

**Molecular Basis for Selective Late-Stage Transformations in Fungal  
Indole Alkaloid Biosynthesis**

by

Amy E. Fraley

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Medicinal Chemistry)  
in the University of Michigan  
2019

Doctoral Committee:

Professor David H. Sherman, Chair  
Professor Janet L. Smith, Chair  
Professor Anna K. Mapp  
Professor Scott D. Larsen  
Professor Emily E. Scott

Amy E. Fraley

[ae fraley@umich.edu](mailto:ae fraley@umich.edu)

ORCID ID: 0000-0002-4950-0538

© Amy E. Fraley 2019

## Acknowledgements

I would first like to thank my co-advisors Dr. David Sherman and Dr. Janet Smith who have both guided and supported me as I developed as a scientist. The experiences that I had in their labs will forever shape the way that I think about science, and their encouragement has helped me to reach my highest potential in graduate school, even when experimental difficulties seemed insurmountable. I would also like to thank Dr. Cherie Dotson and Dr. Ronald Woodard for providing me with the opportunity to participate in the Interdisciplinary Research Experience for Undergraduates (IREU), as well as David for accepting me as an undergraduate researcher in his lab. This opportunity changed the trajectory of my career and allowed me to see my potential as a scientific researcher. Prior to the IREU program, I had no formal research experience, and I am sincerely indebted to Cherie, Ron, and David for providing me with this opportunity.

I have also been very fortunate to work with wonderful lab mates who have made this experience incredibly enriching. The past and present members of both the Sherman and Smith groups have always provided great scientific and professional insight. In particular, I would like to acknowledge Dr. Ashootosh Tripathi, Dr. Jennifer Schmidt, Dr. Dheeraj Khare, and Dr. Qingyun Dan for all of their help. All of the members of both labs have provided a wonderful and comfortable learning environment, and I am particularly thankful for those who have also become good friends. Additionally, a few students have worked with me throughout graduate school, and I am grateful for their dedication to science and their patience as I fine-tuned my teaching skills.

Throughout graduate school, I have had the opportunity to collaborate with many fantastic scientists outside of UM. I would like to acknowledge Dr. Marc Garcia-Borràs and Dr. Ken Houk whose work was key in identifying the mechanism of the MalA halogenase. I would also like to thank all of the team members who I worked with at Novartis Institutes for BioMedical Research including Dr. Hasnain Malik, Dr. Artiom Cernijenko, and David Dunstan. Additionally, I would like to acknowledge Dr. Tyler Beyett, with whom I collaborated on the LSI Cubed project.

I greatly appreciate the guidance from my committee members: Dr. Anna Mapp, Dr. Emily Scott, and Dr. Scott Larsen. I would also like to thank Debbie Lounds and Pam Schultz for keeping the lab running efficiently. My thesis work could not have been accomplished without

funding from the following sources: Richard Thomas Garrett and Marilyn Holtom Garrett Research Fellowship, Department of Education GAANN Fellowship (P200A150164), Life Sciences Institute Cubed Research Grant, Rackham Predoctoral Fellowship, and the NSF Center for Selective C–H Functionalization (CHE-1700982).

Lastly, I want to thank my family for their love, support, and encouragement. Their strength and determination have inspired me to make the most of every opportunity.

## Table of Contents

<b>Acknowledgements</b> .....	<b>ii</b>
<b>List of Figures</b> .....	<b>xii</b>
<b>List of Tables</b> .....	<b>xxii</b>
<b>List of Schemes</b> .....	<b>xxiv</b>
<b>List of Abbreviations</b> .....	<b>xxv</b>
<b>Abstract</b> .....	<b>xxvii</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
<b>1.1 Introduction to Natural Products</b> .....	<b>1</b>
<b>1.2 Evolution of Natural Product Biosynthesis in the Bicyclo[2.2.2]diazaoctane</b>	
<b>Containing Fungal Indole Alkaloids</b> .....	<b>2</b>
1.2.1 Introduction to Fungal Indole Alkaloids .....	2
1.2.2 Distinct Nonribosomal Peptide Synthetases .....	3
1.2.3 Redundant Prenyltransferases .....	5
1.2.4 Enantioselective Diels-Alderases .....	7
1.2.4.1 Isolation and Precursor Incorporation Studies .....	8
1.2.4.2 Enzymology .....	10
1.2.5 Flavin Monooxygenases (FMOs) .....	12
1.2.5.1 Paraherquamides .....	12
1.2.5.2 Notoamides .....	13

1.2.5.3 Taichunamides	16
1.2.5.4 Brevianamides	17
1.2.5.5 Effect on the Diels-Alder Products	17
1.2.5.6 FMO Evolution	19
1.2.5.7 Mechanism of Spirooxindole Formation	20
1.2.6 Dioxepin and Pyran Ring Formation	22
1.2.6.1 Mechanism of Pyran and Dioxepin Ring Formation	23
1.2.7 Late-Stage Halogenation	23
1.2.8 Conclusion	25
<b>1.3 Thesis Overview</b>	<b>26</b>
<b>Chapter 2: Characterization of the Bifunctional Reductase/Diels-Alderase Responsible for the Formation of the Bicyclo[2.2.2]diazaoctane Ring</b>	<b>28</b>
<b>2.1 Introduction</b>	<b>28</b>
<b>2.2 In Vitro Reconstitution of the Malbrancheamide Biosynthetic Pathway</b>	<b>29</b>
2.2.1 Nonribosomal Peptide Synthetase	29
2.2.2 Prenyltransferase	31
2.2.3 IMDA Enzyme Biochemistry	32
2.2.4 IMDA Enzyme Structural Analysis and Mutagenesis	34
2.2.5 Conclusion	36
<b>Chapter 3: Biocatalytic Late-Stage Halogenation</b>	<b>37</b>
<b>3.1 Introduction to Halogenation Biocatalysts and Enzyme Engineering</b>	<b>37</b>
3.1.1 Introduction to Halogenases	37
3.1.2 Bioactive Halogenated Natural Products and Halogenating Enzymes	38

3.1.3 Halogenation and Drug Design .....	39
3.1.4 Halogenase Engineering .....	40
3.1.4.1 Enzyme Stability and Optimal Activity .....	41
3.1.4.2 Structure-based Halogenase Engineering .....	44
3.1.4.3 Directed Evolution for Engineered Halogenases .....	46
3.1.5 Fluorination Biocatalysts .....	48
3.1.6 Summary .....	50
<b>3.2 Characterization of MalA as the Malbrancheamide Late-Stage Halogenase .....</b>	<b>51</b>
3.2.1 Introduction to Malbrancheamide and MalA Halogenase .....	51
3.2.2 Biochemical Characterization of MalA .....	53
3.2.2.1 Isolation of Malbrancheamide from <i>M. aurantiaca</i> .....	53
3.2.2.2 Biochemical Activity of MalA .....	54
3.2.2.3 Kinetic Characterization of MalA .....	55
3.2.3 Structural Characterization of the Substrate Complexes of MalA' .....	56
3.2.4 Molecular Dynamics Simulations and Quantum Mechanics Models .....	58
3.2.5 Discussion .....	62
<b>3.3 Structure-Based Engineering of MalA for Selective Late-Stage Halogenation .....</b>	<b>65</b>
<b>3.4 Analysis of Substrate Scope of MalA and Variants .....</b>	<b>66</b>
<b>Chapter 4: Molecular Basis for Spirooxindole Formation in the Paraherquamide</b>	
<b>Biosynthetic Pathway .....</b>	<b>70</b>
<b>4.1 Introduction to the Paraherquamides and Spirooxindole Formation .....</b>	<b>70</b>
<b>4.2 Biochemical Analysis of PhqK .....</b>	<b>73</b>
<b>4.3 Discussion .....</b>	<b>78</b>

<b>Chapter 5: Perturbation of the Interactions of Calmodulin with GRK5 Using the</b>	
<b>Natural Product Malbrancheamide</b> .....	<b>80</b>
<b>5.1 Introduction</b> .....	<b>80</b>
5.1.1 Introduction to Malbrancheamide as a Chemical Probe .....	80
5.1.2 Introduction to the Ca <sup>2+</sup> ·CaM–GRK5 Complex .....	82
<b>5.2 Biochemical Characterization of Ca<sup>2+</sup>·CaM–GRK5</b> .....	<b>83</b>
5.2.1 A Single Ca <sup>2+</sup> ·CaM Binds the RH-Kinase Domain Junction of GRK5 .....	83
5.2.2 Ca <sup>2+</sup> ·CaM Activates GRK5 and Promotes Autophosphorylation .....	85
5.2.3 Malbrancheamide Binds the C-lobe of Calmodulin and Inhibits	
Ca <sup>2+</sup> ·CaM-Stimulated GRK5 Autophosphorylation .....	87
5.2.4 The N-lobe of Ca <sup>2+</sup> ·CaM Interacts With the N-Terminus of GRK5 to	
Inhibit Receptor Phosphorylation .....	88
5.2.5 Malbrancheamide Inhibits Nuclear Translocation of GRK5 and	
Prevents Hypertrophy .....	89
<b>5.3 Discussion</b> .....	<b>91</b>
<b>Chapter 6: Summary, Discussion, and Future Directions</b> .....	<b>95</b>
<b>6.1 Characterization of the Bifunctional Reductase/Diels-Alderase Responsible</b>	
<b>for the Formation of the Bicyclo[2.2.2]diazaoctane Ring</b> .....	<b>95</b>
6.1.1 Summary and Discussion of Work Completed .....	95
6.1.2 Future Directions .....	95
<b>6.2 Late-Stage Halogenation by Mala</b> .....	<b>96</b>
6.2.1 Summary and Discussion of Work Completed .....	96
6.2.2 Future Directions .....	97



<b>6.3 Molecular Basis for Spirooxindole Formation in the Paraherquamide</b>	
<i>Biosynthetic Pathway</i> .....	<b>98</b>
6.3.1 Summary and Discussion of Work Completed .....	98
6.3.2 Future Directions .....	98
<b>6.4 Perturbation of the Interactions of Calmodulin with GRK5 Using the Natural</b>	
<i>Product Malbrancheamide</i> .....	<b>99</b>
6.4.1 Summary and Discussion of Work Completed .....	99
6.4.2 Future Directions .....	100
<b>Chapter 7: Late-Stage Halogenation by MalA Experimental Section</b> .....	<b>101</b>
<b>7.1 Experimental Methods</b> .....	<b>101</b>
7.1.1 <i>M. graminicola</i> genomic DNA Extraction and Sequencing .....	101
7.1.2 <i>M. aurantiaca</i> cDNA Preparations .....	101
7.1.3 Cloning of <i>mala-pMCSG7</i> .....	101
7.1.4 <i>M. aurantiaca</i> Growth and Extraction of Malbrancheamides .....	102
7.1.5 Expression and Purification .....	103
7.1.5.1 Expression of <i>mala</i> , <i>mala'</i> , and <i>mala/A'</i> Mutants .....	103
7.1.5.2 Expression of <i>mala'</i> for Selenomethionyl MalA' .....	103
7.1.5.3 Protein Purification for Chlorination Assays and	
Large-Scale Reactions .....	103
7.1.5.4 Protein Purification for Bromination Assays and	
Large-Scale Reactions .....	104
7.1.5.5 Protein Purification for Crystallography .....	104
7.1.6 MalA Biochemical Activity Assays .....	105

7.1.7 Cocrystallization of MalA'	105
7.1.7.1 MalA' Crystallization Conditions	105
7.1.7.2 Data Collection of MalA' Crystals	106
7.1.8 Site-Directed Mutagenesis	106
7.1.8.1 MalA' (MalA L276P/R428P)	107
7.1.8.2 MalA K108A	107
7.1.8.3 All Other MalA Mutants	108
7.1.9 MalA Large-Scale Reactions and Isolation of Products	109
7.1.9.1 Chlorination Reaction Conditions and Extraction	109
7.1.9.2 Bromination Reaction Conditions and Extraction	109
7.1.9.3 HPLC Purification	109
7.1.10 Michaelis-Menten Model Kinetics	110
7.1.10.1 Substrates Malbrancheamide B and Isomalbrancheamide B to Product Malbrancheamide	110
7.1.10.2 Substrates Premalbrancheamide to Products Isomalbrancheamide B and Malbrancheamide B	110
7.1.11 Density Functional Theory Calculations	111
7.1.11 Molecular Dynamics Simulations	111
<b>7.1 Tables</b>	<b>113</b>
<b>7.3 Figures</b>	<b>121</b>
<b>7.4 Compound Characterization</b>	<b>130</b>
7.4.1 Mass Spectrometry Data	130
7.4.2 NMR Spectrometry Data	133

## Chapter 8: Molecular Basis for Spirooxindole Formation in the Paraherquamide

<b>Biosynthetic Pathway Experimental Sections</b> .....	<b>153</b>
<b>8.1 Experimental Methods</b> .....	<b>153</b>
8.1.1 Fungal Strains and Culture Conditions.....	153
8.1.2 cDNA Preparation and Cloning of <i>phqK</i> .....	153
8.1.3 Overexpression and Purification for Enzymatic Reactions.....	153
8.1.4 Purification of PhqK for Crystallization and Kinetics.....	154
8.1.5 Crystallization of PhqK.....	154
8.1.6 Data Collection.....	155
8.1.7 Enzymatic Reactions and HPLC Analysis of PhqK Reactions.....	155
8.1.8 Enzymatic Reactions and HPLC Analysis of PhqK Kinetics.....	155
<b>8.2 Tables</b> .....	<b>157</b>
<b>8.3 Figures</b> .....	<b>165</b>
<b>8.4 Compound Characterization</b> .....	<b>174</b>
8.4.1 NMR Characterization.....	174

## Chapter 9: Perturbation of the Interactions of Calmodulin with GRK5 Using the

<b>Natural Product Malbrancheamide Experimental Section</b> .....	<b>201</b>
<b>9.1 Experimental Methods</b> .....	<b>201</b>
9.1.1 Isolation of Malbrancheamide.....	201
9.1.2 Production and Purification of Isomalbrancheamide D.....	201
9.1.3 Protein Expression and Purification.....	202
9.1.4 Determination of the Ca <sup>2+</sup> ·CaM·Malbrancheamide Crystal Structures.....	203
9.1.5 Isothermal Titration Calorimetry (ITC).....	204

9.1.6 Radiometric Kinase Assays	204
9.1.7 Complex Formation and Light Scattering	205
9.1.8 Small-Angle X-Ray Scattering (SAXS)	205
9.1.9 Negative Stain EM	206
9.1.10 Flow Cytometry Protein Interaction Assay (FCPIA)	206
9.1.11 GRK5 Nuclear Translocation Assay	207
9.1.12 GRK5 Hypertrophic Phenotypic Assay	207
<b>9.2 Tables</b>	<b>208</b>
<b>9.3 Figures</b>	<b>211</b>
<b>References</b>	<b>218</b>

## List of Figures

<b>Figure 1.1</b> Select fungal indole alkaloids containing the bicyclo[2.2.2]diazaoctane ring	2
<b>Figure 1.2</b> PhqB R-domain structure	4
<b>Figure 1.3</b> Proposed mechanism for the generation of the methyl proline-containing molecules	5
<b>Figure 1.4</b> Comparison of MalE and MalB	7
<b>Figure 1.5</b> Metabolites isolated from <i>Aspergillus amoenus</i> and <i>Aspergillus protuberus</i>	8
<b>Figure 1.6</b> Hydroxylated notoamide M and halogenated notoamide N	9
<b>Figure 1.7</b> Brevianamide formation	10
<b>Figure 1.8</b> Reactions performed by MalC	11
<b>Figure 1.9</b> Representative spirocyclized fungal indole alkaloids	12
<b>Figure 1.10</b> Secondary metabolites with <i>anti</i> and <i>syn</i> configuration	15
<b>Figure 1.11</b> Taichunamides with <i>anti</i> -relative stereochemistry of the bicyclic ring	17
<b>Figure 1.12</b> <i>Anti</i> bridged bicycles	17
<b>Figure 1.13</b> Brevianamide E, which is generated from deoxybrevianamide E	17
<b>Figure 1.14</b> Conformations conducive to <i>syn</i> or <i>anti</i> bicyclic ring formation	18
<b>Figure 1.15</b> Asperversiamides A-E	19
<b>Figure 1.16</b> Domains of PhqK FMO	20
<b>Figure 1.17</b> Spirotryprostatin and fumitremorgin biosynthetic pathway crosstalk	21
<b>Figure 1.18</b> Halogenated malbrancheamides	24
<b>Figure 1.19</b> Structural comparison of FMOs	25
<b>Figure 2.1</b> Scheme of <i>in vitro</i> reconstitution assays	29
<b>Figure 2.2</b> <i>In vitro</i> reaction of MalG	30

<b>Figure 2.3</b> <i>In vitro</i> reaction of MalG and MalE .....	30
<b>Figure 2.4</b> <i>In vitro</i> reaction of MalG and MalB .....	31
<b>Figure 2.5</b> Chiral separation of racemic premalbrancheamide .....	31
<b>Figure 2.6</b> <i>In vitro</i> reaction of MalG, MalE, and MalC .....	32
<b>Figure 2.7</b> Aerobic and anaerobic MalC-catalyzed reactions .....	33
<b>Figure 2.8</b> Production of (+)-malbrancheamide .....	34
<b>Figure 2.9</b> MalC mutant reaction data .....	35
<b>Figure 2.10</b> Proposed catalytic mechanism for MalC/PhqE .....	36
<b>Figure 3.1</b> Halogenase Engineering Strategies .....	37
<b>Figure 3.2</b> Examples of flavin-dependent halogenases .....	39
<b>Figure 3.3</b> Cl- $\pi$ interaction in MalA halogenase .....	40
<b>Figure 3.4</b> Flavin-dependent halogenase cofactor regeneration .....	42
<b>Figure 3.5</b> Incorporation of FDHs into plants .....	43
<b>Figure 3.6</b> Halogenase structural overlay .....	45
<b>Figure 3.7</b> Biosynthesis of fluorinated molecules .....	49
<b>Figure 3.8</b> Mechanism of fluoride binding .....	50
<b>Figure 3.9</b> Malbrancheamide biosynthetic gene clusters in <i>M. aurantiaca</i> and <i>M. graminicola</i> .....	52
<b>Figure 3.10</b> HPLC traces for MalA <i>in vitro</i> reactions .....	54
<b>Figure 3.11</b> Key gHMBCAD and gCOSY NMR correlations .....	55
<b>Figure 3.12</b> Active site view of MalA' .....	56
<b>Figure 3.13</b> Substrate interactions with Phe489 .....	57
<b>Figure 3.14</b> Reactions data for MalA mutants .....	57
<b>Figure 3.15</b> Structures of MalA' E494D and H253F .....	58

<b>Figure 3.16</b> Computational studies with MalA' complex .....	60
<b>Figure 3.17</b> Computational studies with MalA' complex .....	64
<b>Figure 3.18</b> Spectrum of MalA variants .....	65
<b>Figure 3.19</b> MalA' S129A/P85S cocrystal structure with <b>1.14</b> .....	66
<b>Figure 3.20</b> Workflow for halogenation biocatalysis .....	67
<b>Figure 4.1</b> Various enzyme classes that catalyze spirocycle formation .....	71
<b>Figure 4.2</b> Generation of spiromalbramide via <i>in vitro</i> PhqK reactions .....	74
<b>Figure 4.3</b> Expansion of PhqK active site pocket to accommodate substrates .....	74
<b>Figure 4.4</b> Modeled cofactor dynamics .....	75
<b>Figure 4.5</b> Overlay of PhqK cocrystal structures .....	76
<b>Figure 4.6</b> Essential amino acids in PhqK active site .....	77
<b>Figure 4.7</b> Mutants generated in PhqK .....	78
<b>Figure 4.8</b> Proposed mechanism for PhqK epoxidation and selective spirocyclization .....	78
<b>Figure 5.1</b> Mechanism of malbrancheamide in NO/cGMP-dependent vasorelaxation .....	81
<b>Figure 5.2</b> Biophysical characterization of the Ca <sup>2+</sup> ·CaM–GRK5 complex .....	84
<b>Figure 5.3</b> Modulation of substrate phosphorylation by Ca <sup>2+</sup> ·CaM .....	86
<b>Figure 5.4</b> Malbrancheamide ( <b>1.17</b> ) binds to the C-lobe of Ca <sup>2+</sup> ·CaM .....	87
<b>Figure 5.5</b> Effects of malbrancheamide ( <b>1.17</b> ) and GRK5 peptides on Ca <sup>2+</sup> ·CaM modulation of GRK5 activity .....	89
<b>Figure 5.6</b> Effects of malbrancheamide ( <b>1.17</b> ) on GRK5 nuclear translocation and cardiomyocyte hypertrophy .....	90
<b>Figure 5.7</b> Model for domain-specific regulation of GRK5 by Ca <sup>2+</sup> ·CaM .....	92
<b>Figure 5.8</b> Summary of malbrancheamide biological activity as a selective inhibitor of	

Ca <sup>2+</sup> ·CaM protein-protein interactions .....	94
<b>Figure 7.1</b> MalA kinetics .....	121
<b>Figure 7.2</b> Malbrancheamide numbering scheme .....	121
<b>Figure 7.3</b> MalA omit maps .....	122
<b>Figure 7.4</b> Zn <sup>2+</sup> -binding site of MalA' .....	122
<b>Figure 7.5</b> Chloride binding site in MalA' active site .....	122
<b>Figure 7.6</b> RMSF measured along 500 ns MD simulations for the apo and premalbrancheamide bound MalA' systems .....	123
<b>Figure 7.7</b> MD simulations with MalA complex .....	123
<b>Figure 7.8</b> Lysine chloramine conformations .....	124
<b>Figure 7.9</b> Interactions between Glu494 and substrates .....	124
<b>Figure 7.10</b> Interactions between chloramine and substrates .....	125
<b>Figure 7.11</b> Interactions between chloramine and substrates in H253F .....	125
<b>Figure 7.12</b> Interactions between chloramine and substrates in H253A .....	126
<b>Figure 7.13</b> Interactions between chloramine and substrates in D109A .....	126
<b>Figure 7.14</b> Cl-C-H angle and Cl-C distance for WT and H253A/F .....	127
<b>Figure 7.15</b> Solvation shells in WT and H253A/F .....	127
<b>Figure 7.16</b> pK <sub>a</sub> predictions for Lys108 and Glu494 in the apo state of WT MalA' .....	128
<b>Figure 7.17</b> DFT optimized reaction pathways .....	128
<b>Figure 7.18</b> Optimized structures for the three computational models .....	129
<b>Figure 7.19</b> Premalbrancheamide (fungal) mass spectrometry data .....	130
<b>Figure 7.20</b> Malbrancheamide B ( <i>in vitro</i> assay) mass spectrometry data .....	130
<b>Figure 7.21</b> Isomalbrancheamide B ( <i>in vitro</i> assay) mass spectrometry data .....	130



<b>Figure 7.22</b> Malbrancheamide ( <i>in vitro</i> assay) mass spectrometry data .....	131
<b>Figure 7.23</b> Malbrancheamide C ( <i>in vitro</i> assay) mass spectrometry data .....	131
<b>Figure 7.24</b> Isomalbrancheamide C ( <i>in vitro</i> assay) mass spectrometry data .....	131
<b>Figure 7.25</b> Malbrancheamide D ( <i>in vitro</i> assay) mass spectrometry data .....	132
<b>Figure 7.26</b> Isomalbrancheamide D ( <i>in vitro</i> assay) mass spectrometry data .....	132
<b>Figure 7.27</b> <sup>1</sup> H-NMR of premalbrancheamide isolated from <i>M. aurantiaca</i> (400 MHz, CD <sub>3</sub> OD) .....	133
<b>Figure 7.28</b> <sup>1</sup> H-NMR of malbrancheamide B isolated from <i>M. aurantiaca</i> (400 MHz, -(CD <sub>3</sub> ) <sub>2</sub> SO- <i>d</i> <sub>6</sub> ) .....	134
<b>Figure 7.29</b> <sup>1</sup> H-NMR of isomalbrancheamide B from <i>M. aurantiaca</i> (400 MHz, (CD <sub>3</sub> ) <sub>2</sub> SO- <i>d</i> <sub>6</sub> ) .....	135
<b>Figure 7.30</b> <sup>1</sup> H-NMR of malbrancheamide isolated from <i>M. aurantiaca</i> (400 MHz, CD <sub>3</sub> OD)	136
<b>Figure 7.31</b> <sup>1</sup> H-NMR of malbrancheamide B from MalA <i>in vitro</i> reaction (400 MHz, (CD <sub>3</sub> ) <sub>2</sub> SO- <i>d</i> <sub>6</sub> ) .....	137
<b>Figure 7.32</b> <sup>1</sup> H-NMR of isomalbrancheamide B from MalA <i>in vitro</i> reaction (400 MHz, (CD <sub>3</sub> ) <sub>2</sub> SO- <i>d</i> <sub>6</sub> ) .....	138
<b>Figure 7.33</b> <sup>1</sup> H-NMR of malbrancheamide from MalA <i>in vitro</i> reaction (400 MHz, CD <sub>3</sub> OD)	139
<b>Figure 7.34</b> <sup>1</sup> H-NMR of malbrancheamide C ( <b>5</b> ) from MalA <i>in vitro</i> reaction (700 MHz, CD <sub>3</sub> OD) .....	140
<b>Figure 7.35</b> <sup>13</sup> C-NMR of malbrancheamide C from MalA <i>in vitro</i> reaction (176 MHz, CD <sub>3</sub> OD) .....	141
<b>Figure 7.36</b> <sup>1</sup> H-NMR of isomalbrancheamide C from MalA <i>in vitro</i> reaction (700 MHz, CD <sub>3</sub> OD) .....	142

<b>Figure 7.37</b> $^{13}\text{C}$ -NMR of isomalbrancheamide C from MalA <i>in vitro</i> reaction (176 MHz, $\text{CD}_3\text{OD}$ )	143
<b>Figure 7.38</b> $^1\text{H}$ -NMR of malbrancheamide D from MalA <i>in vitro</i> reaction (700 MHz, $\text{CD}_3\text{OD}$ )	144
<b>Figure 7.39</b> gCOSY correlations of malbrancheamide D from <i>in vitro</i> reaction with MalA (700 MHz, $\text{CD}_3\text{OD}$ )	145
<b>Figure 7.40</b> gHMBCAD correlations of malbrancheamide D from <i>in vitro</i> reaction with MalA (700 MHz, $\text{CD}_3\text{OD}$ )	146
<b>Figure 7.41</b> gHSQCAD correlations of malbrancheamide D from <i>in vitro</i> reaction with MalA (700 MHz, $\text{CD}_3\text{OD}$ )	147
<b>Figure 7.42</b> $^1\text{H}$ -NMR of isomalbrancheamide D from <i>in vitro</i> reaction with MalA (700 MHz, $\text{CD}_3\text{OD}$ )	148
<b>Figure 7.43</b> $^{13}\text{C}$ -NMR of isomalbrancheamide D from <i>in vitro</i> reaction with MalA (176 MHz, $\text{CD}_3\text{OD}$ )	149
<b>Figure 7.44</b> gCOSY correlations of isomalbrancheamide D from <i>in vitro</i> reaction with MalA (700 MHz, $\text{CD}_3\text{OD}$ )	150
<b>Figure 7.45</b> gHMBCAD correlations of isomalbrancheamide D from <i>in vitro</i> reaction with MalA (700 MHz, $\text{CD}_3\text{OD}$ )	151
<b>Figure 7.46</b> gHSQCAD correlations of isomalbrancheamide D from <i>in vitro</i> reaction with MalA (700 MHz, $\text{CD}_3\text{OD}$ )	152
<b>Figure 8.1</b> Numbering Scheme for paraherquamide K ( <b>1.24</b> )	165
<b>Figure 8.2</b> gHMBC and gCOSY correlations for paraherquamide K ( <b>1.24</b> )	165
<b>Figure 8.3</b> Numbering scheme for paraherquamide L ( <b>1.25</b> )	165

<b>Figure 8.4</b> gHMBC and gCOSY correlations for paraherquamide L (1.25) .....	166
<b>Figure 8.5</b> Numbering scheme for paraherquamide M (1.26) .....	166
<b>Figure 8.6</b> gHMBC and gCOSY correlations for paraherquamide M (1.26) in DMSO-d <sub>6</sub> .....	166
<b>Figure 8.7</b> NOESY correlations for paraherquamide M (1.26) in DMSO-d <sub>6</sub> .....	167
<b>Figure 8.8</b> Numbering scheme for paraherquamide N (1.27) .....	167
<b>Figure 8.9</b> gHMBC and gCOSY correlations for paraherquamide N (1.27) in DMSO-d <sub>6</sub> .....	167
<b>Figure 8.10</b> NOESY correlations for paraherquamide N (1.27) in DMSO-d <sub>6</sub> .....	168
<b>Figure 8.11</b> Numbering scheme for malbrancheamide (1.17) .....	168
<b>Figure 8.12</b> Numbering scheme for spiromalbramide (3.1) .....	168
<b>Figure 8.13</b> gHMBC and gCOSY correlations for spiromalbramide (3.1) in CD <sub>3</sub> OD .....	169
<b>Figure 8.14</b> NOESY correlations for spiromalbramide (3.1) in CD <sub>3</sub> OD .....	169
<b>Figure 8.15</b> Michaelis-Menten model kinetics for paraherquamide K (1.24) .....	169
<b>Figure 8.16</b> Michaelis-Menten model kinetics for paraherquamide L (1.25) .....	170
<b>Figure 8.17</b> PhqK reactions with paraherquamide K 1.24 (a.) and paraherquamide L 1.25 (b.) with no enzyme control in blue and the reaction shown in red .....	170
<b>Figure 8.18</b> a.) MFo-DFc map contoured at 1σ for malbrancheamide B cocrystal structure. b.) 2MFo-DFc map contoured at 1σ for malbrancheamide B cocrystal structure ·	171
<b>Figure 8.19</b> a.) MFo-DFc map contoured at 1σ for FAD in malbrancheamide B cocrystal structure. b.) 2MFo-DFc map contoured at 1σ for FAD in malbrancheamide B cocrystal structure .....	171
<b>Figure 8.20</b> a. MFo-DFc map contoured at 1σ for FAD cocrystal structure. b. 2MFo-DFc map contoured at 1σ for FAD cocrystal structure. ....	171
<b>Figure 8.21</b> a.) MFo-DFc map contoured at 1σ for malbrancheamide C cocrystal structure.	

b.) 2MFO-DFc map contoured at 1 $\sigma$ for malbrancheamide C cocrystal structure	172
<b>Figure 8.22</b> a.) MFO-DFc map contoured at 1 $\sigma$ for FAD in malbrancheamide C cocrystal structure. b.) 2MFO-DFc map contoured at 1 $\sigma$ for FAD in malbrancheamide C cocrystal structure	172
<b>Figure 8.23</b> a.) MFO-DFc map contoured at 1 $\sigma$ for paraherquamide K cocrystal structure. b.) 2MFO-DFc map contoured at 1 $\sigma$ for paraherquamide K cocrystal structure	172
<b>Figure 8.24</b> a.) MFO-DFc map contoured at 1 $\sigma$ for FAD in paraherquamide K cocrystal structure. b.) 2MFO-DFc map contoured at 1 $\sigma$ for FAD in paraherquamide K cocrystal structure	173
<b>Figure 8.25</b> a.) MFO-DFc map contoured at 1 $\sigma$ for paraherquamide L cocrystal structure. b.) 2MFO-DFc map contoured at 1 $\sigma$ for paraherquamide L cocrystal structure	173
<b>Figure 8.26</b> a. MFO-DFc map contoured at 1 $\sigma$ for FAD in paraherquamide L cocrystal structure. b. 2MFO-DFc map contoured at 1 $\sigma$ for FAD in paraherquamide L cocrystal structure.	173
<b>Figure 8.27</b> <sup>1</sup> H-NMR of paraherquamide K ( <b>1.24</b> ) (800 MHz, DMSO-d <sub>6</sub> )	174
<b>Figure 8.28</b> <sup>13</sup> C-NMR of paraherquamide K ( <b>1.24</b> ) (800 MHz, DMSO-d <sub>6</sub> )	175
<b>Figure 8.29</b> gHSQCAD correlations of paraherquamide K ( <b>1.24</b> ) (800 MHz, DMSO-d <sub>6</sub> )	176
<b>Figure 8.30</b> gHMBCAD correlations of paraherquamide K ( <b>1.24</b> ) (800 MHz, DMSO-d <sub>6</sub> )	177
<b>Figure 8.31</b> gCOSY correlations of paraherquamide K ( <b>1.24</b> ) (800 MHz, DMSO-d <sub>6</sub> )	178
<b>Figure 8.32</b> <sup>1</sup> H-NMR of paraherquamide L ( <b>1.25</b> ) (600 MHz, DMSO-d <sub>6</sub> )	179
<b>Figure 8.33</b> <sup>13</sup> C-NMR of paraherquamide L ( <b>1.25</b> ) (600 MHz, DMSO-d <sub>6</sub> )	180
<b>Figure 8.34</b> gHSQCAD correlations of paraherquamide L ( <b>1.25</b> ) (600 MHz, DMSO-d <sub>6</sub> )	181
<b>Figure 8.35</b> gHMBCAD correlations of paraherquamide L ( <b>1.25</b> ) (600 MHz, DMSO-d <sub>6</sub> )	182

<b>Figure 8.36</b> gCOSY correlations of paraherquamide L ( <b>1.25</b> ) (600 MHz, DMSO-d <sub>6</sub> )	183
<b>Figure 8.37</b> <sup>1</sup> H-NMR of paraherquamide M ( <b>1.26</b> ) (800 MHz, DMSO-d <sub>6</sub> )	184
<b>Figure 8.38</b> <sup>13</sup> C-NMR of paraherquamide M ( <b>1.26</b> ) (800 MHz, DMSO-d <sub>6</sub> )	185
<b>Figure 8.39</b> gHSQCAD correlations of paraherquamide M ( <b>1.26</b> ) (800 MHz, DMSO-d <sub>6</sub> )	186
<b>Figure 8.40</b> gHMBCAD correlations of paraherquamide M ( <b>1.26</b> ) (800 MHz, DMSO-d <sub>6</sub> )	187
<b>Figure 8.41</b> gCOSY correlations of paraherquamide M ( <b>1.26</b> ) (600 MHz, DMSO-d <sub>6</sub> )	188
<b>Figure 8.42</b> NOE correlations of paraherquamide M ( <b>1.26</b> ) (800 MHz, DMSO-d <sub>6</sub> )	189
<b>Figure 8.43</b> <sup>1</sup> H-NMR of paraherquamide N ( <b>1.27</b> ) (600 MHz, DMSO-d <sub>6</sub> )	190
<b>Figure 8.44</b> <sup>13</sup> C-NMR of paraherquamide N ( <b>1.27</b> ) (800 MHz, DMSO-d <sub>6</sub> )	191
<b>Figure 8.45</b> gHSQCAD correlations of paraherquamide N ( <b>1.27</b> ) (800 MHz, DMSO-d <sub>6</sub> )	192
<b>Figure 8.46</b> gHMBCAD correlations of paraherquamide N ( <b>1.27</b> ) (800 MHz, DMSO-d <sub>6</sub> )	193
<b>Figure 8.47</b> gCOSY correlations of paraherquamide N ( <b>1.27</b> ) (600 MHz, DMSO-d <sub>6</sub> )	194
<b>Figure 8.48</b> NOE correlations of paraherquamide N ( <b>1.27</b> ) (800 MHz, DMSO-d <sub>6</sub> )	195
<b>Figure 8.49</b> <sup>1</sup> H-NMR of fungal malbrancheamide ( <b>1.17</b> ) (600 MHz, CD <sub>3</sub> OD)	196
<b>Figure 8.50</b> <sup>1</sup> H-NMR of spiromalbramide ( <b>3.1</b> ) (600 MHz, CD <sub>3</sub> OD)	197
<b>Figure 8.51</b> gHSQCAD correlations of spiromalbramide ( <b>3.1</b> ) (600 MHz, CD <sub>3</sub> OD)	198
<b>Figure 8.52</b> gHMBCAD correlations of spiromalbramide ( <b>3.1</b> ) (600 MHz, CD <sub>3</sub> OD)	199
<b>Figure 8.53</b> NOE correlations of spiromalbramide ( <b>3.1</b> ) (600 MHz, CD <sub>3</sub> OD)	200
<b>Figure 9.1</b> SAXS data for GRK5 and the Ca <sup>2+</sup> ·CaM-GRK5 complex	211
<b>Figure 9.2</b> Flexibility analysis of GRK5 and the Ca <sup>2+</sup> ·CaM-GRK5 complex	212
<b>Figure 9.3</b> Reconstructed electron densities from SEC-SAXS experiments using DENSS	212
<b>Figure 9.4</b> Representative negative stain micrographs	213
<b>Figure 9.5</b> Representative ITC curves	214

<b>Figure 9.6</b> Ca <sup>2+</sup> ·CaM complex with isomalbrancheamide D·····	214
<b>Figure 9.7</b> Electron density for malbrancheamide (Mal) bound to Ca <sup>2+</sup> ·CaM·····	215
<b>Figure 9.8</b> Effect of malbrancheamide on GRK5 activities in the absence of Ca <sup>2+</sup> ·CaM·····	216
<b>Figure 9.9</b> Determination of terminal peptide binding affinities by flow cytometry protein interaction assay (FCPIA) ·····	217

## List of Tables

<b>Table 3.1</b> Non-proprietary hits from the MalA high-throughput screen .....	68
<b>Table 4.1</b> Michaelis-Menten kinetic parameters for PhqK reactions .....	73
<b>Table 4.2</b> Reaction data for PhqK variants .....	77
<b>Table 7.1</b> MalA' Crystallographic Information .....	113
<b>Table 7.2</b> Kinetic Parameters for MalA Chlorination Reactions .....	115
<b>Table 7.3</b> <sup>1</sup> H-NMR data for malbrancheamide B ( <b>1.15</b> ) fungal and <i>in vitro</i> samples .....	115
<b>Table 7.4</b> <sup>1</sup> H-NMR data for fungal and <i>in vitro</i> isomalbrancheamide B ( <b>1.16</b> ) .....	116
<b>Table 7.5</b> <sup>1</sup> H-NMR data for fungal and <i>in vitro</i> malbrancheamide ( <b>1.17</b> ) .....	116
<b>Table 7.6</b> <sup>1</sup> H-NMR and <sup>13</sup> C-NMR data for <i>in vitro</i> malbrancheamide C ( <b>1.92</b> ) .....	117
<b>Table 7.7</b> <sup>1</sup> H-NMR and <sup>13</sup> C-NMR data for <i>in vitro</i> isomalbrancheamide C ( <b>1.93</b> ) .....	118
<b>Table 7.8</b> <sup>13</sup> C-NMR, <sup>1</sup> H-NMR, gHMBCAD and gCOSY correlations for <i>in vitro</i> malbrancheamide D ( <b>1.94</b> ) .....	119
<b>Table 7.9</b> <sup>13</sup> C-NMR, <sup>1</sup> H-NMR, gHMBCAD, gCOSY correlations for <i>in vitro</i> isomalbrancheamide D ( <b>1.95</b> ) .....	120
<b>Table 8.1</b> Primers for <i>phqK</i> intron removal and amplification .....	157
<b>Table 8.2</b> Crystallographic information .....	158
<b>Table 8.3</b> NMR data for paraherquamide K ( <b>1.24</b> ) .....	159
<b>Table 8.4</b> NMR data for paraherquamide L ( <b>1.25</b> ) .....	160
<b>Table 8.5</b> NMR data for paraherquamide M ( <b>1.26</b> ) .....	161
<b>Table 8.6</b> NMR data for paraherquamide N ( <b>1.27</b> ) .....	162
<b>Table 8.7</b> <sup>1</sup> H-NMR for malbrancheamide isolated from <i>Malbranchea aurantiaca</i> .....	163

<b>Table 8.8</b> $^{13}\text{C}$ -NMR, $^1\text{H}$ -NMR, gHMBCAD, gCOSY, and NOESY correlations for spiromalbramide isolated from PhqK <i>in vitro</i> reaction .....	164
<b>Table 9.1</b> SAXS Data collection parameters .....	208
<b>Table 9.2</b> SAXS Data collection statistics for SEC-SAXS experiments .....	209
<b>Table 9.3</b> Crystallographic data collection for the $\text{Ca}^{2+}$ -CaM complexes .....	210



## List of Schemes

<b>Scheme 1.1</b> Malbrancheamide biosynthetic pathway .....	3
<b>Scheme 1.2</b> Paraherquamide biosynthetic pathway.....	4
<b>Scheme 1.3</b> Notoamide biosynthetic pathway.....	6
<b>Scheme 1.4</b> Proposed mechanism for paraherquamide spirooxindole formation.....	13
<b>Scheme 1.5</b> Proposed mechanism for assembly of notoamides C and D .....	14
<b>Scheme 1.6</b> Aurachin biosynthesis.....	20
<b>Scheme 1.7</b> Proposed mechanisms for dioxepin ring formation .....	23
<b>Scheme 2.1</b> MalG NRPS reaction.....	31
<b>Scheme 2.3</b> Prenylation and IMDA cyclization to generate (+)-premalbrancheamide .....	33
<b>Scheme 3.2</b> Proposed mechanism for MalA halogenation.....	61

## List of Abbreviations

<b>AngII</b>	–	angiotensin II
<b>AST</b>	–	active site tether
<b>CaM</b>	–	calmodulin
<b>cGMP</b>	–	cyclic guanosine monophosphate
<b>CYP</b>	–	cytochrome P450 enzyme
<b>DKP</b>	–	diketopiperazine
<b>DMAPP</b>	–	dimethylallyl pyrophosphate
<b>DNA</b>	–	deoxyribonucleic acid
<b>eNOS</b>	–	endothelial nitric oxide synthase
<b>FAD</b>	–	flavin adenine dinucleotide
<b>FDH</b>	–	Flavin-dependent halogenase
<b>FMO</b>	–	Flavin-dependent monooxygenase
<b>FR</b>	–	flavin reductase
<b>GPCR</b>	–	G-protein-coupled receptor
<b>GRK</b>	–	G-protein-coupled receptor kinase
<b>HDAC</b>	–	histone deacetylase
<b>IMDA</b>	–	intramolecular Diels-Alder
<b>ITC</b>	–	isothermal titration calorimetry
<b>KD</b>	–	kinase domain
<b>LC/MS</b>	–	liquid chromatography/mass spectrometry
<b>MALS</b>	–	multi-angle light scattering
<b>MBP</b>	–	maltose-binding protein
<b>MKP</b>	–	monoketopiperazine
<b>MLCK</b>	–	myosin light chain kinase
<b>MyBP</b>	–	myelin basic protein
<b>NADH</b>	–	nicotinamide adenine dinucleotide
<b>NADPH</b>	–	nicotinamide adenine dinucleotide phosphate

<b>NFAT</b>	–	nuclear factor of activated T-cells
<b>NIBR</b>	–	Novartis Institutes for BioMedical Research
<b>NMR</b>	–	nuclear magnetic resonance
<b>NO</b>	–	nitric oxide
<b>NRPS</b>	–	nonribosomal peptide synthetase
<b>ONB</b>	–	<i>O</i> -nitrobenzyl
<b>PCP</b>	–	peptidyl carrier protein
<b>PDB</b>	–	protein data bank
<b>PDE1</b>	–	phosphodiesterase 1
<b>PE</b>	–	phenylephrine
<b>RH</b>	–	regulator of G-protein signaling homology
<b>ROS</b>	–	rod outer segments
<b>SAM</b>	–	<i>S</i> -adenosyl- <i>L</i> -methionine
<b>SAXS</b>	–	small angle X-ray scattering
<b>SD</b>	–	standard deviation
<b>SDR</b>	–	short chain dehydrogenase reductase
<b>SEC</b>	–	size exclusion chromatography

## Abstract

The class of fungal indole alkaloids containing the bicyclo[2.2.2]diazaoctane ring is comprised of diverse molecules that display a range of biological activities. While much interest has been garnered due to their therapeutic potential, this class of molecules also displays unique chemical functionality, making them intriguing synthetic targets. Many elegant and intricate total syntheses have been developed to generate these alkaloids, but selectivity and yield have always presented barriers to efficient synthesis. Alternatively, if we can understand the molecular mechanisms behind how the fungi produce these complex molecules, we can leverage the power of nature to perform these chemical transformations.

This work has delved into the enzymatic machinery responsible for key biosynthetic transformations in the production of the fungal indole alkaloids. With full mechanistic characterization, these enzymes are now available as biocatalytic tools for generating pre-existing and novel indole alkaloids in the search for the next line of therapeutic molecules.

The studies presented in this thesis start with the discovery of the Diels-Alderase enzymes responsible for forming the characteristic bicyclo[2.2.2]diazaoctane ring. This core component is the signature of this class of molecules and we have found that the enzymes responsible for its formation perform multiple roles in the biosynthesis, including reduction and cyclization.

Halogen atoms are found within a small number of the fungal indole alkaloids, but the halogen moieties on these molecules significantly contribute to the biological activity. This is particularly true in the case of malbrancheamide, for which the dichlorination is required for its biological activity. This work has thoroughly characterized the iterative late-stage halogenase (MalA) which performs dichlorination as the last step in malbrancheamide biosynthesis. Cocrystal structures and computational studies have led to the generation of site-selective variants produced through structure-based engineering of the halogenase. A collaboration with Novartis Institutes for BioMedical Research led to the discovery that this halogenase has a broad substrate scope, and thus great utility as a biocatalyst for late-stage halogenation.

A common functionality of the fungal indole alkaloids is the spirooxindole center. This is another core component of many molecules within this class, and it provides dimension to an

otherwise relatively planar structure. This thesis presents the discovery of a flavin monooxygenase that is responsible for the selective spirocyclization of the potent antihelmintic paraherquamides. The natural substrates of this enzyme were identified and cocrystal structures demonstrated that the enzyme binds the molecules in a manner conducive for stereocontrol. The broad substrate scope demonstrated with PhqK provides further evidence that it could be utilized to develop new therapeutic molecules.

As mentioned, this class of fungal indole alkaloids has displayed a wide range of biological activities. There has been significant interest in the malbrancheamides because they have displayed a potent vasorelaxant effect, and a complete reversal of the swelling of cardiac tissue. With this in mind, the culmination of my thesis work demonstrated the utility of malbrancheamide as a probe for protein-protein interactions that have been implicated in cardiovascular disease. Malbrancheamide was characterized as a selective inhibitor of calmodulin, and its unique binding mode relieved the effects of  $\text{Ca}^{2+}$ -calmodulin-induced cardiac hypertrophy. We used the *in vitro* data with malbrancheamide and the  $\text{Ca}^{2+}$ -CaM·GRK5 complex to validate a much-debated model for this protein-protein interaction, providing valuable insight for the field of  $\text{Ca}^{2+}$ -calmodulin-dependent kinase signaling.

