4	2.54–2.57 (m, ov)	28.4	
5	2.08 (s)	14.6	
1'	–	170.3	
2'	4.39 (q, 7.0)	52.5	
2'-NH	8.57 (s, br)		120.9
3'	1.63–1.69 (m, br) 1.49–1.56 (m, ov)	29.3	
4'	1.42–1.48 (m, ov) 1.49–1.55 (m, ov)	25.0	
5'	3.09 (q, 7.0)	40.4	
5'-NH	7.58 (t, br, 5)		84.8
6'	–	156.7	
1''	–	169.0 <sup>4</sup>	
2''	4.57 (br)	53.2	
2''-NH	7.70 (br) <sup>2</sup>		Not obs.
3''	2.92–2.98 (m, br), 2.78–2.84 (m, br)	36.1	
4''	–	127.0	
5''	6.58 (d, 1.9)	129.8 <sup>5</sup>	
6''	–	126.8	
7''	–	150.2	
7''-OH	9.87 (s)	–	
8''	6.90–6.93 (ov)	116.2	
9''	6.90–6.93 (ov)	129.4	
1'''	–	171.9	
2'''	4.74 (q, 8.1)	50.6	
2'''-NH	8.53 (s, br)		Not obs.
3'''	1.45–1.49 (m, ov) 1.41–1.44 (m, ov)	42.3	
4'''	1.60 (quin, 6.6)	24.3	
5'''	0.91 (d, 6.6) <sup>3</sup>	23.0 <sup>3</sup>	
6'''	0.87 (d, 6.6) <sup>3</sup>	22.0 <sup>3</sup>	
1''''	–	Not obs.	
2''''	4.48 (vbr) <sup>7</sup>	Not obs.	
2''''-NH	8.86 (d, 9.5)		121.3
3''''	2.84–2.91 (m, ov) 3.33–3.38 (m, ov)	26.3	
4''''	–	112.7	
5''''	–	128.4	
6''''	7.60 (d, 7.0)	118.1	
7''''	7.09 (td, 7.0, 1.6)	119.3	
8''''	7.12 (td, 7.0, 1.6)	121.5	
9''''	7.14 (d, 7.0)	112.1	
10''''	–	137.1	
11''''	6.96 (s, br)	128.3	

<sup>1</sup> Tentative assignment, clear signal identified in <sup>13</sup>C spectrum but no 2D correlations observed. <sup>2</sup> Broad signal in proton spectrum baseline. Assigned from NOESY correlation to H2' ( $\delta_H$  4.39) and TOCSY correlation to H3'' ( $\delta_H$  2.92–2.98,  $\delta_H$  2.78–2.84) and H2'' ( $\delta_H$  4.58). <sup>3</sup> Signal assignments may be interchanged. <sup>4</sup> Chemical shift assigned from <sup>13</sup>C-HMBC spectrum. <sup>5</sup> Chemical shift assigned from <sup>13</sup>C-HSQC spectrum. <sup>6</sup> Chemical shifts assigned from <sup>15</sup>N-HSQC spectrum. <sup>7</sup> Tentative assignment of very broad signal. Identified by TOCSY correlations to 2''''-NH ( $\delta_H$  8.86) and H3'''' ( $\delta_H$  2.84–2.91, 3.33–3.38).



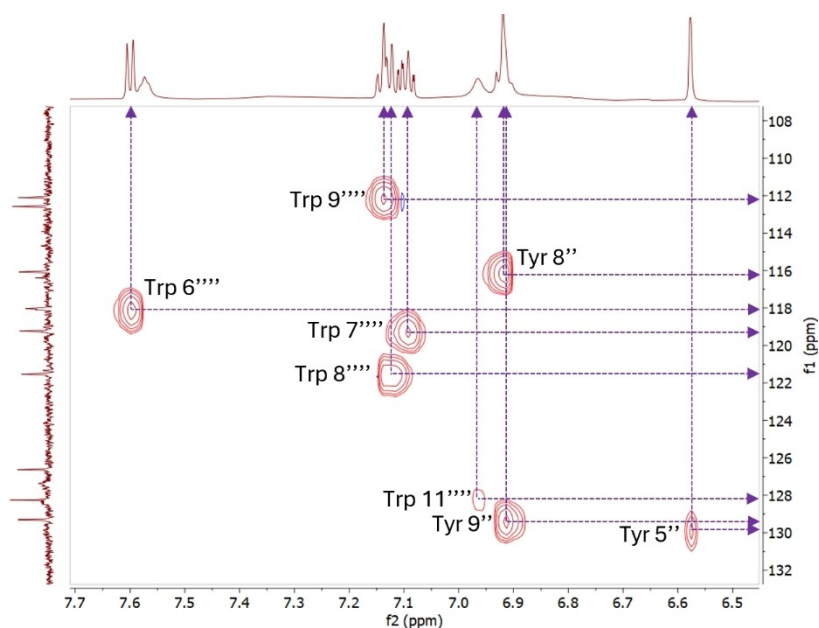


Figure 4. Extract of  $^{13}\text{C}$ -HSQC spectrum of **4** indicating observed  $^1\text{H}$ - $^{13}\text{C}$  correlations of aromatic protons (x-axis) and carbons (y-axis).

integrated to one proton. The identification of eight aromatic CH protons was consistent with heteroatom involvement in one side of the crosslink, producing an A–X–B system as previously indicated by the HDX analysis.<sup>[5a]</sup>

A hydroxyl proton was identified as a sharp singlet at  $\delta_{\text{H}}$  9.87 and demonstrated an HMBC correlation to Tyr-C7'' at the characteristic  $\delta_{\text{C}}$  150.2. The  $\delta_{\text{H}}$  9.87 hydroxyl proton displayed further HMBC correlations to the quaternary position Tyr-C6'' ( $\delta_{\text{C}}$  126.8) and to Tyr-C8'' at  $\delta_{\text{C}}$  116.2 ( $\delta_{\text{H}}$  6.90–6.93). Tyr-H9'' overlapped with the Tyr-H8'' signal but the corresponding Tyr C9'' carbon resonance was identified by  $^{13}\text{C}$ -HSQC at  $\delta_{\text{C}}$  129.4. Tyr-H5'' was located at  $\delta_{\text{H}}$  6.58 ( $\delta_{\text{C}}$  129.8) and presented as a typical  $J = 1.9$  Hz *meta* coupled doublet and was correlated by

HMBC to Tyr-C9''. Further HMBC correlations were identified between Tyr-H5'' and the quaternary positions Tyr-C6'' and Tyr-C7'' confirming the presence of a trisubstituted Tyr residue (Figure 3).

Full analysis of COSY,  $^{13}\text{C}$  HSQC and HMBC spectra identified and allowed assignment of the four CH positions of the six membered ring of the Trp indole moiety. This indicated that none of these positions were involved in the crosslink and ruled out the more commonly observed C–C crosslinks of the types observed in pseudosporamide<sup>[10]</sup> or rubrin.<sup>[3b]</sup> The broad  $\delta_{\text{H}}$  6.96 proton signal was assigned to Trp-H11''' ( $\delta_{\text{C}}$  128.3), this assignment was supported by observation of an nOe correlation between Trp-H11''' and Tyr-H5'' ( $\delta_{\text{H}}$  6.58) (Figure 5). A further

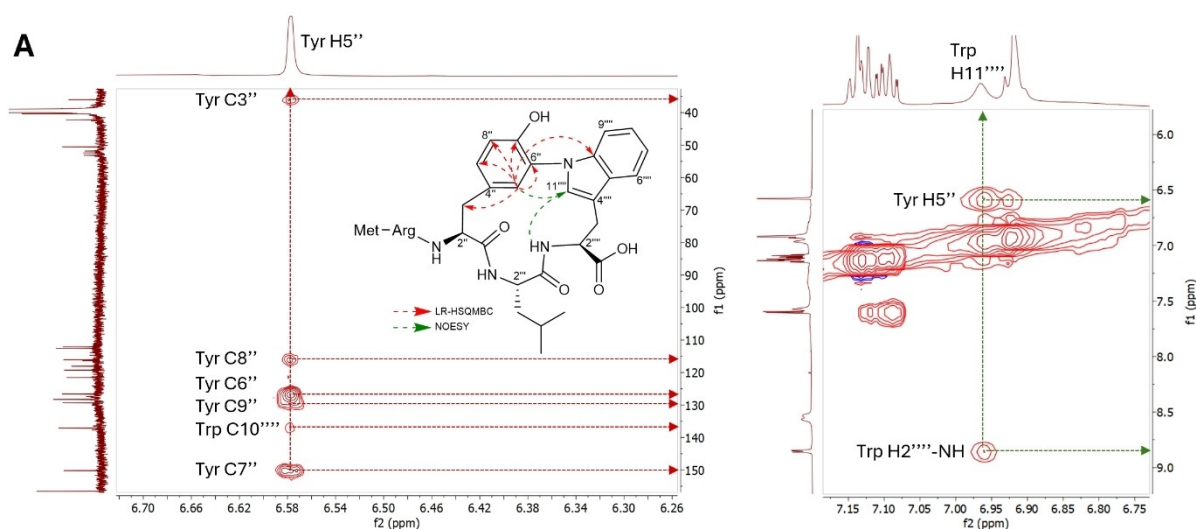


Figure 5. Selected 2D NMR experiments from the characterisation of **4**. (A) LR-HSQMBC correlations observed from Tyr H5'' ( $\delta_{\text{H}}$  6.58) using an LR-selHSQMBC experiment at 6.58 ppm, 40 ms shaped pulse. (B) Extract of NOESY spectrum showing correlations observed from Trp H11'''. Relevant LR-HSQMBC and nOe correlations are illustrated (A inset). For full correlations, assignments and spectra see S1.

nOe correlation was observed between Trp-H11<sup>'''</sup> and Trp-2<sup>'''</sup>-NH. This suggests that the Trp-H11<sup>'''</sup> proton is positioned within the macrocycle; conformational heterogeneity of this strained macrocycle may be responsible for the observed signal broadening. Computational analysis of the most stable conformers of the two possible atropisomers, both with the Trp-H11<sup>'''</sup> proton facing toward the interior of the macrocycle, revealed that the Trp-H11<sup>'''</sup> to Tyr-H5<sup>''</sup> and the Trp-H11<sup>'''</sup> to Trp-2<sup>'''</sup>-NH atom-atom distances fell within 4 Å, consistent with the observation of nOe correlations (See SI). The observation of the five CH protons of the indole moiety combined with the HDX evidence for heteroatom involvement suggests participation of the indole nitrogen in the crosslink and therefore a Tyr C6<sup>''</sup>-Trp C11<sup>'''</sup>-N crosslink. This rare Tyr C-Trp N crosslink was described recently in the peptide lapparbin,<sup>[3b]</sup> in which the analogous Trp-H11<sup>'''</sup> position was reported with comparable chemical shifts to those reported here,  $\delta_{\text{H}}$  7.19,  $\delta_{\text{C}}$  129.9 (D<sub>2</sub>O).<sup>[3b]</sup> The NMR of a truncated version of this structure was contemporaneously reported in DMSO-d<sub>6</sub> with the analogous Trp position identified at  $\delta_{\text{H}}$  7.14,  $\delta_{\text{C}}$  128.8.<sup>5</sup> Additionally, the Tyr C6<sup>''</sup> chemical shift was documented in these two reports at  $\delta_{\text{C}}$  130.0 (D<sub>2</sub>O)<sup>[3b]</sup> and  $\delta_{\text{C}}$  126.6 (DMSO-d<sub>6</sub>),<sup>5</sup> respectively; these chemical shifts are consistent with the  $\delta_{\text{C}}$  126.8 (DMSO-d<sub>6</sub>) reported in this work.

No signals were observed in <sup>13</sup>C-HMBC experiments indicative of correlations through the crosslink. Such correlations would however require observation of <sup>4</sup>J<sub>C-H</sub> coupling not typically observed with an 8 Hz <sup>13</sup>C-HMBC experiment; a 4 Hz <sup>13</sup>C-HMBC experiment was similarly unsuccessful. LR-HSQMBC is known to allow observation of correlations with smaller coupling constants and has been used successfully in cross-linked peptides.<sup>[3d]</sup> However, the LR-selHSQMBC (selective) experiment has been reported to provide improved sensitivity of long range coupling optimised for small J<sub>C-H</sub> values when compared to the standard broadband LR-HSQMBC experiment.<sup>[11]</sup> The LR-selHSQMBC experiment was thus undertaken centred at 6.58 ppm, (Tyr H5<sup>''</sup>) using a 40 ms gaussian 180° shaped pulse and the data acquired using the in-phase mode only.<sup>[12]</sup> This experiment allowed observation of a <sup>4</sup>J<sub>C-H</sub> correlation from Tyr H5<sup>''</sup> ( $\delta_{\text{H}}$  6.58) to Trp C10<sup>'''</sup> ( $\delta_{\text{C}}$  137.1) confirming the presence of the crosslink between the tyrosine and the tryptophan, and is consistent with the proposed Tyr C6<sup>''</sup>-Trp C11<sup>'''</sup>-N crosslink (Figure 5A).

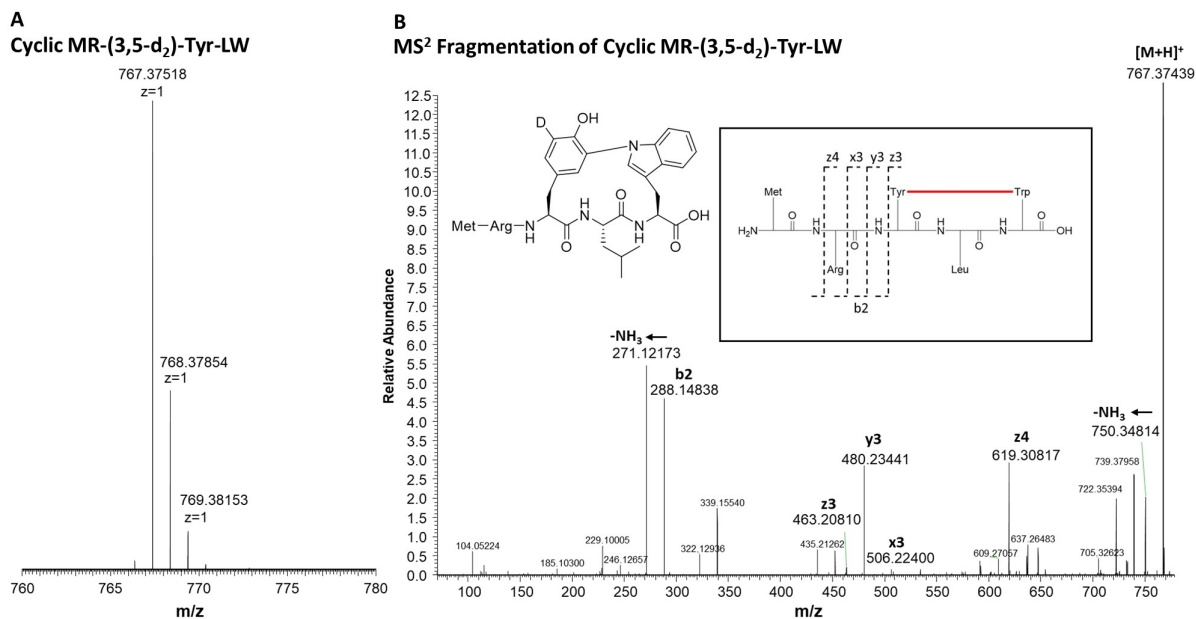
Using this revised NMR data, we undertook a computational NMR analysis. Model systems of the crosslinked MRYLW peptide **4** were created by removing the Met-Arg amino acids and assembling 13 possible C–C, C–O, C–N, and N–O linkages between the indole and phenol (see SI). DP5<sup>[13]</sup> and DP4AI<sup>[14]</sup> analyses of their computational <sup>13</sup>C spectra were the most informative. The system containing a Tyr C6<sup>''</sup>-Trp C11<sup>'''</sup>-N crosslink had the highest degree of matching to the experimental spectrum according to the DP5 metric (71.7% probability, which is considered to represent a high correlation when NMR variance is taken into account<sup>[13]</sup>). It was also the most probable structure among all the computed structures according to the DP4AI metric (77.5% probability). The second most likely structure (with DP5 and DP4AI probability values of 70.0%

and 21.0% respectively) was a C–C linkage between the Tyr-C6<sup>''</sup> and the Trp-C9<sup>'''</sup> atoms; however, this would not align with the observed four-proton spin system of the indole. We also examined the computed <sup>1</sup>H spectra, but found the results to be less conclusive, even when isolated to the aromatic region.

To further verify this revised structure, we performed the enzymatic turnover of a double deuterated 3,5-d<sub>2</sub>-Tyr containing MRYLW peptide using P450<sub>Blt</sub>, as we have recently demonstrated that this is a valuable technique for revealing the nature of the crosslink in such systems.<sup>[9]</sup> The results of this turnover experiment showed the loss of a single deuterium atom in the formation of **4<sub>d</sub>**. This agrees with the crosslink within the YxW peptide as linking via the *m*-position of the Tyr residue (Calculated *m/z* 767.37677, actual *m/z* 767.37518,  $\Delta$  2.1 ppm), further confirming the revised structure of the peptide as determined by NMR (Figure 6).

With the structure of the revised YxW peptide **4** indicating that the crosslinking occurs via the indole nitrogen, we further sought to interrogate this in the context of a recent structure solved for P450<sub>Blt</sub> in complex with an MRYLH peptide substrate.<sup>[8]</sup> In the wildtype P450<sub>Blt</sub>/MRYLH complex, the interactions of a His-residue in the I-helix (His-234) with the His-5 residue of the peptide substrate were shown to be crucial in controlling the nature of the crosslink installed (i.e. C–N vs C–C).<sup>[8]</sup> Inspecting this structure and taking the previously reported binding spectra of the MRYLW peptide into account,<sup>[5a]</sup> it becomes clear that the indole ring can only be accommodated in a manner consistent with peptide binding such that the indole ring is oriented towards the I-helix, and with the indole NH then oriented towards both the heme and the Tyr-3 ring. The added bulk of the Trp residue explains the low affinity binding measured with the MRYLW peptide. This also helps to explain the atypical binding spectra of MRYLW, as the binding spectra for this complex lacks the typical increase in absorbance at ~390 nm and decrease at ~420 nm that is seen in water displacement from the heme iron upon substrate binding.<sup>[5a]</sup> Indeed, the spectral signature observed (increase in absorbance at ~450 nm, decrease at ~420 nm) is closer to that seen with the interaction of the heme iron with a nitrogen ligand, possibly in this case the water bridged indole of the MRYLW peptide. Furthermore, no H-bonding interaction with His-234 would be possible for MRYLW without major rearrangements of the peptide substrate backbone.

Given this, and to explore the hypothesised apparent indifference of biaryllyte P450s for His to Trp substitutions within peptide substrates, we turned to *in vivo* experiments to investigate the crosslinking of an altered MRYLW peptide by BytO, the archetypal biaryllyte P450 previously shown to incorporate a C–C crosslink in its native MRYH peptide encoded by the minimal gene *bytA*. Heterologous expression of *bytAO* in *Streptomyces coelicolor* M1152 utilising the integrating plasmid pSET\_ermE\* was sufficient to produce a crosslinked N-acetylated Tyr-Tyr-His tripeptide.<sup>[7b]</sup> For mutational studies, the plasmid pSET\_ermE\* *bytAO* was mutated with primers in the *bytA* region to change its native amino acid sequence from MRYH to MRYLW. This plasmid was conjugated to *S. coelicolor* M1152, which was then cultivated and extracted according to



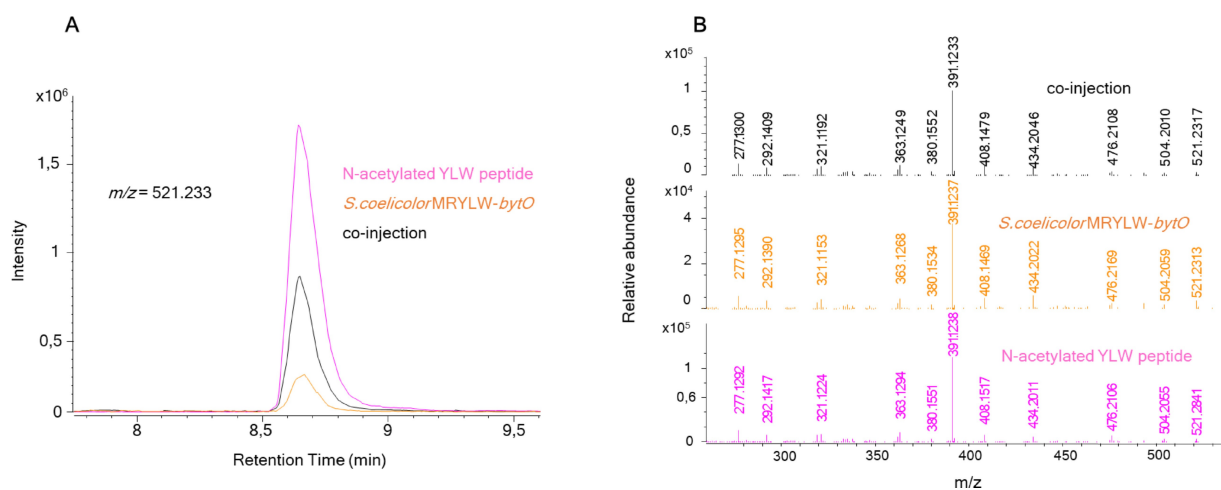
**Figure 6.** Cyclization of MR-(3,5-d<sub>2</sub>)-Tyr-LW by P450<sub>Blt</sub>. (A) HRMS of the MS<sup>1</sup> isotope cluster of 4<sub>d</sub> is indicative of the loss of deuterium during peptide cyclization, supporting the linking via the *m*-position of the Tyr residue in 4<sub>d</sub>. (B) MS<sup>2</sup> fragmentation of cyclic MR-(3,5-d<sub>2</sub>)-Tyr-LW 4<sub>d</sub> is supportive of the formation of a YLW crosslink, with the anticipated loss of deuterium observed in the crosslinked YLW tripeptide fragments.

previously described protocols.<sup>[7b]</sup> The extract from this strain clearly demonstrated the ability of BytO to crosslink the Trp-containing acetylated tripeptide (Figure 7A). To test if the crosslink was the same as that installed by P450<sub>Blt</sub> crosslinked pentapeptide 4 was digested with trypsin to remove the N-terminal MR dipeptide and subsequently acetylated to afford 5. Comparison of this standard 5 with the *in vivo* product of BytO was then performed. Co-injection of these samples showed co-elution of the peaks that also exhibited the same mass and fragmentation patterns, indicating the formation of the same C–N crosslink by both enzymes (Figure 7B). This result indicates that the YxW crosslinking reaction is independent of the

presence of an equivalent His-234 residue within the I-helix of BytO, and further shows that both C–C and C–N type biaryllyte crosslinking P450s tolerate tryptophan in place of histidine well despite there being no report of a natural biaryllyte peptide containing a crosslinked Trp residue.<sup>[3b,7b]</sup>

## Conclusions

P450<sub>Blt</sub> is a versatile biaryllyte enzyme that can install a range of crosslinks between His, Tyr and Trp side chains in pentapeptides. Here, we have reassigned the structure of the crosslinked



**Figure 7.** Mutation of the MRYH precursor peptide BytA into MRYLW remains an effective substrate for P450 BytO *in vivo*. High-resolution extracted ion chromatograms (EIC) displaying co-elution of N-acetylated YLW peptide 5 and *S. coelicolor* MRYLW-bytO product, with a mass-to-charge ratio (*m/z*) of 521.233 (Calculated [C<sub>28</sub>H<sub>33</sub>N<sub>4</sub>O<sub>6</sub>]<sup>+</sup>: 521.2395.) B) Comparison of (*m/z*) 521.233 MS<sup>2</sup> fragmentation from each sample shows identical fragmentation patterns.

MRYLW peptide **4** (Figure 3) as occurring via the Trp indole N to the *m*-position of the Tyr, which is supported by extensive NMR experiments as well as computational and isotope labelling experiments. We have further shown that such crosslinking is not restricted to C–N histidine crosslinking biaryllyl P450s and can also be accomplished by C–C forming enzymes. The ability of P450<sub>BLT</sub> to generate Trp crosslinked peptides is remarkable given the size of the sidechain and the lack of natural examples of Trp-crosslinked substrates and has great potential for the future synthetic generation of a range of constrained, cross-linked tripeptides.

## Experimental

### Computational Methods

Conformational searching on model crosslinked peptides was performed in MacroModel<sup>[15]</sup> using default settings and the OPLS3e<sup>[16]</sup> force field. Geometry optimizations of the resulting structures were performed using Gaussian 16 Revision C.01<sup>[17]</sup> at the B3LYP/6-31G(d)/IEFPCM(DMSO)<sup>[18]</sup> level while single point energies used M06-2X/def2-TZVP/IEFPCM(DMSO)<sup>[18c,19]</sup> and NMR calculations used mPW1PW91/def2-TZVP/IEFPCM(DMSO)<sup>[18c,19b,20]</sup>. The DP4A<sup>[14]</sup> and DP5<sup>[13]</sup> analysis programs were used to statistically determine the predicted most likely structures.

### NMR Assignment

Nuclear magnetic resonance (NMR) spectra of the crosslinked peptide **4** was collected using a Bruker Avance III HD equipped with a 16.4 T magnet and TCI cryoprobe operating at 700 MHz (<sup>1</sup>H), 176 MHz (<sup>13</sup>C), 71 MHz (<sup>15</sup>N).

Chemical shifts ( $\delta$ ) are reported in parts per million (ppm) and referenced to the residual solvent signals. <sup>15</sup>N spectra are referenced to liquid ammonia  $\delta_N$  0 ppm using the Bruker program Xiref. Coupling constants, *J* are reported in Hz.

<sup>13</sup>C-HMBC spectra reported herein were collected using 8 Hz *J*<sub>C-H</sub>. LR-SELHSQMBC spectra were collected using 3 Hz *J*<sub>C-H</sub> and a 40 ms gaussian 180° shaped pulse, centred as indicated.

### Synthesis of MR-d<sub>2</sub>-YLW Peptide

Peptide synthesis was performed manually on 2-chlorotrityl chloride resin (GL Biochem) at 0.025  $\mu$ mol scale. 2-Chlorotrityl chloride resin (25 mg) was washed with DCM (3×2 mL), swelled in DCM (2 mL, 30 min), and washed with DMF (3×2 mL). The first Fmoc protected amino acid (Fmoc-Trp(Boc)-OH, 3 eq) was loaded on the resin using DIEA (6 eq) at room temperature for 1 h. The unreacted chlorotrityl groups were capped with MeOH by washing (3×) with 2 mL DMF/DIEA/MeOH (7:2:1). Fmoc removal was achieved by washing (3×) with 20% piperidine in DMF. Coupling of subsequent amino acids (3 eq) was performed using OxymaPure (3 eq) and DIC (6 eq) in DMF at 50 °C for 10 min, with the deprotection and coupling reactions repeated three times to generate the required pentapeptide sequence. After a final deprotection step using 20% piperidine in DMF, the resin was washed three times with DCM before being subjected to peptide cleavage using 5 mL of TFA/TIS/H<sub>2</sub>O (95:2.5:2.5, v:v:v) solution with shaking for 2 h at room temperature. The separation of resin and cleaved peptide was achieved by filtration. After washing the resin with TFA (2×1 mL), the filtrate was concentrated under N<sub>2</sub> stream to ~1 mL,

followed by precipitation with pre-cooled diethyl ether (20 mL), centrifugation in a spark-free centrifuge (Spintron, GT-175) and washing with diethyl ether (3×10 mL). The peptide obtained was then dissolved in 10 mL of 50% MeCN in water and lyophilized. The peptide was analysed by LC–MS (Shimadzu) with a gradient of MeCN (0.1% FA) from 5% to 95% in 30 minutes.

### P450 Mediated Crosslinking of MR-d<sub>2</sub>-YLW Peptide

The P450 mediated crosslinking reaction was performed as described previously.<sup>[5a]</sup> Briefly, HEPES (50 mM), NaCl (50 mM), P450 enzyme (0.5  $\mu$ M), redox partners PuR (0.5  $\mu$ M) and PuxB A105 V (2.5  $\mu$ M),<sup>[21]</sup> glucose (0.33%), glucose dehydrogenase (0.033 mg/mL), substrate (25  $\mu$ M) and NADH (2 mM) were combined and incubated overnight at 30 °C at 300 rpm. Purification was then performed using solid-phase extraction (Bond elute, Agilent) which had been activated with 500  $\mu$ L of MeOH (0.1% FA) and equilibrated with 1 mL of water (0.1% FA). The reaction was diluted to 1 mL using water (0.1% FA) was applied to the column, washed with 1 mL of water (0.1% FA), and eluted with 600  $\mu$ L of MeOH (0.1% FA). The methanol elute was dried using a concentrator (Eppendorf), and then re-dissolved in 50  $\mu$ L of ACN/H<sub>2</sub>O (1:4, 0.1% FA) prior to analysis by LCMS. LCMS analysis was performed using a Shimadzu ultra-performance liquid chromatography system coupled to a mass spectrometer LCMS-8050 (ESI, operating both in positive and negative mode) equipped with an SPD–M40 photodiode array detector and an LC-40B solvent delivery module. Analytical separation was performed using a Phenomenex Aeris<sup>TM</sup> 1.7  $\mu$ m PEPTIDE XB–C18 100 column.

High-resolution mass spectrometry analysis of the products of the P450 crosslinking reaction was performed on an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled to a Dionex UltiMate 3000 RSLCnano system equipped with a Dionex UltiMate 3000 RS autosampler, an Acclaim PepMap RSLC analytical column (75  $\mu$ m×50 cm, nanoViper, C18, 2  $\mu$ m, 100 Å; Thermo Scientific) and an Acclaim PepMap 100 trap column (100  $\mu$ m×2 cm, nanoViper, C18, 5  $\mu$ m, 100 Å; Thermo Scientific) as previously reported.<sup>[5a]</sup> HRMS data for **Linear MR-(3,5-d<sub>2</sub>)-Tyr-LW** HRMS (ESI) *m/z*: [M+H]<sup>+</sup> Calcd for C<sub>37</sub>H<sub>52</sub>D<sub>2</sub>N<sub>9</sub>O<sub>7</sub>S<sup>+</sup> 770.39870; Found 770.39716; **4 Cyclic MR-(5-d)-Tyr-LW 4<sub>d</sub>** (ESI) *m/z*: [M+H]<sup>+</sup> Calcd for C<sub>37</sub>H<sub>51</sub>DN<sub>9</sub>O<sub>7</sub>S<sup>+</sup> 767.37677; Found 767.37518.

### Synthesis of Cyclic Ac-YLW Tripeptide

The cyclic MRYLW peptide **4** obtained using P450<sub>BLT</sub> (~150  $\mu$ g) was dissolved in 100  $\mu$ L of 50 mM Tris-HCl (pH 8) with a drop of DMSO added to aid with peptide solubility. Trypsin was added in a protease: peptide ratio of 1:20 (w/w) and the mixture was incubated at 37 °C overnight. The cleavage reaction was quenched by the addition of 1 mL of ice-cold methanol, and the precipitated protein and salts were removed by centrifuging at 50,000 *g* for 10 min. The supernatant was dried using a concentrator (Eppendorf), and the obtained tripeptide was then dissolved in 2 mL of DCM. Acetylation was achieved using acetic anhydride (3 eq) and DIEA (6 eq), with the DCM removed under an N<sub>2</sub> stream after acetylation was complete. The removal of the phenolic acetyl groups on Tyr was achieved as previously reported.<sup>[22]</sup> Briefly, the acetylated peptide was dissolved in a methanol/ water mixture (2:1) and the pH was adjusted to pH 9 using a saturated NaHCO<sub>3</sub> solution. The reaction was monitored using LC–MS and was deemed complete after 3 h incubation time at room temperature. The mixture was adjusted to pH 3 and extracted with ethyl acetate (3×). The solvent was removed *in vacuo* to obtain the crude cyclic Ac-YLW tripeptide standard **5**.



**Table 2.** Primers for preparation of *bytA* mutant plasmid pSET\_ermE\*\_MRYLW\_bytO. Mismatching nucleotides introducing mutations are underlined.

Primer	Sequence 5'-3'
pSETbytAO.rev	GCAGTCTGACTCTAGAGAGGCTAGCGGGGAGAAGGACGACC
MRYYH -> MRYLW.for	AGCAACGGAGGTACGGAAGGAGGAGGTGCGATGCGCTAC <u>CTCTGGT</u> GATCCGCCGTTT

### In vivo *bytA* Mutation

To introduce point mutations into *BytA*, mutagenesis primers (Table 2) were designed and the 1303 bp sized PCR fragment cloned into pSET152\_ermE\* via Gibson assembly. As a template, pSET152\_ermE\* *bytA*O was used.<sup>[7b]</sup> Transformation into *Streptomyces coelicolor* M1152, heterologous expression and extraction were performed as described by Zdouc et al., 2021.<sup>[7b]</sup>

High resolution HPLC-MS analyses of the extracts and **5** were conducted as described by Hansen et al., 2024.<sup>[8]</sup>

### Supplementary Information

NMR spectra and computational data. The authors have cited additional references within the Supporting Information.<sup>[13–14,19–21]</sup>

### Acknowledgements

We thank Dr Josep Saurí (IQS Barcelona) for technical advice regarding parameters for the LR-SelHSQMBC experiment and Dr Joel Hooper regarding the synthesis of deuterated Tyr. This study used BPA-enabled (Bioplatforms Australia)/NCRIS-enabled (National Collaborative Research Infrastructure Strategy) infrastructure located at the Monash Proteomics and Metabolomics Platform. Computational resources were provided by the National Computational Infrastructure through the National Computational Merit Allocation Scheme (supported by NCRIS) and by the University of Queensland Research Computing Centre. This work was supported by Monash University and EMBL Australia. L.P. was funded through a PhD scholarship from the German Academic Scholarship Foundation. A.K. acknowledges the support of an Australian Government Research Training Program Scholarship and an American Australian Association Graduate Education Fund Scholarship. This research was conducted by the Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012) and funded by the Australian Government. Open Access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

### Conflict of Interests

The authors declare no conflict of interest.

### Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

- [1] A. Zorzi, K. Deyle, C. Heinis, *Curr. Opin. Chem. Biol.* **2017**, *38*, 24–29.
- [2] a) H. Aldemir, S. Shu, F. Schaefer, H. Hong, R. Richarz, S. Harteis, M. Einsiedler, T. M. Milzarek, S. Schneider, T. A. M. Gulder, *Chem. Eur. J.* **2022**, *28*, e202103389; b) K. Haslinger, M. Peschke, C. Brieke, E. Maximowitsch, M. J. Cryle, *Nature* **2015**, *521*, 105–109; c) R. D. Süßmuth, A. Mainz, *Angew. Chem. Int. Ed.* **2017**, *56*, 3770–3821.
- [3] a) Y. L. Hu, F. Z. Yin, J. Shi, S. Y. Ma, Z. R. Wang, R. X. Tan, R. H. Jiao, H. M. Ge, *J. Am. Chem. Soc.* **2023**, *145*, 27325–27335; b) H. Nam, J. S. An, J. Lee, Y. Yun, H. Lee, H. Park, Y. Jung, K.-B. Oh, D.-C. Oh, S. Kim, *J. Am. Chem. Soc.* **2023**, *145*, 22047–22057; c) J. S. An, H. Lee, H. Kim, S. Woo, H. Nam, J. Lee, J. Y. Lee, S.-J. Nam, S. K. Lee, K.-B. Oh, S. Kim, D.-C. Oh, *Angew. Chem. Int. Ed.* **2023**, *62*, e202300998; d) H. Saad, T. Majer, K. Bhattarai, S. Lampe, D. T. Nguyen, M. Kramer, J. Straetener, H. Brötz-Oesterheld, D. A. Mitchell, H. Gross, *Chem. Sci.* **2023**, *14*, 13176–13183; e) S. Kunakom, H. Otani, D. W. Udway, D. T. Doering, N. J. Mouncey, *Journal of Industrial Microbiology and Biotechnology* **2023**, *50*; f) B.-B. He, J. Liu, Z. Cheng, R. Liu, Z. Zhong, Y. Gao, H. Liu, Z.-M. Song, Y. Tian, Y.-X. Li, *Angew. Chem. Int. Ed.* **2023**, *62*, e202311533; g) S. T. Lima, B. G. Ampolini, E. B. Underwood, T. N. Graf, C. E. Earp, I. C. Khedi, M. A. Pasquale, J. R. Chekan, *Angew. Chem. Int. Ed.* **2023**, *62*, e202218082; h) L. B. Bushin, B. C. Covington, K. A. Clark, A. Caruso, M. R. Seyedsayamdost, *Nat. Chem. Biol.* **2022**, *18*, 1135–1143; i) R. Sugiyama, A. F. L. Suarez, Y. Morishita, T. Q. N. Nguyen, Y. W. Tooh, M. N. H. B. Roslan, J. Lo Choy, Q. Su, W. Y. Goh, G. A. Gunawan, F. T. Wong, B. I. Morinaka, *J. Am. Chem. Soc.* **2022**, *144*, 11580–11593; j) H. Nguyen, I. D. Made Kresna, N. Böhringer, J. Ruel, E. d. I. Mora, J.-C. Kramer, K. Lewis, Y. Nicolet, T. F. Schäberle, K. Yokoyama, *J. Am. Chem. Soc.* **2022**, *144*, 18876–18886; k) P. Nanudorn, S. Thiengmag, F. Biermann, P. Erkoc, S. D. Dirnberger, T. N. Phan, R. Fürst, R. Ueoka, E. J. N. Helfrich, *Angew. Chem. Int. Ed.* **2022**, *61*, e202208361.
- [4] A. Greule, J. E. Stok, J. J. De Voss, M. J. Cryle, *Nat. Prod. Rep.* **2018**, *35*, 757–791.
- [5] a) Y. Zhao, E. Marschall, M. Treisman, A. McKay, L. Padva, M. Crüsemann, D. R. Nelson, D. L. Steer, R. B. Schittenhelm, J. Tailhades, M. J. Cryle, *Angew. Chem. Int. Ed.* **2022**, *61*, e202204957; b) T. N. Dinh, Z. Lingyang, A. M. Douglas, A. v. d. D. Wilfred, *bioRxiv* **2023**, 2023.2010.2030.564719.
- [6] J. J. Hug, J. Dastbaz, S. Adam, O. Revermann, J. Koehnke, D. Krug, R. Müller, *ACS Chem. Biol.* **2020**, *15*, 2221–2231.
- [7] a) J. J. Hug, N. A. Frank, C. Walt, P. Šenica, F. M. R. Panter, *Molecules* **2021**, *26*, 7483; b) M. M. Zdouc, M. M. Alanjary, G. S. Zarazúa, S. I. Maffioli, M. Crüsemann, M. H. Medema, S. Donadio, M. Sosio, *Cell Chem. Biol.* **2021**, *28*, 733–739.e734.
- [8] M. H. Hansen, A. Keto, M. Treisman, V. M. Sasi, L. Coe, Y. Zhao, L. Padva, C. Hess, V. Leichthammer, D. L. Machell, R. B. Schittenhelm, C. J. Jackson, J. Tailhades, M. Crüsemann, J. J. De Voss, E. H. Krenske, M. Cryle, *ACS Catal.* **2024**, *14*, 812–826.
- [9] M. Treisman, L. Coe, Y. Zhao, V. M. Sasi, J. Gullick, M. H. Hansen, A. Ly, V. Leichthammer, C. Hess, D. L. Machell, R. B. Schittenhelm, J. Hooper, C. J. Jackson, J. Tailhades, J. J. De Voss, M. Cryle, *Org. Lett.* **2024**, *26*, 1828–1833.
- [10] S. Saito, K. Atsumi, T. Zhou, K. Fukaya, D. Urabe, N. Oku, M. R. U. Karim, H. Komaki, Y. Igarashi, *Beilstein J. Org. Chem.* **2020**, *16*, 1100–1110.
- [11] K. Motiram-Corral, P. Nolis, J. Sauri, T. Parella, *J. Nat. Prod.* **2020**, *83*, 1275–1282.
- [12] J. Sauri, P. Nolis, T. Parella, *Magn. Reson. Chem.* **2020**, *58*, 363–375.
- [13] A. Howarth, J. M. Goodman, *Chem. Sci.* **2022**, *13*, 3507–3518.

- [14] a) A. Howarth, K. Ermanis, J. M. Goodman, *Chem. Sci.* **2020**, *11*, 4351–4359; b) K. Ermanis, K. E. B. Parkes, T. Agback, J. M. Goodman, *Org. Biomol. Chem.* **2019**, *17*, 5886–5890.
- [15] a) L. Schrödinger, *Schrödinger Release 2019–2 ed.*, New York, NY, **2019**; b) K. S. Watts, P. Dalal, A. J. Tebben, D. L. Cheney, J. C. Shelley, *J. Chem. Inf. Model.* **2014**, *54*, 2680–2696; c) F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **2004**, *11*, 440–467.
- [16] K. Roos, C. Wu, W. Damm, M. Reboul, J. M. Stevenson, C. Lu, M. K. Dahlgren, S. Mondal, W. Chen, L. Wang, R. Abel, R. A. Friesner, E. D. Harder, *J. Chem. Theory Comput.* **2019**, *15*, 1863–1874.
- [17] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Wallingford, CT, **2016**.
- [18] a) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652; b) W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257–2261; c) G. Scalmani, M. J. Frisch, *J. Chem. Phys.* **2010**, *132*, 114110.
- [19] a) Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215–241; b) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- [20] C. Adamo, V. Barone, *J. Chem. Phys.* **1998**, *108*, 664–675.
- [21] S. G. Bell, F. Xu, E. O. D. Johnson, I. M. Forward, M. Bartlam, Z. Rao, L.-L. Wong, *J. Biol. Inorg. Chem.* **2010**, *15*, 315–328.
- [22] C. Brieke, M. J. Cryle, *Org. Lett.* **2014**, *16*, 2454–2457.

Manuscript received: March 9, 2024  
Accepted manuscript online: May 7, 2024  
Version of record online: June 17, 2024

### 7.3 Appendix C

The following pages feature the article “Structural Insights into a Side Chain Cross-Linking Biaryllytide P450 from RiPP Biosynthesis” as it was published in ACS Catalysis by American Chemical Society. This publication is licensed under CC BY 4.0 and therefore free for copy and redistribution of the material in any medium or format or any purpose. A written consent from American Chemical Society Publications is not necessary for this purpose.

The article is reprinted with permission from:

Mathias H. Hansen, Angus Keto, Maxine Treisman, Vishnu Mini Sasi, Laura Coe, Yongwei Zhao, Leo Padva, Caroline Hess, Victor Leichthammer, Daniel L. Machell, Ralf B. Schittenhelm, Colin J. Jackson, Julien Tailhades, Max Crüsemann, James J. De Voss, Elizabeth H. Krenske, Max J. Cryle. *ACS Catalysis* 2024, *14*, 812-826.

Copyright 2024 The Authors. Published by American Chemical Society.

To view the publication and supplementary information online, please follow the link below.

<https://doi.org/10.1021/acscatal.3c05417>

# Structural Insights into a Side Chain Cross-Linking Biaryllytide P450 from RiPP Biosynthesis

Mathias H. Hansen,<sup>\*,†</sup> Angus Keto,<sup>†</sup> Maxine Treisman,<sup>†</sup> Vishnu Mini Sasi, Laura Coe, Yongwei Zhao,<sup>†</sup> Leo Padva, Caroline Hess,<sup>†</sup> Victor Leichthammer,<sup>†</sup> Daniel L. Machell,<sup>†</sup> Ralf B. Schittenhelm,<sup>†</sup> Colin J. Jackson, Julien Tailhades,<sup>†</sup> Max Crüsemann, James J. De Voss, Elizabeth H. Krenske, and Max J. Cryle<sup>\*,†</sup>



Cite This: *ACS Catal.* 2024, 14, 812–826



Read Online

ACCESS |



Metrics & More



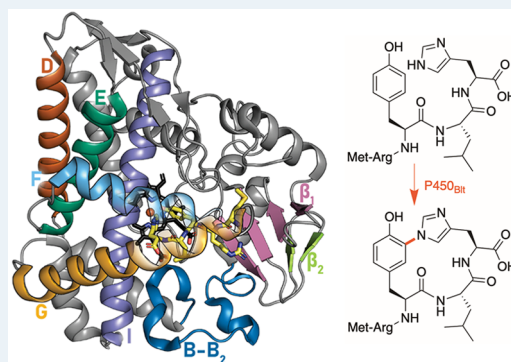
Article Recommendations



Supporting Information

**ABSTRACT:** Peptide side chain cross-linking is an important feature of many natural products, with an increasing number of examples catalyzed by cytochrome P450s being reported from ribosomal biosynthesis pathways in addition to well-known examples from nonribosomal peptide antibiotics. Despite the dramatic recent increase in the number of enzymes and reactions catalyzed, substrate bound structures of such P450s have proven elusive to date. Here, we report the structural characterization of the biaryllytide cross-linking enzyme P450<sub>Blt</sub> in complex with its pentapeptide substrate MRYLH. This structure, in combination with computational and biochemical experiments, shows the importance of key I-helix residues in this P450 in coordinating to the histidine residue of the substrate and further that this appears to be central to the specificity of this enzyme for generating a C–N link between the tyrosine and histidine residues in the MRYLH substrate. The structure of the P450<sub>Blt</sub>-MRYLH complex provides the first insight into how peptide substrates can be accommodated within P450s and offers insights into how other examples of related P450s can accept the varied substrates that have recently been identified using bioinformatic methods.

**KEYWORDS:** cytochrome P450, aromatic crosslinking, peptide, RiPP, biaryllytide



## INTRODUCTION

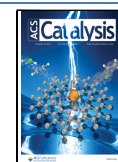
Cytochrome P450s make up a super family of powerful oxidative enzymes that are widely distributed in nature.<sup>1</sup> These haem-containing enzymes can access a range of highly reactive reaction intermediates, including Compound I (an iron(IV) oxo porphyrin cation radical species), via two single electron reductions and the selective delivery of two protons to haem-bound molecular oxygen within the P450 active site.<sup>2</sup> The generation of this powerful electrophilic intermediate requires the maintenance of a haem thiolate ligand, donated from a cysteine residue, at physiological pH, with the P450 active cycle also providing access to other reactive species including ferric hydroperoxyl and superoxide species. The structure of P450s is highly conserved, with a core 4-helix bundle surrounding the haem cofactor and containing the residues required for oxygen protonation (I-helix) and haem ligation (L-helix).<sup>3</sup> Despite this conservation, P450s display great diversity in the substrates they can oxidize, and furthermore can perform a wide variety of oxidative modifications beyond the archetypal P450 reaction, the hydroxylation of unactivated C–H bonds.<sup>3</sup> In biosynthesis pathways, P450s are often important for their ability to selectively install hydroxyl groups at a late stage in natural products, with such modifications seen

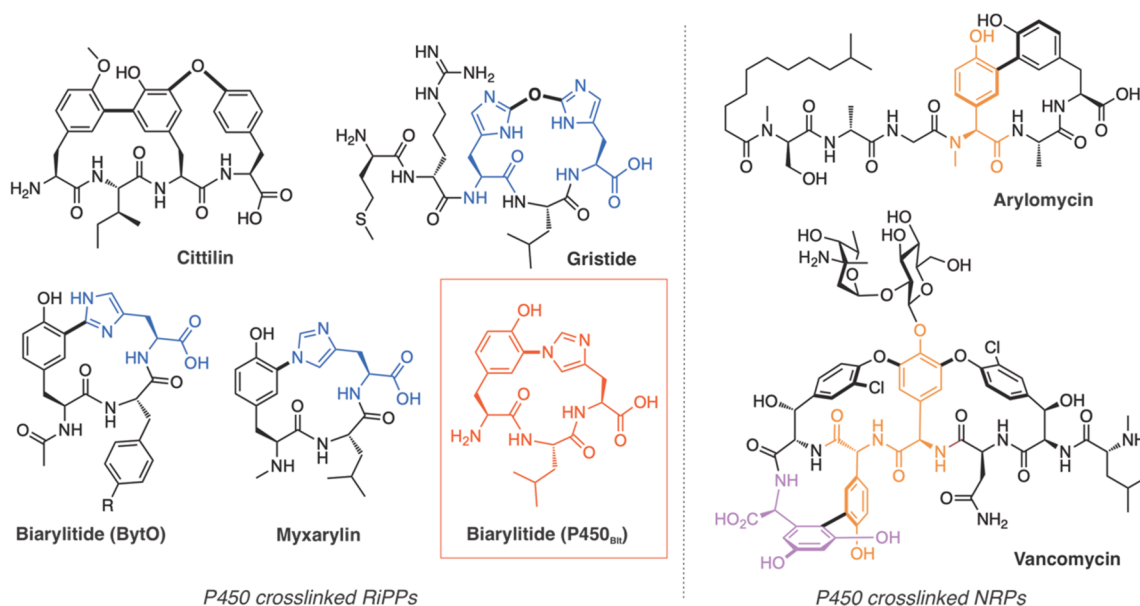
in peptide, polyketide, and terpene biosynthesis.<sup>3</sup> P450s can also generate perhaps more significant changes in structure through the cross-linking of aromatic groups (Figure 1). This is particularly notable in the biosynthesis of several important classes of nonribosomal peptides (NRPs), including the glycopeptide antibiotics (GPAs, such as vancomycin, teicoplanin and corbamyacin)<sup>4</sup> and arylomycin.<sup>5</sup> While P450s from NRP pathways have been characterized,<sup>5–9</sup> and in the case of GPAs also in complex with their essential recruitment domains,<sup>10,11</sup> none of these P450s have been in a substrate bound form, making an assessment of the mechanism utilized and the substrate binding mode employed by these enzymes elusive.<sup>12,13</sup> More recently, ribosomal biosynthesis pathways (Ribosomally synthesized and post-translationally modified peptides (RiPPs)) involving P450 cross-linking reactions have been widely reported and investigated, showing high levels of

Received: November 10, 2023

Accepted: December 12, 2023

Published: January 3, 2024





**Figure 1.** Examples of side chain cross-linked peptides from both ribosomal (RiPPs—citterlin, biarylilides (R = H, OH), myxarylin and gristide) and nonribosomal peptide biosynthesis pathways (arylomycin and vancomycin). Bonds (and atoms in the case of gristide) installed by the P450 enzymes are shown in bold. Residue colors: red – product of P450<sub>Blt</sub>, blue – histidine, orange – 4-hydroxyphenylglycine, pink – 3,5-dihydroxyphenylglycine.

**Table 1.** X-ray Data Collection and Refinement Statistics for P450<sub>Blt</sub> Structures Reported in This Work<sup>a</sup>

Protein	8U3N - P450 <sub>Blt</sub> (apo)	8U2M - P450 <sub>Blt</sub> (MRYLH)	8UKZ - P450 <sub>Blt</sub> (E238A)
Data collection			
Space group	P 1 2 <sub>1</sub> 1	P 1 2 <sub>1</sub> 1	P 1 2 <sub>1</sub> 1
Unit-cell parameters (Å, °)	61.078 95.748 105.217 90 92.629 90	61.674 95.607 105.763 90 92.706 90	61.957 94.48 105.985 90 92.677 90
Resolution range (Å)	46.07–2.15 (2.227–2.15)	43.55–1.79 (1.854–1.79)	46.18–1.95 (2.02–1.95)
Wavelength	0.95365	0.95365	0.95373
Total reflections	450333 (45519)	385962 (38974)	311364 (31378)
Unique reflections	65843 (6446)	114354 (11205)	88780 (8823)
Completeness (%)	99.64 (97.07)	99.07 (97.08)	99.89 (99.58)
R <sub>pim</sub> <sup>b</sup> (%)	5.964 (38.05)	6.294 (67.36)	6.958 (107.5)
(I/σ(I))	7.45 (1.25)	9.93 (0.96)	5.68 (0.89)
CC <sub>1/2</sub> <sup>c</sup> (%)	99.5 (69.3)	99.6 (35.5)	99.5 (25.7)
CC <sup>*d</sup> (%)	99.9 (90.5)	99.9 (72.4)	99.9 (64)
R <sub>merge</sub> <sup>e</sup> (%)	14.45 (93.73)	9.888 (107.4)	11.18 (174.2)
Average multiplicity	6.8 (7.0)	3.4 (3.5)	3.5 (3.6)
Wilson B-factor	34.20	25.65	34.75
Structure refinement			
R <sub>work</sub> /R <sub>free</sub> <sup>f</sup> (%)	19.33/23.48	19.86/22.95	20.54/22.42
RMSD, bond lengths (Å)	0.009	0.003	0.004
RMSD, bond angles (deg)	1.08	0.68	0.67
Ramachandran angles <sup>g</sup>			
Favored	98.58	98.86	98.48
Allowed	1.42	1.05	1.52
Outliers	0.00	0.09	0
B-factors			
Average B-factor	49.38	36.43	47.02
ligands	45.63	33.62	38.13
solvent	41.81	36.81	39.57

<sup>a</sup>Values in parentheses are for the highest resolution shell. <sup>b</sup> $R_{p.i.m.} = \sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl) \rangle|$  <sup>c</sup>

$$CC_{1/2} = \frac{\sum (x - \langle x \rangle)(y - \langle y \rangle)}{[\sum (x - \langle x \rangle)^2 \sum (y - \langle y \rangle)^2]^{1/2}} \quad ^d CC^* = \sqrt{\frac{2CC_{1/2}}{1 + CC_{1/2}}} \quad ^e R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \quad ^f$$

$$R = \frac{\sum |F_{obs}| - |F_{calc}|}{\sum |F_{obs}|} \quad ^g \text{Categories were defined by MolProbity}^{20}$$

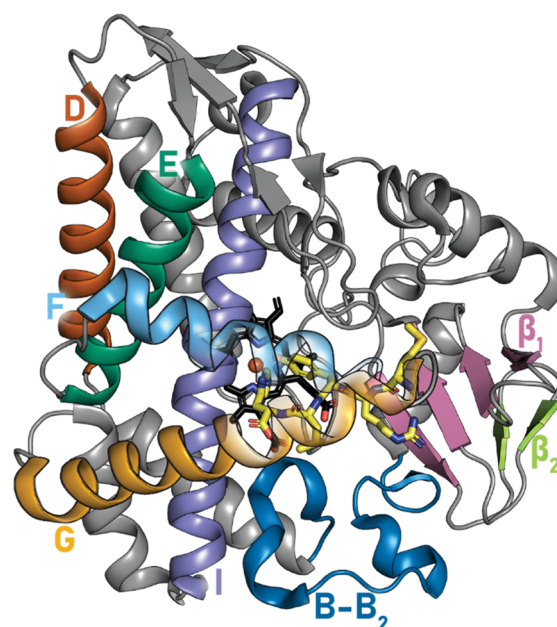


diversity in the reactions catalyzed and the substrates accepted by this intriguing group of P450s.<sup>14</sup> While many of these new classes of P450s utilize large leader sequences relative to their smaller core peptides, the biarylites<sup>15</sup> (and related myxarylins)<sup>16</sup> are remarkably atom efficient in their use of a two amino acid leader appended to a three amino acid core peptide. Biochemical characterization has shown that the biarylite P450s possess the ability to generate alternate (C–C, C–N) cross-links within their Tyr–Xaa–His cores (where Xaa can be highly varied) and furthermore that these enzymes can also tolerate a range of modifications within the cross-link and even the replacement of the terminal His residue with Trp.<sup>17</sup> Given these features and the potential importance of such P450s as biocatalysts to generate cross-linked peptide building blocks, we undertook to gain structural and mechanistic insights into these cross-linking P450s. Here, we report the first substrate bound structure of such a cross-linking P450 from the biarylite ribosomal pathway from *Micromonospora* sp. MW-13. This structure demonstrates the first insights into how these peptides are bound by P450 enzymes and furthermore suggests that the cross-link installed by such P450s is controlled by the protonation of the His-residue within the substrate peptide.

