Chemoenzymatic Synthesis of Macrolides and Directed Evolution of the Pikromycin Thioesterase

By

Maria L. Adrover Castellano

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2024

Doctoral Committee:

Professor David H. Sherman, Chair
Assistant Professor Roland D. Kersten
Professor John Montgomery
Professor Janet L. Smith

Maria L. Adrover-Castellano

madrover@umich.edu

ORCID iD: 0000-0002-6110-7640

© Maria L. Adrover-Castellano 2024

Dedication

I dedicate my thesis to my grandparent Luis R. Castellano-Hernández (Abuelo Luis), for being the most wonderful Abuelo a granddaughter could have. I made a promise to take care of our family and to make you proud. I'm happy to know you have been here with me during my good and bad times. Growing up, you were my cheerleader and the best story/joke teller, until undergrad at the same university as you (University of Puerto Rico, Río Piedras or UPR-RP) when you left me to go to heaven becoming my angel. I know you are proud, and my thesis is dedicated to you, who inspired me to continue with graduate school and to get my Ph.D. I miss you every day, and I will love you forever and always.

Acknowledgements

First, I would like to give thanks and express my gratitude to my research advisor David H. Sherman for accepting me in your lab and for being such an amazing mentor for almost 5 years. I'm happy I decided to join your lab, where I completed various projects in the field of natural products and biocatalysis. You have been a role model, inspiration as a scientist, and I'm honored to have completed my Ph.D. in your research lab. I will always be grateful for your mentorship and guidance throughout my Ph.D. Thank you for believing in me, cheering me up when I faced difficulties with personal matters, or the various research projects that I was involved in, and for your scientific feedback. In addition to celebrating with me every grant, fellowship and award I obtained, as well as your excitement whenever I would get data and results. Also, thank you for allowing me to have a work-life-balance outside the university, by letting me visit numerous times my family in Puerto Rico and even visit Europe for vacation. As a mentor, you were always very enthusiastic to learn about my trips and I appreciated our conversations during our 1-on-1 meetings. Again, thank you David for everything!

Likewise, I want to thank my thesis committee: Roland D. Kersten, John Montgomery and Janet L. Smith. Thank you for your scientific guidance and support for the completion of my doctoral degree. I will always be grateful to John and Janet for taking me as rotation student in your labs. You all helped me immensely to figure out my career path, and you served as inspiration and role models then and now.

Second, I want to give a million thanks to my family back in Puerto Rico: Mami (Mom, Wanda), Papi (Dad, Jaime) and my brother Jaimito. You are all the reason why I'm able to say I

finished my Ph.D.! Without your support, this wouldn't have been possible. Thank you for answering all my phone calls, whether it was by voice or video, those really made my days easier, since being far away from you all was one of the hardest things. I want to thank my grandparents, Abuelo Luis, Abuela Gladys, Abuelo Toño, and Abuela Catalina. You have been a great part of who I am, and I'm grateful to all of you. [Now, I will do a switch to Spanish (my native language) to give special thanks to each member of my immediate family circle.] [Mami, gracias por siempre escucharme, estar pendiente de mí, y por tu gran amor. Gracias por emocionarte por mis logros, y ser mi confidente. También por enviarme regalitos, y comidas típicas puertorriqueñas para Michigan con ayuda de Papi. Muchas veces me sentí como si los hubiese tenido aquí cerca. Papi, gracias por creer en mí, y darme siempre los mejores consejos. Tu personalidad jocosa siempre me sacaba sonrisas cuando más lo necesitaba. Gracias por al igual que mami estar pendiente de hacerme feliz. Ambos son excelentes padres y fueron mi inspiración para mis estudios, desde que comencé en la escuela en kindergarten, siendo mis maestros y motivación para continuar en la UPR-RP y luego el doctorado. A Jaimito, los hermanos Adrover siempre unidos. Gracias por ser el mejor hermano mayor que una hermana menor pudiese tener. Por también estar pendiente de mí, enviarme chistes y videos por texto para alegrar mis días en Michigan. Y por siempre hacer mis visitas en Puerto Rico las mejores, al igual que tus gestos de ayuda y apoyo hacia mí. Mami, Papi y Jaimito los amo, gracias. A Abuela Gladys y Abuelo Luis, quienes siempre dieron lo mejor de ustedes para que yo fuese feliz y tuviese la mejor infancia en Jayuya, con los recuerdos más bellos que una nieta pudiese tener. Gracias por llenar mi corazón de amor y por complacerme tanto. Los amo siempre, gracias, esto es por ustedes también. Y a mis abuelos Toño y Catalina, por ser inspiración para luchar y dar lo mejor de mí para cuidar de la familia.]