# Chapter 1

## Introduction

- ***Introduction to Natural Products***

Extracts from natural sources have long been used as medicines for treating a variety of ailments. Dating back to 2600 B.C., oils from *Cupressus sempervirens* (cypress) and *Commiphora* (myrrh) have been used to treat the symptoms of colds including cough and inflammation. Historically, these extracts were applied without knowledge of the active ingredients, with the only source of validation being relief from the illness. We now know and understand a lot more about these natural medicines, and have even found additional uses for natural products such as biofuels for sustainable energy and pesticides for eco-friendly agricultural use. Of utmost importance, our scientific advances have aided us in identifying these unique natural products and how they are made. With the availability of next generation sequencing, methods for identifying biosynthetic pathways have become more efficient, and we can readily piece together the components required for generating a natural product of interest.

The gene products, or biosynthetic enzymes, each catalyze a specific reaction within a sequence to generate a secondary metabolite. The regio- and stereoselectivity of the enzyme is usually highly regulated, forming complex molecules which are difficult to make via traditional synthetic methods. Enzymes have evolved this selectivity over time to assume an evolutionary advantage for the host organism. In a single organism, the enzymes within a biosynthetic pathway may evolve in concert, building on the selectivity of the upstream enzymes. Additionally, enzymes from different pathways can evolve to react with the same intermediates in a form of crosstalk to further diversify the secondary metabolite pool.

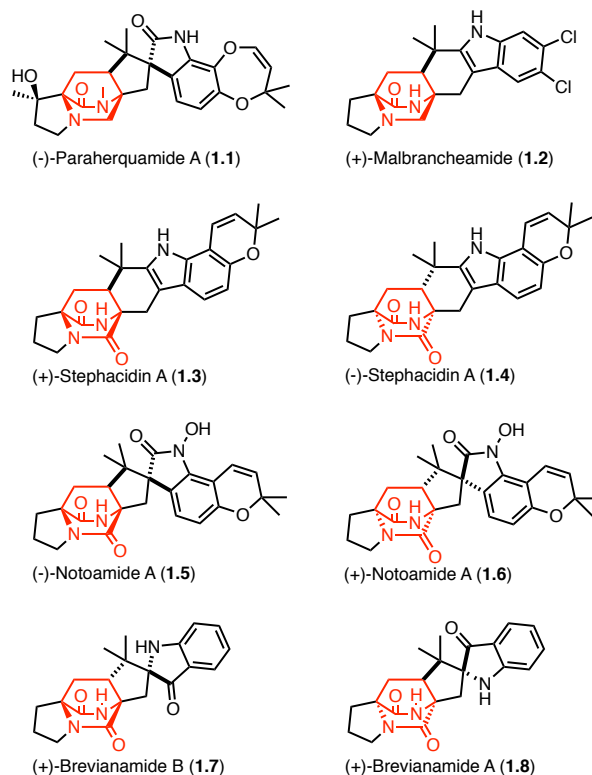
The biosynthetic pathways within fungal systems have demonstrated exquisite evolutionary development with divergent pathways between the various organisms. In some cases, the biosynthetic gene clusters within two different strains can be highly identical, yet they produce antipodal natural products. The mechanistic details behind this selectivity could only be revealed through in depth biochemical and structural characterization of the enzymes involved in these intricate chemical transformations.

## 1.2 Evolution of Natural Product Biosynthesis in the Bicyclo[2.2.2]diazaoctane Containing Fungal Indole Alkaloids\*

### 1.2.1 Introduction to Fungal Indole Alkaloids

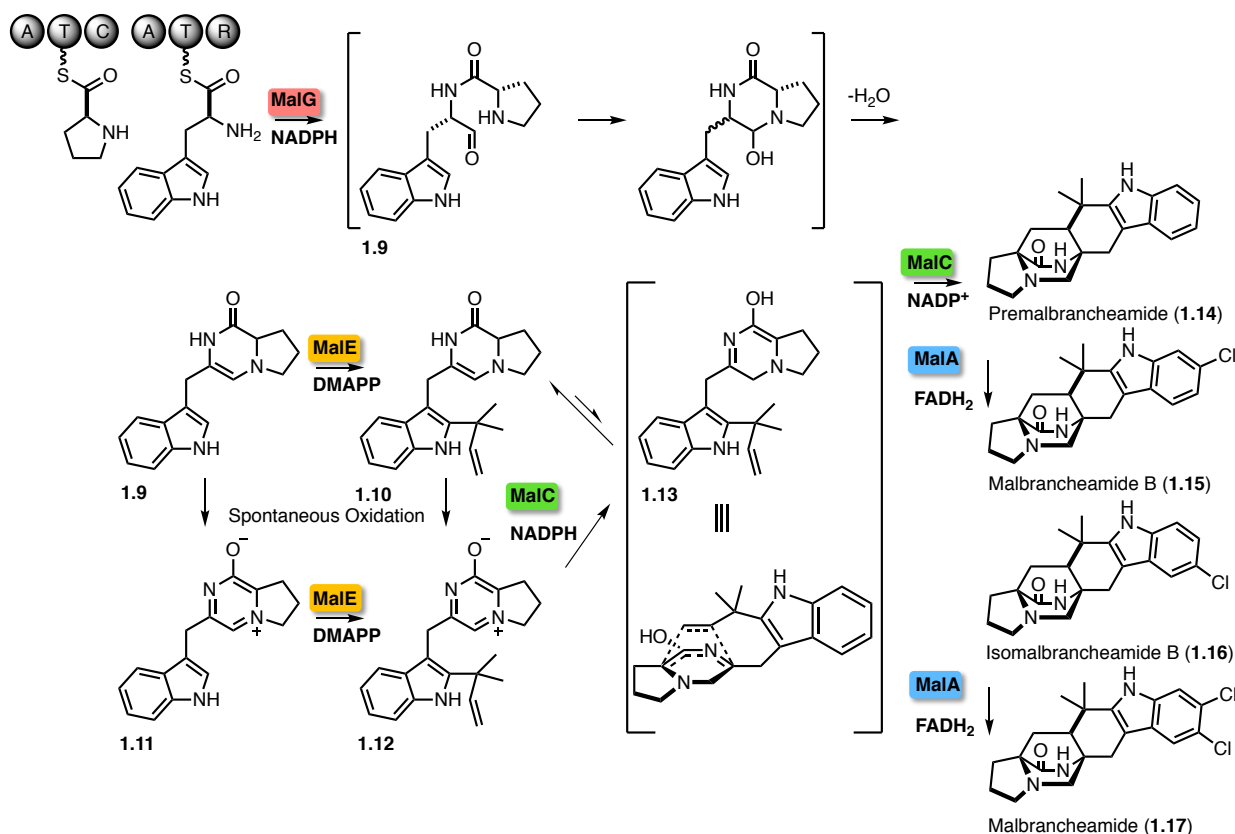
The fungal indole alkaloid class of natural products contains molecules with unique structural properties and a variety of biological activities. In particular, the subgroup typified by the bicyclo[2.2.2]diazaoctane ring has been extensively studied within our group. Initially, feeding studies using isotopically labeled precursors and biomimetic chemical syntheses provided insight into the biosynthetic pathways for these molecules. These efforts were later supplemented with genome sequencing and bioinformatics analyses,<sup>2</sup> greatly advancing our understanding of how these natural products are generated, yet we have only recently utilized our capabilities to address the knowledge gaps with detailed biochemical characterization. Multiple strategies including molecular and structural biology have brought our foundational understanding of these pathways to the forefront of the field.

The extent of our work has mainly focused on the families of monoketopiperazines (MKP) such as the paraherquamides (Phq)<sup>3, 4</sup> and malbrancheamides (Mal),<sup>5</sup> and diketopiperazine (DKP) stephacidins,<sup>6</sup> notoamides (Not),<sup>7-9</sup> and brevianamides (Bvn) (**Figure 1.1**).<sup>10</sup> These molecules display significant biological activities with potential therapeutic value such as antihelmintics,<sup>11-14</sup> vasodilators,<sup>15</sup> and anticancer agents.<sup>16</sup> All bear a similar core scaffold, but vary structurally based on key functionalities such as oxidation, halogenation, and additional ring systems. Thus, our initial bioinformatic analyses have enabled us to link the construction of the common core to the homologous enzymes, while cluster-specific genes provide the basis for structural differences.<sup>2</sup> There is a clear evolutionary



**Figure 1.1** Select fungal indole alkaloids containing the bicyclo[2.2.2]diazaoctane ring.

\* Section 1.2 is included in a manuscript in preparation.

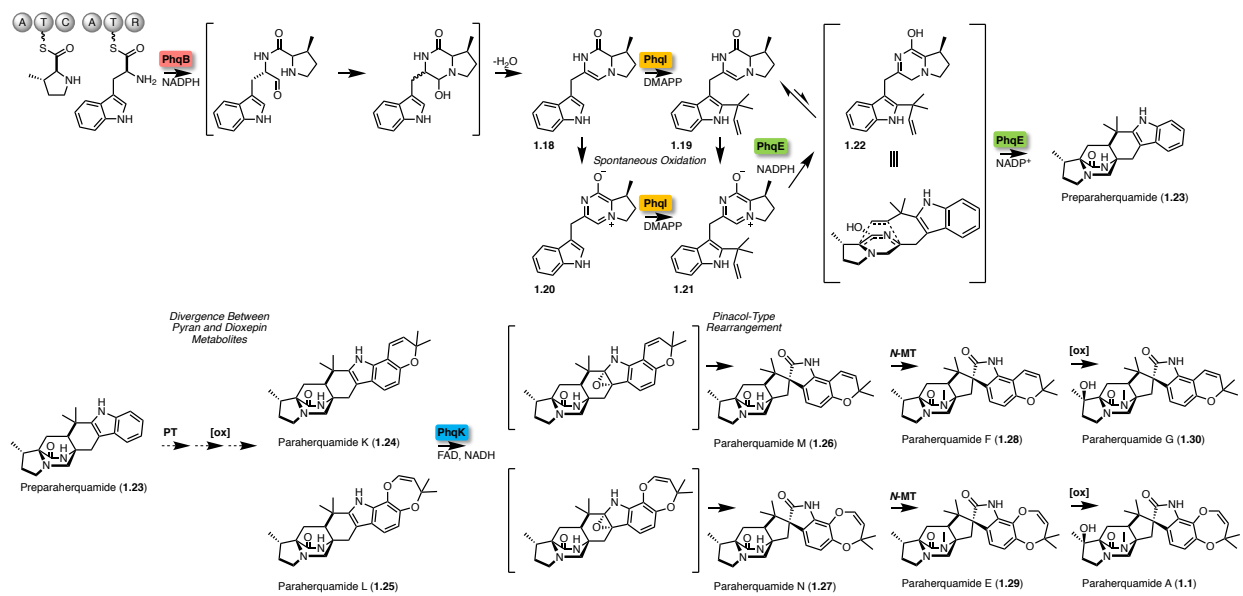


**Scheme 1.1** Malbrancheamide biosynthetic pathway. The pathway includes NRPS MalG, prenyltransferase MalE, intramolecular Diels-Alderase MalC, and flavin-dependent halogenase MalA. divergence between the MKP and DKP containing families within this class. Thus, installation of similar functional groups may be accomplished by different type of enzymes. Additionally, parallel work in pathways that do not contain the core bicycle reveal compelling biosynthetic branch points. Extensive characterization of the fumitremorgin,<sup>17, 18</sup> spirotryprostatin,<sup>19</sup> and aurachin<sup>20, 21</sup> families by a number of groups has provided insights into the homologous transformations within these varied bicyclo[2.2.2]diazaoctane ring containing metabolic systems.

### 1.2.2 Distinct Nonribosomal Peptide Synthetases

The nonribosomal peptide synthetases (NRPSs) are multidomain enzymes that perform the first step in the biosynthesis of these fungal indole alkaloids. The architecture of the NRPS consists of adenylation (A), thiolation (T), condensation (C), and reduction I domains, where the first A domain in this class of compounds is selective for proline, distinguishing these systems from other fungal dimodular NRPS pathways. The domain construction of the terminal module determines the oxidation state of the offloaded dipeptide, and subsequently the core scaffold. The DKP containing molecules involve a bimodular NRPS with A-T-C-A-T-C domain organization,

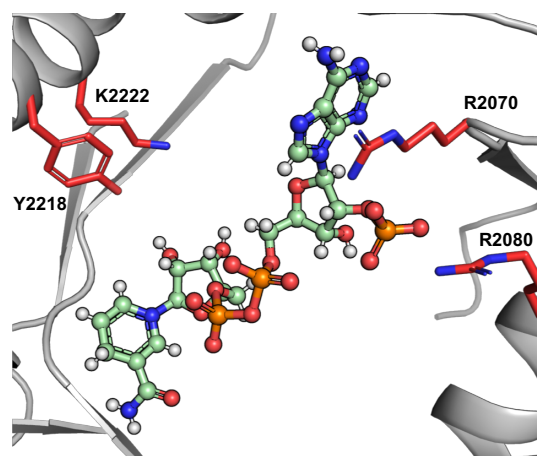




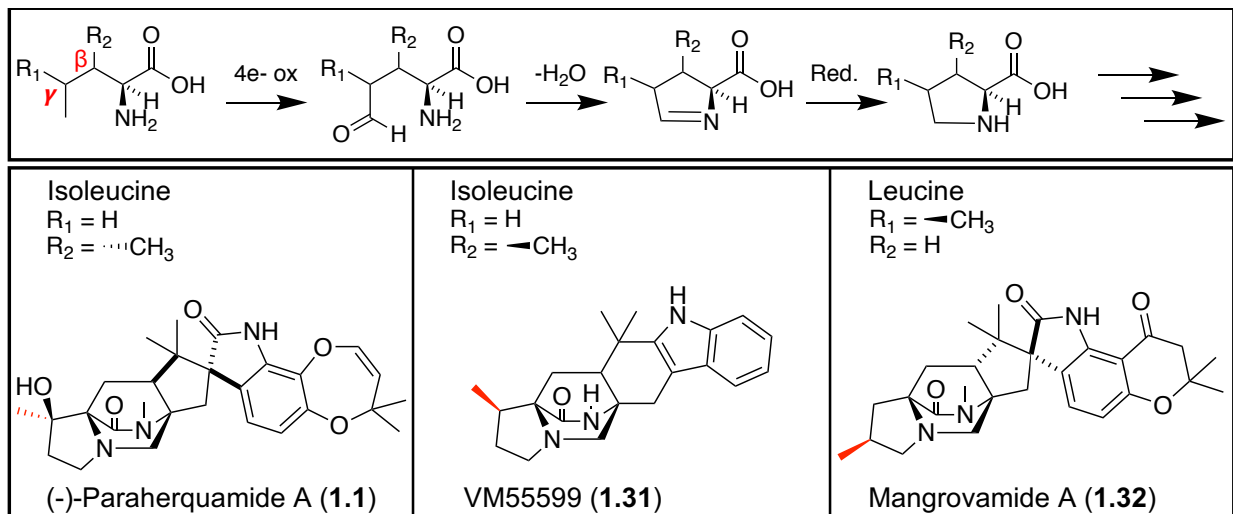
**Scheme 1.2** Paraherquamide biosynthetic pathway. The pathway includes NRPS PhqB, prenyltransferase PhqI, intramolecular Diels-Alderase PhqE, and flavin monooxygenase PhqK. A putative prenyltransferase and oxidative enzyme are proposed to be involved in the formation of the pyran and dioxepin rings.

whereas the MKP containing molecules involve a bimodular NRPS with A-T-C-A-T-R organization. For example, the notoamide (DKP) biosynthetic pathway contains an NRPS with a terminal condensation domain (NotE),<sup>22</sup> while the malbrancheamide and paraherquamide (MKPs) biosynthetic pathways include NRPSs with terminal reductase domains.<sup>2, 23</sup> The different terminal domains and the coinciding mono- or diketopiperazine systems, indicate that the NRPSs offload distinct dipeptide intermediates for the subsequent intramolecular Diels-Alder (IMDA) reaction. Thus, it is now evident that the mechanism for the IMDA cyclization is based on the organization of the upstream NRPS.<sup>2</sup>

The mechanism for the MKP-producing NRPS has been identified based on recent investigations of the Phq and Mal systems by our group. The NRPS couples L-proline and L-tryptophan to produce an L-Pro-L-Trp aldehyde product through cyclization and dehydration (**Scheme 1.1** and **1.2**). Under conditions that contain the NRPS alone, the linear aldehyde product spontaneously oxidizes to a cyclic aromatic zwitterionic species.<sup>23</sup> This catalytic mechanism



**Figure 1.2** PhqB R-domain structure. The structure is in complex with NADPH displaying canonical SDR catalytic residues.



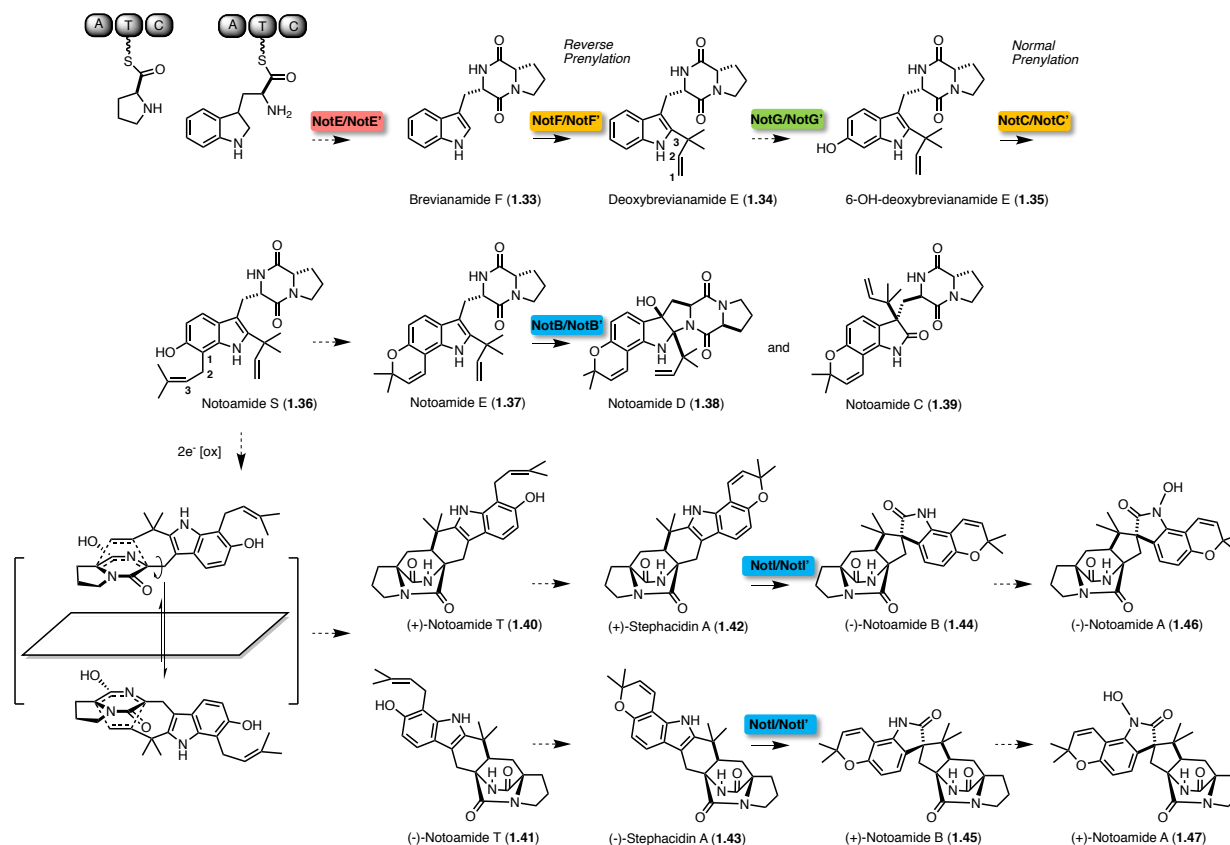
**Figure 1.3** Proposed mechanism for the generation of the methyl proline-containing molecules.

requires the NADPH cofactor to provide the reducing equivalents, and both PhqB and MalG R-domains contain the canonical short chain dehydrogenase reductase (SDR) tyrosine and lysine active site residues (**Figure 1.2**).<sup>23</sup>

The key difference between the Mal and Phq NRPS systems is the incorporation of  $\beta$ -methyl proline into the dipeptide core within the paraherquamide pathway. Through isotopic feeding studies in *Penicillium fellutanum* (ATCC: 20841), it was determined that isoleucine is the precursor to the  $\beta$ -methyl proline,<sup>24</sup> and that this nonproteinogenic amino acid is directly incorporated into paraherquamide A (**1.1**).<sup>25</sup> The enzyme responsible for this amino acid transformation is proposed to perform an oxidative cyclization of the nitrogen atom onto the methyl group (C $\delta$ ), followed by a selective reduction on the same face as the methyl (**Figure 1.3**).<sup>24-26</sup> The subsequent hydroxylation observed on the proline moiety in the paraherquamides occurs later in the biosynthesis, following bicyclic ring formation (**Scheme 1.2**).<sup>27</sup> Interestingly, compound VM55599 (**1.31**) has the opposite stereochemistry of the methyl group,<sup>24, 28</sup> indicating that *Penicillium* sp. IMI 332995 contains an enzyme with the opposite facial selectivity. Additionally, the recently discovered mangrovamide A (**1.32**) contains a unique  $\gamma$ -methyl proline moiety.<sup>29</sup> We hypothesize that this functionality is generated by a leucine-selective enzyme (**Figure 1.3**) involving oxidation, cyclization, and reduction, similar to the pathway involved in the echinocandin biosynthesis.<sup>30</sup>

### 1.2.3 Redundant Prenyltransferases

Prenyltransferases catalyze normal and reverse prenylations in natural product biosynthetic pathways, where the bond is formed between the substrate and C1 of the isoprene unit in normal



**Scheme 1.3** Notoamide biosynthetic pathway. The pathway includes NRPS NotE, prenyltransferases NotF and NotC, putative hydroxylase NotG, a proposed but unidentified intramolecular Diels-Alderase, and flavin monooxygenases NotB and NotI.

prenylation and between the substrate and C3 of the isoprene in reverse prenylation (**Scheme 1.3**). We have found that most of the bicyclo[2.2.2]diazaoctane fungal indole alkaloid gene clusters contain an additional prenyltransferase relative to the number of prenyl units incorporated into the molecule. The initial reverse prenylation occurs at the indole C2 position of the NRPS dipeptide product,<sup>23</sup> while prenylation of the phenolic portion of the indole to form the pyran and dioxepin rings can either occur before (**Scheme 1.2**) or after (**Scheme 1.3**) the formation of the [2.2.2] fused bicycle. Precursor incorporation studies have provided insight on the origin of the prenyl groups. *P. fellutanum*, *P. brevicompactum*, and *Aspergillus ustus* all demonstrated incorporation of [<sup>13</sup>C<sub>2</sub>]-acetate into paraherquamide A, brevianamide A, and austamide, respectively, indicating a mevalonic acid pathway origin of the prenyl group. Additionally, the geminal dimethyl group stemming from the indole C2 position were found to be non-selectively assigned, thus the facial bias of the reverse prenylation in the biosynthesis of these molecules varies by pathway.<sup>31</sup>

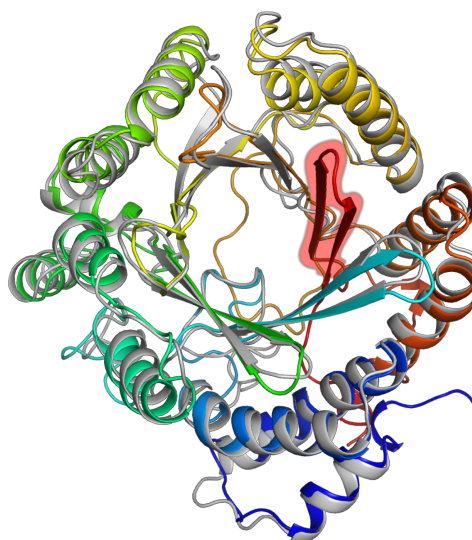
Enzymatic analyses have provided key details for the prenylation reactions within these systems. Within notoamide biosynthesis in the marine-derived *Aspergillus protuberus* (formerly

*Aspergillus* sp. MF297-2), NotF has been characterized as the reverse prenyltransferase or deoxybrevianamide E synthase (**Scheme 1.3**).<sup>22, 32, 33</sup> It was hypothesized that brevianamide F (**1.33**) would be the substrate for NotF based on previous work in fumitremorgin biosynthesis,<sup>17, 18</sup> although in the case of the fumitremorgins, **1.33** underwent normal prenylation at the indole C2 position.<sup>18</sup> The notoamide biosynthetic pathway also contains a normal prenyltransferase (NotC) which is involved in the formation of notoamide S (**1.36**), indicating its significance in pyran ring formation.<sup>22</sup> Additionally, the NotF homolog BrePT from the terrestrial *Aspergillus amoenus* (formerly *A. versicolor* NRRL 35600) has been characterized.<sup>34</sup> Although these strains contain biosynthetic gene clusters that are nearly identical, BrePT displayed a broader substrate scope, while maintaining its indole C2 regioselectivity.

Our investigations in the MKP pathways have led to the characterization of the indole C2-reverse prenyltransferases MalE and PhqI (**Schemes 1.1** and **1.2**).<sup>23</sup> We determined that the free-standing reduced dipeptide product was the favored substrate (**1.9** and **1.18**), rather than either the oxidized zwitterion or an NRPS-carrier protein tethered substrate. Additionally, free tryptophan was not prenylated, eliminating the option of an early-stage modification. A presumed redundant prenyltransferase MalB displayed low activity with the reduced dipeptide (**1.9**) as substrate. Sequence alignments and structural comparisons revealed that MalB is missing two integral strands of the central  $\beta$  barrel likely responsible for its attenuated activity (**Figure 1.4**).

#### 1.2.4 Enantioselective Diels-Alderases

Initial proposals for the biosynthesis of the bicyclo[2.2.2]diazaoctane core involved an intramolecular [4+2] Diels-Alder (IMDA) reaction to form the *exo* and *endo* products.<sup>35</sup> Initial biosynthetic proposals hypothesized that both the MKP and DKP systems were formed through a common DKP intermediate, with the tryptophan carbonyl of the MKPs derived from a net four-electron reduction occurring



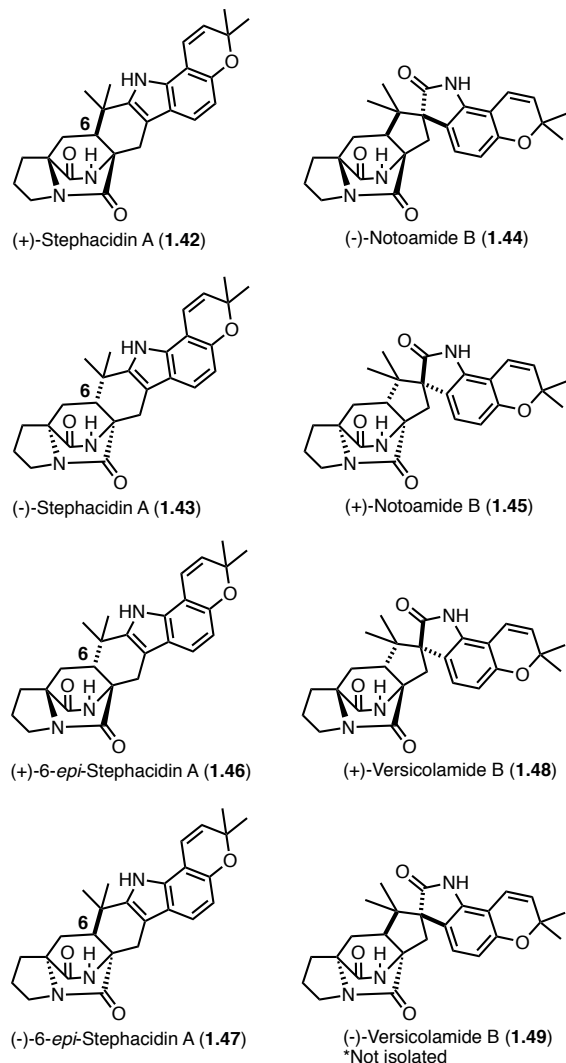
**Figure 1.4** Comparison of MalE and MalB. MalE (rainbow) and MalB (gray) Phyre2<sup>1</sup> models display the missing C-terminus of MalB in red.

after a putative Diels-Alder construction.<sup>35</sup> The search for the Diels-Alderase within this class of molecules was hindered by the inability to identify candidate enzymes through bioinformatic analysis. This was due to the unique identities of the respective Diels-Alderases that result from divergent evolutionary processes. Recent studies have shown that Diels-Alderase enzymes have lost their original function and evolved to control stereo-selectivity for the cycloaddition reaction.<sup>23, 36-42</sup> Additionally, bound cofactors have typically lost a catalytic function (e.g. methyl transfer, redox), and play a structural role in maintaining the proper active site conformation for stereochemical control.<sup>36-42</sup> Moreover, most Diels-Alderases have no identified catalytic residues, leading to the conclusion that these biocatalysts function through a mechanism of constraining the substrate in the proper orientation.<sup>36-42</sup> Within this class of bicyclo[2.2.2]diazaoctane ring containing molecules, the cycloaddition is proposed to be

stereospecific based on the *syn*- or *anti*- configuration of the bridged bicycle relative to C6 (Stephacidin numbering, **Figure 1.5**).<sup>35</sup> The biosynthetic IMDA enzyme is proposed to have strict diastereo- and enantioselectivity, but some organisms have been found to produce both antipodes.<sup>10, 43</sup>

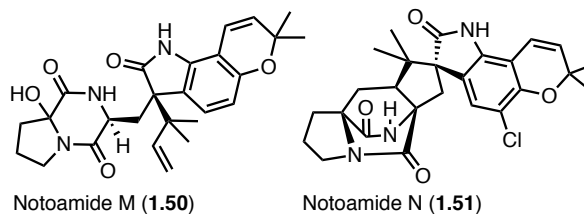
#### 1.2.4.1 Isolation and Precursor Incorporation Studies

Within the notoamide (Not') producing organism *A. amoenus*, both (-)-stephacidin A (**1.43**) and (+)-versicolamide B (**1.48**) have been isolated (**Figure 1.5**). The production of **1.48** suggests that the putative IMDA reaction that leads to the major metabolites within the producing organism may suffer from stereochemical leakage with respect to the facial selectivity of addition



**Figure 1.5** Metabolites isolated from *Aspergillus amoenus* and *Aspergillus protuberus*.

to the reverse isoprene moiety anchored at the indole C2 position.<sup>44</sup> Interestingly, the organism must then produce (+)-6-*epi*-stephacidin A (**1.46**), but not (+)-stephacidin A (**1.42**), indicating the presence of multiple highly selective IMDAases or one semi-selective IMDAase.



**Figure 1.6** Hydroxylated notoamide M and halogenated notoamide N. These molecules which exemplify some of the unique secondary metabolites in notoamide biosynthesis.

The marine-derived *Aspergillus* sp. (Not) also produces intermediates that may be relevant to the Diels-Alder reaction in this organism. The isolation of notoamide M (**1.50**), which is hydroxylated at C17, raised questions regarding the biosynthetic Diels-Alder reaction within the Not pathway (**Figure 1.6**).<sup>32</sup> It was proposed that **1.50** would undergo a dehydration to form the azadiene intermediate for cycloaddition; alternatively, **1.50** could be an artifact of hydrolytic capture of the transient azadiene species.

The notoamides and stephacidins are part of the same family of this class of alkaloids and are often co-isolated from a single strain. A unique aspect of these systems is that *A. protuberus* and *A. amoenus* produce antipodal versions of these molecules. While *A. protuberus* produces *exo* products (+)-stephacidin A (**1.42**)/(-)-notoamide B (**1.44**) and *endo* products (+)-6-*epi*-stephacidin A (**1.46**)/(+)-versicoloamide B (**1.48**), *A. amoenus* produces (-)-stephacidin A (**1.43**)/(+)-notoamide B (**1.45**), where in both systems the *exo*-metabolites are favored (**Scheme 1.3**, **Figure 1.5**).<sup>45</sup> After isolation of notoamide E (**1.37**) from the marine-derived *Aspergillus* sp., doubly <sup>13</sup>C-labeled **1.37** was prepared for feeding studies.<sup>33</sup> Surprisingly, it was not incorporated into the [2.2.2]-bicyclic ring containing molecules, and no bicycle-containing molecules were produced.<sup>33</sup> These data contrast with the brevianamide system in *P. brevicompactum*, where [<sup>3</sup>H]-deoxybrevianamide E (**1.34**) was incorporated into brevianamides A (**1.8**) and B (**1.7**), showing that the bicyclic ring was constructed from the reverse prenyl group and the DKP ring.<sup>46</sup> Alternatively, notoamide E was observed to be incorporated into notoamide C (**1.39**), notoamide D (**1.38**), and trace amounts of 3-*epi*-notoamide C in *A. amoenus* (**Scheme 1.3**).<sup>47</sup> To highlight the differences between the two strains, larger amounts of 3-*epi*-notoamide C were produced within the marine strain,<sup>33</sup> while the formation of the bicyclo[2.2.2]diazaoctane ring was not completely abolished in the terrestrial strain as trace amounts of unlabeled **1.42** and **1.44** were produced.<sup>47</sup> The presence of **1.37** appears to suppress the formation of stephacidin A (**1.42** and **1.43**) and notoamide

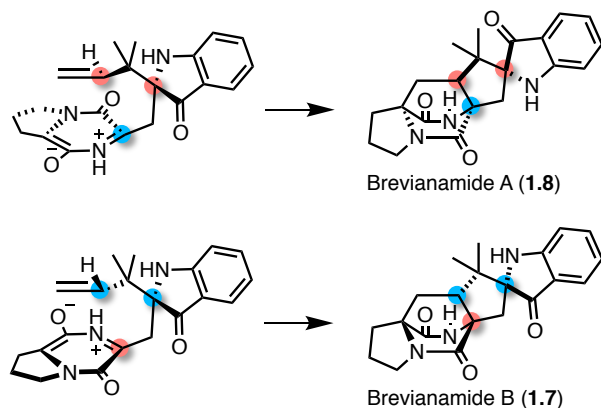
B (**1.44** and **1.45**), suggesting that this compound inhibits or diverts enzymatic machinery responsible for the production of the bicyclo[2.2.2]-diazaoctane-containing metabolites. These findings suggested a branch point in the biosynthetic pathway, just prior to the formation of **1.37**, perhaps after the formation of **1.36** (Scheme 1.3).<sup>47</sup>

Enzyme-controlled facial selectivity appears to be lost within the brevianamide

biosynthetic Diels-Alder reaction in *Penicillium brevicompactum* (Figure 1.7).<sup>46</sup> It was proposed that the pinacol-type rearrangement occurs on the 3-hydroxyindolenine, setting the *R*-absolute stereochemistry at the indoxyl quaternary center. Oxidation of the DKP subunit forms the azadiene of which the major rotamer would form **1.8** and the minor rotamer would form **1.7**.<sup>46</sup> Therefore the preponderance of **1.8** over **1.7** is likely due to intrinsic energy barriers in the cycloaddition reaction (non-enzyme catalyzed), but could potentially arise from the relative activities of two different IMDA enzymes or the affinity of a single enzyme active site for the individual conformers. Through the utilization of frontier molecular orbital theory it was predicted that, for the energy levels of a relatively electron-rich diene (such as that in the DKP) to effectively interact with a dienophile, powerful electron-withdrawing groups would need to be present in the dienophile. Since the prenyl group is an isolated electron-neutral vinyl group, one would not expect the [4+2] cycloaddition in DKPs to be spontaneous without catalysis, but biomimetic syntheses have shown that the azadiene undergoes a spontaneous IMDA to generate a mixture of brevianamides with enantiomorphic bicyclo[2.2.2] ring systems.<sup>48</sup>

#### 1.2.4.2 Enzymology

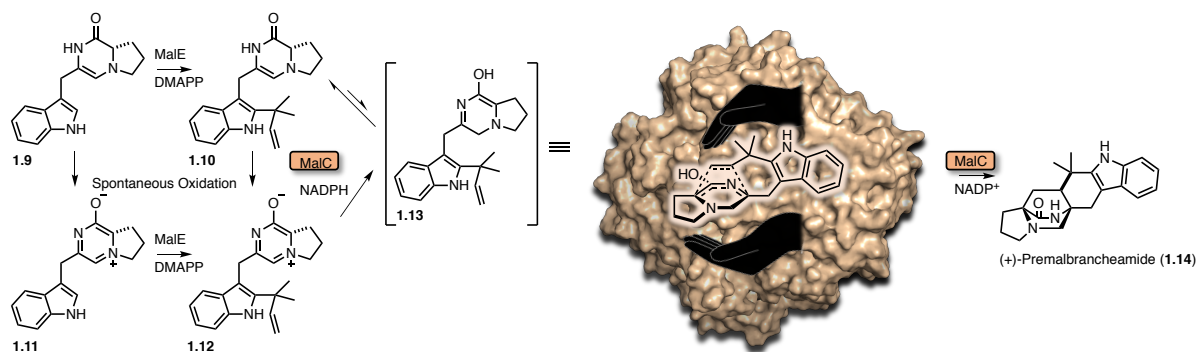
Recent work within our lab has provided key biochemical evidence to support the hypothesis of a biosynthetic Diels-Alder enzyme within MKP systems.<sup>23</sup> MalC was found to catalyze enantioselective cycloaddition as a bifunctional NADPH-dependent reductase/Diels-Alderase. MalC catalyzes NADPH-dependent reduction of the prenylated zwitterion **1.12** generating the reactive azadiene **1.13**, and controls the enantioselectivity of the cycloaddition reaction to generate exclusively (+)-**1.14**. In the absence of the enzyme, chemical reduction of the



**Figure 1.7** Brevianamide formation. Conformations of the azadiene *en route* to the major rotamer brevianamide A and the minor rotamer

prenylated zwitterion leads to racemic premalbrancheamide, thus the enzyme dependent shift in enantiomeric excess and rate increase led to the conclusion that this reaction is enzyme catalyzed (**Scheme 1.1**).

The crystal structures of MalC and the homolog PhqE (from the paraherquamide (Phq) pathway) provided further mechanistic insight. Their closest structural homologs are canonical SDRs, although the Diels-Alderase lack the essential catalytic residues for SDR activity (Tyr, Lys, Asn, Ser).<sup>49</sup> MalC Asp165 forms part of a conserved PDPGW motif and was shown to be required for reduction of the substrate prior to the Diels-Alder reaction. The structure showed that Asp165 has moved about 3Å closer to the substrate than the analogous amino acid in canonical SDRs, highlighting a larger change of the active site. Based on mutagenesis and computational studies, a mechanism for the reduction and Diels-Alder selectivity was proposed. An arginine residue and the 2'-OH of the NADPH ribose initiate the reaction with NADPH acting as the hydride donor and D165 stabilizing the positive charge of the prenyl zwitterion to facilitate the formation of the reactive azadiene intermediate. Stereocontrol seemed to be primarily driven by active site shape complementarity including aromatic amino acids and the cofactor, which hold the substrate in the proper conformation (**Figure 1.8**). The work on the Phq system was consistent with prior feeding studies and predictions, where the methyl-Pro-Trp dipeptide was formed by the NRPS.<sup>24, 27</sup> Disruption of *phqE* in *Penicillium simplicissimum* using the CRISPR-Cas9 system generated a mutant strain that accumulated the methyl-Pro-Trp prenyl zwitterion intermediate **1.12**.<sup>23</sup> With the addition of the PhqE structures in complex with the substrate, the product, and the NADP<sup>+</sup> cofactor, a more thorough mechanistic understanding was achieved. Preference for the NADPH cofactor is explained by an electrostatic interaction between Lys50 and the cofactor 2'-



**Figure 1.8** Reactions performed by MalC. Reduction and stereochemical control are performed by MKP IMDAase MalC which “holds” the substrate in the proper conformation to generate (+)-premalbrancheamide (**1.14**).



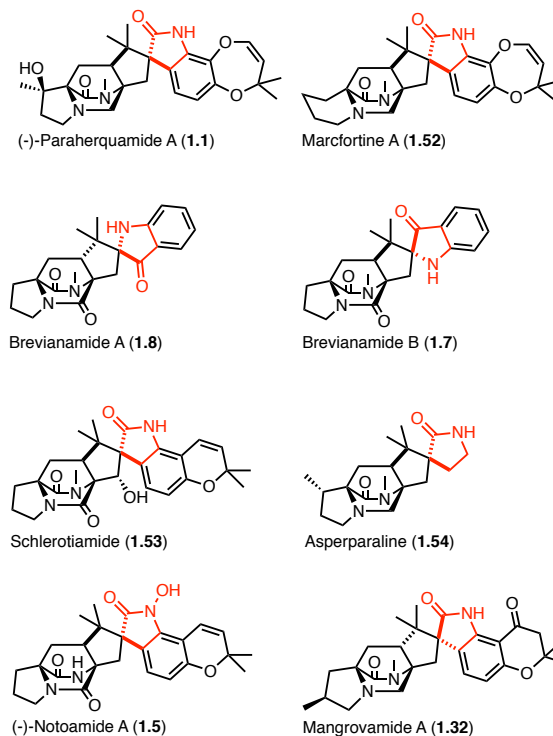
phosphate. Additionally, a relatively short distance between the nicotinamide C4 hydride donor and the presumed deoxy C5 acceptor of the product premalbrancheamide indicated that reduction and cycloaddition are highly coordinated.

Amino acid sequence comparisons, structural data, and NADPH cofactor dependence support the hypothesis that the Diels-Alderase catalyzing formation of the MKP bicyclo[2.2.2]diazaoctane ring system evolved from a short chain dehydrogenase progenitor. The active site pocket in the MKP systems clearly orients the substrate for proper enantiocontrol, while the evolution of the active site for its new function is evident. This spectacular display of protein evolution has led to an enzyme that maintained its reductase function without the canonical catalytic residues, while also providing stereocontrol for the Diels-Alder [4+2] cyclization. Moreover, our studies have revealed that the MKP and DKP pathways have evolved in a convergent manner based on the NRPS terminal domain identity, and redox chemistry leading to the dipeptide precursor ring system.

### 1.2.5 Flavin Monooxygenases (FMOs)

The spirooxindole center is prevalent within members of this indole alkaloid family including paraherquamides,<sup>3, 4</sup> notoamides,<sup>7-9</sup> marcfortines,<sup>50, 51</sup> schlerotiamide,<sup>52</sup> mangrovamides,<sup>29</sup> and asperparaline,<sup>53</sup> among others (**Figure 1.9**). The stereochemistry of the spirocycle is proposed to be controlled by the facial selectivity of the initial indole 2,3-epoxidation, followed by its collapse to a 2-alkoxyindole intermediate, and a pinacol-like rearrangement to generate the final spirocyclized product.<sup>54</sup>

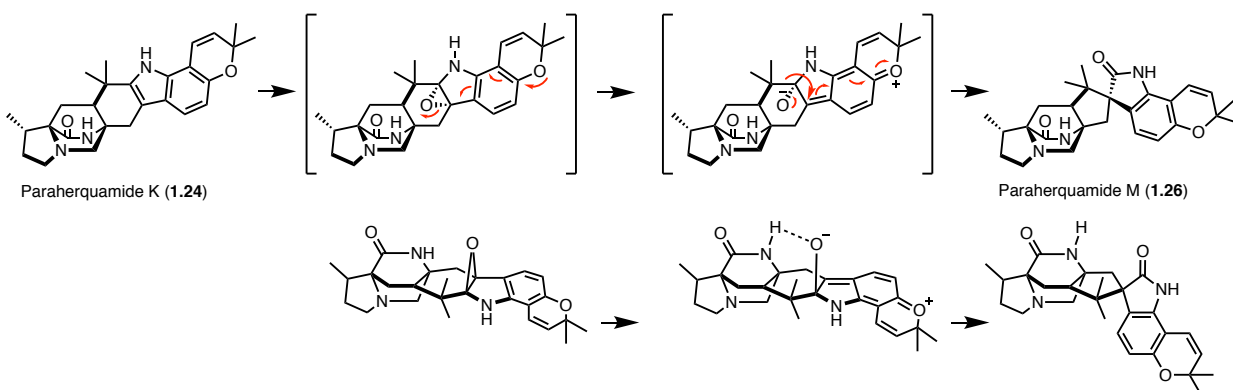
#### 1.2.5.1 Paraherquamides



**Figure 1.9** Representative spirocyclized fungal indole alkaloids. These molecules contain a spirooxindole form of the bicyclo[2.2.2]diazaoctane ring system.

The proposed precursor to the spirooxindole moiety found in the paraherquamides is preparaherquamide (**1.23**). Based on previous data,<sup>27</sup> it was hypothesized that the [2.2.2]bicyclic ring is formed with the non-oxidized tryptophanyl moiety and that oxidations of the indole ring to form both the catechol-derived dioxepin and the spirooxindole must occur after formation of **1.23** (Scheme 1.2).<sup>55</sup> To probe the timing of spirocyclization and dioxepin ring formation, triply deuterium-labelled 7-hydroxy-preparaherquamide was introduced to a culture of *Penicillium fellutanum*; the results indicated that indole C7 oxidation is not the immediate step following the IMDA reaction, but still may occur prior to spirooxindole formation.<sup>55</sup> Recent gene disruption studies in *Penicillium simplicissimum* using the CRISPR-Cas9 system established that the pyran and dioxepin rings are both formed prior to spirocyclization, elucidating two key intermediates *en route* to paraherquamides A and G.

The genetic knock-out data in *Penicillium simplicissimum* have also revealed that the epoxidation and semipinacol rearrangement occur through two parallel pathways. The biosynthetic pathway diverges after the formation of **1.23** to form either a pyran or a dioxepin ring on the indole (see below for a discussion of the proposed mechanism for assembly of the heptacyclic molecules). The PhqK FMO accepts both paraherquamide K (**1.24**) and L (**1.25**) and performs a facially-selective epoxidation, with a controlled collapse of the epoxide to form the spirooxindole (Scheme 1.4). Crystal structures revealed that substrate orientation in the PhqK active site determined the facial selectivity of epoxidation. Collapse of the presumed 2,3-indole epoxide intermediate is strongly influenced by the pyran/dioxepin oxygen, and the presence of a positively charged arginine. Based on Michaelis-Menten kinetic parameters, the dioxepin appears to be the favored substrate, but both paraherquamide K (**1.24**) and L (**1.25**) can undergo pinacol rearrangement to the spirooxindole product (Chapter 4).