## RESULTS AND DISCUSSION

**Structural Characterization of P450<sub>Blt</sub>.** With relatively few instances of bound P450/peptide substrate complexes solved to date, we focused on obtaining a bound structure of P450<sub>Blt</sub> in complex with a pentapeptide substrate. While initial screening yielded several possible conditions, these all proved to be highly challenging to pursue due to large unit cells and relatively low resolution of the diffraction data. Careful inspection of PEG-containing conditions showed the presence of cubic crystals, which we optimized to obtain reproducible crystals with this new condition (0.1 M MES pH 6.5, 12% PEG 400 and 14% PEG 20K). Analysis showed that these diffracted to significantly better resolution than most initial hits, and thus we pursued these crystals, both to solve the structure of P450<sub>Blt</sub> (space group P 1 2<sub>1</sub> 1) to a resolution of 2.15 Å (PDB 8U3N, Table 1) and for use in soaking experiments to obtain peptide-bound complexes of this P450 (SI Figure S1). The structure of P450<sub>Blt</sub> in complex with a pentapeptide MRYLH substrate was finally obtained by soaking crystals supplemented with 50 mM MRYLH and 12% m-erythritol, affording a 1.8 Å data set (PDB 8U2N). The structure was determined using molecular replacement to obtain phases with a poly alanine model of CYP107L2 (PDB: 5CWE)<sup>18</sup> as the search model. While we were also able to initially obtain phases using a model generated by Robetta,<sup>19</sup> subsequent model building was expedited from the model generated from SCWE. The asymmetric unit contains 3 monomers (A, B, and C), with density for the pentapeptide substrate in the P450 active site only found in the A chain. The three monomers are highly similar to a root-mean-square deviation (RMSD) of 0.27–0.4 Å across the backbone atoms of all chains (Table 1).

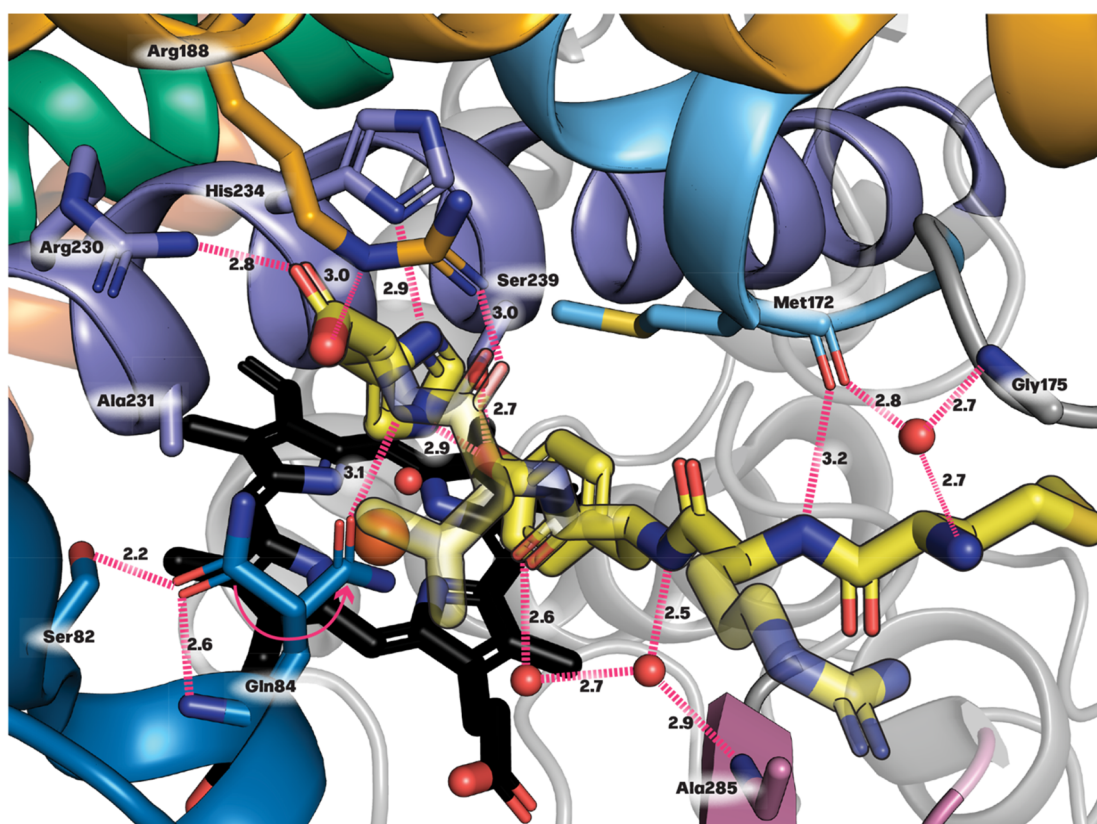
P450<sub>Blt</sub> adopts the canonical P450 fold with predominantly  $\alpha$ -helical secondary structure, several  $\beta$ -sheets, and a thiolate-coordinated haem moiety at the center of the protein (Figure 2). In total, P450<sub>Blt</sub> comprises 18  $\alpha$ -helices (designated A–L) and 3  $\beta$ -sheets ( $\beta$ -1 to  $\beta$ -3), with the central haem moiety sandwiched within the central four-helix bundle (comprising the D, E, I and J helices). The regions of secondary structure surrounding the active site are well-defined in the electron



**Figure 2.** Overall structure of P450<sub>Blt</sub> demonstrating the classic P450 fold ( $\alpha$ -helices A–L) together with three  $\beta$ -sheets ( $\beta$ -1 to  $\beta$ -3). Active site secondary structure elements are colored: F/G-helices that cap the active site in sky blue and orange,  $\beta$ -sheets 1 and 2 in magenta and chartreuse, D-helix in brown, E-helix in green, the B–B<sub>2</sub> loop in blue, and I-helix in slate. The haem moiety is presented in black with its central iron atom as an orange sphere, highlighting the enzyme active site.

density, including the capping F/G-helices and the B–B<sub>2</sub> loop region, which appears to form a stable conformation even in the absence of substrate. The propionate groups of the haem moiety are coordinated by salt bridging interactions from His-91/His-342 and Arg-286, with the haem thiolate ligand donated by the side chain of Cys-344. The typically conserved acid/alcohol pair of residues central to oxygen activation during the P450 active cycle are found in the I-helix above the haem (Glu-238/Ser-239).

**Peptide Binding to P450<sub>Blt</sub>.** The MRYLH substrate contacts all regions of the P450 active site, including the F/G-helices that cover the haem, the central I-helix containing essential catalytic residues as well as the B/C loop, the C-terminal loop, and the  $\beta$ -1 sheet (Figure 3). MRYLH is positioned perpendicular to the I-helix with the peptide C-terminal carboxylate group nested between the I- and G-helices. Such an orientation is similar therefore to that believed to occur in GPA cross-linking P450s,<sup>8,12</sup> although this has yet to be structurally verified. The N-terminal methionine residue of MRYLH is situated in a hydrophobic cavity largely formed by residues in the protein N-terminus (Pro-16),  $\beta$ -1 sheet (Leu-44, Ala-283, Val-306, Phe-308), Ala-380/Thr-381 in the C-terminal loop and Ile-176 at the end of the F-helix, with the side chain stabilized by van der Waals contacts. The terminal amine is not coordinated directly by the P450, with water 563 bridging the substrate N-terminal amino group (2.8 Å) and the backbone carbonyl (2.7 Å) of Met172 in the F-helix. The amino terminus projects into a largely hydrophobic pocket that is sufficiently sized to accommodate small modifications to the amine moiety (such as a formyl group, which is known to be accepted by this enzyme).<sup>17</sup> Instead of contacting the P450, the Met-1 carbonyl moiety is instead oriented toward the side chain of the neighboring Arg-2 residue (3.4 Å). The Tyr-3 and



**Figure 3.** MRYLH peptide bound in the active site of P450<sub>Blt</sub>. MRYLH is rendered in yellow sticks, with residues involved in hydrogen bonding—either directly or water-mediated, dispersed as sticks and labeled. Hydrogen bonds are depicted with dashed light magenta lines, highlighting the network of interactions between MRYLH and the active site residues. Red spheres represent water molecules that participate in this bonding network. The peptide substrate is aligned perpendicular to the I-helix, forming multiple hydrogen bonds; notably, the C-terminal carboxylate of MRYLH is coordinated between Arg-230 on the I-helix and Arg-188 on the G-helix. The N-terminal methionine of the substrate is situated close to Met-172 on the F-helix. Directly above haem pyrrole ring C, Tyr-3 of MRYLH establishes a critical hydrogen bond with Ser-239 on the I-helix. The backbone NH group of Tyr-3 is part of a hydrogen bond network with two water molecules, demonstrating intricate water-mediated interactions. His-5, positioned parallel to helix I, is hydrogen-bonded through its N $\epsilon$ 2 to His-234. The presence of the substrate induces Gln-84 to assume an alternative rotameric state, allowing a hydrogen bond to the substrate, illustrated with a light magenta arrow to indicate this dynamic conformational shift.

His-5 side chains of the MRYLH peptide that are cross-linked by P450<sub>Blt</sub> are oriented toward the haem, where there is still a haem-bound water molecule present (2.1 Å) despite the presence of the substrate. This phenomenon has also previously been reported in the substrate-bound structures of related P450 cross-linking enzymes such as CYP121 and appears to be caused by the cryogenic temperature used for diffraction experiments.<sup>21</sup> The side chain of the leucine residue in position 4 of the peptide substrate projects out from the substrate binding pocket but is within 3.3 Å of the unbound Gln-84 N $\epsilon$ 1, suggesting a possible function for the P4 residue side chain in displacing Gln-84 from its “resting” position in H-bonding Ser-82 to its “functional” position coordinating the substrate peptide backbone (*vide infra*). This hypothesis could also explain the low conversions previously noted for MRYXH substrates bearing modified, polar P4 side chains, while also helping to explain the reduced activity of a Q84L mutant (an order of magnitude reduction in affinity with a concomitant 50% reduction in conversion vs WT, Table 2).<sup>17</sup>

The C-terminus of MRYLH is tightly coordinated by the P450, with one carboxylate oxygen forming a hydrogen bond to N $\eta$ 1 of Arg-230 in the I-helix (2.8 Å) and the other hydrogen bonding to N $\epsilon$  of Arg-188 in the G-helix (3.0 Å; Figure 3). The presence of such interactions with the substrate

**Table 2.** Substrate Binding and Turnover Results for the MRYLH Substrate with Wildtype P450<sub>Blt</sub> and Mutants Prepared in This Study

Enzyme	Binding Type (spin state shift) <sup>a</sup>	K <sub>d</sub> (μM)	Conversion (%)
P450 <sub>Blt</sub> <sup>23</sup>	Type I (5%) <sup>17</sup>	2.1 <sup>17</sup>	85 ± 4.2 <sup>17</sup>
Q84L mutant	Type I (2%)	28 ± 3.0	40 ± 2.4
A231 V mutant	Type I (1%)	5.0 ± 3.0	45 ± 5.7
H234L mutant	Type II (1%)	0.6 ± 0.4	8.8 ± 2.3
E238A mutant	Type II (8%)	4.7 ± 3.0	28 ± 6.8

<sup>a</sup>Type I shift—substrate binding, water displacement from haem iron; Type II — inhibitor type binding through coordination to the haem iron (either directly or via water).

carboxylate moiety explains the loss of two-orders of magnitude binding affinity reported when this carboxylate is replaced with a terminal amide and is in keeping with the loss of a salt bridging interaction. The N $\eta$ 2 of Arg-188 forms another hydrogen bond with the carbonyl oxygen of the peptide bond between Leu-4/His-5 of the MRYLH substrate (3.0 Å). The amide nitrogen of the Leu-4/His-5 peptide bond is hydrogen bonded to O $\epsilon$ 1 of Gln-84 in the B/C-loop (3.1 Å), which is an alternate rotameric state to that seen in the absence of substrate (in which the Gln-84 O $\epsilon$ 1 hydrogen bonds to the



side chain of Ser-82 (2.2 Å). Tyr-3 of MRYLH is positioned primarily above the haem pyrrole ring C with its phenol moiety almost directly above the haem iron (4.3 Å). The phenol forms an intermolecular hydrogen bond to the hydroxyl moiety of the Ser-239 residue (2.7 Å) in the I-helix. Numerous van der Waals contacts are involved in the stabilizing the Tyr-3 benzene ring, which is oriented with its edge toward the pyrrole ring C of the haem and forms contacts via C4 and C5. The other side of the Tyr-3 benzene contacts Leu-382, while all atoms of the benzene ring contact Val-281. Ordered water molecules also interact with substrate backbone at Tyr-3, with water-103 hydrogen bonding the backbone carbonyl (2.6 Å) of Tyr-3 and to water-102 (2.7 Å), which in turn forms a hydrogen bond to the backbone NH group of Tyr-3 (2.5 Å). His-5 of the MRYLH substrate is oriented parallel to the I-helix with van der Waals interactions between C $\beta$ /Ala-231 and the imidazole ring/A-235. Curiously, Ne2 of His-5 hydrogen bonds to the N $\delta$ 1 of His-234 in the I-helix (2.9 Å), with His-234 itself coordinating via Ne2 to the side chain and carbonyl oxygen of Ser-165 (2.9 Å/ 3.1 Å). This implies that there is no abstractable proton on His-5 close to the haem unless this residue is protonated (*vide infra*).

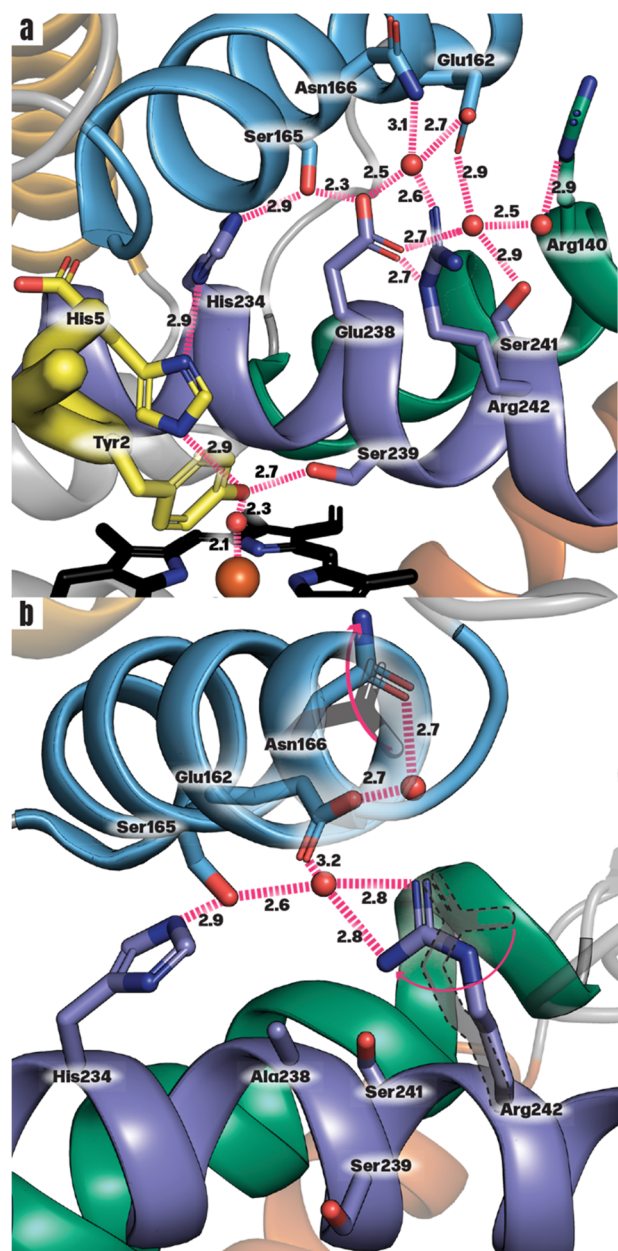
**Assessment of Peptide Binding Using Molecular Dynamics Simulations.** Given that the structure of the MRYLH peptide in the P450<sub>Blt</sub> active site appeared to be competent for catalysis without substantial rearrangement, we next performed molecular dynamics simulations to evaluate the stability of the complex observed in the crystal structure and to predict the formation of an intermolecular interaction over time. To this end, we monitored the stability of the complexes by means of root-mean-square deviation (RMSD), root mean fluctuation (RMSF), and protein–ligand contact mapping, extracted from the respective 200 ns simulation trajectories. These experiments supported the stability of the complex observed, which allowed us to analyze the interactions observed in this structure with greater confidence. The RMSD of the MRYLH-P450<sub>Blt</sub> complex is shown in SI Figure S2, indicating that equilibrium is reached after the initial 25–50 ns of the simulation, after which the RMSD value of the MRYLH-P450<sub>Blt</sub> complex showed deviations less than 1 Å. This suggests that the complex of MRYLH-P450<sub>Blt</sub> is stable during simulations without major structural perturbations. The ligand root mean square fluctuation (RMSF) characterizes changes in the ligand atom positions and showed RMSF values less than 3 Å with the highest RMSF deviations observed at atoms associated with Arg-2 and Leu-4. This is in accordance with these residues being oriented away from the P450<sub>Blt</sub> pocket in addition to the flexible nature of Arg side chains.

To profile protein–ligand interactions, we analyzed the MRYLH-P450<sub>Blt</sub> complex over a 200 ns simulation interval, focusing on hydrogen bonding, hydrophobic interactions, ionic interactions, and water bridge formation (SI Figure S3). The MRYLH-P450<sub>Blt</sub> complex exhibited several notable features. In more than 93% of frames for each of the triplicate simulations, the Tyr-3 phenol of MRYLH was hydrogen bonded to the hydroxyl moiety of the Ser-239 residue in the I-helix, which is the central alcohol residue linked to oxygen activation in P450s. Additionally, the amide nitrogen of the Met-1/Arg-2 peptide bond was hydrogen bonded to the backbone carbonyl of Met-172 in the F-helix (>86% of the simulation time). The amide nitrogen of the Arg-2/Tyr-3 peptide bond was found to hydrogen bond to Asp-284 in >76% of the simulation time in the three simulations, which also participated in a water bridge

to the carbonyl oxygen of the peptide bond between Tyr-3/Leu-4 (>60% of the simulation time in 2 of 3 simulations). The possible sensor residue Gln-84 was found to hydrogen bond with the amide nitrogen of the Leu-4/His-5 peptide bond (>54% of the simulation time in 2 of 3 simulations), and to interact directly or via water mediated interactions with the carbonyl oxygen of the peptide bond between Tyr-3/Leu-4. Arg-230 and Arg-188, implicated in binding the peptide carboxylate, displayed some of the most consistent interactions to the ligand across the simulations, averaging two contacts across the trajectory, with both residues engaging in direct hydrogen bonding and water mediated interactions. However, most of the interactions with the MRYLH carboxylate were by hydrogen bonding from Arg-230 as well as a small fraction of salt bridging interactions. Although Arg-2 of the MRYLH peptide was uncoordinated in the crystal structure, MD simulations suggest that the P450<sub>Blt</sub> residue Phe-69 coordinates Arg-2 through a Pi-cation interaction (>35% of the simulation time in 2 of 3 simulations). These simulations thus reveal that the binding of the MRYLH peptide is dynamic and is characterized by a limited number of interactions between the peptide and P450<sub>Blt</sub>. Instead, numerous direct and indirect interactions facilitated by a complex water network form a highly plastic interaction network responsible for peptide coordination. This analysis is also in agreement with the low spin state change previously demonstrated for the binding of the MRYLH peptide to P450<sub>Blt</sub>, which supports a limited fraction of substrate P450 complexes existing in a catalytically competent state at any one time.<sup>17</sup>

**Clarifying the Role of Polar I-Helix Residues in Peptide Cross-Linking.** Given the information provided by the MD simulations about the orientation of the substrate within the active site of P450<sub>Blt</sub>, we next turned to the analysis of residues central in coordinating the peptide above the haem (by comparing the bound chain A with the unbound chain B). A polar cavity is located above and parallel to the I-helix, which is defined by several residues, including Glu-238, Ser-241, and Arg-242 from the I-helix, Glu-162, Ser-165 and Asn-166 from the F-helix, and Arg-140 from the E-helix (Figure 4). This cavity also contains several ordered waters, including water-511 (2.5 Å) and water-558 (2.7 Å) that are hydrogen bonded to Glu-238 that is part of the acid-alcohol pair. Neither Glu-238 nor Arg-242 are accessible to the active site, being shielded by Val-169 and Leu-382. Notably, Glu-238 forms a short 2.3 Å hydrogen bond with the hydroxyl moiety of Ser-165, indicating a single-well hydrogen bond where the proton is shared symmetrically and can be easily transferred from one residue to the other.<sup>22</sup> To explore the importance of this relay, we designed and expressed two mutants (Glu-238 to Asn or Ala) with the anticipation that the former would be inactive and that the latter would display reduced activity. While the E238N mutant was insoluble (and hence unable to be purified), the E238A mutant displayed 1/3 of the activity of the WT enzyme and an alteration of the P450 binding spectrum (Table 2, Figure 5). Structural characterization of the E238A mutant (1.95 Å, PDB 8UKZ, Table 1) revealed that the proton transfer pathway is restored by the rotation of the Arg-242 side chain, which adopts a new rotamer in the E238A mutant such that the two  $\eta$ -nitrogen atoms within the guanidine moiety hydrogen bond (2.8 Å) to the same catalytic water molecule as Ser-165 (2.6 Å), mirroring the position of water-511 in the WT structure. Thus, the E238A mutation can maintain reduced activity due to the ability of Arg-242 to reconstitute





**Figure 4.** Proton transfer in P450<sub>Blt</sub> active site structures. a) Wild-type P450 structure detailing the proton relay network with key residues and catalytic waters represented as sticks and red spheres, respectively. Dashed light magenta lines show hydrogen bonds with distances indicating the proton transfer pathway. The MRYLH peptide is depicted as a yellow tube, with only Tyr-2 and His-5 shown in stick form due to their ability to participate in the proton transfer network. b) E238A mutant structure, where the absence of Glu-238 results in the rotation of the Arg-242 side chain. This rotation allows Arg-242 to form hydrogen bonds with a catalytic water molecule bridged to Ser-165, compensating for the disrupted original relay. Hydrogen bond distances are again depicted, illustrating the maintenance of a proton transfer pathway, despite the impact of the mutation on proton transfer efficacy. Wildtype Arg-242 and Asn-166 rotamers are indicated as black dashed outlines.

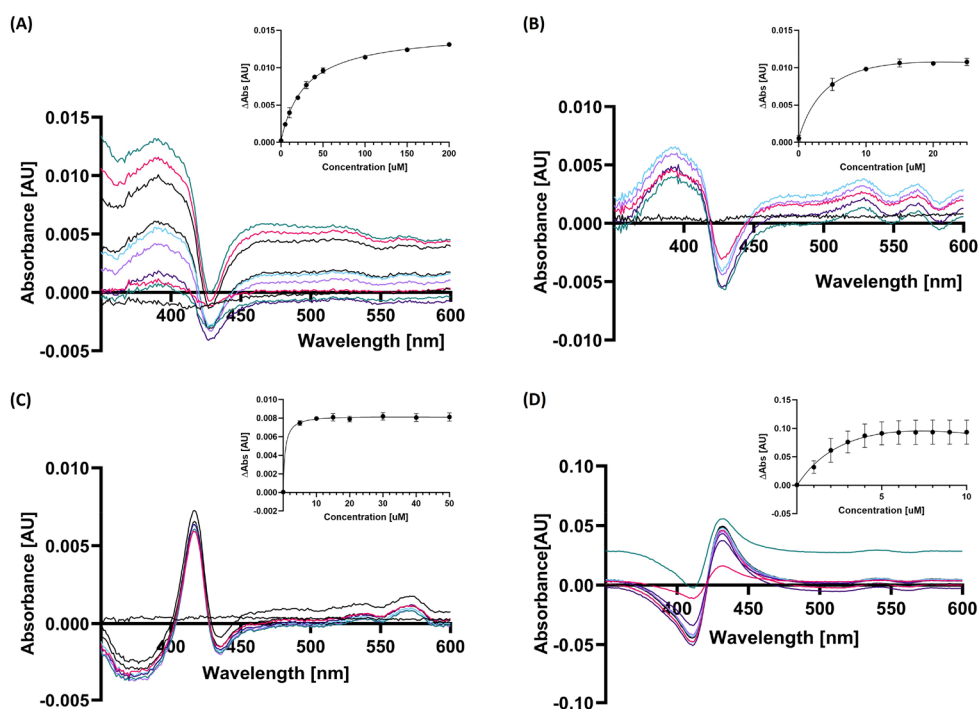
the proton transfer pathway, albeit with a significantly impaired proton transfer.

Beyond interacting with Glu-238, Ser-165 is also found in the hydrogen bonding distance of His-234 (2.9 Å), forming a potential proton relay network from the hydrated cavity to the

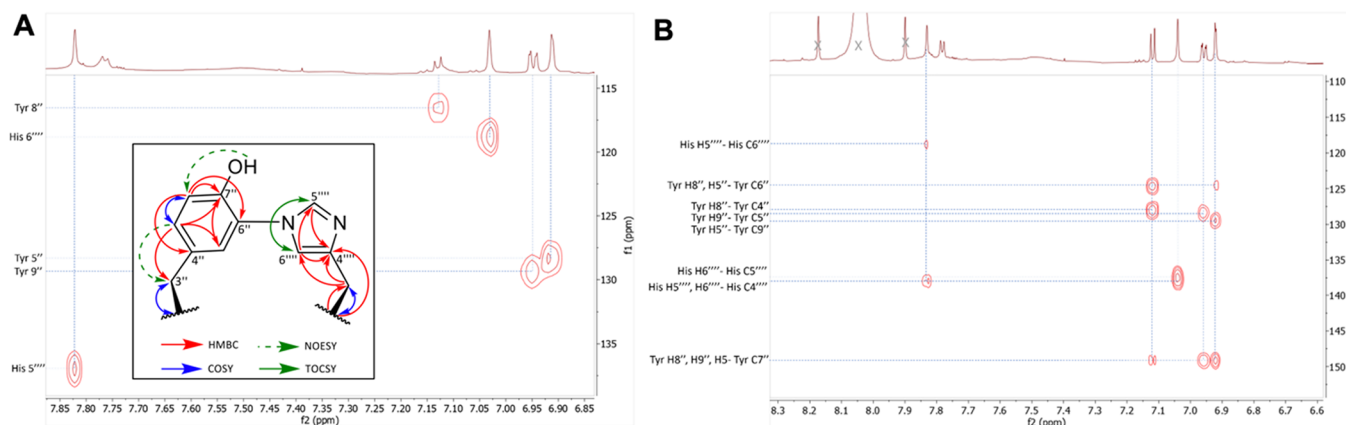
substrate. This network is likely critical not only for catalysis but also for the initial binding of the MRYLH peptide, considering the need to protonate His-234 to enable hydrogen bonding with the His-5 side chain of the MRYLH substrate peptide. Indeed, such hydrogen bonding interactions are comparable in strength to salt bridge and cation- $\pi$  interactions.<sup>24</sup> To probe the importance of this interaction for catalysis, a H234L P450<sub>Blt</sub> mutant was generated and the binding and turnover of the MRYLH substrate explored (Table 2, Figure 5). The importance of His-234 for catalysis was confirmed by an order of magnitude reduction in cross-linking demonstrated by the H234L mutant (>9% vs 85% for WT), and can be explained by the altered binding mode seen with this mutant, which switches to an inhibition-type spectrum that is explicable either via direct or water-mediated His-coordination to the haem of P450<sub>Blt</sub> due to the loss of the interaction with His-234.<sup>25</sup> The marked effect of the H234L mutation is in contrast with the more minor effects of mutating the major nonpolar I-helix residue that interacts with the side chains of the MRYLH peptide (Ala-231) in an A231 V mutant (53% conversion), showing the importance of His-234 for both binding and turnover of the MRYLH substrate.

**Confirmation of C–N Peptide Cross-Link between Tyr and His.** Next, to confirm the nature of the cross-link installed by P450<sub>Blt</sub> we prepared a sample of the Nle-RYLH peptide that had been cross-linked by P450<sub>Blt</sub>, purified this compound, and then performed extensive NMR experiments (1-D <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N NMR and 2-D COSY, TOCSY, HSQC, HMBC, NOESY and <sup>15</sup>N-HSQC). It has previously been demonstrated that substituting the terminal Met residue with an Nle in the peptide resulted in a substrate that was well tolerated. This change is advantageous as it prevents formation of unwanted and complicating sulfoxides and greatly simplifies purification.<sup>17</sup> Utilization of a polar aprotic solvent was preferred to prevent proton exchange. Poor spectral resolution was observed using DMSO-*d*<sub>6</sub>, particularly in the key aromatic region and this was only partially overcome by employing variable temperature experiments. Spectral quality significantly improved at room temperature by utilization of DMF-*d*<sub>7</sub>. This solvent has proved advantageous in prior studies of marine cyclic peptides.<sup>26</sup> Initial observation of the proton spectrum allowed for identification of key spectral features (Figure 6). The five expected  $\alpha$ -protons were identified at  $\delta_{\text{H}}$  4.26, 4.53, 4.71, 4.80, and 4.91. Also identified were amide NH protons at  $\delta_{\text{H}}$  9.01, 7.78, 8.50, and 8.75 (all doublets) and one Arg guanidino NH triplet at  $\delta_{\text{H}}$  8.43. The corresponding nitrogen resonances were identified by cross peaks in the <sup>15</sup>N-HSQC. The N-terminal amino group and the remaining guanidino NH protons were not assigned but are likely responsible for the broad baseline peaks observed in the downfield region. Careful examination of the <sup>13</sup>C-HMBC spectrum revealed the five amide carbonyl carbons at  $\delta_{\text{C}}$  169.5, 170.5, 169.9, 172.4, and 173.0. These features were consistent with the expected Nle-RYLH pentapeptide and full structural assignment was then undertaken, including side chains, using 1-D <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N NMR and 2-D COSY, TOCSY, HSQC, HMBC, NOESY and <sup>15</sup>N-HSQC experiments (SI Table S3).

Determination of the nature of the installed Tyr-His peptide cross-link required careful examination of the aromatic region of the various spectra. A hydroxyl proton at a characteristic  $\delta_{\text{H}}$  10.50 ppm was consistent with the presence of the free Tyr phenolic group (HMBC cross peak to C7'  $\delta_{\text{C}}$  149.2) and excluded the possibility of C–O bond formation. An aromatic



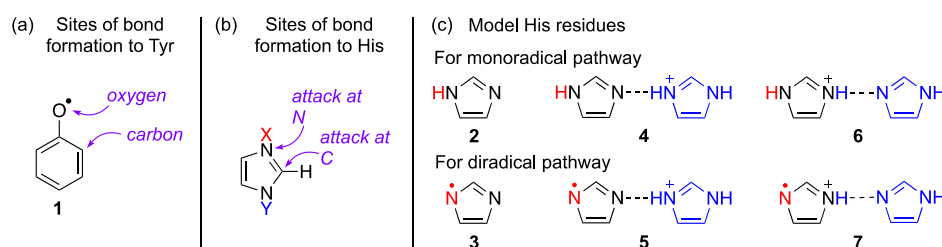
**Figure 5.** UV–visible spectroscopic investigation of the binding of P450<sub>Blt</sub> mutants to the MRYLH–OH substrate. The spectral response of different mutants to varying concentrations of MRYLH–OH with the calculated amplitude of binding (inset) shown for the Q84L (A), A231 V (B), H234L (C) and E238A mutants (D). Amplitudes of the different spectra were extrapolated ( $\Delta A = A_{\max} - A_{\min}$ ) and plotted against the substrate concentration.  $\Delta A_{\text{abs}}$  and  $K_d$  derived from the one-site binding model,  $n = 3$  (see Table 2).



**Figure 6.** Selected NMR spectra used to confirm the C–N cross-link formed by P450<sub>Blt</sub>. Aromatic region of 700 MHz NMR HSQC spectrum of Nle-RYLH using  $^1J_{\text{H–C}} = 170$  Hz (A), with Key HMBC, COSY, TOCSY and NOESY correlations used to establish cross-link between Tyr and His residues (inset). Aromatic region of 700 MHz NMR HMBC spectrum of Nle-RYLH (B).

ABX system was identified, consistent with a trisubstituted Tyr residue. The H8'' doublet at  $\delta_{\text{H}} 7.12$  ( $\text{C8''}$   $\delta_{\text{C}} 116.4$ ) demonstrated a typical 8.2 Hz *ortho* coupling to the neighboring  $\delta_{\text{H}} 6.96$  doublet of doublets (HSQC C9''  $\delta_{\text{C}} 129.2$ ), which demonstrated a further 2.3 Hz *meta* coupling to the doublet at  $\delta_{\text{H}} 6.92$  (HSQC C5''  $\delta_{\text{C}} 128.1$ ). The C8'' chemical shift of  $\delta_{\text{C}} 116.4$  is consistent with an aromatic methine group adjacent to a hydroxylated position, placing  $\delta_{\text{H}} 6.96$  at H9'' and  $\delta_{\text{H}} 6.92$  at H5''. These assignments are supported by observed COSY, HMBC and NOE correlations (Figure 6, SI Table S3). Further, both H5'' and H9'' demonstrated HMBC correlations to the Tyr-C3'' methylene group at  $\delta_{\text{C}} 36.0$ . This substitution pattern strongly suggests the participation of Tyr-C6'' in the cross-link to the His residue.

No HMBC correlation was observed between C5''/H5'' and the imidazole residue and so efforts were focused on identification of the structure of the His residue. A clear proton resonance for H6''' was identified at  $\delta_{\text{H}} 7.04$  (s, br,  $\delta_{\text{C}} 118.4$ ) by HMBC correlation from H3''' (2.85–2.87, m and 3.22–3.25, m). A proton resonance was identified at  $\delta_{\text{H}} 7.83$ , and this resonance appeared significantly broadened under some experimental conditions suggesting sensitivity to minor changes in pH. A 2D-TOCSY correlation was observed between the  $\delta_{\text{H}} 7.83$  proton and H6''' ( $\delta_{\text{H}} 7.04$ ). Examination of the HSQC spectrum indicated a weak and out of phase correlation from  $\delta_{\text{H}} 7.83$  suggesting a carbon signal at  $\delta_{\text{C}} 136.8$ . To confirm the identity of this signal a further  $^1\text{H}$ – $^{13}\text{C}$  HSQC was performed with a narrowed  $^{13}\text{C}$  spectral width focused so as to identify aromatic resonances and optimized for 170 Hz



**Figure 7.** Different reactive intermediates, sites of attack, and hydrogen bonding (or protonation) states explored by means of density functional theory calculations.

**Table 3. Relative Barrier Heights ( $\Delta G_{\text{rel}}^\ddagger$ , kcal/mol) for Different Tyr–His Bond-Forming Pathways**

Pathway	Monoradical			Diradical		
Imidazole protonation or hydrogen bonding	None (2)	Hydrogen bonded, reacting His unprotonated (4)	Hydrogen bonded, reacting His protonated (6)	None (3)	Hydrogen bonded, reacting His unprotonated (5)	Hydrogen bonded, reacting His protonated (7)
C(Tyr)–N <sub>1</sub> (His)	6.2	–	–	6.9	4.7	8.2
C(Tyr)–C <sub>2</sub> (His)	8.1	8.3	10.0	0	2.6	0
O(Tyr)–C <sub>2</sub> (His)	0	0	0	2.2	0	0.1

$^1\text{H}$ – $^{13}\text{C}$  coupling, expected for an  $sp^2$  C–H coupling. A corresponding carbon resonance was identified at  $\delta_{\text{C}}$  136.8 and assigned as C5''' with H5''' at  $\delta_{\text{H}}$  7.83 (Figure 6). This assignment was additionally supported by an HMBC correlation observed between H6''' and C5''' (Figure 6). A further  $^{15}\text{N}$  HSQC experiment was performed using similarly focused parameters, and no further nitrogen resonances were identified as would have been expected if an imidazole NH was indeed present. The position of the cross-link was therefore confirmed to be between His-C5'''–N and Tyr-C6''.

**Analysis of the Mechanism of Tyr–His Cross-Linking by P450<sub>Blt</sub>.** To investigate whether the hydrogen bonding between His-5 of MRYLH and His-234 of the enzyme influences the nature of the cross-link formed, we performed density functional theory calculations (Figure 7). A phenoxyl radical (1) was used to model the intermediate tyrosyl radical formed by hydrogen abstraction from MRYLH by Compound I. Imidazole was used to model the histidine of MRYLH. A variety of possible reaction pathways were evaluated as follows. First, three types of cross-links were considered (Figure 7a,b): (i) a C–N cross-link, as observed with P450<sub>Blt</sub>, (ii) a C–C cross-link, as observed with P450 enzyme BytO from *Planomonospora* sp. that lacks the His-234 residue,<sup>15</sup> and (iii) an O–C cross-link. Second, two chemically different reaction mechanisms were considered: (i) a monoradical pathway in which the tyrosyl radical adds to the histidine of MRYLH, and (ii) a diradical pathway in which a second hydrogen abstraction leads to an intermediate containing tyrosyl and histidiny radical which then undergo coupling. Finally, the role of hydrogen bonding to His-234 was explored by modeling the reacting imidazole or imidazolyl radical (Figure 7c) as being (i) neutral (2, 3), hydrogen bonded to a model imidazolium cation (4, 5), or protonated and hydrogen bonded to a neutral imidazole (6, 7). In the latter two cases, imidazolium or imidazole mimics the role of His-234.

The activation barriers for the different reaction pathways are listed in Table 3. The computations predicted that in all but one case, the preferred mode of cross-linking was the formation of an O–C cross-link. In these cases, and depending on the exact system, O–C cross-link formation was favored by 2.6–10 kcal/mol relative to other cross-linking modes. The exceptions were the diradical pathway in the absence of any

hydrogen bonding (3) and where the protonated histidiny radical is hydrogen bonded to a histidine (7). In these two cases, C–C bond formation was either preferred (by 2 kcal/mol, for 3) or equienergetic compared to C–O bond formation (for 7). The experimentally observed coupling mode (C–N coupling) in most cases was higher in energy by 5 kcal/mol or more than the preferred coupling mode. This represents a sizable energy penalty, which the enzyme would need to overcome to achieve selective C–N cross-link formation. Based on the small model systems calculated here, the formation of a C–N bond appears to be most accessible if it proceeds via a diradical mechanism where the histidiny radical is hydrogen bonded to a histidiny cation (5). In such a reaction, the hydrogen bonding to the histidiny cation provides a 2 kcal/mol benefit to the formation of a C–N cross-link, compared to the corresponding system without hydrogen bonding (3). However, C–N coupling is not inherently preferred and therefore additional substrate- or enzyme-controlled factors (e.g., conformational restriction or specific enzyme–substrate noncovalent interactions) would be required to achieve the preference for C–N bond formation observed experimentally. In the absence of the hydrogen bond to a histidinium ion, C–C cross-linking is preferred instead.

These results are broadly consistent with the experimental observations that P450<sub>Blt</sub> generates a C–N cross-link, but the P450 enzyme BytO from *Planomonospora* sp.,<sup>15</sup> which lacks a His-234 residue, leads to the formation of a C–C cross-link. To investigate the role of this specific residue in *Planomonospora* P450 BytO, we mutated the sequence encoding the corresponding residue Val-219 to His and expressed the mutated *bytAO* in *Streptomyces coelicolor* M1152. This mutant, V219H, exhibited a significant reduction in activity, yielding only 0.6% formation of biaryllyde YYH compared with native BytAO, further underlining the importance of this residue. Additionally, when we mutated Gln-216, the counterpart of the dominant nonpolar I-helix residue Ala-231, to Ala, we observed product formation of 7.6% compared to native BytO, indicating an important but not essential role of Gln-216 in *Planomonospora* BytO (SI Figure S4). This is consistent with the observation of reduced, but not totally abolished conversion of the mutant A231 V in P450<sub>Blt</sub> (SI Table S5).



### Structural Comparisons with Other P450 Enzymes.

Finally, we turned to the analysis of the structure of P450<sub>Blt</sub> in comparison to other P450 structures. The three most similar P450 structures in the Protein Data Bank as ranked by the DALI server<sup>27</sup> were P450 EryF (1Z8P; 2.4 Å RMSD over 369 Cα atoms and 30% sequence identity; A254S mutant, macrolide hydroxylase),<sup>28</sup> CreJ (5GWE; 2.2 Å RMSD over 367 Cα atoms and 25% sequence identity; alkylphenol hydroxylase)<sup>29</sup> and XplA (2WIV; 2.5 Å RMSD over 370 Cα atoms and 21% sequence identity; explosive degradation).<sup>30</sup> However, P450<sub>Blt</sub> shares few similarities in the active site, even when including the typically conserved acid/alcohol residue pair (Glu-238/Ser-239 in P450<sub>Blt</sub>). P450 EryF catalyzes the hydroxylation of 6-deoxyerythronolide B, and while this shares little similarity with the MRYLH peptide, the macrolide 5-OH group is also coordinated to the haem iron and the normally conserved I-helix alcohol residue (Ser-245 in this P450 EryF A254S mutant) in a similar manner as is seen with the Tyr-3 phenol of the biaryllyl peptide substrate. While the phenol of the biaryllyl Tyr-3 residue is in hydrogen-bond distance to the Oγ of Ser-239 the rotamer of the corresponding residue in P450 EryF is oriented away from the substrate, which is anticipated as the role of this serine in P450 EryF is replaced by a hydroxy group within the substrate of this P450.<sup>31,32</sup> XplA also possesses major alterations at these key positions (394-Met/395-Ala), likely due to the different mechanism of this unusual P450. While CreJ contains the conserved acid/alcohol pair (295-Glu/296-Thr), the substrate binding pocket is much smaller and is largely hydrophobic to coordinate the small aromatic substrate of this degradative enzyme.<sup>29</sup>

Given that P450 structures often share high similarity simply by virtue of the highly conserved fold, we next identified structures with similarities in substrate coordination by aligning the sequences of all PDB entries to identify P450s with similar substrates or a mode of substrate binding. Two structurally characterized P450s were identified that shared the equivalent residue to Gln-84, which is one of the few residues to respond to substrate binding in P450<sub>Blt</sub>. These enzymes – AurH (3P3L, 3P3Z (3.1 Å RMSD over 367 Cα atoms and 24% sequence identity), 3P3X, 3P3O)<sup>33</sup> and OxyE (3O1A<sup>8</sup>/3O03,<sup>34</sup> (2.8 Å RMSD over 351 Cα atoms and 24% sequence identity)) – are both involved in the assembly of ring systems: AurH catalyzes the stepwise formation of a homochiral oxygen heterocycle (tetrahydrofuran) that is a key structural and pharmacophoric component of the polyketide antibiotic aureothin,<sup>33</sup> while OxyE is responsible for the phenolic coupling of the aromatic side chains of the first and third peptide residues (F-O-G ring) in the heptapeptide teicoplanin.<sup>8,34,35</sup> While AurH does not accept a peptide substrate, Zocher and colleagues<sup>33</sup> observed a similar movement of Gln-91 (in AurH) when bound to the competitive inhibitor ancymidol via formation of a hydrogen-bond to a nearby serine residue (Ser-66). The two deposited structures for OxyE are both without the cognate heptapeptide substrate<sup>8,34</sup> (that is also carrier protein bound and requires an adjacent P450-recruitment (X) domain),<sup>11,35</sup> but the conserved Gln-80 residue in OxyE adopts a similar rotamer to that seen with Gln-84 in the unbound P450<sub>Blt</sub> structure. This suggests that such glutamine residues can play a role as a sensor in substrate binding and allow these states to be identified based on the rotameric state of these residues for specific biosynthetic P450s.

While no structurally characterized P450s possess the same Arg-188/Arg-230 pair as observed in P450<sub>Blt</sub>, some similarities in the use of I-helix Arg residues for substrate coordination can be found in the CYP152 family, which rely on peroxide and utilize such residues to coordinate to the carboxylate moiety of their fatty acid substrates and present the carbons close to this group for hydroxylation.<sup>36,37</sup> Indeed, these Arg residues in CYP152 enzymes are found in the same place in the I-helix as His-234 in P450<sub>Blt</sub>, which in turn is also linked with a comparable Gln as the sensor Gln-84 P450<sub>Blt</sub> residue via an intermediate I-helix Asn residue. Curiously, Arg-280 is conserved in both CYP199A2 (4DNJ, Arg-246) and CYP199A4 (4DO1, Arg-243), which are highly selective for para-substituted benzoic acids as their substrates.<sup>38</sup> In CYP199A2/4 the I-helix Arg-246/3 residue interacts with the carboxylate via a water mediated interaction in conjunction with Arg-94/2 from the BC-loop that is directly hydrogen-bonded to the substrate carboxylate;<sup>38</sup> this is comparable to the situation for P450<sub>Blt</sub> as noted in the MD simulations. CYP121, which is involved in the cross-linking of cyclo-dityrosine, also has an Arg residue (Arg-386) that is closely involved in water-mediated interactions with the substrate, although in this case the residue projects from the C-terminal loop region of the P450.<sup>21</sup>

Outside of P450s sharing specific active site residues with P450<sub>Blt</sub>, the closest structural similarities were found with P450s that oxidize large substrates, including a range of macrolide modifying P450s (EryK,<sup>39</sup> PikC,<sup>40</sup> MycG,<sup>41</sup> TamI<sup>42</sup>), as well as to those involved in oxidizing ACP-bound substrates (such as P450<sub>Biol</sub><sup>43</sup> and CalO2<sup>44</sup>). When comparing P450<sub>Blt</sub> with P450s that perform comparable (aromatic cross-linking) reactions (such as the CYP158 enzymes involved in flavin biosynthesis)<sup>45,46</sup> and/or that accept peptide-derived substrates such as diketopiperazines (DKPs), several enzymes related to diketopiperazine oxidation were identified, including CYP121 (responsible for the cross-linking of cyclodityrosine<sup>21</sup> together with further possible reactions<sup>47</sup>), TxtC (aromatic and aliphatic hydroxylation),<sup>48</sup> CypX (N-oxidation and aromatization),<sup>49</sup> P450s involved in C–N cross-linking in indolactam biosynthesis<sup>50</sup> as well as dimerization via C–N bond formation (AspB).<sup>51</sup> Overall, these studies have shown that the P450 scaffold is capable of multiple binding modes for such DKP (or related)<sup>52</sup> substrates, which are able to thus undergo multiple reactions and different sites (such as TxtC) or even dimerization of such substrates (AspB),<sup>51</sup> which was recently shown to proceed via a ferric-superoxide intermediate.<sup>53</sup> In terms of the acceptance of larger peptide substrates by P450s, the structure of P450<sub>Blt</sub> represents the first to be solved for a P450 bound to a linear peptide substrate. Comparison of this structure to that of the Oxy enzymes (A,<sup>6</sup> B,<sup>54</sup> C<sup>9</sup> and E<sup>8,34</sup>) from GPA biosynthesis shows how P450<sub>Blt</sub> requires a much more structured and enclosed active site, with the Oxy enzymes all displaying a highly open structure that is presumably due to the required interface with the PCP-domain that presents the peptide to these P450s.<sup>10,11</sup> The direction of the peptide in P450<sub>Blt</sub> running parallel to the I-helix generally matches the expected orientation of the peptide substrate in such Oxy enzymes,<sup>11,55</sup> which match well to the PCP-binding mode seen in the only structure of a PCP-bound P450 solved to date (P450<sub>sky</sub> from skyllamycin biosynthesis).<sup>56</sup> In terms of active site elements, the Oxy enzymes are largely formed from residues with hydrophobic side chains, a feature that is also seen in the

arylomycin cross-linking P450 enzyme AryC,<sup>5</sup> although in this case the P450 is able to function on soluble substrates. However, slight changes to the orientation of the peptide bound to the Oxy enzymes is likely to be expected due to the enlarged oxygen binding groove seen in the I-helices of such enzymes (and that is also seen in AryC). One somewhat curious feature is the replacement of the apparent sensor Gln residue in OxyB with a methionine residue,<sup>7,11</sup> although it is unclear yet what role such a residue could play given the lack of peptide-bound structure for GPA-cross-linking P450 enzymes. The structure of P450<sub>Blt</sub> also demonstrates how the binding mode for the aromatic rings within the substrates parallel to the haem as seen for other well characterized cross-linking P450s such as the CYP158 family enzymes<sup>45,46</sup> is not possible within the confines of side chains linked within a linear peptide backbone, with the P450<sub>Blt</sub> substrate presenting Tyr-3 and His-5 side chains rather perpendicular to the haem plane. This is also in agreement with computational studies investigating the first cross-linking step in GPA biosynthesis performed by OxyB.<sup>12</sup>

## CONCLUSIONS

In this work, we have structurally characterized a biaryllytite cross-linking P450<sub>Blt</sub> in complex with its substrate, the pentapeptide MRYLH.<sup>15,17</sup> In doing so, we have obtained further insights into how P450s control their acceptance of diverse substrates and how residues central to this process can also aid in the specificity of these enzymes for distinct cross-linking outcomes. The residues involved in P450<sub>Blt</sub>—almost all located in the central I-helix that also contains the conserved acid/alcohol pair—form crucial interactions with the C-terminus of the peptide, while also orienting the Tyr/His residues to be cross-linked in the substrate peptide. The importance of Arg residues in coordinating the C-terminus of the peptide substrate, previously implied by biochemical experiments,<sup>17</sup> is clarified in the structure of P450<sub>Blt</sub>. Indeed, comparison of this structure with the sequence alignments of this enzyme and those accepting C-terminally extended biaryllytite peptides<sup>57</sup> shows that these Arg residues present in P450<sub>Blt</sub> are replaced by residues with far smaller side chains in enzymes with extended substrates, thus allowing them to accept longer heptapeptide sequences.

Turning to other residues within the I-helix, His-234 appears to play a central role in controlling the C–N cross-linking outcome mediated by P450<sub>Blt</sub>, which otherwise would be likely to favor formation of a C–C bond as has been shown for other biaryllytite P450 enzymes.<sup>15</sup> Comparison of these central residues across biaryllytite cross-linking P450s more broadly generally shows limited diversity in these positions (Ala-231 to Val, Ser, Thr, Met or Leu; His-234 to Leu, Ile, Asn or Gln),<sup>57</sup> although further studies are required to better understand how the control of cross-linking outcomes can be controlled through such interactions. The acid/alcohol pair Asp-238/Ser-239 found in P450<sub>Blt</sub> suggests that oxygen activation proceeds via the canonical P450 catalytic cycle in these enzymes,<sup>3</sup> and furthermore that despite the prevalence of hydrogen bonding in such peptides that the replacement of these crucial interactions via those mediated by the substrate are not sufficient for these P450s.<sup>31,32,58</sup> The conservation of residues in these acid/alcohol positions that can form hydrogen bond interactions in biaryllytite P450s also suggests that recent reports of P450 cross-linking enzymes using a superoxide anion intermediate for catalysis<sup>53</sup> are unlikely to be

realized for such P450s, although definitive experiments to address this question remain to be performed. Beyond their classic roles in the P450 catalytic cycle, the P450<sub>Blt</sub> complex structure also shows that these residues perform further roles, including mediating crucial interactions with the substrate peptide (Ser-239 to the Tyr-3 phenol) or Asp-238 in aiding in the maintenance of a highly ordered water network likely responsible for acting as a proton delivery channel to the active site.

Comparing the sequences of recently reported His-His RiPP cross-linking P450s<sup>57,59</sup> that are closely related to the biaryllytite Tyr-His cross-linking enzymes is also highly revealing when interrogated through the lens of the P450<sub>Blt</sub>/MRYLH complex. Such an analysis shows that the only major difference in the active site of His-His cross-linking P450s compared to Tyr-His cross-linking P450s is the C-terminal loop residue Leu-382, which is replaced by a Glu in His-His P450s and likely plays a role in coordinating to the His-3 residue present in the altered substrate of these enzymes (MRHxH).<sup>57,59</sup> While the I-helix residues Ala-231/His-234 from P450<sub>Blt</sub> are replaced by Val and Gln in His-His cross-linking enzymes, these in fact match those found in the initially reported biaryllytite C–C bond forming enzyme from *Planomonospora*<sup>15</sup> and are therefore also in agreement with the lack of an N-cross-link in the reported product from such His-His cross-linking enzymes.<sup>59</sup> The intriguing two-step reaction performed by these P450s (presumably initial hydroxylation followed by cross-linking, which would be in agreement with our computational results concerning the energetics of such reactions) yet again demonstrates the versatility of P450s, and furthermore demonstrates that there is no restriction to cross-linking and hydroxylation chemistry enforced by P450s generally (which has also been recently demonstrated by work in engineering an aarylomycin cross-linking P450).<sup>60</sup> This also implies that the exquisite specificity demonstrated by P450s involved in cross-linking over classic hydroxylation chemistry (such as is found in GPA biosynthesis)<sup>4</sup> is therefore something that must be controlled carefully by these enzymes through the appropriate positioning of the substrate peptide. Finally, the ability to interrogate the divergent cross-linking activity of biaryllytite related cross-linking P450s with reference to the structure of the P450<sub>Blt</sub>/MRYLH complex reminds us of the importance of structural characterization of important biosynthetic enzymes to addressing reaction specificity in such systems. It also shows that despite the largely conserved fold of cytochrome P450s, many valuable insights into the nuances of their reactivity and specificity can be obtained through the structural elucidation of key complexes of these fascinating biocatalysts.

## EXPERIMENTAL SECTION

**Chemicals and Reagents.** 2-Chlorotriyl chloride resin (GL Biochem), DCM (Chem-supply), DMF (Ajax Finechem), methanol (Scharlau), TFA (Oakwood chemical), TIPS (Sigma-Aldrich), diethyl ether (Sigma-Aldrich), Oxyma Pure (Novabiochem), DIC (Sigma-Aldrich), KH<sub>2</sub>PO<sub>4</sub> (Sigma-Aldrich), K<sub>2</sub>HPO<sub>4</sub> (Sigma-Aldrich), tryptone (MP Biochemicals), yeast extract (MP Biochemicals), SIGMAFAST protease inhibitor cocktail tablets (EDTA-free; Sigma-Aldrich), formic acid (Sigma-Aldrich), HEPES (Sigma-Aldrich), IPTG (Promega), imidazole (Sigma-Aldrich), Tris-HCl (Sigma-Aldrich),  $\delta$ -aminolevulinic acid (Sigma-Aldrich), NaCl (Sigma-Aldrich), MgCl<sub>2</sub> (Sigma-Aldrich), glucose (Sigma-Aldrich), glucose

dehydrogenase (Sigma-Aldrich), NADH (Sigma-Aldrich), kanamycin sulfate (Sigma-Aldrich), gentamycin sulfate (Sigma-Aldrich), 5- $\alpha$  competent *E. coli* cells (NEB), ArcticExpress (DE3) competent cells (Agilent), Ni-NTA agarose (Macherey-Nagel), Precision Plus Protein Dual Color Standards (Bio-Rad), MES (Sigma-Aldrich), PEG 400 (Sigma-Aldrich), PEG 20k (Sigma-Aldrich), methylerythritol (Sigma-Aldrich), SPE Bond elute (Agilent), C18 Strata-X-SPE-cartridges (Phenomenex).