Third, I want to give thanks to Nick, my significant other. One of the best parts of coming to Michigan was definitely meeting you. Thank you for being there for me every step of the way, and for all of your love and support. You and I are a team. We worked very hard to obtain our Ph.D.'s and now we both get to say we are Doctors! Thank you also for being with me through all of our enjoyable moments and even in adversity, for believing in me, and for helping me grow as a person and as a scientist. You are my favorite person, and I'm proud of our accomplishments and sharing these together. I love you. I also want to say thank you for allowing me to become part of your family in Texas. Rosalinda, Nacho, grandparents and sisters, thank you for your love and support.

Fourth, I want to say special thanks to my uncle Francis, aunt Gladys, cousins Christina and Andy. You guys have been crucial for my career and I'm grateful to all of you. Thank you for opening the doors to your house in Guaynabo, where I stayed while studying for my bachelor's degree in chemistry at UPR-RP. Tío and Tití, thank you for becoming second parents for me, for taking care of me, motivating me, and for allowing me to experience unforgettable moments and trips with you. Since very young taking me to Disney World to taking me to Europe. You guys are awesome. Christina and Andy, you both are like a sister and a brother to me. Thank you both for caring for me, and showing me the most unconditional love. You guys are the sweetest, thank you for always being so patient with me when I was stressed with school, for your jokes, and laughers throughout the years, which made my days better when I needed them the most. Thank you both.

Lastly, I would like to say thanks to my friends. Amalia, thank you for being my best friend since high school, for also cheering for me and caring for me. I have many memories with you in Jayuya and over the years of friendship, which I will never forget, no matter the distance or the time. Thank you, you are the best. Previous Sherman lab people: to Jenn, my research mentor, who also became my dear friend. Thank you for taking me as a rotation student in my first year in the

Sherman lab, for your patience in teaching me about science and research. Thank you for believing in me. Also, for always being there for me when I needed someone the most. I know I can always count on you, and I'm forever grateful to have found you. To Rosa, thank you for also being an awesome mentor and friend. You also made my days in the lab better by supporting me, and being a person that I can rely on. You and I have been partners in crime, and I'm so happy we met. To Robbie, thank you for making me laugh, and being a light in the day when I needed someone to just talk to about random stuff; whether it was what we were up to, or the newest funny music video. Thank you all. Thanks also to previous and current members of the lab who made my lab experience better. I'm grateful to have met wonderful people while in the lab: Alanna, Amy, Andrew, Brian, Carolyn, Casandra, Filipa, Hannah, Jake, Kat, Lauren, Lyanne, Maribel, Mary, Morgan, Natalia, Nikki, Kate, Pam, Sahar, Sam, Sean, Sebastian, Vik, Ying, Yogan and Zach. Everyone's support meant a lot to me during graduate school. Thanks to the many friends I made throughout my undergraduate career at UPR-RP, you guys' rock. I miss you all and cannot wait for our next reunion.

Table of Contents

Dedication ii

Acknowledgements iii

List of Tables xiii

List of Figures xvi

List of Abbreviations xxvii

Abstract xxix

CHAPTER I Introduction and Background 1

I.1 Introduction 1

I.1.1 Natural Products in Medicine 1

I.1.2 Chemical Synthesis of Natural Products and Related Analogs 2

I.1.3 Polyketide Synthases and Nonribosomal Peptide Synthetases 4

I.2 Thioesterases 6

I.2.1 Thioesterase Mechanistic Overview 7

I.2.2 General Cyclization Reactions PKS/NRPS TEs (Applications for Lactones, Lactams, Chemoenzymatic Synthesis and Engineering in these Systems) 9

I.3 Thioesterase Catalyzed Oligomerization 19

I.4 Fungal Thioesterase Domains: C-C bond forming Claisen-Like Cyclases 22

I.5 Summary and Outlook on Thioesterases 24

I.6 Thesis Overview 25

CHAPTER II Chemoenzymatic Synthesis of Macrolactone Rings and Potential Macrolide Antibiotics Utilizing Polyketide Synthases (PKSs) and Post-PKS Tailoring Enzymes 28

II.1 Abstract 28

II.2 Introduction 29

II.2.1 Antibiotic Resistance and Macrolide Antibiotics 29

II.2.2 Macrolide Design and Applications of the Pik Biosynthetic Pathway 31

II.2.3 *In-vitro* Characterization of PikAIII-TE Fusion Protein and PikAIII/AIV System 32

II.2.4 Synthesis of Unnatural Substrates to Interrogate the Flexibility of PikAIII-TE Fusion Protein for the Production of New Macrolactones 35

II.2.5 Synthesis of Unnatural Hexaketides to Investigate the TE as a Catalytic Gatekeeper 37

II.3 Development of New Pik Pentaketide Analogs 38

II.4 Evaluation of New Pentaketides Against PikAIII-TE and PikAIII/AIV 40

II.4.1 PikAIII-TE Enzymatic Reactions 40

II.4.2 PikAIII/AIV Enzymatic Reactions 41

II.5 Identification of a Novel Macrolactone Ring Opening Activity for Pik TE 43

II.6 *In-vivo* Glycosylation and Late-stage Macrolide C-H Oxidation (Post-PKS Tailoring Steps) 48