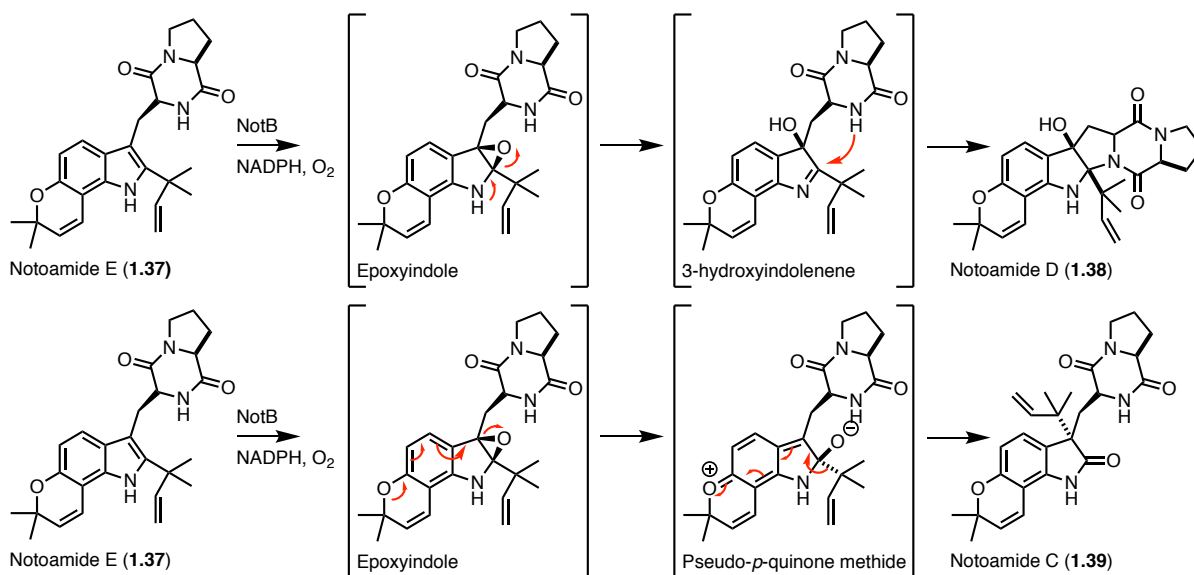


**Scheme 1.4** Proposed mechanism for paraherquamide spirooxindole formation.

### 1.2.5.2 Notoamides

The oxidative conversion of notoamide E (**1.37**) was investigated through biomimetic oxaziridine-based syntheses in which notoamide C (**1.39**)<sup>56</sup> (48%), 3-*epi*-notoamide C<sup>56</sup> (28%), and minor amounts of notoamide D (**1.38**) and 2,3-*epi*-notoamide D were produced (**Scheme 1.3**).<sup>8</sup> The electronic properties of the indole ring were found to influence the regiochemistry of oxidation of the C2–C3 indole bond, resulting in either the pyrroloindole or the oxindole.<sup>8</sup> While the synthetic method generated unequal product distribution, the organism produced almost equal amounts of **1.38** and **1.39**, implying that the biosynthetic enzyme can override the inherent reactivity of the molecule.<sup>8</sup>

Enzymatic characterization of the FMO NotB revealed both **1.38** and **1.39** can be produced from **1.37**, and it was proposed that both can be formed through  $\beta$ -epoxidation of the C2–C3 indole bond (**Scheme 1.5**).<sup>8</sup> Ring-opening of the  $\beta$ -2,3-epoxyindole intermediate to the 3-hydroxyindolenine species followed by N-C ring closure from the diketopiperazinyl NH generates pyrroloindole **1.38** as the major product. The minor product **1.39** is derived from the pseudoquinone methide species and subsequent  $\alpha$ -face migration of the prenyl group from C2 to C3. It was also proposed that an interaction between the indole C2 oxyanion intermediate and DKP N-H determines whether the pyran contributes to the breakdown of the epoxide (**Scheme 1.4** and **1.5**). The N-H can stabilize the negatively charged intermediate, but this interaction does not occur when the epoxide collapses to the indole C3 position, thus the hydroxyl is formed immediately, producing **1.38**.<sup>8</sup> The products resulting from  $\alpha$ -face epoxidation were not observed in the

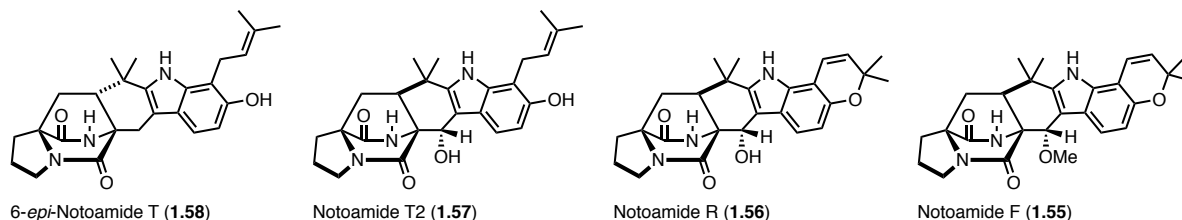


**Scheme 1.5** Proposed mechanism for assembly of notoamides C and D.

enzymatic reaction, thus it was proposed that their isolation in the precursor incorporation study was an artifact of the excessive amounts of **1.37** within the culture medium.<sup>56</sup>

Because **1.37** halted production of bicyclo[2.2.2]diazaoctane ring containing metabolites, it was proposed that another enzyme was required to catalyze formation of the spirooxindole within this family.<sup>8</sup> Intriguingly, it was reported that *A. protuberus* produced (+)-stephacidin A (**1.42**) and (-)-notoamide B (**1.44**),<sup>7</sup> while *A. amoenus* produced the enantiomers (-)-stephacidin A (**1.43**) and (+)-notoamide B (**1.45**) (Scheme 1.3).<sup>44</sup> Due to the antipodal relationship of the metabolites isolated from *A. protuberus* (marine) and *A. amoenus* (terrestrial), an *R*-selective monooxygenase was proposed to exist within the marine strain and an *S*-selective monooxygenase within the terrestrial strain.<sup>44</sup> At first, whether the alternative FMO would act before or after the IMDA remained unclear. It was considered that the breakdown of the indole 2,3-epoxide followed by pinacol rearrangement of the isoprenyl group from C2 to C3 could produce intermediates that undergo oxidation and tautomerization to yield the azadiene intermediates for the IMDA.<sup>32</sup> However, the isolation of notoamide T (**1.40** and **1.41**), stephacidin A (**1.42** and **1.43**),<sup>57</sup> and the precursor incorporation studies led to the biosynthetic hypothesis that the IMDA occurs prior to spirocyclization. Late-stage spirooxindole formation was probed by introducing doubly <sup>13</sup>C-labeled racemic stephacidin A to cultures of *A. amoenus* and *A. protuberus*; analysis of the metabolites revealed enantiospecific incorporation of intact **1.43** into **1.45** in *A. amoenus* and **1.42** into **1.44** in *A. protuberus*, and the unreacted **1.42** and **1.43** were re-isolated from the respective fungal extracts.<sup>57</sup> These data provided evidence of divergent flavoenzymes with opposite *R/S*-substrate selectivity in each strain.

In feeding studies, racemic D,L-[<sup>13</sup>C]<sub>2</sub>-notoamide T (**1.40** and **1.41**) was incorporated into **1.42** and **1.45** in *A. amoenus*,<sup>58</sup> indicating that **1.40** and **1.41** are both converted to **1.42** and **1.43** and **1.43** is rapidly converted to **1.45**, while **1.42** accumulates as a shunt product. In *A. protuberus* D,L-[<sup>13</sup>C]<sub>2</sub>-notoamide T (**1.40** and **1.41**) was incorporated into **1.42/1.43**, **1.44/1.45**, notoamide F



**Figure 1.10** Secondary metabolites with *anti* and *syn* configuration. The *syn* products (**1.57**, **1.56**, **1.55**) are all generated from notoamide T *in vivo*.

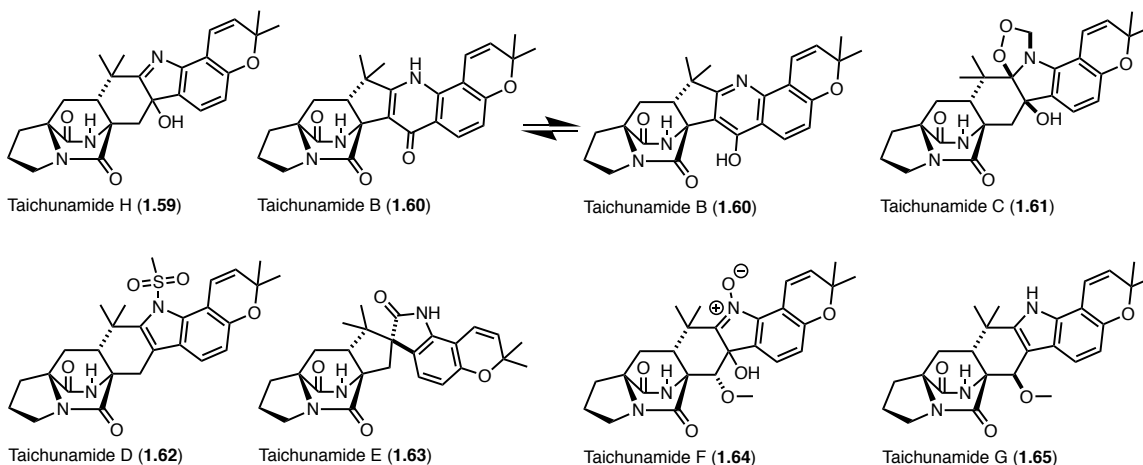
(**1.55**), notoamide R (**1.56**), and notoamide T2 (**1.57**) (all racemic) (**Scheme 1.3** and **Figure 1.10**).<sup>58</sup> This led to a contradictory notion that the gene products in *A. protuberus* do not discriminate between the two enantiomers of notoamide T (**1.40/1.41**) or stephacidin A (**1.42/1.43**) in the oxidative conversion to the final metabolites.<sup>58</sup>

Additionally, the presence of 6-*epi*-isomers raised some ambiguity in the selectivity of the IMDAase (**Figure 1.5**). (+)-Versicolamide B (**1.48**) was isolated from *A. amoenus* and was the first member of this family to possess the *anti*-relative stereochemistry within the bicyclo[2.2.2]diazaoctane ring system (**Figure 1.5**).<sup>44</sup> Racemic 6-*epi*-notoamide T (**1.58**) was converted to 6-*epi*-stephacidin A (**1.46**) and versicolamide B (**1.48**) in *A. protuberus*.<sup>45</sup> The isolation and incorporation of notoamide S (**1.36**) into notoamides C (**1.39**), D (**1.38**), (-)-stephacidin A (**1.43**), (+)-notoamide B (**1.45**), and (+)-versicolamide B (**1.48**) in *A. amoenus*<sup>59, 60</sup> indicated that the FMO can also accept (+)-6-*epi*-stephacidin A (**1.46**). The observation of an enantiomeric mixture of 6-*epi*-stephacidin A enriched with the (-)-isomer from *A. amoenus* indicated that the enzymes upstream of the FMO were not selective. While the *A. amoenus* FMO transforms **1.46** to **1.48**, there is not a suitable oxidase for the (-)-6-*epi*-stephacidin A (**1.47**) shunt metabolite.<sup>45</sup> It is intriguing that the closely related *Aspergillus* species have evolved enantiodivergent pathways to the stephacidins and notoamides, but converged on the production of **1.48**.<sup>45</sup>

*In vitro* work with NotI from *A. protuberus* and NotI' from *A. amoenus* demonstrated the conversion of **1.42** and **1.43** to **1.44** and **1.45**, although a clear preference was observed for **1.43** (**Scheme 1.3**). Both NotI and NotI' converted **1.46** to **1.48** but no reaction was observed with **1.47**. This is consistent with the conversion observed in *A. amoenus* where **1.48** was produced and **1.47** was determined to be a shunt metabolite.<sup>45</sup> The ability to convert **1.42** may be an evolved trait from an ancestral enzyme previously selective for **1.43**, as this reactivity is retained in both NotI and NotI'.

### 1.2.5.3 Taichunamides

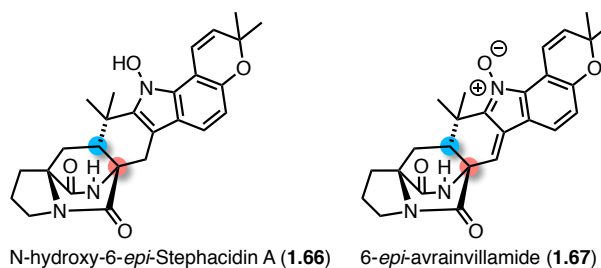
The isolation of fungal natural products with the *anti*-relative stereochemistry of the bicyclic ring generated questions regarding the driving force for selectivity of the IMDA reaction in these families. Taichunamide H (**1.59**), isolated from mangrove-derived fungus *Aspergillus versicolor* HDN11-84, and taichunamides B-G (**1.60-1.65**), isolated from the fungus *Aspergillus taichungensis* (IBT 19404), contain an *anti*-bicyclo[2.2.2]diazaoctane core (**Figure 1.11**).<sup>61, 62</sup>



**Figure 1.11** Taichunamides with *anti*- relative stereochemistry of the bicyclic ring.

They are proposed to share the common precursors notoamide S (1.36), (+)-6-*epi*-notoamide T (1.58), and (+)-6-*epi*-stephacidin A (1.46). Additionally, *A. taichungensis* ZHN-7-07 produces 1.46, N-hydroxy-6-*epi*-stephacidin A, and 6-*epi*-avrainvillamide, all with *anti*- relative configuration (Figure 1.12).<sup>63</sup> This indicates

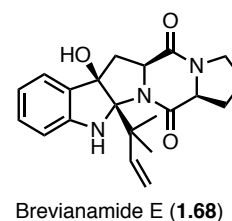
that the taichunamide biosynthetic enzymes share some overlap or crosstalk with those from the notoamide biosynthetic pathway in these producing organisms, but in all cases they seem to be selective for the *anti*-form of the bicyclic ring configuration.



**Figure 1.12** *Anti* bridged bicycles.

#### 1.2.5.4 Brevianamides

The *anti*-relative stereochemistry has been proposed to be generated from the oxindole intermediate. [<sup>3</sup>H]-labeled deoxybrevianamide E (1.34)(Scheme 1.3) was shown to be efficiently incorporated into brevianamides A (1.8), B (1.7) (Figure 1.9), and E (1.68) (Figure 1.13) It was proposed that 1.68 may be a biosynthetic precursor to 1.7 and 1.8, but [<sup>3</sup>H]-labeled 1.68 was not incorporated into either, disproving its intermediacy within the pathway.<sup>46</sup> Thus, it is likely that oxidation of the indole occurs at C3 prior to the IMDA. An R-selective hydroxylation is proposed to generate the 3-hydroxyindolenine which can either cyclize to form 1.68 or undergo a pinacol-type rearrangement to form the C2-prenylated intermediate. Oxidation to the azadiene would provide the proper substrate for the IMDA.<sup>46</sup>

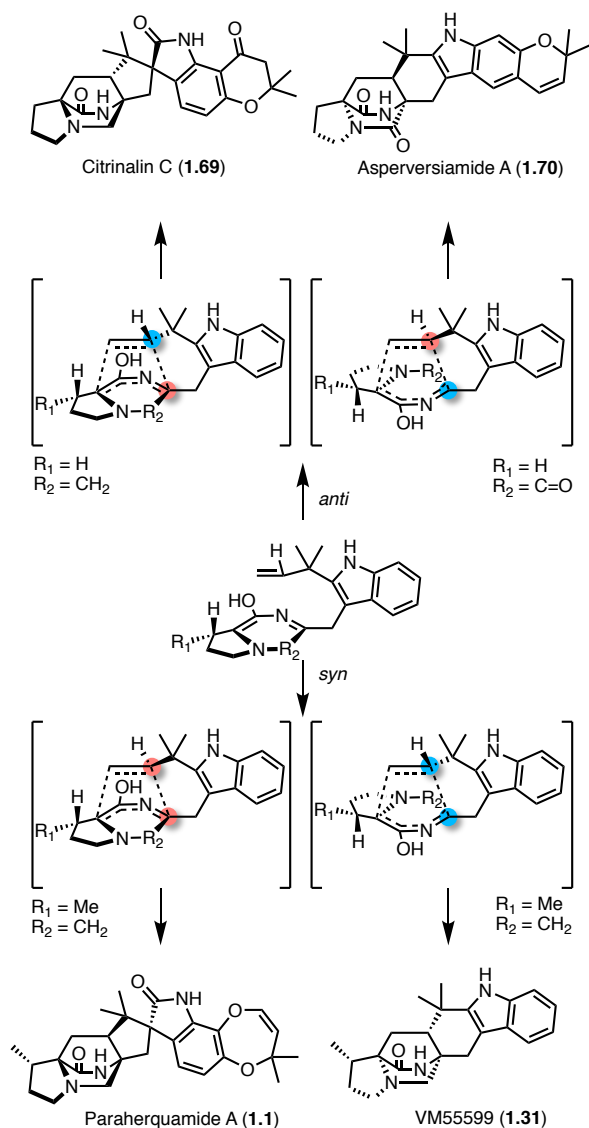


**Figure 1.13** Brevianamide E which is generated from deoxybrevianamide E.

### 1.2.5.5 Effect on the Diels-Alder Product

With the knowledge that the spirooxindole is differentially installed preceding or following the IMDA reaction in various strains, the implications of this distinction regarding other reactions within the biosynthetic sequence must be considered. Density functional theory computational studies have been performed to elucidate the stereochemical features of the cyclization step involved in paraherquamide A (**1.1**) and VM55599

(**1.31**) biosynthesis (**Figure 1.14**).<sup>64</sup> It was determined that the non-oxidized indole favors *syn* diastereoselectivity. This agrees with precursor incorporation experiments by Williams et al.<sup>31, 46</sup> concluding that the *syn* cycloaddition could take place via the nonoxidized tryptophanyl moiety and that the oxidation of the indole ring occurs after the IMDA cycloaddition in paraherquamide biosynthesis.<sup>64</sup> In the computational study, *anti*-selectivity was found for the cycloadditions of the oxindole-based derivatives, indicating that the FMO reaction could precede the IMDA in cases where *anti*-bicyclo[2.2.2]diazaoctane rings are observed.<sup>64</sup> Contrary to this hypothesis, nonspirocyclized *anti*-MKP members of this family have also been isolated,<sup>65-67</sup> indicating that there may be various divergent routes to the *syn/anti*-bicyclo[2.2.2]diazaoctane ring. Monoxygenases have evolved to perform the presumed epoxidation and pinacol rearrangement at different points within the biosynthesis, affecting the stereochemical outcome of the Diels-Alder reaction.<sup>46</sup> While the computational study presents a substrate-controlled mechanism for directing the IMDA, the enzyme active site conformation may



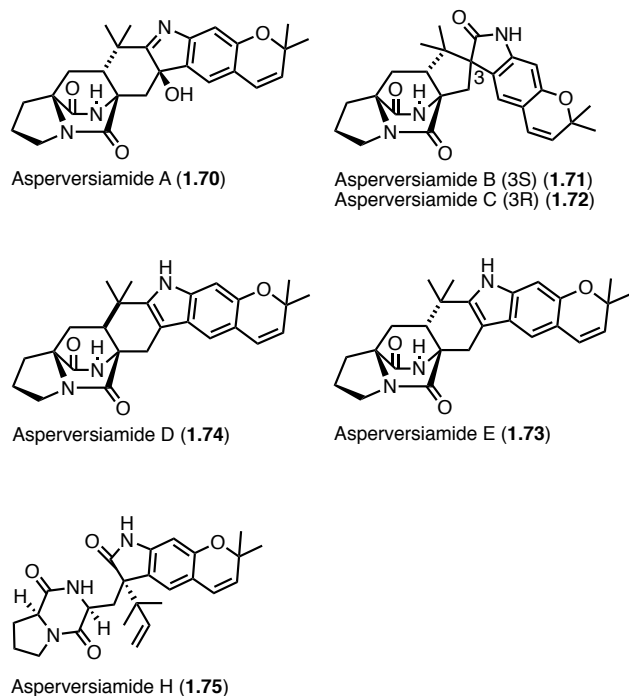
**Figure 1.14.** Conformations conducive to *syn* or *anti* bicyclic ring formation. Examples from each subgroup of the fungal indole alkaloids are shown.

provide a catalyst-controlled means to overcome the innate reactivity. With this in mind, IMDA enzymes may be selective for either the oxidized or non-oxidized substrate, and the orientation in which they bind may direct the reaction. If this is the case, there may have been a parallel evolution of the IMDA enzyme and the flavin monooxygenase (FMO) in each of the biosynthetic gene clusters.

Interestingly, in the case of the aspersiamides, both *syn* and *anti* products have been isolated from the same organism. While aspersiamides A-C (1.70-1.72) and E (1.73) contain the *anti*-bicyclo[2.2.2]diazaoctane ring, aspersiamide D (1.74) contains the *syn*-ring (Figure 1.15). With the isolation of the prenylated precursor aspersiamide H (1.75), one can infer that the Diels-Alder reaction is the last step in the pathway to 1.70-1.73, following the oxindole formation, leading to an *anti*-configuration.<sup>43</sup> This system is unique in that it must contain two different monooxygenases since one would result in the pinacol rearrangement and the other would generate a hydroxylated product.<sup>43</sup> Additionally, one FMO would act at an early stage, and the other at a late stage with various options for the IMDA. The IMDA enzyme may be non-selective and accept both substrates, or it may have an evolved active site that can override the inherent selectivity of the molecule. Among numerous possibilities, there may even be two different IMDA enzymes, serving to generate the diastereomeric products.

1.2.5.6 FMO Evolution

There are a variety of different types of FMOs that have diverged into two groups defined by whether they require a reductase partner.<sup>68</sup> For standalone enzymes, such as those involved in spirocyclization, the reactions with the electron donor and oxygen are catalyzed by a single protein. Over time, the flavin monooxygenases have undergone domain fusion events, eventually leading



**Figure 1.15** Aspersiamides A-E. Aspersiamides A-C and E display *anti*-relative stereochemistry and aspersiamide D displays *syn*-relative stereochemistry.

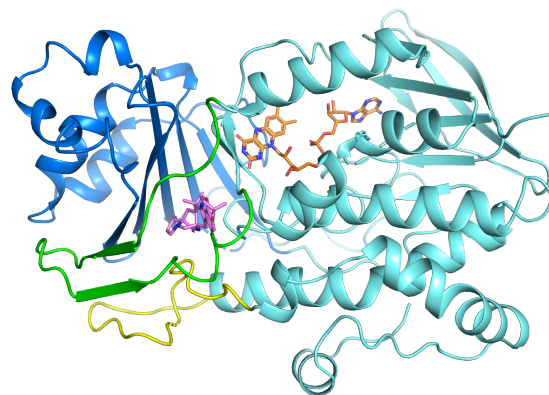


to the development of a new domain for substrate binding (**Figure 1.16**). While some FMOs have two distinct cofactor binding domains, others (such as PhqK) contain a single nucleotide binding domain for the FAD, and have developed a NADPH binding groove on the surface of the protein.

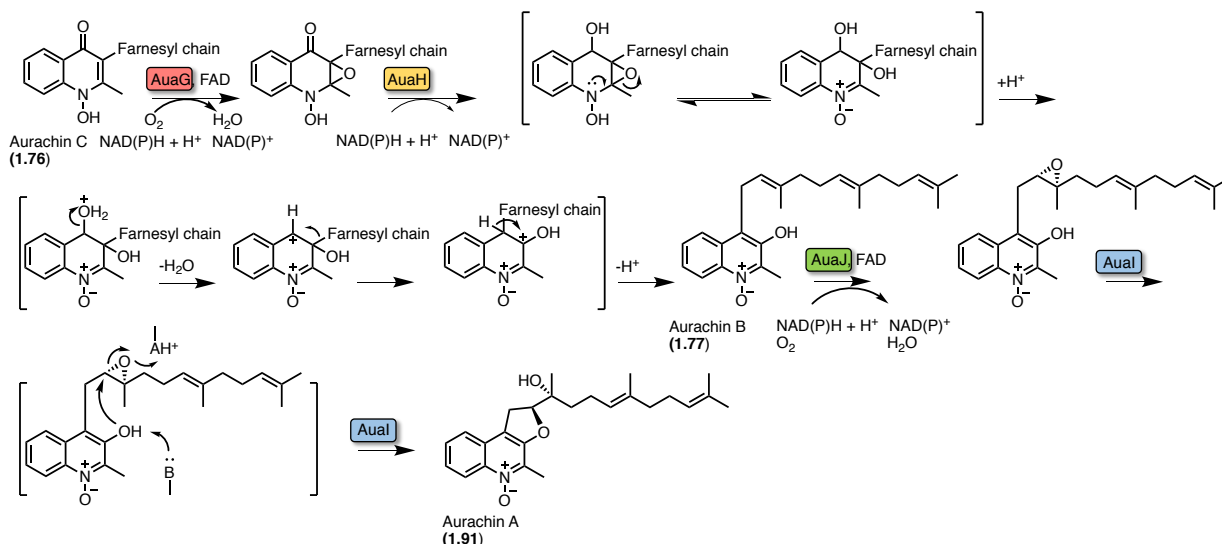
#### 1.2.5.7 Mechanism of Spirooxindole Formation

Mechanistic implications for spirooxindole formation can be gleaned from systems outside the bicyclo[2.2.2]diazaoctane ring containing family. The conversion of aurachin C (**1.76**) to B (**1.77**) is proposed to involve the migration of the prenyl group from position C3 to C4 through a

pinacol-type rearrangement,<sup>21</sup> similar to that proposed for the spirocyclization in the fungal indole alkaloids (**Scheme 1.6**). The proposed mechanism involves an epoxidation of the quinoline core of **1.76** by AuaG followed by AuaH-catalyzed ring opening of the epoxide and semipinacol rearrangement, while the proposed mechanism for the fungal indole alkaloids is initiated with an epoxidation of the indole C2=C3, and the breakdown of the epoxide leads to a spontaneous pinacol



**Figure 1.16** Domains of PhqK FMO. PhqK FAD-binding domain (cyan) and three insertions that form the substrate-binding domain (blue, green, yellow) are highlighted. The substrate is shown in magenta and the FAD cofactor is shown in orange.

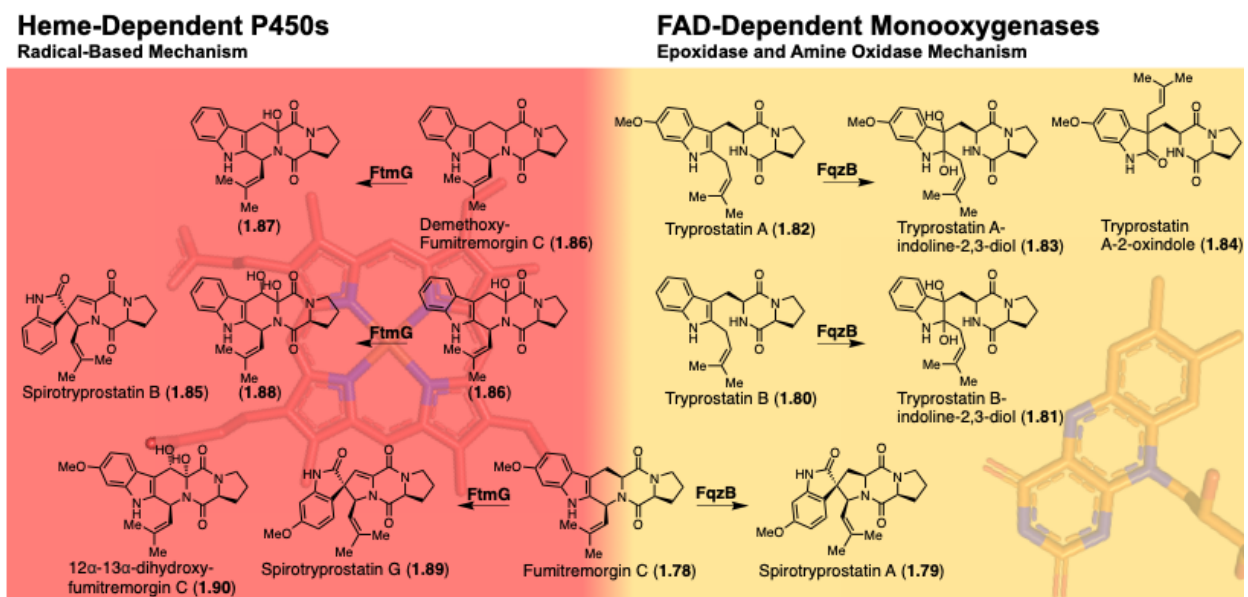


**Scheme 1.6** Aurachin biosynthesis. AuaG/AuaH performs epoxidation and semi-pinacol rearrangement like PhqK, and AuaJ/AuaI catalyzes epoxidation and cyclization like the pyran and dioxepin formation in fungal indole alkaloid biosynthesis.

rearrangement to generate the spirocycle. In the aurachin pathway two enzymes catalyze the pinacol-type rearrangement where the main product of AuaG appears to be an energetically less-favored isomer that is immediately stabilized by reduction and subsequent aromatization catalyzed by AuaH reductase.<sup>21</sup> In contrast, the fungal indole alkaloid systems have evolved a single enzyme to perform both functions.

A detailed study of spiro-ring formation in spirotryprostatins derived from the fumitremorgin biosynthetic pathway has also shed light on this topic (**Figure 1.17**).<sup>19</sup> FqzB is an FAD-dependent monooxygenase from the unrelated fumiquinazoline biosynthetic pathway which catalyzes the spirocycle formation on fumitremorgin C (**1.78**) to generate spirotryprostatin A (**1.79**) via an epoxidation route (**Figure 1.17**).<sup>19</sup> The conversion of tryprostatin B (**1.80**) to tryprostatin B-indoline-2,3-diol (**1.81**) and conversion of tryprostatin A (**1.82**) to tryprostatin A-indoline-2,3-diol (**1.83**) and tryprostatin A-2-oxindole (**1.84**) indicates that FqzB catalyzes epoxidation of the C2-C3 olefin of the indole on tryprostatin A and B. The formation of **1.84** indicates that the mechanism involves a semi-pinacol-type rearrangement of the epoxidized intermediate.<sup>19</sup>

Within the same organism, a cytochrome P450 FtmG from the fumitremorgin biosynthetic pathway catalyzes the spiro-ring formation observed in spirotryprostatin B (**1.85**).<sup>19</sup> FtmG converts 21renylate-fumitremorgin C (**1.86**) into spirotryprostatin B (**1.85**) as well as mono-hydroxyl (**1.87**) and diol (**1.88**) forms of 21renylate-fumitremorgin C.<sup>19</sup> Based on the intermediate structures, a



**Figure 1.17** Spirotryprostatin and fumitremorgin biosynthetic pathway crosstalk.

reaction mechanism was proposed involving a P450 heme-catalyzed initial radical formation and subsequent two rounds of hydroxylation of **1.86**.<sup>19</sup> The production of **1.85** by FtmG is proposed to occur through radical migration from C8 to C2 and subsequent hydroxylation at C2 can set up the molecule for a semipinacol-type rearrangement involving a concomitant spiro-ring formation with dehydration to yield **1.84**. Additionally, FtmG crosses into the fumitremorgin biosynthetic pathway to convert fumitremorgin C to spirotryprostatin G (**1.89**) (**Figure 1.17**).<sup>19</sup>

*A. fumigatus* maintains two orthogonal spiro-ring formation systems in its secondary metabolite biosynthetic pathway: an FAD-dependent route (for the formation of **1.79**) catalyzed by the FMO FqzB and a radical route (for the formation of **1.85** and **1.89**) catalyzed by the P450 FtmG.<sup>19</sup> The fumiquinazoline FMO FqzB crosses into the fumitremorgin biosynthetic pathway to generate **1.79** whose olefin adjacent to the spirocenter is fully reduced (biosynthetic pathway crosstalk).<sup>19</sup> This study highlights the versatile role of oxygenating enzymes in the biosynthesis of structurally complex natural products and indicates that cross-talk of different biosynthetic pathways promotes product diversification in natural product biosynthesis.<sup>19</sup>

#### 1.2.6 Dioxepin and Pyran Ring Formation

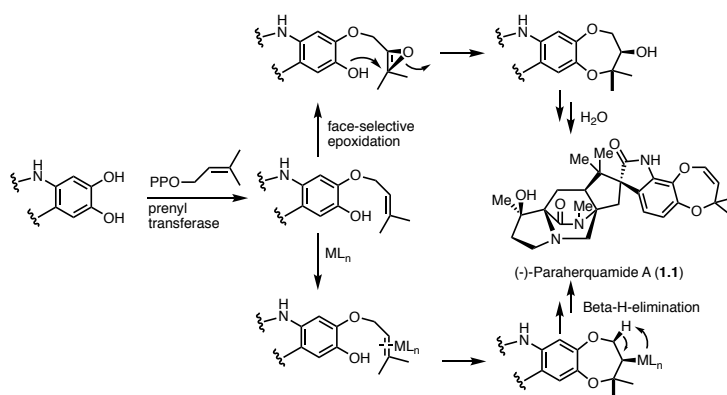
Within *P. fellutanum*, isoprene units from the mevalonic acid (mevalonate) pathway were incorporated into the dioxepin ring of paraherquamide A (**1.1**).<sup>31</sup> It was proposed that the formation of the dioxepin is due to a facially selective epoxidation of the olefin followed by a stereospecific ring-opening of the epoxide and dehydration.<sup>31</sup> Another alternative for the dioxepin formation would be face-selective complex with a transition metalloprotein to the olefinic  $\pi$ -system followed by stereospecific intramolecular nucleophilic addition and reductive elimination to the enol ether (**Scheme 1.7**).<sup>31</sup> The hydroxylation and epoxidation are both proposed to be performed by currently unidentified oxidative enzymes. This fungal system contrasts with the non-heme iron dependent enzyme deguelin cyclase found in plants.<sup>69, 70</sup>

The aspersiamides have a unique connectivity with the isoprenyl substituent located on the indole C5 to form the pyrano-[3,2-*f*]indole moiety as opposed to the angularly fused pyrano[2,3-*g*]indole (**Figures 1.14** and **1.15**).<sup>43</sup> The prenyltransferase responsible for this transformation must have C5 selectivity as opposed to typically observed indole C7 prenylation. Consequently, the downstream enzymes, such as those responsible for catalyzing pyran ring formation must have evolved in parallel to perform their function on the new metabolite.

Within the notoamide-producing strain *A. amoenus*, [ $^{13}\text{C}$ ]<sub>2</sub>-notoamide T (**1.40** and **1.41**) was incorporated into (+)-stephacidin A (**1.42**) and (+)-notoamide B (**1.45**) (**Scheme 1.3**).<sup>58</sup> This indicates that both **1.42** and **1.43** were accepted by the pyran ring forming enzyme to generate **1.42** and **1.43**, respectively. Thus the enzyme responsible for pyran formation must have evolved to accept both enantiomers.

### 1.2.6.1 Mechanism of Pyran and Dioxepin Ring Formation

Similar biosynthetic enzymes involved in aurachin production in the myxobacterium *Stigmatella aurantiaca* Sg a15 have provided insight into the potential mechanism to form the pyran and dioxepin moieties. This involved the 3,4-migration of the farnesyl chain to produce A-type aurachins which was proposed to proceed via 2,3-indole

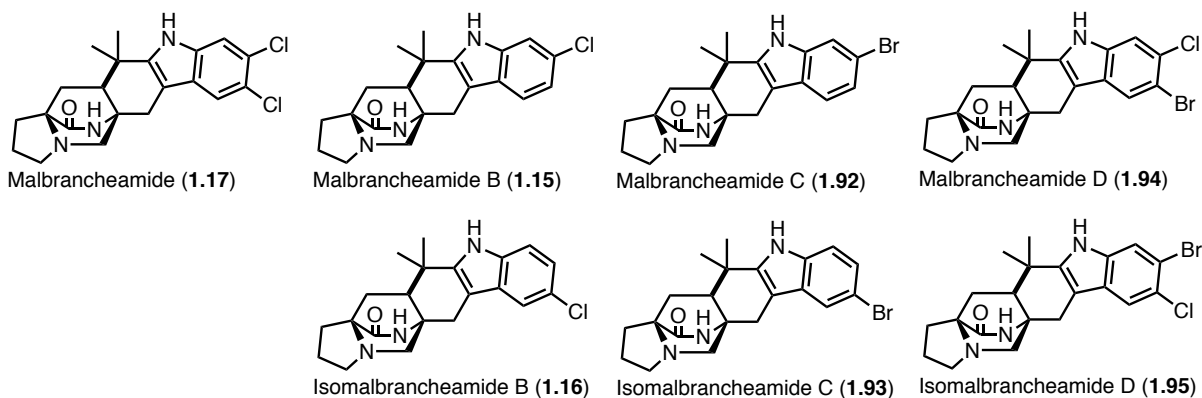


**Scheme 1.7** Proposed mechanisms for dioxepin ring formation. This is performed by an oxidative enzyme with a metal cofactor ( $\text{ML}_n$ ) and a prenyltransferase.

epoxidation (AuaG) and subsequent reduction by AuaH as described above (**Scheme 1.6**).<sup>20</sup> AuaJ is proposed to perform an epoxidation that is followed by AuaI-catalyzed hydrolysis to form aurachin A (**1.91**), which functions similar to the pyran and dioxepin ring forming enzymes within fungal indole alkaloid biosynthesis.<sup>20</sup> In this case, two enzymes (AuaG and AuaH) perform the proposed function of PhqK and NotI in paraherquamide and notoamide biosynthesis, respectively. PhqK and NotI have evolved as dual-function enzymes to perform the epoxidation and selective pinacol-type rearrangement in their respective pathways. It is possible that the epoxidation and cyclization to form the pyran and dioxepin rings may also be mediated by a single fungal enzyme, but the mechanism remains unclear (**Scheme 1.7**).

### 1.2.7 Late-Stage Halogenation

Halogenation of the fungal bicyclo[2.2.2.] indole alkaloids appears to be specific to the *Malbranchea* strains *M. aurantiaca*<sup>5</sup> and *M. graminicola*.<sup>71</sup> The halogenases MalA (*M. aurantiaca*) and MalA' (*M. graminicola*) are responsible for the iterative chlorination to produce the natural product malbrancheamide (**1.17**).<sup>5, 71, 72</sup> These halogenases catalyze chlorination and bromination reactions to generate monochlorinated malbrancheamide B (**1.15**) and



**Figure 1.18** Halogenated malbrancheamides. These molecules are generated by the halogenases MalA and MalA'.

isomalbrancheamide B (1.16), monobrominated malbrancheamide C (1.92) and isomalbrancheamide C (1.93), and the mixed halogen malbrancheamide D (1.94) and isomalbrancheamide D (1.95) (Figure 1.18). The monobrominated molecules had previously been isolated from *M. graminicola*, and *in vitro* work with MalA and MalA' demonstrated that both strains were capable of generating the brominated and chlorinated molecules. The MalA/A' study included the first computational analysis of the flavin-dependent halogenase (FDH) family, and demonstrated a modified mechanism for fungal FDHs. Additionally, it was determined that MalA is stereospecific, distinguishing a single antipodal substrate. Thus, when incubated with racemic premalbrancheamide, only (+)-premalbrancheamide (1.14) was converted to (+)-malbrancheamide (1.17).<sup>23</sup>

The notoamide family also contains a halogenated secondary metabolite, notoamide N, which contains an indole C5 chlorination (Figure 1.6).<sup>32</sup> Based on what we have observed with the production of spiromalbramide,<sup>71</sup> and in reactions with NotI/PhqK, we can hypothesize that the pyran ring formation and the halogenation of the [2.2.2] bicyclic product occur prior to the spirocyclization. The molecule could be halogenated first through a similar mechanism to MalA, then the pyran ring would be formed, followed by spirocyclization. However, the corresponding halogenating enzyme has yet to be identified.

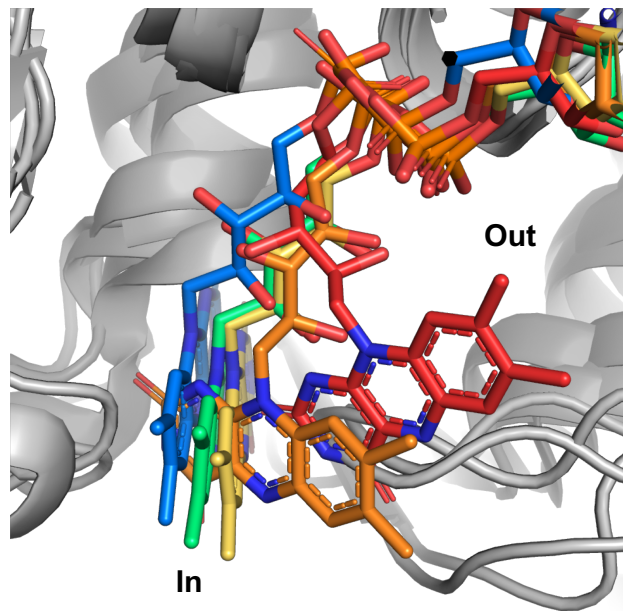
The spirooxindole-forming enzymes and FDHs are both part of the overarching FMO family. While the enzymes that catalyze spirooxindole formation are typically discrete proteins, the FDHs require a FAD reductase partner to initiate catalysis. The stand-alone enzymes have a highly dynamic FAD cofactor that moves between “in” and “out” conformation as it is reduced and oxidized, respectively. After a structure-based evolutionary analysis of FMOs, it was clear that

the halogenases have evolved to bind the FAD in a conformation that resembles the “in” conformation of the stand-alone enzymes. Along this evolutionary trajectory, the enzyme also refashioned the FAD redox chemistry to produce hypohalous acid and direct this halogenating agent through the production of a lysine haloamine intermediate (**Figure 1.19**).

### 1.2.8 Conclusion

The bicyclo[2.2.2]diazaoctane ring containing fungal indole alkaloids are a unique family of molecules and biosynthetic enzymes. The NRPSs bear terminal condensation or reductase domains, which lead to the formation of diketopiperazine or monoketopiperazine bicyclic ring systems, respectively. Additionally, the incorporation of  $\beta$ -methyl proline and  $\gamma$ -methyl proline has been accommodated through enzymes which have evolved to assemble them from isoleucine and leucine, respectively. Following the NRPS reaction, a selective reverse prenylation of the indole C2 position occurs.

The IMDA enzyme performs a central role to form the bicyclic core common to this class of molecules. Initially annotated as a short chain dehydrogenase, the MKP IMDAases MalC and PhqE perform an initial reduction prior to controlling the selective [4+2] cycloaddition reaction. This is an example of the only characterized Diels-Alder enzymes that retain their ancestral functionality.<sup>40, 73</sup> Interestingly, DFT calculations suggest that *syn-/anti-* diastereospecificity of the IMDA reaction can be substrate-driven based on the oxidation state of the indole ring, where presence of the FMO-derived spirooxindole moiety reverses the intrinsic diastereospecificity. Natural examples of pre- and post-IMDA spirocyclization have been identified in these pathways. Lastly, halogenation is performed by a flavin-dependent halogenase, which has only been identified in the *Mal* gene clusters. In each of the aforementioned diversification mechanisms, a



**Figure 1.19** Structural comparison of FMOs. Some FMOs catalyze oxidation reactions and some FMOs catalyze halogenation reactions. PhqK (red), PqsL FMO, PDB: 6fho (orange), MalA halogenase, PDB: 5wgr (yellow), PltA halogenase, PDB: 5dbj (green), and decarboxylative hydroxylase, PDB: 5eow (blue).

unique family of molecules with intriguing bioactivities has been generated. This pool of biosynthetic knowledge, which our group and others have generated, will contribute to our ability to recombine these enzymes and utilize nonnatural substrates to reach the full potential of this class of broadly bioactive molecules.

### **1.3 Thesis overview**

My doctoral thesis work has focused on the elucidation of key steps in the biosynthesis of therapeutically relevant fungal indole alkaloids. A majority of my efforts were devoted to characterizing late-stage tailoring steps in the biosynthesis of malbrancheamide and paraherquamide. Biochemical and structural information regarding the halogenation and spirooxindole formation led us to believe that these enzymes have evolved unique selectivities to generate their respective secondary metabolites. Through in depth structural and mechanistic characterization, my work has built a foundation for future development of these enzymes as biocatalysts for late-stage transformations.

I was also part of a collaborative effort to elucidate the mechanism behind the stereocontrolled [4+2] IMDA cyclization that generates the bicyclo[2.2.2]diazaoctane ring. This work involved the development of a chiral separation for the enantiomers of *syn*-premalbrancheamide, leading to the discovery that the IMDA enzyme controls the cyclization to produce only (+)-premalbrancheamide. Our collaborative efforts probed the intricate details of MalC/PhqE to identify catalytic amino acids and a proposed reaction mechanism.

After the IMDA cyclization, (+)-premalbrancheamide is halogenated by a flavin-dependent halogenase, MalA, which I thoroughly characterized through biochemical methods and structural studies. MalA was found to both chlorinate and brominate the natural substrate at two positions, and structures complexed with the natural substrates (+)-premalbrancheamide, (+)-malbrancheamide B, and (+)-isomalbrancheamide B were utilized in computational studies in collaboration with the Houk group at UCLA. We used a structure-based approach to generate the first site-selective mutants of this enzyme, and established a new reaction mechanism for fungal flavin-dependent halogenases, which relies on a serine residue in the active site to direct halogenation to the inherently less reactive carbon. This mechanism allowed for further enzyme engineering, which generated around twenty variants with broad site-selectivity. We then initiated a collaboration with Novartis Institutes for BioMedical Research, where we found that MalA and its variants are suitable biocatalysts for late-stage halogenation of a broad range of substrates. This

work has served to develop an environmentally friendly biocatalyst for late-stage halogenation, while also demonstrating the characterization of the final step of the malbrancheamide biosynthetic pathway.

My second major project involved the characterization of the enzyme responsible for forming the spirooxindole moiety in the paraherquamide biosynthetic pathway. Collaborative efforts within the Sherman group at the University of Michigan identified the natural substrates for the spirocycle-forming enzyme. My work focused on the *in vitro* characterization of the flavin monooxygenase responsible for this transformation (PhqK). I found that PhqK has two native substrates, paraherquamide K and L, and that it stems from an earlier divergence within this pathway to generate molecules containing either a pyran or dioxepin ring appended to the indole. PhqK and its downstream enzymes have evolved to accept substrates with both ring systems to generate the secondary metabolites paraherquamides A and G. I demonstrated the substrate flexibility of PhqK in reactions with halogenated malbrancheamide analogs, and through the solution of numerous cocrystal structures and one structure with no ligand bound, we can visualize how the active site of this enzyme changes to accommodate the various molecules.

Lastly, I was fortunate to be a part of a very innovative project supported by the LSI Cubed grant. I collaborated with a graduate student in the Chemical Biology Program, Tyler Beyett, to investigate the therapeutic potential of malbrancheamide in targeting protein-protein interactions between calmodulin and G-protein-coupled receptor kinase 5 (GRK5). We found that malbrancheamide selectively inhibits only the C-terminus of calmodulin, leaving the N-terminus of the protein free to interact with GRK5. This selective inhibition halts autophosphorylation and nuclear translocation, decreasing the effect of cardiac hypertrophy. After our initial *in vitro* studies, we tested malbrancheamide in cells in collaboration with the Koch lab at Temple University. We found that malbrancheamide decreased the swelling of cardiac tissue that would normally swell under induction with hypertrophic agonist phenylephrine. Not only were we able to identify the intricate mechanism of action of this halogenated natural product *in vitro*, but we demonstrated its utility in cardiac tissue.