Peptide analysis was conducted on a Shimadzu HPLC-MS system (LCMS-2020, ESI operating in positive and negative mode; or LCMS-8050, ESI operating in positive and negative mode). The mobile phases used were water +0.1% FA and ACN + 0.1% FA for analytical runs. P450-mediated cyclization analyses were conducted using HRMS on an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled to a Dionex UltiMate 3000 RSLCnano system.

**Peptide Synthesis.** Peptide synthesis was performed as automatically as previously described, using a Liberty Blue Automated Microwave Peptide Synthesizer (CEM) at a 0.1 mmol scale.<sup>17</sup>

**Plasmid Construction.** The P450<sub>Blt</sub> gene was synthesized with codon optimization for *E. coli* and ligated into a pET28a(+) vector (Twist Bioscience) following a previously reported protocol.<sup>17</sup>

**Generation of P450 Mutants.** Site-directed mutagenesis was performed to generate several P450<sub>Blt</sub> mutants according to an optimized Quikchange Site-Directed Mutagenesis (Agilent Technologies) protocol. To generate the mutants, pairs of primers were designed, and the vector amplified by PCR (SI Table S1). High fidelity PCR was carried out with 12.5  $\mu$ L of Phusion High Fidelity DNA Polymerase MasterMix (NEB), 1  $\mu$ L of 10  $\mu$ M P450<sub>Blt</sub> plasmid as the template, 1.25  $\mu$ L of 10  $\mu$ M forward and reverse primers, and 0.75  $\mu$ L of DMSO and nuclease-free water to make up a 25  $\mu$ L reaction. The PCR reaction was then placed in a T100 thermocycler (Bio-Rad). An initial denaturation step (30 s, 98  $^{\circ}$ C) was then followed by 32 cycles of denaturation (10 s, 98  $^{\circ}$ C), annealing (30 s, 60  $^{\circ}$ C) and elongation (30 s/1000 bps, 72  $^{\circ}$ C). After completion of the cycles a final elongation step was performed (2 min, 72  $^{\circ}$ C). Parental template DNA was digested using the *Dpn*I restriction endonuclease (NEB). The PCR products were transformed into DH5 $\alpha$  competent *E. coli* (as described in ref 17). Clones were selected for using 1% LB agar plates supplemented with kanamycin. Plasmids were isolated and purified using the GeneJET Plasmid Miniprep kit and protocol. The desired mutations were verified by Sanger sequencing (Garvan Institute).

**Protein Expression and Purification.** Test expressions as well as large scale expression and 3-step purification (Ni-NTA affinity, anion exchange, gel filtration) of P450 mutant enzymes were performed as has been previously reported with minor modifications to the buffer pH utilized (SI Table S2).<sup>17</sup>

**UV-vis Spectroscopy.** The UV-visible spectra of the P450 enzymes were obtained using a Jasco V-750 spectrophotometer as has been previously reported.<sup>17</sup> The spectra were measured until they reached a stable peak. CO difference spectra were obtained using a Jasco V-750 spectrophotometer at 30  $^{\circ}$ C as has been previously described.<sup>17</sup> The UV-vis spectra were measured between 390–600 nm every 5 min until the spectra was stable.<sup>10,17</sup> Substrate interaction with the P450s was measured using UV-visible absorbance spectroscopy carried out on a Jasco V-750 spectrophotometer according to

previously reported protocol.<sup>17</sup> Amplitudes of the different spectra were extrapolated ( $\Delta A = A_{\text{max}} - A_{\text{min}}$ ) and plotted against the substrate concentration (1). Curve fitting of the resulting data points was performed using GraphPad Prism 9 by applying a one-site binding model:  $Y = B_{\text{max}} * X / (K_d + X) + \text{NS} * X + \text{Background}$  ( $Y = \text{delta Abs}$ ,  $B_{\text{max}}$  = the maximum specific binding,  $K_d$  = equilibrium dissociation constant, NS = slope of nonspecific binding, Background = nonspecific binding with no added ligand).<sup>17</sup> Each enzyme–substrate pair was assayed in triplicate, and the average  $K_d$  determined.

**Enzyme Assays.** Peptide cyclization was performed using the P450 enzyme as well as redox partners as has been previously reported.<sup>17</sup> Peptides were synthesized using reported protocols.<sup>61,62</sup> The reaction was purified using a solid-phase extraction column (Bond elute, Agilent) as performed previously. Following concentration, the samples were analyzed on LCMS using an Agilent ZORBAX 300SB-C18 5  $\mu$ m column as previously reported.<sup>17</sup> Each reaction was carried out in triplicate, and data was analyzed using Excel and GraphPad Prism 9. Large-scale enzymatic cyclization was carried out to obtain sufficient material for analysis by NMR as has been previously reported.<sup>17</sup>

**HRMS Analysis.** High-resolution mass spectrometry measurements were performed on an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled to a Dionex UltiMate 3000 RSLCnano system equipped with a Dionex UltiMate 3000 RS autosampler, an Acclaim PepMap RSLC analytical column (75  $\mu$ m  $\times$  50 cm, nanoViper, C18, 2  $\mu$ m, 100  $\text{\AA}$ ; Thermo Scientific) and an Acclaim PepMap 100 trap column (100  $\mu$ m  $\times$  2 cm, nanoViper, C18, 5  $\mu$ m, 100  $\text{\AA}$ ; Thermo Scientific) as previously reported.<sup>17</sup>

**Crystallization.** To identify crystallization conditions, initial broad matrix screens were performed at the Monash Molecular Crystallization Platform (MMCP). Crystals of P450<sub>Blt</sub> were obtained at 10 mg/mL from both the PACT premier screen (drop G4:20% w/v PEG 3350, 0.2 M KSCN, 0.1 M BIS-TRIS prop 7.5 pH) and the Shotgun SG1 screen (drop B2:0.1 M MES, 6.5 pH, 12% w/v PEG 20K) via sitting drop at 20  $^{\circ}$ C. To enhance crystal quality, conditions were refined. Fine screen crystallization drops: 1.5  $\mu$ L of protein, 0.75  $\mu$ L of precipitant, equilibrated against 300  $\mu$ L precipitant. P450<sub>Blt</sub> yielded rod-shaped crystal aggregates. The rods, belonging to space group P3<sub>1</sub>, diffracted to 2.6  $\text{\AA}$ . These crystals had translational noncrystallographic symmetry present at a level complicating refinement. A single cubic crystal grew from the optimized condition 0.1 M Bis-Tris pH 7, 22% PEG3350, 0.05 KSCN which belonged to space group P 1 2<sub>1</sub> 1 and diffracted to 2.15  $\text{\AA}$ . Through optimization, we consistently produced cubic crystals in the space group P 1 2<sub>1</sub> 1 grown from the condition 0.1 M MES pH 6.5, 12% PEG400, and 14% PEG 20K. To obtain the substrate-bound structure, the crystal was soaked for 25 min in the reservoir solution supplemented with 30% sucrose and 10 mM MRYLH (acetic acid salt) before being mounted on cryo-loops and vitrified in liquid N<sub>2</sub> prior to X-ray data collection. All synchrotron diffraction data was collected at 100 K on the MX2<sup>63</sup> beamline at the Australian Synchrotron (Clayton, Victoria, Australia) equipped with an Eiger detector (Dectris). Data were processed and scaled using the routines XDS, Pointless, and Aimless from the CCP4 suite.<sup>64</sup> Data collection statistics are given in Table 1. The phases for structure determination were obtained by molecular replacement using PHASER from the PHENIX package and a poly alanine model of CYP107L2



(PDB: 5CWE)<sup>18</sup> as the search for 8U3N. Using the Phaser solution obtained, we conducted Cartesian-simulated annealing with PHENIX to avoid phase bias.<sup>65</sup> The final refined model emerged from iterations between manual real-space refinement in COOT<sup>66</sup> and automated refinement in PHENIX.<sup>65</sup> Initial stages of refinement primarily involved manual rebuilding, employing basic refinement options such as reciprocal and real space refinement and individual atom isotropic B factors with default NCS restraints. X-ray data for geometry weights and atomic displacement factors were automatically determined using the “optimize X-ray/stereochemistry weight” and “optimize X-ray/ADP weight” functions, respectively. Model validation was carried out using COOT<sup>66</sup> and MolProbity.<sup>20</sup> Statistics were generated in PHENIX through the “Generate Table 1 for journal” function. Chain A from the final model of 8U3N was used as a search model for 8U2M and 8UKZ during molecular replacement, followed by the same refinement strategy as for 8U3N. Structurally related P450s were identified using the DALI server,<sup>27</sup> using PDB entries deposited until November 2022.

**Molecular Dynamics Simulations.** The crystal structure of wild-type P450<sub>Blt</sub> in complex with its peptide substrate MRYLH (PDB 8U2M) was selected for molecular dynamics (MD) studies.

The protein preparation wizard module from Schrödinger<sup>67</sup> was used to refine the structure, encompassing the addition of missing hydrogens and side chains, removal of water molecules, assignment of bond orders, creation of zero-order bonds to metals, and generation of heteroatom states at a physiological pH of 7.4. The structure was further optimized using PROPKA at pH of 7.4 and minimized using OPLS4 force field, with the heavy atoms converged to RMSD of 0.3 Å.

The Desmond program from Schrödinger<sup>68</sup> was used to perform all MD simulations. The TIP5P solvent model was used to design the simulation system, with the system box shape set to Orthorhombic and a buffer distance of 15 Å. The system's total charge was neutralized using appropriate Na<sup>+</sup>/Cl<sup>−</sup>, and 0.15 M NaCl was added, with the OPLS4<sup>69</sup> force field incorporated to prepare the final system. Following an initial 1 ns minimization, 300 ns simulations were recorded by using the NPT ensemble for each system in triplicate, utilizing the Molecular Dynamics module. The time step was set to 2.0 fs, and temperature (300 K) and pressure (101.325 kPa) controls were established using the Nosé–Hoover Chain thermostat and Martyna–Tobias–Klein barostat, respectively. Short-range Coulombic interactions were calculated using the cutoff method. The generated trajectories were analyzed using the simulation event analysis and simulation interaction diagram modules.

**QM/MM Calculations.** Density functional theory calculations were carried out at the M06-2X/def2-TZVP// $\omega$ B97X-D/6-31+G(d)<sup>70–73</sup> level using Gaussian 16 Rev C.01.<sup>74</sup> Free energies are reported at standard states of 298.15 K and 1 mol/L. Conformational searching of each TS was conducted manually.

**NMR Analysis.** Nuclear magnetic resonance (NMR) spectra were collected using a 700 MHz Bruker Avance III HD instrument equipped with a TCI cryoprobe. Chemical shifts are given in parts per million (ppm) and referenced to the DMF-*d*<sub>7</sub> residual solvent peaks using  $\delta_{\text{H}}$  8.03 and  $\delta_{\text{C}}$  163.15. <sup>15</sup>N spectra are referenced to liquid ammonia  $\delta_{\text{N}}$  of 0 ppm using the Bruker program Xiref. Coupling constants, *J*, are reported in Hz.

**In Vivo Experiments.** To introduce point mutations into specific residues of BytO, mutagenesis primers were designed including complementary sequences for cloning into pSET152\_ermE\* via Gibson assembly, as well as mismatching nucleotides to amplify sequences with desired changes (SI Table S4). As a template, pSET152\_ermE\* bytAO was used.<sup>11</sup> Primer pairs A1 and A2 resulted in a 530 bp fragment and B1/B2 and B3 in a 791 bp fragment, which were subsequently assembled with a PCR-linearized vector. The following steps, including cloning, transformation, heterologous expression, and extraction were performed as described by Zdouc et al., 2021.<sup>15</sup>

Mass spectra were recorded on an Agilent 6530 Accurate Mass Q-TOF with ESI-source coupled with an Agilent 1100 HPLC using a Waters Atlantis T3, 5  $\mu$ m, 4.6 mm  $\times$  50 mm column. The column temperature was 25 °C. MS data were acquired over a range from 50 to 2500 *m/z* in positive mode. Auto MS/MS fragmentation was achieved with rising collision energy (35–50 keV over a gradient from 500 to 2000 *m/z*) with a frequency of 3 Hz for all ions over a threshold of 500. Elution started with a flow of 0.5 mL/min at 90% H<sub>2</sub>O containing 0.1% acetic acid (A) and 10% acetonitrile containing 0.1% acetic acid (B). After 2 min, a linear gradient to 100% B was run for 10 min, followed by an isocratic step with 100% B for 6 min. Five  $\mu$ L of a 2 mg/mL sample solution were injected per run.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscatal.3c05417>.

Primers for preparation of P450<sub>Blt</sub> mutants, buffers used for protein purification steps, complete NMR assignment for the product of cyclization of Nle-RYLH by P450<sub>Blt</sub>, mutagenesis primers used in this study for P450 BytO. Mismatching nucleotides introducing mutations are underlined, areas under curve for biaryllytite YYH heterologous expression in *S. coelicolor*, OMIT map and polder map, molecular dynamics simulations, molecular dynamics interaction diagram, extracted ion chromatograms, supplementary structures, supplementary references (PDF)

X-ray data (CIF)

X-ray data (CIF)

X-ray data (CIF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Mathias H. Hansen** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia; [orcid.org/0000-0003-4746-8519](https://orcid.org/0000-0003-4746-8519); Email: [mathias.hansen1@monash.edu](mailto:mathias.hansen1@monash.edu)

**Max J. Cryle** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia; [orcid.org/0000-0002-9739-6157](https://orcid.org/0000-0002-9739-6157); Email: [max.cryle@monash.edu](mailto:max.cryle@monash.edu)

## Authors

**Angus Keto** – School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia

**Maxine Treisman** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia

**Vishnu Mini Sasi** – ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia; Research School of Chemistry, The Australian National University, Acton, Australian Capital Territory 2601, Australia

**Laura Coe** – School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia; [orcid.org/0000-0002-3799-3865](https://orcid.org/0000-0002-3799-3865)

**Yongwei Zhao** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia

**Leo Padva** – Institute of Pharmaceutical Biology, University of Bonn, 53115 Bonn, Germany

**Caroline Hess** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; EMBL Australia, Monash University, Clayton, Victoria 3800, Australia; [orcid.org/0009-0001-7436-5067](https://orcid.org/0009-0001-7436-5067)

**Victor Leichthammer** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; EMBL Australia, Monash University, Clayton, Victoria 3800, Australia

**Daniel L. Machell** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia

**Ralf B. Schittenhelm** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; Monash Proteomics and Metabolomics Platform, Monash University, Clayton, Victoria 3800, Australia; [orcid.org/0000-0001-8738-1878](https://orcid.org/0000-0001-8738-1878)

**Colin J. Jackson** – ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia; Research School of Chemistry, The Australian National University, Acton, Australian Capital Territory 2601, Australia; ARC Centre of Excellence in Synthetic Biology, Australian National University, Canberra 2601, Australia; Research School of Biology, Australian National University, Acton, Australian Capital Territory 2601, Australia; [orcid.org/0000-0001-6150-3822](https://orcid.org/0000-0001-6150-3822)

**Julien Tailhades** – Department of Biochemistry and Molecular Biology, The Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia; ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia; [orcid.org/0000-0003-0391-2526](https://orcid.org/0000-0003-0391-2526)

**Max Crüsemann** – Institute of Pharmaceutical Biology, University of Bonn, 53115 Bonn, Germany; [orcid.org/0000-0001-6660-2715](https://orcid.org/0000-0001-6660-2715)

**James J. De Voss** – School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia; [orcid.org/0000-0002-2659-5140](https://orcid.org/0000-0002-2659-5140)

**Elizabeth H. Krenske** – School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia; [orcid.org/0000-0003-1911-0501](https://orcid.org/0000-0003-1911-0501)

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acscatal.3c05417>

## Author Contributions

<sup>†</sup>M.H.H., A.K., and M.T. contributed equally.

## Funding

This work was supported by Monash University and EMBL Australia. L.P. was funded through a Ph.D. scholarship from the German Academic Scholarship Foundation. A.K. acknowledges the support of an Australian Government Research Training Program Scholarship and an American Australian Association Graduate Education Fund Scholarship. This research was conducted by the Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012) and funded by the Australian Government.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This study used BPA-enabled (Bioplatforms Australia)/NCRIS-enabled (National Collaborative Research Infrastructure Strategy) infrastructure located at the Monash Proteomics and Metabolomics Platform. Computational resources were provided by the National Computational Infrastructure through the National Computational Merit Allocation Scheme (supported by NCRIS) and by the University of Queensland Research Computing Centre.

## REFERENCES

- (1) Ariyasu, S.; Stanfield, J. K.; Aiba, Y.; Shoji, O. Expanding the applicability of cytochrome P450s and other haemoproteins. *Curr. Opin. Chem. Biol.* **2020**, *59*, 155–163.
- (2) Yosca, T. H.; et al. Iron(IV)hydroxide pK(a) and the role of thiolate ligation in C-H bond activation by cytochrome P450. *Science* **2013**, *342*, 825–829.
- (3) Greule, A.; Stok, J. E.; De Voss, J. J.; Cryle, M. J. Unrivalled diversity: the many roles and reactions of bacterial cytochromes P450 in secondary metabolism. *Nat. Prod. Rep.* **2018**, *35*, 757–791.
- (4) Hansen, M. H.; Stegmann, E.; Cryle, M. J. Beyond vancomycin: recent advances in the modification, reengineering, production and discovery of improved glycopeptide antibiotics to tackle multidrug-resistant bacteria. *Curr. Opin. Biotech.* **2022**, *77*, No. 102767.
- (5) Aldemir, H.; Shu, S.; Schaefer, F.; Hong, H.; Richarz, R.; Harteis, S.; Einsiedler, M.; Milzarek, T. M.; Schneider, S.; Gulder, T. A. M. Carrier Protein-Free Enzymatic Biaryl Coupling in Arylomycin A2 Assembly and Structure of the Cytochrome P450 AryC. *Chem.—Eur. J.* **2022**, *28*, No. e202103389.
- (6) Haslinger, K.; Cryle, M. Structure of OxyAtei: completing our picture of the glycopeptide antibiotic producing Cytochrome P450 cascade. *FEBS Lett.* **2016**, *590*, 571–581.



- (7) Haslinger, K.; Maximowitsch, E.; Brieke, C.; Koch, A.; Cryle, M. J. Cytochrome P450 OxyBte catalyzes the first phenolic coupling step in teicoplanin biosynthesis. *ChemBioChem* **2014**, *15*, 2719–2728.
- (8) Cryle, M. J.; Staaden, J.; Schlichting, I. Structural characterization of CYP16SD3, a cytochrome P450 involved in phenolic coupling in teicoplanin biosynthesis. *Arch. Biochem. Biophys.* **2011**, *507*, 163–173.
- (9) Pylypenko, O.; Vitali, F.; Zerbe, K.; Robinson, J. A.; Schlichting, I. Crystal Structure of OxyC, a Cytochrome P450 Implicated in an Oxidative C-C Coupling Reaction during Vancomycin Biosynthesis. *J. Biol. Chem.* **2003**, *278*, 46727–46733.
- (10) Greule, A.; Izore, T.; Iftime, D.; Tailhades, J.; Schoppet, M.; Zhao, Y.; Peschke, M.; Ahmed, L.; Kulik, A.; Adamek, M.; Goode, R. J. A.; Schittenhelm, R. B.; Kaczmarek, J. A.; Jackson, C. J.; Ziemert, N.; Krenske, E. H.; De Voss, J. J.; Stegmann, E.; Cryle, M. J. Kistamicin biosynthesis reveals the biosynthetic requirements for production of highly crosslinked glycopeptide antibiotics. *Nat. Commun.* **2019**, *10*, 2613.
- (11) Haslinger, K.; Peschke, M.; Brieke, C.; Maximowitsch, E.; Cryle, M. J. X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* **2015**, *521*, 105–109.
- (12) Ali, H. S.; Henchman, R. H.; de Visser, S. P. Cross-linking of aromatic phenolate groups by cytochrome P450 enzymes: a model for the biosynthesis of vancomycin by OxyB. *Organic & Biomolecular Chemistry* **2020**, *18*, 4610–4618.
- (13) Forneris, C. C.; Nguy, A. K. L.; Seyedsayamdost, M. R. Mapping and Exploiting the Promiscuity of OxyB toward the Biocatalytic Production of Vancomycin Aglycone Variants. *ACS Catal.* **2020**, *10*, 9287–9298.
- (14) Kunakom, S.; Otani, H.; Udway, D. W.; Doering, D. T.; Mouncey, N. J. Cytochromes P450 involved in bacterial RiPP biosyntheses. *Journal of Industrial Microbiology and Biotechnology* **2023**, *50*, kuad005.
- (15) Zdouc, M. M.; Alanjary, M. M.; Zarazua, G. S.; Maffioli, S. I.; Crusemann, M.; Medema, M. H.; Donadio, S.; Sosio, M. A biaryl-linked tripeptide from *Planomonospora* reveals a widespread class of minimal RiPP gene clusters. *Cell Chem. Biol.* **2021**, *28*, 733–739.e734.
- (16) Hug, J. J.; et al. Genome-Guided Discovery of the First Myxobacterial Biaryllylids Reveals Distinct C-N Biaryl Crosslinking in RiPP Biosynthesis. *Molecules* **2021**, *26*, 7483.
- (17) Zhao, Y.; Marschall, E.; Treisman, M.; McKay, A.; Padva, L.; Crusemann, M.; Nelson, D. R.; Steer, D. L.; Schittenhelm, R. B.; Tailhades, J.; Cryle, M. J. Cytochrome P450Blt Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks. *Angew. Chem.* **2022**, *61*, No. e202204957.
- (18) Han, S.; et al. Structural insights into the binding of lauric acid to CYP107L2 from *Streptomyces avermitilis*. *Biochem. Biophys. Res. Commun.* **2017**, *482*, 902–908.
- (19) Baek, M.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876.
- (20) Chen, V. B.; et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D* **2010**, *66*, 12–21.
- (21) Belin, P.; et al. Identification and structural basis of the reaction catalyzed by CYP121, an essential cytochrome P450 in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 7426–7431.
- (22) Kemp, M. T.; Lewandowski, E. M.; Chen, Y. Low barrier hydrogen bonds in protein structure and function. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2021**, *1869*, No. 140557.
- (23) Zhao, Y.; Ho, Y. T. C.; Tailhades, J.; Cryle, M. Understanding the Glycopeptide Antibiotic Crosslinking Cascade: In Vitro Approaches Reveal the Details of a Complex Biosynthesis Pathway. *ChemBioChem* **2021**, *22*, 43–51.
- (24) Iyer, A. H.; Krishna Deepak, R. N. V.; Sankaramakrishnan, R. Imidazole Nitrogens of Two Histidine Residues Participating in N-H...N Hydrogen Bonds in Protein Structures: Structural Bioinformatics Approach Combined with Quantum Chemical Calculations. *The. J. Phys. Chem. B* **2018**, *122*, 1205–1212.
- (25) Mohamed, H.; Ghith, A.; Bell, S. G. The binding of nitrogen-donor ligands to the ferric and ferrous forms of cytochrome P450 enzymes. *J. Inorg. Biochem.* **2023**, *242*, No. 112168.
- (26) Bewley, C. A.; He, H. Y.; Williams, D. H.; Faulkner, D. J. Aciculitins A-C: Cytotoxic and antifungal cyclic peptides from the lithistid sponge *Aciculites orientalis*. *J. Am. Chem. Soc.* **1996**, *118*, 4314–4321.
- (27) Holm, L.; Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **2010**, *38*, W545–W549.
- (28) Nagano, S.; Cupp-Vickery, J. R.; Poulos, T. L. Crystal Structures of the Ferrous Dioxygen Complex of Wild-type Cytochrome P450EryF and Its Mutants, A245S and A245T: investigation of the proton transfer system in P450EryF. *J. Biol. Chem.* **2005**, *280*, 22102–22107.
- (29) Du, L.; Dong, S.; Zhang, X.; Jiang, C.; Chen, J.; Yao, L.; Wang, X.; Wan, X.; Liu, X.; Wang, X.; Huang, S.; Cui, Q.; Feng, Y.; Liu, S.-J.; Li, S. Selective oxidation of aliphatic C–H bonds in alkylphenols by a chemomimetic biocatalytic system. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E5129–E5137.
- (30) Sabbadin, F.; et al. The 1.5 Å Structure of XplA-heme, an Unusual Cytochrome P450 Heme Domain That Catalyzes Reductive Biotransformation of Royal Demolition Explosive. *J. Biol. Chem.* **2009**, *284*, 28467–28475.
- (31) Cupp-Vickery, J. R.; Han, O.; Hutchinson, C. R.; Poulos, T. L. Substrate-assisted catalysis in cytochrome P450eryF. *Nat. Struct. Biol.* **1996**, *3*, 632–637.
- (32) Cupp-Vickery, J. R.; Poulos, T. L. Structure of cytochrome P450eryF involved in erythromycin biosynthesis. *Nat. Struct. Biol.* **1995**, *2*, 144–153.
- (33) Zocher, G.; Richter, M. E. A.; Mueller, U.; Hertweck, C. Structural Fine-Tuning of a Multifunctional Cytochrome P450 Monooxygenase. *J. Am. Chem. Soc.* **2011**, *133*, 2292–2302.
- (34) Li, Z.; Rupasinghe, S. G.; Schuler, M. A.; Nair, S. K. Crystal structure of a phenol-coupling P450 monooxygenase involved in teicoplanin biosynthesis. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 1728–1738.
- (35) Peschke, M.; Brieke, C.; Cryle, M. J. F-O-G Ring Formation in Glycopeptide Antibiotic Biosynthesis is Catalysed by OxyE. *Sci. Rep* **2016**, *6*, 35584.
- (36) Shoji, O.; et al. Understanding substrate misrecognition of hydrogen peroxide dependent cytochrome P450 from *Bacillus subtilis*. *JBIC Journal of Biological Inorganic Chemistry* **2010**, *15*, 1331–1339.
- (37) Fujishiro, T.; et al. Crystal Structure of H<sub>2</sub>O<sub>2</sub>-dependent Cytochrome P450SPa with Its Bound Fatty Acid Substrate. *J. Biol. Chem.* **2011**, *286*, 29941–29950.
- (38) Bell, S. G.; et al. The crystal structures of 4-methoxybenzoate bound CYP199A2 and CYP199A4: structural changes on substrate binding and the identification of an anion binding site. *Dalton Trans* **2012**, *41*, 8703–8714.
- (39) Savino, C.; et al. Investigating the Structural Plasticity of a Cytochrome P450: three dimensional structures of P450 EryK and binding to its physiological substrate. *J. Biol. Chem.* **2009**, *284*, 29170–29179.
- (40) Sherman, D. H.; et al. The Structural Basis for Substrate Anchoring, Active Site Selectivity, and Product Formation by P450 PikC from *Streptomyces venezuelae*. *J. Biol. Chem.* **2006**, *281*, 26289–26297.
- (41) Li, S.; et al. Substrate Recognition by the Multifunctional Cytochrome P450 MycG in Mycinamicin Hydroxylation and Epoxidation Reactions. *J. Biol. Chem.* **2012**, *287*, 37880–37890.
- (42) Newmister, S. A.; et al. Molecular Basis of Iterative C–H Oxidation by TamI, a Multifunctional P450 Monooxygenase from the Tirandamycin Biosynthetic Pathway. *ACS Catal.* **2020**, *10*, 13445–13454.
- (43) Cryle, M. J.; Schlichting, I. Structural insights from a P450 Carrier Protein complex reveal how specificity is achieved in the

- P450<sub>Biol</sub> ACP complex. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 15696–15701.
- (44) McCoy, J. G.; et al. Structural characterization of CalO2: A putative orsellinic acid P450 oxidase in the calicheamicin biosynthetic pathway. *Proteins: Struct., Funct., Bioinf.* **2009**, *74*, 50–60.
- (45) Zhao, B.; et al. Binding of Two Flavinol Substrate Molecules, Oxidative Coupling, and Crystal Structure of *Streptomyces coelicolor* A3(2) Cytochrome P450 158A2. *J. Biol. Chem.* **2005**, *280*, 11599–11607.
- (46) Zhao, B.; et al. Different Binding Modes of Two Flavinol Substrate Molecules in Cytochrome P450 158A1 (CYP158A1) Compared to CYP158A2. *Biochemistry* **2007**, *46*, 8725–8733.
- (47) Nguyen, R. C.; Yang, Y.; Wang, Y.; Davis, I.; Liu, A. Substrate-Assisted Hydroxylation and O-Demethylation in the Peroxidase-like Cytochrome P450 Enzyme CYP121. *ACS Catal.* **2020**, *10*, 1628–1639.
- (48) Alkhalaf, L. M.; et al. Binding of Distinct Substrate Conformations Enables Hydroxylation of Remote Sites in Thaxtomin D by Cytochrome P450 TxtC. *J. Am. Chem. Soc.* **2019**, *141*, 216–222.
- (49) Cryle, M. J.; Bell, S. G.; Schlichting, I. Structural and Biochemical Characterization of the Cytochrome P450 CypX (CYP134A1) from *Bacillus subtilis*: A Cyclo-l-leucyl-l-leucyl Dipeptide Oxidase. *Biochemistry* **2010**, *49*, 7282–7296.
- (50) He, F.; et al. Molecular basis for the P450-catalyzed C–N bond formation in indolactam biosynthesis. *Nat. Chem. Biol.* **2019**, *15*, 1206–1213.
- (51) Shende, V. V.; Harris, N. R.; Sanders, J. N.; Newmister, S. A.; Khatri, Y.; Movassaghi, M.; Houk, K. N.; Sherman, D. H. Molecular Dynamics Simulations Guide Chimeraogenesis and Engineered Control of Chemoselectivity in Diketopiperazine Dimerases. *Angew. Chem.* **2023**, *62*, No. e202210254.
- (52) Makino, M.; et al. Crystal structures and catalytic mechanism of cytochrome P450 StaP that produces the indolocarbazole skeleton. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11591–11596.
- (53) Gering, H. E.; et al. A Ferric-Superoxide Intermediate Initiates P450-Catalyzed Cyclic Dipeptide Dimerization. *J. Am. Chem. Soc.* **2023**, *145*, 19256–19264.
- (54) Zerbe, K.; et al. Crystal Structure of OxyB, a Cytochrome P450 Implicated in an Oxidative Phenol Coupling Reaction during Vancomycin Biosynthesis. *J. Biol. Chem.* **2002**, *277*, 47476–47485.
- (55) Cryle, M. J.; Meinhart, A.; Schlichting, I. Structural Characterization of OxyD, a Cytochrome P450 Involved in b-Hydroxytyrosine Formation in Vancomycin Biosynthesis. *J. Biol. Chem.* **2010**, *285*, 24562–24574.
- (56) Haslinger, K.; Briek, C.; Uhlmann, S.; Sieverling, L.; Sussmuth, R. D.; Cryle, M. J. The Structure of a Transient Complex of a Nonribosomal Peptide Synthetase and a Cytochrome P450 Monooxygenase. *Angew. Chem. Int. Ed.* **2014**, *53*, 8518.
- (57) Nam, H.; et al. Exploring the Diverse Landscape of Biaryl-Containing Peptides Generated by Cytochrome P450 Macrocyclases. *J. Am. Chem. Soc.* **2023**, *145*, 22047–22057.
- (58) Zhang, S.; et al. P450-mediated dehydrotyrosine formation during WS9326 biosynthesis proceeds via dehydrogenation of a specific acylated dipeptide substrate. *Acta Pharmaceutica Sinica B* **2023**, *13*, 3561–3574.
- (59) He, B.-B.; Liu, J.; Cheng, Z.; Liu, R.; Zhong, Z.; Gao, Y.; Liu, H.; Song, Z.-M.; Tian, Y.; Li, Y.-X. Bacterial Cytochrome P450 Catalyzed Post-translational Macrocyclization of Ribosomal Peptides. *Angew. Chem.* **2023**, *62*, No. e202311533.
- (60) Molinaro, C.; et al. Engineered Cytochrome P450-Catalyzed Oxidative Biaryl Coupling Reaction Provides a Scalable Entry into Arylomycin Antibiotics. *J. Am. Chem. Soc.* **2022**, *144*, 14838–14845.
- (61) Briek, C.; Cryle, M. J. A Facile Fmoc Solid Phase Synthesis Strategy To Access Epimerization-Prone Biosynthetic Intermediates of Glycopeptide Antibiotics. *Org. Lett.* **2014**, *16*, 2454–2457.
- (62) Tailhades, J.; et al. A route to diastereomerically pure phenylglycine thioester peptides: crucial intermediates for investigating glycopeptide antibiotic biosynthesis. *Chem. Commun.* **2018**, *54*, 2146–2149.
- (63) McPhillips, T. M.; et al. Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *Journal of Synchrotron Radiation* **2002**, *9*, 401–406.
- (64) Winn, M. D.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D* **2011**, *67*, 235–242.
- (65) Adams, P. D.; et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **2010**, *66*, 213–221.
- (66) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **2010**, *66*, 486–501.
- (67) Schrödinger Release 2021-3: Protein Prep Wizard; Schrödinger: New York, NY, 2021).
- (68) Schrödinger Release 2021-3: Desmond; Schrödinger: New York, NY, 2021).
- (69) Lu, C.; et al. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *J. Chem. Theory Comput.* **2021**, *17*, 4291–4300.
- (70) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (71) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (72) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- (73) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (74) Gaussian 16 Revision. C.01; Gaussian, Inc.: Wallingford, CT, 2016.

## 7.4 Appendix D

The following pages feature the article “Ribosomal pentapeptide nitration for non-ribosomal peptide antibiotic precursor biosynthesis” as well as its supplemental information as it was published in Chem by Elsevier Inc. This publication is licensed under the Creative Commons CC-BY-NC license and permits non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited. A written consent from Elsevier Inc. is not necessary for this purpose.

The article is reprinted with permission from:

Leo Padva, Lukas Zimmer, Jemma Gullick, Yongwei Zhao, Vishnu Mini Sasi, Ralf B. Schittenhelm, Colin J. Jackson, Max Cryle, Max Crüsemann. *Chem* 2025, 11, 102438.

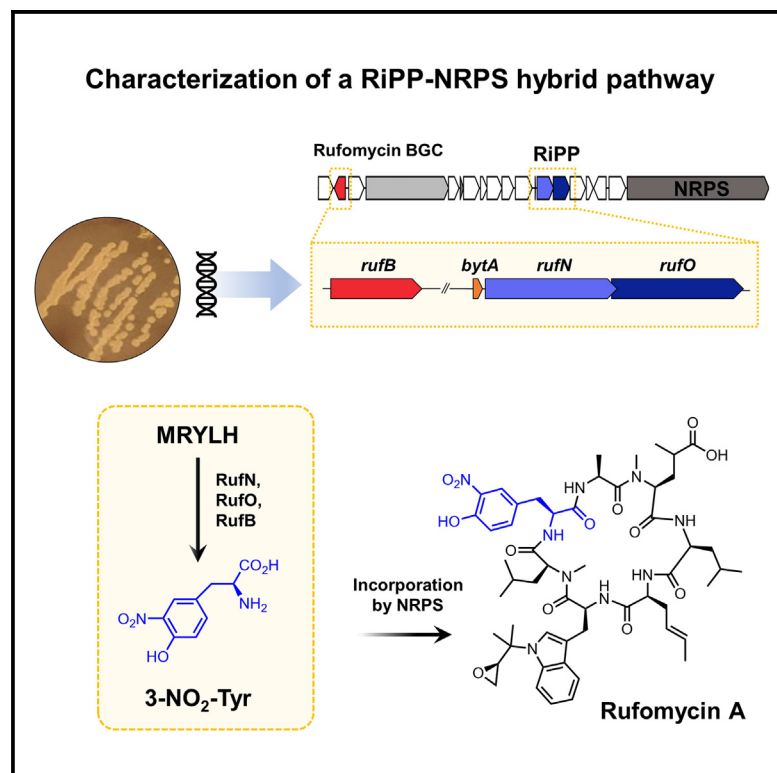
Copyright 2025 The Authors. Published by Elsevier Inc.

To view the publication and supplementary information online, please follow the link below.

<https://doi.org/10.1016/j.chempr.2025.102438>

# Ribosomal pentapeptide nitration for non-ribosomal peptide antibiotic precursor biosynthesis

## Graphical abstract



## Authors

Leo Padva, Lukas Zimmer, Jemma Gullick, ..., Colin J. Jackson, Max J. Cryle, Max Crüsemann

## Correspondence

max.cryle@monash.edu (M.J.C.),  
cruesemann@em.uni-frankfurt.de (M.C.)

## In brief

We discovered and characterized the biosynthesis of the noncanonical 3-nitrotyrosine building block found in the potent non-ribosomal antituberculosis peptide antibiotic rufomycin. In this pathway, a ribosomally synthesized pentapeptide is nitrated by a dedicated P450 monooxygenase that evolved from a biaryl-crosslinking enzyme. 3-nitrotyrosine is then released by a cluster-encoded protease before its incorporation by the non-ribosomal rufomycin megasynthetase. This example represents an unprecedented interplay of ribosomal and non-ribosomal peptide biosynthesis pathways.

## Highlights

- Biosynthesis and incorporation of NO<sub>2</sub>-Tyr residue in rufomycin was characterized
- A ribosomal pentapeptide is nitrated and digested to release the 3-NO<sub>2</sub>-Tyr precursor
- A crosslinking P450 monooxygenase evolved to perform direct pentapeptide nitration
- First characterized NRPS-RiPP hybrid biosynthesis pathway



Padva et al., 2025, Chem 11, 102438  
June 12, 2025 © 2025 The Author(s). Published by Elsevier Inc.  
<https://doi.org/10.1016/j.chempr.2025.102438>

Article

# Ribosomal pentapeptide nitration for non-ribosomal peptide antibiotic precursor biosynthesis

Leo Padva,<sup>1</sup> Lukas Zimmer,<sup>1,10</sup> Jemma Gullick,<sup>2,3,4,10</sup> Yongwei Zhao,<sup>2,3,4</sup> Vishnu Mini Sasi,<sup>4,5</sup> Ralf B. Schittenhelm,<sup>2,6</sup> Colin J. Jackson,<sup>4,5,7,8</sup> Max J. Cryle,<sup>2,3,4,\*</sup> and Max Crüsemann<sup>1,4,9,11,\*</sup>

<sup>1</sup>Institute of Pharmaceutical Biology, University of Bonn, 53115 Bonn, Germany

<sup>2</sup>The Monash Biomedicine Discovery Institute, Monash University, Clayton, VIC 3800, Australia

<sup>3</sup>EMBL Australia, Monash University, Clayton, VIC 3800, Australia

<sup>4</sup>ARC Centre of Excellence for Innovations in Peptide and Protein Science, Acton, ACT 2061, Australia

<sup>5</sup>Research School of Chemistry, The Australian National University, Acton, ACT 2601, Australia

<sup>6</sup>The Monash Proteomics and Metabolomics Platform, Monash University, Clayton, VIC 3800, Australia

<sup>7</sup>ARC Centre of Excellence in Synthetic Biology, Australian National University, Canberra, ACT, Australia

<sup>8</sup>Research School of Biology, Australian National University, Acton, ACT 2601, Australia

<sup>9</sup>Institute of Pharmaceutical Biology, Goethe University of Frankfurt, 60438 Frankfurt, Germany

<sup>10</sup>These authors contributed equally

<sup>11</sup>Lead contact

\*Correspondence: [max.cryle@monash.edu](mailto:max.cryle@monash.edu) (M.J.C.), [cruesemann@em.uni-frankfurt.de](mailto:cruesemann@em.uni-frankfurt.de) (M.C.)

<https://doi.org/10.1016/j.chempr.2025.102438>

**THE BIGGER PICTURE** Peptides are one of the major natural product classes harboring enormous structural diversity and complexity. They are biosynthesized by two major pathway types, either assembled by non-ribosomal peptide synthetase megaenzymes, or as ribosomally produced and post-translationally modified peptides. In this study, we elucidated the biosynthesis of the noncanonical 3-nitrotyrosine building block in the rufomycin class of cyclic peptide antibiotics. This residue has been shown to be crucial for the potent antibiotic activity of the rufomycins against the causative agent of tuberculosis. We found that the source of this building block is a previously overlooked ribosomal pentapeptide encoded in the rufomycin gene cluster. The nitrating monooxygenase acts on this ribosomal peptide and is closely related to crosslinking enzymes, with its activity able to be switched to catalyze biaryl bond formation by just two active-site mutations. This study shows an unprecedented interplay of non-ribosomal and ribosomal peptide biosynthesis to construct a potent antibiotic natural product and reveals novel opportunities for pathway and enzyme engineering for the generation of improved bioactive metabolites.

## SUMMARY

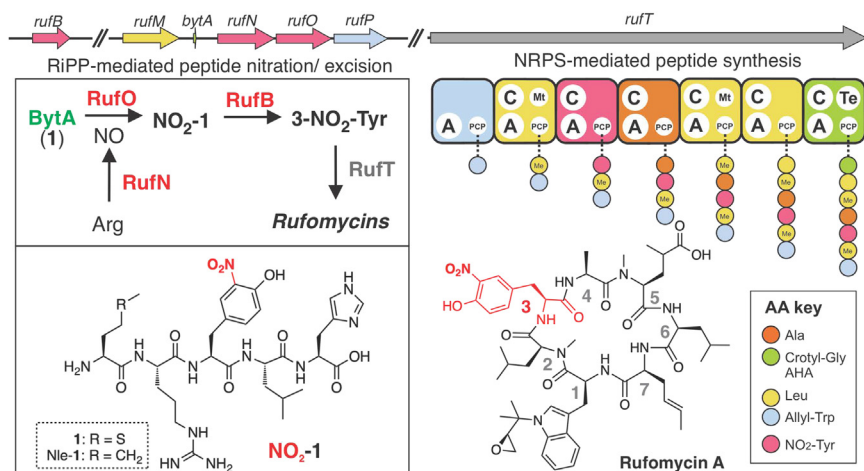
Peptide natural products possess a fascinating array of complex structures and diverse biological activities. Central to this is a repertoire of structurally modified amino acid building blocks, which stem from fundamentally different biosynthetic pathways for peptides of non-ribosomal and ribosomal origins. Given these origins, the integration of non-ribosomal and ribosomal peptide biosynthesis has previously been thought implausible. Now, we report how nature has synergized ribosomal and non-ribosomal peptide pathways in the biosynthesis of the rufomycins, exceptionally potent antitubercular antibiotics. In this pathway, a biaryl-type ribosomal pentapeptide precursor is nitrated by a modified cytochrome P450 biaryl-crosslinking enzyme. The nitrated residue, key for antibiotic activity, is liberated by a dedicated protease before activation and peptide incorporation by the non-ribosomal rufomycin synthetase assembly line. This resolves the enigmatic origins of 3-nitrotyrosine within rufomycin biosynthesis and unveils a novel function for ribosomally synthesized peptides as templates for biosynthesis of modified non-ribosomal peptide building blocks.

## INTRODUCTION

Specialized metabolism generates a wealth of structurally complex molecules with diverse biological functions. Peptides are

one of the major natural product classes, and their highly variable structures and potent biological activities make peptides important molecules for application in human medicine, for example, as antibiotics.<sup>1</sup> Natural peptides are predominantly produced





**Figure 1. Rufomycin BGC harbors a RiPP pathway to afford 3-nitrotyrosine**

The rufomycin biosynthetic gene cluster (BGC), an NRPS-polyketide synthase (PKS) hybrid, features a RiPP biosynthetic pathway, with *bytA* encoding the pentapeptide MRYLH (1) located upstream of the *rufNO* operon. RufN generates nitric oxide (NO) from Arg, with RufO using NO to nitrate Tyr in 1. RufB cleaves NO<sub>2</sub>-1 to release 3-NO<sub>2</sub>-Tyr, which is incorporated into the non-ribosomal peptide chain by the RufT A<sub>3</sub> domain (MLP RufH not shown). Color code links genes to residues in rufomycin with genes shown to scale, except for *rufT*. For complete cluster details, see Figure S1. Peptide structures are shown in Figure S9.

by two major pathways categorized by either the use or exclusion of the ribosome. Ribosomally independent peptide biosynthesis is diverse and includes the common and well-investigated modular thiotemplated non-ribosomal peptide synthetase (NRPS) megaenzymes that utilize condensation (C) domains for peptide bond formation between carrier protein-bound amino acids.<sup>2</sup> In addition, non-ribosomal peptide biosynthesis may feature other ways of amide bond formation, catalyzed e.g., by transglutaminases, ligases, or tRNA-dependent peptide synthetases. These may work independently of carrier proteins and are not typically organized into multidomain modules.<sup>3</sup> While ribosome-independent biosynthetic machineries are not limited to the selection of proteinogenic amino acids, ribosomally synthesized and post-translationally modified peptides (RiPPs)<sup>4</sup> require the incorporation of proteinogenic residues that could be seen as a limiting factor for their broader structural diversity. However, recent studies reporting highly modified peptide natural products of ribosomal origin have demonstrated that there is unprecedented capacity for structural diversity within such RiPP pathways,<sup>5–11</sup> thus revealing that the biosynthetic scope of these different pathway types is closer than was previously envisioned.<sup>12</sup>

Within peptide natural product biosynthesis, one important modification that has been thought to be restricted to non-ribosomal pathways is the nitration of amino acids.<sup>13</sup> Nitration is a common modification in higher eukaryotes, where it is often a marker of oxidative stress caused by the diffusion of nitric oxide (NO) and reactive oxygen species.<sup>14</sup> The role of nitration in plants is also an area of active investigation, especially given the importance of nitrogen metabolism in these systems and the interplay of soil bacteria and plants in the soil microenvironment. Within peptide biosynthesis pathways, amino acid nitration has been reported for Trp- and Tyr-containing peptides, although the direct nitration of amino acids has been restricted to the biosynthesis pathways of the phytotoxin thaxtomin, assembled by a tRNA-dependent cyclodipeptide synthase<sup>15,16</sup> and the NRPS-derived rufomycin/ilamycin antibiotics.<sup>17,18</sup> These pathways both utilize a NO synthase (TxtD/RufN) that generates NO from Arg paired with a cytochrome P450 enzyme (TxtE/RufO) that cat-

alyzes nitration.<sup>13,16</sup> Cytochrome P450s are a superfamily of diverse oxidative enzymes involved in a range of complex chemical transformations.<sup>19</sup> Direct nitration however remains an unusual transformation that lies outside of the classic oxidative chemistry catalyzed by the highly reactive P450 intermediate compound I.<sup>20</sup> Such alternate reactivity makes the study of these nitrating P450s of great interest, and while TxtE, the enzyme from thaxtomin biosynthesis, has been characterized,<sup>15,16</sup> the function of RufO from rufomycin biosynthesis had remained enigmatic.

The rufomycins, also known as ilamycins, are a family of more than 50 characterized congeners of cyclic heptapeptides known for their potent bioactivity against *Mycobacteria*, the causative agent of tuberculosis.<sup>21</sup> Rufomycins exhibit improved bioactivity compared with existing anti-tuberculosis agents, positioning them as compelling candidates for drug development efforts by targeting the protease ClpC1 in *Mycobacterium tuberculosis*.<sup>22,23</sup> Structure-activity relationships of the rufomycins have revealed moieties central to their antibacterial activity, the most prominent of which is the NO<sub>2</sub> group of the unique 3-NO<sub>2</sub>-Tyr residue.<sup>24</sup> The rufomycins are synthesized by a heptamodular NRPS RufT, with the non-proteinogenic L-2-amino-4-hexenoic acid (AHA) moiety synthesized by the polyketide synthase RufE and incorporated by the last module of the rufomycin NRPS (Figure 1).<sup>17,18</sup> Due to the high homology of the genes *rufNO* to the *txtDE* system, which was shown to perform direct nitration of tryptophan in thaxtomin biosynthesis,<sup>16</sup> the 3-NO<sub>2</sub>-Tyr residue in rufomycin has been postulated to be synthesized by an analogous mechanism involving direct nitration of free Tyr.<sup>18</sup> However, although deletion experiments showed that both genes are essential for rufomycin (ilamycin) biosynthesis,<sup>17</sup> this hypothesis has not been supported by reported data, and recent structural studies have suggested that RufO may instead act on Tyr bound to a carrier protein domain following selection and activation by the NRPS machinery.<sup>25–27</sup>

Our groups have been investigating the biosynthesis of the biarylites, a family of ribosomally synthesized, biaryl-linked cyclic tripeptides recently discovered from the actinomycete genus *Planomonospora*.<sup>28</sup> Bioinformatic mining studies by

both ourselves and others have revealed that additional biaryl-  
lityde-type pathways are present in a number of actinobacterial  
genomes. Many of these clusters encode additional biosynthetic  
enzymes, suggesting a wide distribution and diverse chemistry  
encoded in this RiPP family.<sup>28–30</sup> Among the reported expansion  
of P450s involved in the biosynthesis of crosslinked RiPPs,<sup>29–31</sup>  
the biarylitydes remain unique due to their biosynthesis being  
centered on translation of the precursor pentapeptide (typically  
MxYxH) encoded by *bytA*, the smallest known functional gene  
in the tree of life.<sup>28</sup> This linear peptide undergoes intramolecular  
aromatic crosslinking between Tyr-3 and His-5, a transformation  
catalyzed by the P450 BytO, consistent with other P450 en-  
zymes catalyzing the formation of a range of crosslinked tripep-  
tides.<sup>32–34</sup> Here, we reveal and characterize the unprecedented  
involvement of a biarylityde RiPP pathway in generating the  
non-proteinogenic 3-NO<sub>2</sub>-Tyr residue during biosynthesis of  
the NRPS-derived rufomycin antibiotics.

## RESULTS

During our search for novel, noncanonical biarylityde pathways,  
we were intrigued to identify a P450 with 97.97% sequence iden-  
tity to RufO encoded in one of our in-house strains, *Strepto-  
myces atratus* S3\_m208\_1 (isolated from a Belgian soil sam-  
ple<sup>35</sup>), located next to a *bytA* homolog encoding a MRYLH  
pentapeptide upstream of the *rufNO* operon (Figure 1). Remark-  
ably, this biarylityde biosynthetic gene cluster (BGC) is located in  
the center of the rufomycin BGC, with homologous enzymes also  
identified in all other rufomycin BGCs found in the databases  
(Table S1; Figures S1 and S2). We thus started to explore the  
role of this RiPP pathway in rufomycin biosynthesis.

### Cytochrome P450 RufO nitrates a RiPP precursor peptide

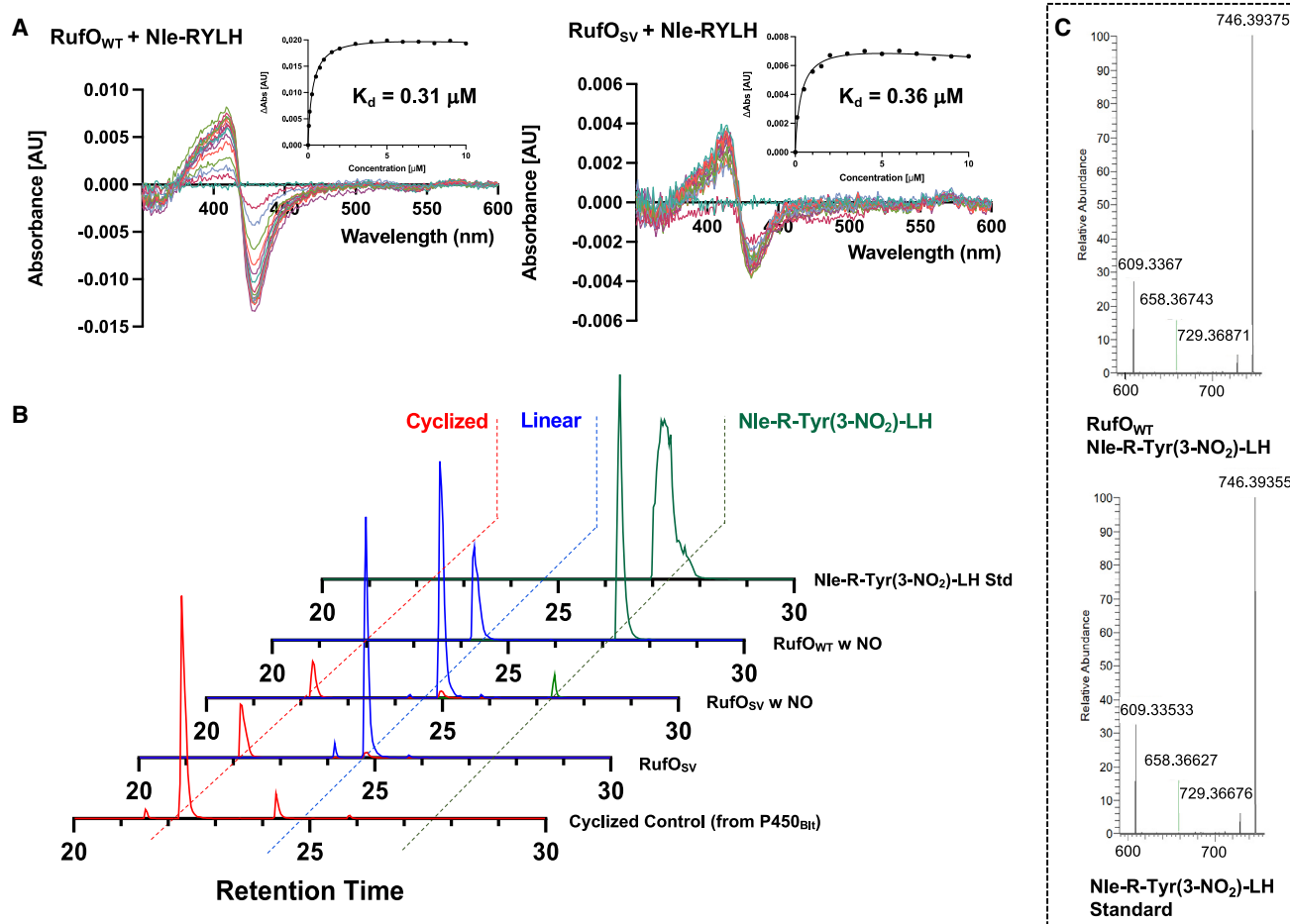
Our initial investigations into the function of RufO stemmed from  
comparisons between its structure<sup>26,27</sup> and that of the biarylityde  
crosslinking P450<sub>Bit</sub> that we had recently characterized.<sup>32</sup> Our  
analysis of these structures revealed that they were highly similar  
(41.6% amino acid ID, C- $\alpha$  RMSD 1.8 Å to 8SPC; Figure S3),  
apart from differences at a key section of the central I-helix of  
RufO (Figure S4). The I-helix in P450 enzymes runs across the  
prosthetic heme group within the enzyme active site (Figure S5,  
cyan), with key residues—typically a conserved acid/alcohol  
pair—involved in oxygen activation during the P450 active cy-  
cle.<sup>33</sup> The identities of the residues found at the typical acid/  
alcohol positions in RufO (P450<sub>Bit</sub>: Ser239, Val240 vs. RufO:  
Val240, Pro241; Figures S4 and S5) were indicative of a major  
alteration in the transformation performed by this P450, given  
what has been seen in the chemistry performed by other  
P450s (Table S2).<sup>36,37</sup> Given the unexpected presence of the  
*bytA* RiPP precursor gene within the *ruf* BGC, we postulated  
that this pentapeptide may in fact be the substrate for RufO.  
To explore this, we characterized the interaction of RufO with  
this putative pentapeptide substrate (Nle-RYLH [Nle-1]; where  
norleucine [Nle] was used to avoid non-enzyme catalyzed Met  
oxidation; Figures 2 and S6–S11; Tables S3–S6). We used bind-  
ing experiments that monitor the change in spin state of the iron  
within the P450 by UV/visible spectroscopy upon addition of

increasing amounts of substrate peptide to show that the affinity  
of RufO for Nle-1 was comparable to that seen with P450<sub>Bit</sub> (0.31  
vs. 1.1  $\mu$ M; Figure S7), albeit with a shift in the Soret maximum on  
binding to 410 nm (Figures S7 and S8). Such tight binding was  
somewhat unexpected given the bulky residue Val replaces the  
active-site Ser in P450<sub>Bit</sub> that is involved in key H-bonding inter-  
actions with the substrate peptide in P450<sub>Bit</sub>.<sup>32</sup> The effects of this  
alteration in binding were investigated by further analyzing the in-  
teractions of 1 with P450<sub>Bit</sub> and the S239V/V240P mutant using  
molecular dynamics simulations (MDS), which showed 1 adopt-  
ing a conformation in the mutant in which the C-terminal residues  
of 1 orient away from the I-helix. Little alteration in position was  
observed with the Tyr-3 residue despite the mutation of  
Ser239 to Val within these simulations, which was in keeping  
with Tyr-3 being the site of activity of RufO (Figure S12).