## Chapter 2\*

# Characterization of the Bifunctional Reductase/Diels-Alderase Responsible for the Formation of the Bicyclo[2.2.2]diazaoctane Ring

### 2.1 Introduction

Fungal indole alkaloids containing the bicyclo[2.2.2]diazaoctane core have attracted considerable interest due to their wide spectrum of biological activities and their compelling biosynthetic chemistry. The formation of the bicycle has been of considerable interest because it is proposed to arise from an intramolecular [4+2] Diels-Alder (IMDA) reaction, a well-known synthetic transformation which is rarely found in natural product biosynthesis.<sup>35</sup> There are two distinct groups within this class of molecules composed of the diketopiperazines (DKP) and the monoketopiperazines (MKP), each of which is proposed to be generated through a divergent mechanism. These groups are also well-correlated with the domain organization of the upstream NRPS, where they have either a terminal condensation or reductase domain. The offloaded product of the reductase domain is a reduced dipeptide such as Pro-Trp in the malbrancheamide biosynthetic pathway.<sup>2</sup> The reduced Pro-Trp intermediate (**1.9**) is then reverse prenylated at the indole C2 position followed by the IMDA reaction.<sup>22</sup> This cycloaddition is stereospecific based on the relative position of the diene and the dienophile. While variation in stereocontrol has been observed in some strains, only (+)-malbrancheamide (**1.17**) has been isolated from *Malbranchea aurantiaca*, indicating strict diastereo- and enantioselectivity of the biosynthetic IMDA.<sup>5</sup> We were intrigued by the enzymatic control of this reaction, thus we initiated an investigation of the enzyme responsible for this transformation.

The existing Diels-Alderases all seem to have evolved from divergent ancestors to perform their new functions.<sup>36-42</sup> Among the progenitors are *S*-adenosyl-L-methionine (SAM)-dependent SpnF<sup>36, 74</sup> and LepI,<sup>40, 75</sup>  $\beta$ -barrel containing PyrI4<sup>76</sup> and AbyU,<sup>77</sup> and the flavin-dependent PyrE3.<sup>73</sup>

---

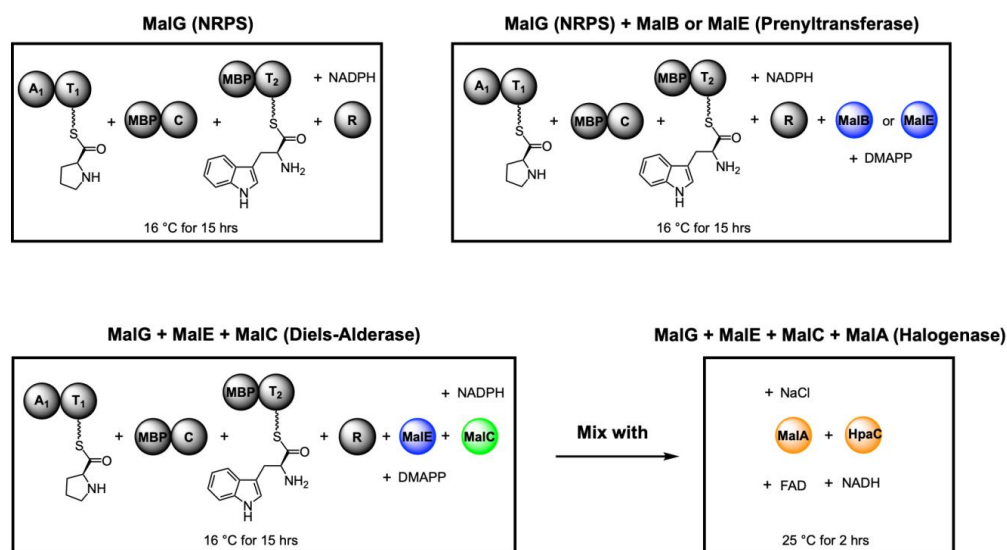
\*Chapter 2 is included in a manuscript: Dan, Q.; Newmister, S. A.; Klas, K.; Fraley, A. E.; McAfoos, T. J.; Somoza, A. D.; Sunderhaus, J. D.; Ye, Y.; Shende, V. V.; Yu, F.; Sanders, J. N.; Brown, W. C.; Zhao, L.; Paton, R. S.; Houk, K. N.; Smith, J. L.; Sherman, D. H.; Williams, R. M. Fungal Indole Alkaloid Biogenesis Through Evolution of a Bifunctional Reductase/Diels-Alderase. **Under Review**.

While some of these maintain the bound cofactor, it no longer serves its traditional role, and instead facilitates a catalytically competent conformation for the regio- and stereoselective [4+2] pericyclic reaction. Since most of the IMDA reactions are spontaneous, the sole responsibility of these enzymes is to hold the substrate in the proper orientation for the selective cyclization, and until recently, no catalytic residues had ever been identified. Through a collaborative effort, we investigated the putative IMDA mechanisms in malbranchamide and paraherquamide biosynthesis,<sup>23</sup> and identified the bifunctional reductase/Diels-Alderase enzyme responsible for the stereocontrolled construction of the MKP bicyclic core. Consequently, we demonstrated that this is the first IMDA enzyme to retain its ancestral short-chain dehydrogenase (SDR) function by utilizing a bound NADPH cofactor for reduction and the oxidized NADP<sup>+</sup> for maintaining proper stereocontrol.

## 2.2 In Vitro Reconstruction of the Malbranchamide Biosynthetic Pathway

### 2.2.1 Nonribosomal Peptide Synthetase

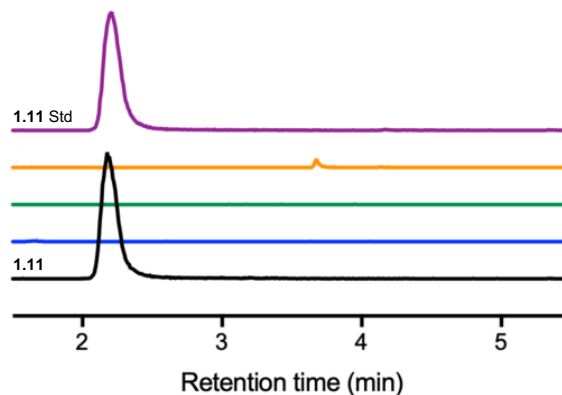
The first step in decoding the intricate details of the malbranchamide biosynthetic pathway involved the tedious process of obtaining the full NRPS enzyme. We hypothesized the first step of malbranchamide biosynthesis involved the coupling of L-proline and L-tryptophan by



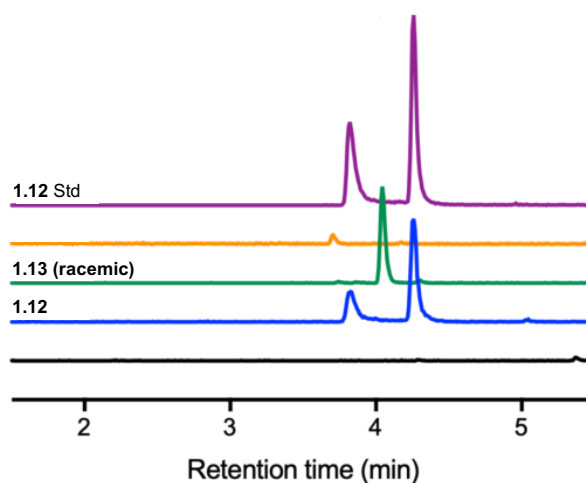
**Figure 2.1** Scheme of *in vitro* reconstitution assays. The flavin-dependent MalA halogenase requires an exogenous FAD reductase to reduce FAD to FADH<sub>2</sub> after each catalytic cycle, here HpaC reductase + NADH.<sup>1</sup> MBP = maltose binding protein.

MalG, a dimodular NRPS containing six domains ( $A_1$ - $T_1$ - $C$ - $A_2$ - $T_2$ - $R$ , **Scheme 1.1**), to produce L-Pro-L-Trp aldehyde through reductive off-loading. Qingyun Dan in the Smith group at the University of Michigan successfully expressed and purified individual components of the NRPS. Since the full-length protein could not be produced in soluble form, Dan identified domain boundaries in MalG, developed expression constructs for the excised  $A_1$ - $T_1$ ,  $C$ ,  $T_2$  and  $R$  domains, and loaded the putative amino acid substrates onto the MalG  $T_1$  and  $T_2$  domains (**Figure 2.1**). Phosphopantetheinylated MalG  $A_1$ - $T_1$  was loaded with L-proline in the presence of ATP and  $Mg^{2+}$ , consistent with our functional annotation. With no access to soluble MalG  $A_2$ , L-tryptophan was loaded onto MalG  $T_2$  using Sfp,<sup>30</sup> a nonspecific 4'-phosphopantetheinyl transferase, and L-Trp-coenzyme A (CoA) which was synthesized by Sean Newmister in

the Sherman group at the University of Michigan. L-Pro  $A_1$ - $T_1$  and L-Trp  $T_2$  were incubated with the MalG  $C$  domain and  $R$  domain with the presumed NADPH cofactor. Product formation was determined by LC/MS and comparison with authentic standards. Instead of the proposed dipeptidyl aldehyde-derived product **1.9**, Dan identified the aromatic zwitterion **1.11** as the main product (**Schemes 1.1** and **2.1**, and **Figure 2.2**). We hypothesized that **1.11** was produced from spontaneous oxidation of **1.9**. This was confirmed by chemical synthesis of **1.9** by Kimberly Klas in the Williams group at Colorado State University, which spontaneously and irreversibly converted to **1.11**. After anaerobic experiments, it was concluded that the malbrancheamide NRPS



**Figure 2.2** *In vitro* reaction of MalG. This produced the zwitterion **1.11** by spontaneous oxidation of the dipeptidyl aldehyde **1.9**.



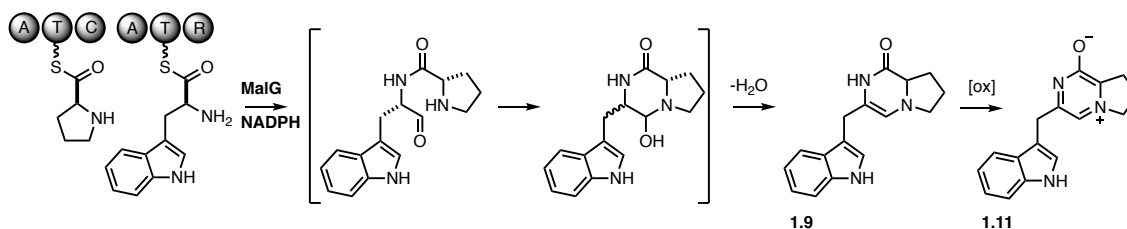
**Figure 2.3** *In vitro* reaction of MalG and MalE. This produced the prenylated zwitterion **1.12** and racemic premalbrancheamide.

product rapidly cyclized and dehydrated to **1.9** and subsequently spontaneously oxidized to **1.11** under aerobic (i.e. physiological) conditions (Scheme 2.1).

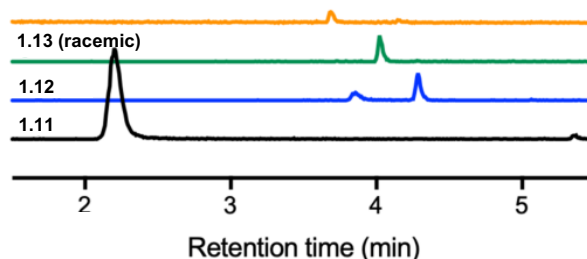
### 2.2.2 Prenyltransferase

We proposed that the NRPS product **1.9** would undergo a reverse prenylation as the next biosynthetic step, thereby installing the dienophile for the IMDA reaction. Two genes, *malE* and *malB* (from the malbrancheamide gene cluster) encode putative prenyltransferases. Dan incubated MalB or MalE with substrate-loaded MalG domains, NADPH and dimethylallyl pyrophosphate (DMAPP), the prenyl donor. MalE readily catalyzed a C2 reverse prenyl transfer reaction to produce the prenylated zwitterion **1.12** (Figure 2.3 and Scheme 2.3) whereas MalB displayed modest activity, suggesting that *malB* may be a redundant gene in the pathway (Figure 2.4 and Figure 1.4).

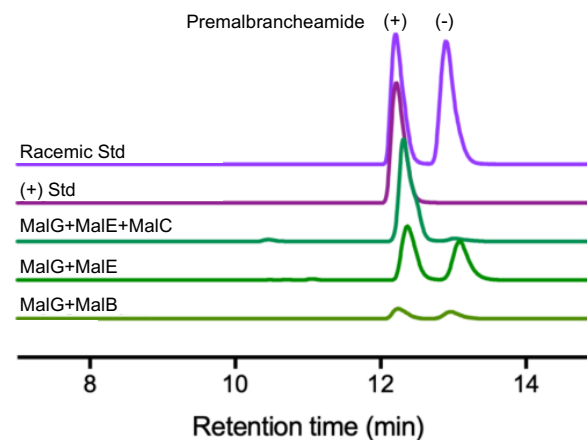
Because we could not distinguish in this assay whether **1.9** or **1.11** was the MalE substrate, synthetic **1.9** was produced under anaerobic conditions by UV irradiation of an *O*-nitrobenzyl (ONB) photo-protected dipeptide aldehyde and subjected to the prenyltransferase assay. This substrate was rapidly prenylated by MalE in contrast to synthetic **1.11**, which showed low levels of conversion with the enzyme, indicating that **1.9** is the native substrate for MalE (Scheme 1.1).



**Scheme 2.1** MalG NRPS reaction. This produces the aldehyde-derived **1.9** which oxidizes to zwitterion **1.11**.

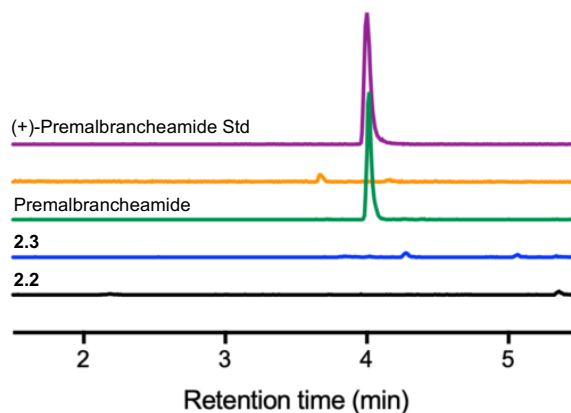


**Figure 2.4** *In vitro* reaction of MalG and MalB. This produced the prenylated zwitterion **1.12** and racemic premalbrancheamide.



**Figure 2.5** Chiral separation of racemic premalbrancheamide. The separation of (+)-premalbrancheamide indicates that MalC provides enantioselectivity for the IMDA.

This raised the question regarding how MalE accesses substrate **1.9** prior to its rapid oxidation to **1.11**. Thus, we considered the possibility that C2 reverse prenylation occurs with the substrate tethered to the NRPS T<sub>2</sub> domain. To address this question, we tested whether MalE or MalB could 32renylated L-Trp, L-Trp-loaded MalG T<sub>2</sub>, or MalG T<sub>2</sub> loaded with a dipeptidyl-CoA analog where the prolyl nitrogen was replaced with an oxygen to prevent nucleophilic addition. In all cases, no product was detected, confirming that the prenyl transfer reaction occurred on free substrate following the NRPS-catalyzed reaction.



**Figure 2.6** *In vitro* reaction of MalG, MalE, and MalC. This produces only (+)-premalbrancheamide.

We noticed low levels of premalbrancheamide in the reconstitution assays with MalG and MalE or MalB. I developed a chiral LC/MS method to analyze whether the products were optically pure, and revealed a 1:1 racemic mixture of ( $\pm$ )-premalbrancheamide (**Figure 2.5**), in agreement with the biomimetic synthesis described previously.<sup>23</sup> Further investigation using synthetic **1.12** revealed that racemic premalbrancheamide arose through non-enzymatic reduction of **1.12** by NADPH to azadiene **1.13**, which undergoes spontaneous cycloaddition in the reaction buffer, thereby explaining the background accumulation of the Diels-Alder products ( $\pm$ )-premalbrancheamide from *in vitro* assays. From these studies we ascertained that MalG and MalE are the minimal components required for premalbrancheamide biosynthesis, albeit lacking stereocontrol in the IMDA reaction.

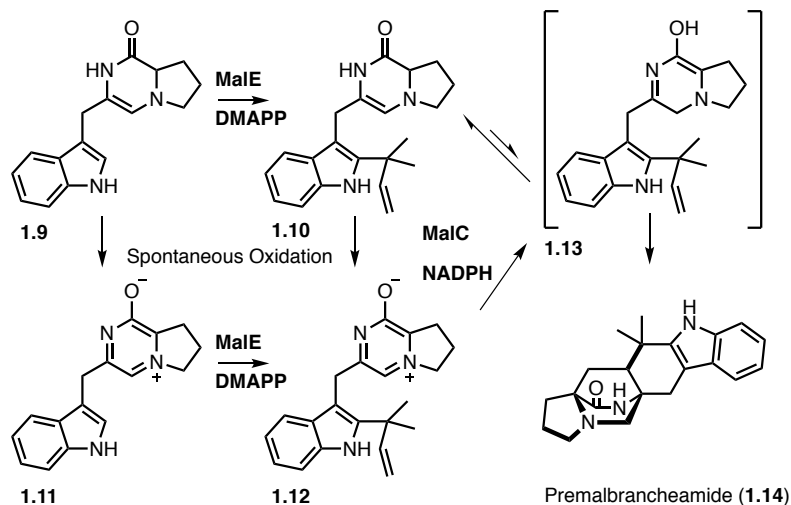
### 2.2.3 IMDA Enzyme Biochemistry

(+)-Premalbrancheamide isolated from *Malbranchea aurantiaca* is optically pure, which strongly implicates enzymatic control in the IMDA reaction. Known Diels-Alderase have diverse origins, thus bioinformatic analyses of the paraherquamide and malbrancheamide gene clusters provided little evidence as to the enzyme of interest. Nonetheless, Dan tested whether MalC, annotated as a short-chain dehydrogenase/reductase (SDR), could function as the presumed Diels-

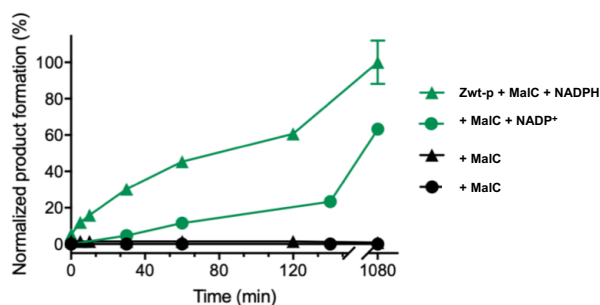
Alderase. When MalC was incubated with substrate-loaded MalG and MalE (NADPH and DMAPP included), neither aromatic zwitterion intermediate, **1.11** or **1.12**,

was detected in the chiral LC/MS; instead the sole product was (+)-premalbrancheamide, confirming that MalC functions as an intramolecular [4+2] Diels-

Alderase (**Figure 2.6**). To our surprise, when MalC was added to the reaction mixture after significant amounts of **1.12** had accumulated, the oxidized intermediate was converted to (+)-premalbrancheamide, indicating that MalC possessed the ability to reduce the zwitterion **1.12** to the reactive azadiene **1.13** prior to conducting the diastereo- and enantio-controlled cycloaddition reaction (**Scheme 2.3**). This unexpected reactivity of MalC was confirmed using synthetic **1.12** and NADPH. To our knowledge, this is a unique example where reduction regenerates the biosynthetic substrate from an oxidized (aromatic) intermediate to provide a productive mode for cycloaddition. The fact that **1.12** is a MalC substrate indicates that it is an authentic pathway intermediate, and motivated us to address whether an aerobic or anaerobic biosynthetic route is operative in the producing organism. This question was interrogated in two ways; first by performing MalC assays under anaerobic conditions with synthetic **1.10**, which was generated by photo-deprotection of ONB prenyl dipeptidyl aldehyde. Conversion to (+)-premalbrancheamide was observed only in the presence of MalC and NADP<sup>+</sup> (**Figure 2.7**). However, the efficiency of this reaction was attenuated compared to the MalC-catalyzed conversion of **1.12** to (+)-premalbrancheamide under aerobic conditions, indicating that the dienamine

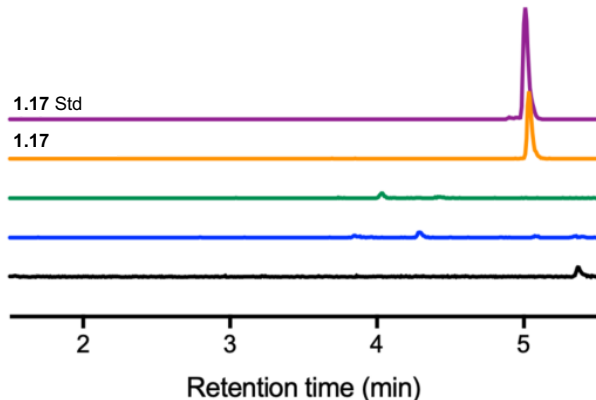


**Scheme 2.3** Prenylation and IMDA cyclization to generate (+)-premalbrancheamide.



**Figure 2.7** Aerobic and anaerobic MalC-catalyzed reactions. Under aerobic conditions the prenylated zwitterion (**1.12**) was the substrate and under anaerobic conditions the dienamine tautomer (**1.10**) was the substrate.

tautomer **1.10** is not optimally recognized by MalC. It is unknown whether MalC can play a role in tautomerization of **1.10**; notably, background conversion of **1.10** to racemic premalbrancheamide was not detected under these conditions. Second, Ying Ye in the Sherman group at the University of Michigan conducted gene disruption of the *malC* homolog *phqE* in the paraherquamide-producing strain *Penicillium simplicissimum* using a CRISPR-Cas9 system.<sup>31</sup> Extracts from the *phqE* mutant strain grown on CYA medium showed the presence of the expected (methyl-Pro-Trp prenyl) zwitterion (**1.20**) intermediate by LC/MS analysis and co-injection with a synthetic standard, confirming the accumulation of this oxidized metabolite *in vivo*. Taken together, these data indicate that **1.12** is the native substrate for 34renyla route to (+)-premalbrancheamide.



**Figure 2.8** Production of (+)-malbrancheamide. (+)-Malbrancheamide (**1.17**) was produced as the final pathway product when MalA selectively halogenated (+)-premalbrancheamide.

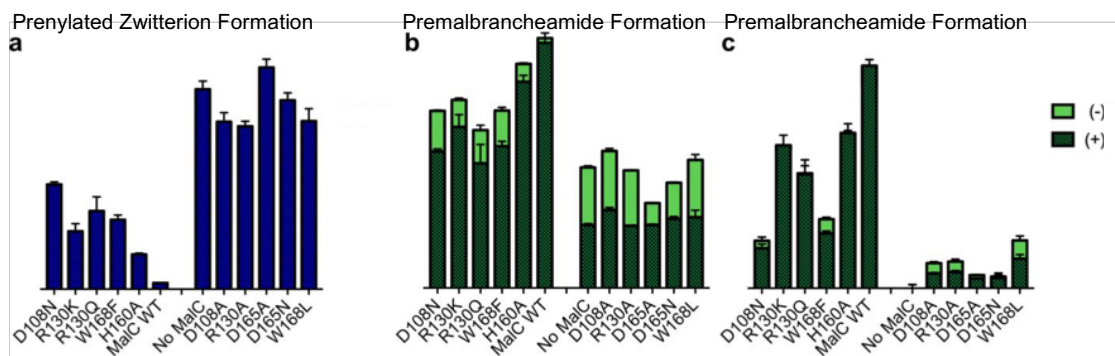
For the MalC-catalyzed reduction of **1.12**, NADH or NADPH are effective as the cofactor. However, NADP<sup>+</sup> is required for strict stereocontrol of the IMDA reaction, as MalC produced a 63:37 mixture of (+)- and (-)-premalbrancheamide (**1.14**) when using NADH. This is consistent with the anaerobic experiment in which NADP<sup>+</sup> was required to generate (+)-**1.14**, and further indicates that NADPH plays an important role in the IMDA stereocontrol.

To complete the biosynthetic pathway, flavin-dependent halogenase MalA was employed to add chlorine atoms on C8 and C9 of (+)-**1.14** to provide (+)-**1.17**.<sup>32</sup> We incubated MalA with its pathway partners (L-Pro and L-Trp MalG, MalE and MalC, NADPH, DMAPP, NaCl and FADH<sub>2</sub>) and identified (+)-**1.17** as the final product (**Figure 2.8**). We also found that MalA is stereospecific: when incubated with a racemic mixture of **1.14**, MalA chlorinated only the natural (+) enantiomer.

#### 2.2.4 IMDA Enzyme Structural Analysis and Mutagenesis

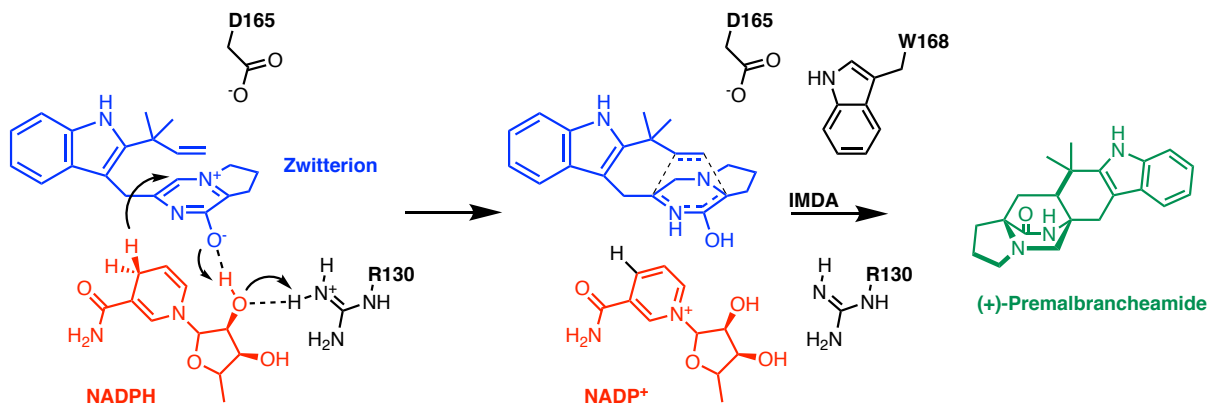
The high-resolution crystal structures of both MalC and PhqE solved by Dan and Sean, respectively, have provided the foundation for mechanistic analyses. In both cases, the catalytic tyrosine and lysine and essential asparagine and serine amino acids typically found in SDRs were missing.<sup>78, 79</sup> While complexes with MalC were not obtained, PhqE was crystallized in complex with the oxidized NADP<sup>+</sup> cofactor and the product (+)-premalbrancheamide (**1.14**) as well as NADP<sup>+</sup> and the prenylated zwitterion substrate **1.12**. After analyzing the structures and performing molecular dynamics simulations (Sean and the Houk group at UCLA), the reaction mechanism was probed by site-directed mutagenesis, and analyzed via chiral LC/MS.

Based on this information, we probed the reaction mechanism by site-directed mutagenesis. MalC was chosen for this analysis to directly compare results with the previous reconstitution assays. All of the targeted amino acids are conserved in MalC and PhqE. With the *in vitro* reconstitution assay, MalC variants were assayed in the presence of the MalG NRPS and MalE prenyltransferase (**Figure 2.9 a and b**). NRPS reductase activity was assessed by the levels of oxidized intermediate **1.12**: higher levels indicated less reductase activity. The effect on the IMDA reaction was determined by measuring levels of unnatural (-)-premalbrancheamide as a percent of all premalbrancheamide: ~50% (-)-premalbrancheamide formation indicates loss of enzymatic IMDA function. We identified five MalC substitutions that abolished reductase activity (D108A, R130A, D165A, D165N, W168L), and found that loss of function is highly correlated with loss of stereocontrol in the IMDA reaction. A single exception is MalC D165A, which produced mainly (+)-premalbrancheamide, suggesting that Asp165 is required for reduction but not the IMDA reaction. The activity of MalC variants was also measured in assays with **1.12** (**Figure 2.9 c**). In agreement with the reconstitution assay, Asp108, Arg130 and Asp165 were required for reduction.



**Figure 2.9** MalC mutant reaction data. a) and b) Product profiles of prenylated zwitterion **1.12** and premalbrancheamide **1.14** formed in the MalG+MalE+MalC reconstitution assay with MalC variants. C) MalC variants assessed by conversion of synthetic **1.12**.





**Figure 2.10** Proposed catalytic mechanism for MalC/PhqE. MalC amino acid numbers are used.

Based on the biochemical, structural, and computational data, we proposed a catalytic mechanism for MalC/PhqE (**Figure 2.10**). Arginine 130 is proposed to interact with the 2'-OH of NADPH ribose, initiating a hydride transfer to complete the first reduction step. Aspartate 165 can stabilize the positive charge of **1.12** during the hydride transfer, leading to the formation of the unstable azadiene intermediate **1.13**. The enzyme controls the diastereo- and enantioselectivity of the IMDA reaction by packing the substrate close to the NADP<sup>+</sup> and an aromatic tryptophan in the active site.

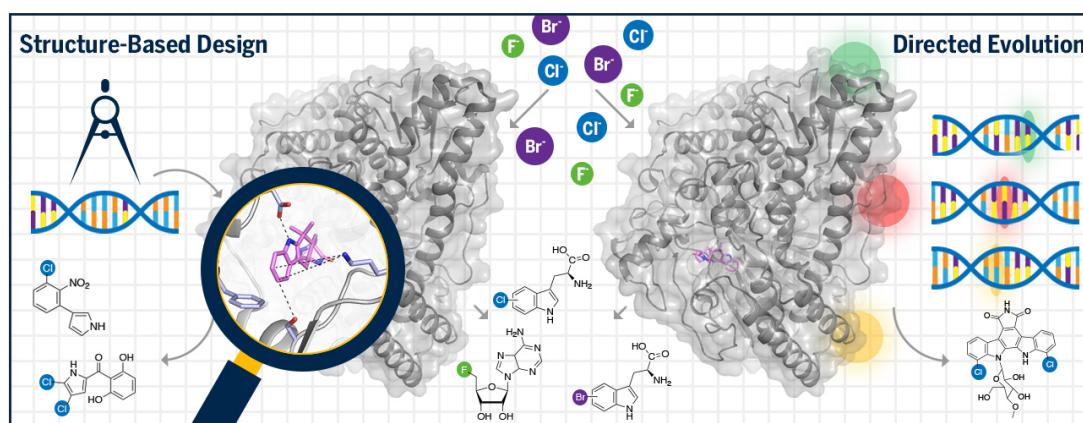
### 2.2.5 Conclusion

This work has unraveled complex scientific questions which were first described nearly half a century ago by Birch.<sup>10, 80</sup> Our collaborative efforts have led us to discover the origin of the polycyclic core, and combined with our previous biosynthetic knowledge, we could piece together the full enzymatic pathway to generate malbrancheamide. The bifunctional reductase and intramolecular [4+2] Diels-Alderase MalC/PhqE was found to provide diastereo- and enantiocontrol through the accommodation of an aromatic zwitterion substrate. MalC/PhqE defines a new class of IMDA enzymes where the aromatic zwitterion substrate is “rescued” from causing an otherwise premature termination of the pathway. This is an elegant example of evolutionary selection to protect the structural and stereochemical integrity of this complex family of indole alkaloids.

## Chapter 3

### Biocatalytic Late-Stage Halogenation

#### 3.1 Introduction to Halogenation Biocatalysts and Enzyme Engineering\*



**Figure 3.1** Halogenase Engineering Strategies. Halogenation is commonly used in medicinal chemistry to improve the potency of pharmaceutical leads. While synthetic methods for halogenation present selectivity and reactivity challenges, halogenases have evolved over time to perform selective reactions under benign conditions. The optimization of halogenation biocatalysts has utilized enzyme evolution and structure-based engineering alongside biotransformation in a variety of systems to generate stable site-selective variants.

##### 3.1.1 Introduction to Halogenases

Enzymes have evolved to catalyze reactions with high efficiency and selectivity, and can overcome inherent chemical biases with complex substrates. Halogenation is particularly important in medicinal chemistry and drug design, and with the aid of biocatalysts, chemists can implement efficient methods for halogenation that are environmentally benign. Since the discovery of halogenating enzymes (halogenases), an effort has been made to engage their selective reactivity in syntheses of complex molecules such as natural products. Thus, the collective knowledge gained from the characterization of these enzymes has led to a major effort in halogenase discovery and engineering for biocatalyst development.<sup>81-84</sup> This has included methodologies in structure-based engineering<sup>85, 86</sup> and directed evolution,<sup>87</sup> where biocatalysts with scalable capabilities<sup>88</sup> have been

\* Section 3.1.1 is included in a manuscript: Fraley, A. E.; Sherman, D. H. Halogenase Engineering and Its Utility in Medicinal Chemistry. *Bioorg. Med. Chem. Lett.* **2018**, 28 (11), 1992-1999.

generated. Additionally, halogenases have been implemented *in vivo* using various microbial expression systems to optimize biotransformation of complex molecules.

Most halogenases are currently grouped into four classes including 1) haloperoxidases (heme-containing and vanadium-containing), 2) iron (II)/2-(oxo)-glutarate ( $\text{Fe}^{\text{II}}/2\text{OG}$ )-dependent halogenases, 3) flavin-dependent halogenases (FDHs), and 4) fluorinases. Haloperoxidases tend to be nonselective due to the freely diffusing hypohalous acid that acts as the halogenating agent, whereas  $\text{Fe}^{\text{II}}/2\text{OG}$ -dependent halogenases proceed through a radical mechanism to halogenate aliphatic carbons on a variety of substrates. FDHs utilize hypohalous acid similarly to the haloperoxidases, but they implement a selectivity mechanism involving the formation of chloramine on a catalytic lysine in the enzyme active site.<sup>89</sup> Enzymatic fluorination requires unique conditions that enable desolvation of the fluoride ion, which subsequently performs a nucleophilic attack on the substrate. Numerous studies relating to the mechanism and function of FDHs have been reported, from which a foundation for halogenase engineering has been built. Recent studies have included the first application of computational chemistry to analyze the mechanism of FDH catalyzed chlorination.<sup>72</sup>

### 3.1.2 Bioactive Halogenated Natural Products and Halogenating Enzymes

Many of the over 4,000 known halogenated natural products have exhibited therapeutic potential. The antitumor agent rebeccamycin,<sup>90</sup> antibiotic vancomycin,<sup>91, 92</sup> and vasorelaxant malbrancheamide<sup>5, 15</sup> are structurally complex compounds, and strategies toward their synthesis would benefit from new halogenation methods (**Figure 3.2**). The challenges involved in preparing these materials through synthetic methods have motivated researchers to investigate the biosynthetic machinery involved in the assembly and late-stage tailoring of these secondary metabolites. Current chemical halogenation methods are rarely selective and tend to involve the use of toxic reagents. The rebeccamycin halogenase RebH has been extensively characterized<sup>89, 93, 94</sup> and the wealth of biochemical information has been utilized toward enzyme engineering efforts. RebH binds tryptophan, a biosynthetic precursor, and catalyzes chlorination at the Trp C7 position prior to incorporation into the metabolite by additional biosynthetic enzymes (**Figure 3.2a**). The halogenase involved in vancomycin biosynthesis (VhaA) performs a late-stage dehalogenation of the carrier protein bound hexapeptide precursor prior to cyclization and offloading of the final product (**Figure 3.2b**).<sup>95</sup> Late-stage halogenation of complex molecules is particularly challenging, but tethering a substrate to a carrier protein presents its own biosynthetic demands.

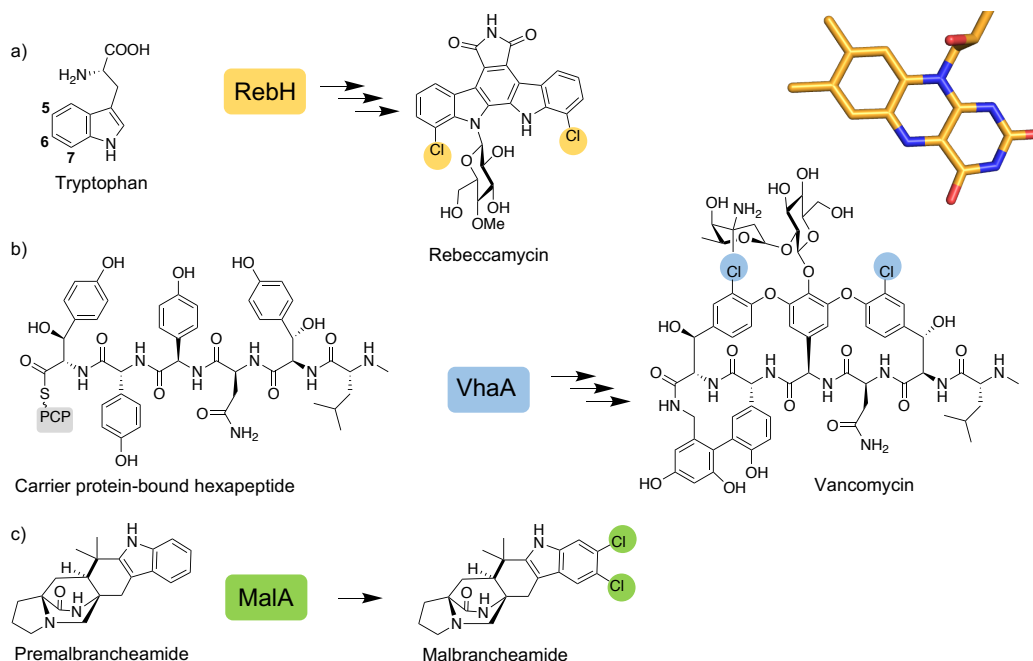


Figure 3.2 Examples of flavin-dependent halogenases. These enzymes are involved in the production of therapeutically relevant natural products. PCP denotes peptidyl carrier protein.

Numerous fungal FDHs that react with freestanding complex molecules have been identified,<sup>96-101</sup> including MalA which is a late-stage halogenase involved in malbrancheamide biosynthesis (**Figure 3.2c**).<sup>72</sup> Chloramphenicol is a well-known antibiotic that is chlorinated by the FDH CmlS, which bears a rare covalently bound flavin cofactor, and its biosynthetic halogenation remains under investigation.<sup>102-105</sup> Molecular dissection has enabled engineering strategies directed toward improving stability, expanding substrate scope, and altering the site-selectivity of FDHs.

### 3.1.3 Halogenation and Drug Design

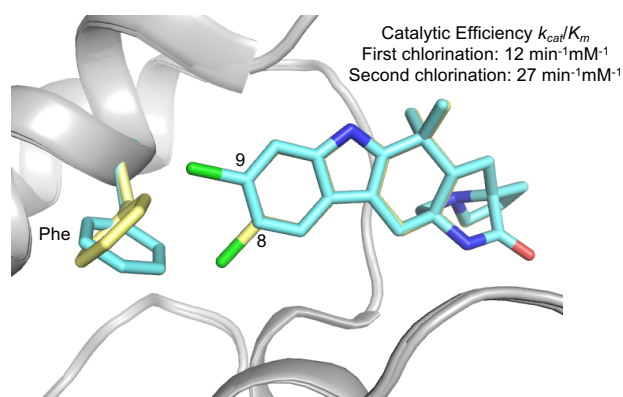
One third of the drugs currently in clinical trials are halogenated, which underscores the significance of methods for introducing halogen atoms into pharmaceuticals.<sup>106</sup> Halogens can aid in controlling metabolism and improving pharmacological properties such as lipophilicity and permeability. Additionally, the halogen functionality can facilitate target binding which was originally thought to be a consequence of electron-withdrawing and steric effects. More recently, investigators have begun studying the “halogen bond” and its utility in generating more selective therapeutics.<sup>107</sup> Due to the anisotropic nature of the halogen atom, a region of positive potential (sigma hole) can interact with a negatively charged region on the protein to form a halogen bond,<sup>108, 109</sup> the strength of which directly correlates with the size of the halogen itself. In fact, the strongest halogen bond (between iodine and a backbone carbonyl) is comparable in energy to the weakest hydrogen bond.<sup>110</sup> Typically, halogen bonds have been observed either between the ligand

halogen and aromatic amino acids or a backbone carbonyl.<sup>32</sup> The interactions with aromatic side chains are considered either edge-on or face-on, where the halogen approaches the aromatic ring from the periphery (edge-on) or interacts directly with the center of the aromatic ring (face-on).<sup>106,107</sup> Structural analysis of proteins bound to halogenated ligands has indicated that the edge-on interaction is more common and tends to involve phenylalanine and histidine side chains. In contrast, the face-on interaction frequently involves the side chains of tryptophan and tyrosine residues, potentially due to their higher  $\pi$  densities.<sup>107</sup> When the halogen atom is attached to an aromatic region of the ligand, these bond strengths are potentiated.

The pharmacokinetic properties of drugs can be modulated by the insertion of a halogen substituent such as fluorine, which is commonly substituted for aromatic hydrogens that are prone to metabolism. In a study of *N*-benzylamide-based thrombin inhibitors, *para* substitution of fluorine showed the highest increase in potency presumably due to the decrease in metabolism.<sup>111</sup> The halogen bond has also been utilized to facilitate the selective binding of ligands to certain cytochrome P450 enzymes (CYPs).<sup>106</sup> Cl- $\pi$  interactions between the cardiovascular therapeutics ticlopidine and amlodipine and CYP2B4 stabilize a conformation conducive to metabolism that is not observed when the molecules bind CYP2B6. The Cl- $\pi$  interaction is not only significant for therapeutic targets, but it can also potentiate reactivity of some biosynthetic halogenases involved in producing halogenated natural products. Recent work that focused on the characterization of an iterative halogenase, MalA, found that a Cl- $\pi$  interaction facilitated the binding of the monochlorinated ligands leading to a higher catalytic efficiency for a second halogenation reaction (**Figure 3.3**).<sup>72</sup> Detailed mechanistic and structural analyses of halogenases have been instrumental in the development of these enzymes as biocatalysts. With insight into how these proteins work, engineering efforts have produced selective halogenases that perform chemical transformations on therapeutically relevant small molecules.

### 3.1.4 Halogenase Engineering

Enzyme engineering has long been employed in efforts to provide robust and selective biocatalysts for applications in



**Figure 3.3** Cl- $\pi$  interaction in MalA halogenase. The C9 chlorinated substrate is shown in blue (PDB ID: 5WGW) and the C8 chlorinated substrate is shown in yellow (PDB ID: 5WGZ).

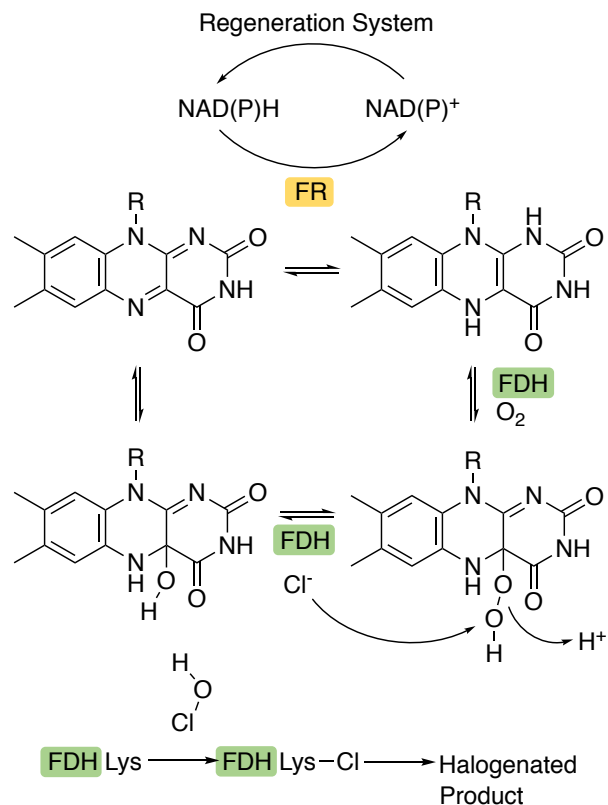
synthetic chemistry. Engineering strategies have included directed evolution<sup>112</sup> as well as structure-guided protein engineering,<sup>85</sup> both of which have been utilized to generate improved biocatalysts for halogenation. Directed evolution involves selecting a parent enzyme that performs a reaction of interest and randomly mutating the polypeptide to produce a library of selective variants. In contrast, a structure-guided approach relies on knowledge of the protein structure and mechanism for rational selection of mutations that are predicted to alter the reactivity. Although unpredictable, directed evolution has produced biocatalysts with increased thermostability,<sup>113</sup> varied substrate scope,<sup>114</sup> and altered site-selectivity.<sup>87</sup> Structure-guided engineering presents a more rational approach that can be used for well-characterized proteins. This strategy has produced biocatalysts with reactivity comparable to traditional synthetic methods for halogenation,<sup>88</sup> as well as with improved reactivity<sup>115</sup> and site-selectivity.<sup>85, 86</sup> A key difference between these two methods is that developing a hind-sight rationale for enhancement in enzyme activity is often difficult for directed evolution.

#### *3.1.4.1 Enzyme Stability and Optimal Activity*

In order to utilize enzymes for effective chemical transformations, the biocatalyst must be highly efficient. In principle, a more efficient biocatalyst would exhibit both increased stability and higher reactivity. Based on this concept, initial studies using directed evolution to develop a thermostable FDH provided a RebH variant with a melting temperature 18 °C higher and an optimal reaction temperature 5 °C higher than the wild-type enzyme.<sup>113</sup> The thermostable RebH variant also had improved conversion and a longer half-life at higher reaction temperatures. At lower reaction temperatures that were optimal for the wild-type enzyme, the catalytic efficiency of the variant was much lower than that of wild-type RebH. This work demonstrated that the stability of an enzyme is not necessarily a reflection of its activity. In this example, the thermostable RebH variant may have been less active at physiological temperature due to constrained conformational flexibility.<sup>113</sup> Flavin-dependent proteins in general have been reported to exhibit broad conformational flexibility between the oxidized and reduced cofactor bound forms. The reduced flavin cofactor is required for halogenase activity, and it has been supplemented into the system using a variety of methods,<sup>88, 116</sup> as described below.