Next, we turned to P450-catalyzed turnover experiments to  
clarify the activity of RufO (Tables S3–S6; Figures S13 and  
S34). Attempts to crosslink Nle-1 using standard bacterial redox  
partner proteins (PuxB [ferredoxin]/PuR [ferredoxin reductase])<sup>38</sup>  
together with RufO were not successful, which stands in contrast  
to the high levels of crosslinking seen with these redox partners  
with P450<sub>Bit</sub> (Table S3).<sup>34</sup> Based on the lack of crosslinking activ-  
ity and the putative role of RufO in rufomycin biosynthesis, we  
next explored the potential for RufO to nitrate peptide Nle-1. Us-  
ing protocols reported for the diketopiperazine nitrating P450  
TxE (and making use of the nitrating reagent diethylamine NON-  
Oate [DEANO]<sup>16</sup>), we could observe significant nitration of Nle-1,  
with HR-MS<sup>2</sup> and comparison to authentic standards demon-  
strating that nitration occurs exclusively on the Tyr-3 residue of  
Nle-1 (Figures 2 and S13; Table S3). Determination of the reac-  
tion kinetics (Table S4; Figure S14,  $K_M$  9.6  $\pm$  2.0  $\mu$ M,  $k_{cat}$   
0.25  $\pm$  0.02 min<sup>−1</sup>) revealed values in keeping with what has  
been reported for biosynthetic transformations of bacterial and  
human steroid metabolizing P450 enzymes,<sup>39–41</sup> albeit slower  
than those reported for the peptide crosslinking enzyme OxyB.<sup>42</sup>

The mechanism of TxE has been investigated and is believed  
to be catalyzed by a heme iron(III)-peroxynitrite active species,  
the result of a single-electron reduction of the P450.<sup>16,43</sup> This dif-  
fers to typical P450 chemistry, which is largely the result of the  
activity of a highly electrophilic iron(IV)-oxo porphyrin cation  
radical species (termed compound I) and whose formation re-  
quires two electrons to be delivered from redox partners.<sup>20</sup> To  
explore whether a similar intermediate might be responsible for  
the nitration activity of RufO, we also tested for nitration of  
Nle-1 in the presence of DEANO and dithionite (a single-electron  
reductant) in place of redox partner proteins (Table S5;  
Figures S15–S17).<sup>34</sup> Nitration of Nle-1 was detected from these  
assays, and peroxide was also found to be able to sustain nitration  
activity of RufO, albeit at levels far lower than supported by redox  
partners. Thus, while it is plausible that the nitration mechanism  
of RufO occurs prior to the formation of compound I, and would  
thus be similar to the mechanism reported for TxE, further ex-  
periments are required to explicitly define this.<sup>16,34,37</sup> In this re-  
gard, it is interesting to note that despite the similar reactivity  
of RufO and TxE, these P450s do not contain the same muta-  
tions in their conserved catalytic residues (Table S2).

To further explore the activity of RufO, we next generated a se-  
ries of modified pentapeptide substrates (2–6, Figures S18–S22)



**Figure 2. In vitro characterization of RufO-catalyzed nitration of Nle-1**

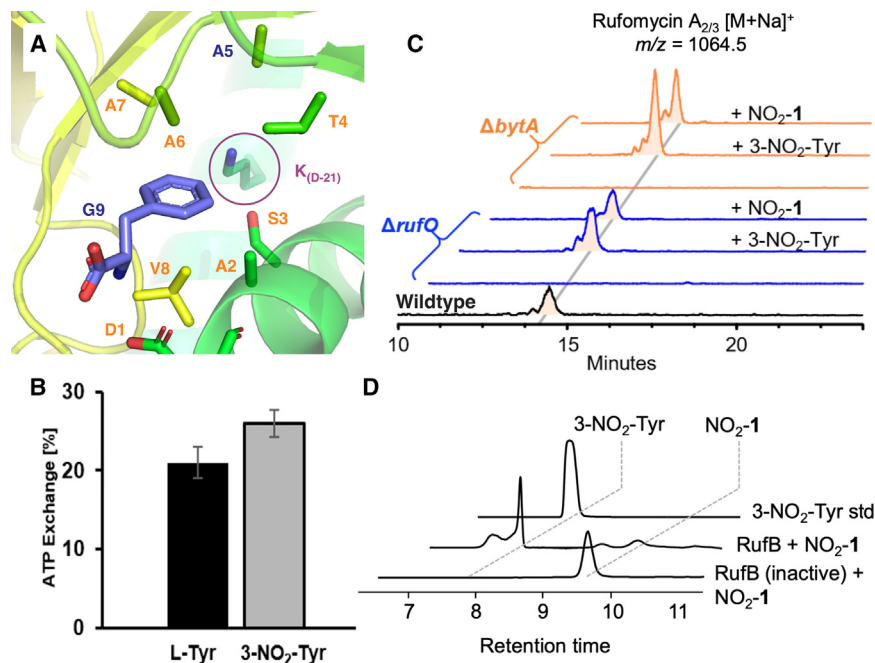
Substrate-binding traces with affinities for RufO and RufO<sub>SV</sub> mutant for Nle-1 (A); turnover traces for RufO, RufO<sub>SV</sub>, and P450<sub>Blt</sub> with Nle-1 (B); and comparison of MS<sup>2</sup> fragmentation of nitrated Nle-1 from RufO turnover with synthetic standard Nle/NO<sub>2</sub>-1 (C).

and tested their acceptance by this P450. The substitution of His-5 for Trp-5 in peptide **2** (MRYLW) demonstrated a reverse in activity for this enzyme, with a lack of nitration detected and instead the enzyme now displaying crosslinking activity with **2** (Table S7; Figure S23). This supports the crucial nature of the altered peptide-binding mode seen with **1** and the P450<sub>Blt</sub> mutant in MD simulations regarding the peptide C terminus. As aromatic residues at position 5 of the peptide are only essential for crosslinking activity in such P450s, we also tested whether the replacement of His-5 with non-aromatic residues (Asn and Gln) could still support nitration (Table S7; Figures S24 and S25). However, turnover experiments with these two probes (**3** and **4**) showed greatly reduced nitration activity, indicating that the positioning of the central Tyr-3 residue is in part maintained by the presence of His-5. This may also explain the strict sequence conservation of the MRYLH peptide in rufomycin BGCs (Table S8). Replacement of the central Tyr-3 residue with Phe-3 (**5**) led to a total loss in nitration activity (Table S7; Figure S26), most likely due to the importance of the phenol moiety for catalysis, while replacement of Leu-4 with Tyr-4 (**6**) revealed nitration remained specific to Tyr-3

despite the presence of an additional Tyr residue in this peptide, albeit with the increase in size of this residue leading to an increase crosslinking seen with **6** (Table S7; Figures S27 and S28).

Having seen the switch in chemistry occurring with RufO and peptide **2** (and to some extent **6**), we next tested the effect of mutation of the altered I-helix residues in RufO to those seen in the crosslinking enzyme P450<sub>Blt</sub>. While a single Val240 to Ser mutant was unable to be produced in soluble form (suggesting the mutation was structurally destabilizing), the Val240Ser/Pro241Val (RufO<sub>SV</sub>) double mutant could be successfully produced, albeit with somewhat reduced yield. The effect of this double mutation was reversion of activity from nitration of Nle-1 to crosslinking (20% conversion; 1:5 nitration to crosslinking ratio when DEANO was present), although curiously the rate enhancement seen with the addition of DEANO in wild-type RufO was maintained in this mutant despite the lack of nitration (Table S3; Figures S29–S34). Low levels of cyclized and nitrated Nle-1 were also detected in these assays, showing an ability for the enzyme to perform sequential transformations. The reverse experiment, using the double mutant





**Figure 3. Evaluating incorporation of NO<sub>2</sub>-Tyr into rufomycin and release from NO<sub>2</sub>-1**  
AlphaFold model of the RufT A<sub>3</sub> domain with overlay of Phe from PheA A-domain (PDB: 1AMU) showing key selection residues 1–9 plus unusual Lys residue (A);  $\gamma^{18}\text{O}_4$ -ATP exchange assay of substrate activation for L-Tyr and 3-NO<sub>2</sub>-Tyr by RufT A<sub>3</sub> domain. Experiments in triplicate, error bars shown. (B); Rufomycin A<sub>2/3</sub> production in *S. atratus* S3 m208\_1 wild-type (black)  $\Delta\text{rufO}$  (blue) and in  $\Delta\text{bytA}$  (orange) mutant strains, with rufomycin biosynthesis restored in the mutants on supplementation with either 3-NO<sub>2</sub>-Tyr or peptide NO<sub>2</sub>-1 (C); RufB specifically digests NO<sub>2</sub>-1 to release the central, modified amino acid 3-NO<sub>2</sub>-Tyr (D). Retention times shown in minutes.

of P450<sub>Blt</sub> bearing the I-helix mutations seen in RufO (Ser239Val, Val240Pro) was not successful, however, with this enzyme unable to catalyze nitration or crosslinking (Table S3). Taken together, these results demonstrate the importance of the typically conserved catalytic I-helix residues in controlling P450 catalysis in biaryllyte pathways toward either nitration or crosslinking and furthermore implied an unexpected route to generate 3-NO<sub>2</sub>-Tyr in rufomycin biosynthesis.

#### RufT module 3 A-domain accepts 3-NO<sub>2</sub>-Tyr

With evidence for the formation of 3-NO<sub>2</sub>-Tyr via RufO-mediated activity toward the peptide Nle-1, we next sought to validate the incorporation of 3-NO<sub>2</sub>-Tyr by the A<sub>3</sub> adenylation domain of the NRPS enzyme RufT. Structural modeling using AlphaFold<sup>44</sup> and analysis of the substrate-binding pocket together with comparison to Tyr-accepting A-domains pointed to an enlarged substrate-binding pocket due to a Cys-to-Gly mutation in the penultimate position of the A-domain specificity code. This mutation would provide additional space for the pendant NO<sub>2</sub> group of 3-NO<sub>2</sub>-Tyr and allowing favorable interactions with an unusual Lys residue found in the A<sub>3</sub> domain and made accessible by this mutation (Figures 3A, S35, and S36). We designed constructs of A<sub>3</sub> with varying lengths and including regions of the adjacent C or peptidyl carrier protein (PCP) domains, which were co-produced with the MbH-like protein (MLP) RufH encoded in the rufomycin BGC.<sup>45</sup> One A-PCP construct was successfully co-produced with RufH and was purified to evaluate its adenylating activity using an  $\gamma^{18}\text{O}_4$ -ATP exchange assay.<sup>46</sup> Unexpectedly, both L-Tyr and L-3-NO<sub>2</sub>-Tyr were adenylated at similar rates (Figure 3B), indicating promiscuity of this A-domain; this also agrees with the large number of natural

and mutasynthesis-derived rufomycin analogs with variations in the Tyr position.<sup>47</sup> Such promiscuity in Tyr-activating domains has previously been reported for the A<sub>6</sub> domain from teicoplanin biosynthesis<sup>48</sup> and suggests possible analogous proofreading by the RufT C<sub>2</sub> domain during peptide extension. These results supported the direct

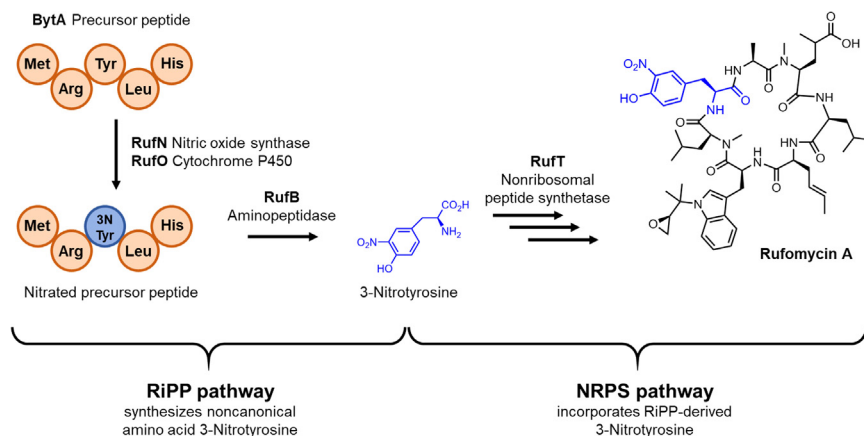
activation of 3-NO<sub>2</sub>-Tyr by the RufT A<sub>3</sub> domain and so we turned to *in vivo* assays to clarify the biosynthesis of 3-NO<sub>2</sub>-Tyr.

#### Role of the BytA peptide in rufomycin biosynthesis

We used *S. atratus* S3\_m208\_1 to test its ability to produce rufomycins under laboratory conditions and further to examine the role of *bytA*, which we reasoned must be fundamental for generation of 3-NO<sub>2</sub>-Tyr and thus for rufomycin biosynthesis. We designed a plasmid for homologous recombination, incorporating a 12-bp deletion in *bytA* and cloned this into pYH7.<sup>49</sup> After conjugation to *S. atratus* S3\_m208\_1 and selective passaging, we obtained clones showing desired recombination. Crucially, *S. atratus* S3\_m208\_1  $\Delta\text{bytA}$  lost the ability to produce the rufomycins, supporting the essential nature of *bytA* for rufomycin biosynthesis. Addition of 3-NO<sub>2</sub>-Tyr or NO<sub>2</sub>-1 to cultures of the mutant could complement this deletion mutant and restore rufomycin biosynthesis, confirming the essential role of the MRYLH peptide 1 (Figures 3C and S37). We then used CRISPR-Cas9-cBest<sup>50</sup> to introduce a specific stop codon into *rufO* and generated the *S. atratus* S3\_m208\_1  $\Delta\text{rufO}$  mutant. This mutant also showed a loss of rufomycin biosynthesis in the same manner as its substrate peptide 1 encoded by *bytA*. Supplementation of 3-NO<sub>2</sub>-Tyr or NO<sub>2</sub>-1 to the  $\Delta\text{rufO}$  mutant could re-establish rufomycin production, further supporting the role of this biaryllyte RiPP pathway in rufomycin biosynthesis (Figures 3C and S38).

#### RufB is an aminopeptidase that cleaves the BytA pentapeptide 1

The availability of 3-NO<sub>2</sub>-Tyr as a substrate for the biosynthesis of rufomycin requires the release of the modified amino acid from the precursor pentapeptide NO<sub>2</sub>-1. Leader peptide removal during RiPP biosynthesis is a diverse process and can be catalyzed



**Figure 4. Summarizing overview of 3-NO<sub>2</sub>-Tyr biosynthesis and its incorporation into rufomycin**

*rufN/rufO* genes (all *ruf* BGCs) and those where these genes are physically separated. In clusters with separated *rufN* and *rufO* homologs, *bytA* is consistently positioned directly upstream of the P450 gene. Among the predominantly two-modular NRPS genes in these clusters, tyrosine is predicted as one of the substrates (70%), supporting the likely incorporation of nitrotyrosine into the respective natural products. The peptidases encoded within these clusters belong to diverse serine aminopeptidase families, explaining why searches based on the peptidase RufB only identified closely related *ruf* BGCs. Some of the detected BGCs lack an encoded peptidase entirely, suggesting that they may utilize peptidases encoded elsewhere in their genomes.

by either specific or nonspecific proteases within the cytoplasm, during secretion or extracellularly.<sup>51</sup> Given the essential role of the *bytA* encoded peptide for rufomycin biosynthesis, we hypothesized that a specific mechanism would be involved in the release of 3-NO<sub>2</sub>-Tyr from the pentapeptide NO<sub>2</sub>-1. Inspection of the rufomycin gene cluster revealed the presence of the RufB enzyme annotated as an  $\alpha/\beta$  fold hydrolase containing a serine aminopeptidase domain (IPR022742). A reported deletion mutant of the *rufB* homolog *ilaC* showed that production of ilamycin (rufomycin) was halved compared with the wild-type producer,<sup>17</sup> although the exact function of this enzyme was not determined. To test our hypothesis that RufB was involved in 3-NO<sub>2</sub>-Tyr release from NO<sub>2</sub>-1, we produced RufB fused with an N-terminal maltose-binding protein tag in *E. coli*<sup>52</sup> that was purified and used in digestion assays with NO<sub>2</sub>-1. HPLC-MS analysis showed that RufB completely digested NO<sub>2</sub>-1 for concomitant release of 3-NO<sub>2</sub>-Tyr (Figures 3D and S39). These results reveal the role of RufB as a BGC-encoded peptidase acting upon the nitrated peptide NO<sub>2</sub>-1, a requirement for the release of 3-NO<sub>2</sub>-Tyr from the BytA peptide and allowing subsequent incorporation by the rufomycin NRPS machinery.

### Genome mining reveals further nitrotyrosine encoding pathways

Finally, we aimed to identify related BGCs that may also encode the biosynthesis of nitrotyrosine and would thus represent further examples for the integration of RiPP and NRPS systems. We conducted a bioinformatic analysis for the NO synthase RufN and the P450 monooxygenase RufO using position-specific iterative BLAST (PSI-BLAST).<sup>53</sup> Due to its small size, the pentapeptide precursor gene *bytA* required manual inspection of candidate clusters rather than conventional BLAST searches. In addition to a number of previously unreported rufomycin BGCs, we detected 19 BGCs containing homologs of the nitrotyrosine biosynthetic system of *bytA* (with always the same sequence MRYLH), *rufN* and *rufO* alongside NRPS and PKS genes, as well as additional biosynthetic genes (Table S8; Figures S2 and S40). Intriguingly, all of the detected clusters exhibit no significant similarity to characterized BGCs. These clusters can be classified into two main groups based on the genomic arrangement of *rufN* and *rufO*: those with overlapping

## DISCUSSION

Natural product biosynthesis demonstrates remarkable examples of biosynthetic ingenuity. While specific biosynthetic pathways can already produce compounds with great chemical complexity, synergizing assembly lines, as seen in NRPS/PKS hybrid systems, can afford yet even more structural diversity. While increasing examples of such synergy have been identified across pathways, one example that has proved more elusive is the integration of peptide biosynthesis from both ribosomal and non-ribosomal pathways. Here, we have investigated the generation of the noncanonical 3-NO<sub>2</sub>-Tyr, critical for antitubercular activity, in rufomycin biosynthesis and could demonstrate how these pathways can be synergized to deliver noncanonical amino acid building blocks for complex peptide natural products (Figure 4).

Recent examples of combining ribosomal peptide biosynthesis with additional pathways include, for example, polycyclic terpene moieties added to RiPP building blocks<sup>11</sup> and the peptide-amino acyl tRNA ligase (PEARL) enzymes.<sup>54–56</sup> These act on ribosomally synthesized peptides to catalyze the attachment of amino acids on the C terminus of a scaffold peptide in an ATP- and tRNA-dependent manner to install previously unprecedented biosynthetic transformations. The biosynthesis of rufomycin represents a new instance of comparable biosynthetic ingenuity and is the first report of a complete RiPP biosynthesis pathway merged within non-ribosomal peptide biosynthesis. Here, the biaryllyl RiPP precursor has acquired a new function as template for noncanonical non-ribosomal peptide building block biosynthesis. This is a shared theme with the pearlyn biosynthetic pathways, where the ribosomal peptide is also used as template for biosynthesis of other natural product building blocks. In the case of pearlyn however, the biosynthetic scaffold may be recycled, which stands in contrast to rufomycin

biosynthesis, where the ribosomal peptide is digested by another dedicated enzyme in the BGC, indicating significantly higher metabolic costs of this pathway.

Of note, this study also establishes a new paradigm for non-proteinogenic amino acid biosynthesis in NRPS pathways. These were so far either found to be synthesized from free proteinogenic precursors before activation by A-domains or synthesized after loading of proteinogenic precursors to carrier proteins by discrete NRPS domains in *cis*, or by modifying enzymes specifically interacting with assembly line domains in *trans*.<sup>57,58</sup>

This study furthermore reveals that P450 enzymes within RiPP pathways can perform specific nitration on peptide substrates with only minimal changes to the active site. The evolution of a peptide crosslinking P450—of which 20 classes have recently been identified<sup>29</sup>—into one capable of selective nitration shows parallels to the ability of radical SAM dependent enzymes to rapidly diversify their function with minimal changes to the enzyme.<sup>59</sup> The identification of mutations to key P450 residues as a fingerprint for alternate catalysis<sup>37</sup> offers important clues to identify examples of diversified P450 function, while demonstrating the power of the P450 active cycle to access different intermediates during catalysis and underlining the versatility of P450s in biosynthetic pathways. The previously failed detection of the *bytA* gene in the rufomycin BGC demonstrates the challenges facing biosynthetic discovery due to the small size of such genes and the current lack of bioinformatic detection pipelines for these minimal RiPPs.

The fact that a RiPP pathway now serves to synthesize a rare amino acid building block for an NRPS—a system classically prized for its ability to install non-proteinogenic amino acids—and that the P450 enzyme central to the pathway has evolved to perform a new catalytic function to accomplish this—once more highlights how nature has developed unexpected solutions for complex biosynthetic problems. It also demonstrates the importance of continued study to identify and characterize such novel examples of biocatalysis, of which there are no doubt many more examples to discover.

## METHODS

### Identification and comparison of ilamycin/rufomycin BGCs and other *rufNO*/*bytA*-containing systems

Amino acid sequences of RufN (XCX13212.1) and RufO (XCX13213.1) encoded in the rufomycin BGC from *Streptomyces atratus* S3\_m208\_1 (PQ034618.1) were used as queries to detect BGCs harboring homologs of NO synthase and the nitrating cytochrome P450 gene using cblaster with default settings.<sup>60</sup> The genome sequences of the obtained hits were collected from the NCBI database and were analyzed using antiSMASH 7.0 to determine whether the respective organisms possess a rufomycin/ilamycin BGC.<sup>61</sup> Additionally, these sequences were manually inspected for the presence of the *bytA* precursor peptide sequence. Rufomycin BGCs were visualized with CAGECAT using a minimum alignment sequence identity of 0.45.<sup>62</sup> To identify more distantly related systems, a PSI-BLAST search<sup>53</sup> using the RufO amino acid sequence (XCX13213.1) and the active-site I-helix motif RAXXHAGXEVPRFL with subsequent manual inspection for the presence of *bytA* and a NO synthase gene was

performed. Obtained hits were collected from the NCBI database and analyzed using antiSMASH 7.0.<sup>61</sup>

### Synthesis of L-tyrosine-(phenyl-3,5-d<sub>2</sub>) and Fmoc-L-tyrosine-(phenyl-3,5-d<sub>2</sub>)

The syntheses of L-tyrosine-(phenyl-3,5-d<sub>2</sub>) and Fmoc-L-tyrosine-(phenyl-3,5-d<sub>2</sub>) were performed according to the previously reported protocol.<sup>33</sup> All spectra matched that previously reported.<sup>33</sup>

### Peptide synthesis

Peptide synthesis was performed manually at 0.1 mmol scale using *N,N'*-diisopropylcarbodiimide (DIC) and Oxyma Pure as coupling reagents. 2-chlorotrityl chloride resin (97 mg) was washed with DCM (3 × 2 mL), swelled in DCM (2 mL, 30 min), and washed again with DCM (3 × 2 mL). The first Fmoc protected amino acid (0.15 mmol) was coupled to the resin overnight using *N,N'*-diisopropylethylamine (DIEA). The unreacted chlorotrityl groups were capped using MeOH by washing with 3 × 2 mL DMF/DIEA/MeOH (7:2:1). After the Fmoc removal step using 20% piperidine in DMF, Fmoc-amino acid coupling (3 equiv) was performed using DIC (3 equiv) and Oxyma Pure (6 equiv) at 50°C for 10 min. The deprotection and coupling reaction was repeated until the sequence was finished, and a final deprotection was performed to remove the Fmoc protecting group of the final amino acid. The cleavage of peptide from resin was performed using 5 mL of TFA/TIS/H<sub>2</sub>O (95:2.5:2.5, v:v:v) solution with shaking for 2 h at room temperature. The separation of resin from the solution was achieved by filtration. After washing the resin with TFA (2 × 1 mL), the filtrate was concentrated under a stream of N<sub>2</sub> to ~1 mL, followed by peptide precipitation with pre-cooled diethyl ether (20 mL), centrifugation in a spark-free centrifuge (Spintron, GT-175) and washing with diethyl ether (3 × 10 mL). Peptides were analyzed by LC-MS (Shimadzu) using a gradient of ACN (0.1% FA) from 0% to 50% over 30 min. For structures see Figures S9 and S18, for analysis of peptides see Figures S10, S11, and S19–S22.

### Site-directed mutagenesis

Point mutations in this study are generated by site-directed mutagenesis. Primer pairs containing the desired mutation were designed and then purchased from Sigma-Aldrich. These primer pairs are used in PCR to generate mutated plasmid containing nicks using Phusion high-fidelity PCR Master Mix with HF Buffer (NEB) or Q5 Hot Start High-Fidelity DNA Polymerase (NEB). The PCR products were treated either with *DpnI* endonuclease (NEB) to digest the methylated parental DNA template when sticky ends were generated, or with KLD Enzyme Mix (NEB) perform the ligation as well as template digestion when blunt ends are generated. After the reaction, the mixture was used for transformation of NEB 10-β competent *E. coli* cells. Following overnight incubation, plasmid DNA from individual colonies was isolated and sequenced to confirm the correct construction of target proteins. Primer sequences can be found in Table S9.

### Expression and purification of P450 enzymes and mutants

The expression and purification of P450<sub>Blt</sub> and mutants were performed as previously described.<sup>34</sup> The expression and purification of RufO also used the same protocol.<sup>34</sup>

### Peptide cyclization and nitration assay

The cyclization assay was performed as previously described.<sup>34</sup> Briefly, HEPES (50 mM), NaCl (50 mM), P450 (0.5  $\mu$ M), PuR (0.5  $\mu$ M), PuxB A105V mutant (2.5  $\mu$ M)<sup>38</sup>, glucose (0.33%), glucose dehydrogenase (0.033 mg/mL), peptide substrate (50  $\mu$ M), and NADH (2 mM) were combined and incubated overnight at 30°C with 300 rpm shaking. For the nitration reaction NO was also supplied in the form of diethylamine NONOate (DEANO, 0.5 mM). For the reaction with sodium dithionite, no PuR, PuxB A105V, glucose or glucose dehydrogenase were added, and sodium dithionite (200  $\mu$ M) was added instead. For reactions with catalase or superoxide dismutase these were present at a final concentration of 1,000 U/mL. For reactions with hydrogen peroxide this was present at a concentration of 50  $\mu$ M. Purification was then performed by solid-phase extraction (Bond Elute, Agilent) that had been activated with 500  $\mu$ L of MeOH (0.1% FA) and equilibrated with 1 mL of water (0.1% FA). The reaction was diluted to 1 mL using water (0.1% FA) before being applied to the column, subsequently washed with 1 mL of water (0.1% FA) and finally eluted with 600  $\mu$ L of MeOH (0.1% FA). The methanol elution fraction was dried using a concentrator (Eppendorf), before being re-dissolved in 50  $\mu$ L of ACN/H<sub>2</sub>O (1:4, 0.1% FA) before LC-MS analysis. All LC-MS analyses were carried out on a Shimadzu high-performance liquid chromatography system coupled to mass spectrometer LC-MS-2020 (ESI, operating both in positive and negative mode) equipped with an SPD-20A Prominence photo diode array detector and an LC-20AD solvent delivery module. Analytical separations were performed using an Agilent ZORBAX 300SB-C<sub>18</sub> 5  $\mu$ m column. All cyclic pentapeptide products were confirmed by high-resolution mass spectrometry (HRMS) analysis.

### NADH consumption assay

The consumption of NADH by P450s was measured using a Jasco V-750 spectrophotometer at 30°C. A cyclization or nitration assay was set up as described above with no addition of glucose, glucose dehydrogenase, NADH, and the peptide substrate. A blank measurement was taken prior to the addition of NADH to a final concentration of 100  $\mu$ M. The absorbance was then measured at 340 nm for 2 min to obtain a background slope ( $Abs_{back}$ ) before peptide substrate (25  $\mu$ M) was added and the absorbance measured for 3 min. The slope of the curve after NADH addition was then determined ( $Abs_{smp}$ ). The  $V_{max}$  and  $K_{cat}$  values were calculated using the molar extinction coefficient of NADH at 340 nm and the following equations:  $V_{max} = (Abs_{smp} - Abs_{back})/0.00622 \times 1$  and  $K_{cat} = V_{max}/[P450]$ .

### Michaelis-Menten kinetic experiments

Nitration reactions at varying concentrations of peptide (3.125–100  $\mu$ M) were set up as described above with RufO<sub>WT</sub> (0.2  $\mu$ M), PuxB AV+ (1  $\mu$ M), PuR (0.2  $\mu$ M) and DEANO (10 times excess to peptide concentration). Samples were taken at time points of 0, 10, 20, and 30 min and reactions were quenched with methanol + 0.1% FA. The samples were then dried and analyzed by LC-MS to determine the initial rate of reaction at each concentration of peptide. The initial rate of reaction ( $v_0$ ) was plotted against substrate concentration ( $[S]$ ) and a

Lineweaver-Burk plot was constructed. The kinetic measurements,  $K_M$ ,  $k_{cat}$ , and  $K_M/k_{cat}$  were then determined.

### P450 CO difference spectra

Spectra were obtained using a Jasco V-750 spectrophotometer at 30°C. The enzyme was diluted to 2.5  $\mu$ M using Tris-HCl (50 mM, pH 7.4), and a baseline spectrum obtained. CO was generated by the addition of a small volume of disodium boranocarbonate, and the P450 reduced using 20  $\mu$ L of freshly prepared saturated sodium dithionite solution. The UV-vis spectra of CO-complex was then measured until the spectra was stable.

### P450/substrate binding determined via UV-vis spectroscopy

The interaction of substrates with P450s were measured using UV-visible absorbance spectroscopy on a Jasco V-750 spectrophotometer at 30°C. The P450 enzyme was diluted to 0.5  $\mu$ M for RufO and 1  $\mu$ M for P450<sub>BIT</sub> using Tris-HCl (50 mM, pH 8.0) and split into two cuvettes. After being equilibrated to 30°C, a baseline measurement was obtained, and the initial spectra measured between  $\lambda = 350$  and 600 nm. The substrate was then added into the sample cuvette in a stepwise manner, with the same amount of buffer added into the reference cuvette. The spectra were measured each step after a 2 min equilibration period and normalized against the 600-nm absorbance. The resulting amplitudes of the different spectra were extrapolated ( $\Delta A = A_{max} - A_{min}$ ) and plotted against the ligand concentration. Curve fitting of the resulting data points was performed using GraphPad Prism 9 with a one-site binding model:  $Y = B_{max} \times X/(K_D + X) + NS \times X + Background$  ( $Y = \Delta A$ ,  $B_{max}$  = the maximum specific binding,  $K_D$  = equilibrium dissociation constant,  $NS$  = slope of nonspecific binding,  $Background$  = nonspecific binding with no added ligand).

### HRMS analysis of *in vitro* turnovers

High-resolution mass spectrometry measurements were performed on an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled to a Dionex UltiMate 3000 RSLCnano system equipped with a Dionex UltiMate 3000 RS autosampler, an Acclaim PepMap RSLC analytical column (75  $\mu$ m  $\times$  50 cm, nanoViper, C18, 2  $\mu$ m, 100 Å; Thermo Scientific) and an Acclaim PepMap 100 trap column (100  $\mu$ m  $\times$  2 cm, nanoViper, C18, 5  $\mu$ m, 100 Å; Thermo Scientific). The peptides were separated by increasing concentrations of 80% acetonitrile/0.1% formic acid at a flow of 250 nL.min<sup>-1</sup> over 30 min. The instrument was operated in alternating data-dependent acquisition (DDA) and parallel reaction monitoring (PRM) cycles in such that 5 MS<sup>2</sup> scans were triggered per survey MS<sup>1</sup> scan followed by several targeted MS<sup>2</sup> scans to ensure fragmentation of predefined, sample-dependent  $m/z$  precursors. Each survey MS<sup>1</sup> scan (300–1,800  $m/z$ ) was acquired with a resolution of 240,000 and a normalized AGC (automatic gain control) target of 200%. Dynamic exclusion was set to 10 s after one occurrence. The 5 most intense ions were selected for HCD fragmentation (fixed collision energy mode, 24% HCD Collision Energy) with a resolution of 15,000, a normalized AGC target of 200% and a fixed first mass of 100  $m/z$ . Subsequent targeted MS<sup>2</sup> scans were acquired with essentially identical settings. The raw data files



were analyzed with QualBrowser (XCalibur 3.0.63, Thermo Scientific) to view spectra and to generate extracted ion chromatograms.

### Protein preparation

All the modeling work, including MDS, was performed in Maestro (Schrödinger Release 2021–3: Maestro, Schrödinger, LLC, New York, NY, 2021). Protein preparation wizard (Schrödinger Release 2021–3: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2021; Impact, Schrödinger, LLC, New York, NY, 2021; Prime, Schrödinger, LLC, New York, NY, 2021) was accessed to perform structural edits on one monomer of the protein, which involved correcting missing hydrogens and bond orders, adding missing sidechains and zero-order bonds to heme Fe<sup>2+</sup> ion and finally assigning protonation states at a pH of 7.4 using Epik. The hydrogen bond network of the prepared structure was further optimized using PROPKA and minimized by converging the het atoms at an RMSD of 0.3 Å.

### Molecular modeling studies

The Build module (Schrödinger Release 2021–3: Maestro, Schrödinger, LLC, New York, NY, 2021) was utilized to selectively mutate the Ser239 and Val240 to Val239 and Pro240 in the P450<sub>BIT</sub>–MRYLH complex, respectively. The Ser239Val/Val240Pro mutant model was energy-minimized using the Prime (Schrödinger Release 2021–3: Prime, Schrödinger, LLC, New York, NY, 2021) module to generate the starting representative models for MDS.

### Molecular Dynamics Simulations

The representative model was assigned a simulation box using the System Builder module (Schrödinger Release 2021–3: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2021). The TIP4P water model was utilized, with periodic boundary conditions involving an orthorhombic box shape using the buffer method (12 × 12 × 12 Å). After minimizing the volume, the net charge was neutralized by adding 3 Na<sup>+</sup> ions, and the system was supplemented with 0.15 M NaCl. The system was minimized for 1 ns before the simulation. Each system was simulated for 300 ns (in three replicates), generating 1,002 frames each, at a temperature of 300 K and a pressure of 1.01325 bar. The system was first equilibrated using the NPT ensemble, with the Nose-Hoover chain thermostat and the Martyna-Tobias-Klein barostat chosen for the simulation. The time step was set to 2.0 fs, and the cutoff method (cutoff radius 9.0 Å) was selected to define the short-range Coulombic interactions. All replicates were randomized, and the trajectories were centered around the substrate. All the molecular visualizations were performed using PyMOL (Schrödinger, LLC, [2020]. The PyMOL Molecular Graphics System, version 2.5).

### Cloning, expression, and peptide digestion assay of RufB

Genomic DNA from *S. atratus* S3\_m208\_1 was extracted using the Wizard Genomic DNA Purification Kit (Promega) according to the manufacturer's instructions. *E. coli* NEB10β cells were made chemically competent using the Inoue method.<sup>63</sup> The *rufB* DNA sequence was codon-optimized for *E. coli* (Thermo Fisher Scientific) and amplified using Q5-Polymerase (NEB). The pET28a-MBP-TEV vector (Addgene) was linearized with PCR. The purified linear vector and insert were mixed at a 1:3 ratio with 15 μL of Gibson assembly master mixture (1 ×

ISO buffer, 10 U/L T5 exonuclease, 2 U/L Phusion polymerase, and 40 U/μL Taq ligase and incubated at 50°C for 1 h to assemble pET28a-MBP-TEV-RufB. Then, 5 μL of the reaction mix was used to transform chemically competent *E. coli* NEB10β cells, which were plated on Luria-Bertani (LB) agar containing 50 μg/mL apramycin. Positive clones containing the plasmid pET28-RufB were screened by colony PCR and confirmed by Sanger sequencing. For expression, a 3 mL inoculum was added to 1 L of LB medium containing 50 μg/mL kanamycin and grown at 180 rpm and 37°C for ~4 h until the optical density at 600 nm reached 0.6. The culture was then placed on ice for 30 min and induced with 0.05 mM IPTG, followed by incubation at 18°C for 16 h at 180 rpm. Cells were harvested by centrifugation, and the cell pellet was resuspended in lysis buffer (50 mM Tris, 150 mM NaCl, 10 mM imidazole, 1 mM DTT, pH 7.5), supplemented with 1 mg/mL lysozyme and disrupted by sonication. Cell debris was removed by centrifugation (30 min, 15,000 rcf, 4°C). The supernatant was incubated with 3 mL of Ni-NTA resin equilibrated with lysis buffer for 1 h at 4°C, washed with five column volumes of wash buffer I (50 mM Tris, 150 mM NaCl, 20 mM imidazole, pH 7.5) and wash buffer II (wash buffer I with 35 mM imidazole). The protein was eluted with 4 mL of elution buffer (50 mM Tris, 150 mM NaCl, 300 mM imidazole, pH 7.5), concentrated and buffer-exchanged to remove imidazole using a 30 kDa MWCO centrifuge filter unit (Amicon) into storage buffer (50 mM Tris, 150 mM NaCl, 5% glycerol, pH 7.5), and treated with 1.5 U thrombin to 200 μg protein and incubated for 16 h at 4°C. 0.5 mL of Ni-NTA resin was added to remove impurities MBP-RufB was collected in the flow through fraction and flash-frozen in liquid nitrogen until further use. For the peptide digestion assay, reaction components were combined 100 μL total volume and incubated at 30°C shaking at 300 rpm with the following final concentrations: RufB (5 μM), NO<sub>2</sub>-1 (100 μM), Tris (50 mM), and NaCl (50 mM). After 16 h, 40 μL assay solution was mixed with 80 μL water + 0.1% FA and 40 μL acetonitrile + 0.2% FA and the protein was removed by centrifugation. For analysis of peptide fragments, 10 μL samples were added to 90 μL H<sub>2</sub>O + 0.1% FA and extracted with Pierce C18 (Thermo Fisher Scientific) according to the manufacturer's instruction. 50 μL were injected and analyzed on the HR-LC-MS. For detection of single amino acids, the 10 μL sample was added to 20 μL water + 0.1% FA and 10 μL acetonitrile and centrifuged for 10 min at 15,000 rcf. The supernatant was filtered and analyzed by LC-MS using an Agilent 6550 Accurate Mass Q-TOF with ESI-source coupled with an Agilent 1100 HPLC using a Waters Atlantis T3, 5 μm, 4.6 × 50 mm column. The column temperature was 25°C. MS data were acquired over a range from 50–2,500 m/z in positive mode. Auto MS/MS fragmentation was achieved with rising collision energy (35–50 keV over a gradient from 500–2,000 m/z) with a frequency of 3 Hz for all ions over a threshold of 500. Elution started with a flow of 0.5 mL/min at 90% H<sub>2</sub>O containing 0.1% acetic acid (A) and 10% acetonitrile containing 0.1% acetic acid (B). After 2 min, a linear gradient to 100% B was run for 10 min, followed by an isocratic step with 100% B for 6 min. 20–50 μL were injected per run depending on the concentration of the sample.

### Cloning, expression, and $\gamma^{18}\text{O}_4$ -ATP exchange assay of RufTA<sub>3</sub>PCP<sub>3</sub>/RufH

DNA sequences of RufTA<sub>3</sub>PCP<sub>3</sub> and RufH were amplified from genomic DNA with Q5-Polymerase and cloned into pET28 and pRSFDuet-1 (Novagen) for co-expression in a single *E. coli* host using Gibson assembly as described above, to generate pET28- RufTA<sub>3</sub>PCP<sub>3</sub> and pRSFDuet-1-RufH, which were confirmed for correct assembly by Sanger sequencing. Protein expression and purification was performed analogously to RufB described above. For the  $\gamma^{18}\text{O}_4$ -ATP exchange assay,<sup>34</sup> three solutions were prepared: Solution 1 with 3 mM  $\gamma^{18}\text{O}_4$ -ATP and 15 mM MgCl<sub>2</sub> in 20 mM Tris (pH 7.5); Solution 2 with 3 mM amino acid and 15 mM Na-pyrophosphate in 20 mM Tris (pH 7.5); and Solution 3 with 5  $\mu\text{M}$  of the enzyme in 5% (v/v) glycerol and 20 mM Tris (pH 7.5). Each solution (2  $\mu\text{L}$ ) was mixed and incubated at 22°C for 90 min. The negative control contained Na-pyrophosphate but no amino acid in solution 2. The reaction was stopped by adding 6  $\mu\text{L}$  of 9-aminoacridine in acetone (10 mg/mL) and centrifuged at 15,000  $\times$  g for 5 min to remove precipitated enzymes. The assay was analyzed by MALDI-TOF mass spectrometry using an AutoFlex III MALDI-ToF/ToF (Bruker) in negative mode, with measurements taken from  $m/z$  400–1,200 and data analyzed with DataAnalysis 4.2 (Bruker). Absolute substrate conversion was calculated from the integrals of peaks at  $m/z$  506 for  $\gamma^{16}\text{O}_4$ -ATP and  $m/z$  508, 510, 512, 514 for  $\gamma^{18}\text{O}_4$ -ATP, using the formula absolute substrate conversion =  $^{16}\text{O}/(^{16}\text{O}+^{18}\text{O})$ . Given the equilibrium molar ratio of unlabeled pyrophosphate to  $\gamma^{18}\text{O}_4$ -ATP in the assay is 5:1, an 83.33% detected substrate conversion corresponds to 100% substrate exchange, calculated as % exchange = (absolute substrate conversion)/0.8333.

### Deletion mutants of *rufO* and *bytA*

The generation of knockout mutant strains of *S. atratus* S3\_m208\_1 was performed using homologous recombination and CRISPR-cBEST base editing.<sup>64</sup> Initially, Cas9 single guide RNA protospacer sequences were selected with CRISPy-web.<sup>65</sup> The top three sequences were ordered as oligonucleotides with cloning restriction sites and primer sequences for sgRNA amplification. These sgRNAs were PCR-amplified, digested with *NcoI* and *SnaBI*, ligated into similarly digested pCRISPR-cBEST-v2-PermE\*, and transformed into *E. coli* NEB10 $\beta$ . Correct assembly was confirmed by Sanger sequencing, followed by transformation into *E. coli* ET12567 pUZ8002 for conjugation into *S. atratus* S3\_m208\_1. Apramycin-resistant exconjugants were screened for the stop codon introduction by PCR and confirmed by Sanger sequencing. Due to the small size of *bytA*, the use of CRISPR based systems was not practical; therefore, we decided to use the *Streptomyces-E.-coli* shuttle vector pYH7.<sup>49</sup> After amplification of homologous regions excluding a 12-bp sequence of *bytA*, the fragments were cloned into pYH7 with Gibson assembly (to yield pYH7\_Δ*bytA*). This plasmid was transformed into *E. coli* NEB10 $\beta$  and transformed into *Streptomyces atratus* S3\_m208\_1 via conjugation.

### Feeding experiments

For experiments with deletion mutants *Streptomyces atratus* S3\_m208\_1 *rufO*<sup>stop</sup> and *Streptomyces atratus* S3\_m208\_1

Δ*bytA*, a preculture was inoculated from mycelium in tryptic soy broth and cultivated for 3 days at 30°C and 180 rpm. Cultivation of main cultures were performed in 5 mL GSS medium (soluble starch 30g/L, glucose 20 g/L, soybean flour 10 g/L, peptone 5g/L, NaCl 3 g/L, CaCO<sub>3</sub> 5g/L, corn steep liquor 10 mL) and addition of substrate of interest (0.5 mM final concentration) in 25 mL Erlenmeyer flasks for 7 days at 30°C and 200 rpm. Cultures were extracted three times with ethyl acetate. Extracts were dissolved in methanol for a final concentration of 2 mg/mL and analyzed with HPLC-MS using a Waters Alliance e2695 Separation Module, equipped with a RP-18 column (Waters XBridge BEH Shield RP18 Column; 3.5  $\mu\text{m}$ , 2.1  $\times$  100 mm). A binary solvent mixture of A: 90% acetonitrile + 10% ammonium acetate (5 mM) pH 7.4 and B: 10% acetonitrile + 90% ammonium acetate (5 mM) pH 7.4 was used, with a flow rate of 0.3 mL/min, starting at minute 0–5 with isocratic elution of 10% B, a gradient 10% B  $\rightarrow$  100% at minute 5–25, and again isocratic conditions of 100% B at minute 25–30.

### RESOURCE AVAILABILITY

#### Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Max Crüsemann ([cruesemann@em.uni-frankfurt.de](mailto:cruesemann@em.uni-frankfurt.de)).

#### Data availability

See [supplemental information](#).

#### Materials availability

The mass spectrometry data for *in vitro* P450 enzyme assays have been deposited into the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD053820. The sequence of the rufomycin BGC has been deposited to GenBank under accession number PQ034618.

### ACKNOWLEDGMENTS

We thank Aurélien Carlier (INRAE, Toulouse) for providing *S. atratus* S3\_m208\_1; Tilmann Weber (DTU Copenhagen) for providing CRISPR-Cas9-cBest plasmids; Mathias Hansen (Monash) and James De Voss (UQ) for helpful discussions; and Stefan Kehraus (Uni Bonn) for LC-MS measurements. We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for funding through the Heisenberg-Programme, project number 495740318 (to M.C.); the German Academic Scholarship Foundation for a PhD scholarship (to L.P.); the Deutsche Bundesstiftung Umwelt (DBU) for a PhD scholarship (to L.Z.); and Monash University and EMBL Australia (to M.J.C.). This study used Bioplatforms Australia (BPA)-enabled/National Collaborative Research Infrastructure Strategy (NCRIS)-enabled infrastructure located at the Monash Proteomics and Metabolomics Platform. This research was conducted by the Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012) and funded by the Australian Government.

### AUTHOR CONTRIBUTIONS

L.P.: *in vitro* P450 assays, *in vivo* experiments, *in vitro* RufB and A-domain assays, and figure and manuscript preparation. L.Z.: gene cluster analysis, *in vivo* experiments, and *in vitro* RufB and A-domain assays. J.G. and Y.Z.: synthesis, *in vitro* P450 assays and analysis. R.B.S.: HRMS analysis. V.M.S. and C.J.J.: MD simulations. M.C. and M.J.C.: analysis and figure and manuscript preparation.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.chempr.2025.102438>.

Received: September 13, 2024

Revised: December 17, 2024

Accepted: January 14, 2025

Published: February 20, 2025

## REFERENCES

- Wang, L., Wang, N., Zhang, W., Cheng, X., Yan, Z., Shao, G., Wang, X., Wang, R., and Fu, C. (2022). Therapeutic peptides: current applications and future directions. *Signal Transduct. Target. Ther.* 7, 48. <https://doi.org/10.1038/s41392-022-00904-4>.
- Süssmuth, R.D., and Mainz, A. (2017). Nonribosomal peptide synthesis – challenges and prospects. *Angew. Chem. Int. Ed. Engl.* 56, 3770–3821. <https://doi.org/10.1002/anie.201609079>.
- Dell, M., Dunbar, K.L., and Hertweck, C. (2022). Ribosome-independent peptide biosynthesis: the challenge of a unifying nomenclature. *Nat. Prod. Rep.* 39, 453–459. <https://doi.org/10.1039/d1np00019e>.
- Montalbán-López, M., Scott, T.A., Ramesh, S., Rahman, I.R., van Heel, A.J., Viel, J.H., Bandarian, V., Dittmann, E., Genilloud, O., Goto, Y., et al. (2021). New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* 38, 130–239. <https://doi.org/10.1039/d0np00027b>.
- Bhushan, A., Egli, P.J., Peters, E.E., Freeman, M.F., and Piel, J. (2019). Genome mining- and synthetic biology-enabled production of hypermodified peptides. *Nat. Chem.* 11, 931–939. <https://doi.org/10.1038/s41557-019-0323-9>.
- Bösch, N.M., Borsa, M., Greczmiel, U., Morinaka, B.I., Gugger, M., Oxenius, A., Vagstad, A.L., and Piel, J. (2020). Landomamides: Antiviral ornithine-containing ribosomal peptides discovered through genome mining. *Angew. Chem. Int. Ed. Engl.* 59, 11763–11768. <https://doi.org/10.1002/anie.201916321>.
- Hubrich, F., Bösch, N.M., Chepkirui, C., Morinaka, B.I., Rust, M., Gugger, M., Robinson, S.L., Vagstad, A.L., and Piel, J. (2022). Ribosomal derived lipopeptides containing distinct fatty acyl moieties. *Proc. Natl. Acad. Sci. USA* 119, e2113120119. <https://doi.org/10.1073/pnas.2113120119>.
- Ayikpoe, R.S., Zhu, L., Chen, J.Y., Ting, C.P., and van der Donk, W.A. (2023). Macrocyclization and backbone rearrangement during RiPP biosynthesis by a SAM-dependent domain-of-unknown-function 692. *ACS Cent. Sci.* 9, 1008–1018. <https://doi.org/10.1021/acscentsci.3c00160>.
- Acedo, J.Z., Bothwell, I.R., An, L., Trouth, A., Frazier, C., and van der Donk, W.A. (2019). O-methyltransferase-mediated incorporation of a  $\beta$ -amino acid in lanthipeptides. *J. Am. Chem. Soc.* 141, 16790–16801. <https://doi.org/10.1021/jacs.9b07396>.
- An, L., Cogan, D.P., Navo, C.D., Jiménez-Osés, G., Nair, S.K., and van der Donk, W.A. (2018). Substrate-assisted enzymatic formation of lysinoalanine in duramycin. *Nat. Chem. Biol.* 14, 928–933. <https://doi.org/10.1038/s41589-018-0122-4>.
- Hubrich, F., Kandy, S.K., Chepkirui, C., Padhi, C., Mordhorst, S., Moosmann, P., Zhu, T., Gugger, M., Chekan, J.R., and Piel, J. (2024). Ribosomal peptides with polycyclic isoprenoid moieties. *Chem* 10, 3224–3242. <https://doi.org/10.1016/j.chempr.2024.07.026>.
- Mordhorst, S., Ruijter, F., Vagstad, A.L., Kuipers, O.P., and Piel, J. (2023). Emulating nonribosomal peptides with ribosomal biosynthetic strategies. *RSC Chem. Biol.* 4, 7–36. <https://doi.org/10.1039/d2cb00169a>.
- Caranto, J.D. (2019). The emergence of nitric oxide in the biosynthesis of bacterial natural products. *Curr. Opin. Chem. Biol.* 49, 130–138. <https://doi.org/10.1016/j.cbpa.2018.11.007>.
- Ferrer-Sueta, G., Campolo, N., Trujillo, M., Bartesaghi, S., Carballal, S., Romero, N., Alvarez, B., and Radi, R. (2018). Biochemistry of peroxynitrite and protein tyrosine nitration. *Chem. Rev.* 118, 1338–1408. <https://doi.org/10.1021/acs.chemrev.7b00568>.
- Kers, J.A., Wach, M.J., Krasnoff, S.B., Widom, J., Cameron, K.D., Bukhalid, R.A., Gibson, D.M., Crane, B.R., and Loria, R. (2004). Nitration of a peptide phytotoxin by bacterial nitric oxide synthase. *Nature* 429, 79–82. <https://doi.org/10.1038/nature02504>.
- Barry, S.M., Kers, J.A., Johnson, E.G., Song, L., Aston, P.R., Patel, B., Krasnoff, S.B., Crane, B.R., Gibson, D.M., Loria, R., and Challis, G.L. (2012). Cytochrome P450-catalyzed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis. *Nat. Chem. Biol.* 8, 814–816. <https://doi.org/10.1038/nchembio.1048>.
- Ma, J., Huang, H., Xie, Y., Liu, Z., Zhao, J., Zhang, C., Jia, Y., Zhang, Y., Zhang, H., Zhang, T., and Ju, J. (2017). Biosynthesis of ilamycins featuring unusual building blocks and engineered production of enhanced anti-tuberculosis agents. *Nat. Commun.* 8, 391. <https://doi.org/10.1038/s41467-017-00419-5>.
- Tomita, H., Katsuyama, Y., Minami, H., and Ohnishi, Y. (2017). Identification and characterization of a bacterial cytochrome P450 monooxygenase catalyzing the 3-nitration of tyrosine in rufomycin biosynthesis. *J. Biol. Chem.* 292, 15859–15869. <https://doi.org/10.1074/jbc.M117.791269>.
- Greule, A., Stok, J.E., De Voss, J.J., and Cryle, M.J. (2018). Unrivalled diversity: the many roles and reactions of bacterial cytochromes P450 in secondary metabolism. *Nat. Prod. Rep.* 35, 757–791. <https://doi.org/10.1039/c7np00063d>.
- Rittle, J., and Green, M.T. (2010). Cytochrome P450 compound I: capture, characterization, and C-H bond activation kinetics. *Science* 330, 933–937. <https://doi.org/10.1126/science.1193478>.
- Zhou, B., Shetye, G., Yu, Y., Santarsiero, B.D., Klein, L.L., Abad-Zapatero, C., Wolf, N.M., Cheng, J., Jin, Y., Lee, H., et al. (2020). Antimycobacterial rufomycin analogues from *Streptomyces atratus* strain MJM3502. *J. Nat. Prod.* 83, 657–667. <https://doi.org/10.1021/acs.jnatprod.9b01095>.
- Wolf, N.M., Lee, H., Choules, M.P., Pauli, G.F., Phansalkar, R., Anderson, J.R., Gao, W., Ren, J., Santarsiero, B.D., Lee, H., et al. (2019). High-resolution structure of ClpC1-rufomycin and ligand binding studies provide a framework to design and optimize anti-tuberculosis leads. *ACS Infect. Dis.* 5, 829–840. <https://doi.org/10.1021/acsinfecdis.8b00276>.
- Park, C.R., Paik, S., Kim, Y.J., Kim, J.K., Jeon, S.M., Lee, S.-H., Whang, J., Cheng, J., Suh, J.-W., Cao, J., et al. (2021). Rufomycin exhibits dual effects against *Mycobacterium abscessus* infection by inducing host defense and antimicrobial activities. *Front. Microbiol.* 12, 695024. <https://doi.org/10.3389/fmicb.2021.695024>.
- Choules, M.P., Wolf, N.M., Lee, H., Anderson, J.R., Grzelak, E.M., Wang, Y., Ma, R., Gao, W., McAlpine, J.B., Jin, Y.-Y., et al. (2019). Rufomycin targets ClpC1 proteolysis in *Mycobacterium tuberculosis* and *M. abscessus*. *Antimicrob. Agents Chemother.* 63, e02204-18. <https://doi.org/10.1128/AAC.02204-18>.
- Haslinger, K., Brieke, C., Uhlmann, S., Sieverling, L., Süssmuth, R.D., and Cryle, M.J. (2014). The structure of a transient complex of a nonribosomal peptide synthetase and a cytochrome P450 monooxygenase. *Angew. Chem. Int. Ed. Engl.* 53, 8518–8522. <https://doi.org/10.1002/anie.201404977>.
- Jordan, S., Li, B., Traore, E., Wu, Y., Usai, R., Liu, A., Xie, Z.-R., and Wang, Y. (2023). Structural and spectroscopic characterization of RufO indicates a new biological role in rufomycin biosynthesis. *J. Biol. Chem.* 299, 105049. <https://doi.org/10.1016/j.jbc.2023.105049>.
- Dratch, B.D., McWhorter, K.L., Blue, T.C., Jones, S.K., Horwitz, S.M., and Davis, K.M. (2023). Insights into substrate recognition by the unusual nitrating enzyme RufO. *ACS Chem. Biol.* 18, 1713–1718. <https://doi.org/10.1021/acscmbio.3c00328>.