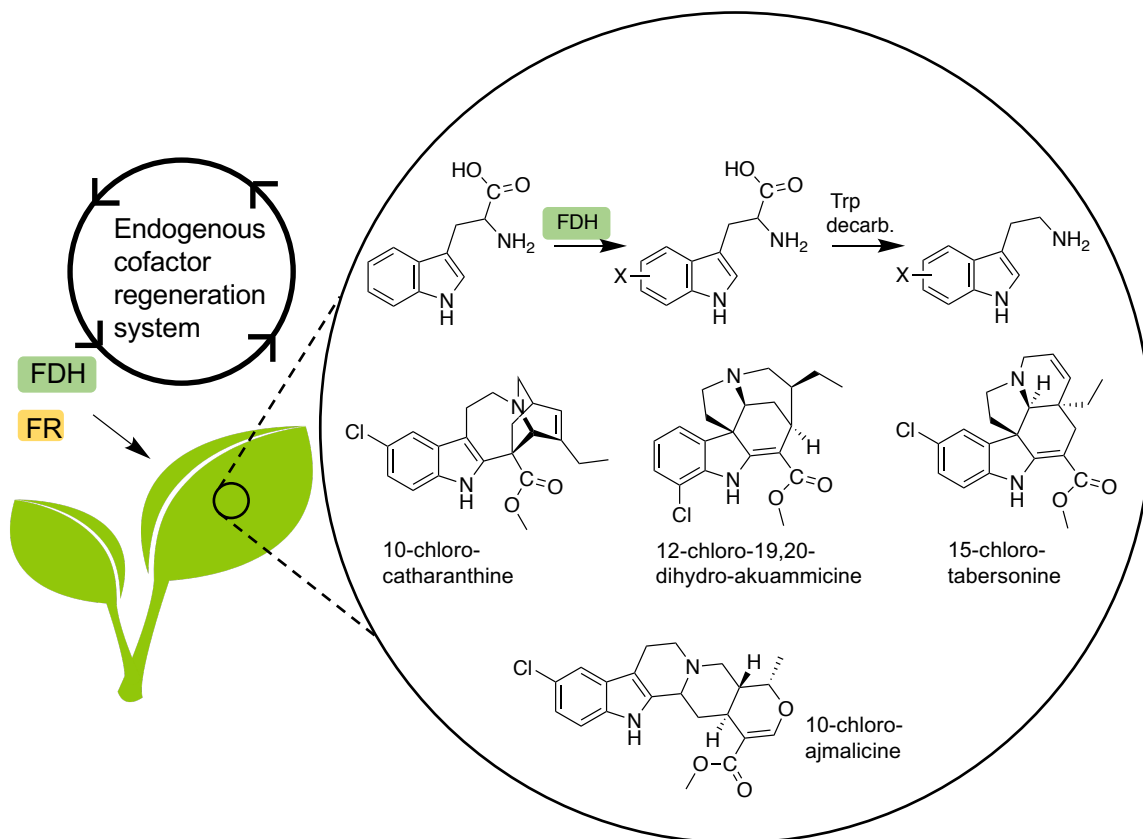
The use of FDHs as biocatalysts can be challenging due to the enzymatic systems required to generate the reduced flavin (FADH<sub>2</sub>) that is necessary for their activity. A nonspecific NAD(P)H-dependent flavin reductase (FR), such as RebF in rebeccamycin biosynthesis,<sup>93</sup> often

provides the FADH<sub>2</sub> for these reactions. In the most efficient scenarios, the NAD(P)H is regenerated through a recycling system consisting of either glucose dehydrogenase<sup>116</sup> or alcohol dehydrogenase (Figure 3.4).<sup>88</sup> One hypothesis is that by bringing the enzymes in this system into close proximity, they could exchange cofactors with each other at a faster rate. Frese, et al. immobilized these macromolecular components into aggregates and achieved gram-scale bromination of tryptophan by RebH.<sup>88</sup> The aggregates could not only be reused up to ten times, but they were also amenable to storage for up to four months at 4°C with no apparent decrease in activity. Immobilization of the enzyme components also streamlined the purification of the brominated product. Moreover, in contrast to *in vivo* bioconversion methods, chloride ion contamination could be avoided in the immobilized system, limiting the production of chlorinated molecules when bromination was desired.



**Figure 3.4** Flavin-dependent halogenase cofactor regeneration.

Rather than link all of the enzymes in the system together, Andorfer, et al.<sup>117</sup> applied a technique previously shown to be successful in generating productive P450/reductase fusion enzymes.<sup>118</sup> The autologous FR was linked to RebH with the expectation that the increased local concentration of FADH<sub>2</sub> would increase enzyme activity. While reductase activity was maintained, the kinetic parameters of the halogenation reaction for the fusion protein were diminished. When excess reductase was added to the reaction mixture, the halogenation catalyzed by the fusion protein increased, indicating that the FR must be present in a higher concentration than the halogenase to provide adequate levels of FADH<sub>2</sub> for efficient catalysis. When the halogenase and reductase were present in a 1:1 ratio the activity was decreased in both the fusion and the standalone systems, indicating that the local concentration of reductase in proximity to the halogenase was not pertinent to increasing the activity. While *in vitro* reactions of the fusion protein did not produce higher conversion rates, the *in vivo* fusion protein assays demonstrated



**Figure 3.5** Incorporation of FDHs into plants. The halogenated precursors are then incorporated into complex indole alkaloids.

higher conversion, which may have been due to the presence of an endogenous FR enhancing generation of the reduced cofactor.

Cofactor regeneration can also be achieved by using an *in vivo* system for biotransformation. By expressing Trp 7-halogenase RebH and Trp 5-halogenase PyrH,<sup>119</sup> alongside FR in Madagascar periwinkle (*Catharanthus roseus*), the incorporation of halogenated indole into various medicinally relevant alkaloids was observed (**Figure 3.5**).<sup>120</sup> The halogenated Trp was converted to the tryptamine precursor by the endogenous Trp-decarboxylase for the targeted indole alkaloid products. While wild-type *C. roseus* produced minor amounts of 19,20-dihydroakuammicine, the chlorinated analog 12-chloro-19,20-dihydroakuammicine was produced in high quantities in the RebH/RebF<sup>93</sup> variant, demonstrating the inherent flexibility of the downstream enzymes to accept the 7-chlorotryptamine precursor. A similar study where RebH and Trp-6-halogenase SttH<sup>121</sup> were expressed in *Nicotiana benthamiana* revealed that the localization of the FDH/FR system to the chloroplasts could aid cofactor regeneration.<sup>122</sup> Localization of halogenases to the chloroplasts in the absence of exogenous FR led to comparable product



formation indicating that the plastids provide a sufficient supply of reducing agents from native metabolic pathways.<sup>123</sup> The targeted coexpression of the Trp decarboxylase was required for production of halogenated tryptamine derivatives as the chloroplasts are devoid of this enzyme. Although most of the FDHs discussed are derived from bacterial hosts, the concept of transferring these biocatalysts to plants is particularly intriguing with respect to their less well-studied fungal counterparts, as these may be more effectively expressed in a eukaryotic host.

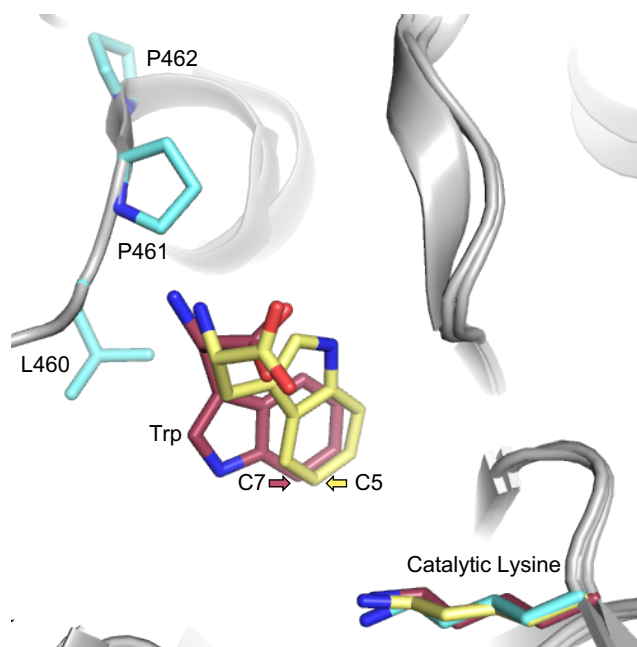
#### 3.1.4.2 Structure-based Halogenase Engineering

The halogenases involved in preparation of bioactive natural products are most synthetically useful when a clear understanding of their function has been determined. The first structure of the FDH PrnA<sup>124</sup> led to a mechanistic proposal that continues to be refined, and initial attempts to modulate the selectivity of these enzymes were structure-based. In the case of PrnA, a structural analysis of substrate complexes led to the mutation of a phenylalanine residue that appeared to interact directly with the indole ring of the tryptophan substrate.<sup>86</sup> This mutation induced a shift in the site-selectivity of the chlorination and bromination reactions from solely C7-halogenation to both C7 and C5 due to differential substrate positioning near the chloramine adduct that the enzyme is thought to employ for selective halogenation.<sup>89</sup>

Further structure-based engineering of RebH resulted in a mutant form that preferentially accepted tryptamine over its native tryptophan substrate.<sup>125</sup> The introduction of this mutant into *C. roseus* led to the selective chlorination of tryptamine and its incorporation into 12-chloro-19,20-dihydroakummicine. The *in vivo* biosynthetic applications of the chlorinated tryptamine precursor included preparation of halogenated analogs of molecules with therapeutic potential (**Figure 3.5**). The further utility of the tryptamine-selective mutant in this system depended on the flexibility of the downstream biosynthetic machinery toward accepting the halogenated precursor. Rational engineering of halogenase function has also been applied to the Trp 7-FDH PrnA and Trp 5-FDH PyrH.<sup>115</sup> The substrate scope of these two halogenases was assessed on small molecules that were aromatic, but lacked the indole ring present in tryptophan. A key substrate with more mobility in the active site, anthranilic acid, was selected to query engineered PrnA. Basic residues were substituted into the active site to coordinate the molecule and position it for selective halogenation. Single point mutations led to an increase in activity and selectivity between the *para* and *ortho* positions. By combining two of the most significant mutations, a large increase in activity was observed. X-ray crystal structures of the mutants displayed shifts in residues important

for substrate binding, but the mechanism of altered selectivity remained ambiguous due to the lack of bound substrate in the structures. The structure of the first Trp 6-halogenase, SttH, guided rational engineering efforts toward more selective halogenation.<sup>85</sup> The factors contributing to site-selective halogenation were determined by comparing the structures of the Trp 7-halogenases (RebH and PrnA) and the Trp 5-halogenase PyrH to that of SttH (**Figure 3.6**). A triple mutant was prepared by targeting residues in a region close to the core of the Trp substrate. Introduction of these mutations tuned the selectivity of the Trp 6-halogenase, which was then able to chlorinate at the C5 position as well. Upon utilizing a substrate with more flexibility in the active site (3-indolepropionic acid), the SttH triple mutant produced primarily 5-chloro-3-indolepropionic acid. Alternatively, when the corresponding triple mutant of PyrH was produced in an attempt to shift its selectivity from the C5 to C6 position, the protein was inactive. This work provided a rational strategy for switching the site-selectivity of a FDH without decreasing the catalytic efficiency of the enzyme. The observed differences between the different FDHs were exploited to design halogenation biocatalysts with tailored site-selectivity. This approach was applied to additional substrates and, combined with the immobilization methods of Frese et al.,<sup>88</sup> demonstrated the synthetic utility of these biocatalysts on a preparative scale.

A similar comparison of the crystal structures of the tetrahalogenase Bmp2 and 45renylated45e Mpy16 led to the identification of structural elements responsible for controlling the degree of pyrrole halogenation.<sup>126</sup> A Bmp2 triple mutant with the corresponding residues from Mpy16 produced a monohalogenated product, indicating that the substrate may have been repositioned. The same substitution in Mpy16 led to insoluble protein, thus the ability to shift the



**Figure 3.6** Halogenase structural overlay. Structures of Trp 7-halogenase PrnA (pink, PDB ID: 2JKC), Trp 6-halogenase SttH (cyan, PDB ID: 5HY5), and Trp 5-halogenase PyrH (yellow, PDB ID: 2WET) were aligned. Residues 460-462 of SttH were mutated to modulate the regioselectivity toward the Trp substrate.

selectivity of FDHs may require modulation of amino acids outside the immediate vicinity of the active site.

Structure-based engineering efforts have also been implemented to change the type of C–H functionalization reaction performed by an enzyme. After acquiring a detailed mechanistic understanding of Fe<sup>II</sup>/2OG-dependent oxygenases and halogenases, an effort to convert an oxygenase to the latter was undertaken.<sup>127</sup> Fe<sup>II</sup>/2OG halogenases are members of the oxygenase family and are highly similar to related enzymes that catalyze hydroxylation reactions. Unlike FDHs, Fe<sup>II</sup>/2OG-dependent halogenases perform chlorination or bromination reactions on aliphatic carbons through a radical-based mechanism. The structure of the halogenase WelO5 provided information about substrate binding that was critical to this engineering effort.<sup>128, 129</sup> The mechanism in both Fe<sup>II</sup>/2OG-dependent oxygenases and halogenases proceeds through a ferryl (Fe<sup>IV</sup>-O) intermediate, which abstracts a hydrogen atom from the substrate, leaving behind a substrate carbon-centered radical and Fe<sup>III</sup>-OH. In hydroxylases, the substrate radical reacts with the Fe<sup>III</sup>-OH species to yield a hydroxylated product. The iron coordination differs between hydroxylases, which utilize a carboxylate residue, and halogenases where this residue is substituted with an alanine/glycine. In the latter, an additional coordination site is open for a halide ion to bind and react with the substrate radical. With the aid of the WelO5 structure, it was determined that in order for halogenation to occur, the substrate must be located away from the ferryl intermediate and in a conformation unique to Fe<sup>II</sup>/2OG-dependent halogenases. A hydroxylase with high structural similarity to WelO5 was identified and the Fe<sup>II</sup> carboxylate ligand was mutated to glycine. With a structurally similar protein, it was hypothesized that the substrate would bind optimally for halogenation, and only the single amino acid substitution would be required for conversion to a halogenase. Both chlorinated and brominated products were acquired using this mutant, while some hydroxylation was still observed. This work provided a proof of concept that a thorough understanding of reaction mechanism could be exploited to achieve a switch in reaction type.

### *3.1.4.3 Directed Evolution for Engineered Halogenases*

In contrast to structure-based protein engineering strategies, RebH has also been subjected to a thorough directed evolution campaign. The pioneering effort to stabilize the enzyme toward a more tractable biocatalyst led to an increased melting temperature, and a more active enzyme.<sup>113</sup> The substrate scope of the thermostable RebH was expanded, and the mutants were shown to

chlorinate a variety of complex unnatural substrates.<sup>114</sup> Previous structure-based work had demonstrated the ability of a RebH mutant (Y455W) to selectively halogenate tryptamine over tryptophan.<sup>125</sup> This idea was utilized in subsequent work where RebH was engineered for optimal activity on tryptamine (N470S) to facilitate the employment of a high-throughput screening strategy for assessing halogenation selectivity.<sup>87</sup> The method involved assaying RebH mutants (generated through directed evolution) with 7-deuterotryptamine to assess halogenation at alternative sites by mass spectrometry. Error-prone PCR was employed to generate a library of RebH mutants with excellent selectivity for chlorinating the C6- and C5-positions of the tryptamine indole ring. Mutations identified from previous work were incorporated into the evolution process,<sup>113</sup> and degenerate codons were utilized at sites that showed particular selectivity when varied. Although the process appears relatively straightforward, the system became more complex when the authors found that reverting some of the original mutations in the newer variants led to higher reactivity while maintaining selectivity. After preparation of the selective mutants, the authors attempted to rationalize the mechanism of selectivity through docking models. For each mutant, the docking simulations produced a variety of substrate poses, some of which were consistent with the observed selectivity. For the C7- and C5-selective halogenases, the docking results included poses that were highly similar to those found in the crystal structures of the Trp 7-halogenase RebH and Trp 5-halogenase PyrH complexes with substrate. The engineered C6-halogenase provided mixed results, with no clear indication of the mechanism of selectivity. Shortly after this work was performed, the first structure of a Trp 6-halogenase (SttH) was solved, further aiding this analysis.<sup>85</sup>

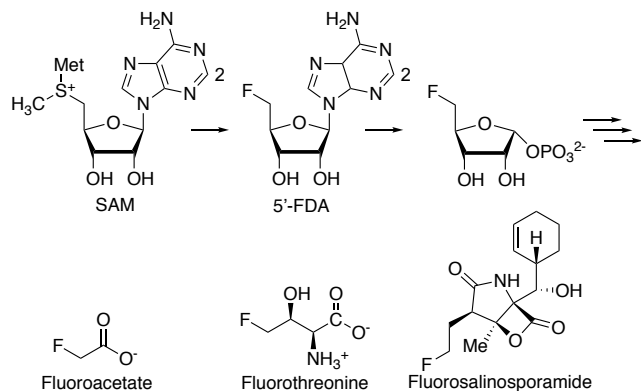
To expand the inquiry into FDH utility in synthetic approaches, a collaborative academic-industry partnership enabled a large survey of substrate scope for a number of natural and engineered FDHs.<sup>130</sup> The biocatalysts included wild-type RebH (Trp-7-halogenase), Thal (Trp-6-halogenase), several engineered RebH mutants, and a set of marginally characterized fungal FDHs. The RebH mutants were selected based on previous work that identified enzymes exhibiting increased activity on nonnative substrates,<sup>114, 115</sup> as well as altered site selectivity on tryptamine.<sup>87</sup> Through this screen, the investigators determined that fungal FDHs preferentially halogenated substrates that were less nucleophilic than bacterial Trp FDHs. Using computational methods, the authors predicted that there was a minimal electronic activation (HalA) required for halogenation by FDHs through the proposed electrophilic aromatic substitution. Accordingly, potential

substrates could be more accurately predicted. In previous directed evolution studies with these proteins, it was observed that the electronic activation is not the only factor contributing to their site-selectivity. Docking studies were useful for developing hypotheses regarding substrate binding, but developments using these methods were not always insightful. While a broad range of substrates were analyzed in this work,<sup>130</sup> FDH activity could only be correlated with the calculations of HalA and the key component of structural validation remained to be analyzed.

An interesting analysis was conducted between the chlorination reactions catalyzed by the FDHs compared to the common small-molecule reagent *N*-chlorosuccinimide (NCS).<sup>130</sup> The halogenases mediated a site-selective reaction, whereas NCS typically produced mixtures of regioisomers. Additionally, the use of NCS sometimes led to dihalogenated molecules, which was not observed for the corresponding FDHs, and the authors concluded that this was due to the decreased activity on halogenated substrates. Monohalogenation has been uniquely overcome by the fungal FDH MalA which displayed a higher catalytic efficiency for the second of two chlorination reactions on the complex indole substrate prenalbranchamide.<sup>72</sup> The surprising variation within this class of enzymes indicates that discovery of new FDHs will continue to broaden the repertoire of halogenating biocatalysts.

### 3.1.5 Fluorination Biocatalysts

The elusive fluorinase is of particular interest in medicinal chemistry due to the frequent use of fluorine to improve potency and selectivity of pharmaceutical lead compounds. Fluorine is frequently substituted for hydrogen in therapeutic molecules during the development phase because it tends to maintain or enhance target binding. Fluorine can also effectively modulate the metabolic profile and pharmacokinetic properties of the molecule.<sup>111, 131, 132</sup> Synthetic methods for fluorination represent a discrete focus of medicinal chemistry because of the challenges associated with fluoride ion desolvation. Without the use of harsh synthetic strategies, a fluorinase enzyme from *Streptomyces cattleya* was found to produce the free F<sup>-</sup> ion, which could perform a substitution reaction on S-adenosyl-L-methionine (SAM) to produce 5'-fluorodeoxyadenosine (5'-FDA) and 4-fluoro-threonine.<sup>133-135</sup> *S. cattleya* also produced fluoroacetate which could then be incorporated into therapeutically relevant natural products such as polyketides and steroids, via fluoromalonyl-CoA (**Figure 3.7**).<sup>132</sup> Novel fluorinated natural products have been produced through feeding fluorinated precursors into biosynthetic systems, as well as by genetically engineering fluorinases into the producing organism of interest.<sup>136</sup> The fluorinase gene from *S.*



**Figure 3.7** Biosynthesis of fluorinated molecules.

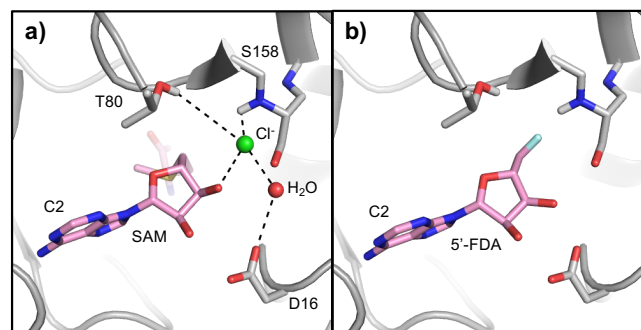
*cattleya* was genetically incorporated into *Salinospora tropica* to produce a fluorinated salinosporamide A, which is a naturally chlorinated anticancer agent.<sup>136, 137</sup> These strategies are only favorable when the downstream biosynthetic machinery can accommodate the fluorinated intermediates.

In the *S. tropica* system, endogenous enzymes were found to prefer the natural chlorinated substrates over the unnatural fluorinated molecules, and a disruption of the chlorinase gene was generated for optimal fluorine incorporation. Following discovery of the first fluorinase, an *in silico* genome mining effort was conducted to identify four additional enzymes in this family.<sup>138, 139</sup>

The fluorinase from *S. cattleya* has also been used in the preparation of <sup>18</sup>F-labeled imaging reagents for medical applications.<sup>140, 141</sup> Biocatalytic fluorination presents a significant advancement in <sup>18</sup>F chemistry because this radioisotope is usually prepared using a cyclotron with [<sup>18</sup>O]-H<sub>2</sub>O, leaving the final <sup>18</sup>F ion in aqueous conditions that are not suitable for the subsequent chemistry.<sup>142</sup> Fluorinase-catalyzed functionalization can utilize the <sup>18</sup>F ion under aqueous conditions efficiently without having to deviate from ambient temperature and neutral pH. In an application of biocatalytic <sup>18</sup>F fluorination, a fluorinase was paired with a nucleoside hydrolase to produce 5-[<sup>18</sup>F]-fluoro-5-deoxyribose which was then used in tumor imaging.<sup>63</sup>

Initial efforts to engineer fluorinase enzymes started with a comparison of the chlorinase in *S. tropica* to the fluorinase in *S. cattleya* which have 35% amino acid identity. Upon comparison of the halide ion binding site, an attempt was made to convert the chlorinase to a fluorinase by substituting the respective halide-binding residues.<sup>133, 143</sup> Structural<sup>135, 144</sup> and computational studies<sup>145</sup> of the fluorinase and chlorinase provided insights into the mechanism of halide selection. Crystal structures of the fluorinase in complex with chloride ion (since F<sup>-</sup> is indistinguishable from water) showed that the halide coordination is reliant on S158 and T80 in the fluorinase which are

substituted with G131 and Y70 in the chlorinase, respectively (**Figure 3.8**).<sup>134, 135,67</sup> Although, unsuccessful, this endeavor provided evidence that the halide binding site was not the only mechanism of halide selectivity. In recent advances in biocatalytic fluorination, more complex analogs of the traditional 5'-FDA were prepared through late-stage fluorination of C2-substituted molecules, including those bearing click-chemistry handles.<sup>66,146</sup> One



**Figure 3.8** a) Mechanism of fluoride binding. The mechanism of halide binding was demonstrated in crystal structures of the *S. cattleya* fluorinase with SAM and Cl<sup>-</sup>. The halide is proposed to be further coordinated by S158 to desolvate the F<sup>-</sup> ion for fluorination activity (PDB ID: 2V7U). b) Product complex with 5'-FDA. (PDB ID: 2V7V).

shortcoming of biocatalysis is that enzymes tend to exhibit reduced catalytic efficiency with unnatural substrates, which has been observed in the previously discussed halogenases, and fluorinases are no exception. Saturation mutagenesis was used to generate more active mutants, while reaction conditions such as temperature and leaving group options on the substrate were optimized.<sup>70</sup> Further understanding of the molecular determinants of substrate specificity based on the work of Sun, et al. led to an expanded substrate scope and higher reactivity on unnatural substrates.<sup>66</sup> The field of fluorine chemistry has benefited greatly from advances in fluorination biocatalysis, adding to the significance of biohalogenation as a whole.

### 3.1.6 Summary

Halogenases have been engineered to perform halogenation reactions on molecules relevant to medicinal chemistry and drug development. Enzyme engineering efforts have produced halogenases with greater stability and reactivity bringing the field closer to industrial applicability. The work toward plant-based systems for *in vivo* halogenation of small molecules has succeeded in bringing the complexity of the system (including cofactors and multiple accessory proteins) down to a practical level. This has also been realized through the use of cross-linked enzyme aggregates and fusion proteins that produced a more viable system for large-scale catalysis. As biocatalytic halogenation systems become more robust and scalable, they will become increasingly attractive to academic, medicinal, and process chemists working on drug development as well as total syntheses.

## 3.2 Characterization of MalA as the Malbrancheamide Late-Stage Halogenase\*

### 3.2.1 Introduction to the Malbrancheamides and MalA Halogenase

The prevalence of halogenated natural products has led to significant advances in understanding various classes of halogenases involved in secondary metabolism. Most halogenases characterized thus far can be placed into three classes: haloperoxidases (heme-containing and vanadium-containing), non-heme Fe(II) $\alpha$ -ketoglutarate-dependent enzymes, and flavin-dependent enzymes. Haloperoxidases are generally nonselective and perform halogenation through a purely chemical mechanism utilizing hypohalous acid. By contrast, Fe(II) $\alpha$ -ketoglutarate-dependent halogenases proceed through a radical mechanism, typically halogenating aliphatic, unactivated carbons.<sup>84</sup> Flavin-dependent halogenases (FDHs) also proceed through a hypohalous acid intermediate, with the reactive reagent captured by a lysine residue that appears to control the regioselectivity of halogenation on aromatic substrates.<sup>89, 124</sup> The FDH-derived hypohalous acid is generated through a reaction between the flavin C4a-peroxide adduct and the bound chloride ion. FDHs are thought to proceed through an electrophilic aromatic substitution where the catalytic lysine residue provides the chloramine halogenating agent and a catalytic glutamate facilitates the reaction by deprotonating the positively charged intermediate generated during catalysis.<sup>124</sup>

The majority of previously characterized FDHs are of bacterial origin, with relatively few reported from eukaryotes,<sup>98, 101, 147-150</sup> and fewer still characterized biochemically.<sup>98, 101, 147</sup> The bacterial FDHs have been found to catalyze reactions on both free,<sup>93, 119, 124, 151</sup> and carrier-protein bound substrates,<sup>126, 152</sup> including precursor amino acids in natural product biosynthesis. The well-characterized eukaryotic FDHs Rdc2<sup>101</sup> and ChlA<sup>147</sup> catalyze late-stage C–H functionalization reactions in the biosynthesis of halogenated metabolites. However, structural data for these two enzymes have not been reported, and it has remained unclear how they control late-stage site-selective halogenation on large, structurally complex substrates.

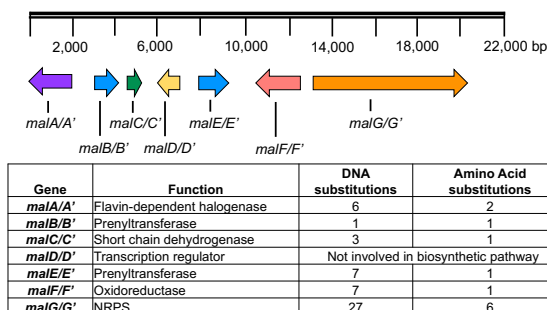
Malbrancheamide (**1.17**) is a complex halogenated indole alkaloid produced by the terrestrial fungus *Malbranchea aurantiaca* RRC1813<sup>5</sup> and the marine-derived fungus

---

\* Sections 3.2-3.5 are included in a manuscript: Fraley, A. E.; Garcia-Borràs, M.; Tripathi, A.; Khare, D.; Mercado-Marin, E. V.; Tran, H.; Dan, Q.; Webb, G. P.; Watts, K. R.; Crews, P.; Sarpong, R.; Williams, R. M.; Smith, J. L.; Houk, K. N.; Sherman, D. H. Function and Structure of MalA/MalA', Iterative Halogenases for Late-Stage C-H Functionalization of Indole Alkaloids. *J. Am. Chem. Soc.* **2017**, *139* (34), 12060-12068.



*Malbranchea graminicola* (086937A).<sup>71</sup> The discovery of **1.17** was enabled by a search for calmodulin antagonists, and several studies have characterized its significant vasorelaxant effect.<sup>15, 153</sup> Along with malbrancheamide, a close structural relative, spiromalbramide (**3.1**), was isolated from a marine fungus *Malbranchea graminicola*.<sup>71</sup> The two strains are highly related, with 99% sequence identity overall, and their biosynthetic pathways for **1.17** are proposed to be identical (**Figure 3.9**). **1.17** belongs to a family of prenylated indole alkaloids formed through peptide coupling by a nonribosomal peptide synthetase (NRPS), addition of an isoprene unit by a prenyltransferase, and a proposed [4+2] Diels-Alder cycloaddition to form the characteristic bicyclo[2.2.2]diazaoctane ring of premalbrancheamide (**1.14**) (Scheme 1.1).<sup>2, 23, 154-157</sup> **1.14** is then proposed to be 52renylated52ed through an iterative mechanism, but whether this halogenation was performed by one or two halogenases remained to be determined.<sup>157</sup> Since the chlorination of the indole ring differentiates this molecule from the rest of its class and significantly contributes to its biological activity,<sup>15</sup> we were motivated to probe the mechanism of its iterative late-stage halogenation at two adjacent positions on the indole ring system.



**Figure 3.9** Malbrancheamide biosynthetic gene clusters in *M. aurantiaca* and *M. graminicola*.

In earlier efforts to elucidate the malbrancheamide biosynthetic pathway, precursor incorporation studies were performed in *M. aurantiaca*. This led to the conclusion that **1.14** is indeed incorporated into the monochlorinated malbrancheamide B (**1.15**) and that both compounds are natural metabolites of *M. aurantiaca*.<sup>157</sup> We had previously proposed that there is a site-selective chlorination of the C9 position prior to functionalization of C8 for production of **1.17**.<sup>157</sup> However, the isolation of both C8 (isomalbrancheamide B (**1.16**)) and C9 (**1.15**) monochlorinated metabolites from *M. aurantiaca*<sup>153</sup> and from *M. graminicola*<sup>71</sup> conflicted with the proposed C9 selectivity hypothesis, providing further motivation to investigate the malbrancheamide halogenation process.

Previous genome sequencing and bioinformatic analyses of *M. aurantiaca* and *M. graminicola* led to the identification of MalA and MalA', respectively. These two FDHs are 99% identical, differing by only two amino acids, and are proposed to catalyze 52renylated52ed52 as the last step in the malbrancheamide biosynthetic pathways of the respective organism (Scheme

**1.1).**<sup>2</sup> The late-stage halogenation of free substrate by a flavin-dependent halogenase from an NRPS-containing gene cluster is rare. Indeed, halogenation typically occurs as the first step, prior to activation of an amino acid in bacterial non-ribosomal peptide biosynthesis.<sup>84</sup> Of the previously characterized flavin-dependent halogenases, most act on subunit substrates such as single amino acids,<sup>93, 119, 121, 124</sup> or carrier protein-tethered small molecules.<sup>105, 126, 152, 158</sup> In terms of late-stage activity on a complex polycyclic substrate, the closest comparison to MalA is the cyanobacterial-derived WelO5 non-heme Fe(II) $\alpha$ -ketoglutarate-dependent halogenase, which acts on fischerindole and hapalindole alkaloids.<sup>128</sup> In addition to substrate scope analyses, halogen selectivity has also been explored in flavin-dependent halogenases, and the majority were found to catalyze both chlorination and bromination reactions. In precursor incorporation studies using high bromide salt concentrations in the marine fungal strain *M. graminicola*, bromination of **1.14** led to the production of malbrancheamide C (**1.92**) and isomalbrancheamide C (**1.93**)<sup>71</sup> (**Figure 1.18**).

The ability to selectively halogenate C–H bonds in highly complex molecules through synthetic methods has posed a formidable challenge due to the abundance of chemically equivalent C–H bonds, and the inability to overcome inherent steric or electronic bias for reactivity.<sup>159, 160</sup> The large number of biologically active natural products that undergo late-stage functionalization by tailoring enzymes provides a unique opportunity to leverage the power of halogenating enzymes to perform difficult chemical transformations. Early efforts to modulate the selectivity of these enzymes achieved shifts in site-selectivity and slight modifications to substrate scope,<sup>86, 125</sup> while recent efforts have succeeded in engineering FDHs for a broad range of site-selectivities and substrates.<sup>87, 114, 116</sup> Through this work, we sought to identify and characterize the versatile halogenase involved in malbrancheamide biosynthesis and demonstrate its potential as a biocatalyst for late-stage halogenation of the structurally complex substrate, premalbrancheamide (**1.14**).

### 3.2.2 Biochemical Characterization of MalA

#### 3.2.2.1 Isolation of Malbrancheamides from *M. aurantiaca*

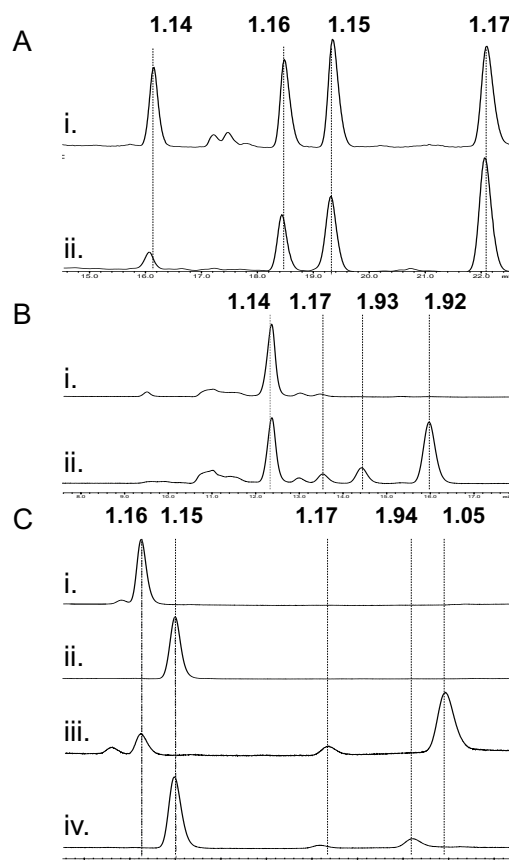
From a 1.5 L growth of *M. aurantiaca*, we obtained the following yields of the naturally occurring malbrancheamides: 1.6 mg/L premalbrancheamide (**1.14**), 2.6 mg/L isomalbrancheamide B (**1.16**), 4.4 mg/L malbrancheamide B (**1.15**), and 5.8 mg/L

malbrancheamide (**1.17**). These materials enabled subsequent biochemical and structural studies of MalA and MalA’.

### 3.2.2.2 Biochemical Activity of MalA

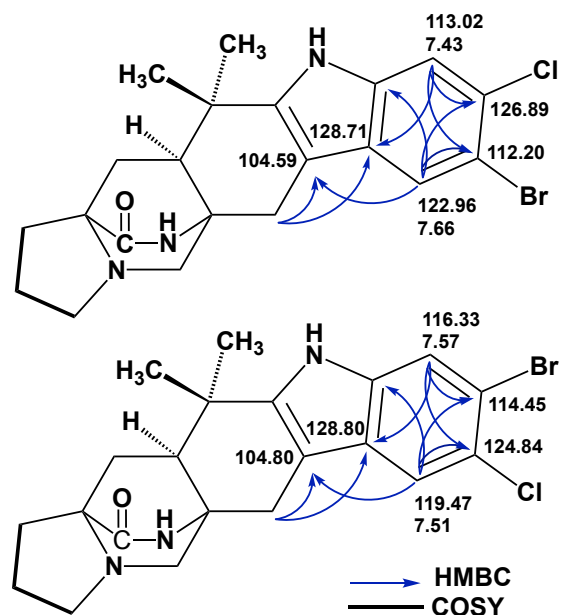
Purification of MalA by Ni-affinity chromatography and gel filtration provided pure protein for *in vitro* assays, and MalA was found to catalyze the iterative chlorination and monobromination of the natural precursor premalbrancheamide (**1.14**). In reactions with the monochlorinated malbrancheamide B (**1.15**) and isomalbrancheamide B (**1.16**). MalA was also able to chlorinate and brominate these compounds to generate malbrancheamide C (**1.92**), isomalbrancheamide C (**1.93**) malbrancheamide D (**1.94**), and isomalbrancheamide D (**1.95**), of which the latter two are novel indole alkaloids. The chlorination reaction of MalA was confirmed by co-elution with standards isolated from *M. aurantiaca* (Figure 3.10). <sup>1</sup>H-NMR analysis confirmed the halogenation site on the indole ring and the high-resolution mass spectrometry data for all compounds were consistent with the expected masses and isotope peak patterns for the halogenated products (Experimental section 7.4.1).

The chlorination of **1.14** to produce **1.15**, **1.16**, and **1.17** in a 5 mg *in vitro* reaction showed 34%, 26%, and 24% conversions respectively with slight selectivity for monochlorination at the C9 position. The bromination reaction was also found to be slightly selective for the C9 position, where a 4 mg reaction generated 23% **1.92** and 18% **1.93** as determined by isolated yields. The methodology for separation of these monohalogenated intermediates by HPLC is well resolved compared to previous reports, thus



**Figure 3.10** HPLC traces (240 nm) for the MalA *in vitro* reactions. (A) chlorination of premalbrancheamide (**1.14**) (ii) compared to standards from the fungal extract (i), (B) bromination of **1.14** (ii) compared to no enzyme control (NEC) (i), and (C) bromination of isomalbrancheamide B (**1.16**) (iii) compared to NEC (i) and bromination of malbrancheamide B (**1.15**) (iv) compared to NEC (ii).

the NMR data for the individual molecules significantly adds to the literature of these brominated indole alkaloids. MalA was also used as a biocatalyst for the generation of two novel bromo-chloro-malbrancheamide analogs malbrancheamide D (**1.94**) and isomalbrancheamide D (**1.95**) (Figures 3.10 and 3.11, Tables 7.8 and 7.9). The structural assignments of **1.94** and **1.95** were confirmed through extensive 1D and 2D NMR analyses. The structures were confirmed using key gHMBCAD correlations (Figure 3.11) where a significant downfield shift was observed for the brominated carbon as opposed to the chlorinated carbon. The positions of the halogens on the indole ring were confirmed by the two singlet peaks observed in the <sup>1</sup>H-NMR aromatic region of each spectrum. The bromination led to a downfield shift of the neighboring proton.



**Figure 3.11** Key gHMBCAD and gCOSY NMR correlations. These are from the indole region used to determine the sites of halogenation of new analogs malbrancheamide D (**1.94**, top) and isomalbrancheamide D (**1.95**, bottom).

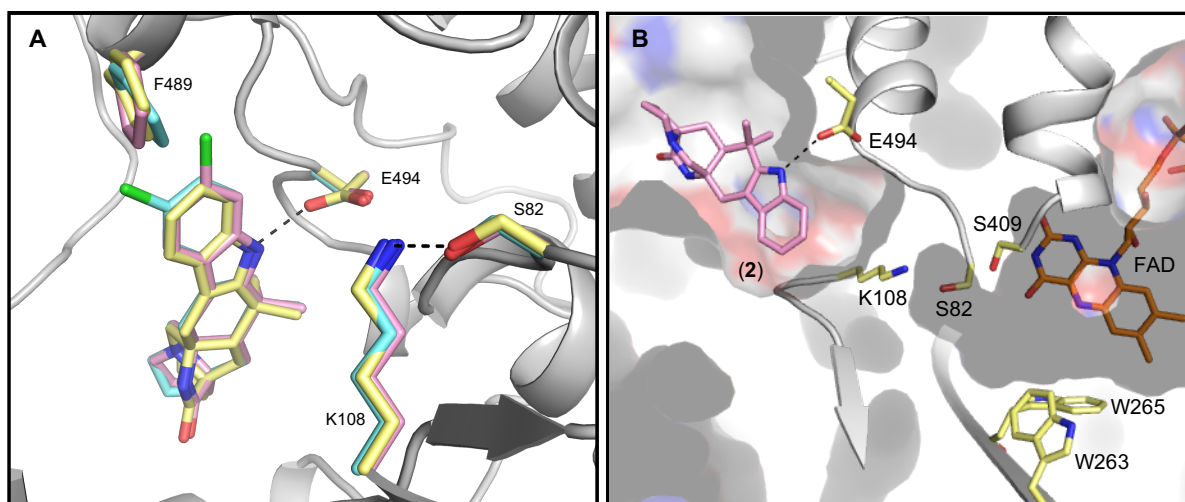
### 3.2.2.3 Kinetic Characterization of MalA

Michaelis-Menten kinetic parameters were determined for the natural chlorination reactions of MalA. They revealed that the enzyme has similar  $k_{cat}$  and  $K_m$  values for both the initial and second chlorination reactions. The  $k_{cat}$  from premalbrancheamide (**1.14**) to malbrancheamide B (**1.15**) and isomalbrancheamide B (**1.16**), malbrancheamide B (**1.15**) to malbrancheamide (**1.17**), and isomalbrancheamide B (**1.16**) to malbrancheamide (**1.17**) were 0.08, 0.09, 0.12, and 0.12 min<sup>-1</sup> (Table 7.2, Figure 7.1), respectively, which are comparable values to those of FDH PrnA (0.10 min<sup>-1</sup>).<sup>124</sup> The  $K_m$  values for **1.14** to **1.15** and **1.16**, **1.15** to **1.17**, and **1.16** to **1.17** were 7.0, 7.5, 4.4, and 4.0 μM. The catalytic efficiencies were calculated for each of the four reactions resulting in the  $k_{cat}/K_m$  values of 11.5, 12.0, 27.3, and 29.7 min<sup>-1</sup>mM<sup>-1</sup>, respectively (Table 7.2, Figure 7.1). These catalytic efficiencies are fairly high compared to those of the eukaryotic FDH Rdc2 which are 2.93 min<sup>-1</sup>mM<sup>-1</sup> for the initial chlorination reaction and 0.11 min<sup>-1</sup>mM<sup>-1</sup> for the second chlorination reaction.<sup>101</sup>

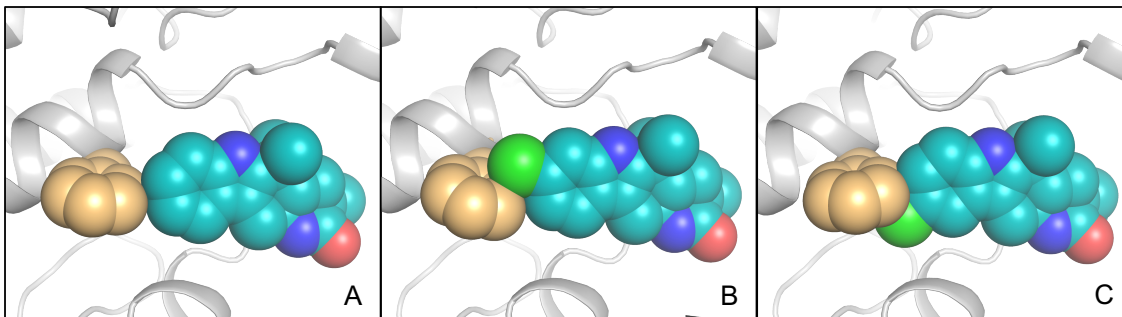
### 3.2.3 Structural Characterization of the Substrate Complexes of MalA'

To further elucidate the unique functionality of MalA/MalA' halogenase reactivity at two sites, the co-crystal structures of a MalA' in complex with premalbrancheamide (**1.14**), malbrancheamide B (**1.15**), and isomalbrancheamide B (**1.16**) were determined. MalA and MalA' are 99% identical, differing at only two amino acid positions (Leu276 and Arg428 in MalA; Pro276 and Pro428 in MalA'), and have comparable catalytic activities, but only MalA' was amenable to crystallization. To verify that MalA' was a viable substitute for MalA, the activities of the two were compared, and it was determined that MalA' was essentially identical to MalA under the reaction conditions tested. The structures of the ternary complexes with FAD, chloride ion, and each of the three substrates (**1.14-1.16**) were determined at 2.36 Å, 2.09 Å, and 2.04 Å, respectively. MalA' has a similar overall structure to bacterial FDHs with the addition of a few unique motifs including a Zn<sup>2+</sup>-binding C-terminus and a large active site capable of accommodating the structurally complex substrates. The natural substrates **1.14**, **1.15**, and **1.16** have similar binding modes in the MalA' active site (**Figure 3.12**). Specific interactions include a hydrogen bond of the indole nitrogen to Glu494 and a Cl- $\pi$  interaction of **1.15** and **1.16** with Phe489 (**Figures 3.12** and **3.13**).

The roles of amino acids in the active site were analyzed through site-directed mutagenesis (**Figures 3.13** and **3.14**), and Lys108 was determined to be necessary for catalytic activity. Trp263 and Trp265 form a characteristic flavin-binding motif proposed to aid in cofactor binding



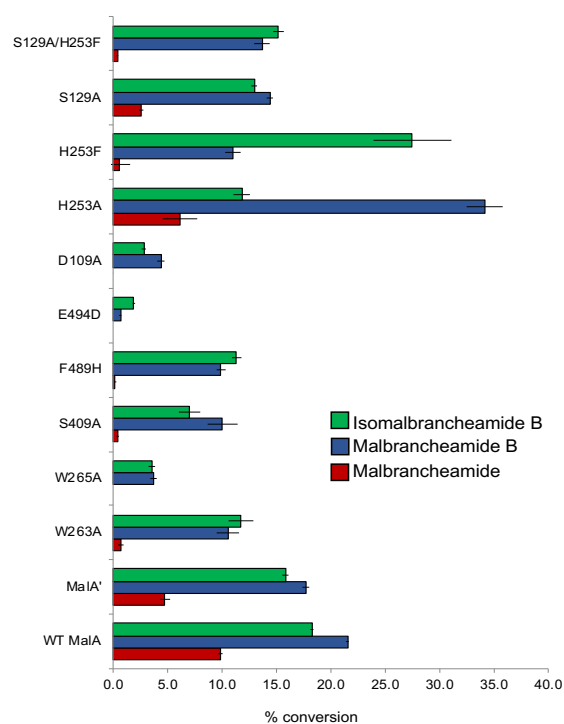
**Figure 3.12** Active site view of MalA'. A) MalA' active site overlay of complexes with substrates premalbrancheamide (**1.14**) (yellow), malbrancheamide B (**1.15**) (pink), and isomalbrancheamide B (**1.16**) (cyan). B) Active site of MalA'.



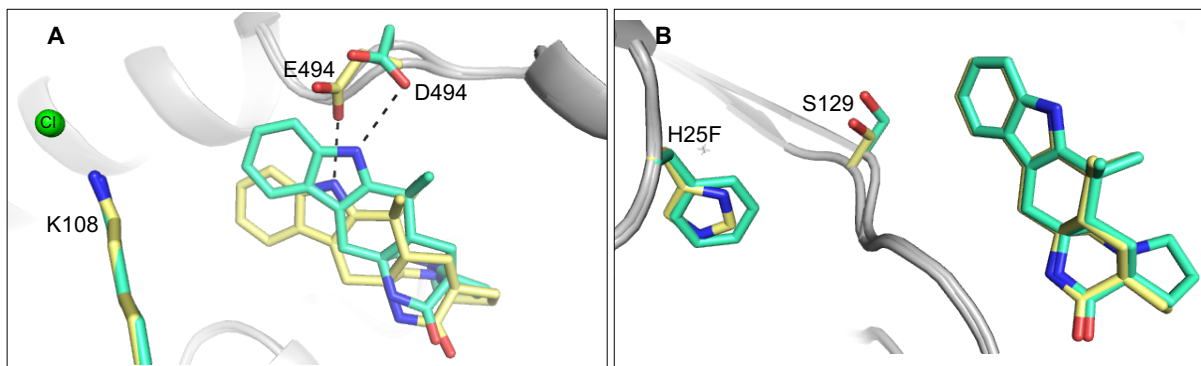
**Figure 3.13** Substrate interactions with Phe489. Interactions are observed between MalA' Phe489 (orange) and substrates (A) premalbrancheamide (**1.14**), (B) malbrancheamide B (**1.15**), and (C) isomalbrancheamide B (**1.16**).