28. Zdouc, M.M., Alanjary, M.M., Zarazúa, G.S., Maffioli, S.I., Crüsemann, M., Medema, M.H., Donadio, S., and Sosio, M. (2021). A biaryl-linked tripeptide from *Planomonospora* reveals a widespread class of minimal RiPP gene clusters. *Cell Chem. Biol.* **28**, 733–739.e4. <https://doi.org/10.1016/j.chembiol.2020.11.009>.
29. Nam, H., An, J.S., Lee, J., Yun, Y., Lee, H., Park, H., Jung, Y., Oh, K.-B., Oh, D.-C., and Kim, S. (2023). Exploring the diverse landscape of biaryl-containing peptides generated by cytochrome P450 macrocyclases. *J. Am. Chem. Soc.* **145**, 22047–22057. <https://doi.org/10.1021/jacs.3c07140>.
30. Hu, Y.L., Yin, F.Z., Shi, J., Ma, S.Y., Wang, Z.R., Tan, R.X., Jiao, R.H., and Ge, H.M. (2023). P450-modified ribosomally synthesized peptides with aromatic cross-links. *J. Am. Chem. Soc.* **145**, 27325–27335. <https://doi.org/10.1021/jacs.3c07416>.
31. He, B.-B., Liu, J., Cheng, Z., Liu, R., Zhong, Z., Gao, Y., Liu, H., Song, Z.-M., Tian, Y., and Li, Y.-X. (2023). Bacterial cytochrome P450 catalyzed post-translational macrocyclization of ribosomal peptides. *Angew. Chem. Int. Ed. Engl.* **62**, e202311533. <https://doi.org/10.1002/anie.202311533>.
32. Hansen, M.H., Keto, A., Treisman, M., Sasi, V.M., Coe, L., Zhao, Y., Padva, L., Hess, C., Leichthammer, V., Machell, D.L., et al. (2024). Structural insights into a side chain cross-linking biaryllytic P450 from RiPP biosynthesis. *ACS Catal.* **14**, 812–826. <https://doi.org/10.1021/acscatal.3c05417>.
33. Treisman, M., Coe, L., Zhao, Y., Sasi, V.M., Gullick, J., Hansen, M.H., Ly, A., Leichthammer, V., Hess, C., Machell, D.L., et al. (2024). An engineered biaryllytic cross-linking P450 from RiPP biosynthesis generates alternative cyclic peptides. *Org. Lett.* **26**, 1828–1833. <https://doi.org/10.1021/acs.orglett.3c04366>.
34. Zhao, Y., Marschall, E., Treisman, M., McKay, A., Padva, L., Crüsemann, M., Nelson, D.R., Steer, D.L., Schittenhelm, R.B., Tailhades, J., and Cryle, M.J. (2022). Cytochrome P450Blt enables versatile peptide cyclisation to generate histidine- and tyrosine-containing crosslinked tripeptide building blocks. *Angew. Chem. Int. Ed. Engl.* **61**, e202204957. <https://doi.org/10.1002/anie.202204957>.
35. Qi, S.S., Bogdanov, A., Cnockaert, M., Acar, T., Ranty-Roby, S., Coenye, T., Vandamme, P., König, G.M., Crüsemann, M., and Carlier, A. (2021). Induction of antibiotic specialized metabolism by co-culturing in a collection of phyllosphere bacteria. *Environ. Microbiol.* **23**, 2132–2151. <https://doi.org/10.1111/1462-2920.15382>.
36. Lee, D.S., Nioche, P., Hamberg, M., and Raman, C.S. (2008). Structural insights into the evolutionary paths of oxylipin biosynthetic enzymes. *Nature* **455**, 363–368. <https://doi.org/10.1038/nature07307>.
37. Gering, H.E., Li, X., Tang, H., Swartz, P.D., Chang, W.C., and Makris, T.M. (2023). A Ferric-superoxide intermediate initiates P450-catalyzed cyclic dipeptide dimerization. *J. Am. Chem. Soc.* **145**, 19256–19264. <https://doi.org/10.1021/jacs.3c04542>.
38. Bell, S.G., Xu, F., Johnson, E.O.D., Forward, I.M., Bartlam, M., Rao, Z., and Wong, L.L. (2010). Protein recognition in ferredoxin-P450 electron transfer in the class I CYP199A2 system from *Rhodospseudomonas palustris*. *J. Biol. Inorg. Chem.* **15**, 315–328. <https://doi.org/10.1007/s00775-009-0604-7>.
39. Yoshimoto, F.K., and Auchus, R.J. (2016). Rapid kinetic methods to dissect steroidogenic cytochrome P450 reaction mechanisms. *J. Steroid Biochem. Mol. Biol.* **161**, 13–23. <https://doi.org/10.1016/j.jsbmb.2015.10.005>.
40. Han, S., Pham, T.-V., Kim, J.-H., Lim, Y.-R., Park, H.-G., Cha, G.-S., Yun, C.-H., Chun, Y.-J., Kang, L.-W., and Kim, D. (2015). Functional characterization of CYP107W1 from *Streptomyces avermitilis* and biosynthesis of macrolide oligomycin A. *Arch. Biochem. Biophys.* **575**, 1–7. <https://doi.org/10.1016/j.abb.2015.03.025>.
41. Yao, Q., Ma, L., Liu, L., Ikeda, H., Fushinobu, S., Li, S., and Xu, L.-H. (2017). Hydroxylation of compactin (ML-236B) by CYP105D7 (SAV\_7469) from *Streptomyces avermitilis*. *J. Microbiol. Biotechnol.* **27**, 956–964. <https://doi.org/10.4014/jmb.1610.10079>.
42. Peschke, M., Haslinger, K., Brieke, C., Reinstein, J., and Cryle, M.J. (2016). Regulation of the P450 Oxygenation Cascade Involved in Glycopeptide Antibiotic Biosynthesis. *J. Am. Chem. Soc.* **138**, 6746–6753. <https://doi.org/10.1021/jacs.6b00307>.
43. Mondal, P., Udukale, D., Mohamed, A.A., Wong, H.P.H., de Visser, S.P., and Wijeratne, G.B. (2024). A Cytochrome P450 TxtE Model System with Mechanistic and Theoretical Evidence for a Heme Peroxynitrite Active Species. *Angew. Chem. Int. Ed. Engl.* **63**, e202409430. <https://doi.org/10.1002/anie.202409430>.
44. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
45. Kreidler, D.F., Gemmell, E.M., Schaffer, J.E., Wenciewicz, T.A., and Gulick, A.M. (2019). The structural basis of N-acyl- $\alpha$ -amino- $\beta$ -lactone formation catalyzed by a nonribosomal peptide synthetase. *Nat. Commun.* **10**, 3432. <https://doi.org/10.1038/s41467-019-11383-7>.
46. Phelan, V.V., Du, Y., McLean, J.A., and Bachmann, B.O. (2009). Adenylation enzyme characterization using  $\gamma$ -18O-ATP pyrophosphate exchange. *Chem. Biol.* **16**, 473–478. <https://doi.org/10.1016/j.chembiol.2009.04.007>.
47. Wang, Y., He, J., Alam, M.S., Wang, F., Shang, Z., Chen, Y., Sun, C., Lu, Z., Gao, Y., Zhang, T., et al. (2024). Efficient mutasynthesis of “non-natural” antitubercular ilamycins with low cytotoxicity. *ACS Synth. Biol.* **13**, 930–941. <https://doi.org/10.1021/acssynbio.3c00730>.
48. Kaniusaite, M., Tailhades, J., Marschall, E.A., Goode, R.J.A., Schittenhelm, R.B., and Cryle, M.J. (2019). A proof-reading mechanism for non-proteinogenic amino acid incorporation into glycopeptide antibiotics. *Chem. Sci.* **10**, 9466–9482. <https://doi.org/10.1039/c9sc03678d>.
49. Sun, Y., He, X., Liang, J., Zhou, X., and Deng, Z. (2009). Analysis of functions in plasmid pHZ1358 influencing its genetic and structural stability in *Streptomyces lividans* 1326. *Appl. Microbiol. Biotechnol.* **82**, 303–310. <https://doi.org/10.1007/s00253-008-1793-7>.
50. Tong, Y., Whitford, C.M., Blin, K., Jørgensen, T.S., Weber, T., and Lee, S.Y. (2020). CRISPR-Cas9, CRISPRi and CRISPR-BEST-mediated genetic manipulation in streptomycetes. *Nat. Protoc.* **15**, 2470–2502. <https://doi.org/10.1038/s41596-020-0339-z>.
51. Eslami, S.M., and van der Donk, W.A. (2024). Proteases Involved in Leader Peptide Removal during RiPP Biosynthesis. *ACS Bio Med Chem Au.* **4**, 20–36. <https://doi.org/10.1021/acsbiochemchem.3c00059>.
52. Trowitzsch, S., Viola, C., Scheer, E., Conic, S., Chavant, V., Fournier, M., Papai, G., Ebong, I.-O., Schaffitzel, C., Zou, J., et al. (2015). Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. *Nat. Commun.* **6**, 6011. <https://doi.org/10.1038/ncomms7011>.
53. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
54. Ting, C.P., Funk, M.A., Halaby, S.L., Zhang, Z., Gonen, T., and van der Donk, W.A. (2019). Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science* **365**, 280–284. <https://doi.org/10.1126/science.aau6232>.
55. Daniels, P.N., Lee, H., Splain, R.A., Ting, C.P., Zhu, L., Zhao, X., Moore, B.S., and van der Donk, W.A. (2022). A biosynthetic pathway to aromatic amines that uses glycyl-tRNA as nitrogen donor. *Nat. Chem.* **14**, 71–77. <https://doi.org/10.1038/s41557-021-00802-2>.
56. Yu, Y., and van der Donk, W.A. (2024). PEARL-catalyzed peptide bond formation after chain reversal during the biosynthesis of non-ribosomal peptides. *ACS Cent. Sci.* **10**, 1242–1250. <https://doi.org/10.1021/acscentsci.4c00044>.



57. Hedges, J.B., and Ryan, K.S. (2020). Biosynthetic pathways to nonproteinogenic  $\alpha$ -amino acids. *Chem. Rev.* **120**, 3161–3209. <https://doi.org/10.1021/acs.chemrev.9b00408>.
58. Ehinger, F.J., and Hertweck, C. (2024). Biosynthesis and recruitment of reactive amino acids in nonribosomal peptide assembly lines. *Curr. Opin. Chem. Biol.* **81**, 102494. <https://doi.org/10.1016/j.cbpa.2024.102494>.
59. Precord, T.W., Ramesh, S., Dommaraju, S.R., Harris, L.A., Kille, B.L., and Mitchell, D.A. (2023). Catalytic site proximity profiling for functional unification of sequence-diverse radical S-adenosylmethionine enzymes. *ACS Bio Med Chem Au.* **3**, 240–251. <https://doi.org/10.1021/acsbiomedchemau.2c00085>.
60. Gilchrist, C.L.M., Booth, T.J., van Wersch, B., van Grieken, L., Medema, M.H., and Chooi, Y.-H. (2021). cblaster: a remote search tool for rapid identification and visualization of homologous gene clusters. *Bioinform. Adv.* **1**, vbab016. <https://doi.org/10.1093/bioadv/vbab016>.
61. Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.J.N., et al. (2023). antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* **51**, W46–W50. <https://doi.org/10.1093/nar/gkad344>.
62. van den Belt, M., Gilchrist, C., Booth, T.J., Chooi, Y.-H., Medema, M.H., and Alanjary, M. (2023). CAGECAT: The CompArative GENE Cluster Analysis Toolbox for rapid search and visualisation of homologous gene clusters. *BMC Bioinform.* **24**, 181. <https://doi.org/10.1186/s12859-023-05311-2>.
63. Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene.* **96**, 23–28. [https://doi.org/10.1016/0378-1119\(90\)90336-p](https://doi.org/10.1016/0378-1119(90)90336-p).
64. Tong, Y., Whitford, C.M., Robertsen, H.L., Blin, K., Jorgensen, T.S., Klitgaard, A.K., Gren, T., Jiang, X., Weber, T., and Lee, S.Y. (2019). Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. *Proc. Natl. Acad. Sci. USA.* **116**, 20366–20375. <https://doi.org/10.1073/pnas.1913493116>.
65. Blin, K., Pedersen, L.E., Weber, T., and Lee, S.Y. (2016). CRISPy-web: An online resource to design sgRNAs for CRISPR applications. *Synth. Syst. Biotechnol.* **1**, 118–121. <https://doi.org/10.1016/j.synbio.2016.01.003>.

**Chem, Volume 11**

**Supplemental information**

**Ribosomal pentapeptide nitration for non-ribosomal  
peptide antibiotic precursor biosynthesis**

**Leo Padvá, Lukas Zimmer, Jemma Gullick, Yongwei Zhao, Vishnu Mini Sasi, Ralf B. Schittenhelm, Colin J. Jackson, Max J. Cryle, and Max Crüsemann**

## Supplemental Tables

**Table S1.** Predicted functions of ORFs in the *ruf* BGC in *S. atratus* S3\_m208\_1 (naming based on the nomenclature by Tomita et al. 2017).<sup>1</sup>

ORF	Amino acids	Proposed function
<i>rufA</i>	353	Streptomycin biosynthesis operon regulator
<i>rufB</i>	267	Serine aminopeptidase
<i>rufC</i>	367	Cytochrome P450
<i>rufD</i>	absent	
<i>rufE</i>	4819	Type I polyketide synthase
<i>rufF</i>	252	Thioesterase
<i>rufG</i>	absent	
<i>rufH</i>	80	MbtH-like protein
<i>rufI</i>	380	Aminotransferase
<i>rufJ</i>	161	Hypothetical protein
<i>rufK</i>	359	ABC transporter
<i>rufL</i>	277	ABC transporter
<i>rufM</i>	402	Cytochrome P450
<i>bytA</i>	5	RiPP precursor peptide
<i>rufN</i>	396	Nitric oxide synthase
<i>rufO</i>	394	Cytochrome P450
<i>rufP</i>	373	Tryptophan N-dimethylallyltransferase
<i>rufQ</i>	167	Hypothetical protein
<i>rufR</i>	296	Luciferase-like monooxygenase
<i>rufS</i>	399	Cytochrome P450
<i>rufT</i>	8012	Nonribosomal peptide synthetase

**Table S2.** Selected examples of Cytochrome P450s displaying alternate acid/ alcohol I-helix residues together with their functions. Unusual residues are emboldened in red.

P450	Function	Acid	Alcohol	Notes	Ref.
P450 <sup>Blt</sup> (CYP1251C3)	MRYLH crosslinking	Glu	Ser	Tyr-His C-N bond	2
RufO (CYP1251)	MRYLH Nitration	Glu	<b>Val</b>	Tyr nitration	3
P450cam (CYP101A1)	C-H 5-exo-hydroxylation of (1s)-camphor	Asp	Thr	Archetypal P450 hydroxylase	4
P450BM3 (CYP102A1)	C-H hydroxylation of fatty acids (mostly $\omega$ -1 to $\omega$ -3)	Glu	Thr	Fused to redox partner domain	5
P450SP $\alpha$ (CYP152B1)	$\alpha$ C-H hydroxylation of fatty acids	<b>Val</b>	<b>Ala</b>	Peroxide driven reaction	6
OxyB (CYP165B7)	C-O-D crosslinking in teicoplanin biosynthesis	Asp	Asn	X-domain recruitment, PCP-bound substrate	7
EryF (CYP107A1)	Macrolide hydroxylation of erythromycin	Glu	<b>Ala</b>	Substrate-mediated activation	8
Cyp17A1	Hydroxylation and C-C bond cleavage of steroids	Glu	Thr	Peroxyanion intermediate for cleavage; cytochrome B5 interaction	9
AspB	DKP crosslinking during aspergilazine A biosynthesis;	<b>Pro</b>	Thr	Ferric superoxide intermediate	10
CypX (CYP134A1)	DKP aromatisation and N-oxidation	Glu	<b>Pro</b>	Multiple step oxidation	11
XplA (CYP177A1)	Reductive denitration of RDX	<b>Met</b>	<b>Ala</b>	Reductive transformation	12
TxtE	Nitration of Trp in DKP biosynthesis of thaxtomin A	<b>Pro</b>	Thr	Uses NO + postulated ferric superoxide intermediate	13
Sas16 (CYP)	Desaturation of Tyr in acyl-Thr-Tyr dipeptide in WS9326A biosynthesis	Glu	<b>Phe</b>	PCP-bound substrate	14
CYP79A1	Oxime formation from Tyr	Asp	Asn	Multiple step oxidation	15
Cyp74A1	Allene oxide synthase	<b>Gly</b>	<b>Gly</b>	Rearrangement of fatty acid peroxide	16

**Table S3.** Product conversion for cytochrome P450 enzymes with Nle-1.

Enzyme	Cyclised (%)	Nitrated Tyrosine (%)
RufO <sub>WT</sub>	0	-
RufO <sub>WT</sub> + NO	0	78.1 ± 8.5
RufO <sub>SV</sub>	19.2 ± 4.1	-
RufO <sub>SV</sub> + NO	14.9 ± 3.7	3.3 ± 1.3
P450 <sub>Blt</sub> WT	85 (17)	-
P450 <sub>Blt</sub> WT + NO	85	0
P450 <sub>Blt</sub> VP	0	-
P450 <sub>Blt</sub> VP + NO	0	0

**Table S4.** K<sub>M</sub>, k<sub>cat</sub> and k<sub>cat</sub>/K<sub>M</sub> for RufO<sub>WT</sub> with Nle-1.

K <sub>M</sub> (μM)	k <sub>cat</sub> (min <sup>-1</sup> )	K <sub>cat</sub> /K <sub>M</sub> (μM <sup>-1</sup> min <sup>-1</sup> )
9.6 ± 2.0	0.25 ± 0.02	0.03

**Table S5.** Product conversion for RufO<sub>WT</sub> with Nle-1 under varying conditions.

Conditions*	Nitrated Tyrosine (%)
No redox partners + sodium dithionite	39.6
No redox partners + sodium dithionite + SOD	6.5
No redox partners + sodium dithionite + catalase*	0
No redox partners + hydrogen peroxide	18.6

\* All experiments performed aerobically

**Table S6.** NADH consumption assay data for cytochrome P450 enzymes.

Enzyme	Substrate	Average K <sub>cat</sub> (min <sup>-1</sup> )
RufO <sub>SV</sub>	Nle-1	0.76 ± 0.18
RufO <sub>SV</sub> + NO	Nle-1	1.47 ± 0.06
RufO <sub>WT</sub>	Nle-1	0.86 ± 0.06
RufO <sub>WT</sub> + NO	Nle-1	1.64 ± 0.25
P450 <sub>Blt</sub>	1	3.49 ± 0.17

**Table S7.** Product conversion for RufO<sub>WT</sub> for peptides 2-6 in the presence of DEANO.

Peptide	Cyclised (%)	Nitrated Tyrosine (%)
2	11.4	0
2 (RufO <sub>SV</sub> )	7.6	0
3	0	1.2
4	0	2.9
5	0	0
6	6.5	13.5

**Table S8.** Two different types of BGCs harbour NRPS/PKS genes and homologous genes of the nitric oxide synthase *rufN*, cytochrome P450 *rufO* and the precursor peptide *bytA*. Identity percentages of homologous RufNO/IlaMN sequences in comparison to the reference sequence of *Streptomyces atratus* S3\_m208\_1. Accession numbers for BGC and P450 enzymes (RufO) are provided alongside the identified precursor peptide sequence for each organism, which is exclusively MRYLH. Some of the clusters are on contig edges, therefore no Stachelhaus code is available.

Strain	BGC accession no.	BGC type	Precursor	P450 accession	RufN identity	RufO identity	A domain	Stachelhaus
<i>Streptomyces atratus</i> S3_m208_1	PQ034618.1	NRPS, T1PKS	MRYLH	XCX13213.1	100	100	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces atratus</i> ATCC 14046	LC257593.1	NRPS, T1PKS	MRYLH	BBA20962.1	96.8	98.0	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces atratus</i> SCSIO_ZH16	KY173348.1	NRPS, T1PKS	MRYLH	ASX95236.1	97.7	98.0	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces atratus</i> NPDC058404	NZ_JBHXKC010000117.1	NRPS, T1PKS	MRYLH	WP_382228932.1	94.9	98.0	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces atratus</i> MJM3502	MN088369.1	NRPS, T1PKS	MRYLH	WPW27215.1	98.2	97.7	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces atratus</i> NPDC057618	NZ_JBHVGW010000156.1	NRPS, T1PKS	MRYLH	WP_322108603.1	95.1	97.7	-	-
<i>Streptomyces</i> sp. A108	NZ_JACBYP010000131.1	NRPS, T1PKS	MRYLH	MBU6535743.1	83.1	92.4	-	-
<i>Streptomyces</i> sp. NPDC006320	NZ_JBEXRI010000031.1	NRPS, T1PKS	MRYLH	WP_356755524.1	91.2	92.4	-	-
<i>Streptomyces</i> sp. NPDC006174	NZ_JBEXTJ010000021.1	NRPS, T1PKS	MRYLH	WP_356755524.1	91.2	92.4	-	-
<i>Streptomyces achromogenes</i> B2110	JAUSYH010000001.1	NRPS, T1PKS	MRYLH	MDQ0834929.1	87.6	92.1	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces achromogenes</i> W4119-2	JAUSYA010000001.1	NRPS, T1PKS	MRYLH	MDQ0687727.1	87.6	92.1	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces</i> sp. B4113	JAUSZG010000003.1	NRPS, T1PKS	MRYLH	MDQ0957474.1	87.1	92.1	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces</i> sp. NPDC089173	JBFBHR010000007.1	NRPS, T1PKS	MRYLH	MEW1629036.1	91.2	92.1	-	-
<i>Streptomyces</i> sp. NPDC058011	NZ_JBHVQZ010000050.1	NRPS, T1PKS	MRYLH	WP_366423644.1	91.2	92.1	-	-
<i>Streptomyces</i> sp. NPDC057245	NZ_JBHUXO010000038.1	NRPS, T1PKS	MRYLH	WP_383477013.1	83.1	92.1	-	-
<i>Streptomyces</i> sp. NPDC006261	JBEXSI010000087.1	NRPS, T1PKS	MRYLH	MEU0163393.1	90.2	91.9	-	-
<i>Streptomyces</i> sp. NPDC012616	JBIBDV010000039.1	NRPS, T1PKS	MRYLH	MFF6807741.1	84.5	91.9	-	-
<i>Streptomyces</i> sp. NPDC005898	NZ_JBEXWC010000001.1	NRPS, T1PKS	MRYLH	WP_358909266.1	85.7	91.1	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces</i> sp. NPDC023327	JBFIAB010000017.1	NRPS, T1PKS	MRYLH	MEU4798741.1	83.4	90.9	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces</i> sp. NPDC005900	JBEXWA010000149.1	NRPS, T1PKS	MRYLH	MEU1133601.1	83.2	90.6	-	-
<i>Kitasatospora</i> sp. NPDC097643	NZ_JBEOSA010000066.1	NRPS, T1PKS	MRYLH	WP_350602965.1	78.0	83.8	-	-
<i>Streptomyces luteosporus</i> JCM 4542	BAAASL010000005.1	NRPS, T1PKS	MRYLH	GAA2712158.1	68.3	74.3	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces</i> sp. NPDC029704	NZ_JBEZWC010000033.1	NRPS, T1PKS	MRYLH	WP_344434043.1	66.5	73.9	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces</i> sp. NPDC049555	JBFAPL010000049.1	NRPS, T1PKS	MRYLH	MEV4743485.1	65.5	73.9	3NO2-Tyr (100%)	DASTAAAVGK
<i>Streptomyces</i> sp. NPDC049585	JBFAPE010000063.1	NRPS, T1PKS	MRYLH	MEV4437330.1	65.7	73.9	-	-
<i>Actinoplanes missouriensis</i> NPDC048962	JBEZFM010000058.1	NRPS, T1PKS	MRYLH	MEU8244545.1	63.1	72.4	-	-
<i>Nonomuraea solani</i> CGMCC 4.7037	FNVT01000022.1	NRPS, T1PKS	MRYLH	SEH01838.1	65.9	67.2	?	DAADVAEVWK DAALLGLVWK
<i>Nonomuraea cypriaca</i> K274	JADOGI010000057.1	NRPS, T1PKS	MRYLH	MBF8188026.1	64.6	66.0	-	-
<i>Nonomuraea basaltis</i> 160415	VCJS010000381.1	NRPS, T1PKS	MRYLH	TMR88472.1	65.4	65.9	-	-

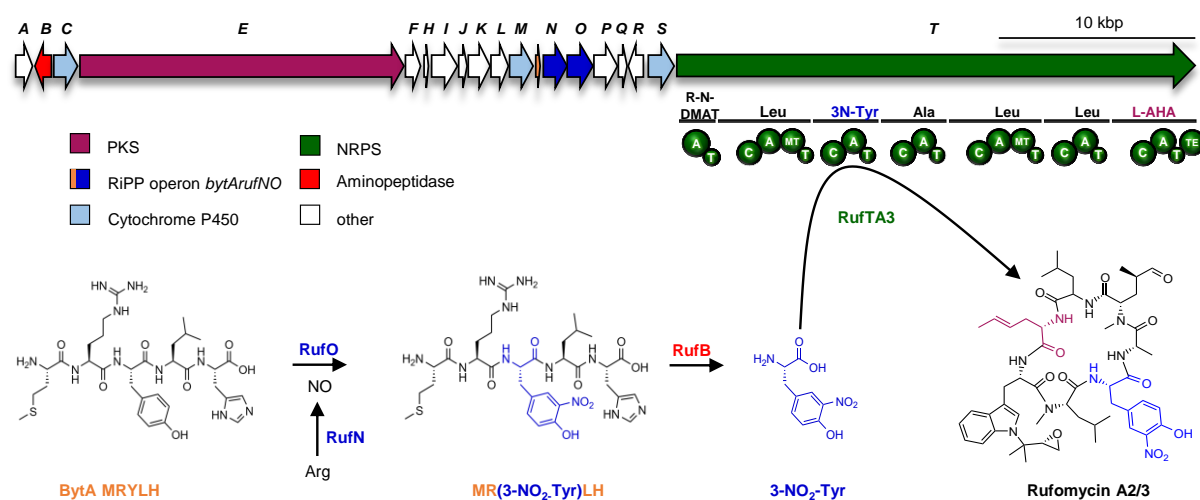
<i>Amycolatopsis magusensis</i> strain 40447	JASCSI010000093.1	NRPS, T1PKS	MRYLH	MDI5978862.1	-	62.1	Tyr (70%)	DAATSGDVWK
<i>Fodinicola feengrottensis</i> JCM 14718	BAAANY010000038.1	NRPS, T1PKS	MRYLH	GAA1714241.1	-	56.7	Tyr (70%)	DAATAGDVWK
<i>Nonomuraea</i> sp. NPDC049152	JBEZDO010000013.1	NRPS, T1PKS	MRYLH	MEU7895959.1	51.5	53.1	Tyr (70%)	DAATSGDVWK
<i>Micromonospora</i> sp. NPDC007271	JBEYER010000041.1	NRPS, T1PKS	MRYLH	MEU2613213.1	-	53.0	Tyr (70%)	DAATSGDVWK
<i>Nonomuraea africana</i> JCM 6240	BAAASY010000040.1	NRPS, T1PKS	MRYLH	GAA2416812.1	51.0	52.8	Tyr (70%)	DAATSGDVWK
<i>Nonomuraea africana</i> NPDC050248	JBFAKZ010000022.1	NRPS, T1PKS	MRYLH	MEV2271487.1	50.1	52.7	Tyr (70%)	DAATSGDVWK
<i>Polymorphospora</i> sp. NPDC050346	JBFAKM010000001.1	NRPS, T1PKS	MRYLH	MEV0727369.1	52.5	52.6	Tyr (70%)	DAATSGDVWK
<i>Polymorphospora</i> sp. NPDC051019	JBFBCH010000017.1	NRPS, T1PKS	MRYLH	MEV7230529.1	-	52.6	Tyr (70%)	DAATSGDVWK
<i>Polymorphospora</i> sp. 2-325	JBCGDC010000059.1	NRPS, T1PKS	MRYLH	MFB6395443.1	-	52.6	Tyr (70%)	DAATSGDVWK
<i>Polymorphospora rubra</i> JCM 14904	BAAAOJ010000053.1	NRPS, T1PKS	MRYLH	GAA2056963.1	-	52.3	Tyr (70%)	DAATSGDVWK
<i>Polymorphospora rubra</i> NPDC050307	JBFAKQ010000036.1	NRPS, T1PKS	MRYLH	MEV0610869.1	-	52.3	Tyr (70%)	DAATSGDVWK
<i>Polymorphospora rubra</i> NPDC050590	JBFAIZ010000001.1	NRPS, T1PKS	MRYLH	MEV0391707.1	-	52.3	Tyr (70%)	DAATSGDVWK
<i>Nonomuraea</i> sp. NPDC003804	JBEXMT010000025.1	NRPS, T1PKS	MRYLH	MET9342177.1	51.4	52.0	Tyr (70%)	DAATSGDVWK
<i>Micromonospora echinofusca</i> MPMI6	NZ_WVUH01000043.1	NRPS, T1PKS	MRYLH	WP_208812380.1	-	50.9	Tyr (70%)	DAATSGDVWK
<i>Micromonospora</i> sp. CPCC 206060	NZ_JBAKHU010000023.1	NRPS, T1PKS	MRYLH	WP_334436711.1	-	50.6	Tyr (70%)	DAATSGDVWK
<i>Actinoplanes</i> sp. NPDC023801	JBFAAI010000003.1	NRPS, T1PKS	MRYLH	MEU4620085.1	49.2	48.7	Tyr (70%)	DAATSGDVWK

**Table S9.** Primers used in this study.

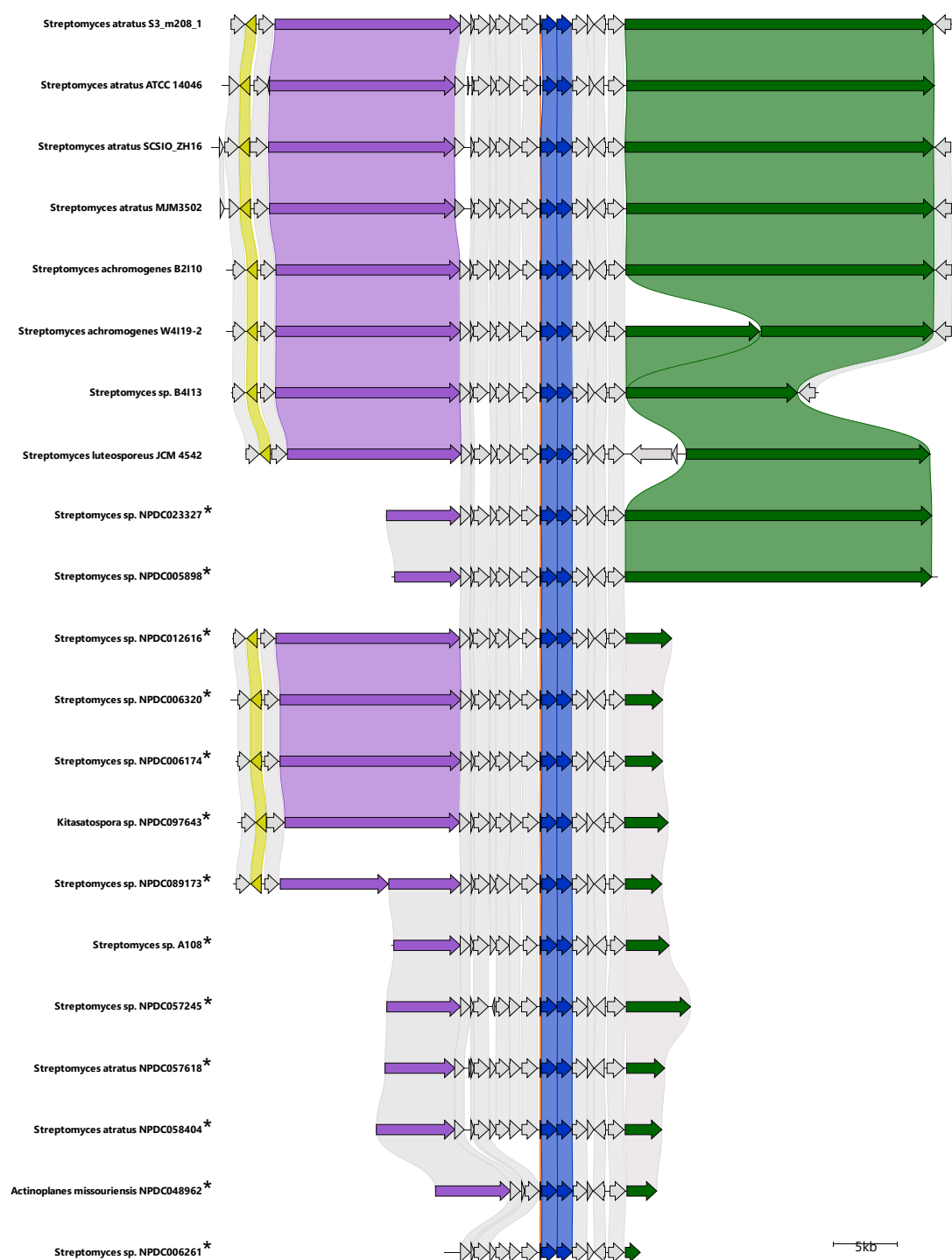
Primer	Sequence
rufO_V240S_fwd	CGAGTCCGTGTCCCGATTCTCGGC
rufO_V240S_rev	CTCGAAGCCTGCATGAAAAACGCTC
rufO_V240SP241V_fwd	CGAGTCCGTGTCCCGATTCTCGGC
rufO_V240SP241V_rev	GGACACGGACTCGAAGCCTGCATG
P450Blt_S239V_fwd	ACGAGGTGGTGTCCCGTCTGCTGGG
P450Blt_S239V_rev	GACACGTGCTCGTAACCAGCGTGCA
P450Blt_S239VV240P_fwd	CGAGGTGCCGTCCCGTCTGCTGGGC
P450Blt_S239VV240P_rev	CGGGACGGCACCTCGTAACCAGCGTGCA
pYH7_ΔbytA_up_for	GCTTGCGGCAGCGTGAAGCTGATCTTCCGGAAGTCGGTG
pYH7_ΔbytA_up_rev	TGAAAGTAGAAACGAGACC
pYH7_ΔbytA_down_for	TGGTCTCGTTTCTACTTTCACATATCGATTACCTCCCCTG
pYH7_ΔbytA_down_rev	GACCTGCAGGCATGCAAGCTCACCGCATCGGAAATCGTC
pYH7_ΔbytA_screen_for	CGTTGTAGGCCAGCCGCAG
pYH7_ΔbytA_screen_rev	TCGAGGAGCTGATGCGCTG
rufO_sgRNA1_for	CATGCCATGGGGCGTGCCAGTAGACCGGCT GTTTTAGAGCTAGAAATAGC
rufO_sgRNA2_for	CATGCCATGGCGGCGTGCCAGTAGACCGGC GTTTTAGAGCTAGAAATAGC
rufO_sgRNA3_for	CATGCCATGGCATCAGCCAGGAGTCGAGGC GTTTTAGAGCTAGAAATAGC
sgRNA_rev	ACGCCTACGTAAAAAAGCACCGACTCGGTGCC
RufB_MBP_TEV_for	CTGAGAACCTGTACTTCCAATCCATGGATTTCCGACACGTC
RufB_MBP_TEV_rev	TGCTGTCCACCAGTCATGCTAGCTTACGGAGCATCCAGAAATGC
RufB_codonopti_for	TGGTGCCGCGCGGCAGCCATATGGATTTCCGACACGTCAGG
RufB_codonopti_rev	GAGTGCGGCCGCAAGCTTGGTTACGGAGCATCCAGAAATGCAC G
pet28a_linearized_for	CCAAGCTTGCGGCCGC
pet28a_RufB_linearized_rev	ATGGCTGCCGCGCGG
rufB_screen_for	CCATGAACGCTCGGAACG
rufB_screen_rev	TCAGGGGGCGTCGAGG
rufTA3PCP3_for	TAGAAGCTTACCCCTCCGACAGGTGCCG
rufTA3PCP3_for	TATCATATGCTGCTTGCCCCGCAGGAGCG
rufN/O_screen_for	GTGAGTTCTTTCGCAGG
ΔbytA_screen_for	TTTCTACTTTCACATATCG
rufM_screen_for	CGTACCGCTCAGCCTGC
rufM_screen_rev	CCAGCTCGGTGTCGGTG



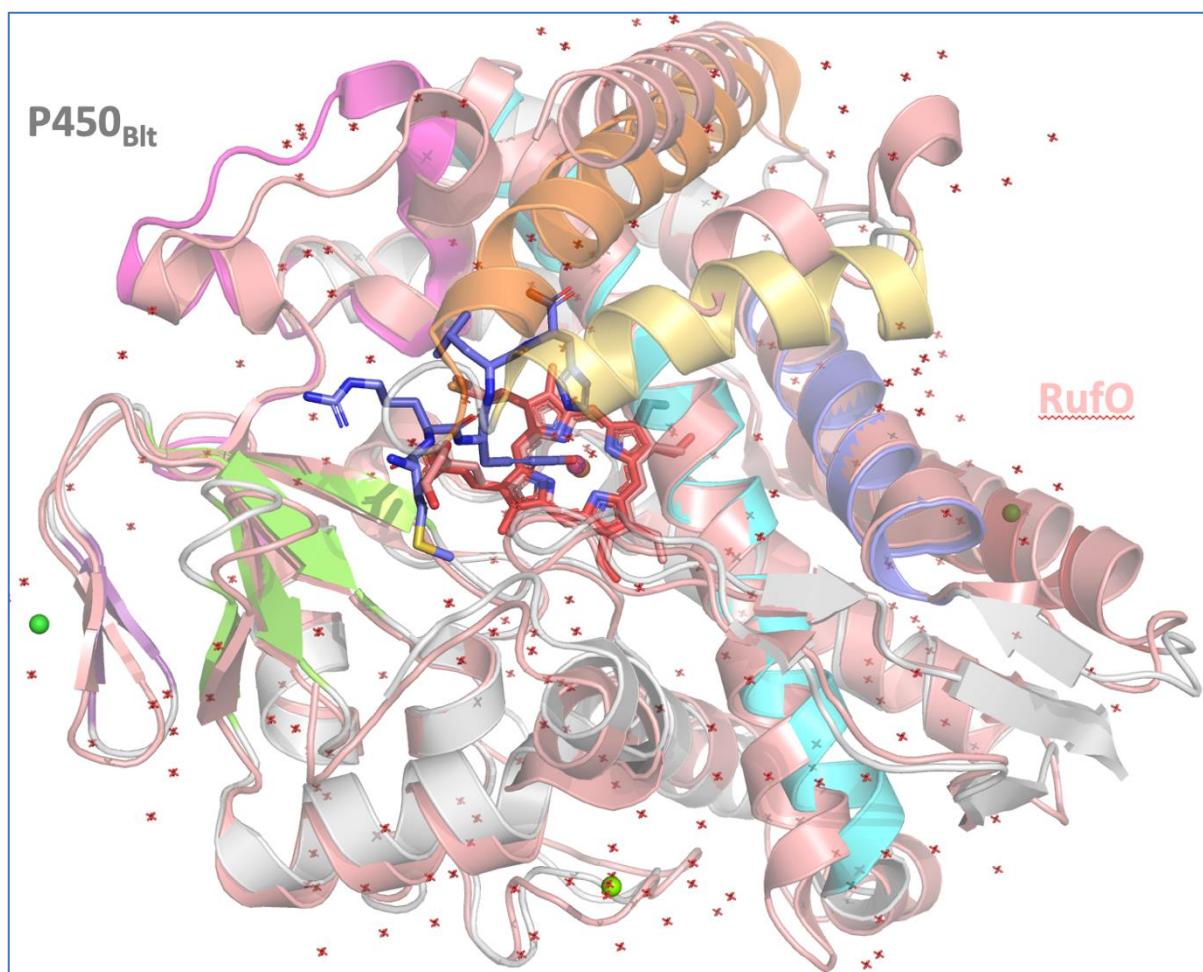
## Supplemental Figures



**Figure S1. The rufomycin BGC harbors a unique pathway for 3-NO<sub>2</sub>-Tyr biosynthesis.** Schematic overview on the *ruf* BGC from *S. atratus* S3\_m208\_1 and 3-NO<sub>2</sub>-Tyr biosynthesis and incorporation.



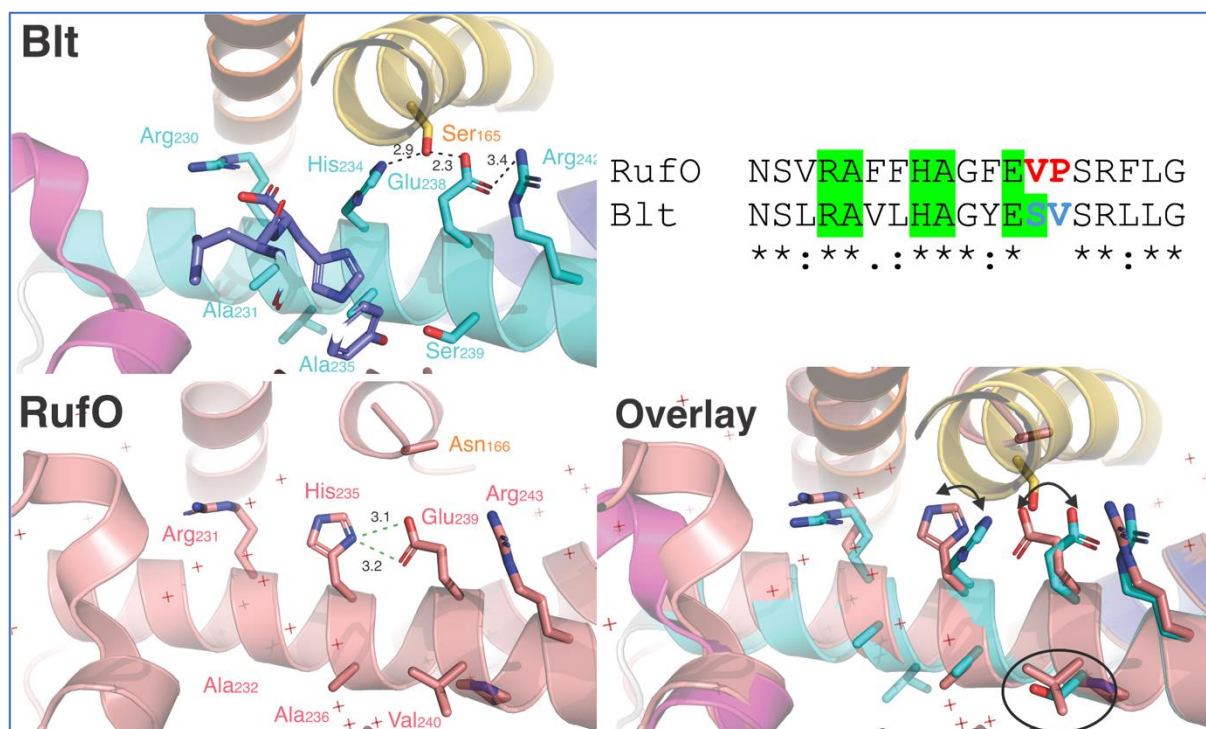
**Figure S2. The biarylittide minimal precursor gene *bytA* is present in all detected homologous rufomycin/ilamycin gene clusters.** All BGCs encode the precursor peptide MRYLH directly upstream of *rufNO*. Homologous genes in each gene cluster are linked. Asterisks are added to regions on contig edges. Figure to scale.



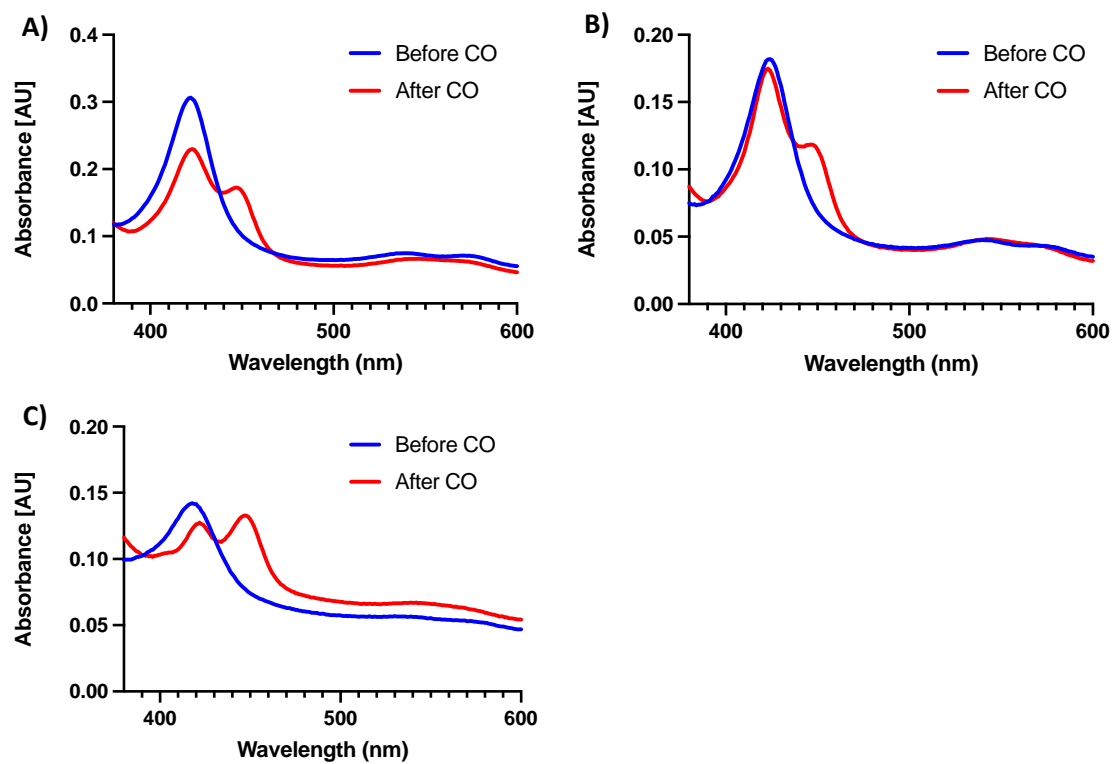
**Figure S3.** Structure of RufO (PDB: 8SPC, chain A)<sup>17</sup> overlaid on the structure of P450<sub>Blt</sub> (PDB: 8U2M, chain A).<sup>2</sup> RufO – salmon; P450<sub>Blt</sub> – grey, heme – red sticks, MRYLH peptide substrate – blue sticks. Specific P450<sub>Blt</sub> secondary structure elements coloured (B/C loop – pink, D-helix – brick red, E-helix – slate blue, F-helix – yellow, G-helix – orange, I-helix – turquoise,  $\beta_1$ -sheet – green,  $\beta_2$ -sheet – purple).

P450 <sub>Blt</sub>	MADEATAAVRDPIYDPLAPSVIADPYPFYRKLRETNTVHWHEFLDSWVVTGYAECRQVLG	60
RufO	MTVSPAPEHTDPLFSPLDPAVLADPYPVYRRLRETHPVYWHAGLDSWLMTRHADCTAILR	60
	*: . :        **::.** *:*:*****.**:*****: *:**        *****:* *:*:        :*	
P450 <sub>Blt</sub>	DTTNFGSDFRRIDVEIPDTQLSVQSLDPPEHGAIHLLVSALHEQPLSTVRQQFAAIAAQ	120
RufO	DPGRFSTDFRKIDIPTPTLLSLQTLDPDQTPRLRHLALDAVRAQDLALRKELTLFADQ	120
	*    *.*:*****:    * * *:*:*****:    :*** :.**:    * *.**:*****:    :* *	
P450 <sub>Blt</sub>	HLAELSGQPGTVDLVSRLFARFVALRTITAFGLGVPPD-GAGFEQWSNAIVRSMDAGIEPA	179
RufO	LLDELAD-RESFDFIHDYADVFTLRAITRFIGVEPPETDEAFARFND <sup>8</sup> DLDHSMDAQLD <sup>8</sup> PD	179
	* **:    :.**:    :*    .:***** *:** **:    . *.    :.**:    : :***** :*:	
P450 <sub>Blt</sub>	RAEPGNQARAELSRVLVTHWLAEADERGFVGAARRAARAQD-VPAAVLANSIRAVLHAGYE	238
RufO	AEEPGLRARAHFNDLVRSWLGDPGPHGVLPDVVRLLPGSGVEADDVLVNSVRAFFHAGFE	239
	*** :***:.    **    **:    . :*:    . *    ...        **.**:*.**:***:*	
P450 <sub>Blt</sub>	<sup>SV</sup> SRLLGGVLRARLVRHPELLAGPATRD-ADEALVDELIRLDGPVQADARVCVRDQPVGAQ	297
RufO	<sup>VP</sup> SRFLGNALAALLATPGAWQLVRGDVGLDTAVEELIRYVGPVQALARACLQDTELGGM	299
	**:**.** *:    *        .    *    . :*:    *:*****        ***** **:*:    :*.	
P450 <sub>Blt</sub>	LVRRGDVLVLFIAAANRDPVFPDPDAVRLTRRRGLHLAFGRGAHACLGLATLQLREV	357
RufO	AVKEGQVVTALIGAANRDPDQFPDPETLRLDRKPNHNLGFGRGHACLGLNVARIEAHVT	359
	*:.*:*.    :*.*****        *****:*** *:    . **.*****:*** :.* :. : .	
P450 <sub>Blt</sub>	LGALRAGGLRLAPAGPAAYEPTATLRGLAELPVSVRQPHRTD	399
RufO	LGALLRH-PGVRSAEPVVRPNGTLRGLSRLPLTLG-----	394
	*****        :    **    . *.*****:*****:	

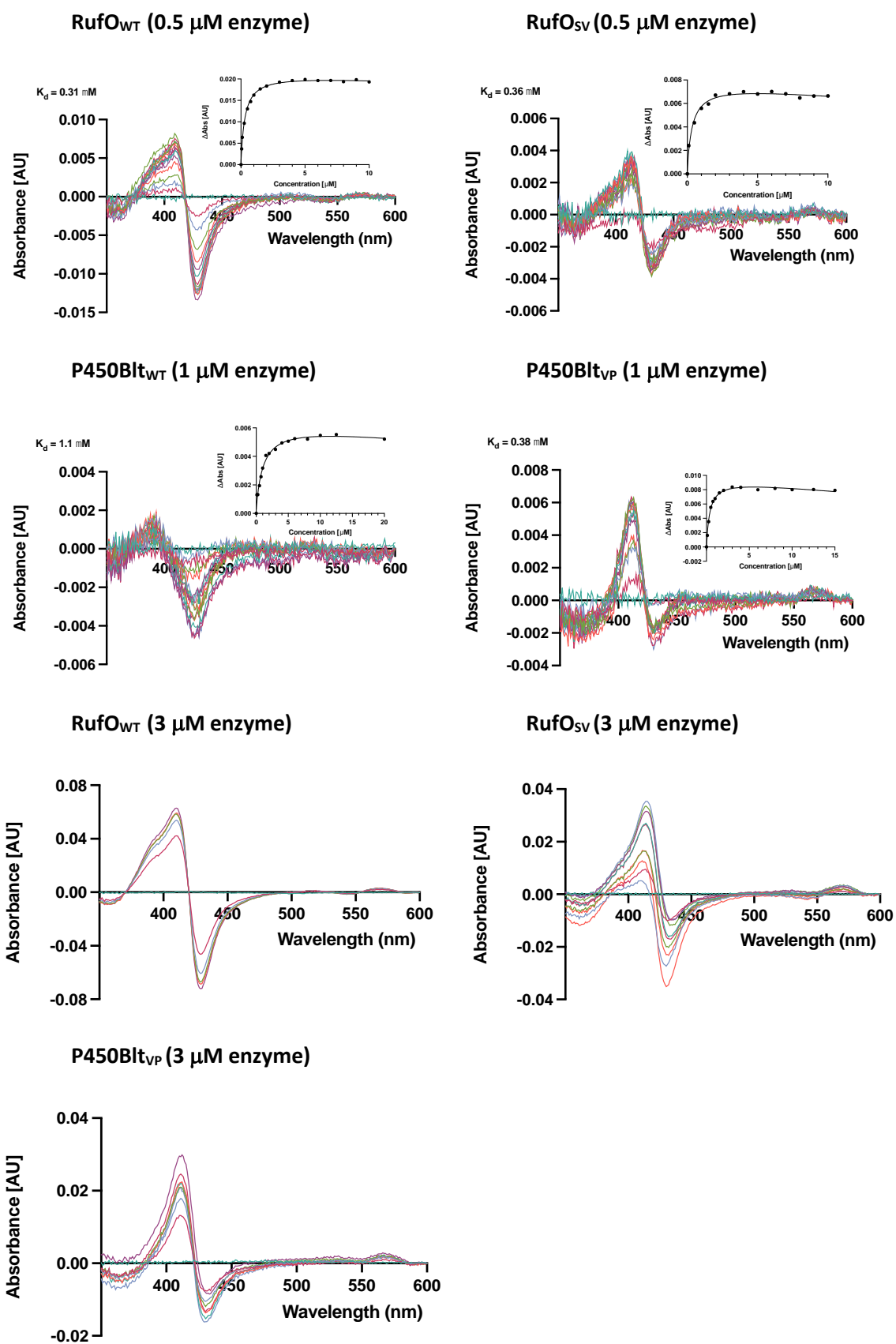
**Figure S4.** Alignment of RufO<sup>1</sup> and P450<sub>Blt</sub><sup>2</sup> with key mutations in the I-helix coloured red and highlighted in yellow. Additional residue possibly contributing alterations in RufO I-helix rotamers (PDB: 8SPC, chain A) is highlighted in yellow, although incomplete dentistry does not allow this to be unambiguously assigned. Lack of density in the F/G-helices of the RufO structure are highlighted in grey.



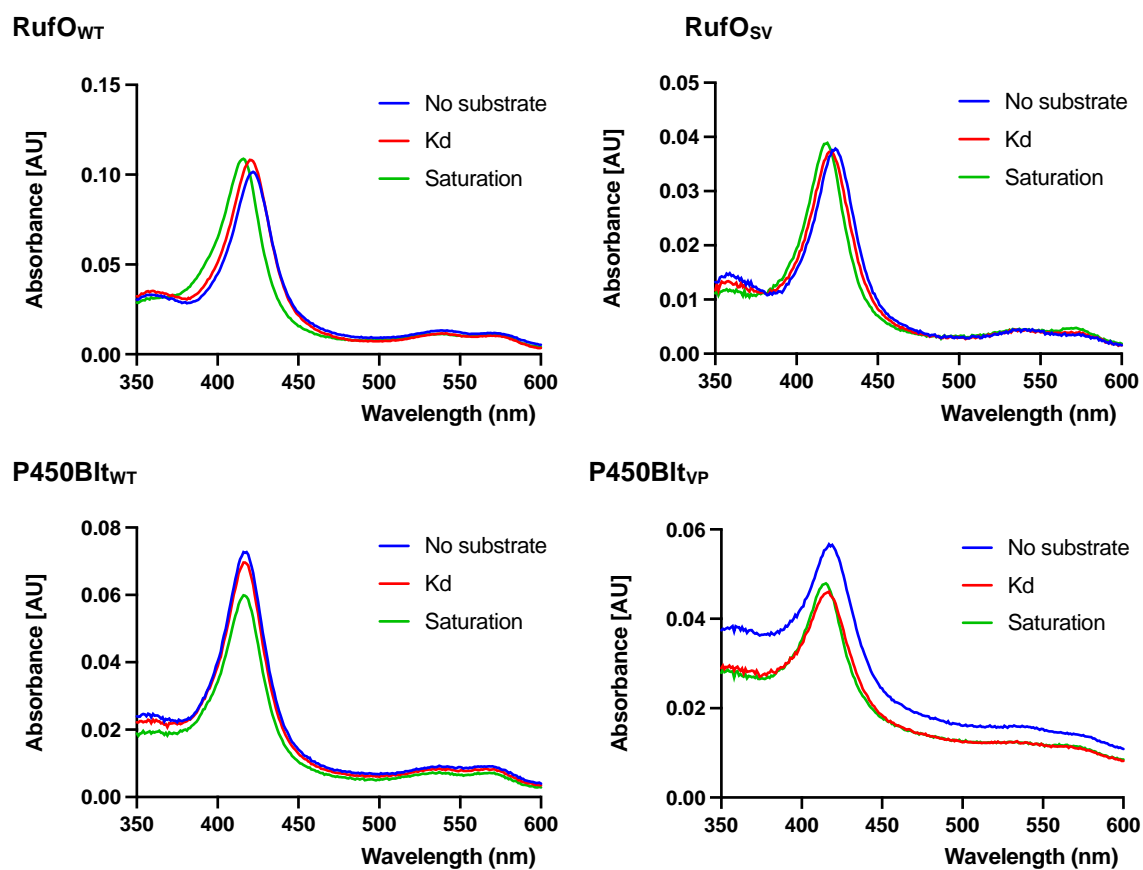
**Figure S5.** Structural comparison of the active site I-helix residues of P450<sub>Blt</sub> (upper left; PDB: 8U2M, chain A)<sup>2</sup> and RufO (lower left; PDB: 8SPC, chain A)<sup>17</sup>; the colour scheme is the same as in the previous figure, residues are labelled and H-bonding distances shown in Angstrom. Comparison of the key I-helix residues via alignment (upper right) and overlay of the P450<sub>Blt</sub> and RufO structures with differences highlighted (lower right).



**Figure S6.** CO difference spectra of the P450s in this study; A) RufO<sub>WT</sub>, B) RufO<sub>SV</sub>, C) P450Blt<sub>WT</sub> as previously published.<sup>18</sup>

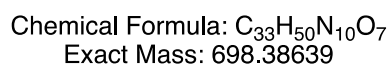
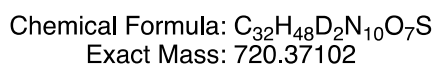
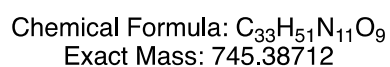
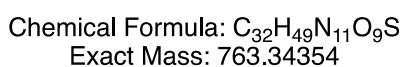
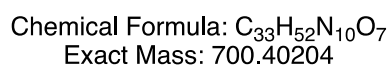
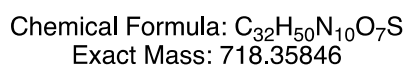


**Figure S7.** Titration of the cytochrome P450s RufO, RufO<sub>SV</sub>, P450Blt<sub>WT</sub> and P450Blt<sub>VP</sub>, with Nle-1.  $K_d$  calculated as described in methods. Curves shown at 2 enzyme concentrations for better depiction of Soret shift.

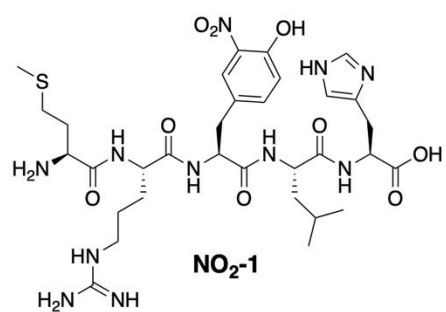


**Figure S8.** Absolute spectra of cytochrome P450s RufO, RufO<sub>SV</sub>, P450Blt<sub>WT</sub> and P450Blt<sub>VP</sub>, with either no substrate, or Nle-1 at K<sub>d</sub> or at saturation.

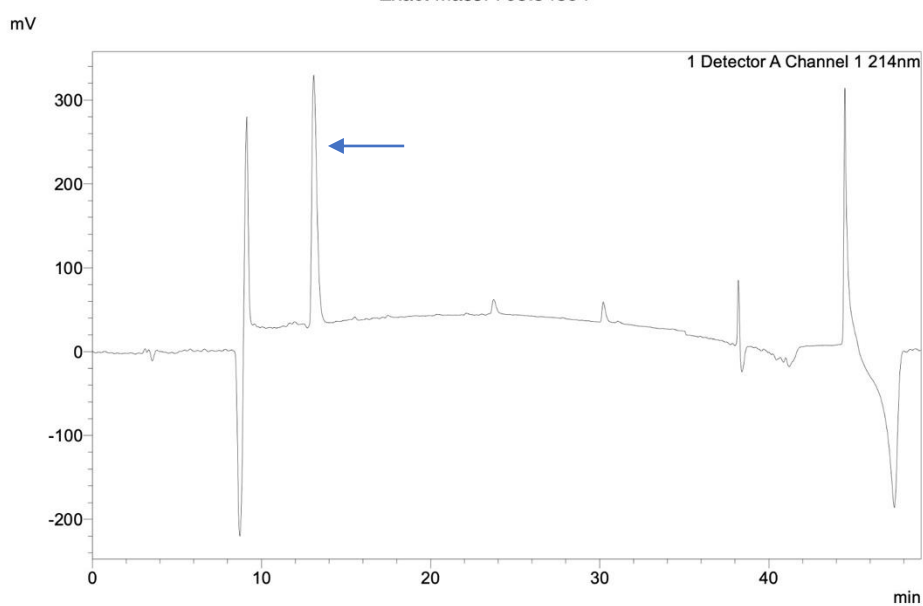




**Figure S9.** Structures of the MRYLH substrate peptides used in this study. Syntheses/characterisation of **1**<sup>18</sup> Nle-**1**<sup>18</sup>, Tyr-(3,5-d<sub>2</sub>)-**1**<sup>19</sup>, and characterisation of enzyme cyclised Nle-**1**<sup>2</sup> as previously reported.

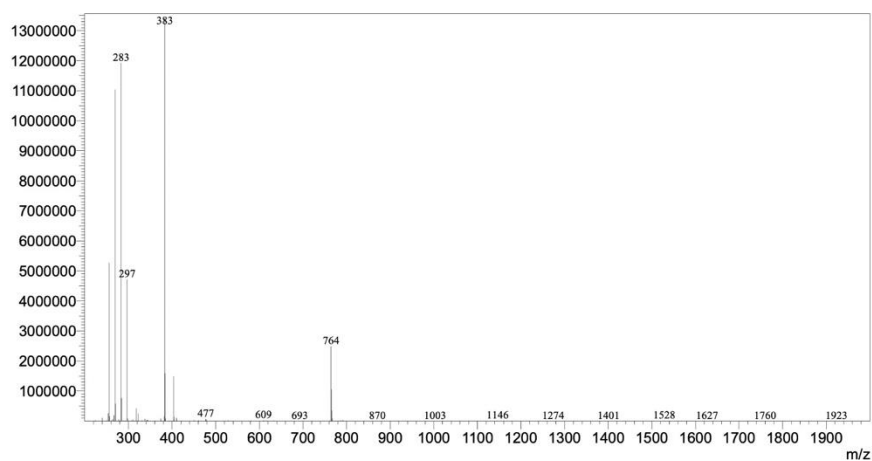


Chemical Formula: C<sub>32</sub>H<sub>49</sub>N<sub>11</sub>O<sub>9</sub>S  
Exact Mass: 763.34354

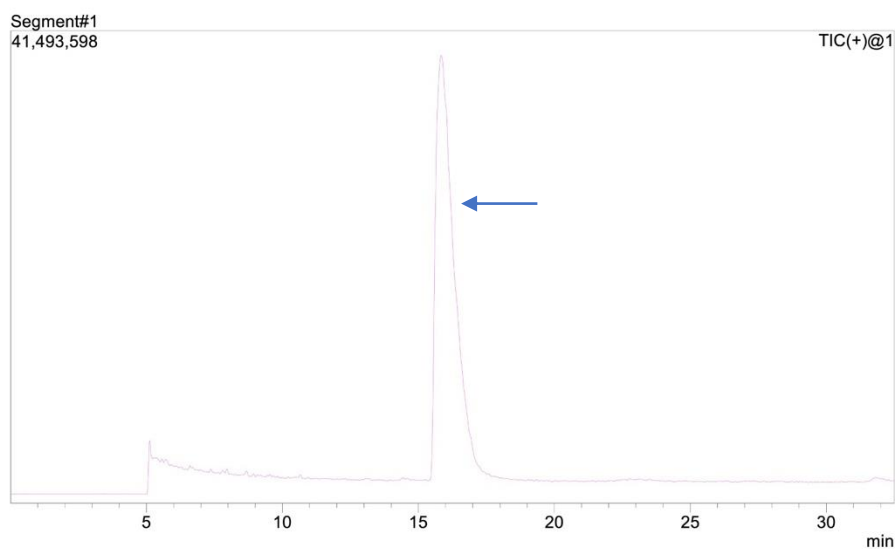
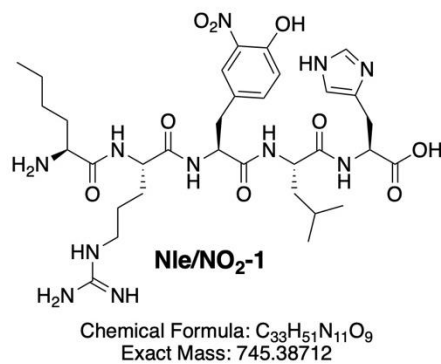


### <Spectrum>

Line#:1 R.Time:----(Scan#:----)  
MassPeaks:1799  
RawMode:Averaged 13.033-13.300(783-799) BasePeak:383(13428102)  
BG Mode:None Segment 1 - Event 1

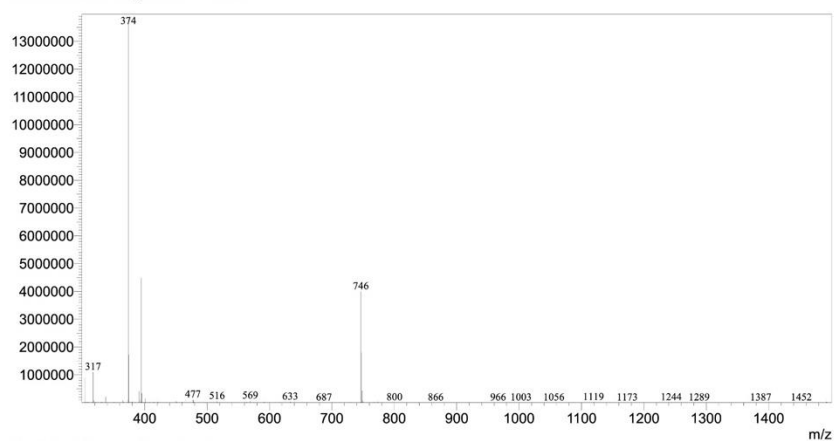


**Figure S10.** LC and MS analysis of synthetic peptide NO<sub>2</sub>-1.

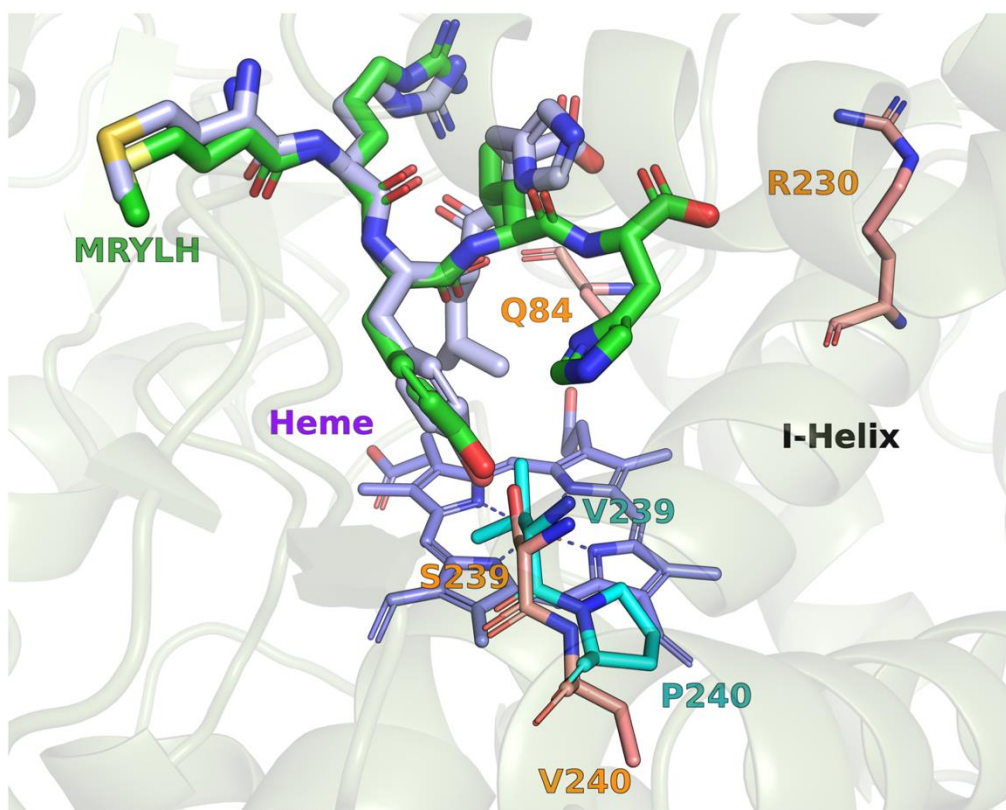


### <Spectrum>

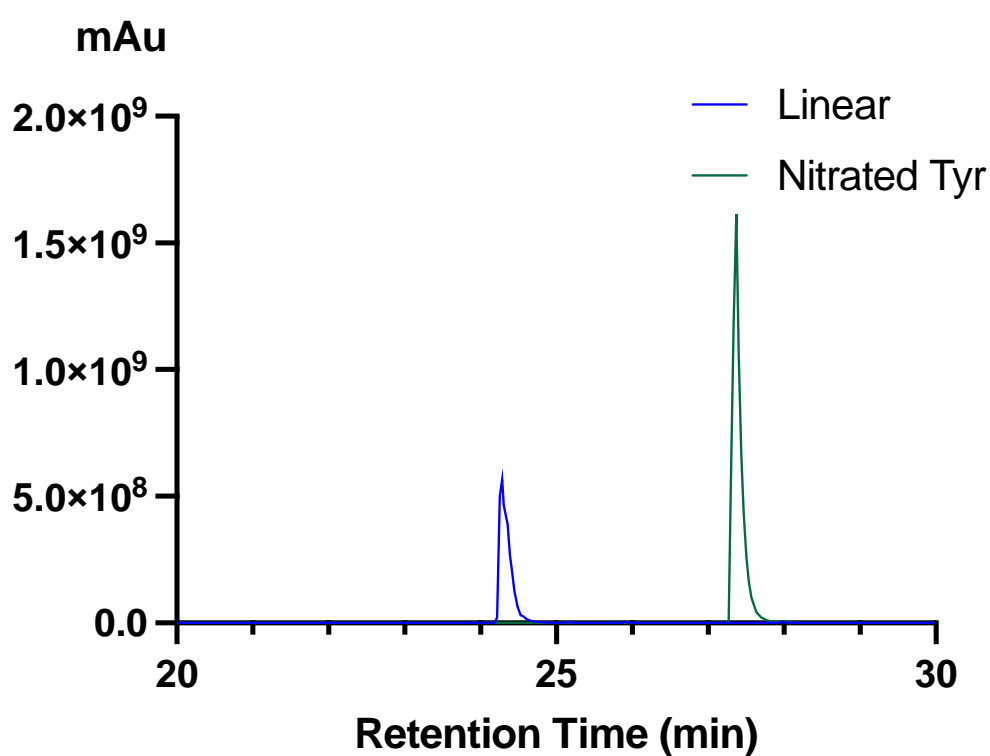
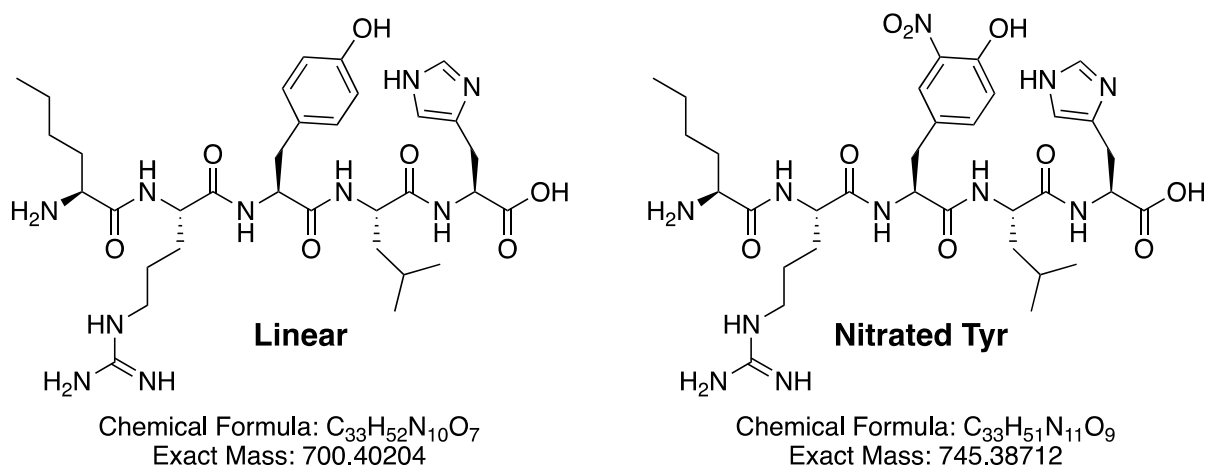
Line#:1 R.Time:----(Scan#:----)  
MassPeaks:1201  
RawMode:Averaged 15.667-16.300(941-979) BasePeak:374(13834656)  
BG Mode:None Segment 1 - Event 1



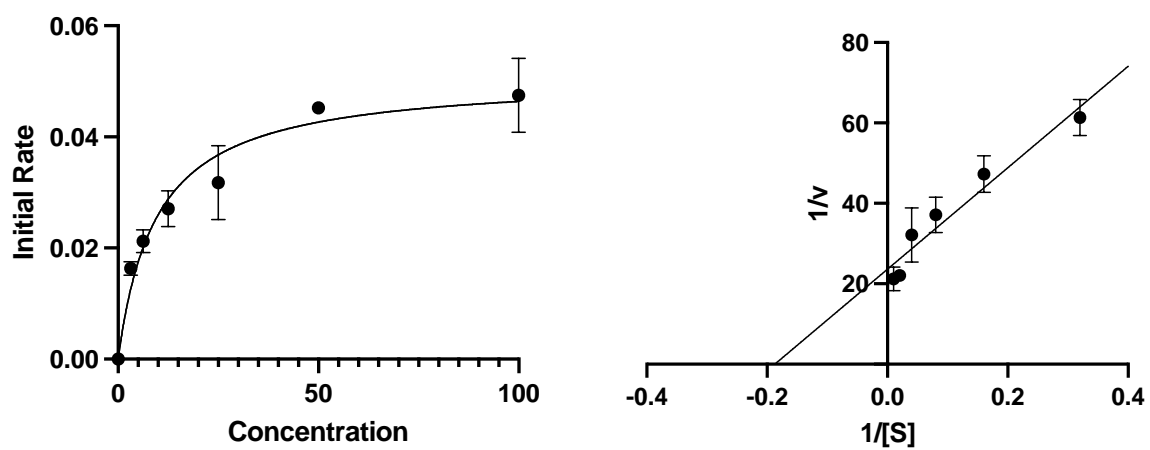
**Figure S11.** TIC and MS analysis of synthetic peptide Nle/NO<sub>2</sub>-1.



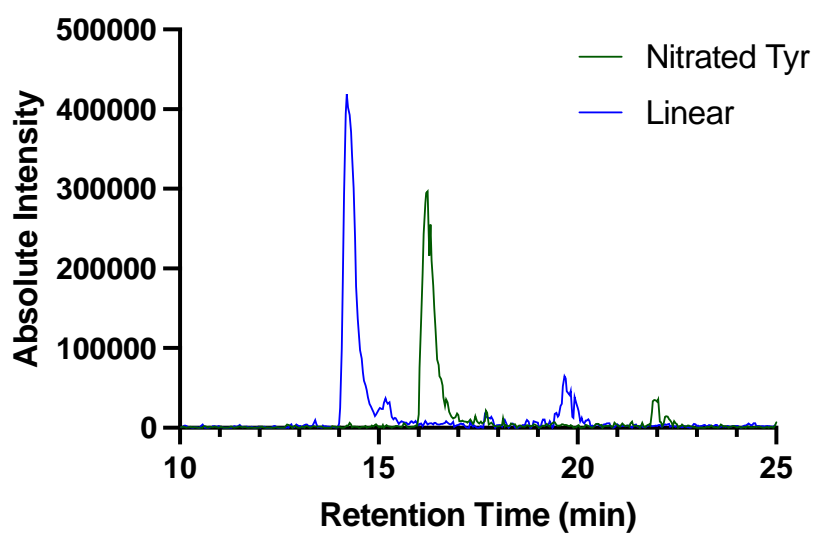
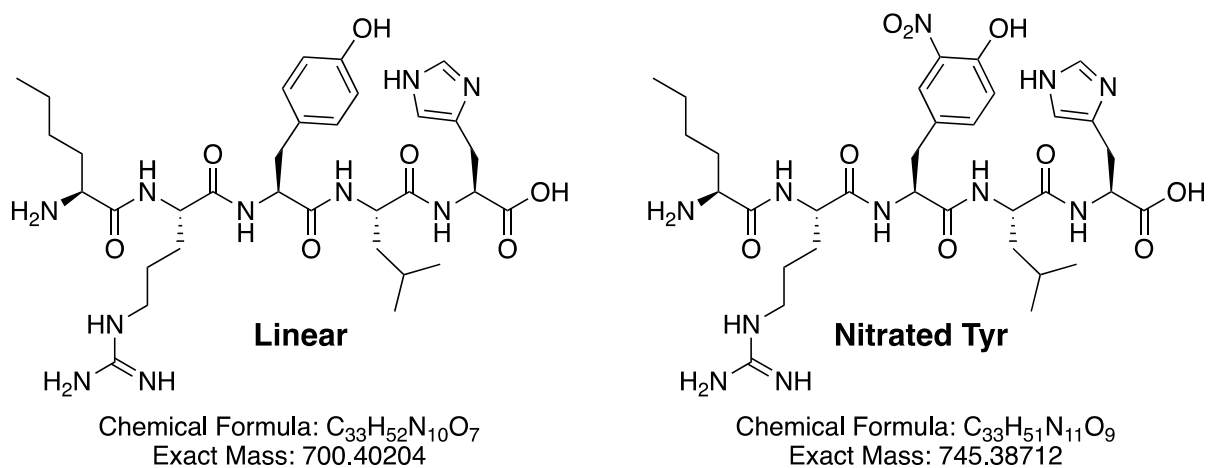
**Figure S12.** Overlapping binding orientations of representative poses from molecular dynamics simulations of MRYLH peptide **1** bound in the active site of P450<sub>BLT</sub> (green) and the Ser239Val/Val240Pro variant (grey) after 300 nanoseconds. Mutated residues Val239 and P240 are shown in cyan, while wild-type residues are shown in orange. Note the major shifts in position of the Leu-4 and His-5 residues of **1**.



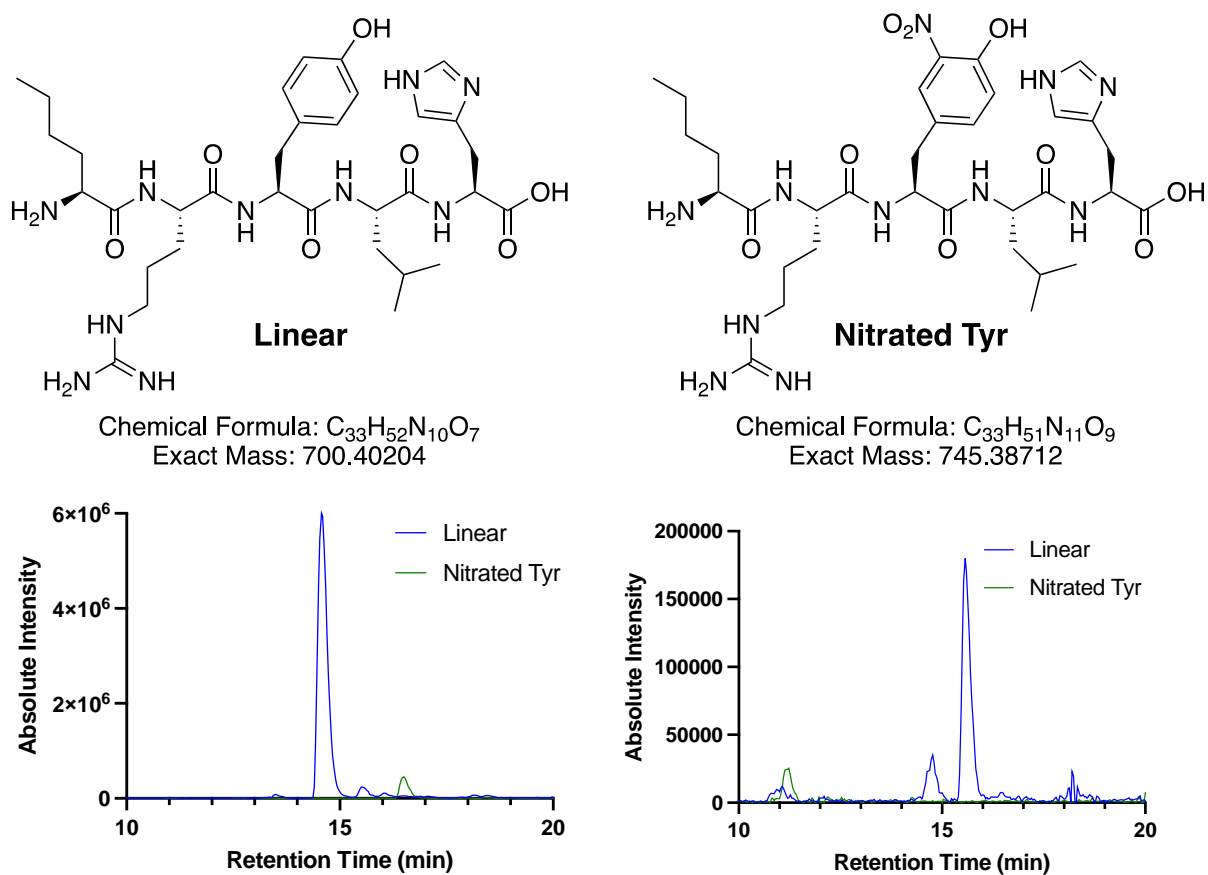
**Figure S13.** Nitration data for Nle-1 using RufO<sub>WT</sub> in the presence of nitric oxide. The structures of linear and nitrated tyrosine-containing peptide are shown. Overlapping ESI traces for linear and nitrated tyrosine-containing peptides based on HRMS analysis.



**Figure S14.** Michaelis-Menten kinetics curve and Lineweaver Burk plot for RufO<sub>WT</sub> for Nle-1.

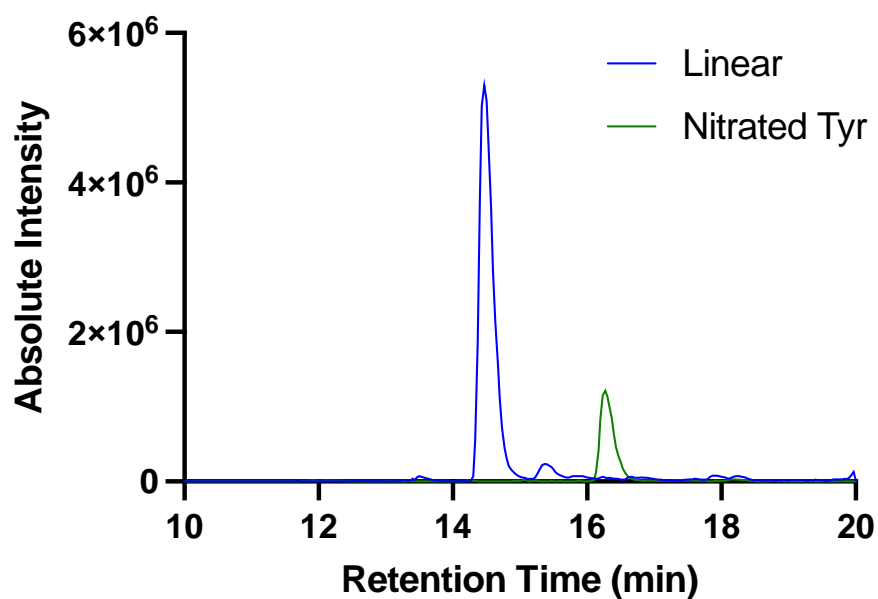
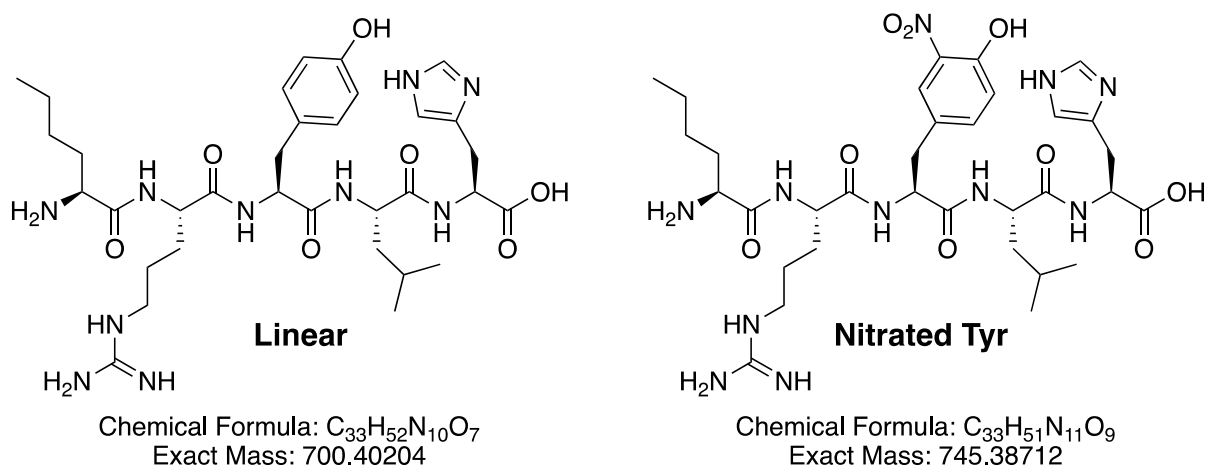


**Figure S15.** Nitration data for Nle-1 using RufO<sub>WT</sub> in the presence of nitric oxide with sodium dithionite and no redox partners. The structures of linear and nitrated tyrosine-containing peptide are shown. Overlapping ESI traces for linear and nitrated tyrosine-containing peptides based on LCMS analysis.

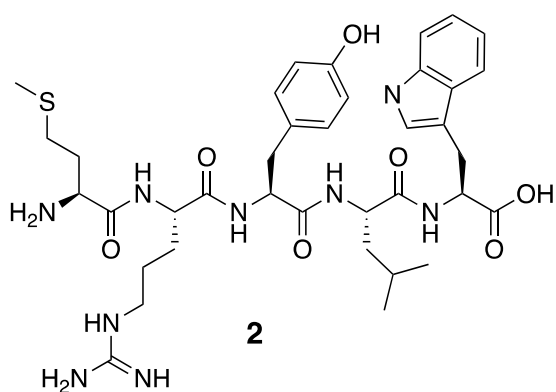


**Figure S16.** Nitration data for Nle-1 using RufO<sub>WT</sub> in the presence of nitric oxide with sodium dithionite and no redox partners with superoxide dismutase (left) and catalase (right). The structures of linear and nitrated tyrosine-containing peptide are shown. Overlapping ESI traces for linear and nitrated tyrosine-containing peptides based on LCMS analysis.

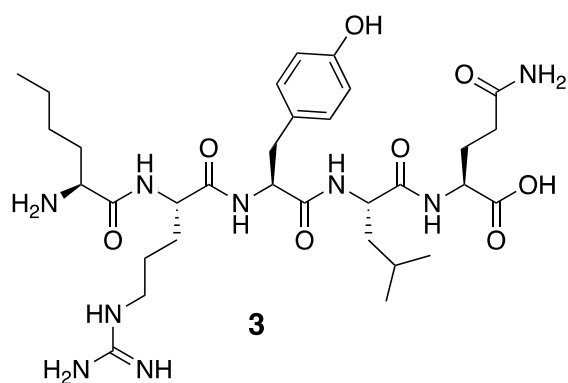




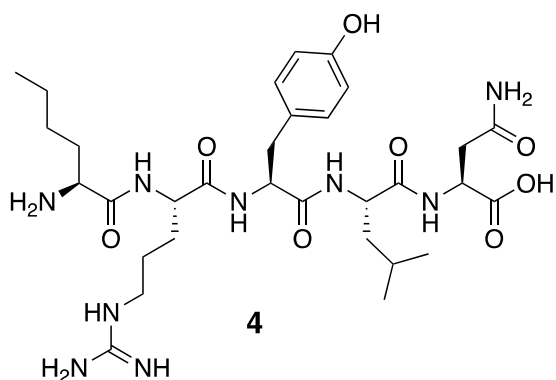
**Figure S17.** Nitration data for Nle-1 using RufO<sub>WT</sub> in the presence of nitric oxide with no redox partners and hydrogen peroxide. The structures of linear and nitrated tyrosine-containing peptide are shown. Overlapping ESI traces for linear and nitrated tyrosine-containing peptides based on LCMS analysis.



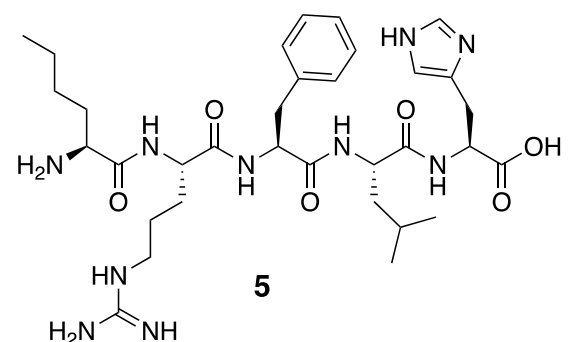
Chemical Formula:  $C_{37}H_{52}N_9O_7S$   
Exact Mass: 766.37104



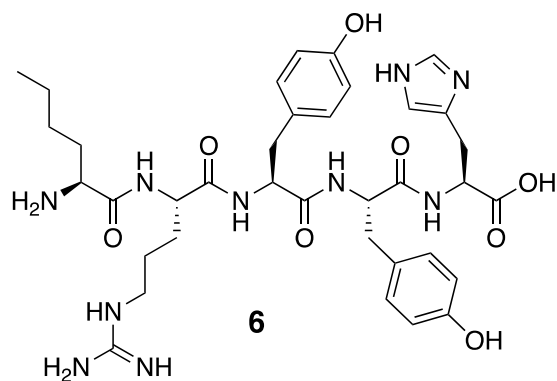
Chemical Formula:  $C_{32}H_{53}N_9O_8$   
Exact Mass: 691.40171



Chemical Formula:  $C_{31}H_{51}N_9O_8$   
Exact Mass: 677.38606

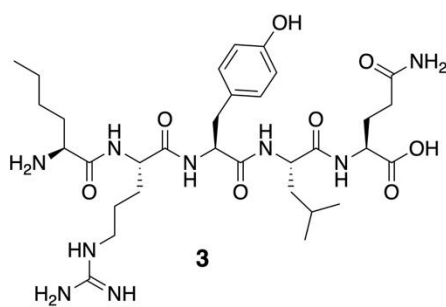


Chemical Formula:  $C_{33}H_{52}N_{10}O_6$   
Exact Mass: 684.40713

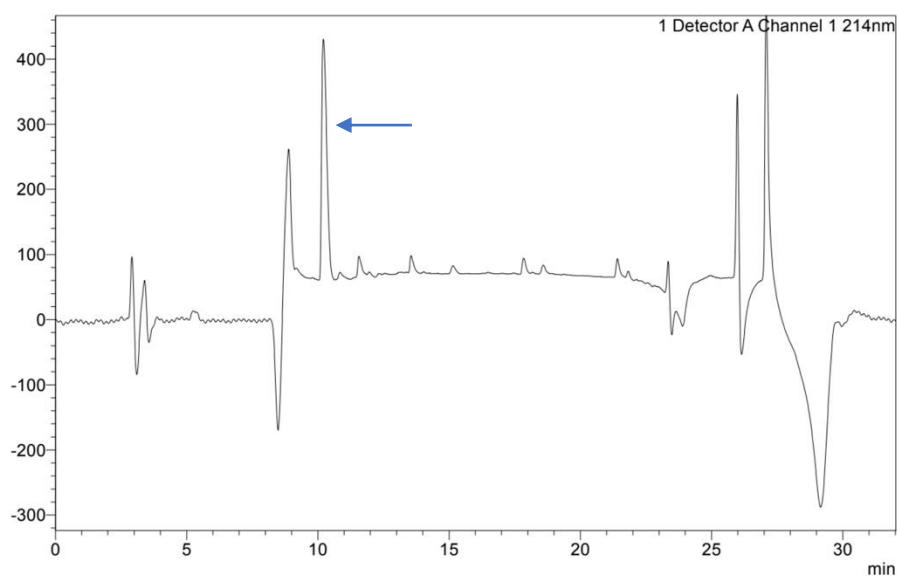


Chemical Formula:  $C_{36}H_{50}N_{10}O_8$   
Exact Mass: 750.38131

**Figure S18.** Structures of the other peptides (**2-6**) used in this study. Synthesis/ characterisation of **2**<sup>18</sup> as previously reported, synthesis of **3-6** can be found below.

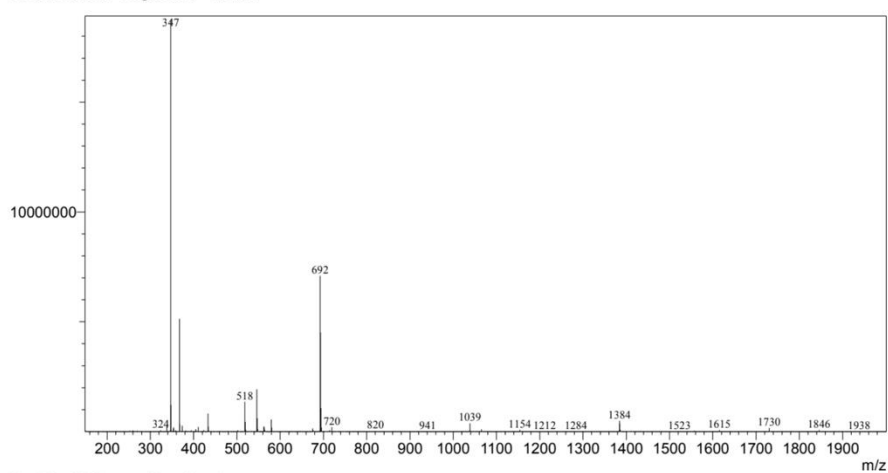


Chemical Formula:  $C_{32}H_{53}N_9O_8$   
Exact Mass: 691.40171

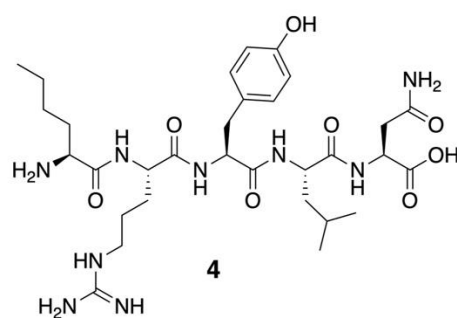


### <Spectrum>

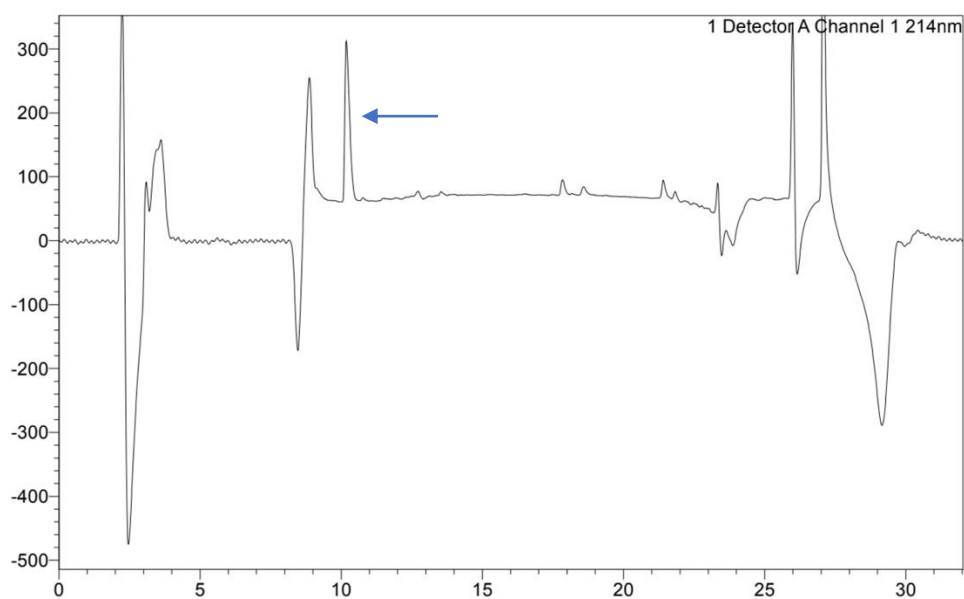
Line#:1 R.Time:----(Scan#:----)  
MassPeaks:1851  
RawMode:Averaged 10.233-10.667(615-641) BasePeak:347(18733525)  
BG Mode:None Segment 1 - Event 1



**Figure S19.** LC and MS analysis of synthetic peptide **3**.

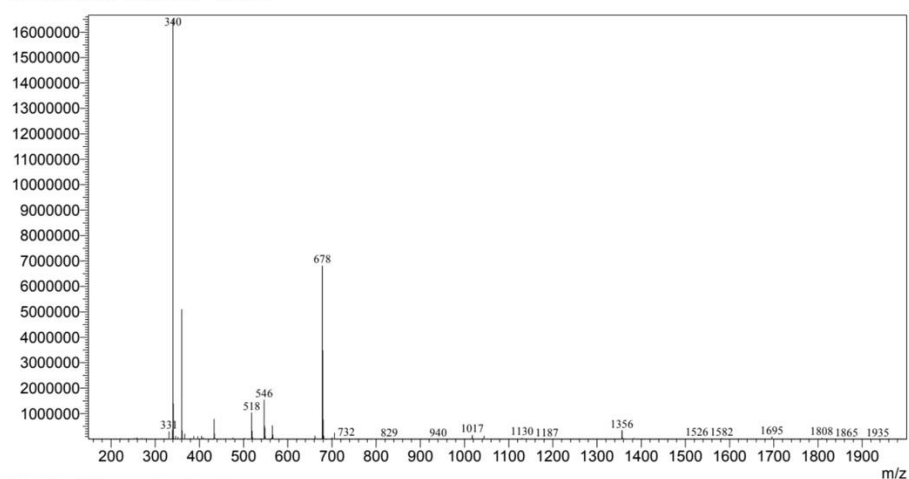


Chemical Formula:  $C_{31}H_{51}N_9O_8$   
Exact Mass: 677.38606

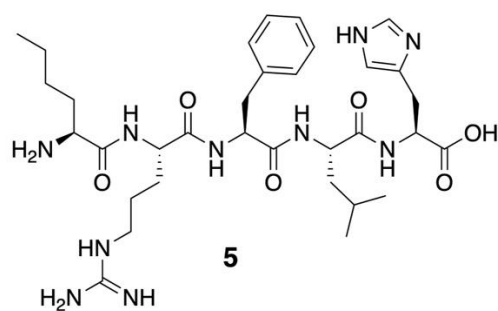


### <Spectrum>

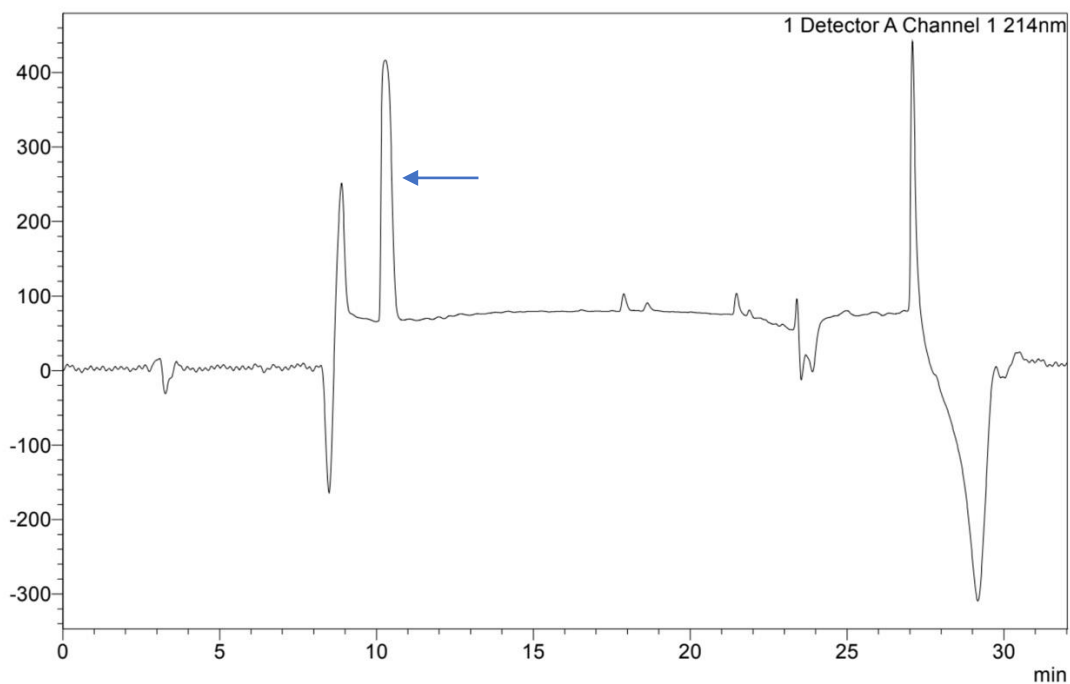
Line#:1 R.Time:----(Scan#:----)  
MassPeaks:1851  
RawMode:Averaged 10.200-10.500(613-631) BasePeak:340(16501847)  
BG Mode:None Segment 1 - Event 1



**Figure S20.** LC and MS analysis of synthetic peptide **4**.

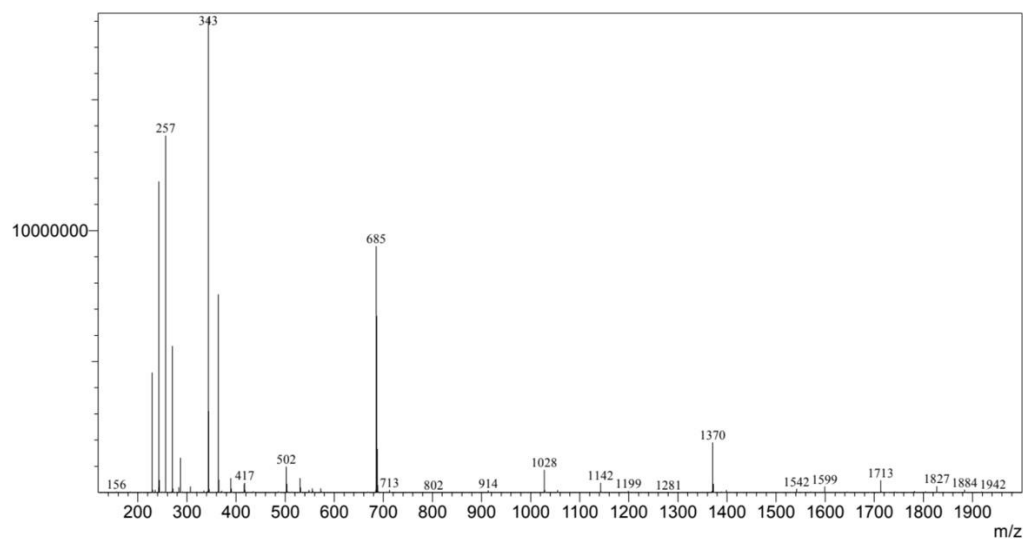


Chemical Formula:  $C_{33}H_{52}N_{10}O_6$   
Exact Mass: 684.40713

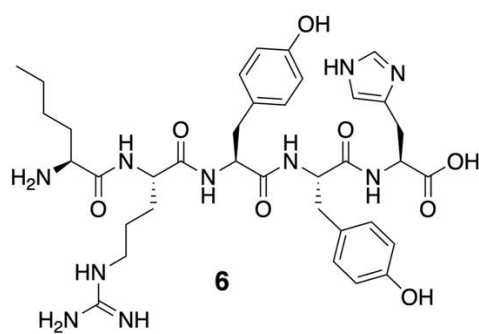


### <Spectrum>

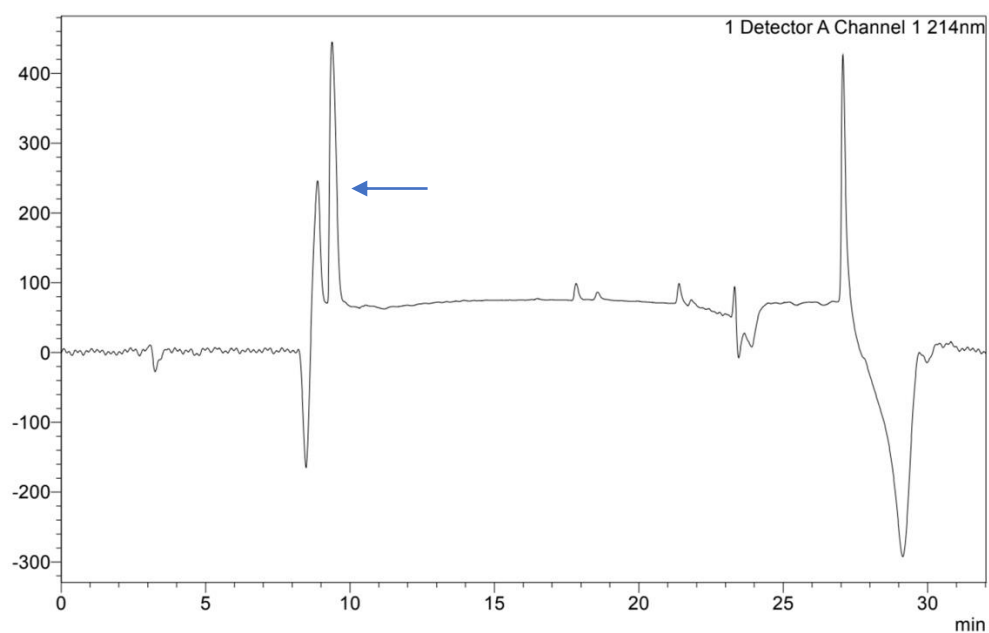
Line#:1 R.Time:----(Scan#----)  
MassPeaks:1879  
RawMode:Averaged 10.333-10.533(621-633) BasePeak:343(18123416)  
BG Mode:None Segment 1 - Event 1