(**Figure 3.12**). While MalA W265A showed a drastic decrease in activity, MalA W263A showed a more modest decrease in product formation (60% production of isomalbrancheamide B (**1.16**), 50% production of malbrancheamide B (**1.15**), and 5% production of malbrancheamide (**1.17**)) relative to wild-type MalA. A residue key to binding the substrate, Phe489, was substituted with histidine to ascertain the significance in the activity of MalA and was found to decrease the activity as well. Phe489 is analogous to the phenylalanine whose mutagenesis altered the site-selectivity in the FDH PrnA.<sup>86</sup>

In a preliminary effort to probe the mechanism of MalA, Glu494 was substituted with a variety of other residues including alanine, glutamine, and aspartate. While E494A and E494Q inactivated MalA, E494D maintained slight activity. Glu494 forms a hydrogen bond with the proton of the indole nitrogen, facilitating binding of the substrate. The substitution of aspartate at this position shifted the substrate away from the catalytic lysine, thus decreasing the activity (**Figures 3.14** and **3.15**). Initial efforts to probe the mechanism of site-selectivity in MalA



**Figure 3.14** Reaction data for MalA mutants. Percent conversion are shown for mutants versus wild-type MalA in reactions with premalbrancheamide (**1.14**) to produce isomalbrancheamide B (**1.16**), (green), malbrancheamide B (**1.15**) (blue), and malbrancheamide (**1.17**) (red). MalA K108A and E494A/Q were inactive and MalA S82A was insoluble, precluding functional analysis.



**Figure 3.15** Structures of MalA' E494D and H253F. A.) Comparison of wild-type MalA' (yellow) and MalA' E494D (green) co-crystallized with premalbrancheamide (**1.14**). B.) Comparison of wild-type MalA' (yellow) and MalA' H253F (green) co-crystallized with premalbrancheamide (**1.14**).

included substitution of His253 with alanine, phenylalanine, and other amino acids. MalA H253A was selective for producing the C9-chlorinated **1.15**, while MalA H253F displayed selectivity for producing the C8-chlorinated isomalbrancheamide B (**1.16**). Co-crystal structures of MalA' H253A in complex with premalbrancheamide (**1.14**) and malbrancheamide B (**1.15**) revealed no evident changes in the protein that would lead to the observed site-selectivity. On the other hand, the co-crystal structure of MalA' H253F in complex with premalbrancheamide (**1.14**) revealed a shift in S129, a residue near the indole ring of the substrate (**Figure 3.15**). When S129 was substituted with alanine, the C8 selectivity of MalA H253F was abolished, leading to the conclusion that Ser129 is involved in the selectivity induced by the Phe substitution at position 253 (**Figure 3.13**).

The structures of MalA' also revealed a unique zinc site with coordination by four cysteine residues (Cys597, Cys600, Cys613, Cys616) near the C-terminus. The  $Zn^{2+}$  ion was identified using anomalous scattering experiments with diffraction data recorded at X-ray energies bracketing the zinc K-edge (9.6586 keV). Anomalous difference density was present only in the map using data from the energy above the edge (**Figure 3.15**, **Figure 7.4**). A double mutant MalA C613S/C616S was prepared and the expressed protein was insoluble, thus no biochemical activity assays were performed in the absence of  $Zn^{2+}$ .

### 3.2.4 Molecular Dynamics Simulations and Quantum Mechanics Models

Molecular dynamics (MD) simulations were conducted by Marc Garcia-Borràs in the Houk group at UCLA to gain further insights into the structure and activity of the protein, starting from

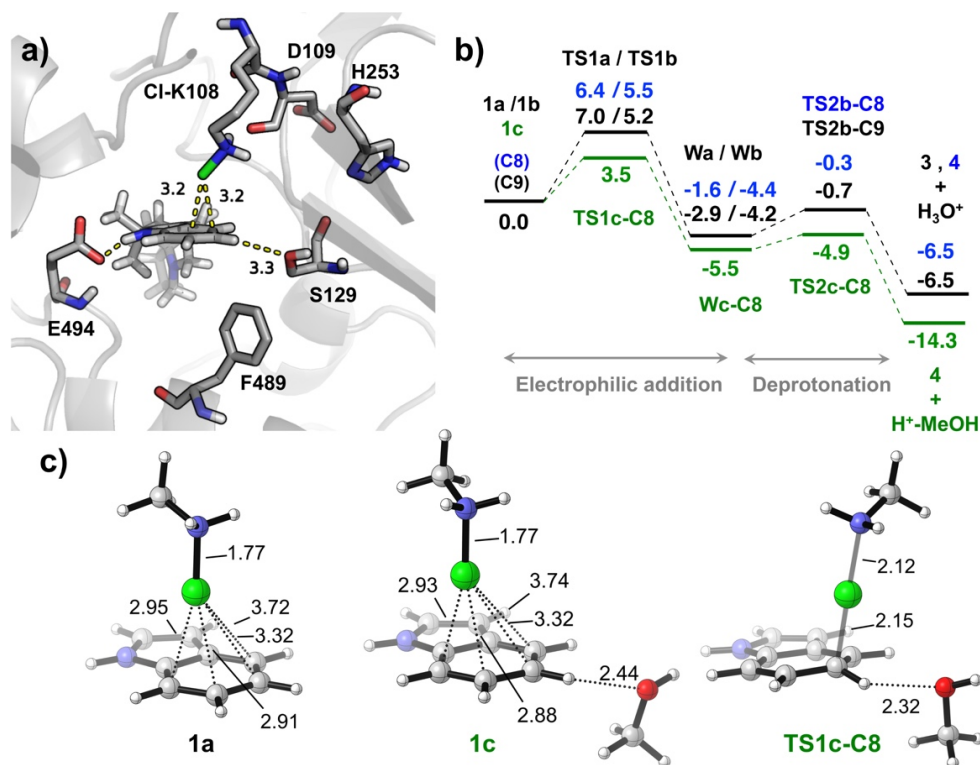
the MalA' crystal structures in their apo and substrate bound forms. In the latter case, the Lys108 chloramine intermediate has been considered integral to the mechanism discussed below.

The analysis of the MD trajectories suggested the structural role played by the Zn<sup>2+</sup> counterion in the protein structure. Residues within the Zn<sup>2+</sup> binding region (597-616) exhibited a low root-mean-square fluctuation (RMSF) compared to the very flexible adjacent region (621-646). These simulations indicate that the flexible region acts as a substrate channel lid, having two main open/closed conformations that were explored during the 500 ns MD simulations in both the apo and substrate-bound states (**Figure 7.6** and **7.7**). From the apo state trajectories, the pK<sub>a</sub> of the Lys108 and Glu494 side chains was estimated (**Figure 7.16**). Glu494 has a relatively high pK<sub>a</sub> ( $\approx$ 6.0-7.5), while Lys108 has an estimated pK<sub>a</sub> of 7.2-8.3 (**Figure 7.16**).

The analysis of possible polar interactions between the substrate and the enzyme active site suggested that, although the Gly131 and Ala132 backbone carbonyls could potentially interact with the substrate amide nitrogen, these interactions are not as important as the Glu494 – H(N-indole) hydrogen bond. The latter corresponds to the main and stronger interaction between the substrate and protein active site residues, and it is observed in all bound substrates (premalbrancheamide (**1.14**), malbrancheamide B (**1.15**) and isomalbrancheamide B (**1.16**)) during the entirety of the MD trajectory simulations (**Figure 7.9**). The basicity of Glu494 can thus enhance the hydrogen bonding between the carboxyl side chain and the indole ring of the substrate, positioning it to effectively interact with the catalytic Lys108 residue.

MD simulations including the proposed active chloramine Cl-Lys108 species suggested that when the **1.14** is bound into the active site and the channel lid is closed, the Cl atom is placed very close to the C8/C9 positions of the substrate, due in part to the positioning of the substrate by the Glu494 residue (**Figure 3.17a**). When the lid is in its open conformation, the substrate binding is slightly displaced although still H-bonding with Glu494, but then the Cl-Lys108 side chain conformation changes to place the Cl atom closer to the FAD cofactor (see **Figure 7.8**). This indicates that the protein conformational change between the substrate bound open/closed states could also control the positioning of Cl-Lys108 active species, which could explore two main conformations to bring the Cl atom from the flavin cofactor to the substrate (**Figure 7.8**). This observation suggests that Cl-Lys108 is the actual chlorinating species. When Cl-Lys108 is in this catalytically competent pose, the distances (Cl-C) and angles (Cl-C-H) measured for both C8 and





**Figure 3.16** a) Computational studies with MalA' complex. Representative snapshot (at 40 ns) taken from the 500 ns MD simulation on MalA' and substrate **1.14** bound complex, including the chloramine adduct at Lys108. b) Computed DFT reaction pathways for the three models: **1a**- indole and methyl chloramine; **1b**- indole, methyl chloramine and a water molecule coordinating at the C8 or C9 positions, respectively; **1c**- indole, methyl chloramine, and MeOH coordinating to H-C8. c) DFT optimized geometries of some representative structures (see Chapter 7 for additional details). Gibbs energies are given in kcal/mol, and distances in Å.

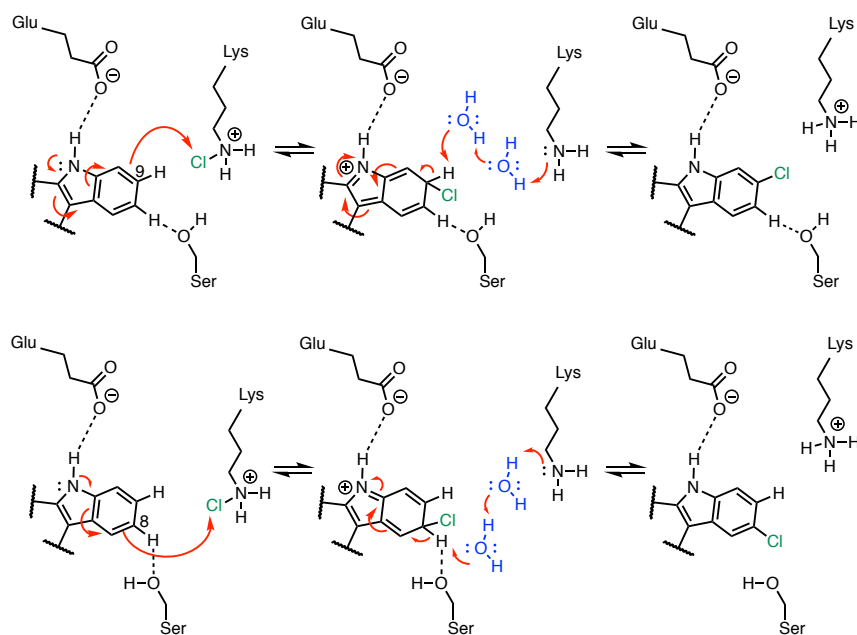
C9 positions are very similar, indicating that Cl-Lys108 is preorganized to chlorinate both positions (**Figure 7.14**).

MD simulations also revealed a key interaction between Cl-Lys108 and the backbone carbonyl of neighboring Asp109 residue, effectively positioning Cl-Lys108 towards a catalytically competent arrangement (**Figure 3.16a**). The Cl atom from Lys108 chloramine active species can only be placed close to C8/C9 positions when this H-bond is present (**Figure 7.10**). The essential role of the Asp109 was further explored by mutagenesis experiments: the D109A mutant has very low activity (as shown in **Figure 3.14**). MD simulations on this particular mutant showed that D109A backbone carbonyl is pointing towards a different position than in the original Asp109, which would eliminate the key hydrogen bond interaction. This is caused by the different conformation of the amino acid side chain, which is pointing outside the protein cavity and exposed to the solvent in Asp109, while in D109A it is displaced (**Figure 7.13**). Indeed, in our 500 ns

simulations for the D109A mutant, Cl-Lys108 never explores any conformation in which the Cl atom approaches the C8/C9 positions to affect chlorination (**Figure 7.13**).

MD simulations on MalA' Cl-Lys108 with bound premalbrancheamide (**1.14**) highlighted the arrangement of the Ser129 side chain with respect H-C8 in **1.14**. The distance between Ser129 O $\gamma$  and H-C8 is particularly short (between 2.5-3.5 Å) when the substrate is in an appropriate orientation for the electrophilic aromatic substitution (**Figure 7.10**). This interaction was not observed along the MD trajectory of MalA' H253A, but was prominent in the MalA' H253F trajectory (**Figures 7.11** and **7.12**). This is quite important since Ser129 is one of the few polar residues in a very apolar region of the active site pocket. Along the MD trajectories, the solvation shell estimated for the C8 and C9 positions of **1.14** along the MD trajectories showed a more apolar environment for wild-type MalA' and MalA' H253A (i.e. fewer surrounding water molecules) than for MalA' H253F (**Figure 7.15**).

Based on the experimental evidence and computational modeling of the enzyme active site, we propose a mechanism for MalA halogenase (**Figure 3.16, Scheme 3.2**) involving the formation of a Lys108-chloramine intermediate active species, which then interacts with C8 or C9 to carry out an electrophilic aromatic substitution (EAS), with generation of a Wheland intermediate (W) before a final deprotonation step (**Scheme 3.2**). This deprotonation can be effected by a water molecule acting as a base, or in the case of C8 by the well-positioned Ser129. To gain insight on



**Scheme 3.2** Proposed mechanism for MalA halogenation. This mechanism involves serine 129 and active site waters.

the proposed reaction mechanism, density functional theory (DFT) calculations were conducted, using three different computational models (see **Chapter 7** for details). The first model considers only an indole ring and a protonated methyl chloramine as the active species (1a); the second adds a water molecule close to H-C8 or H-C9 positions (1b); and the third model includes a methanol molecule near H-C8 position to mimic Ser129 (1c). Our calculations show that the intrinsic rate-limiting step of the reaction is the initial chlorination, while the deprotonation step occurs slightly faster. A water molecule or methanol interacting with the C8/C9 protons accelerates the chlorination steps because hydrogen bonding enhances the nucleophilicity of these carbons. The computed reaction barrier for C8-chlorination (TS1a-C8) was decreased from 6.4 to 5.5 kcal/mol by a coordinating H<sub>2</sub>O (TS1b-C8), and further decreased to 3.5 kcal/mol when MeOH coordinates to the H-C8 (TS1c-C8). On the other hand, the computed barrier for C9 chlorination (TS1a-C9) decreased from 7.0 to 5.2 kcal/mol by a coordinating H<sub>2</sub>O molecule at H-C9 (TS1b-C9), and to 4.9 kcal/mol when MeOH interacted with H-C8 (TS1c-C9), as shown in **Figure 3.16**. This highlights the role of Ser129 in directing the selectivity towards the formation of the C8 chlorinated product.

The Wa-C9 Wheland intermediate is 1.3 kcal/mol more stable than Wa-C8, but they become almost isoenergetic when coordinating water molecules are considered (Wb-C8 and Wb-C9). An apolar environment favors the formation of the C9 chlorinated malbrancheamide B (**1.15**). Finally, once the Wheland intermediates are formed, re-aromatization by deprotonation occurs rapidly. The computed deprotonation barriers for the two positions are 4.1 and 3.5 kcal/mol for C8 (TS2b-C8) and C9 (TS2b-C9) respectively when a water molecule acts as the base, and 0.6 kcal/mol for C8 when MeOH acts as a base (TS2c-C8).

The DFT optimized structures for the reactant complexes and transition states are highly similar, and the catalytically competent arrangement of Cl-Lys108 near C8 and C9 was found in our MD simulations (represented in **Figures 3.16a, 3.16c**). Taking together our QM models, MD simulations, and the pre-organization of the Cl-Lys108 versus the substrate previously described, we conclude that the proposed reaction mechanism involving a Cl-Lys108 intermediate is the most plausible for the MalA flavin-dependent halogenase.

### 3.2.5 Discussion

These results provide an example of a unique subclass of flavin-dependent halogenases that performs iterative late-stage halogenation of complex substrates independent of a carrier protein. MalA is encoded in a gene cluster containing an NRPS, but the evidence provided herein demonstrates that this protein catalyzes effective late-stage functionalization on free substrates. We propose a new mechanism, involving Ser129, for deprotonation of the positively charged Wheland intermediate in MalA/MalA' halogenation (**Scheme 3.2**). The hydrogen bond between Glu494 and the indole nitrogen is proposed to increase the nucleophilicity of the aromatic ring. This facilitates the EAS reaction, producing the positively charged intermediate. A water molecule or serine side chain can then deprotonate the Wheland intermediate, leading to re-aromatization of the indole ring system (**Scheme 3.2**).

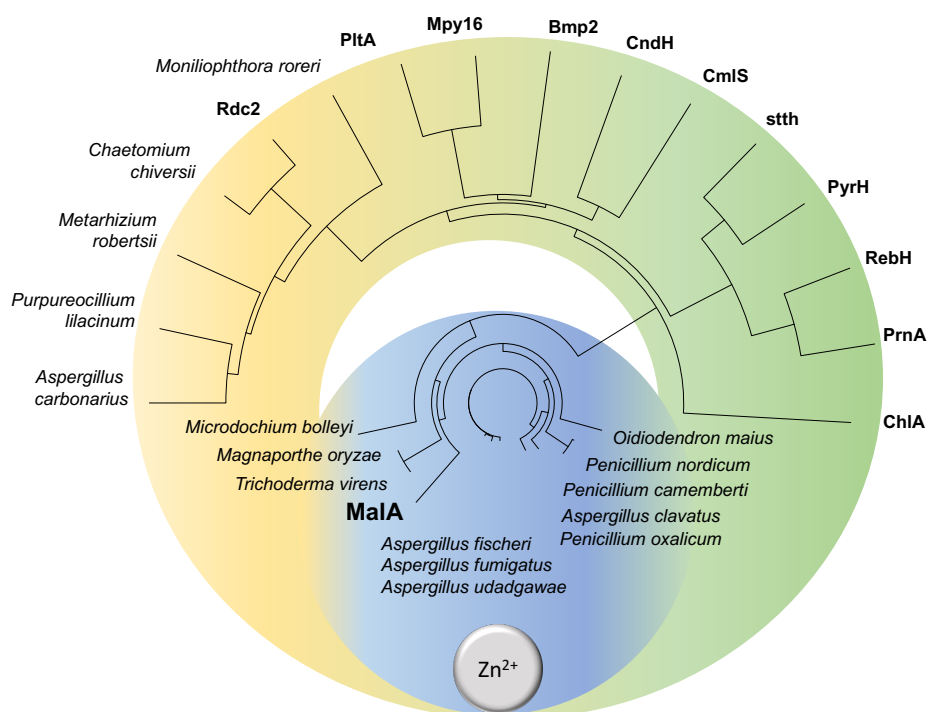
Two critical residues in the active site significantly contribute to substrate binding, Glu494 and Phe489. Glu494 hydrogen bonds with the indole nitrogen and Phe489 facilitates a favorable hydrophobic interaction with the aromatic ring of the indole. The activity of MalA F489H was significantly decreased relative to the wild-type, especially for the second chlorination reaction. We hypothesize that the phenylalanine residue in the back of the active site aids in substrate binding, and maintains interactions with the monochlorinated products to facilitate a second halogenation reaction.

The Michaelis-Menten kinetic parameters displayed equal rates of monochlorination at both the C9 and C8 sites of premalbrancheamide (**1.14**), and equal rates for chlorination of malbrancheamide B (**1.15**) and isomalbrancheamide B (**1.16**), leading to the conclusion that MalA is equally selective for both sites of the indole ring. Interestingly, the catalytic efficiency ( $k_{cat}/K_m$ ) values for the second chlorination were twice those observed for the first chlorination, thus the initial chlorination is proposed to prime the substrate for the second halogenation. This effect can be correlated with the structural data where the chlorine atom on the indole substrates interacts favorably with Phe489, potentially facilitating a better mode of binding. Additionally, the 63renylated63ed malbrancheamide (**1.17**) did not bind in crystals, suggesting a faster dissociation rate for the 63renylated63ed than either of the monochlorinated products, which bound readily in MalA' crystals.

In efforts to elucidate the mechanism of selectivity in MalA, a histidine residue near the catalytic lysine was used to probe the active site region. MalA H253A displayed selectivity for chlorination at the C9 position of premalbrancheamide (**1.14**), while MalA H253F was selective

for the C8 position of premalbrancheamide (**1.14**). MD simulations and DFT calculations demonstrated how key interactions between polar amino acid side chains (Ser129) or water molecules in the active site with the different C–H positions of the indole can control the selectivity of the chlorination reaction. This is accomplished by enhancing the nucleophilicity of the carbon atom during the EAS, but also by pre-organizing a base to carry out the final deprotonation step. An alanine substitution at His253 prevented the interaction of Ser129 and H-C8, leading to an overall apolar active site environment, favoring chlorination at C9. On the other hand, the C8 selectivity observed in MalA H253F can be explained by a more effective Ser129 interaction with H-C8. These findings demonstrate that even small modulations in and around the active site region can lead to significant changes in the site-selectivity of the halogenation reaction.

The designation of MalA as a new class of flavin-dependent halogenases is exemplified not only by its reactivity but also by its unique structural motifs: a Zn<sup>2+</sup>-binding C-terminus and an

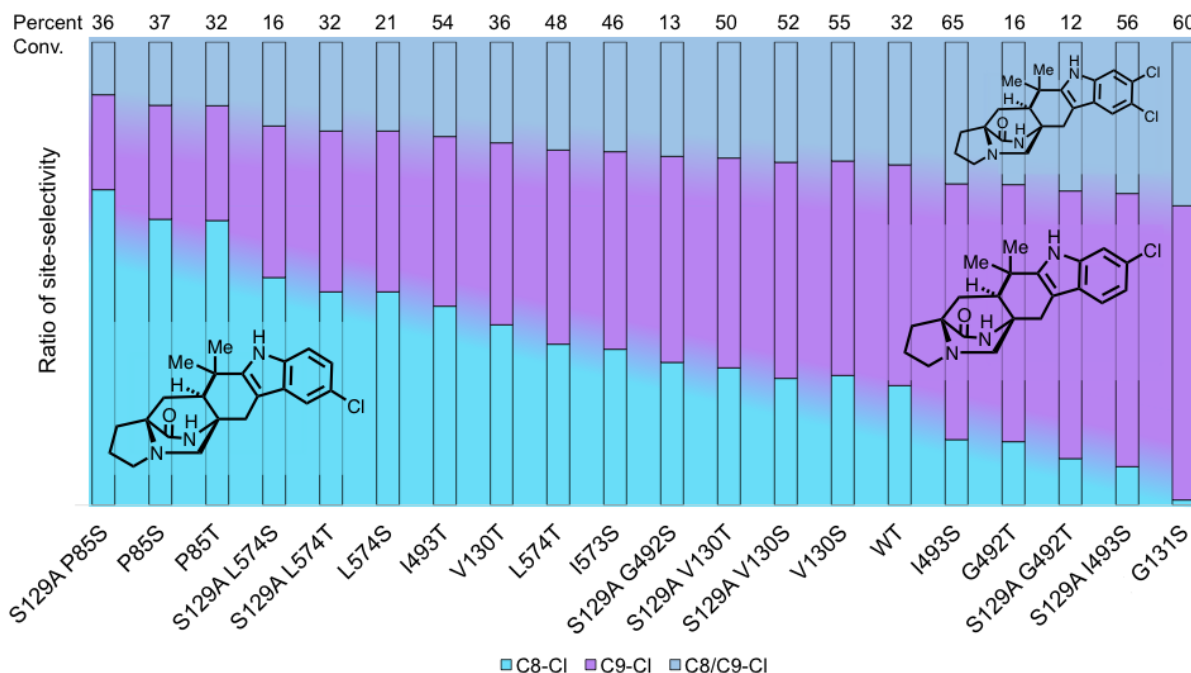


**Figure 3.17** Phylogenetic tree of characterized and uncharacterized (labeled with producing organism) FDHs. Putative fungal FDHs with high amino acid identity (>50%) to MalA cluster in the blue region and are proposed to contain the Zn<sup>2+</sup>-binding motif based on sequence alignments. The green region contains previously characterized bacterial FDHs (the exception is the eukaryotic ChIA) and the yellow region denotes a new subclass of putative FDHs from fungi. Phylogenetic tree prepared using Lasergene MegAlignPro (DNASTAR).

expansive active site, able to accommodate complex substrates. The discovery of this Zn<sup>2+</sup>-binding motif provides a fingerprint for use in mining sequence data for MalA homologs in pursuit of biocatalysts for late-stage halogenation (**Figure 3.17**).<sup>161</sup> Investigation of MalA/MalA' has provided new insights into the biocatalytic mechanism for iterative late-stage halogenation of complex substrates. With a starting model that we have shown is equally reactive toward C8 and C9 of premalbrancheamide (**1.14**), future efforts will focus on rational engineering of MalA to mediate its regioselectivity, substrate scope, and iterative function.

### 3.3 Structure-Based Engineering of MalA for Selective Late-Stage Halogenation

The malbrancheamides are complex hexacyclic fungal indole alkaloids with biological activity as calmodulin antagonists, and the halogenation of the indole ring significantly contributes to the potency of the molecules.<sup>15</sup> MalA has been characterized as an iterative late-stage halogenase that produces brominated and chlorinated malbrancheamide analogs. Our recent collaboration with the Houk group has delved into the mechanism of halogenation that is centered on a pivotal serine residue within the active site.<sup>72</sup> We have utilized this knowledge base to engineer a range of MalA variants for selective halogenation on the natural substrate premalbrancheamide. Structural and computational analysis of the wild-type and mutant MalA have aided in determining a mechanism for modulating the selectivity of the chlorination reaction.



**Figure 3.18** Spectrum of MalA variants. MalA mutants display a broad range of selectivity for the C9 or C8 positions of premalbrancheamide.

The selectivity of the chlorination reaction was modulated by substituting serine and threonine for residues within close proximity of the bound substrate (**Figure 3.18**). The mutations that produced the most significant effect were S129A/I493S, S129A/P85S, and G131S. MalA and MalA' S129A/I493S showed selectivity for the C9 position with 32.9% C9-Cl and 4.6% C8-Cl. MalA and MalA' S129A/P85S showed selectivity for the C8 position with 24.6% C8-Cl and 7.4% C9-Cl. MalA and MalA' G131S showed selectivity for the C9 position with 38.2% C9-Cl. It is hypothesized that the 66renylated66ed malbrancheamide (**1.17**) that is produced is formed from malbrancheamide B (**1.15**) in the case of S129A/I493S and G131S since this monochlorinated molecule is made in excess.

Current efforts are focused on obtaining structures of the mutants to gain insight into the mechanism of the observed switches in selectivity. Only one of the mutants has been crystallized thus far, S129A/P85S, and the substrate appears to shift in the active site pocket to facilitate the C8 chlorination (**Figure 3.19**). When the catalytic serine is mutated to an alanine, the  $\pi$ -interaction between the substrate and F489 is disrupted.

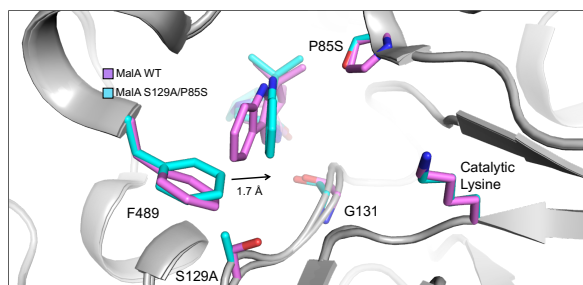
This allows the substrate to move more freely, although G131 still appears to anchor the back end of the substrate in place. The indole ring of the substrate moves 1.7 Å closer to the catalytic lysine to produce the C8-selective chlorination.

The proposed anchoring effect of G131 may shed light on the C9-selective chlorination

observed for MalA G131S. When the glycine backbone no longer anchors the [2.2.2]-diazaoctane ring, the substrate can move more freely in the active site and the inherently more reactive C9 position is halogenated. Crystal structures of the wild-type enzyme compared to variants have aided the visualization of how these mutations change the binding pocket and provide insight on the potential accommodation of unnatural substrates.

### 3.4 Analysis of the Substrate Scope of MalA and Variants

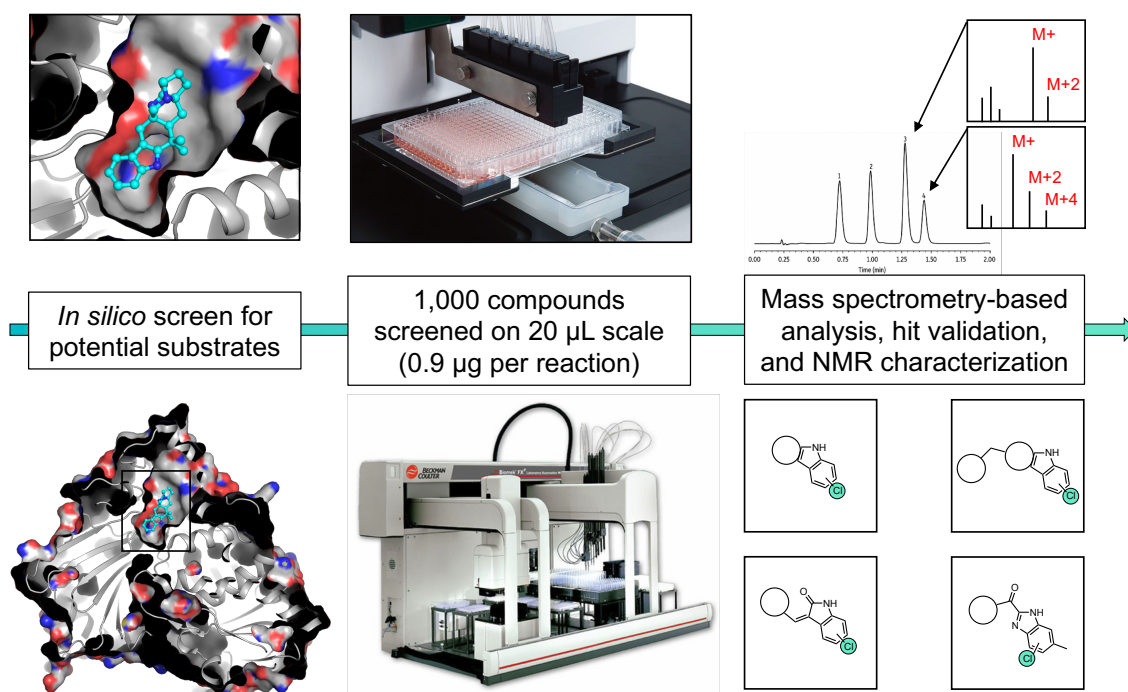
The development of MalA as a biocatalyst for late-stage halogenation of complex heterocycles included a collaboration with Novartis Institutes for BioMedical Research (NIBR) to analyze the substrate scope of the MalA-catalyzed reaction. I worked at NIBR for one month and



**Figure 3.19** MalA' S129A/P85S cocrystal structure with **1.14**.

screened 1,000 computationally predicted substrates from the Novartis library against MalA and a few selected mutants (C9-selective mutants MalA G131S and S129A/I493S and C8-selective mutant MalA S129A/P85S (**Figure 3.19**). The high-throughput assay that was optimized for this work produced a screen with the largest number of reactions on the smallest scale ever performed at NIBR. We ran our reactions on a 20  $\mu\text{L}$  (200  $\mu\text{M}$  substrate) scale in 384-well plates and analyzed the conversion with a three-minute method by LC/MS (**Figure 3.20**).

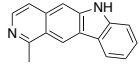
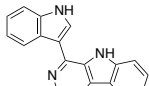
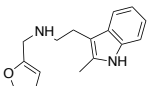
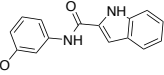
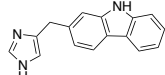
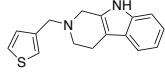
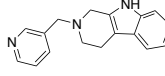
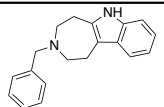
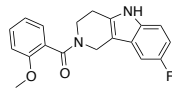
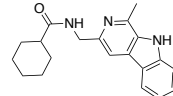
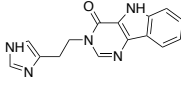
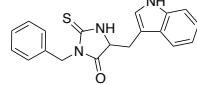
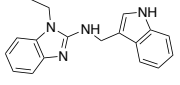
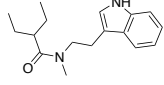
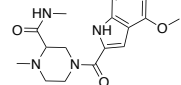
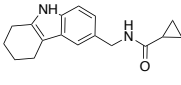
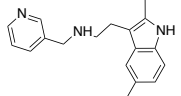
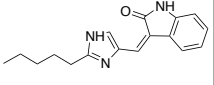
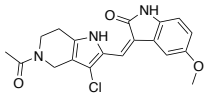
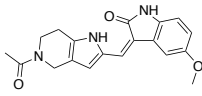
Of the 1,000 compounds screened, 131 were chlorinated to varying degrees. Twenty-three of the compounds displayed 5-10% conversion, thirty displayed 10-20% conversion, twenty displayed 20-30% conversion, and six displayed 30-50% conversion by integration of the extracted ion chromatogram. Future work will focus on characterizing the new chlorinated products by NMR. For the substrates with high conversion, attempts will be made to obtain crystal structures in complex with MalA' and the MalA' mutants.



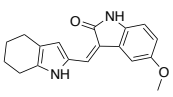
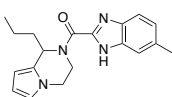
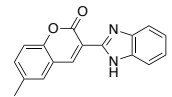
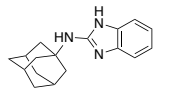
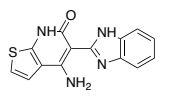
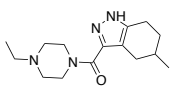
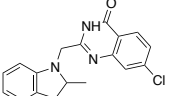
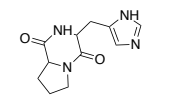
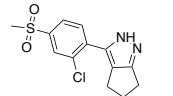
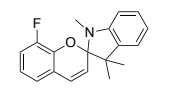
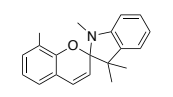
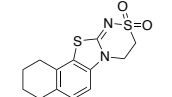
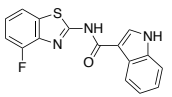
**Figure 3.20** Workflow for halogenation biocatalysis. One thousand substrates were computationally predicted for reactions which were screened in 384-well plates and analyzed by mass spectrometry.



**Table 3.1** Non-proprietary hits from the MaIA high-throughput screen. Monochlorinated products were observed.

					
Formula	C <sub>16</sub> H <sub>12</sub> N <sub>2</sub>	C <sub>19</sub> H <sub>13</sub> N <sub>3</sub>	C <sub>16</sub> H <sub>18</sub> N <sub>2</sub> O	C <sub>15</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	C <sub>16</sub> H <sub>13</sub> N <sub>3</sub>
% Conv.	10	12	3	19, 25, 3	19
Variant	WT	WT	WT	WT, S129A/P85S, G131S	WT
					
Formula	C <sub>16</sub> H <sub>16</sub> N <sub>2</sub> S	C <sub>17</sub> H <sub>17</sub> N <sub>3</sub>	C <sub>19</sub> H <sub>20</sub> N <sub>2</sub>	C <sub>19</sub> H <sub>17</sub> FN <sub>2</sub> O <sub>2</sub>	C <sub>19</sub> H <sub>21</sub> N <sub>3</sub> O
% Conv.	23, 1	23	3, 10, 1, 3	2	2
Variant	WT, G131S	WT	WT, S129A/P85S, S129A/I493S, G131S	WT	WT
					
Formula	C <sub>15</sub> H <sub>13</sub> N <sub>5</sub> O	C <sub>19</sub> H <sub>17</sub> N <sub>3</sub> OS	C <sub>18</sub> H <sub>18</sub> N <sub>4</sub>	C <sub>17</sub> H <sub>24</sub> N <sub>2</sub> O	C <sub>17</sub> H <sub>22</sub> N <sub>4</sub> O <sub>3</sub>
% Conv.	1	4	1	1	3
Variant	WT	WT	WT	WT	WT
					
Formula	C <sub>17</sub> H <sub>20</sub> N <sub>2</sub> O	C <sub>18</sub> H <sub>21</sub> N <sub>3</sub>	C <sub>16</sub> H <sub>17</sub> N <sub>3</sub> O	C <sub>19</sub> H <sub>18</sub> ClN <sub>3</sub> O <sub>3</sub>	C <sub>19</sub> H <sub>19</sub> N <sub>3</sub> O <sub>3</sub>
% Conv.	4	22	3	1	33, 3, 1
Variant	WT	WT	WT	WT	WT, S129A/P85S, G131S

**Table 3.1** Non-proprietary hits from the biocatalytic halogenation high-throughput screen continued.

					
Formula	C <sub>18</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>	C <sub>19</sub> H <sub>22</sub> N <sub>4</sub> O	C <sub>17</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	C <sub>18</sub> H <sub>21</sub> N <sub>3</sub>	C <sub>14</sub> H <sub>10</sub> N <sub>4</sub> OS
% Conv.	3	27, 11	1	5, 1, 1, 1	18, 1
Variant	S129A/P85S	S129A/P85S, G131S	S129A/I493S	WT, S129A/I493S, S129A/P85S, G131S	WT, S129A/P85S
					
Formula	C <sub>15</sub> H <sub>24</sub> N <sub>4</sub> O	C <sub>18</sub> H <sub>16</sub> ClN <sub>3</sub> O	C <sub>11</sub> H <sub>14</sub> N <sub>4</sub> O <sub>2</sub>	C <sub>13</sub> H <sub>13</sub> ClN <sub>2</sub> O <sub>2</sub> S	C <sub>19</sub> H <sub>18</sub> FNO
% Conv.	3	2	2	13, 2, 3, 3	1
Variant	S129A/P85S	WT	WT	WT, S129A/I493S, S129A/P85S, G131S	WT
					
Formula	C <sub>20</sub> H <sub>21</sub> NO	C <sub>13</sub> H <sub>14</sub> N <sub>2</sub> O <sub>2</sub> S <sub>2</sub>	C <sub>16</sub> H <sub>10</sub> FN <sub>3</sub> OS		
% Conv.	1	6	4, 3		
Variant	S129A/P85S	WT	S129A/P85S, G131S		

## Chapter 4

### Molecular Basis for Spirooxindole Formation in the Paraherquamide Biosynthetic Pathway\*

#### 4.1 Introduction to the Paraherquamides and Spirooxindole Formation

The paraherquamide family of natural products contains metabolites isolated from various species of *Penicillium* including *P. paraherquei*,<sup>3</sup> *P. charlesii*,<sup>162</sup> *P. cluniae*,<sup>4</sup> *P. fellutanum*,<sup>55</sup> and *P. IMI 332995*.<sup>28</sup> These molecules are of particular interest due to their antihelmintic therapeutic potential, and various analogs have been developed to improve their pharmacological properties.<sup>12, 14, 163, 164</sup> Derquantel, or 2-deoxy-paraherquamide A, is currently used in combination with established antihelmintic abamectin to combat gastrointestinal nematode infections in sheep.<sup>165</sup> The significant biological activity of this family of molecules is complemented by the intriguing biosynthetic chemistry including late-stage N-oxide formation, unusual oxidative substitutions, and spirooxindole formation.

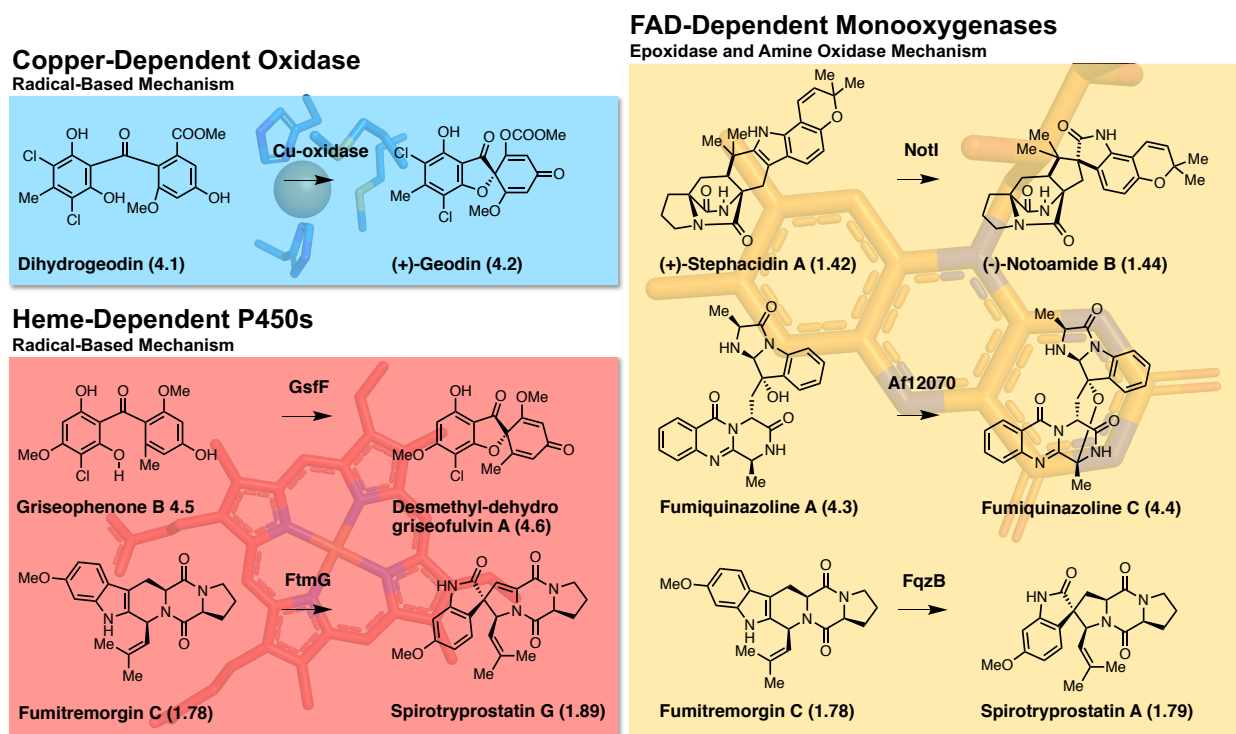
Paraherquamide-type molecules belong to a family of fungal indole alkaloids containing a unique bicyclo[2.2.2]diazaoctane ring, which is formed through a [4+2] intramolecular Diels-Alder (IMDA) reaction.<sup>23</sup> The spirocycle is a common functionality of many molecules within this family including marcfortines A-C,<sup>50, 51</sup> brevianamides A-E,<sup>10, 80</sup> antiinsectan schlerotiamide,<sup>52</sup> and cytotoxic agents citrinadin<sup>166</sup> and notoamides A-C (**Figure 1.9**).<sup>7</sup> The respective enzymes involved in this transformation are proposed to generate an initial C2=C3 indole epoxide with facial selectivity toward the ring system. Controlled collapse of the epoxide generates the spirocycle variants observed within this family. Very few enzymes responsible for this type of reaction within the bicyclo-ring containing molecules have been characterized.<sup>56</sup> NotB catalyzes the 2,3- $\beta$ -face epoxidation of notoamide E to generate notoamides C (**1.39**) and D (**1.38**), which are terminal metabolites, and not substrates for the IMDA to produce the bicyclo[2.2.2]diazaoctane ring (**Scheme 1.3** and **Figure 1.5**). In this pathway, another flavin-dependent monooxygenase NotI performs a presumed 2,3- $\alpha$ -face epoxidation of stephacidin A (**1.42**, **1.43**) in the process of

---

\* Chapter 4 is included in a manuscript in preparation.

generating the bioactive natural product notoamide A (**1.46**, **1.47**). The redundancy of flavin-monooxygenases within this pathway demonstrates that indole epoxidation can occur either before or after the IMDA cyclization. In the pre-IMDA oxidation case, such as that proposed for the brevianamide biosynthetic pathway,<sup>46</sup> the pinacol rearrangement product can still undergo [4+2] cyclization, and the presence of the oxindole significantly contributes to the configuration of the bicyclic product.<sup>64</sup>

Of the currently characterized natural products that bear the spirooxindole functionality, most structures fall outside of this family and the predicted biosynthetic enzymes utilize either metal-dependent radical-based mechanisms or flavin-dependent redox chemistry (**Figure 4.1**). A copper-dependent oxidase is proposed to convert dihydrogeodin (**4.1**) into (+)-geodin (**4.2**) through a radical-based mechanism.<sup>167</sup> An alternative approach involves a flavin-dependent amine oxidase, which converts fumiquinazoline A (**4.3**) to the spirohemiaminal fumiquinazoline C (**4.4**).<sup>168</sup> This is similar to the biosynthesis of tryptoquialanine,<sup>169</sup> which is proposed to be initiated through an indole C2=C3 epoxidation by a flavin-dependent monooxygenase homologous to NotB, followed by oxidation of an amine. For griseofulvin, the unique spirocyclic scaffold is critical for its biological activity. The enzyme responsible is proposed to be a P450,<sup>96</sup> and the diradical



**Figure 4.1** Various enzyme classes that catalyze spirocycle formation. These include copper-dependent oxidases, P450 enzymes, and FAD-dependent monooxygenases.

mechanism involved in the spirocenter formation is similar to that in geodin biosynthesis.<sup>167</sup> Alternatively, the enzyme could perform an epoxidation, and subsequent dehydration and cyclization would convert griseophenone B (**4.5**) into the final product desmethyl-dehydrogriseofulvin (**4.6**). Of the studies detailing spirocycle formation in natural products biosynthesis, work on the spirotryprostatins revealed intricate enzymatic details for this conversion.<sup>19</sup> Two distinct routes for the formation of this spirocyclic moiety involve either a flavin monooxygenase (with homology to NotB), which employs an indole epoxidation route to generate spirotryprostatin A (**1.79**), or a P450 that involves a radical route to generate spirotryprostatins B (**1.85**) and G (**1.89**). Cross talk between the flavin monooxygenase from the fumiquinazoline biosynthetic pathway and enzymes from the fumitremorgin biosynthetic pathway have led to generation of unique diketopiperazine-containing spirocyclic metabolites in *Aspergillus fumigatus*. While these functional studies provided important insights regarding the various mechanisms of formation for the spirocyclized natural products, structural data for the key biocatalysts has remained elusive.

The paraherquamide class of natural products contains a spirooxindole center within a complex heptacyclic ring system. The biosynthesis of these molecules starts with the peptide coupling of L-tryptophan and L- $\beta$ -methyl-proline<sup>24-26</sup> followed by prenylation<sup>27, 31</sup> and Diels-Alder cyclization<sup>23</sup> to produce preparaherquamide (**1.23**)(Scheme 1.2). While precursor incorporation studies demonstrated that **1.23** is incorporated into the final product (-)-paraherquamide A,<sup>27</sup> it was unclear whether the spirocyclization occurred before or after the formation of the pyran and dioxepin rings. Through disruption of *phqK* in *Penicillium simplicissimum* using a CRISPR-Cas9 system (performed by Ying Ye in the Sherman group at the University of Michigan), we were able to identify that **1.23** undergoes a series of oxidation and prenylation reactions to generate paraherquamides K and L (**1.24** and **1.25**) prior to spirocyclization by PhqK.<sup>27, 31, 55</sup> This enzyme has been characterized *in vitro* as a flavin-dependent monooxygenase, which acts on two native substrates to generate the corresponding spirocyclized products. The spirooxindole moiety is stereospecifically formed within the paraherquamides, and based on this observation, we reasoned that the enzyme responsible would bind the substrate in a particular conformation to mediate the highly selective reaction. As previously hypothesized, the molecule is proposed to undergo indole  $\alpha$ -epoxidation, and collapse of the putative epoxide at C3 would generate the pseudo-*para*-quinone methide species.<sup>8</sup> Subsequent migration of the reverse prenyl group from C2 to C3 would generate

the 2-oxindole product. The work described here provides the first structural data for this unique class of spirocycle forming flavin monooxygenases, and sheds light on biocatalyst control of epoxide formation and mode of pinacol rearrangement.