**Figure S21.** LC and MS analysis of synthetic peptide **5**.

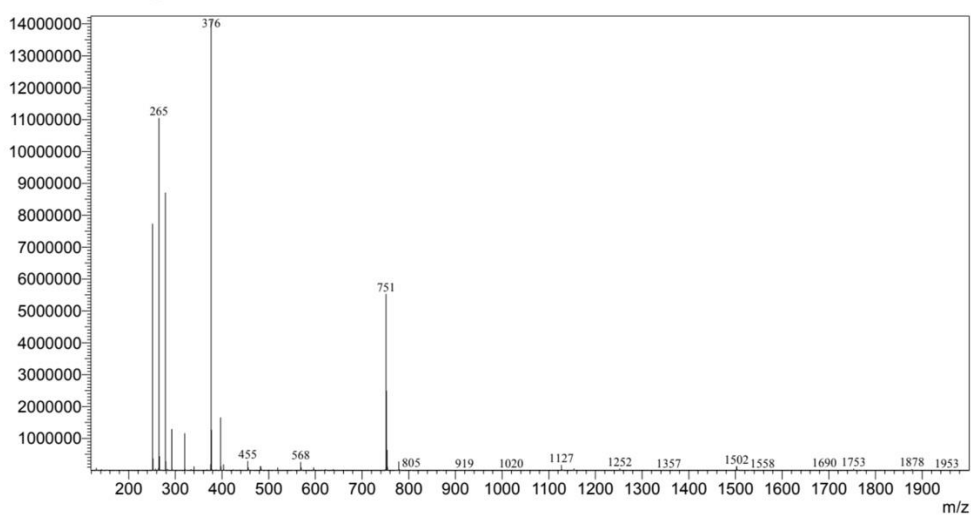


Chemical Formula:  $C_{36}H_{50}N_{10}O_8$   
Exact Mass: 750.38131

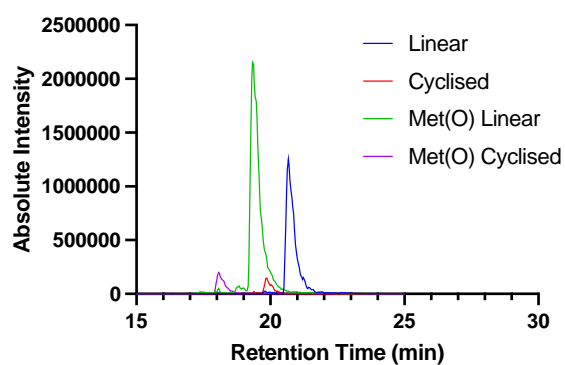
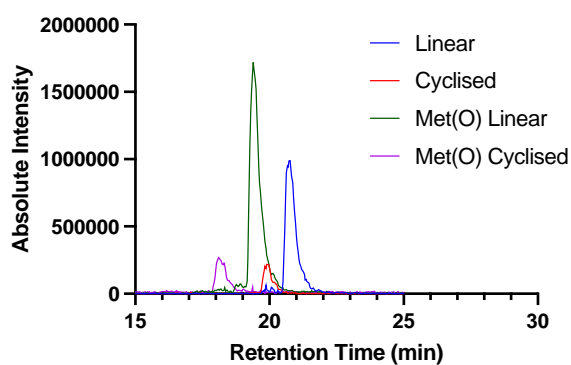
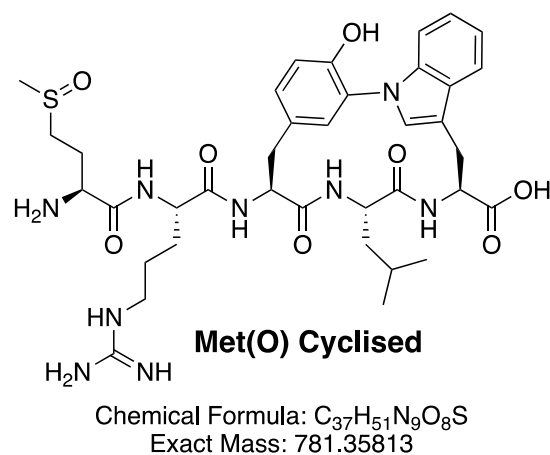
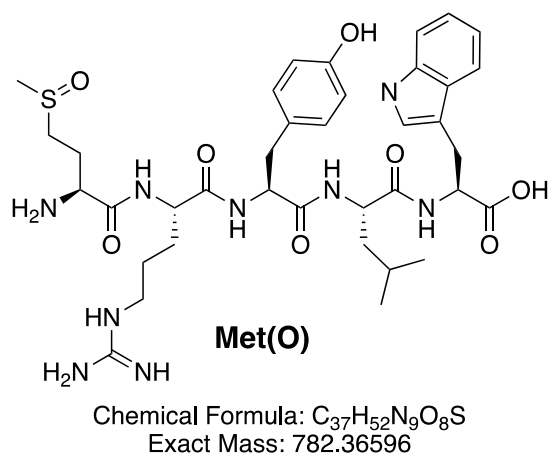
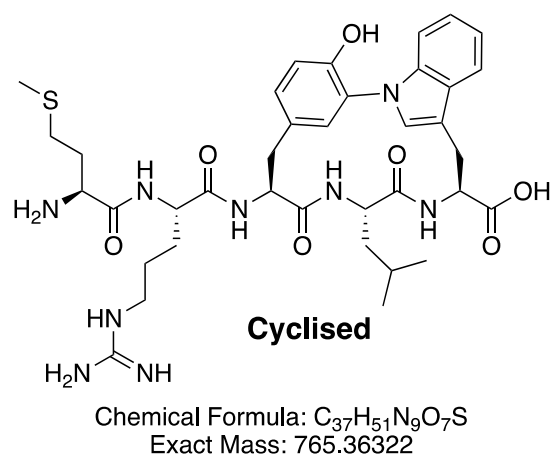
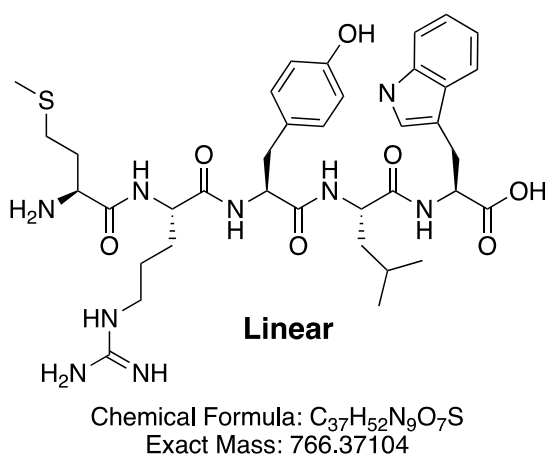


### <Spectrum>

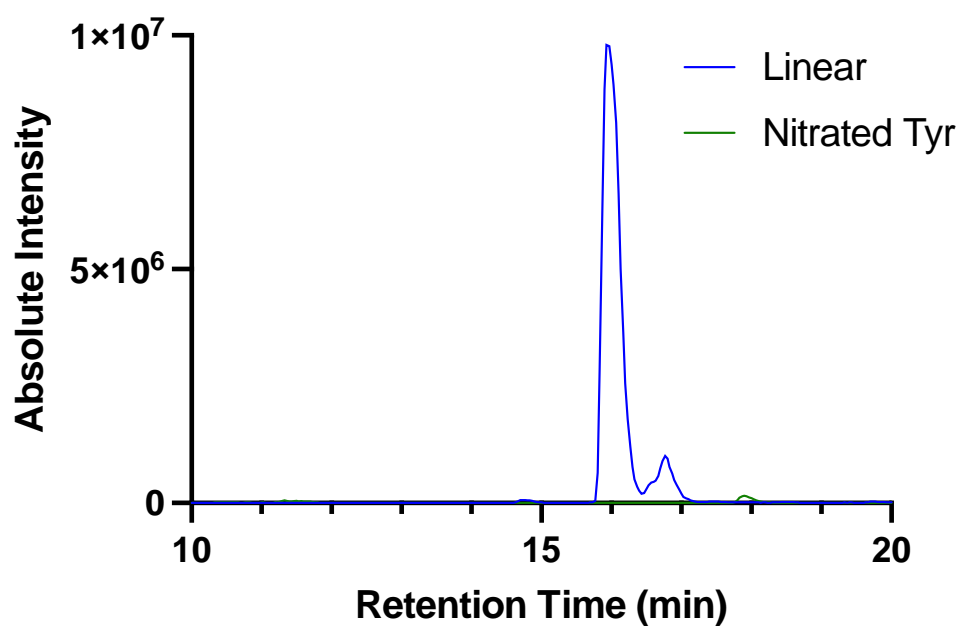
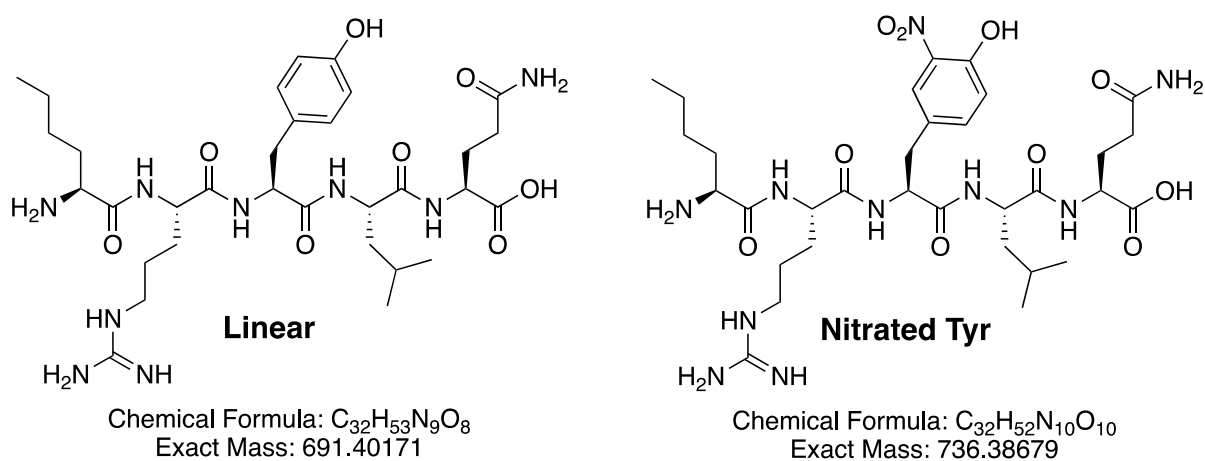
Line#:1 R.Time:----(Scan#:----)  
MassPeaks:1881  
RawMode:Averaged 9.433-9.633(567-579) BasePeak:376(14102981)  
BG Mode:None Segment 1 - Event 1



**Figure S22.** LC and MS analysis of synthetic peptide **6**.

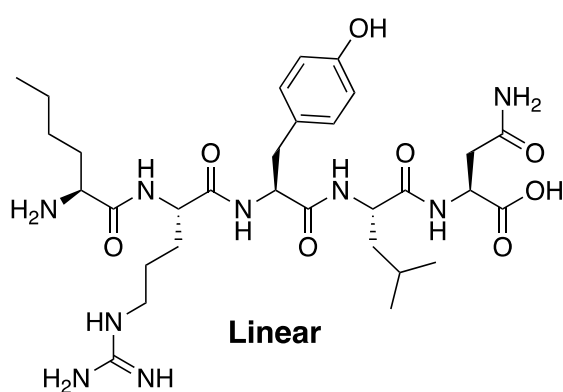


**Figure S23.** Cyclisation data for **2** using RufO<sub>WT</sub> (left) and RufO<sub>SV</sub> (right). The structures of linear and cyclised peptides along with the oxidised methionine versions of these peptides are shown. Overlapping ESI traces for the peptides based on LCMS analysis.

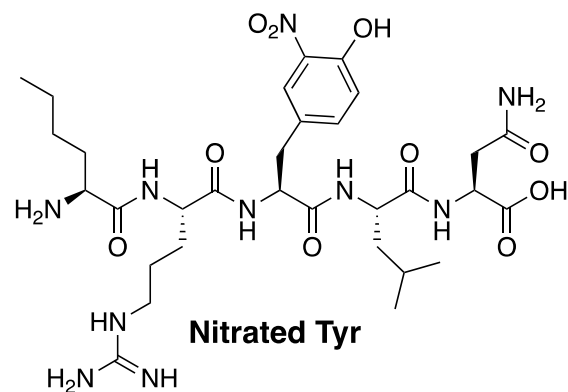


**Figure S24.** Nitration data for **3** using RufO<sub>WT</sub>. The structures of linear and nitrated tyrosine peptides are shown. Overlapping ESI traces for the peptides based on LCMS analysis.

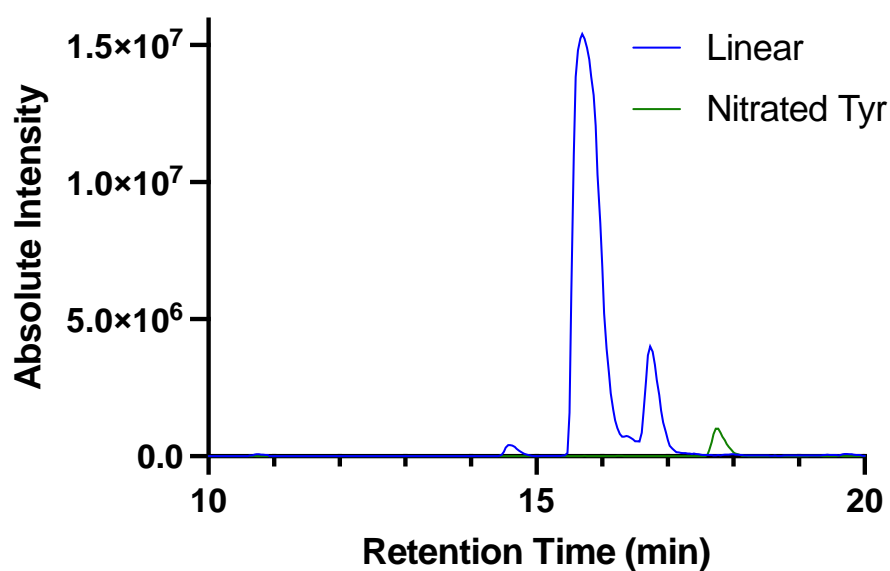




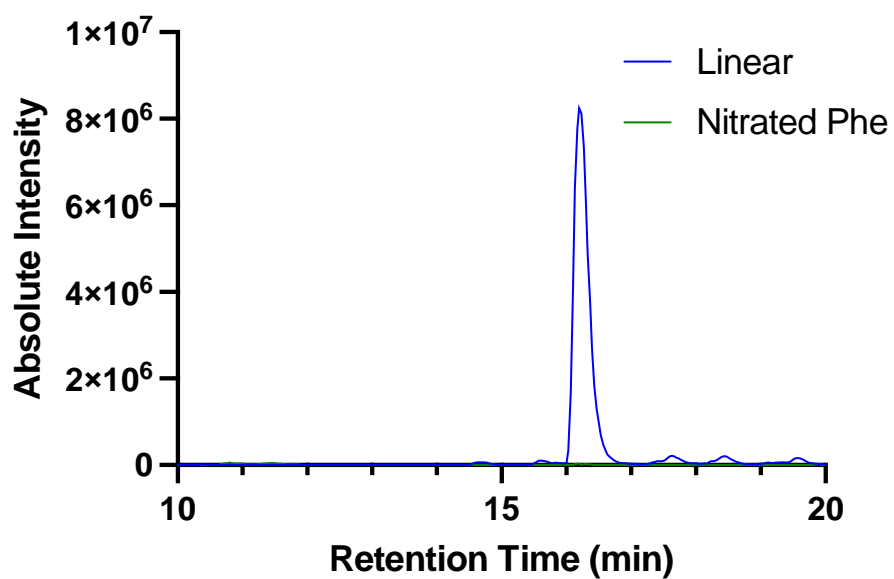
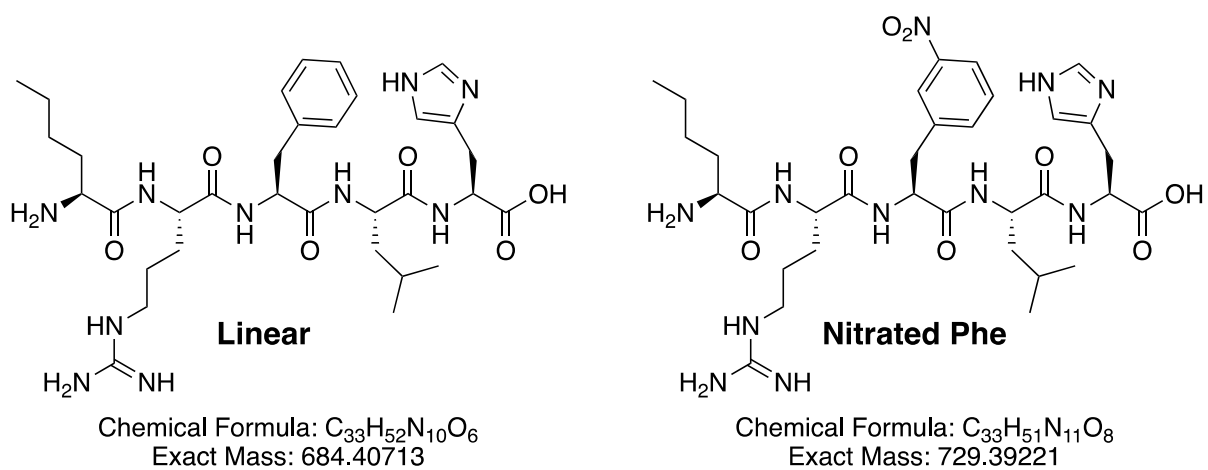
Chemical Formula:  $C_{31}H_{51}N_9O_8$   
Exact Mass: 677.38606



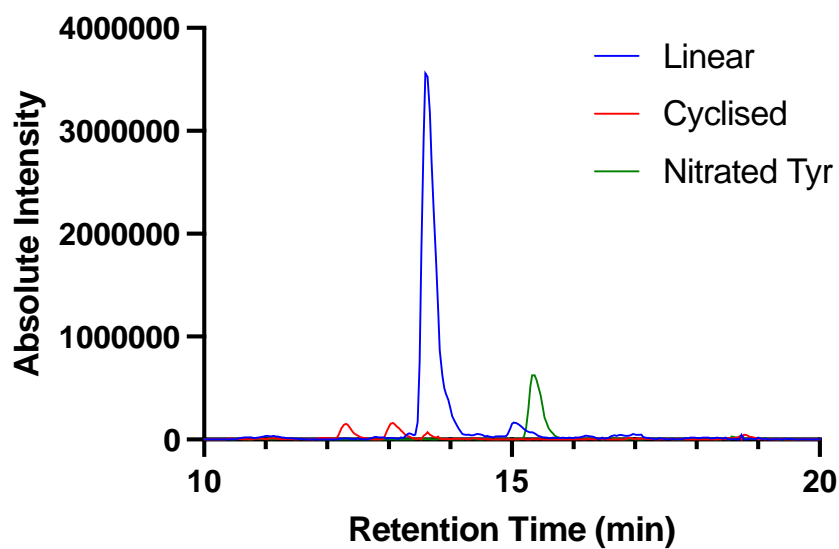
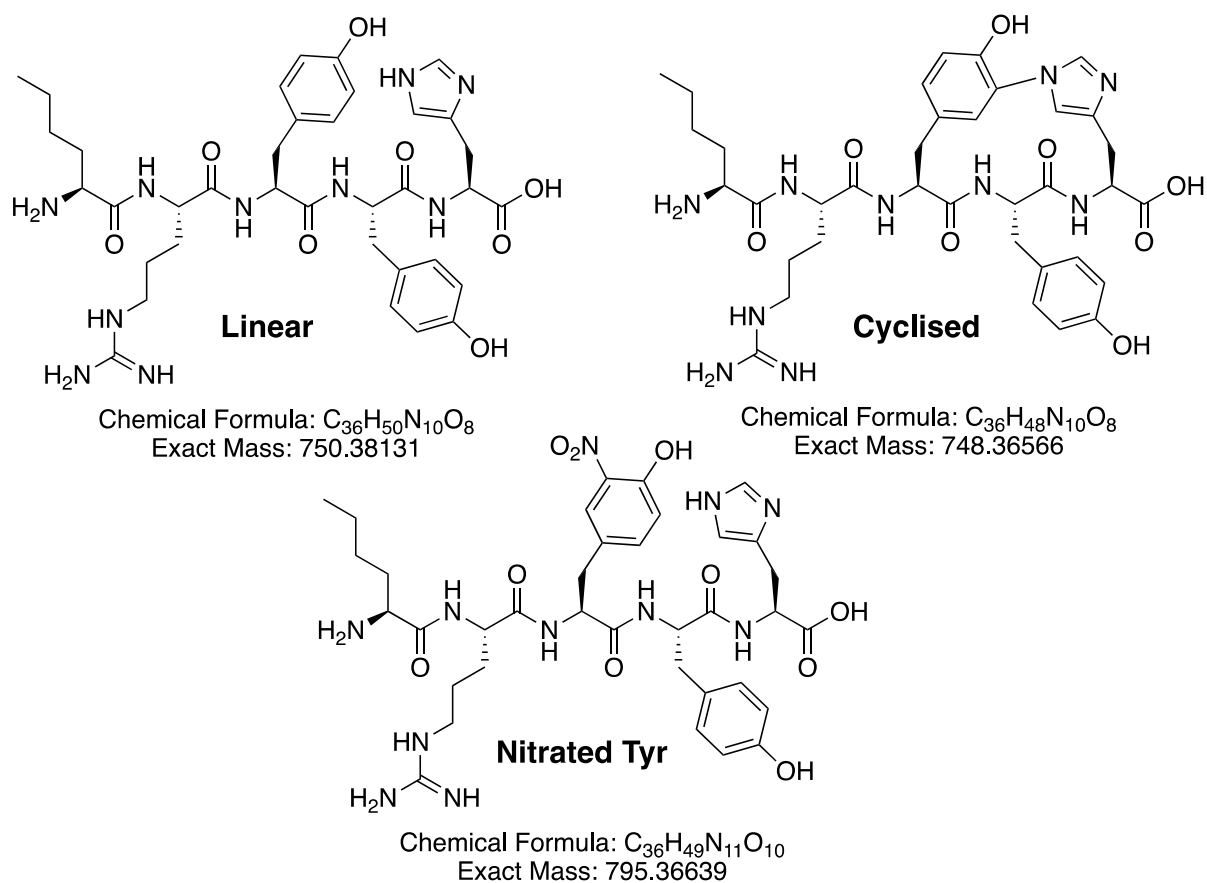
Chemical Formula:  $C_{31}H_{50}N_{10}O_{10}$   
Exact Mass: 722.37114



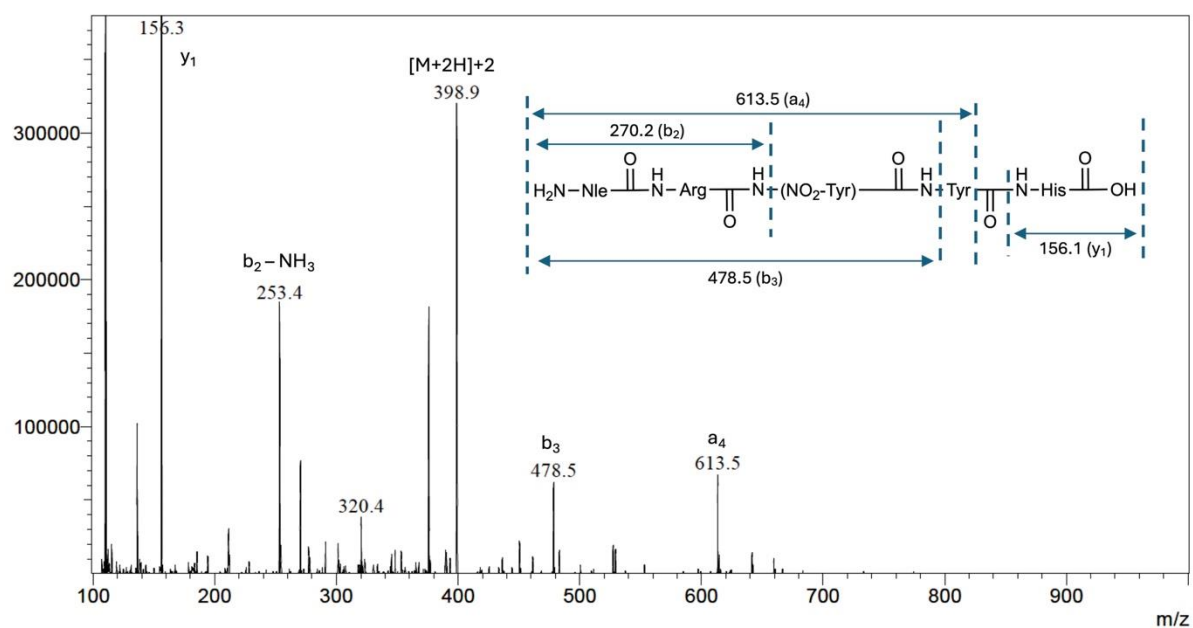
**Figure S25.** Nitration data for **4** using RufO<sub>WT</sub>. The structures of linear and nitrated tyrosine peptides are shown. Overlapping ESI traces for the peptides based on LCMS analysis.



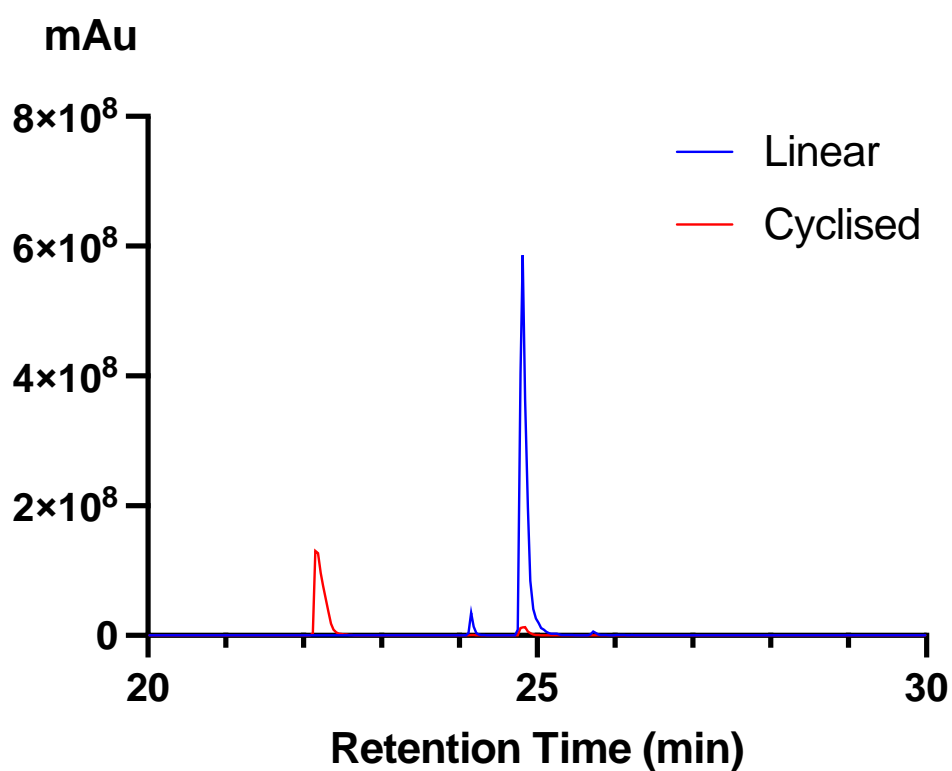
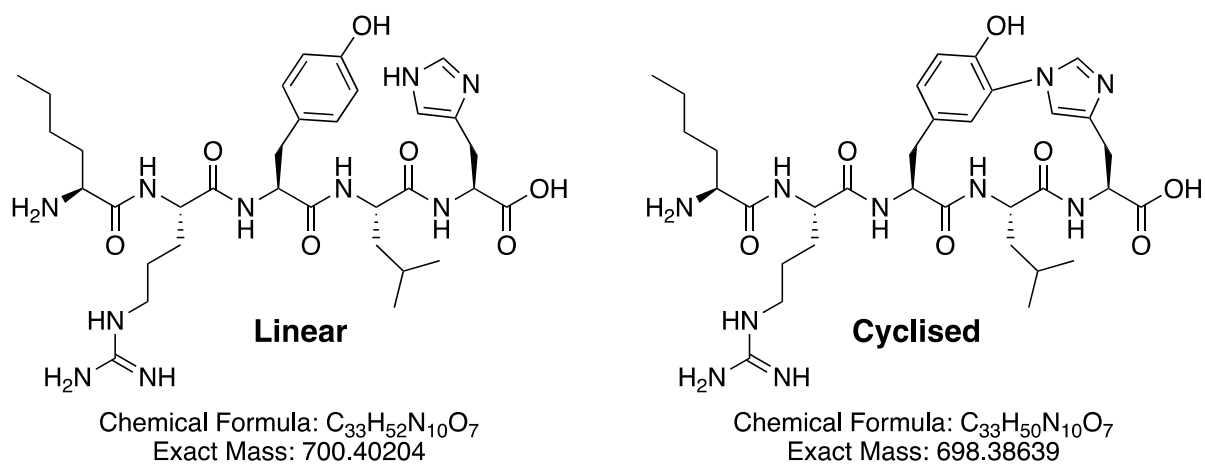
**Figure S26.** Nitration data for **5** using RufO<sub>WT</sub>. The structures of linear and nitrated phenylalanine peptides are shown. Overlapping ESI traces for the peptides based on LCMS analysis.



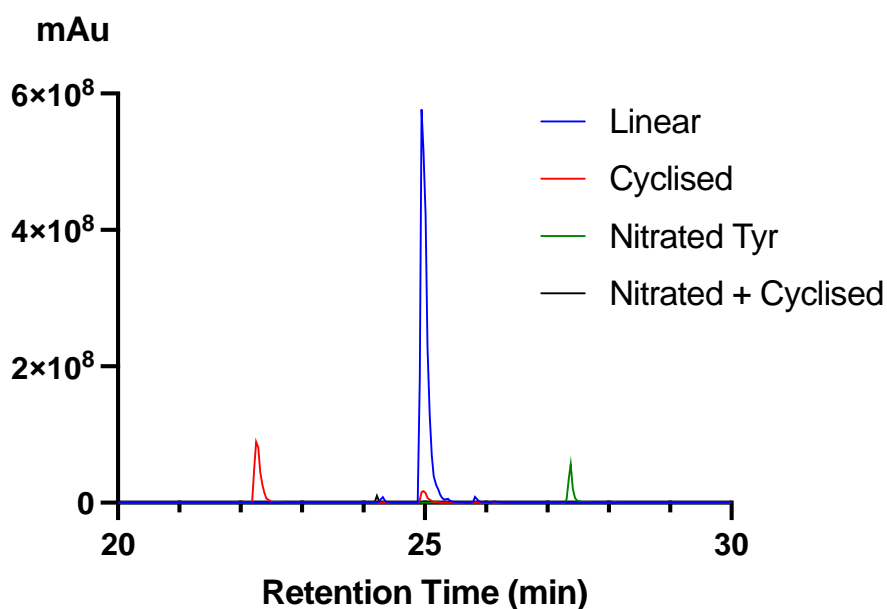
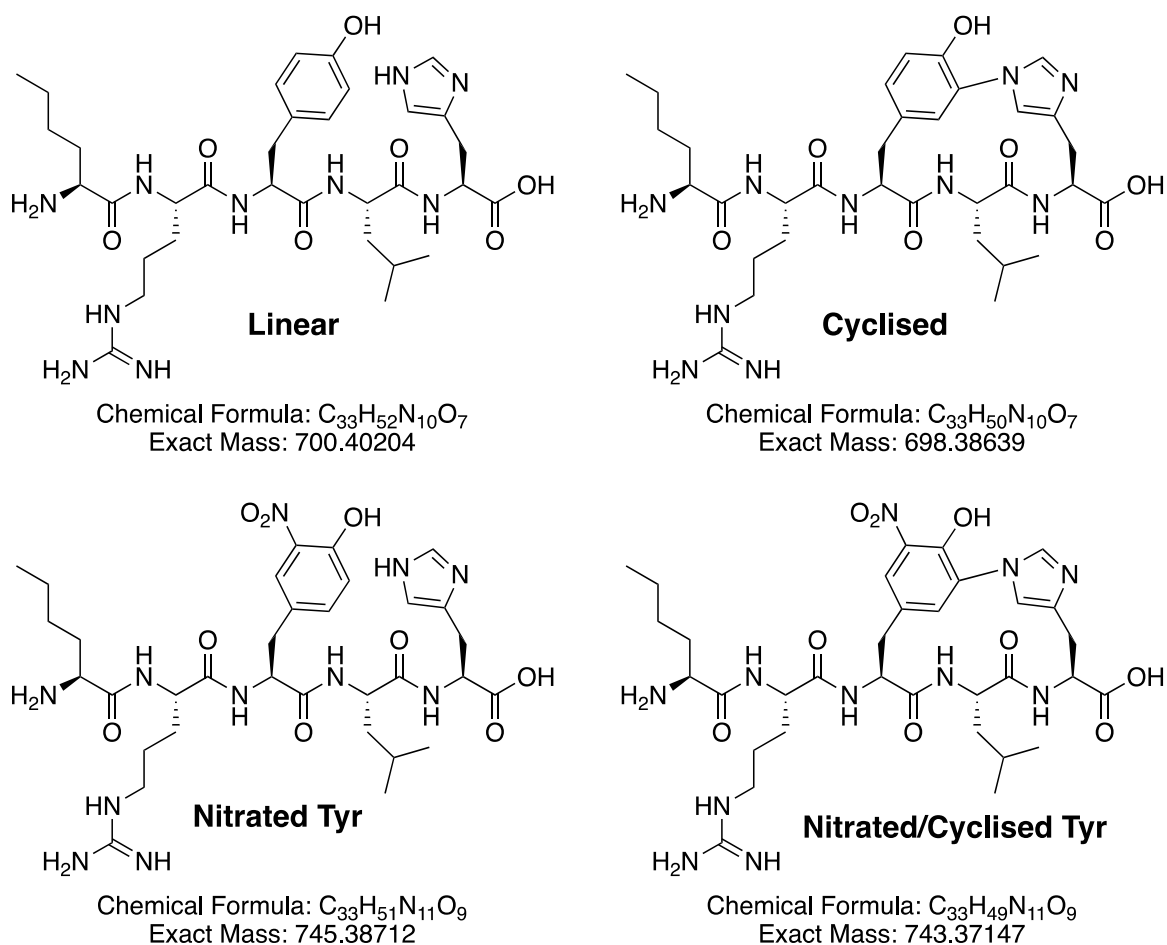
**Figure S27.** Cyclisation and nitration data for **6** using RufO<sub>WT</sub>. The structures of linear and nitrated peptides are shown. Overlapping ESI traces for the peptides based on LCMS analysis.



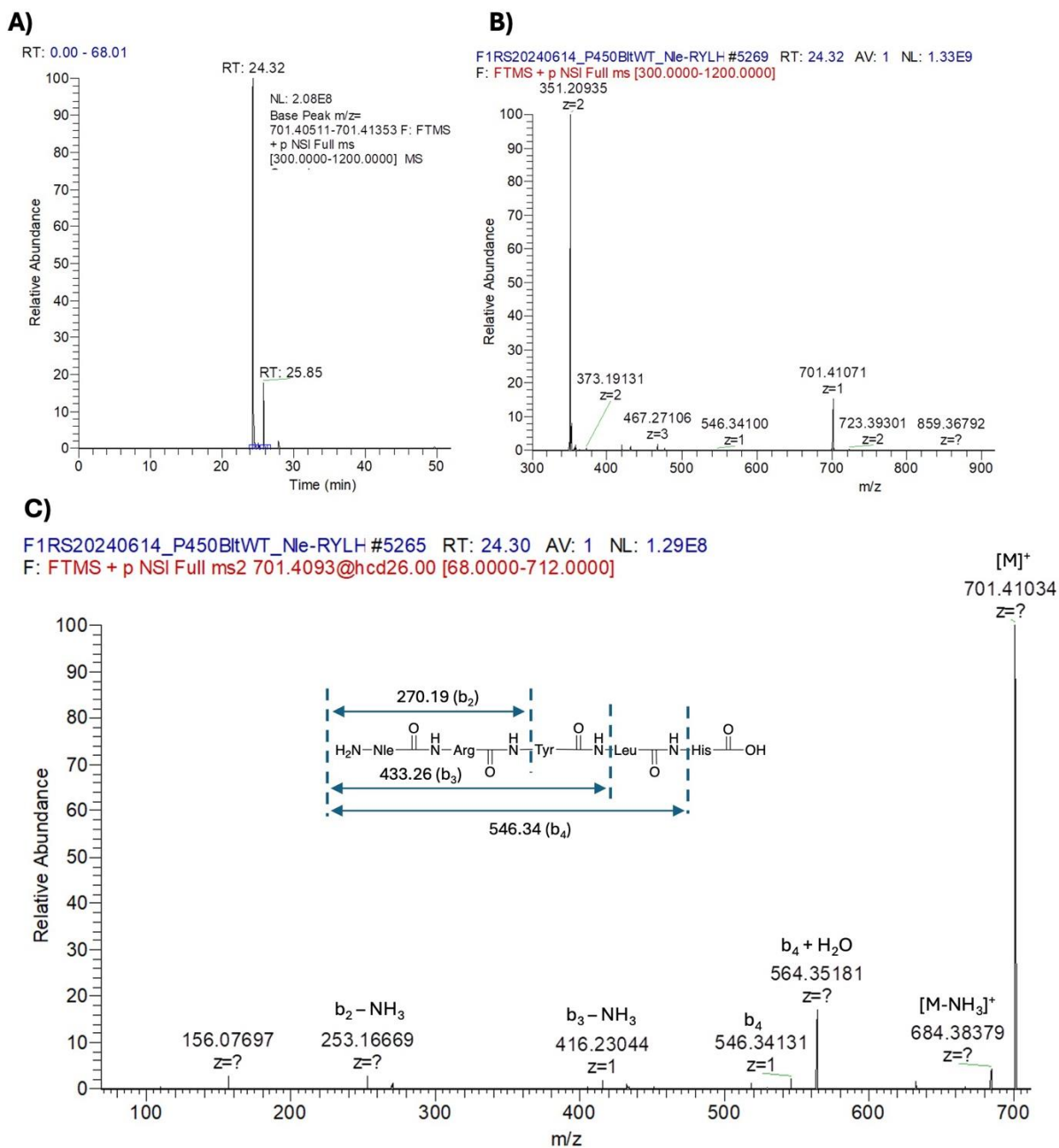
**Figure S28.** Annotated MS<sup>2</sup> data for nitrated peptide 6.



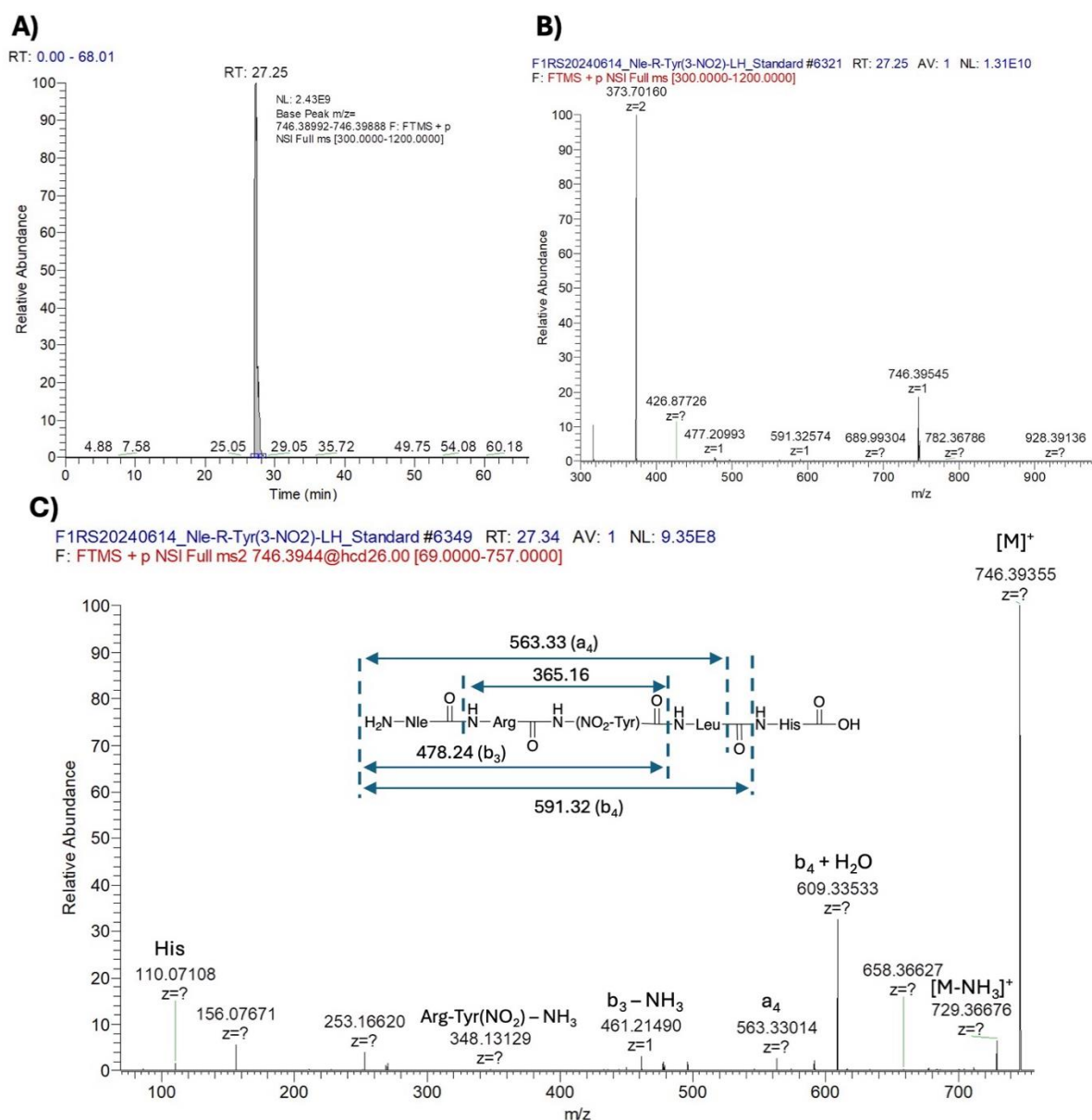
**Figure S29.** Cyclisation data for Nle-1 using  $RufO_{sv}$  in the absence of nitric oxide. The structures of linear and cyclised peptide are shown. Overlapping ESI traces for linear and cyclised peptides based on HRMS analysis.



**Figure S30.** Cyclisation data for Nle-1 using  $RuFO_{SV}$  in the presence of nitric oxide. The structures of linear, cyclic, nitrated tyrosine-containing peptide and nitrated tyrosine and cyclised peptides are shown. Overlapping ESI traces for linear, cyclic, nitrated tyrosine-containing peptide and nitrated tyrosine and cyclised peptides based on HRMS analysis.

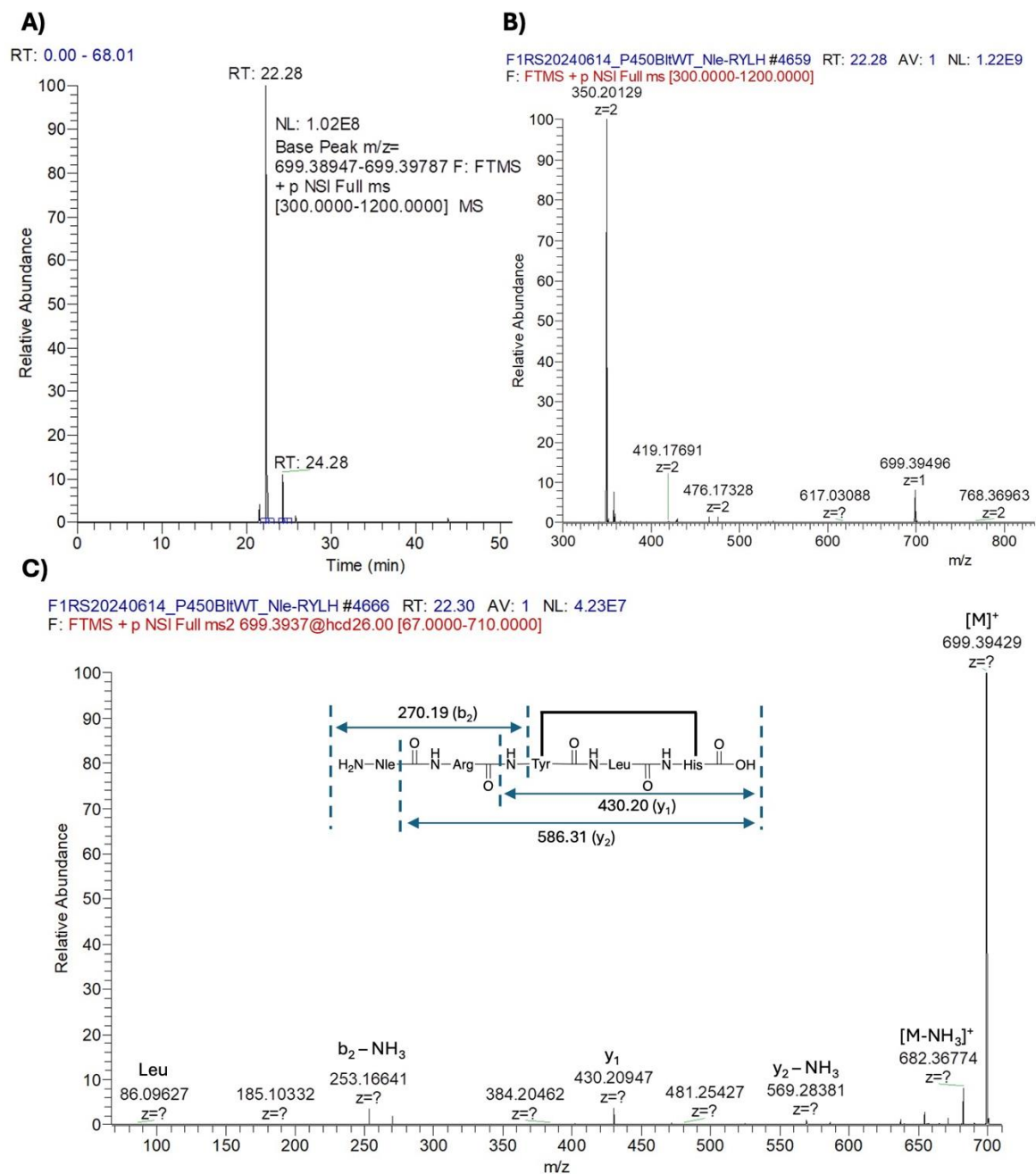


**Figure S31.** HRMS and MS<sup>2</sup> data for linear peptide Nle-1. A) EIC trace of Nle-1. B) HRMS spectrum for Nle-1. C) Annotated MS<sup>2</sup> for Nle-1.

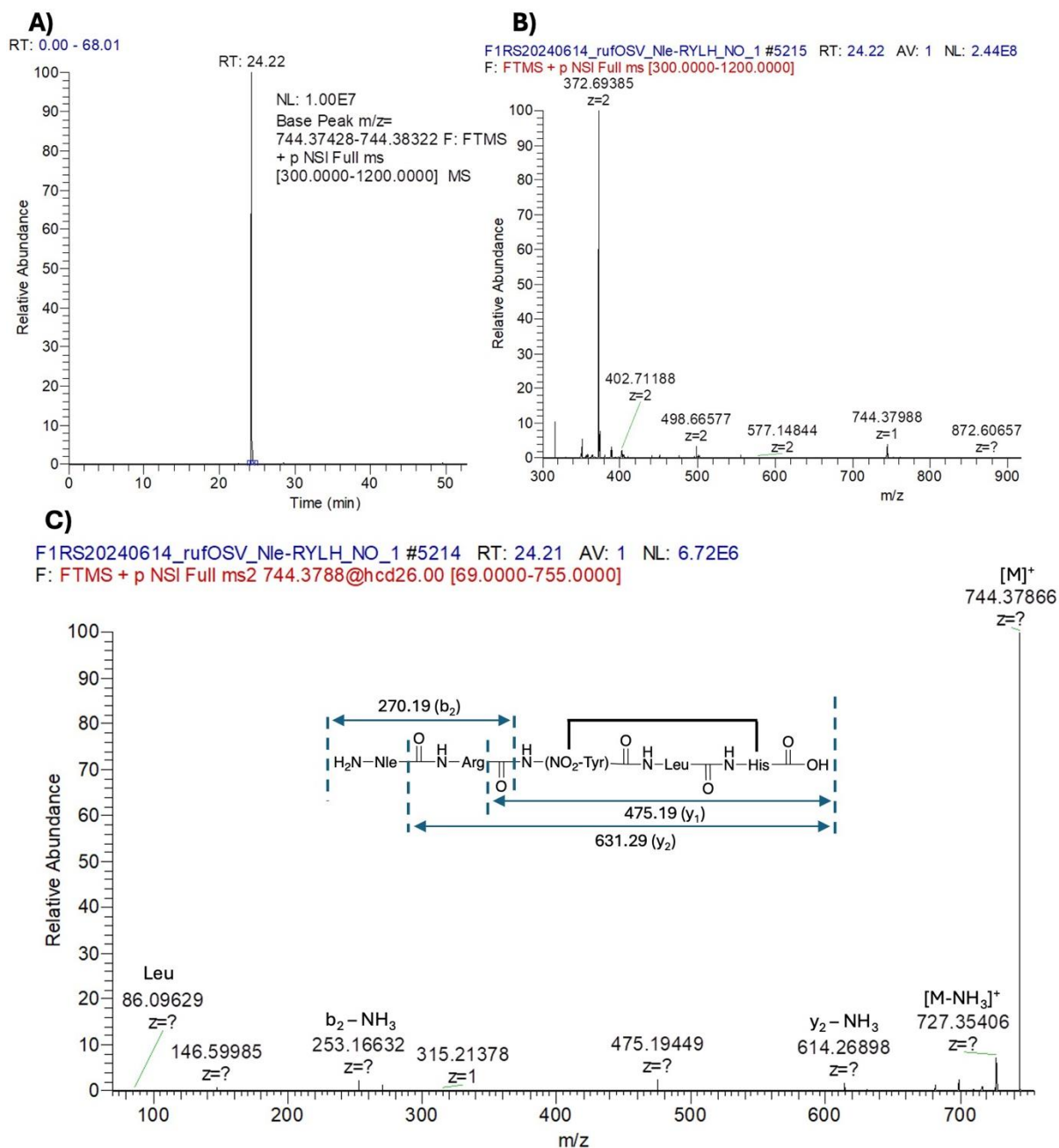


**Figure S32.** HRMS and MS<sup>2</sup> data for nitrated tyrosine-containing peptide Nle/NO<sub>2</sub>-1. A) EIC trace of Nle/NO<sub>2</sub>-1. B) HRMS spectrum for Nle/NO<sub>2</sub>-1. C) Annotated MS<sup>2</sup> for Nle/NO<sub>2</sub>-1.





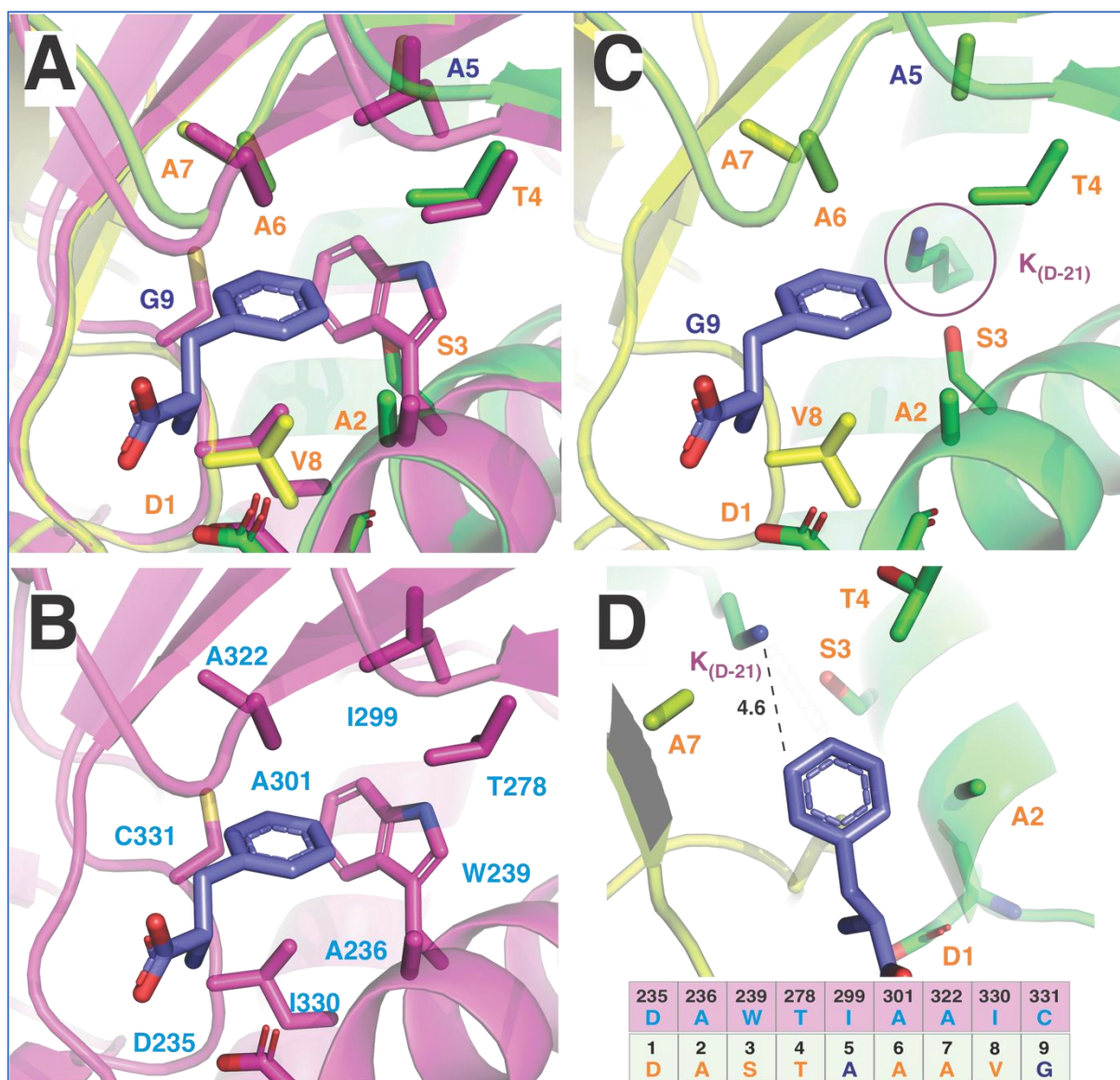
**Figure S33.** HRMS and MS<sup>2</sup> data for cyclic peptide cyclised Nle-1. A) EIC trace of cyclised Nle-1. B) HRMS spectrum for cyclised Nle-1. C) Annotated MS<sup>2</sup> for cyclised Nle-1.



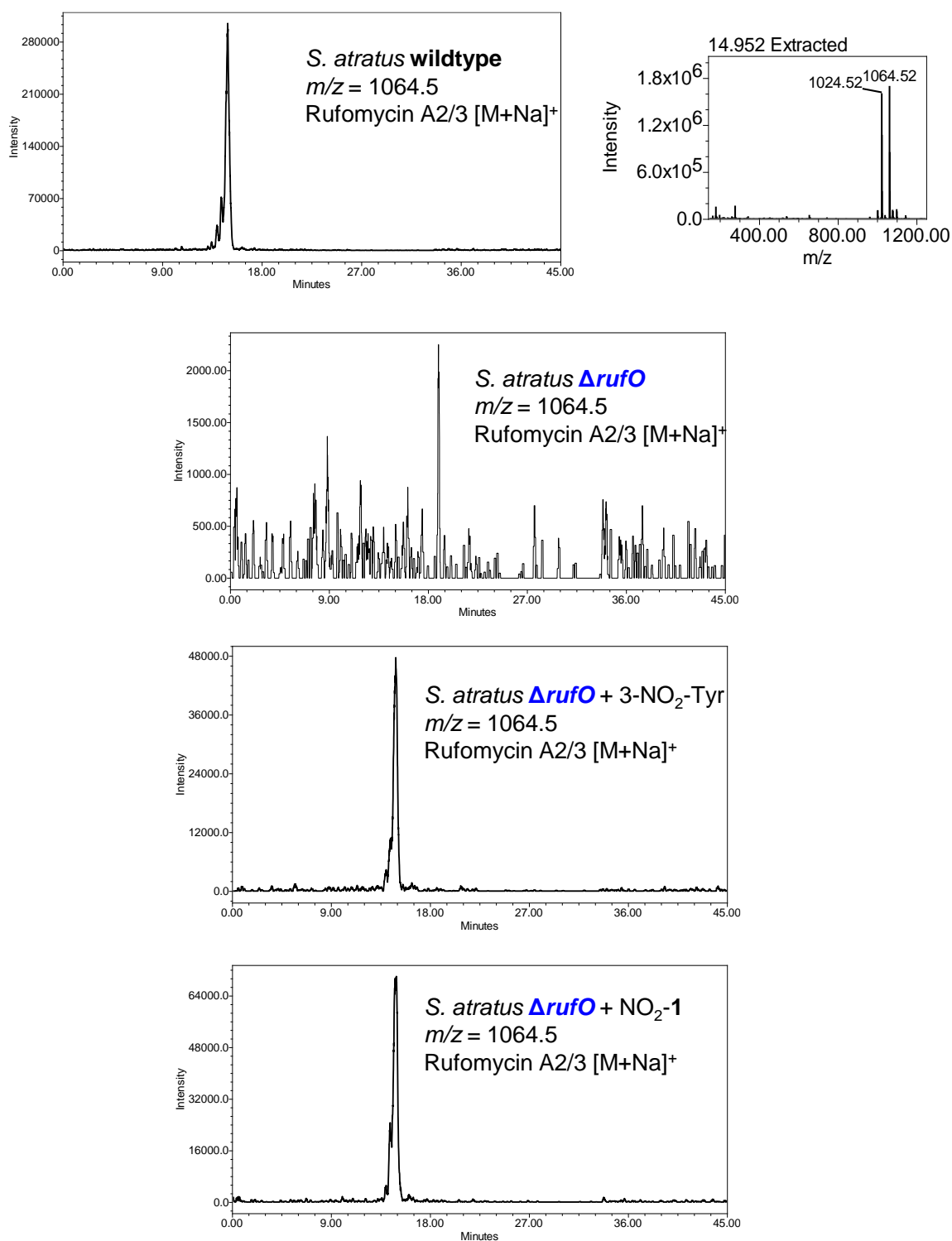
**Figure S34.** HRMS and MS<sup>2</sup> data for cyclised and nitrated tyrosine-containing peptide Nle/NO<sub>2</sub>-1. A) EIC trace of cyclised Nle/NO<sub>2</sub>-1. B) HRMS spectrum for cyclised Nle/NO<sub>2</sub>-1. C) Annotated MS<sup>2</sup> for cyclised Nle/NO<sub>2</sub>-1.

Protein <sup>20</sup>	A <sub>6</sub> selectivity code <sup>21</sup>										Description
A <sub>6</sub> Teicoplanin	D	A	S	T	I	A	G	V	C	K	Tyr permissive pocket (teicoplanin-type)
A <sub>6</sub> UK-68597	D	A	S	T	I	A	G	V	C	K	
A <sub>6</sub> Balhimycin	D	A	S	T	L	G	A	I	C	K	Tyr permissive pocket (balhimycin-type)
A <sub>6</sub> Vancomycin	D	A	S	T	L	G	A	I	C	K	
A <sub>6</sub> Chloroerythromycin	D	A	S	T	L	G	A	I	C	K	
A <sub>6</sub> Arylomycin	D	A	S	T	V	A	A	V	C	K	Tyr specific pocket
A <sub>6</sub> A40926	D	A	S	T	V	A	A	V	C	K	
A <sub>6</sub> Complestatin	D	A	S	T	V	A	A	V	C	K	
A <sub>6</sub> Kistamicin	D	A	S	T	V	A	A	V	C	K	
A <sub>6</sub> Kistamicin-mut	D	A	S	T	I	A	G	V	C	K	Tyr permissive pocket
A <sub>3</sub> Rufomycin (1)	D	A	S	T	A	A	A	V	G	K	NO <sub>2</sub> -Tyr pocket

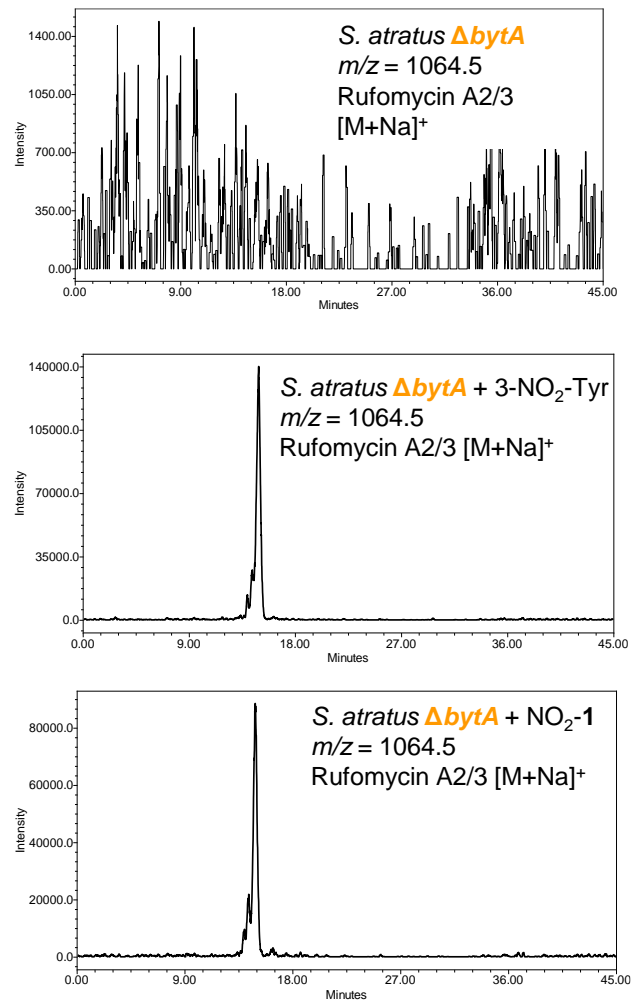
**Figure S35.** Comparative analysis of the RufT A<sub>3</sub> domain specificity-conferring code with other notable tyrosine activating pockets. These include those with tolerance for  $\beta$ -hydroxylation and tolerance for modifications of the Tyr ring (teicoplanin-type permissive pockets), tolerance for  $\beta$ -hydroxylation (balhimycin-type permissive pockets) and those with stringent specificity for Tyr (specific pockets). Important changes in the amino acids within the specificity code are coloured, with those conserved with Tyr-specific pockets shown in green, changes seen in Tyr-permissive pockets coloured yellow (teicoplanin-type) and grey (balhimycin type) and the rufomycin A<sub>3</sub> domain shown in purple; of note in this regard in the Cys to Gly alteration at position 9 of the specificity code.



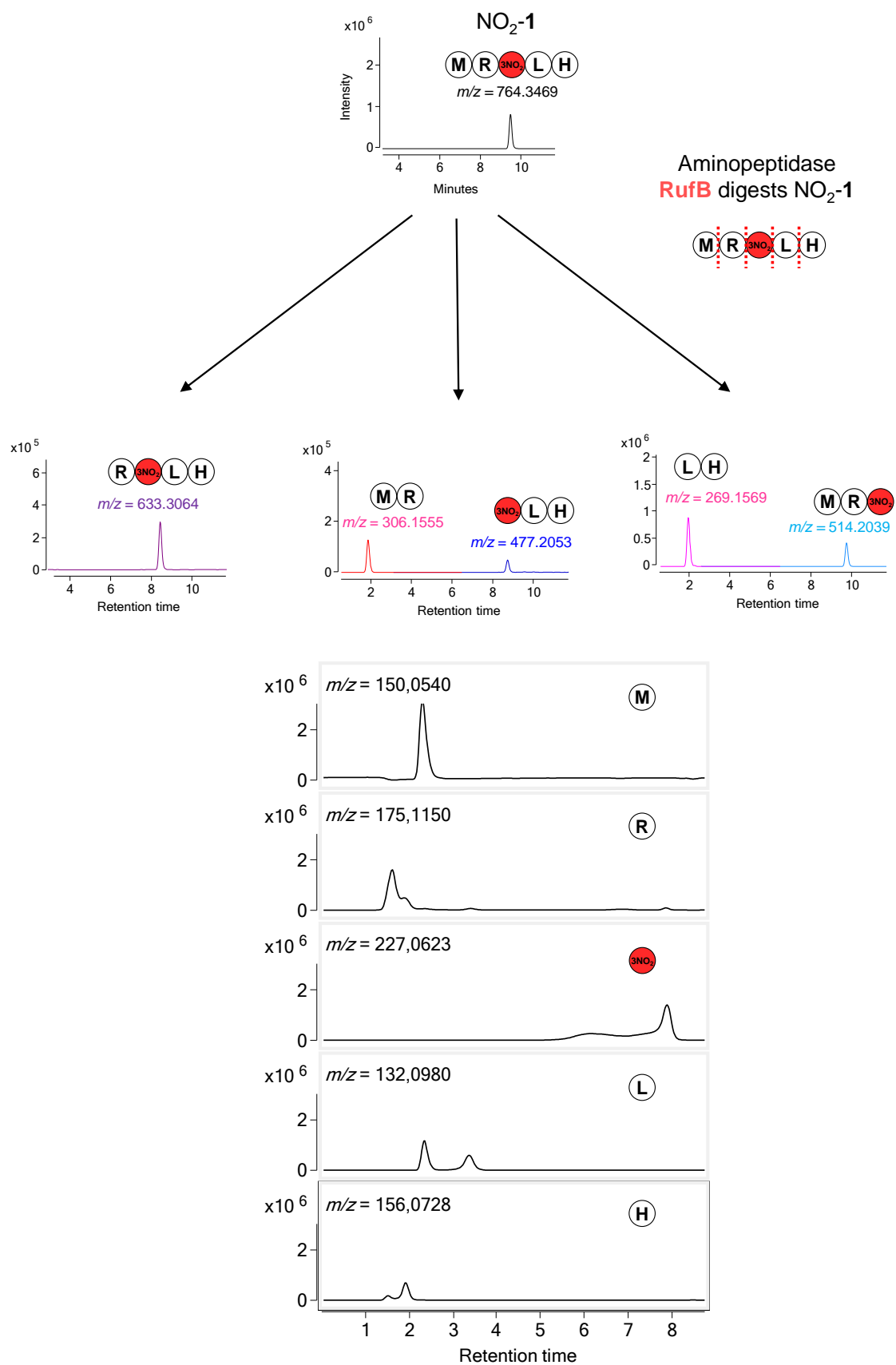
**Figure S36.** Structural analysis of the RufT A<sub>3</sub> selection pocket using an AlphaFold mode<sup>22</sup> in comparison with the phenylalanine activating A-domain PheA (PDB: 1AMU, Chain A).<sup>23</sup> Overlay of the structure (A) with the selectivity encoding residues indicated for RufT (A). Structure of PheA in complex with Phe, with selectivity encoding residues numbered (B). Predicted structure of the RufT A<sub>3</sub> pocket from an AlphaFold model indicating the key residues in the specificity code together with the unusual Lys residue found at D-21 (C). An alternate view of the RufT A<sub>3</sub> domain model showing the distance (in Angstrom) from Lys<sub>D-21</sub> to the *m*-position of Phe in the PheA structure (D). Specificity code residues are list at the bottom right. PheA structure shown in magenta, with phenylalanine shown in blue sticks; RufT A<sub>3</sub> model is shown as Jones' rainbow from *N*- to the C-termini.



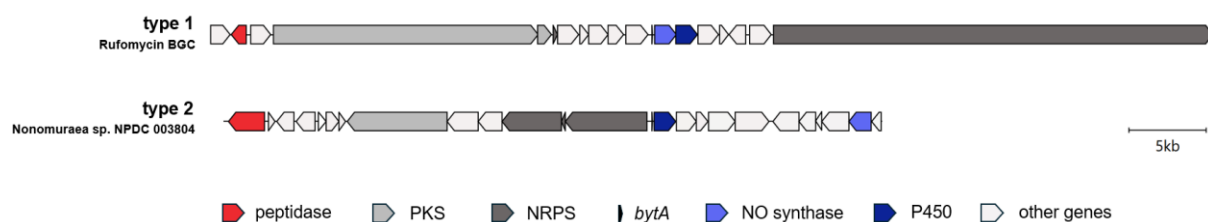
**Figure S37.** Extracted Ion Chromatogram (EIC) data testing rufomycin production in *S. atratus* wildtype and  $\Delta$ *rufO* mutants with and without feeding of 3-NO<sub>2</sub>-Tyr and NO<sub>2</sub>-1. Measured by  $m/z = 1064.5$  corresponding to rufomycin A2/3 [M+Na]<sup>+</sup>. Mass spectrum (top right) displaying the most abundant ions at 14.9 minutes.



**Figure S38.** Extracted Ion Chromatogram (EIC) data testing rufomycin production in *S. atratus*  $\Delta\text{bytA}$  mutants with and without feeding of 3- $\text{NO}_2$ -Tyr and  $\text{NO}_2$ -1. Measured by  $m/z = 1064.5$  corresponding to rufomycin A2/3  $[M+Na]^+$ .



**Figure S39.** RufB completely digests the modified pentapeptide NO<sub>2</sub>-1. Extracted ion chromatograms for corresponding amino acids, di- tri- and tetrapeptides after incubation of NO<sub>2</sub>-1 with RufB are shown.



**Figure S40.** Genome mining results for BGCs harbouring homologs of the necessary genes for nitrotyrosine biosynthesis. Genes involved in the synthesis and incorporation of 3-NO<sub>2</sub>-Tyr are colour-coded. Type 1 includes all rufomycin BGCs (see **Figure S2, Table S8**), where *bytA* and *rufNO* (blue) are colocalized. In type 2 BGCs, *rufN* homologs are located elsewhere in the cluster.



## Supplemental References

1. Tomita, H., Katsuyama, Y., Minami, H., and Ohnishi, Y. (2017). Identification and characterization of a bacterial cytochrome P450 monooxygenase catalyzing the 3-nitration of tyrosine in rufomycin biosynthesis. *J. Biol. Chem.* 292, 15859–15869. 10.1074/jbc.M117.791269.
2. Hansen, M.H., Keto, A., Treisman, M., Sasi, V.M., Coe, L., Zhao, Y., Padva, L., Hess, C., Leichthammer, V., Machell, D.L., et al. (2024). Structural insights into a side chain cross-linking biaryllyl P450 from RiPP biosynthesis *ACS Catal.* 14, 812–826. 10.1021/acscatal.3c05417.
3. Jordan, S., Li, B., Traore, E., Wu, Y., Usai, R., Liu, A., Xie, Z.-R., and Wang, Y. (2023). Structural and spectroscopic characterization of RufO indicates a new biological role in rufomycin biosynthesis. *J. Biol. Chem.* 299, 105049. 10.1016/j.jbc.2023.105049.
4. Schlichting, I., Berendzen, J., Chu, K., Stock, A.M., Maves, S.A., Benson, D.E., Sweet, R.M., Ringe, D., Petsko, G.A., and Sligar, S.G. (2000). The catalytic pathway of cytochrome p450cam at atomic resolution. *Science* 287, 1615-1622. 10.1126/science.287.5458.1615.
5. Boddupalli, S.S., Hasemann, C.A., Ravichandran, K.G., Lu, J.Y., Goldsmith, E.J., Deisenhofer, J., and Peterson, J.A. (1992). Crystallization and preliminary x-ray diffraction analysis of P450terp and the hemoprotein domain of P450BM-3, enzymes belonging to two distinct classes of the cytochrome P450 superfamily. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5567-5571. 10.1073/pnas.89.12.5567.
6. Fujishiro, T., Shoji, O., Nagano, S., Sugimoto, H., Shiro, Y., and Watanabe, Y. (2011). Crystal Structure of H<sub>2</sub>O<sub>2</sub>-dependent Cytochrome P450<sub>SPa</sub> with Its Bound Fatty Acid Substrate: Insight into the Regioselective Hydroxylation of Fatty Acids at the alpha Position. *J. Biol. Chem.* 286, 29941-29950 10.1074/jbc.M111.245225.
7. Haslinger, K., Peschke, M., Brieke, C., Maximowitsch, E., and Cryle, M.J. (2015). X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* 521, 105-109. 10.1038/nature14141.
8. Cupp-Vickery, J.R., and Poulos, T.L. (1995). Structure of cytochrome P450eryF involved in erythromycin biosynthesis. *Nat. Struct. Biol.* 2, 144-153. 10.1038/nsb0295-144.
9. Barnes, H.J., Arlotto, M.P., and Waterman, M.R. (1991). Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 88, 5597-5601. 10.1073/pnas.88.13.5597.
10. Gering, H.E., Li, X., Tang, H., Swartz, P.D., Chang, W.C., and Makris, T.M. (2023). A Ferric-superoxide intermediate initiates P450-catalyzed cyclic dipeptide dimerization. *J. Am. Chem. Soc.* 145, 19256-19264. 10.1021/jacs.3c04542.
11. Cryle, M.J., Bell, S.G., and Schlichting, I. (2010). Structural and Biochemical Characterization of the Cytochrome P450 CypX (CYP134A1) from *Bacillus subtilis*: A Cyclo-L-leucyl-L-leucyl Dipeptide Oxidase. *Biochemistry* 49, 7282-7296. 10.1021/bi100910y.
12. Sabbadin, F., Jackson, R., Haider, K., Tampi, G., Turkenburg, J.P., Hart, S., Bruce, N.C., and Grogan, G. (2009). The 1.5-Å structure of XplA-heme, an unusual cytochrome P450 heme domain that catalyzes reductive biotransformation of royal demolition explosive. *J. Biol. Chem.* 284, 28467-28475. 10.1074/jbc.M109.031559.
13. Barry, S.M., Kers, J.A., Johnson, E.G., Song, L., Aston, P.R., Patel, B., Krasnoff, S.B., Crane, B.R., Gibson, D.M., Loria, R., and Challis, G.L. (2012). Cytochrome P450-catalyzed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis. *Nat. Chem. Biol.* 8, 814–816. 10.1038/nchembio.1048.
14. Zhang, S., Zhang, L., Greule, A., Tailhades, J., Marschall, E., Prasongpholchai, P., Leng, D.J., Zhang, J., Zhu, J., Kaczmarek, J.A., et al. (2023). P450-mediated dehydrotyrosine formation during WS9326 biosynthesis proceeds via dehydrogenation of a specific acylated dipeptide substrate. *Acta Pharm. Sin. B.* 13, 3561-3574. 10.1016/j.apsb.2023.03.021.
15. Tattersall, D.B., Bak, S., Jones, P.R., Olsen, C.E., Nielsen, J.K., Hansen, M.L., Høj, P.B., and Møller, B.L. (2001). Resistance to an Herbivore Through Engineered Cyanogenic Glucoside Synthesis. *Science* 293, 1826-1828. 10.1126/science.1062249.

16. Lee, D.S., Nioche, P., Hamberg, M., and Raman, C.S. (2008). Structural insights into the evolutionary paths of oxylipin biosynthetic enzymes. *Nature* 455, 363-368. 10.1038/nature07307.
17. Dratch, B.D., McWhorter, K.L., Blue, T.C., Jones, S.K., Horwitz, S.M., and Davis, K.M. (2023). Insights into substrate recognition by the unusual nitrating enzyme RufO. *ACS Chem. Biol.* 18, 1713–1718. 10.1021/acscchembio.3c00328.
18. Zhao, Y., Marschall, E., Treisman, M., McKay, A., Padva, L., Crüsemann, M., Nelson, D.R., Steer, D.L., Schittenhelm, R.B., Tailhades, J., and Cryle, M.J. (2022). Cytochrome P450Blt enables versatile peptide cyclisation to generate histidine- and tyrosine-containing crosslinked tripeptide building blocks. *Angew. Chem. Int. Ed.* 61, e202204957. 10.1002/anie.202204957.
19. Treisman, M., Coe, L., Zhao, Y., Sasi, V.M., Gullick, J., Hansen, M.H., Ly, A., Leichthammer, V., Hess, C., Machell, D.L., et al. (2024). An engineered biaryllytide cross-linking P450 from RiPP biosynthesis generates alternative cyclic peptides. *Org. Lett.* 26, 1828-1833. 10.1021/acs.orglett.3c04366.
20. Kaniusaite, M., Tailhades, J., Marschall, E.A., Goode, R.J.A., Schittenhelm, R.B., and Cryle, M.J. (2019). A proof-reading mechanism for non-proteinogenic amino acid incorporation into glycopeptide antibiotics. *Chem. Sci.* 10, 9466-9482. 10.1039/c9sc03678d.
21. Stachelhaus, T., Mootz, H.D., Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* 6, 493-505. 10.1016/S1074-5521(99)80082-9.
22. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunkool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589. 10.1038/s41586-021-03819-2.
23. Conti, E., Stachelhaus, T., Marahiel, M.A., Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* 16, 4174-4183. 10.1093/emboj/16.14.4174.

## 7.5 Appendix E

**Table 20** Biarylthide P450s identified through machine learning based genome mining. The table presents protein reference IDs, bacterial species, precursor sequences, number of precursors, and co-occurring proteins for each P450 enzyme. The associated tree is presented in Figure 9.

Protein reference	Species	Precursor sequence	Number of precursors	Co-occurring proteins
AIK38645.1	<i>Bacillus pseudomycooides</i>	MNYNH*	1	
CUB11028.1	<i>Bacillus cereus</i>	MIYDH*	1	
GAA1025571.1	<i>Amycolatopsis albidoflavus</i>	MLYRH*	1	
GAA2260327.1	<i>Actinomadura glauciflava</i>	MRYLH*	1	
GAA4057093.1	<i>Actinomadura miaoliensis</i>	MKYIW*	1	
GAA4139197.1	<i>Actinomadura keratinilytica</i>	MKYLW*	1	
GAA4372804.1	<i>Actinomadura verrucosospora</i>	MRYLH*	1	
GAA4493708.1	<i>Actinoallomurus oryzae</i>	MRYVH*	1	
GGV42769.1	<i>Actinomadura cremea</i>	MRYLH*	1	
PEP54637.1	<i>Bacillus pseudomycooides</i>	MNYLH*	1	
PGX89698.1	<i>Bacillus cereus</i>	MNYIH*	1	
TNP01825.1	<i>Bacillus pacificus</i>	MNYNH*	1	
WP_015623560.1	<i>Actinoplanes</i> sp. N902-109	MQYDH*	1	Methyltransferase

## Appendix

WP_026422380.1	<i>Actinokineospora inagensis</i> DSM 44258	MSYDH*	1	PGM_ligase
WP_034520669.1	<i>Actinomadura rifamycini</i> DSM 43936	MRYLH*	1	Hydroxylase, Oxidoreductase, Oxidoreductase, Methyltransferase, Acetyltransferase, Protease/Peptidase, Sulfotransferase
WP_067473018.1	<i>Actinomadura hibisca</i> NBRC 15177	MKYWH*	1	additional_P450, Hydroxylase, Methyltransferase
WP_067804628.1	<i>Actinomadura formosensis</i> NBRC 14204	MRYLH*	1	Hydroxylase, Methyltransferase, Protease/Peptidase
WP_075974605.1	<i>Actinokineospora bangkokensis</i>	MSYDHIS*	1	additional_P450
WP_080039781.1	<i>Actinomadura parvosata subsp. kistnae</i>	MRYRH*	1	Hydroxylase
WP_098623608.1	<i>Bacillus pseudomycooides</i>	MNYNH*	1	
WP_113692517.1	<i>Amycolatopsis albispora</i>	MDYPH*	1	Methyltransferase, Protease/Peptidase
WP_117356064.1	<i>Actinomadura logoneensis</i>	MKYWH*	1	
WP_122195090.1	<i>Actinomadura harenae</i>	MKYWH*	1	Hydroxylase, Methyltransferase
WP_125616065.1	<i>Actinomadura</i> sp. WAC 06369	MRYLH*	1	Hydroxylase, Methyltransferase, Acetyltransferase, Protease/Peptidase, Sulfotransferase
WP_127361668.1	<i>Actinacidiphila soli</i>	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_138637966.1	<i>Actinomadura geliboluensis</i>	MRYLH*	1	Hydroxylase, Methyltransferase, Protease/Peptidase, Sulfotransferase
WP_149261402.1	<i>Actinomadura</i> sp. K4S16	MRYLH*	1	Hydroxylase, Methyltransferase, Protease/Peptidase, Sulfotransferase
WP_157429744.1	<i>Actinomadura oligospora</i> ATCC 43269	MKYWH*	1	Hydroxylase, Methyltransferase

## Appendix

WP_160826862.1	<i>Actinomadura</i> sp. J1-007	MRYRH*	1	Hydroxylase, Acetyltransferase, Protease/Peptidase
WP_169813928.1	<i>Actinomadura kijaniata</i> NBRC 14229	MRHVH*	1	Methyltransferase, Sulfotransferase
WP_179846876.1	<i>Actinomadura luteofluorescens</i>	MRYLH*	1	Hydroxylase, Methyltransferase, Protease/Peptidase, Sulfotransferase
WP_182847545.1	<i>Actinomadura namibiensis</i>	MRHVH*	1	Methyltransferase, Sulfotransferase
WP_185026858.1	<i>Actinomadura coerulea</i>	MRYLH*	1	Hydroxylase, Methyltransferase, Protease/Peptidase, Sulfotransferase
WP_187242675.1	<i>Actinomadura alba</i>	MRYIH*	1	Hydroxylase, Aminotransferase, Methyltransferase, Protease/Peptidase, Sulfotransferase
WP_203405941.1	<i>Archangium violaceum</i>	MNYLH*	1	Methyltransferase
WP_203658450.1	<i>Actinocatenispora rupis</i>	MRYWY*	1	Hydroxylase, Methyltransferase, Protease/Peptidase, Sulfotransferase
WP_208263840.1	<i>Actinomadura barringtoniae</i>	MRYWY*	1	Oxidoreductase, Methyltransferase, Sulfotransferase
WP_220506603.1	<i>Amycolatopsis dendrobii</i>	MLYVH*	1	Acetyltransferase
WP_221320770.1	<i>Actinoplanes</i> sp. L3-i22	MRYFH*	1	Methyltransferase
WP_329238118.1	<i>Actinoallomurus</i> sp. NBC_01490	MRYVH*	1	Protease/Peptidase
WP_097787522.1	<i>Bacillus</i> sp. FSL K6-6038	MNYNH*	1	
WP_088119720.1	<i>Bacillus thuringiensis</i>	MNYNH*	1	Oxidoreductase
WP_212211904.1	<i>Bradyrhizobium</i> sp.	MLWPH*	1	Protease/Peptidase

## Appendix

WP_131770571.1	<i>Candidatus Protofrankia californiensis</i>	MIYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_122147575.1	<i>Cellulomonas triticagri</i>	MTYNHIV*	1	PGM_ligase, Protease/Peptidase
GAA1714241.1	<i>Fodinicola feengrottensis</i>	MRYLH*	1	
WP_073388255.1	<i>Jatrophihabitans endophyticus</i>	MTYVH*	1	Oxidoreductase, Protease/Peptidase
WP_054292478.1	<i>Kibdelosporangium phytohabitans</i>	MNYRH*	1	
WP_052478416.1	<i>Kibdelosporangium</i> sp. MJ126-NF4	MHYKH*	1	Aminotransferase
WP_123563405.1	<i>Kitasatospora cineracea</i>	MSYDH*	1	Aminotransferase, PGM_ligase
WP_184914010.1	<i>Kitasatospora gansuensis</i>	MRYDY*	1	Methyltransferase
WP_184933437.1	<i>Kitasatospora kifunensis</i>	MSYDH*	1	Hydroxylase, Oxidoreductase, PGM_ligase
GAA1228896.1	<i>Kitasatospora nipponensis</i>	MGYIH*	1	
WP_329486261.1	<i>Kitasatospora</i> sp. NBC_01246	MSYNH*	2	Protease/Peptidase
WP_329561375.1	<i>Kitasatospora</i> sp. NBC_01266	MSYDH*	1	Methyltransferase, PGM_ligase
WP_063857100.1	<i>Kitasatospora</i> sp. Root107	MRYDY*	1	Methyltransferase, Protease/Peptidase
WP_051817038.1	<i>Kitasatospora</i> sp. YST-16	MSYDH*	1	PGM_ligase
WP_221339774.1	<i>Kutzneria kofuensis</i>	MSYDH*	1	Oxidoreductase
WP_330854646.1	<i>Lentzea</i> sp.	MSYDH*	1	Protease/Peptidase
WP_089955640.1	<i>Lentzea xinjiangensis</i>	MRYLH*	1	Hydroxylase, Methyltransferase

## Appendix

SKB57425.1	<i>Lysinibacillus</i> sp. AC-3	MQWNY*	3	
WP_197142288.1	<i>Lysinibacillus sphaericus</i>	MQWNY*	3	
TQR30875.1	<i>Lysinibacillus sphaericus</i>	MQWNY*	4	
WP_323657460.1	<i>Lysinibacillus xylanilyticus</i>	MQWNY*	6	
GDX38889.1	<i>Methylocystaceae bacterium</i>	MKYLY*	1	
WP_166795220.1	<i>Micrococcus yunnanensis</i>	MTYDHIL*	1	PGM_ligase, additional_P450
RQX01605.1	<i>Micromonospora arida</i>	MSYDH*	1	
WP_208601958.1	<i>Micromonospora eburnea</i>	MRYIH*	1	Protease/Peptidase
WP_208812380.1	<i>Micromonospora echinofusca</i>	MRYLH*	1	Oxidoreductase, Acetyltransferase
WP_145776703.1	<i>Micromonospora olivasterospora</i>	VRYLH*	1	Methyltransferase, Protease/Peptidase
WP_184536939.1	<i>Micromonospora polyrhachis</i>	MRHYY*	2	Aminotransferase, Methyltransferase, Hydroxylase, Protease/Peptidase
WP_128137096.1	<i>Micromonospora provocatoris</i>	MRYAY*	1	Hydroxylase, Protease/Peptidase
WP_201752802.1	<i>Micromonospora rubida</i>	MRYLH*	1	Protease/Peptidase
WP_216932370.1	<i>Micromonospora</i> sp. b486	MRYAY*	1	Hydroxylase, Acetyltransferase
WP_165878048.1	<i>Micromonospora</i> sp. MW-13	MRYWH*	1	Hydroxylase, Methyltransferase
WP_207892810.1	<i>Micromonospora</i> sp. MW-13	MRYLH*	1	Protease/Peptidase
WTY81129.1	<i>Micromonospora</i> sp. NBC_01405	MRYLH*	1	Protease/Peptidase

## Appendix

WTY49331.1	<i>Micromonospora</i> sp. NBC_01412	MRYLH*	1	Hydroxylase, Protease/Peptidase, Sulfotransferase
WP_128137097.1	<i>Micromonospora tulbaghia</i>	MRYAY*	1	Hydroxylase, Methyltransferase, Protease/Peptidase
WP_325687722.1	<i>Mycobacterium</i> sp.	MFYWH*	1	
WP_101951728.1	<i>Mycobacterium</i> sp. 3519A	MFYWH*	1	
WP_085304054.1	<i>Mycolicibacillus koreensis</i>	MFYWH*	1	
WP_163895064.1	<i>Mycolicibacterium hippocampi</i>	MFYWH*	1	
WP_195096709.1	<i>Nocardia blacklockiae</i>	MSYDHVAV*	1	Oxidoreductase
WP_203236935.1	<i>Nocardia panacis</i>	MDYPH*	1	Aminotransferase, Acetyltransferase, Protease/Peptidase
WP_225731383.1	<i>Nocardia</i> sp. JCM 34519	MRHEY*	1	Methyltransferase
WTL55286.1	<i>Nocardia</i> sp. NBC_01499	MMYWH*	1	Methyltransferase
WP_017622485.1	<i>Nocardiopsis chromatogenes</i> YIM 90109	MRYIH*	1	Hydroxylase
WP_192775534.1	<i>Nonomuraea africana</i>	VRYLH*	1	Oxidoreductase
WP_138720392.1	<i>Nonomuraea basaltis</i>	MRYLH*	1	NO_synthase, Methyltransferase, Protease/Peptidase
WP_195896979.1	<i>Nonomuraea cypriaca</i>	MRYLH*	1	NO_synthase, Methyltransferase, Protease/Peptidase
GAA3599678.1	<i>Nonomuraea rosea</i>	MRYWY*	1	



## Appendix

WP_103962916.1	<i>Nonomuraea solani</i>	MRYLH*	1	NO_synthase, Methyltransferase, Protease/Peptidase, Oxidoreductase
WP_103964457.1	<i>Nonomuraea solani</i>	MRYRH*	1	Aminotransferase, Hydroxylase
QLG40373.1	<i>Paenibacillus</i> sp. E222	MIYTH*	3	PGM_ligase
WP_203790048.1	<i>Paractinoplanes rishiriensis</i>	MSYDH*	1	
WP_197938166.1	<i>Phytohabitans flavus</i>	MRYIH*	1	Hydroxylase, Methyltransferase, Acetyltransferase, Protease/Peptidase
BCB75145.1	<i>Phytohabitans flavus</i>	MRYIH*	1	Methyltransferase, Acetyltransferase, Protease/Peptidase
WP_196450865.1	<i>Planomonospora</i> ID82291	MRYYH*	1	Oxidoreductase, Protease/Peptidase
WP_136921689.1	<i>Polyangium aurulentum</i>	MSYDH*	2	
WP_212826894.1	<i>Polymorphospora rubra</i>	MRYLH*	1	Oxidoreductase
WP_163988023.1	<i>Pyxidicoccus caerfyrddinensis</i>	MNYLH*	1	Methyltransferase
WP_169344400.1	<i>Pyxidicoccus fallax</i>	MNYLH*	1	Methyltransferase, Protease/Peptidase
WP_095343844.1	<i>Rothia dentocariosa</i>	MQYAH*	1	Oxidoreductase, additional_P450
WP_070535711.1	<i>Rothia</i> sp. HMSC058E10	MQYAH*	1	Oxidoreductase, additional_P450
WP_203920145.1	<i>Rugosimonospora africana</i>	MKYWH*	1	Hydroxylase, Methyltransferase
WP_211898415.1	<i>Saccharopolyspora erythraea</i>	MDYPH*	1	Protease/Peptidase
EQD85547.1	<i>Saccharopolyspora erythraea</i> D	MDYPH*	1	

## Appendix

WP_168587632.1	<i>Saccharopolyspora</i> sp. ASAGF58	MVYDH*	1	Hydroxylase, PGM_ligase, Acetyltransferase
GGI73495.1	<i>Saccharopolyspora subtropica</i>	MSYIH*	1	
WP_121006811.1	<i>Saccharothrix australiensis</i>	MRYLH*	1	
WP_015100824.1	<i>Saccharothrix espanaensis</i> DSM 44229	MDYPH*	1	Methyltransferase, Protease/Peptidase
GAA3463537.1	<i>Saccharothrix longispora</i>	MDYPH*	1	
WP_219441052.1	<i>Saccharothrix obliqua</i>	MDYPH*	1	Methyltransferase, Protease/Peptidase
WP_201435482.1	<i>Saccharothrix</i> sp. 6-C	MYYVH*	1	Oxidoreductase, Protease/Peptidase
WP_143531888.1	<i>Saccharothrix</i> sp. ALI-22-I	MFYDH*	1	Acetyltransferase, PGM_ligase
WP_073897089.1	<i>Saccharothrix</i> sp. CB00851	MRHLH*	1	Protease/Peptidase
WP_053716375.1	<i>Saccharothrix</i> sp. NRRL B-16348	MYYLH*	1	Oxidoreductase, Acetyltransferase, Protease/Peptidase
WP_033431254.1	<i>Saccharothrix syringae</i>	MDYPH*	1	Protease/Peptidase
WP_211348027.1	<i>Saccharothrix texasensis</i>	MYYVH*	1	Oxidoreductase, Protease/Peptidase
RKT69193.1	<i>Saccharothrix variisporea</i>	MRYVH*	1	Hydroxylase, Methyltransferase
WP_061608438.1	<i>Sorangium cellulosum</i>	MKYCW*	1	Hydroxylase, Oxidoreductase
WP_204011461.1	<i>Sphaerimonospora thailandensis</i>	MRYLH*	1	Hydroxylase, Methyltransferase
WP_176497114.1	<i>Sphingomonas</i> sp. HMP6	MVHPH*	1	Oxidoreductase
WUU87918.1	<i>Spirillospora</i> sp. NBC_01491	MRYWY*	1	Methyltransferase

## Appendix

WP_147135745.1	<i>Stackebrandtia albiflava</i>	MRYWY*	1	Hydroxylase, Methyltransferase, Acetyltransferase, Sulfotransferase
WP_013017934.1	<i>Stackebrandtia nassauensis</i> DSM 44728	MRYWY*	1	Hydroxylase, Methyltransferase, Glycosyltransferase, Sulfotransferase
WP_196194857.1	<i>Streptacidiphilus fuscans</i>	MRYVH*	1	Methyltransferase, Protease/Peptidase
WP_052809996.1	<i>Streptomonospora alba</i>	MRYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase, additional_P450
WTQ78865.1	<i>Streptomyces achromogenes</i>	MSYDH*	1	Hydroxylase, PGM_ligase
WWM30798.1	<i>Streptomyces acidiscabies</i>	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
GHC29544.1	<i>Streptomyces albogriseolus</i>	MRHIH*	1	
TGG76237.1	<i>Streptomyces albus</i>	MRYVH*	1	
WP_055499314.1	<i>Streptomyces albus</i>	MRYVH*	1	Protease/Peptidase
WP_164439288.1	<i>Streptomyces albus</i>	MRYVH*	1	Protease/Peptidase
WP_143648821.1	<i>Streptomyces alkaliterrae</i>	MSYDH*	1	Hydroxylase, PGM_ligase
BBA20962.1	<i>Streptomyces atratus</i>	MRYLH*	1	Aminotransferase, NO_synthase, additional_P450
AXE76435.1	<i>Streptomyces atratus</i>	MRYLH*	1	Aminotransferase, NO_synthase, additional_P450
WP_159041763.1	<i>Streptomyces aureus</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_107082903.1	<i>Streptomyces avermitilis</i>	MRYVH*	1	Hydroxylase, Methyltransferase, Acetyltransferase

## Appendix

BAU77552.1	<i>Streptomyces avermitilis</i> MA-4680	MRYVH*	1	Hydroxylase, Methyltransferase
WP_061920834.1	<i>Streptomyces bungoensis</i>	MRYAY*	1	Hydroxylase, Protease/Peptidase
WP_030819383.1	<i>Streptomyces caelestis</i>	MSYDH*	1	Hydroxylase, PGM_ligase
WP_191207733.1	<i>Streptomyces chumphonensis</i>	MRYLH*	2	Hydroxylase, Methyltransferase
AXU14123.1	<i>Streptomyces clavuligerus</i>	MRHEY*	1	Methyltransferase
GAA4934895.1	<i>Streptomyces coeruleoprunus</i>	MKYWH*	1	
WP_114025389.1	<i>Streptomyces diacarni</i>	MRYVH*	1	Protease/Peptidase
WP_191867971.1	<i>Streptomyces diastaticus</i>	MRYLH*	2	Hydroxylase, Methyltransferase
WP_190056927.1	<i>Streptomyces echinoruber</i>	MSYDH*	1	Hydroxylase, PGM_ligase
WP_024761848.1	<i>Streptomyces exfoliatus</i> DSM 41693	MKYWH*	1	Protease/Peptidase
WP_190042383.1	<i>Streptomyces filamentosus</i>	MSYDH*	1	Hydroxylase, PGM_ligase
WP_051819776.1	<i>Streptomyces flavochromogenes</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
QPK49168.1	<i>Streptomyces gardneri</i>	MKYWH*	1	Protease/Peptidase
WP_167534015.1	<i>Streptomyces gardneri</i>	MKYWH*	1	Protease/Peptidase
ARF53562.1	<i>Streptomyces gilvosporeus</i>	MRYIH*	1	Hydroxylase, Methyltransferase, Protease/Peptidase
WP_190144208.1	<i>Streptomyces glebosus</i>	MRYAY*	1	
WP_158718955.1	<i>Streptomyces globisporus</i>	MKYWH*	1	Methyltransferase, additional_P450

## Appendix

GGU84509.1	<i>Streptomyces gougerotii</i>	MRYLH*	2	
WP_086758420.1	<i>Streptomyces griseiscabiei</i>	MRYLH*	2	Hydroxylase, Methyltransferase
WP_051865888.1	<i>Streptomyces griseus</i>	MRHAH*	2	Aminotransferase, Oxidoreductase, Methyltransferase
TQE28776.1	<i>Streptomyces ipomoeae</i>	MRYLH*	2	
TQE17623.1	<i>Streptomyces ipomoeae</i>	MRYLH*	2	
WP_153526325.1	<i>Streptomyces jumonjinensis</i>	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_189599827.1	<i>Streptomyces lateritius</i>	MKYWH*	1	Methyltransferase, Acetyltransferase
WP_158071316.1	<i>Streptomyces lateritius</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_190158429.1	<i>Streptomyces litmocidini</i>	MKYWH*	1	Methyltransferase, additional_P450
GAA2712158.1	<i>Streptomyces luteosporus</i>	MRYLH*	1	
WP_042158038.1	<i>Streptomyces lydicamycinicus</i>	MRYAY*	1	
WP_069570164.1	<i>Streptomyces lydicus</i>	MRHRW*	1	
WP_159674842.1	<i>Streptomyces mexicanus</i>	MGYDH*	1	Hydroxylase, PGM_ligase
WUS51385.1	<i>Streptomyces mirabilis</i>	MSYDH*	1	
WP_189512238.1	<i>Streptomyces narbonensis</i>	MKYWH*	1	Protease/Peptidase
WP_181924436.1	<i>Streptomyces nymphaeiformis</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_209238816.1	<i>Streptomyces oryzae</i>	MRYFY*	1	

## Appendix

WP_135337917.1	<i>Streptomyces palmae</i>	MRYLH*	3	Protease/Peptidase
OSY48349.1	<i>Streptomyces platensis</i>	MRYAY*	1	
WUK00188.1	<i>Streptomyces platensis</i>	MRYAY*	1	
QEV56021.1	<i>Streptomyces platensis</i>	MRYAY*	1	
WP_074999637.1	<i>Streptomyces qinglanensis</i>	MRYIH*	1	Oxidoreductase, Protease/Peptidase
WP_114013457.1	<i>Streptomyces reniochalinae</i>	MRYIH*	1	Methyltransferase, Protease/Peptidase
WP_048478122.1	<i>Streptomyces roseus</i>	MSYDH*	1	Hydroxylase, PGM_ligase
WP_193453809.1	<i>Streptomyces rutgersensis</i>	MRYLH*	2	Hydroxylase, Methyltransferase
WP_198550307.1	<i>Streptomyces silvensis</i>	MSYDH*	1	Hydroxylase, PGM_ligase
WP_129248903.1	<i>Streptomyces sioyaensis</i>	MRYAY*	1	
WP_203525348.1	<i>Streptomyces sioyaensis</i>	MRYAY*	1	
WP_143627967.1	<i>Streptomyces</i> sp. 1-11	MRYAY*	1	Hydroxylase, Protease/Peptidase
WP_167744107.1	<i>Streptomyces</i> sp. 2BBP-J2	MRHIH*	1	Protease/Peptidase
RKT08294.1	<i>Streptomyces</i> sp. 3211.6	MSYDH*	1	
WP_216724382.1	<i>Streptomyces</i> sp. A108	MRYLH*	1	Aminotransferase, NO_synthase, additional_P450
WP_136214634.1	<i>Streptomyces</i> sp. A1136	MSYDH*	1	Hydroxylase, PGM_ligase
WP_019357243.1	<i>Streptomyces</i> sp. AA1529	MRYIH*	1	Protease/Peptidase

## Appendix

WP_111332071.1	<i>Streptomyces</i> sp. AC1-42T	MSYDH*	1	Hydroxylase, Glycosyltransferase, PGM_ligase
WP_217245349.1	<i>Streptomyces</i> sp. AC602_WCS936	MRHAH*	1	Protease/Peptidase
WP_148006904.1	<i>Streptomyces</i> sp. adm13	MKYWH*	1	Protease/Peptidase, additional_P450
RPF29694.1	<i>Streptomyces</i> sp. Ag109_G2-6	MSYDH*	1	
WP_221907285.1	<i>Streptomyces</i> sp. BHT-5-2	MRHVH*	1	
WP_143204839.1	<i>Streptomyces</i> sp. CB02009	MKYWH*	1	
WP_187621367.1	<i>Streptomyces</i> sp. CB02980	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
WP_188273567.1	<i>Streptomyces</i> sp. CBMA152	MRHVH*	2	
WP_058928901.1	<i>Streptomyces</i> sp. CdTB01	MTYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_047017281.1	<i>Streptomyces</i> sp. CNQ-509	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
SCD88565.1	<i>Streptomyces</i> sp. di50b	MRHAH*	1	
PLW72023.1	<i>Streptomyces</i> sp. DJ	MSYDH*	1	
QIS75529.1	<i>Streptomyces</i> sp. DSM 40868	MSYDH*	1	Methyltransferase, PGM_ligase
WP_121715848.1	<i>Streptomyces</i> sp. E5N91	MRHAH*	1	
WP_227728090.1	<i>Streptomyces</i> sp. ET3-23	MTYYY*	1	
WP_093764026.1	<i>Streptomyces</i> sp. F-7	MRHIH*	1	Protease/Peptidase
WP_026246867.1	<i>Streptomyces</i> sp. HmicA12	MTYDH*	1	PGM_ligase

## Appendix

TXC99749.1	<i>Streptomyces</i> sp. ISID311	MRHAH*	1	
WP_215071590.1	<i>Streptomyces</i> sp. ISL-36	MKYWH*	1	additional_P450
WP_215017310.1	<i>Streptomyces</i> sp. ISL-94	MRYDHCL*	1	Methyltransferase, PGM_ligase
WP_093682451.1	<i>Streptomyces</i> sp. LamerLS-31b	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_158685075.1	<i>Streptomyces</i> sp. LaPpAH-108	MRYAY*	1	Oxidoreductase
WP_081222223.1	<i>Streptomyces</i> sp. M41(2017)	MTYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_200723493.1	<i>Streptomyces</i> sp. MBT49	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase, Acetyltransferase
WP_172388015.1	<i>Streptomyces</i> sp. MNP-20	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WTR14400.1	<i>Streptomyces</i> sp. NBC_00138	MSYDH*	1	Hydroxylase, PGM_ligase
WTR07800.1	<i>Streptomyces</i> sp. NBC_00144	MRHAH*	1	Methyltransferase
WP_328941248.1	<i>Streptomyces</i> sp. NBC_00250	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_328330918.1	<i>Streptomyces</i> sp. NBC_00455	MRHAH*	1	Methyltransferase
WUG07218.1	<i>Streptomyces</i> sp. NBC_00466	MRHAH*	1	Methyltransferase
WUD46218.1	<i>Streptomyces</i> sp. NBC_00513	MSYDH*	1	Hydroxylase, PGM_ligase
WTX68379.1	<i>Streptomyces</i> sp. NBC_00647	MTYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_329285285.1	<i>Streptomyces</i> sp. NBC_00691	MKYWH*	1	Methyltransferase, Protease/Peptidase
WTJ09092.1	<i>Streptomyces</i> sp. NBC_00715	MTYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase



## Appendix

WSX67477.1	<i>Streptomyces</i> sp. NBC_00932	MRHAH*	1	Methyltransferase
WSV42528.1	<i>Streptomyces</i> sp. NBC_01077	MKYWH*	1	Methyltransferase, Protease/Peptidase
WSU39749.1	<i>Streptomyces</i> sp. NBC_01089	MSYDH*	1	
WSS51411.1	<i>Streptomyces</i> sp. NBC_01180	MRHAH*	1	Methyltransferase
WP_327393381.1	<i>Streptomyces</i> sp. NBC_01186	MRYIH*	1	Hydroxylase, Oxidoreductase, Protease/Peptidase
WSS45373.1	<i>Streptomyces</i> sp. NBC_01187	MRYIH*	1	Hydroxylase, Oxidoreductase, Protease/Peptidase
WP_331720860.1	<i>Streptomyces</i> sp. NBC_01241	MSYDH*	1	Hydroxylase, PGM_ligase
WP_329619874.1	<i>Streptomyces</i> sp. NBC_01255	MKYWH*	1	Methyltransferase, Protease/Peptidase
WTY58779.1	<i>Streptomyces</i> sp. NBC_01411	MRHAH*	1	Methyltransferase
WSD37542.1	<i>Streptomyces</i> sp. NBC_01750	MRYVH*	1	Hydroxylase, Methyltransferase
WP_326806865.1	<i>Streptomyces</i> sp. NBC_01775	MRYIH*	1	Hydroxylase, Oxidoreductase, Protease/Peptidase
WSB05146.1	<i>Streptomyces</i> sp. NBC_01794	MRYVH*	1	Hydroxylase, Methyltransferase
WP_326686421.1	<i>Streptomyces</i> sp. NBC_01795	MRYIH*	1	Hydroxylase, Oxidoreductase, Protease/Peptidase
WP_225799993.1	<i>Streptomyces</i> sp. NK15101	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
WP_209497897.1	<i>Streptomyces</i> sp. NPDC090301	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_143664724.1	<i>Streptomyces</i> sp. NRRL B-24572	MKYWH*	1	Methyltransferase

## Appendix

WP_030800649.1	<i>Streptomyces</i> sp. NRRL S-337	MSYDH*	1	Oxidoreductase
WP_185094699.1	<i>Streptomyces</i> sp. PanSC19	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
WP_158992425.1	<i>Streptomyces</i> sp. QHH-9511	MKYWH*	1	Methyltransferase, Acetyltransferase
WP_056562044.1	<i>Streptomyces</i> sp. Root66D1	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
WP_168482770.1	<i>Streptomyces</i> sp. RPA4-5	MKYWH*	1	
WP_164657610.1	<i>Streptomyces</i> sp. S1A1-7	MKYWH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_164405930.1	<i>Streptomyces</i> sp. S1D4-23	MKYWH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_329853522.1	<i>Streptomyces</i> sp. SP18ES09	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_162689260.1	<i>Streptomyces</i> sp. ST1020	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
WP_055718270.1	<i>Streptomyces</i> sp. ST1020	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
PZT70810.1	<i>Streptomyces</i> sp. SW4	MRHAH*	1	
WP_143060225.1	<i>Streptomyces</i> sp. TLI_105	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
WP_129279950.1	<i>Streptomyces</i> sp. TM32	MRYAY*	1	
WP_187821553.1	<i>Streptomyces</i> sp. TRM68367	MTYDH*	1	Hydroxylase, Oxidoreductase, PGM_ligase
WP_125523660.1	<i>Streptomyces</i> sp. WAC 05379	MSYDH*	1	
WP_125741628.1	<i>Streptomyces</i> sp. WAC01280	MKYWH*	1	Protease/Peptidase

## Appendix

WP_143098833.1	<i>Streptomyces</i> sp. yr375	MSYDH*	1	Hydroxylase, Methyltransferase, PGM_ligase
OEJ36126.1	<i>Streptomyces subutilus</i>	MSYDH*	1	
WP_079154530.1	<i>Streptomyces subutilus</i>	MSYDH*	1	Hydroxylase, PGM_ligase
WP_189804110.1	<i>Streptomyces tanashiensis</i>	MKYWH*	1	
WP_190106205.1	<i>Streptomyces tanashiensis</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
QES18500.1	<i>Streptomyces venezuelae</i>	MKYWH*	1	Protease/Peptidase
WP_150162815.1	<i>Streptomyces venezuelae</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
QES09683.1	<i>Streptomyces venezuelae</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_145953730.1	<i>Streptomyces venezuelae</i> ATCC 10712	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
WP_015037861.1	<i>Streptomyces venezuelae</i> ATCC 10712	MKYWH*	1	Methyltransferase, Protease/Peptidase, additional_P450
WP_175439194.1	<i>Streptomyces vilmorinianum</i>	MKYWH*	1	Methyltransferase
WP_188342106.1	<i>Streptomyces xanthii</i>	MKYWH*	1	Hydroxylase, PGM_ligase
WP_328489878.1	<i>Streptomyces zaomyceticus</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
WP_327165416.1	<i>Streptomyces zaomyceticus</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
WTR68839.1	<i>Streptomyces zaomyceticus</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase
WTT42013.1	<i>Streptomyces zaomyceticus</i>	MKYWH*	1	Methyltransferase, Protease/Peptidase

---

## Appendix

---

WP_157995908.1	<i>Thermomonospora amylolytica</i>	MRYIH*	1	Hydroxylase, Glycosyltransferase, Protease/Peptidase, Sulfotransferase
WP_182708337.1	<i>Thermomonospora cellulosilytica</i>	MRYIH*	1	Hydroxylase, Glycosyltransferase, Protease/Peptidase, Sulfotransferase

---

## 8 Acknowledgements

Auf dem Weg meiner Promotion wurde ich von vielen Menschen begleitet und geprägt, denen ich an dieser Stelle meinen herzlichen Dank aussprechen möchte.

An erster Stelle möchte ich mich bei meinem Doktorvater Prof. Dr. Max Crüsemann für die wertvollen fachlichen Anregungen und den Freiraum bei der Gestaltung meiner Forschung bedanken. Ebenso danke ich meinem Zweitbetreuer Prof. Dr. Martin Baunach für die Übernahme des Zweitgutachtens und seine Diskussionsfreude. Prof. Dr. Finn Hansen und Prof. Dr. Gabriele Bierbaum danke ich für ihre Mitarbeit in meiner Prüfungskommission.

Der Studienstiftung des deutschen Volkes danke ich sehr für die finanzielle und vor allem ideelle Förderung während meiner Promotion. Der regelmäßige Austausch mit den inspirierenden Stipendiatinnen und Stipendiaten haben meinen Horizont erweitert.

Im täglichen wissenschaftlichen Umfeld waren es die Kolleginnen und Kollegen am Institut für Pharmazeutische Biologie, denen ich für die positive Atmosphäre dankbar bin. Insbesondere möchte ich mich bei Dr. Sophie Klöppel für die vielen lustigen Momente im Büro und bei Lukas Zimmer für die äußerst produktive und intellektuell anregende Zusammenarbeit bedanken. Den Studierenden, die ich betreuen durfte, danke ich für die wertvollen Erfahrungen, wo ich nicht zuletzt auch selber viel gelernt habe.

Ich bin sehr dankbar für die professionelle und ertragreiche Zusammenarbeit mit Prof. Dr. Max Cryle, der mich für einen Forschungsaufenthalt in seiner Arbeitsgruppe in Melbourne aufnahm. Dr. Yongwei Zhao danke ich für die engagierte Betreuung in Australien.

Die Menschen in meinem persönlichen Umfeld haben mir den notwendigen Rückhalt für das Gelingen dieser Arbeit gegeben. Mein besonderer Dank gilt meiner Familie, meinen Freunden und allen, die mir nahe stehen.

Zuletzt möchte ich Vivi meinen unermesslichen Dank aussprechen. Durch ihre bedingungslose Liebe und emotionale Unterstützung musste ich nie alleine durch schwierige Zeiten gehen und konnte jeden Moment der Freude teilen.