#### 4.2 Biochemical Analysis of PhqK

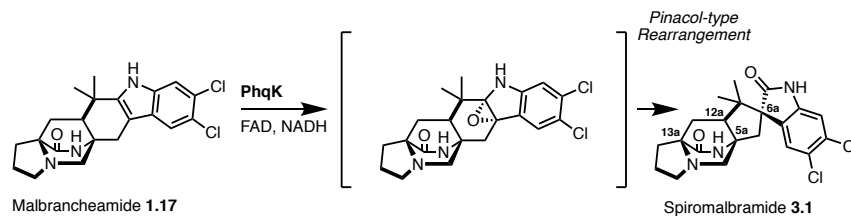
Parallel efforts were pursued *in vitro* and *in vivo* to identify the enzyme responsible for the spirocyclization reaction in paraherquamide biosynthesis. Paraherquamide K (**1.24**) and paraherquamide L (**1.25**) accumulated from the  $\Delta phqK$  strain of *Penicillium simplicissimum*, indicating that these are the natural substrates of the enzyme. The *in vitro* reactions of these new molecules with PhqK confirmed their conversion into the respective spirocyclized products paraherquamide M (**1.26**) and paraherquamide N (**1.27**). This information provided us with evidence to support a new biosynthetic scheme where the pyran and dioxepin ring are both formed prior to the spirocyclization (**Scheme 1.2**). The identification of two natural substrates for PhqK indicates that functionalization of **1.23** is the divergent step between two parallel pathways leading to paraherquamides G (**1.30**) and A (**1.1**). In the first case, a putative prenyltransferase and oxidative enzyme construct a pyran ring, whereas the latter involves the slightly more complex generation of a dioxepin ring.<sup>31</sup> PhqK accepts both substrates, thus the downstream *N*-methyltransferase and  $\beta$ -methyl proline hydroxylase must also display the same level of substrate flexibility.

To obtain further information about the selectivity of PhqK, we employed Michaelis-Menten kinetic analysis (**Table 4.1** and Figures **8.15** and **8.16**). The pyran-containing **1.24** displayed a higher  $V_{\max}$  and  $k_{cat}$ , which is representative of a higher reaction velocity under saturated conditions. In the natural system, however, these secondary metabolites would not be present in solution at such high concentrations, so we turned to the  $K_m$  and catalytic efficiency for insight. The  $K_m$  for **1.25** ( $19.4 \pm 5.2 \mu\text{M}$ ) was significantly lower than that of **1.24** ( $92.1 \pm 29.4 \mu\text{M}$ ). In addition, the  $k_{cat}/K_m$  for **1.25** ( $0.05 \pm 0.01 \mu\text{M}^{-1}\text{min}^{-1}$ ) was slightly higher than for **1.24** ( $0.04 \pm 0.01 \mu\text{M}^{-1}\text{min}^{-1}$ ), indicating that **1.25** may be the favored substrate within the producing organism.

**Table 4.1** Michaelis-Menten kinetic parameters and standard deviations for PhqK reactions.

Michaelis-Menten Kinetic Parameters				
	$V_{\max}$ ( $\mu\text{M}/\text{min}$ )	$k_{cat}$ ( $\text{min}^{-1}$ )	$K_m$ ( $\mu\text{M}$ )	$k_{cat}/K_m$ ( $\mu\text{M}^{-1}\text{min}^{-1}$ )
Paraherquamide K ( <b>1.24</b> )	$0.99 \pm 0.06$	$0.99 \pm 0.06$	$19.37 \pm 5.18$	$0.05 \pm 0.01$
Paraherquamide L ( <b>1.25</b> )	$3.29 \pm 0.36$	$3.29 \pm 0.36$	$92.12 \pm 29.36$	$0.04 \pm 0.01$

The high yields and activity of the enzyme indicated its potential for use as a biocatalyst, thus the substrate scope was

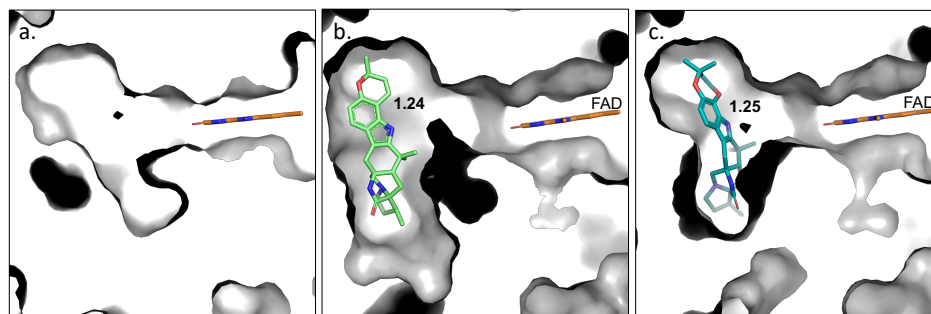


**Figure 4.2** Generation of spiromalbramide via *in vitro* PhqK reaction.

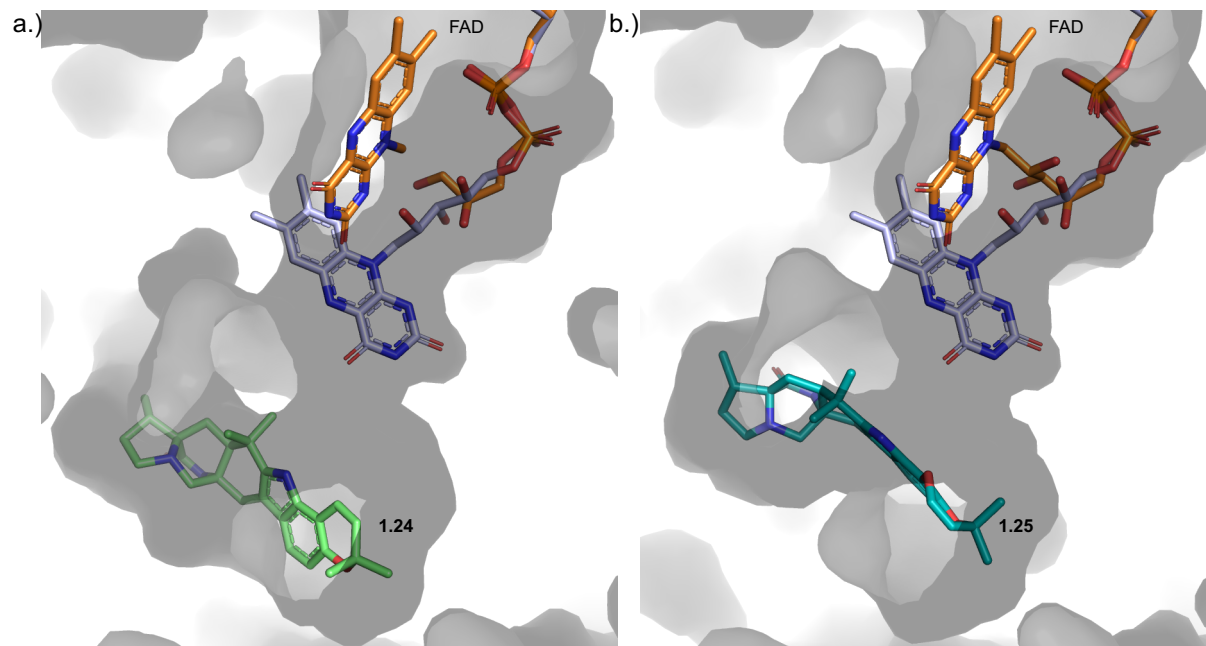
analyzed on a series of structurally related malbrancheamide analogs. The halogenated malbrancheamides are produced by two fungal strains, including the terrestrial strain *Malbranchea aurantiaca*<sup>5</sup> and the marine-derived fungus *Malbranchea graminicola*.<sup>71</sup> The corresponding biosynthetic gene clusters for malbrancheamide and spiromalbramide from these two fungi possess 99% amino acid sequence identity.<sup>72</sup> Intriguingly, the spirocyclized malbrancheamide analog spiromalbramide has only been isolated from the marine strain.<sup>71</sup> In an effort to shed light on this mystery, we tested the reactivity of PhqK on the halogenated malbrancheamides B-D (**1.15**, **1.92**, **1.94**),<sup>72</sup> isomalbrancheamides B-D (**1.16**, **1.93**, **1.95**),<sup>72</sup> and nonhalogenated premalbrancheamide (**1.14**). The reactions generated a variety of spirocyclized malbrancheamides including the natural product spiromalbramide (**3.1**), indicating that PhqK may be highly similar to the flavin monooxygenase involved in *M. graminicola* spiromalbramide biosynthesis (**Figures 4.2**, **8.12**, **8.13**, and **8.14**).<sup>71</sup> PhqK catalyzes formation of products with a relative stereochemical outcome matching that identified in the natural metabolites (5a*R*, 6a*S*, 12a*S*, 13a*S*) and demonstrated its ability to accept unnatural substrates such as the malbrancheamides. With this foundational knowledge, we proceeded to elucidate the molecular basis for stereo-selectivity by analyzing the structure of the enzyme.

PhqK was cocrystallized as four distinct substrate complexes, bearing the natural pyran-containing **1.24** (1.89 Å) and dioxepin-containing **1.25** (2.09 Å), and non-native substrates

monochlorinated malbrancheamide B (**1.15**) (1.69 Å), and monobrominated malbrancheamide C (**1.92**) (1.25 Å). A ligand-free



**Figure 4.3** Expansion of PhqK active site pocket to accommodate substrates. a.) no ligand, b.) parahequamide K **1.24**, c.) parahequamide L substrate **1.25**.

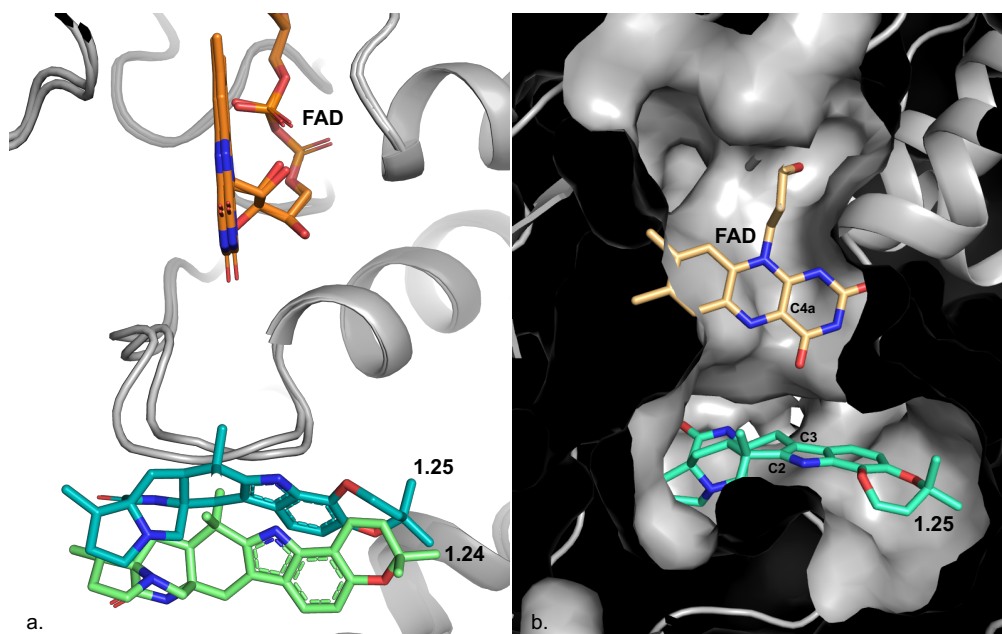


**Figure 4.4** Modeled cofactor dynamics. Flavin dynamics were modeled from HpxO crystal structure (lavender) (PDB: 3rp7) to depict the “in” conformation of the flavin relative to the substrate position. The surface representation of PhqK is shown along with the substrate and FAD from the PhqK structure (orange). Substrates **1.24** (a.) and **1.25** (b.) are shown in green and teal.

structure was also obtained (1.71 Å), aiding in the visualization of how the active site changes to accommodate the hexa- and heptacyclic alkaloids (**Figure 4.3**). The FAD binding pocket and the substrate binding pocket are connected by a long tunnel that appears to put significant distance between the cofactor and the substrate. This can be explained by the well characterized “75renyl” flavin” dynamics described in previous work.<sup>170</sup>

Flavin monooxygenases commonly display a dynamic cofactor with an orientation that is dependent on discrete parameters, including substrate binding or the oxidation state of the flavin. In both the substrate bound and unbound PhqK structures, the flavin was observed in the “out” conformation. It is hypothesized that the “in” conformation may be induced by reduction of the FAD cofactor, but our attempts to obtain a structure bound to reduced flavin using sodium dithionite and NADH were unsuccessful. As an alternative, we modeled the “in” conformation of the cofactor from HpxO into the PhqK structure (**Figures 4.4** and **4.5**).<sup>171</sup> HpxO is an FAD-dependent urate oxidase from *Klebsiella pneumoniae* and its secondary structure overlays well with PhqK (26% sequence identity, RMSD 3.65), with the major difference attributed to the FAD conformation. The divergence of the FAD monooxygenase family is based on evolutionary domain fusion events leading to differences in the substrate binding pocket and reactivity.<sup>68</sup> When





**Figure 4.5** a.) Overlay of PhqK cocystal structures. The structures are complexed with **1.25** (teal) and **1.24** (green) substrates. b.) The “in” FAD cofactor was modeled from the HpxO structure into the PhqK cocystal structure with the **1.25**, where the FAD-C4a is perfectly aligned to perform the epoxidation of the indole C2=C3.

inspecting the overlay of PhqK with HpxO, we observed that the main changes are localized to the substrate-binding region where two loops, one near the cofactor and one near the substrate, are altered in HpxO. Whereas the FAD in the PhqK structure reveals the “out” position, the FAD in HpxO is found as the “in” position, making it a suitable model for generating a visual representation of this phenomenon in PhqK. The “in” flavin from HpxO was modelled into the PhqK structure and fits well in the active site tunnel. In this orientation, the C4a-OOH-FAD would reside sufficiently close to the substrate (**1.25**) to catalyze epoxidation at the indole C2=C3 position (**Figures 4.4 and 4.5**).

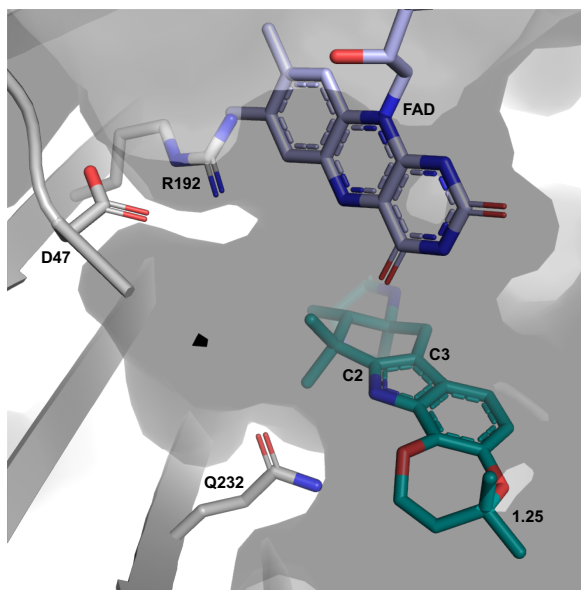
Inspection of the substrates in the PhqK structural complexes shows that each one binds in the same manner except for **1.25**, which is oriented for direct  $\alpha$ -epoxidation. It is possible that the individual substrates sample both conformations in solution. Since **1.25** appears to be the favored substrate, it may have a more prolonged residence in the active conformation, thus we were able to capture that pose in the crystal structure. Computational studies sought to address this question using molecular dynamics simulations.

Site-directed mutagenesis experiments revealed the significance of specific active site amino acid residues (Figures 4.6 and 4.7, Table 4.2). Most of the mutants led to significant losses in flavin incorporation within the protein (indicated as %FAD in Table 4.2) and a decrease in enzyme concentration (indicated as [PhqK] in Table 4.2). A few of the mutants abolished activity, indicating that they perform some catalytic role. Thus, Q232, which interacts with the indole N-H, is necessary for conversion of both the pyran (1.24) and dioxepin (1.25) substrates to the corresponding spirooxindole forms. Substitutions of alanine and glutamate at this site abrogated catalytic

activity toward 1.25, while 1.24 exhibited about 2% conversion. Additionally, R192 and D47 are two essential amino acid residues present within the tunnel between the substrate and cofactor.

When R192 was substituted with alanine and lysine, 1.25 was not converted to product, while 1.24 showed 5-10% conversion. The substitutions of alanine and asparagine for D47 also led to a complete loss of activity.

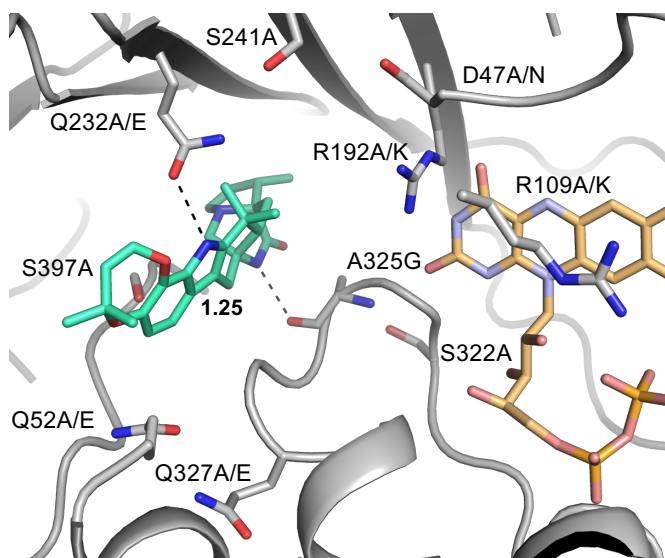
All of these residues are located on the indole N-H side of the substrate and thus



**Figure 4.6** Essential amino acids in PhqK active site. Cocystal structure of PhqK with paraherquamide L (1.25) displaying the active site residues that are pertinent to the activity. The FAD is modeled in the “in” orientation.

**Table 4.2** Reaction data for PhqK variants

Variant	Conversion %		%FAD	[PhqK]
	Paraherquamide K 1.24	Paraherquamide L 1.25		
WT	97	83	100	100
D47A	3	0	22	58
D47N	0	0	48	84
Q52E	20	23	13	33
R192A	11	0	22	81
R192K	5	0	24	98
F219A	35	42	11	49
Q232A	3	0	10	48
Q232E	2	0	10	40
S241A	14	20	22	40
S322A	9	14	20	62
Q327A	22	35	11	31
Q327E	12	29	14	59
S397A	76	89	56	100



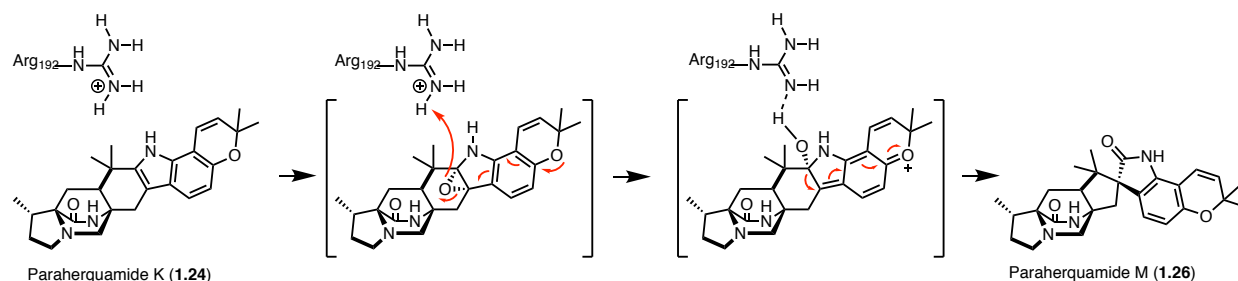
**Figure 4.7** Mutants generated in PhqK.

could play a role in directing the collapse of the epoxide and pinacol rearrangement to this side of the molecule. R192 clashes with the modeled “in” flavin, indicating that it must move in concert with the swing of the cofactor. We hypothesize that R192 moves closer to the substrate upon epoxidation and stabilizes the collapse of the epoxide (**Figure 4.8**). The distance between R192 and the indole C2 is around 7 Å. With a full extension

of R192, a hydrogen bond interaction may be possible with the hydroxyl, stabilizing it to the proper side of the ring system and generating the native chirality of the spirooxindole center (**Figures 4.6 and 4.8**).

### 4.3 Discussion

The spirooxindole moiety occurs in a variety of natural products and its biosynthesis is of significant interest due to the therapeutic potential of these metabolites.<sup>96</sup> The spirocyclic scaffold has increasingly been utilized in the drug development process due to its unique structural attributes.<sup>172</sup> Additionally, the chemistry of the reaction is intriguing from a biochemical standpoint due to the numerous factors contributing to the selectivity. A variety of enzymes have been found to catalyze spirocycle forming reactions including copper-dependent oxidases,



**Figure 4.8** Proposed mechanism for PhqK epoxidation and selective spirocyclization.

cytochromes P450, and FAD-dependent monooxygenases. In the metal dependent enzymes, the reaction is proposed to proceed through a radical-based mechanism, while the flavin -dependent enzymes proceed through FAD redox chemistry. We have characterized the FAD-dependent monooxygenase PhqK as the spirocycle generating enzyme within the pharmaceutically relevant paraherquamide biosynthetic pathway. Through elucidating this key step, we have developed a new biosynthetic scheme for the production of paraherquamides A and G in which spirocyclization occurs after formation of the pyran and dioxepin moieties. We demonstrated that PhqK accepts both **1.24** and **1.25**, confirming the existence of two parallel pathways (branching from **1.23**) within the paraherquamide-producing *Penicillium* strains. While both substrates are accepted by the enzyme, the kinetic data indicate that **1.25**, bearing the dioxepin, is the favored substrate under biologically relevant conditions. This provides evidence that the change in the orientation of **1.25** observed in the PhqK cocrystal structure is due to its persistence in the catalytically competent pose, whereas the less favored substrates might occupy a resting pose until the flavin redox chemistry is initiated. The orientation of the molecule within the active site sets selectivity for  $\alpha$ -epoxidation as shown by the **1.25**-bound structure. Moreover, directed collapse of the initial 2,3-indole epoxide is likely mediated by a positively charged arginine residue, leading to migration of the prenyl group to C3 with oxidation at C2 of the indole.

This work has provided new insights into the structure and mechanism of flavin-dependent spirocycle-generating enzymes, and the evidence supports a selective epoxidation of the indole C2=C3 followed by a controlled collapse of the epoxide to achieve stereo-selectivity. While PhqK catalyzes spirocyclization at a late-stage in paraherquamide biosynthesis, homologous enzymes within the brevianamide, notoamide, and other indole alkaloid pathways can generate pinacol-rearrangement products with varying stereochemistry, or bearing non-spirocyclic functionality. These enzymes are proposed to perform a similar reaction through initial indole epoxidation and subsequent pinacol rearrangement, but the mechanistic basis for selectivity remains to be determined.

## Chapter 5

# Perturbation of the Interactions of Calmodulin with GRK5 Using the Natural Product Malbrancheamide\*

## 5.1 Introduction

### 5.1.1 Introduction to Malbrancheamide as a Chemical Probe

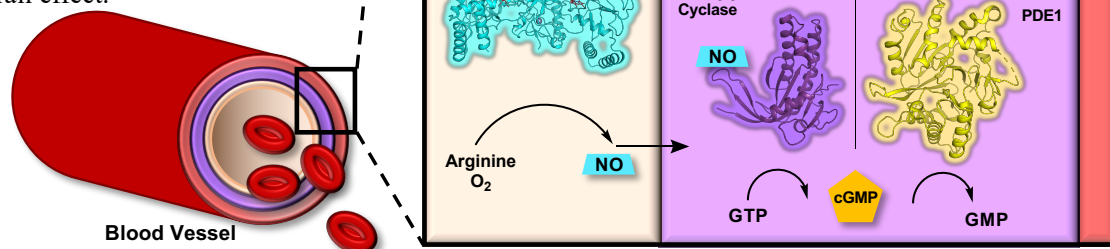
Malbrancheamide (**1.17**) (**Scheme 1.1**) is a natural product that is known to bind calmodulin (CaM),<sup>153</sup> and it has been shown to induce a significant vasorelaxant effect.<sup>15</sup> The mechanism of action of **1.17**-mediated vasorelaxation was proposed to proceed through three possible interactions: CaM–endothelial nitric oxide synthase (eNOS), CaM–80renylated80ed80ase 1 (PDE1), or CaM–myosin light chain kinase (MLCK). Through its interactions with CaM, **1.17** was shown to inhibit CaM-dependent PDE1 and CaM-dependent MLCK,<sup>15,153</sup> inducing vasorelaxation through cGMP/nitric oxide (NO)-dependent pathways. It was hypothesized that **1.17** induces a conformation of CaM that is preferential to CaM–eNOS complex formation, while precluding the formation of complexes with PDE1 and MLCK (**Figure 5.1**).

In terms of analog design, it was determined that the relative stereochemistry of the bicyclo[2.2.2]diazaoctane core and its oxidation states, and the chlorination of the compound significantly impacted the inhibitory biological activity against the CaM–PDE1 complex.<sup>153</sup> Fluorescence titration experiments were used to determine binding constants of **1.17** to CaM, which could be compared to those of approved drugs (e.g. chlorpromazine) that inhibit CaM at the same proposed site. Circular dichroism and protein NMR experiments provided evidence of significant secondary structure changes and particular resonance shifts of CaM upon binding **1.17**. **1.17** was found to exhibit a significant vasorelaxant effect with comparable EC<sub>50</sub> values to known CaM antagonists, and a maximal response of almost 100% relaxation, which is highly unusual for natural products.<sup>15</sup> With the addition of an eNOS inhibitor, the **1.17**-induced vasorelaxation was significantly decreased to about 44%, indicating that the mechanism of action is partially NO-

---

\* Chapter 5 is included in a manuscript: Beyett, T. S.; Fraley, A. E.; Labudde, E.; Patra, D.; Coleman, R. C.; Eguchi, A.; Glukhova, A.; Williams, R. M.; Koch, W. J.; Sherman, D. H.; Tesmer, J. J. G. Perturbation of the Interactions of Calmodulin with GRK5 Using a Natural Product Chemical Probe. **Under Review**.

**Figure 5.1** Mechanism of malbrancheamide in NO/cGMP-dependent vasorelaxation. Malbrancheamide is proposed to allow the binding of  $\text{Ca}^{2+}$ -CaM to eNOS (PDB ID: 3nos) and disrupt the binding of  $\text{Ca}^{2+}$ -CaM to MLCK (PDB ID: 2X4F) and PDE1 (PDB ID: 40JV). cGMP is produced by guanylyl cyclase (PDB ID: 2w01) and is redistributed to myosin light chain (MLC) kinase (PDB ID: 2X4F) and phosphatase pathways, where (de)phosphorylation plays a key role in vasorelaxation. The binding of cGMP to cGMP-dependent protein kinase (PDB ID: 3oD0) changes the intracellular  $\text{Ca}^{2+}$  levels, strengthening the overall effect.



dependent. Previous studies showed that **1.17** also disrupted the formation of CaM–MLCK complex, contributing to the overall vasorelaxation effect.<sup>173</sup>

Various compounds that antagonize CaM via the inhibition of protein-protein interactions with phosphodiesterase 1 and myosin light chain kinase have been previously identified.<sup>174</sup> Many of them are known to bind CaM, and are thought to bind in the same hydrophobic binding pocket, but co-crystal structures with the small molecules presented conflicting results. One of the most potent antipsychotic phenothiazine-type drugs, trifluoperazine, binds to CaM in this hydrophobic pocket and various groups have reported different stoichiometries of binding for these compounds.<sup>175, 176</sup> Another prevalent group of CaM-binding compounds contains a naphthalene core, and includes compounds W-5, J-8, W-7, and A-7 varying in chlorination position and sulfonamide chain length.<sup>174</sup> An NMR solution structure of CaM in complex with W-7 identified the position of binding for these compounds, providing a proposed mechanism of action for their antagonistic effect on CaM.<sup>177</sup> The therapeutics described above are for the treatment of mental illnesses, with side effects coinciding with vasorelaxation, such as decreased blood pressure.<sup>178</sup> These phenothiazine- and naphthalene-type drugs bind CaM and have an inhibitory effect on its

interactions with PDE1 and MLCK, but the previous work neglected to address the interactions of CaM with other proteins such as kinases which are important targets in cardiac therapy. The innovation in this work stemmed from the alternative approach to targeting CaM and its protein-protein interactions, aiming to selectively modulate interactions with various therapeutically relevant enzymes. It was proposed that **1.17** would bind to  $\text{Ca}^{2+}$ -CaM in a mode that would either inhibit or allow protein-protein interactions, leading to an overall vasorelaxant effect.

### *5.1.2 Introduction to the $\text{Ca}^{2+}$ -CaM-GRK5 Complex*

G protein-coupled receptors (GPCR) are the largest family of cell-surface receptors in humans and regulate numerous physiological processes through the activation of heterotrimeric G proteins, GPCR kinases (GRKs), and arrestins. In order to maintain proper physiology, mechanisms to attenuate or cease downstream signaling have evolved. The principal desensitization mechanism is initiated by GRKs, which phosphorylate the third intracellular loops or C-terminal tails of active GPCRs. This process facilitates the recruitment of arrestin which sterically occludes G protein binding and promotes receptor internalization.<sup>179, 180</sup>

As membrane localization and recognition of activated receptors is critical for efficient receptor desensitization, several regulatory mechanisms have evolved to attenuate GRK activity and prolong receptor activation. The ubiquitous  $\text{Ca}^{2+}$ -binding protein calmodulin (CaM) binds six of the seven GRKs and inhibits both membrane localization and receptor phosphorylation in a  $\text{Ca}^{2+}$ -dependent manner.<sup>181-183</sup> The interaction between  $\text{Ca}^{2+}$ -CaM and GRK5 has been the most studied due to its high affinity ( $K_d=8$  nM by surface plasmon resonance) and 50-fold greater potency ( $\text{IC}_{50}=40-60$  nM toward rhodopsin phosphorylation) compared to other GRKs.<sup>181, 182, 184</sup> Immunoprecipitation experiments identified N- and C-terminal helical regions of GRK5 (residues 1-17 and 552-561, respectively) and a stretch of basic residues (residues 20-39) near the junction between the kinase and regulator of G protein signaling homology (RH) domains as regions involved in  $\text{Ca}^{2+}$ -CaM binding.<sup>183, 185</sup>

GRK5 inhibition by  $\text{Ca}^{2+}$ -CaM is particularly significant because it plays a role in the development of maladaptive cardiac hypertrophy, which affects approximately one third of hypertensive individuals and the majority of individuals following a heart attack.<sup>186</sup> After dissociating from membranes, the  $\text{Ca}^{2+}$ -CaM-GRK5 complex translocates to the nucleus where GRK5 phosphorylates histone deacetylase 5 (HDAC5), leading to decreased nuclear levels of HDAC5 and derepression of genes controlled by the transcription factors myocyte enhancer factor-

2 (MEF2),<sup>187</sup> and nuclear factor of activated T-cells (NFAT).<sup>188, 189</sup> Therefore, blocking the translocation of Ca<sup>2+</sup>·CaM–GRK5 to the nucleus represents a potential therapeutic strategy for ameliorating heart failure.<sup>190</sup>

Conflicting models for the Ca<sup>2+</sup>·CaM–GRK5 complex have been proposed. One depicts two Ca<sup>2+</sup>·CaM simultaneously binding distal N- and C-terminal elements of GRK5 in a fashion that would interfere with its membrane binding and receptor phosphorylation.<sup>183</sup> Alternatively, recent structures of GRK5 have revealed the N- and C-terminal elements to be in close proximity to one another, and thus a single Ca<sup>2+</sup>·CaM may bind both simultaneously.<sup>191-193</sup> To better understand the molecular mechanism of inhibition of GRK5 by Ca<sup>2+</sup>·CaM, multiple biophysical methods were used to determine the architecture of the Ca<sup>2+</sup>·CaM–GRK5 complex. Our data are consistent with a single Ca<sup>2+</sup>·CaM binding elements of the kinase critical for receptor recognition, allosteric activation by phospholipids, and membrane association. We confirm that the binding of Ca<sup>2+</sup>·CaM promotes an active conformation of GRK5 resulting in increased phosphorylation of non-receptor substrates. The Ca<sup>2+</sup>·CaM–GRK5 interaction was probed with **1.17**, a natural product with known activity as a Ca<sup>2+</sup>·CaM inhibitor.<sup>5, 153, 194</sup>

Co-crystallization of **1.17** in complex with Ca<sup>2+</sup>·CaM confirmed that the natural product binds to the C-terminal lobe while leaving the N-terminal lobe free to interact with other proteins. Addition of **1.17** selectively inhibited GRK5 autophosphorylation without affecting substrate phosphorylation, thus preventing nuclear translocation of GRK5 and inhibiting the hypertrophic response in cardiomyocytes. This method of inhibition allows **1.17** to suppress maladaptive Ca<sup>2+</sup>·CaM-stimulated GRK5 membrane dissociation without relieving the beneficial ability of Ca<sup>2+</sup>·CaM to suppress receptor phosphorylation.

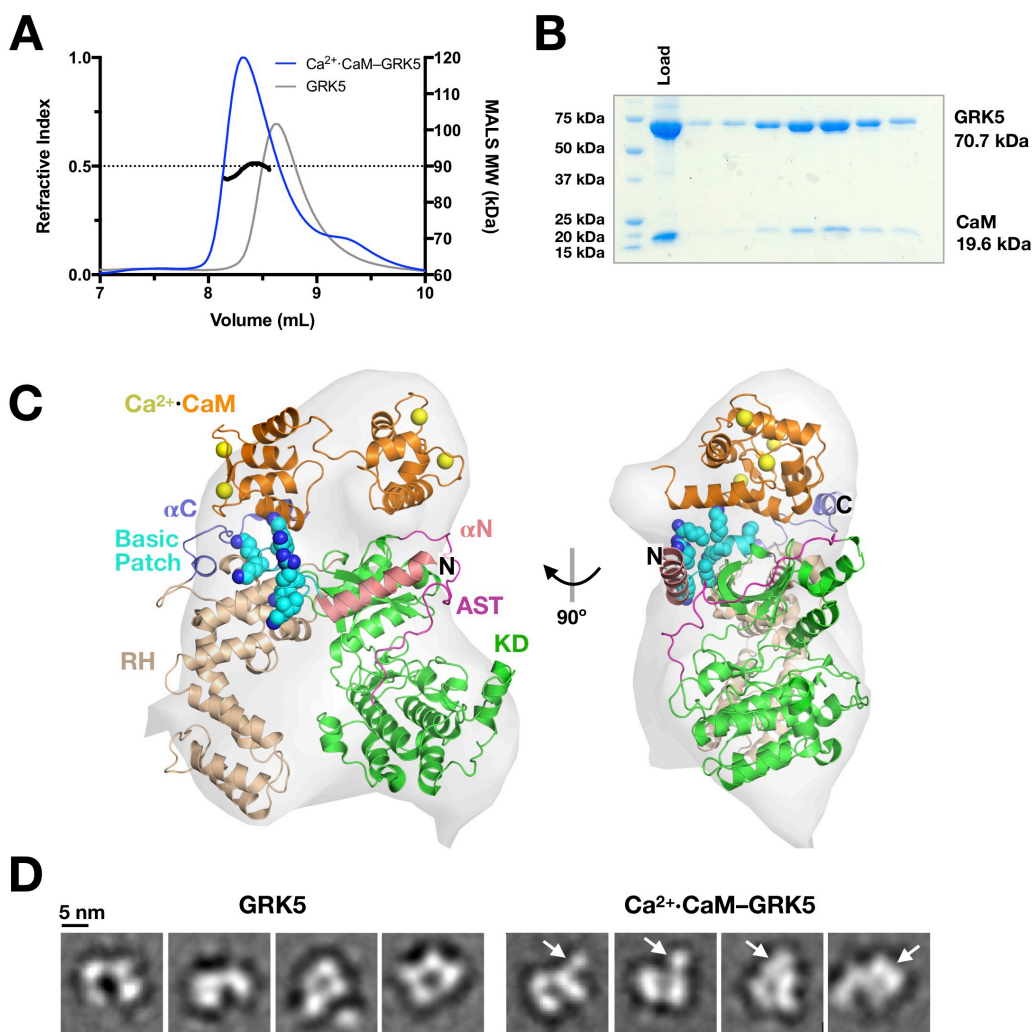
## **5.2 Results**

### *5.2.1 A single Ca<sup>2+</sup>·CaM Binds the RH-kinase Domain Junction of GRK5*

The stoichiometry of the Ca<sup>2+</sup>·CaM–GRK5 complex was determined by Tyler Beyett in the Tesmer group at the University of Michigan via analytical size exclusion chromatography (SEC) coupled to a multi-angle light scattering (MALS) detector. MALS analysis of the peak indicated a molecular weight for the complex of approximately 90 kDa indicating a 1:1 stoichiometry (**Figure 5.2A-B**).



To determine the arrangement of  $\text{Ca}^{2+}\cdot\text{CaM}$  and GRK5 in complex, Tyler performed SEC-small-angle X-ray scattering (SEC-SAXS) (**Figure 9.1, Table 9.1**). Unbiased density for GRK5 alone and the  $\text{Ca}^{2+}\cdot\text{CaM}$ -GRK5 complex were reconstructed *ab initio* (**Figure 9.3**). The crystal structure of  $\text{Ca}^{2+}\cdot\text{CaM}$  (PDB 5J03) and GRK5 (PDB 4WVK)<sup>193</sup> with a modeled N-terminal helix ( $\alpha\text{N}$ ) and active site tether (AST) loop from GRK6 (PDB 3NYN)<sup>195</sup> were automatically docked



**Figure 5.2** Biophysical characterization of the  $\text{Ca}^{2+}\cdot\text{CaM}$ -GRK5 complex. **A)** Representative chromatogram of GRK5  $\pm$   $\text{Ca}^{2+}\cdot\text{CaM}$ . MALS determined the molecular weight of the complex to be 90 kDa (solid line inside the peak), the approximate weight of a 1:1 complex ( $N=2$ ). The molecular weight of His-tagged GRK5 and CaM are 70.7 and 19.6 kDa, respectively. **B)** SDS-PAGE analysis of the SEC chromatogram peak corresponding to the  $\text{Ca}^{2+}\cdot\text{CaM}$ -GRK5 complex. **C)** Reconstructed density at  $3\sigma$  calculated via DENSS from SEC-SAXS scattering curves of the complex at  $10\text{ mg mL}^{-1}$ . GRK5 (PDB entry 4WVK) with terminal regions from GRK6 (PDB entry 3NYN) and  $\text{Ca}^{2+}\cdot\text{CaM}$  (PDB entry 5J03) docked into the envelope. The RGS homology (RH) domain, kinase domain (KD), active-site tether (AST), basic phospholipid binding patch, and terminal helices ( $\alpha\text{N}$  and  $\alpha\text{C}$ ) are labeled. The model fits to the RAW SAXS data well as assessed by CRYSOLO. **D)** Selected negative stain class averages of the  $\text{Ca}^{2+}\cdot\text{CaM}$ -GRK5 complex. Arrows indicate density attributed to  $\text{Ca}^{2+}\cdot\text{CaM}$ .

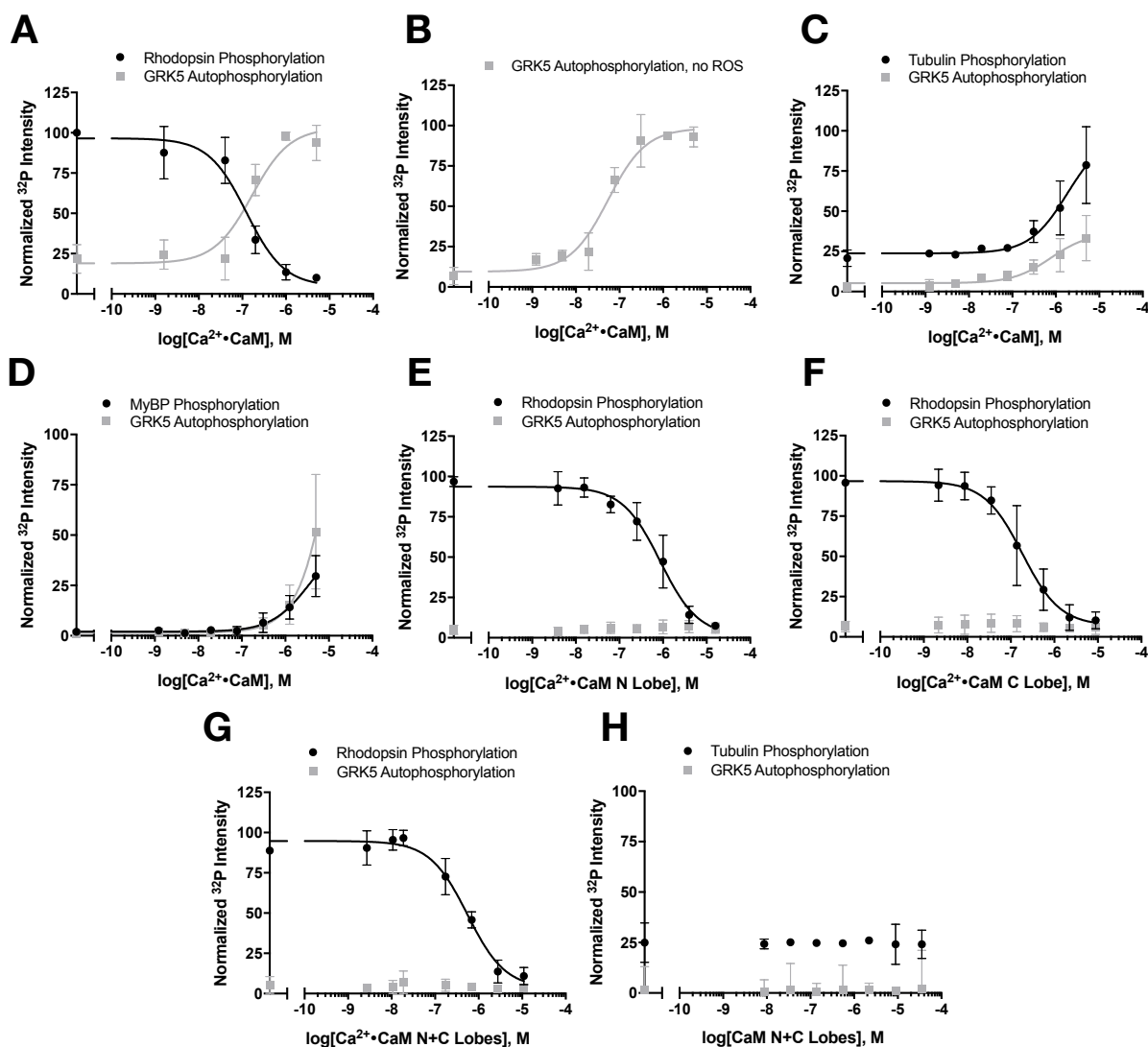
into the density in an unbiased fashion (**Figure 5.2 C**). The C-terminal  $\text{Ca}^{2+}$ -CaM binding site ( $\alpha\text{C}$ ) of GRK5 would be in close proximity to  $\text{Ca}^{2+}$ -CaM given the trajectory of the preceding residues in the GRK5 structure and has been modeled as such. Compared to the density for GRK5 alone (**Figure 9.3**), extra density was observed upon the addition of  $\text{Ca}^{2+}$ -CaM. The surplus density was automatically docked with  $\text{Ca}^{2+}$ -CaM at the junction of the regulator of G protein signaling homology (RH) and kinase domains (KD) of GRK5. This area is near known binding sites of  $\text{Ca}^{2+}$ -CaM, including the N-terminus and the basic phospholipid binding patch of GRK5.<sup>196</sup> The positioning of  $\text{Ca}^{2+}$ -CaM in this configuration does not obscure the substrate or ATP binding pockets, which is consistent with the observation that GRK5 retains catalytic activity toward soluble substrates and autophosphorylation when in complex with  $\text{Ca}^{2+}$ -CaM.<sup>181</sup> Dhableswar Patra characterized GRK5 and the  $\text{Ca}^{2+}$ -CaM-GRK5 complex by negative stain EM and generated reference-free 2D class averages (**Figure 9.4**). Particles of the  $\text{Ca}^{2+}$ -CaM-GRK5 complex revealed additional protein density in a similar position to that indicated in the SAXS experiments, along the surface spanning RH and kinase domains and N-lobe of the kinase domain (**Figure 5.2D**).

#### 5.2.2 $\text{Ca}^{2+}$ -CaM Activates GRK5 and Promotes Autophosphorylation

As previously reported,  $\text{Ca}^{2+}$ -CaM inhibits GRK5-mediated receptor phosphorylation while simultaneously promoting autophosphorylation (**Figure 5.3A**) and enhances phosphorylation of synucleins.<sup>182, 197</sup> Our structural analyses suggest inhibition of receptor phosphorylation is likely the result of  $\text{Ca}^{2+}$ -CaM interacting with regions of GRK5 critical for both receptor interaction and/or membrane association. However, autophosphorylation is not simply a consequence of membrane dissociation, as  $\text{Ca}^{2+}$ -CaM stimulates autophosphorylation in the absence of membranes or other substrates (**Figure 5.3B**). Whereas  $\text{Ca}^{2+}$ -CaM inhibits receptor phosphorylation, GRK5 activity towards soluble substrates is enhanced suggesting that  $\text{Ca}^{2+}$ -CaM binding stabilizes an active kinase conformation (**Figure 5.3C-D**), while autophosphorylation is attenuated due to competition between the soluble substrate and the C-terminus of GRK5.  $\text{Ca}^{2+}$ -CaM is even able to promote phosphorylation of myelin basic protein (MyBP), which is not phosphorylated by GRK5 in the absence of  $\text{Ca}^{2+}$ -CaM (**Figure 5.3D**) further demonstrating the activating effect of this protein-protein interaction.

To dissect the roles of the individual domains of  $\text{Ca}^{2+}$ -CaM, Tyler studied the effects of its independent N- and C-terminal lobes (residues 1-75 and 78-149, respectively) which each contain a pair of  $\text{Ca}^{2+}$ -binding EF hand motifs. Both less potently inhibited receptor phosphorylation than

full-length  $\text{Ca}^{2+}\cdot\text{CaM}$ , consistent with the fact that binding of even a portion of  $\text{Ca}^{2+}\cdot\text{CaM}$  along the putative membrane interacting surface would interfere with membrane association or receptor binding and diminish receptor phosphorylation (**Figure 5.3E-F**). Addition of both independent

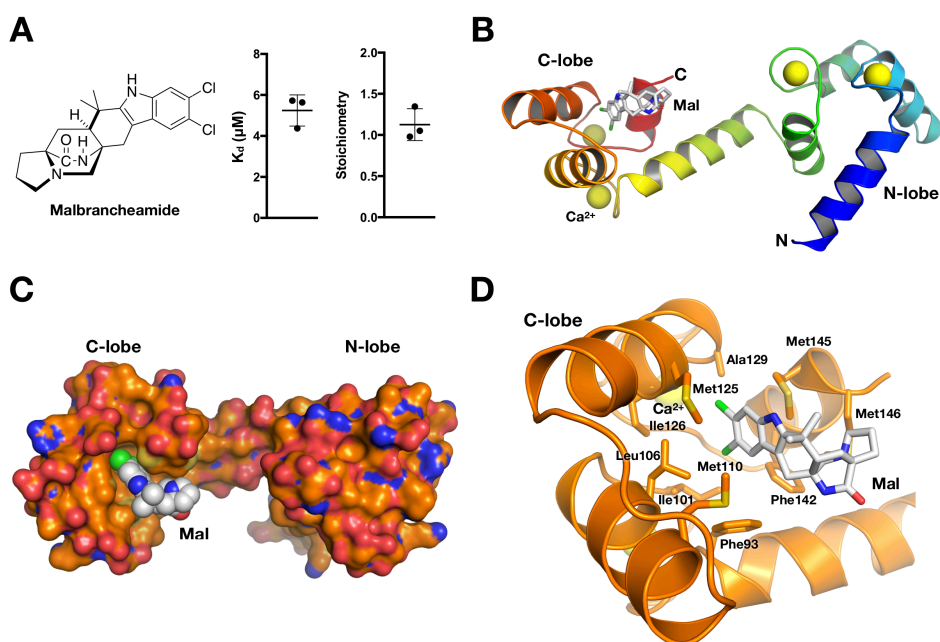


**Figure 5.3** Modulation of substrate phosphorylation by  $\text{Ca}^{2+}\cdot\text{CaM}$ . **A**) Rhodopsin in rod outer segments (ROS, 5  $\mu\text{M}$ , light activated) phosphorylation ( $\text{IC}_{50}=120 \pm 40$  nM) and concurrent GRK5 (50 nM) autophosphorylation ( $\text{EC}_{50}=170 \pm 70$  nM) in the presence of  $\text{Ca}^{2+}\cdot\text{CaM}$ . **B**) Stimulation of GRK5 (50 nM) autophosphorylation by  $\text{Ca}^{2+}\cdot\text{CaM}$  ( $\text{EC}_{50}=50 \pm 20$  nM) in the absence of ROS or other substrates. **C**) Tubulin (10  $\mu\text{M}$ ) phosphorylation in the presence of  $\text{Ca}^{2+}\cdot\text{CaM}$ . Decreased levels of concurrent autophosphorylation are likely due to competition between the C-terminus of GRK5 and tubulin. **D**) Phosphorylation of myelin basic protein (MyBP, 7  $\mu\text{M}$ ) presence of  $\text{Ca}^{2+}\cdot\text{CaM}$  and concurrent autophosphorylation. **E**) Inhibition of rhodopsin phosphorylation in the presence of the  $\text{Ca}^{2+}\cdot\text{CaM}$  N-lobe ( $\text{IC}_{50}=860 \pm 230$  nM), **F**) C-lobe ( $\text{IC}_{50}=180 \pm 60$  nM), or **G**) a 1:1 molar combination of N and C lobes ( $\text{IC}_{50}=600 \pm 170$  nM). All radiometric kinase assays were performed three times and reported as mean  $\pm$  SD and fit to a sigmoidal dose-response model in GraphPad Prism with the Hill slope constrained to 1.

lobes of  $\text{Ca}^{2+}\cdot\text{CaM}$  was not sufficient to stimulate autophosphorylation (**Figure 5.3G**) or increase GRK5 activity toward soluble substrates (**Figure 5.3H**). Together, these data suggest that the binding of full-length  $\text{Ca}^{2+}\cdot\text{CaM}$  allosterically activates GRK5 and induces an autophosphorylation-prone conformation.

### 5.2.3 Malbrancheamide Binds the C-lobe of Calmodulin and Inhibits $\text{Ca}^{2+}\cdot\text{CaM}$ -Stimulated GRK5 Autophosphorylation

To dissect how each lobe of CaM is involved in binding the N- and C-terminal domains of GRK5, we turned to malbrancheamide, a halogenated fungal indole alkaloid that antagonizes  $\text{Ca}^{2+}\cdot\text{CaM}$  (**Figure 5.4A**)<sup>153</sup> and is predicted to bind to the C-terminal lobe, inducing a significant vasorelaxation effect.<sup>5, 153, 194</sup> We first used isothermal titration calorimetry (ITC) to confirm the binding affinity and stoichiometry of malbrancheamide for  $\text{Ca}^{2+}\cdot\text{CaM}$ . The affinity ( $K_d$ ) of malbrancheamide was determined to be  $5.2 \pm 0.8 \mu\text{M}$ , consistent with prior reports,<sup>153</sup> and it was found to bind  $\text{Ca}^{2+}\cdot\text{CaM}$  with a 1:1 stoichiometry (**Figure 5.4A, Figure 9.5**).



**Figure 5.4** Malbrancheamide (**1.17**) binds to the C-lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$ . **A**) Chemical structure of **1.17** (Mal) and ITC binding results (average  $\pm$  S.D.,  $N=3$ ). **B**) Rainbow ribbon representation of the  $\text{Ca}^{2+}\cdot\text{CaM}\cdot\mathbf{1.17}$  structure in a *trans* conformation.  $\text{Ca}^{2+}$  are shown as yellow spheres (PDB entry 6EEB). **C**) Surface representation of  $\text{Ca}^{2+}\cdot\text{CaM}$  with **1.17** shown as spheres. **D**) C-lobe hydrophobic pocket with **1.17** bound. Side chains making van der Waals contacts with **1.17** are labeled and are similar to those contacting isomalbrancheamide D (**1.95**) (**Figure 9.6**).

We next determined the co-crystal structure of the  $\text{Ca}^{2+}\cdot\text{CaM}\cdot\text{malbrancheamide}$  (**Table 9.3**), wherein a single malbrancheamide molecule was bound with its dichlorinated indole region buried in the hydrophobic pocket of the C-lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$  (**Figure 5.4B-D**, **Figure 9.6**). The halogenated indole is positioned between Leu106, Met125, and Met145 and the C-8 chlorine forms a Cl- $\pi$  interaction with Phe93. Only Met145, which occupies the same location as malbrancheamide in ligand-free structures of  $\text{Ca}^{2+}\cdot\text{CaM}$ , is poorly resolved in the C-lobe binding pocket. Malbrancheamide buries approximately 250 Å<sup>2</sup> of surface area in the hydrophobic pocket (PISA server),<sup>194</sup> consistent with ligands of similar affinities.<sup>198</sup> The N-terminal lobe, however, remains unbound and is free to interact with other proteins (**Figure 5.4B**). Addition of malbrancheamide to phosphorylation reactions containing  $\text{Ca}^{2+}\cdot\text{CaM}$ -GRK5 and rhodopsin resulted in inhibition of GRK5 autophosphorylation but did not recover receptor phosphorylation (**Figure 5.5A-B**, **Figure 9.7**), consistent with  $\text{Ca}^{2+}\cdot\text{CaM}$  being simultaneously bound to malbrancheamide and GRK5. Malbrancheamide also had no effect on phosphorylation of the soluble substrate tubulin (**Figure 5.5C**), but similarly inhibited  $\text{Ca}^{2+}\cdot\text{CaM}$ -stimulated autophosphorylation in these reactions.

#### *5.2.4 The N-lobe of $\text{Ca}^{2+}\cdot\text{CaM}$ Interacts with the N-terminus of GRK5 to Inhibit Receptor Phosphorylation*

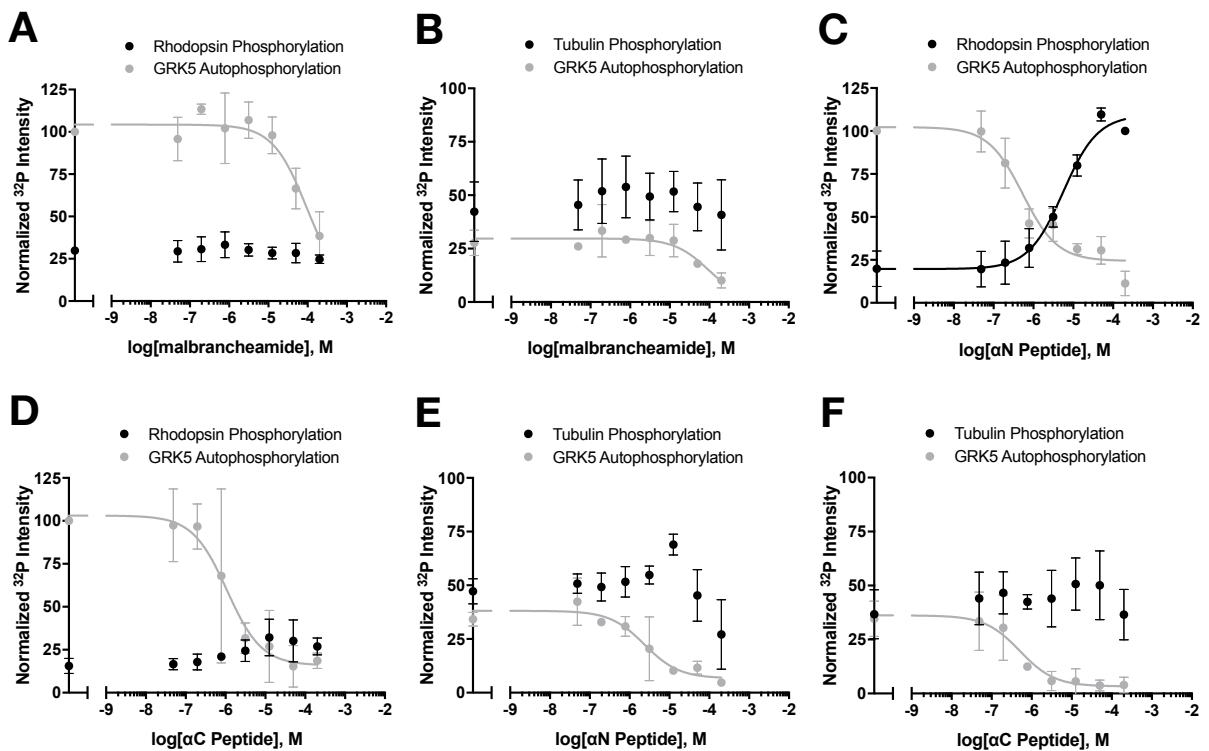
To dissect how each lobe of CaM is involved in binding the N- and C-terminal binding sites of GRK5, we turned to **1.17**, a halogenated fungal indole alkaloid that antagonizes  $\text{Ca}^{2+}\cdot\text{CaM}$  (**Figure 5.4A**)<sup>153</sup> and that is predicted to bind to the C-terminal lobe, inducing a significant vasorelaxation effect.<sup>5, 153, 194</sup> We first used ITC to confirm the binding affinity and stoichiometry of **1.17** for  $\text{Ca}^{2+}\cdot\text{CaM}$ . The affinity ( $K_d$ ) of **1.17** was determined to be  $5.2 \pm 0.8 \mu\text{M}$ , consistent with prior reports,<sup>153</sup> and binds  $\text{Ca}^{2+}\cdot\text{CaM}$  with a 1:1 stoichiometry (**Figure 5.4A**, **Figure 9.5**).

We next determined the co-crystal structure of the  $\text{Ca}^{2+}\cdot\text{CaM}\cdot\text{1.17}$  complex (**Table 9.3**), wherein a single **1.17** molecule is observed bound with its dichlorinated indole region buried in the hydrophobic pocket of the C-lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$  (**Figure 5.4B-D**, **Figure 9.6**). The halogenated indole is positioned between Leu106, Met125, and Met145 and the C-8 chlorine forms a Cl- $\pi$  interaction with Phe93. Only Met145, which occupies the same location as **1.17** in ligand-free structures of  $\text{Ca}^{2+}\cdot\text{CaM}$ , is poorly resolved in the C-lobe binding pocket. **1.17** buries approximately 250 Å<sup>2</sup> of surface area in the hydrophobic pocket (PISA server),<sup>199</sup> consistent with ligands of similar affinities.<sup>198</sup> The N-terminal lobe, however, is free to interact with other proteins (**Figure**

**5.4B).** Addition of **1.17** to phosphorylation reactions containing  $\text{Ca}^{2+}\cdot\text{CaM-GRK5}$  and rhodopsin resulted in inhibition of GRK5 autophosphorylation but did not recover receptor phosphorylation (**Figure 5.5A-B, Figure 9.7**), consistent with  $\text{Ca}^{2+}\cdot\text{CaM}$  being simultaneously bound to **1.17** and GRK5. **1.17** also had no effect on phosphorylation of the soluble substrate tubulin (**Figure 5.5C**), but similarly inhibited  $\text{Ca}^{2+}\cdot\text{CaM}$ -stimulated autophosphorylation in these reactions.

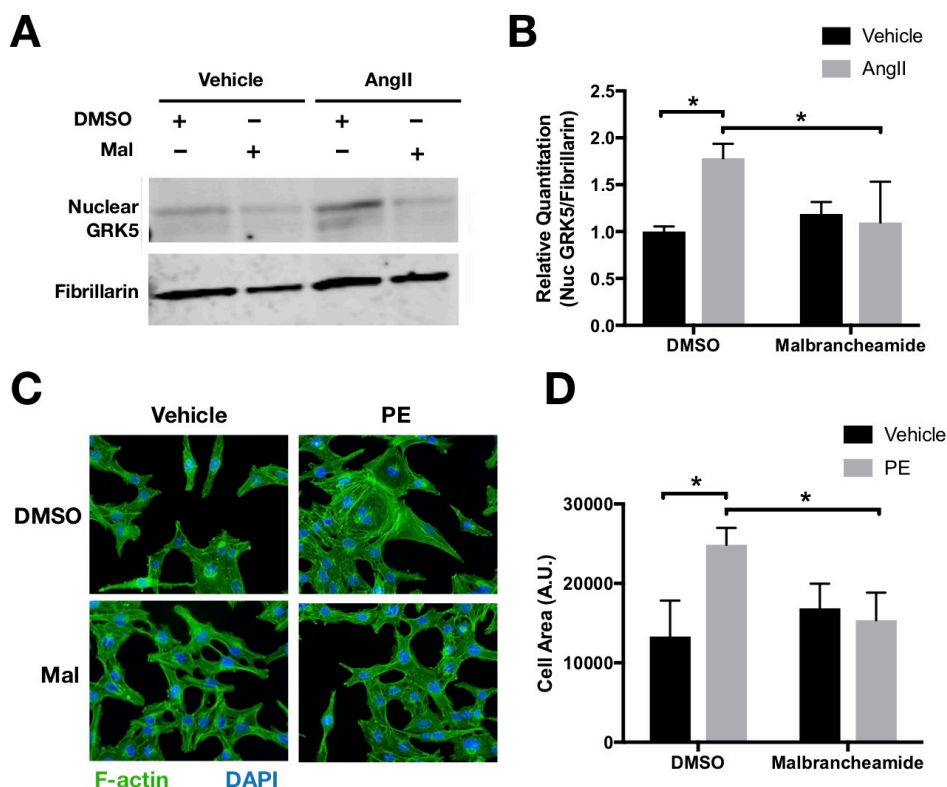
### 5.2.5 Malbrancheamide Inhibits Nuclear Translocation of GRK5 and Prevents Hypertrophy

We hypothesized that **1.17** binding to the C-lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$  would release the membrane binding C-terminus of GRK5 (**Figure 5.2C**) and reduce its nuclear translocation where its primary non-canonical activity is to promote cardiac cell hypertrophy.<sup>190</sup> To test this, we collaborated with Ryan Coleman and Akito Eguchi in the Koch group at Temple University. They first cultured



**Figure 5.5** Effects of malbrancheamide (**1.17**) and GRK5 peptides on  $\text{Ca}^{2+}\cdot\text{CaM}$  modulation of GRK5 activity. **A)** Rhodopsin (5  $\mu\text{M}$ , light activated) phosphorylation in the presence of 500 nM  $\text{Ca}^{2+}\cdot\text{CaM}$  and increasing concentrations of **1.17**. In these assays,  $\text{Ca}^{2+}\cdot\text{CaM}$  was preincubated with **1.17** prior to the addition of GRK5. **B)** Tubulin (5  $\mu\text{M}$ ) phosphorylation and in the presence of 500 nM  $\text{Ca}^{2+}\cdot\text{CaM}$  and varying concentrations of **1.17**. **C)** Rhodopsin phosphorylation in the presence of 500 nM  $\text{Ca}^{2+}\cdot\text{CaM}$  and varying concentrations of GRK5  $\alpha\text{N}$  (residues 2-31,  $\text{IC}_{50,\text{autophosphorylation}}=540 \pm 240$  nM,  $\text{EC}_{50,\text{rhodopsin}}=5.6 \pm 1.6$   $\mu\text{M}$ ) or **D)**  $\alpha\text{C}$  (residues 546-565,  $\text{IC}_{50}=1.1 \pm 0.7$   $\mu\text{M}$ ) peptides. **E)** Tubulin phosphorylation in the presence of 500 nM  $\text{Ca}^{2+}\cdot\text{CaM}$  and varying concentrations of GRK5  $\alpha\text{N}$  ( $\text{IC}_{50}=2.4 \pm 1.7$   $\mu\text{M}$ ) or **F)**  $\alpha\text{C}$  peptides ( $\text{IC}_{50}=500 \pm 440$  nM). All radiometric kinase assays were performed three times and reported as mean  $\pm$  SD and fit to a sigmoidal dose-response model in GraphPad Prism with the Hill slope constrained to 1.

primary neonatal rat cardiac fibroblasts and pretreated cells with 1  $\mu$ M **1.17** or DMSO as a control. This was followed by induction of  $\text{Ca}^{2+}$ -CaM-GRK5-dependent nuclear translocation and accumulation with the GPCR-dependent hypertrophic agonist, angiotensin II (AngII).<sup>200</sup> **1.17** treatment prevented nuclear accumulation of GRK5 after addition of AngII compared to non-



**Figure 5.6** Effects of malbrancheamide (**1.17**) on GRK5 nuclear translocation and cardiomyocyte hypertrophy. **A**) Representative western blot of nuclear GRK5 following pre-treatment with **1.17** (Mal) and stimulation with the  $\text{Ca}^{2+}$ -CaM activator and hypertrophic agonist angiotensin II (AngII). Rat neonatal cardiac fibroblasts were pretreated with 1  $\mu$ M Mal (or DMSO as control) prior to treatment with 5  $\mu$ M AngII (or vehicle as control) for 90 min. Nuclear extracts were then prepared and GRK5 was blotted with fibrillarlin as a nuclear marker and protein loading control. **B**) Quantitation of western blot results showing significant nuclear localization of GRK5 after AngII (compared to vehicle) and this was completely blocked by pre-treatment with Mal (average  $\pm$  SD,  $n=3$  biological replicates, two-way ANOVA with Tukey's post hoc correction, \*,  $p<0.05$ ). **C**) Representative image of F-actin fluorescent stained (green) cultured cardiomyocytes following pre-treatment with Mal (or DMSO as negative control) and then treated with the hypertrophic  $\alpha$ -adrenergic agonist phenylephrine (PE). Myocyte nuclei are stained with DAPI (blue). After overnight pre-treatment of 1  $\mu$ M of Mal (or DMSO), AC16 cardiomyocytes were stimulated with 50  $\mu$ M PE (or vehicle for non-stimulated controls) for 48 h to induce cellular hypertrophy. Larger cells are clearly seen only in the PE-DMSO sections. **D**) Quantitation of myocyte cell area was performed using F-actin stained cells and Image-J software. Three independent experiments were carried out and 50-100 cells per experiment were measured for each included condition. Shown is the average  $\pm$  SD (\*  $p<0.05$ , two-way ANOVA with Tukey's post hoc correction).

stimulated vehicle-treated controls, as determined by quantitative Western blotting of purified nuclear fractions of cardiac fibroblasts (**Figure 5.6A-B**).

In a second set of experiments to confirm **1.17** inhibition of GRK5 nuclear accumulation, Ryan and Akito examined cardiomyocyte hypertrophy in response for the GPCR hypertrophic agonist, phenylephrine (PE).<sup>200</sup> Pretreatment of a cardiomyocyte cell line with **1.17** or DMSO as a control was followed with 48 hrs of PE or vehicle treatment and myocyte hypertrophy was determined by F-actin staining and imaging quantitation of cell area (**Figure 5.6C-D**). As shown, **1.17** indeed could block the cell hypertrophy induced by PE. These data are consistent with a mechanism in which **1.17** releases the sequestered C-terminus of GRK5 and thereby promotes the retention of GRK5 at the plasma membrane in a state that is unable to phosphorylate GPCRs.

### 5.3 Discussion

The role of GRK5 in heart failure and cardiac hypertrophy has stimulated interest in both therapeutic development and the underlying molecular mechanisms of disease.<sup>190</sup> Recent studies have shed light on the role of GRK5 regulation by  $\text{Ca}^{2+}$ -CaM at the cellular and physiological levels concerning non-canonical GRK5-mediated cardiac hypertrophy,<sup>187, 188, 190,201</sup> but the molecular basis of the interaction has remained poorly understood. Here, we present structural and functional data that reveal a single  $\text{Ca}^{2+}$ -CaM binds to GRK5 in a manner that interferes with receptor recognition and membrane association while priming the kinase for phosphorylation of soluble substrates which may contribute to maladaptive hypertrophy.

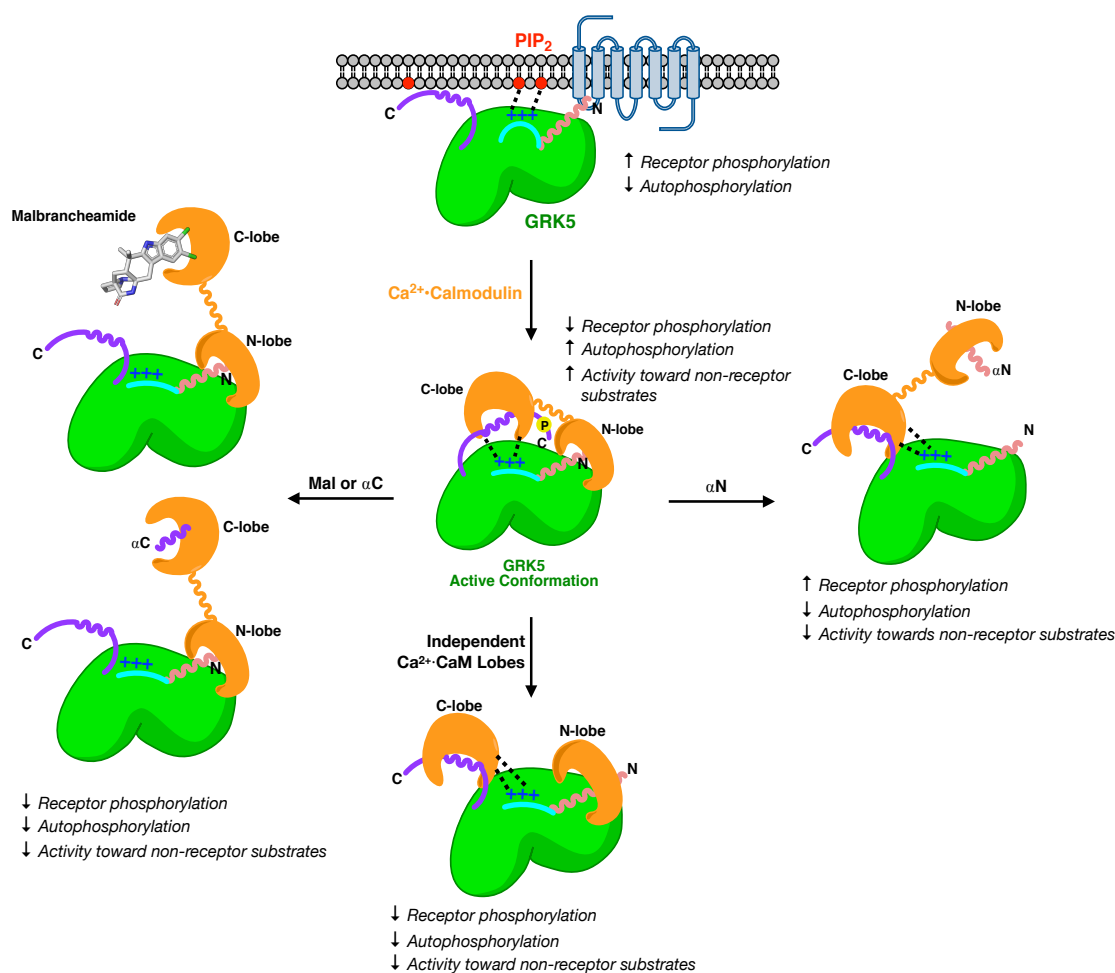
Substrates of  $\text{Ca}^{2+}$ -CaM often contain amphipathic helices with bulky hydrophobic side chains which anchor the hydrophobic pockets of each lobe of  $\text{Ca}^{2+}$ -CaM.<sup>201, 202</sup> The terminal helices and basic patch of GRK5 resemble canonical  $\text{Ca}^{2+}$ -CaM substrates and thus, our model suggests that terminal hydrophobic residues of the  $\alpha\text{N}$  and  $\alpha\text{C}$  helices are likely involved in binding to  $\text{Ca}^{2+}$ -CaM with additional contributions from electrostatic interactions with the acidic helices of  $\text{Ca}^{2+}$ -CaM (**Figure 5.7**). The  $\alpha\text{N}$  helix has been proposed in some models to dock into the hydrophobic core of activated GPCRs, and is well known to be critical for receptor recognition and phosphorylation.<sup>203, 204</sup> It is also thought to bind to the kinase domain and stabilize a closed, active conformation.<sup>195</sup>

Our results demonstrate that  $\text{Ca}^{2+}$ -CaM binds to the small lobe of GRK5 in a manner that receptors. A C-terminal amphipathic helix in GRK5 significantly contributes to membrane association,<sup>205, 206</sup> and our data suggest that the GRK5  $\alpha\text{C}$  helix is close enough to be sequestered



by  $\text{Ca}^{2+}\cdot\text{CaM}$  which is bound by one lobe to the kinase domain. This alters membrane association, which is important for efficient GRK5-mediated phosphorylation of GPCRs.<sup>207</sup>

Although  $\text{Ca}^{2+}\cdot\text{CaM}$  potentially inhibits receptor phosphorylation, the activity of GRK5 towards soluble substrates is enhanced. It has been previously reported that  $\text{Ca}^{2+}\cdot\text{CaM}$  stimulates phosphorylation of synucleins,<sup>197</sup> and this appears to apply to additional soluble substrates suggesting that the binding of  $\text{Ca}^{2+}\cdot\text{CaM}$  stabilizes an activated state of GRK5. As suggested by the GRK6 crystal structure in complex with the nucleoside inhibitor sangivamycin,<sup>195</sup> the activated state of GRKs is stabilized by a network of contacts between the  $\alpha\text{N}$  helix, AST loop, and RH domains. Since our SAXS and EM data provide evidence that one lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$  binds near



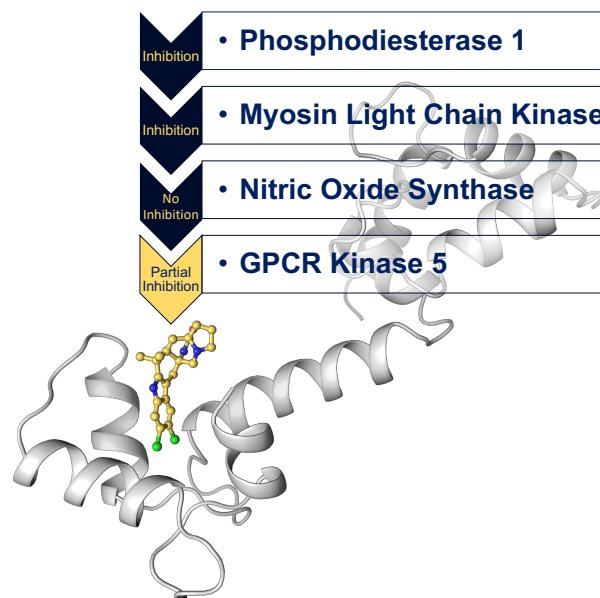
**Figure 5.7** Model for domain-specific regulation of GRK5 by  $\text{Ca}^{2+}\cdot\text{CaM}$ . Features critical for receptor docking ( $\alpha\text{N}$  helix shown as pink wavy line) and membrane association (purple C-terminus of GRK5 and the N-terminal basic patch, shown as a blue line with + signs) are sequestered by  $\text{Ca}^{2+}\cdot\text{CaM}$  but also stabilize an active conformation of the kinase. Both lobes in the intact protein are required for this activation because simultaneous addition of the individual lobes does not promote an activated state of GRK5. The addition of  $\alpha\text{N}$  peptide,  $\alpha\text{C}$  peptide, or malbrancheamide (**1.17**) are all able to relieve various aspects of regulation through differential binding to the two lobes of  $\text{Ca}^{2+}\cdot\text{CaM}$ .

these structural elements, stabilization of this interaction network likely explains the increase in GRK5 activity towards its soluble substrates. Addition of the isolated lobes of  $\text{Ca}^{2+}\cdot\text{CaM}$ , however, did not activate kinase activity at any concentration tested, indicating that full-length  $\text{Ca}^{2+}\cdot\text{CaM}$  is required for stabilizing this active conformation of GRK5.

$\text{Ca}^{2+}\cdot\text{CaM}$ -stimulated autophosphorylation has been mapped to non-canonical residues in the extreme C-terminus of GRK5 (residues 579-590) that, similarly to phosphorylation by protein kinase C,<sup>208</sup> further reduces the ability of GRK5 to associate with membranes. This may represent a mechanism to prolong inhibition of receptor phosphorylation and nuclear targeting after  $\text{Ca}^{2+}\cdot\text{CaM}$  has dissociated.<sup>209</sup> Cleavage of the helix connecting the two lobes of  $\text{Ca}^{2+}\cdot\text{CaM}$  abolished  $\text{Ca}^{2+}\cdot\text{CaM}$ -stimulated autophosphorylation, suggesting that intact  $\text{Ca}^{2+}\cdot\text{CaM}$  positions the terminal elements of GRK5 in a specific conformation that may enable *cis*-autophosphorylation. A likely mechanism for this “repositioning” may simply be loss of membrane interactions of the C-terminal region of GRK5 via the binding of  $\text{Ca}^{2+}\cdot\text{CaM}$ , which would free this disordered region to instead enter the active site of the kinase domain. However, in the absence of membranes and  $\text{Ca}^{2+}\cdot\text{CaM}$ , GRK5 still does not display robust autophosphorylation. Thus, it is likely that  $\text{Ca}^{2+}\cdot\text{CaM}$  both redirects the path of the C-terminus toward the active site while stabilizing an activated conformation of GRK5 leading to robust autophosphorylation.

Determination of the roles of the individual lobes of  $\text{Ca}^{2+}\cdot\text{CaM}$  was aided by **1.17**, a natural product that we show binds solely to the C-lobe. This implies that addition of the compound would block the interactions of GRK5 with the C-lobe. Indeed, the addition of **1.17** or the  $\alpha\text{C}$  peptide inhibited kinase autophosphorylation, but did not affect the inhibition of receptor phosphorylation mediated by  $\text{Ca}^{2+}\cdot\text{CaM}$ . Therefore, the C-lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$  interacts with the C-terminus of GRK5, which dissociates the kinase from membranes, whereas the N-lobe interacts with the  $\alpha\text{N}$  helix to interfere with receptor phosphorylation. It has been previously noted that rhodopsin phosphorylation is similarly inhibited by an antibody that recognizes epitopes in the N-terminus of GRK1 without affecting soluble substrate phosphorylation.<sup>210</sup> Thus,  $\text{Ca}^{2+}\cdot\text{CaM}$  may be considered a GRK5 agonist and GPCR mimic with respect to the phosphorylation of soluble substrates and the primary inducer of its non-canonical nuclear effects.

Extraction of GRK5 from the plasma membrane by  $\text{Ca}^{2+}$ ·CaM is thought to allow the nuclear localization sequence located in the large lobe of the kinase to promote nuclear translocation and pathological gene activation.<sup>205, 211</sup> The ability of  $\text{Ca}^{2+}$ ·CaM to activate GRK5 against soluble targets indicates that it likely augments its activity in the nucleus. **1.17** has a defined binding mode that only blocks C-lobe interactions. Our data show that this binding mode allows **1.17** to block nuclear translocation of GRK5 at low micromolar concentrations induced by AngII, a known hypertrophic agonist that previously was shown to be a potent stimulator of nuclear GRK5 accumulation.<sup>200</sup> Further, **1.17** was also shown to prevent cardiomyocyte hypertrophy induced by PE, where GRK5 has been shown to potentiate this effect due to its non-canonical nuclear activity.<sup>200</sup> Thus, **1.17** represents an effective chemical probe for studying  $\text{Ca}^{2+}$ ·CaM–GRK5-dependent hypertrophic signaling in cells, and may facilitate delineation of the specific regulatory contribution of the C-lobe in other  $\text{Ca}^{2+}$ ·CaM protein-protein interactions. Of note,  $\text{Ca}^{2+}$ ·CaM is known to contribute to cardiomyocyte hypertrophy in other pathways, thus the specific mechanisms by



**Figure 5.8** Summary of malbrancheamide biological activity as a selective inhibitor of  $\text{Ca}^{2+}$ ·CaM protein-protein interactions.

which **1.17** prevents AngII or PE induced cellular hypertrophy (either GRK5-dependent or other) require further investigation. Previous work has demonstrated the biological activity of **1.17** as an inhibitor of  $\text{Ca}^{2+}$ ·CaM-PDE1  $\text{Ca}^{2+}$ ·CaM-MLCK, inducing vasorelaxation through cGMP/NO-dependent pathways (**Figure 5.8**).<sup>15, 153</sup> This work demonstrates that the selective inhibitory effect of this natural product on  $\text{Ca}^{2+}$ ·CaM-GRK5 may act in concert with these previously characterized activities to combat cardiac hypertrophy.

## Chapter 6

### Summary, Discussion, and Future Directions

#### ***6.1 Characterization of the Bifunctional Reductase/Diels-Alderase Responsible for the Formation of the Bicyclo[2.2.2]diazaoctane Ring***

##### *6.1.1 Summary and Discussion of Work Completed*

After many years of proposals and predictions for the biosynthesis of the bicyclo[2.2.2]diazaoctane ring, we have discovered an unexpected bifunctional reductase/Diels-Alderase which catalyzes the [4+2] cyclization. An *in vitro* reconstitution of the full biosynthetic pathway for malbrancheamide revealed each step required for the formation of these unique monoketopiperazine bicycle-containing molecules. By utilization of chiral LC/MS, we were able to resolve an enantiomeric mixture of *syn*-premalbrancheamide, indicating that while the synthetic route produces racemic premalbrancheamide, the fungi produce only (+)-premalbrancheamide. Thus, the biosynthetic machinery was also predicted to perform enantioselective synthesis, and this hypothesis was validated through *in vitro* assays. Reductive cleavage of an L-Pro-L-Trp dipeptide from the MalG NRPS followed by reverse prenylation and a cascade of post-NRPS reactions culminated in an intramolecular [4+2] hetero-Diels-Alder cyclization to furnish the bicyclo[2.2.2]diazaoctane scaffold. Enzymatic assembly of optically pure (+)-premalbrancheamide involves an unexpected zwitterionic intermediate upon which MalC catalyzes enantioselective cycloaddition as a bifunctional NADPH-dependent reductase/Diels-Alderase. Structures of MalC and the homolog PhqE provided key details which were validated through site-directed mutagenesis. These IMDA enzymes are the first to demonstrate catalytic residues, dependence on a cofactor, and the retention of their original function for the selective cyclization process.

##### *6.1.2 Future Directions*

The reactions involved in malbrancheamide biosynthesis all involve complex enzymes that direct their selectivity. While the NRPS would not be a suitable substitute for traditional synthetic peptide coupling methods, the other enzymes in the pathway have elicited much interest for the development as biocatalysts. The prenyltransferase can selectively reverse 95renylated the indole

C2 position, and Diels-Alder cyclization is the most commonly used reaction and the most studied mechanism in organic chemistry, thus the IMDA enzymes MalC and PhqE could be useful biocatalysts for stereo-controlled [4+2] cyclization. The substrate scope of these enzymes has yet to be tested, and in the event that they do not display a broad substrate scope, the crystal structures of the enzymes will provide invaluable insight into potential mutations for structure-based engineering. Additionally, these enzymes both generate the *syn* bicycle, and the characterization of an *anti*-selective IMDA enzyme from this family of molecules would provide strong structural evidence for the observed selectivity.

It is believed that the citrinalin IMDA enzyme produces secondary metabolites with the *anti* configuration. After initial experiments with CitL in the reconstitution assay, it was determined that a racemic mixture of premalbrancheamide was produced. This indicates that CitL may perform the IMDA on a different substrate from MalC and PhqE, but it does not eliminate the concept that CitL may still perform the reduction just not bind the reduced product in a manner conducive to selective cyclization. Based on previous computational studies, it was determined that an oxindole substrate had a higher propensity to generate the *anti* bicycle in a substrate-controlled manner. It is possible that there is an “early-stage” FMO in this pathway which generates the pinacol-rearranged oxindole which would then serve as the substrate for the IMDA enzyme CitL. CitC is the likely candidate for this epoxidation/pinacol rearrangement reaction and it has relatively high homology to PhqK (36% DNA sequence identity and 53% amino acid identity), NotI (42% DNA sequence identity and 61% amino acid identity), and NotB (44% DNA sequence identity, and 63% amino acid identity). CitC has a slightly higher identity to the early-stage FMO NotB, providing evidence that it may indeed act at an early stage as well. Reactions with premalbrancheamide and preparaherquamide performed by Stela Romminger, a postdoctoral researcher in the Sherman group at the University of Michigan, showed no conversion, providing further evidence that CitC and CitL are unique from the enzymes involved in the formation of the *syn* bicycle containing molecules. I propose that CitC acts on the prenyldipeptide, which then serves as the substrate for the IMDA enzyme to generate the *anti* relative bicycle in citrinalin. This enzyme system provides the potential to further diversify the class of fungal indole alkaloids.

## **6.2 Late-stage Halogenation by MalA**

### *6.2.1 Summary and Discussion of Work Completed*

The malbrancheamides are complex fungal indole alkaloids with therapeutic potential as calmodulin antagonists, and the halogenation of the indole ring significantly contributes to the biological activity of these molecules. MalA has been characterized as an iterative late-stage halogenase that provides the halogen moieties to produce brominated and chlorinated malbrancheamide analogs. While synthetic methods have achieved limited late-stage halogenation of premalbrancheamide, inserting a bromine atom at the C9 position, MalA can both chlorinate and brominate the C9 and C8 positions as well as perform a second halogenation reaction. We have delved into the mechanism of enzymatic halogenation through structural and computational studies, and utilized this knowledge base to engineer a range of MalA variants for selective halogenation on the natural substrate premalbrancheamide. To demonstrate the utility of halogenation biocatalysts, we analyzed the substrate scope of MalA and various mutants on 1,000 substrates from the Novartis compound library. Halogenation was achieved on a broad range of substrates with variation in number of products, degree of halogenation, and percent conversion.

### 6.2.2 Future Directions

Through the high throughput screen at Novartis Institutes for BioMedical Research (NIBR), we identified that MalA displays a broad substrate scope and are currently working on characterizing these new halogenated molecules. These scale-up reactions can be performed through the already optimized *in vitro* method or a method for whole cell catalysis could be applied. Whole cell catalysis has been used previously in our lab for large scale conversions by P450 enzymes. Like MalA, these enzymes require a reductase partner, indicating that the requirement for a partner protein may not hinder our efforts for whole cell reactions. Additionally, cofactor regeneration systems could be incorporated to make sure that the necessary reaction components are readily available. We had previously attempted MalA reactions in lysate with and without cofactors or the reductase partner and found similar low conversions in all cases indicating that the low concentration of the protein in the lysate may pose issues requiring resolution by the *in vitro* system.

Ultimately, the MalA work goes beyond just the structural characterization of these new molecules. This enzyme's fascinating display of activity for halogenating complex drug-like compounds has demonstrated its utility as a biocatalyst and generated interest by industrial synthetic biology teams. Compared to other halogenating enzymes, MalA is unique in that it naturally halogenates complex molecules in an iterative fashion. This is an excellent foundation

for initiating an enzyme evolution campaign to generate a suite of halogenating biocatalysts which would greatly serve the chemistry community. The synthetic biology team at NIBR has indicated an interest in continuing this project and utilizing their technology to bring this project to an industrially relevant scale. While we have optimized large-scale protein production via a 30L fermenter, our efforts for large-scale catalytic conversion will be greatly aided by the team at NIBR. I have generated a small library of halogenase variants through structure-based engineering of MalA, but the high-throughput capacity at NIBR would allow for the production and analysis of an infinite number of mutants. This efficient methodology would facilitate the development of selective halogenation biocatalysts for complex substrates for which synthetic methods cannot achieve selectivity.

### ***6.3 Molecular Basis for Spirooxindole Formation in the Paraherquamide Biosynthetic Pathway***

#### *6.3.1 Summary and Discussion of Work Completed*

The paraherquamides are potent antihelmintic natural products with complex heptacyclic scaffolds. One key feature of these molecules is the spirooxindole moiety, which lends a strained 3D conformation to these otherwise fairly planar molecules. The flavin monooxygenase PhqK was found to catalyze spirocycle formation through two parallel pathways in the biosynthesis of paraherquamides A and G. Two new paraherquamides (K and L) were isolated upon generation of the  $\Delta phqK$  strain of *Penicillium simplicissimum*, and subsequent *in vitro* reactions with PhqK generated two additional metabolites paraherquamides M and N. While the enzyme accepted both paraherquamides K and L, kinetic analyses indicated that the dioxepin-containing paraherquamide L was the favored substrate, and cocrystal structures of PhqK in complex with various substrates provided a foundation for mechanistic analyses and computational studies. The mechanistic studies are ongoing and continue to provide us with insight into the selectivity of spirocycle formation. We have identified key amino acids that are essential for catalysis, including an arginine residue that may facilitate the breakdown of the epoxide.

#### *6.3.2 Future Directions*

PhqK performs a late-stage oxidative rearrangement, which adds a unique scaffold to the family of fungal indole alkaloids. There are additional enzymatic systems, such as those involved in citrinalin and brevianamide biosynthesis, which perform the epoxidation and rearrangement at

an early stage. Structural and biochemical characterization of these enzymes will provide valuable insight into the variations which dictate the early or late-stage reactivity.

We have also found that PhqK can accept unnatural substrates including the halogenated malbrancheamides. With this display of substrate scope and efficient reactivity, it is believed that PhqK could be utilized as a biocatalyst for spirooxindole formation. Additionally, the ability to engineer this enzyme through a structure-based approach has only briefly been ascertained, and further mutagenesis could be utilized to generate a variant that either accepts substrates with a broader range of complexity, or generates the opposite stereochemical outcome. The wild-type PhqK and variants can be used in conjunction with the other characterized enzymes from these biosynthetic families to generate a library of novel and complex indole alkaloids. This family of molecules already displays a wide array of biological activities and the library of molecules could be tested for antihelmintic, anticancer, and vasodilation activity.

#### ***6.4 Perturbation of the Interactions of Calmodulin with GRK5 Using the Natural Product Malbrancheamide***

##### *6.4.1 Summary and Discussion of Work Completed*

GRKs are responsible for initiating desensitization of activated GPCRs. Considered to be a “keystone protein,” GRK5 is known to modulate and interconnect multiple signaling pathways involved in cardiovascular and neurodegenerative disorders. GRK5 is potently inhibited by the calcium-sensing protein calmodulin (CaM), which leads to nuclear translocation of GRK5 and promotion of cardiac hypertrophy. This work reported the architecture of the  $\text{Ca}^{2+}\cdot\text{CaM}\text{-GRK5}$  complex determined by small-angle X-ray scattering and negative stain electron microscopy.  $\text{Ca}^{2+}\cdot\text{CaM}$  was found to bind primarily to the small lobe of the kinase domain of GRK5 near elements critical for receptor interaction and membrane association, thereby inhibiting receptor phosphorylation while activating the kinase for phosphorylation of soluble substrates. To define the roles of each lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$ , we utilized the natural product malbrancheamide as a chemical probe. Malbrancheamide had previously been shown to inhibit  $\text{Ca}^{2+}\cdot\text{CaM}$ , but the intricate details of its inhibitory effect have only just been realized through this work. We first used X-ray crystallography to show that malbrancheamide selectively binds to the C-terminal lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$ , and then employed a battery of functional assays to show that each lobe of  $\text{Ca}^{2+}\cdot\text{CaM}$  serves a specific role in regulating GRK5, with the C-terminal lobe regulating membrane binding and the N-terminal lobe regulating receptor phosphorylation and kinase domain activation. In cells,



malbrancheamide attenuated GRK5 nuclear translocation and effectively blocked the hypertrophic response, demonstrating the utility of this natural product and its derivatives in probing  $\text{Ca}^{2+}$ -CaM-dependent hypertrophy.

#### 6.4.2 Future Directions

The modulation of the  $\text{Ca}^{2+}$ -CaM-GRK5 interaction with malbrancheamide demonstrated the utility of this complex natural product as a unique probe for  $\text{Ca}^{2+}$ -CaM-dependent enzymes. Malbrancheamide is the first example of a lobe-specific  $\text{Ca}^{2+}$ -CaM inhibitor and can be used to provide intricate details regarding other protein-protein interactions with the modulatory enzyme. We have identified an analog of malbrancheamide, isomalbrancheamide D, which displayed a higher affinity for  $\text{Ca}^{2+}$ -CaM and could serve as a more efficient probe for these studies. Additional malbrancheamide analogs remain to be tested, including molecules with varying halogen substituents. While malbrancheamide has demonstrated a significant biological effect, it lacks traditional drug-like properties and is challenging to make synthetically. For further analysis or drug development, large-scale fungal growths would need to be performed to obtain enough material. On the other hand, malbrancheamide could serve as a lead molecule for the development of more synthetically tractable analogs. Our cocrystal structure also provides valuable insight for potential structure-based drug design, in particular targeting the Cl- $\pi$  interaction in the binding pocket. The main challenge will be to maintain the stabilizing effect of the *trans*  $\text{Ca}^{2+}$ -CaM conformation, which may be aided by computational predictions.