II.7 Conclusion/Summary 51

II.8 Experimental 51
II.8.1 General Information, Materials and Instrumentation 51
II.8.2 Synthetic Experimentals and Characterizations 54
II.8.3 Protein Biochemistry 54
II.9 NMR Spectra 66
CHAPTER III Directed Evolution of the Pikromycin Thioesterase (Pik TE) 101
III.1 Abstract 101
III.2 Introduction 102
III.3 Chemical Synthesis of Amide Hexaketide and Site Saturation Mutagenesis (Round 1) 105
III.3.1 Isolation of Novel 12-membered Macrolactone/Lactam Ring 106
III.3.2 Pik TE Directed Evolution Strategy 107
III.4 Error Prone PCR (epPCR) Methodology (Rounds 2-3) and Generation of a TE Enzyme Mutant Library 108
III.5 Total Turnover Numbers (TTNs), Thermal Shift Assay and Initial Rates for Substrate Consumption and Macrolactone Formation 110
III.6 Pik TE Variants Cysteine 148 to Serine Reversions 112
III.7 Discussion 114
III.8 Conclusion/Summary 118
III.9 Experimental 118
III.9.1 General Information, Materials and Instrumentation 118

III.9.2 Amide Hexaketide Synthetic Procedure 119

III.9.3 Protein Biochemistry 125

III.9.4 DNA Cloning and Mutagenesis 130

III.9.5 High-throughput Screening Protein Expression and Enzymatic Reactions 138

III.9.6 Enzyme Variants Identification, Expression, Purifications and Enzymatic Reactions 140

III.9.7 Total Turnover Numbers (TTNs), Thermal Shift Assay and Initial Rates for Substrate Consumption and Macrolactone Formation 148

III.9.8 X-ray Crystallography for Hybrid 12-membered Macrolactone/Lactam Ring Molecule 163

III.10 NMR Spectra 174

CHAPTER IV Evaluation of PKS Modules PikAIII-TE and PikAIII/AIV for the Production of 12- and 14-membered Macrolactone Rings 185

IV.1 Abstract 185

IV.2 Introduction 186

IV.3 Analysis of Amide and Ester Pentaketides in the Modules PikAIII-TE WT and PikAIII-TE S148C187

IV.4 Where are these SNAc Byproducts Being Off-Loaded From? The ACP or the TE? 190

IV.5 PikAIII-TE WT and PikAIII-TE S148C Enzymatic Reactions using MeMal-CoA as the Extender Unit 192

IV.6 SNAc Transthioesterification Reactions to Test Substrate Activating Group Reactivity Differences Against Pik TE 194

IV.7 Mass Spectrometry Analysis of SNAc in the Pik TE S148C Active Site 196

IV.8 Ligations of Pik TE Mutants from Directed Evolution into PKS Modules 203

IV.8.1 Enzymatic Reactions of Amide Pentaketide Against PikAIII-TE S148C, A217T, G222V (R1) Mutant and PikAIII-TE S148C, A217T, G222V, M271V, Y25C, L126V (R2) 203

IV.8.2 Enzymatic Reactions of Amide/Ester Pentaketide Against PikAIII/AIV-TE S148C, A217T, G222V (R1) Mutant and PikAIII/AIV-TE S148C, A217T, G222V, M271V, Y25C, L126V (R2) 206

IV.9 Conclusion/Summary 209

IV.10 Experimental 211

IV.10.1 General Information, Materials and Instrumentation 211

IV.10.2 Synthetic Experimentals and Characterizations 211

IV.10.3 Protein Biochemistry 211

IV.10.4 DNA Cloning and Mutagenesis 212

CHAPTER V Summary, Discussion and Future Directions 215

V.1 Chemoenzymatic Synthesis of Macrolactone Rings and Potential Macrolide Antibiotics Utilizing PKSs and Post-PKS Tailoring Enzymes 215

V.1.1 Summary and Discussion of Work Completed 215

V.1.2 Future Directions 218

V.2 Directed Evolution of the Pik TE 219

V.2.1 Summary and Discussion of Work Completed 219

V.2.2 Future Directions 222

V.3 Evaluation of PKS Module PikAIII-TE and PikAIII/AIV for the Production of 12- and 14-membered Macrolactone Rings 224

V.3.1 Summary and Discussion of Work Completed 224

V.3.2 Future Directions 226

References 228

List of Tables

Table II.1 Tests with different reaction conditions and thiols for investigation of spontaneous amide macrolactone ring opening. 59

Table II.2 14-desmethyl-narbonolide (**2.28**) major isomer NMR Data C_6D_6 66

Table II.3 14-desmethyl-narbonolide (**2.28**) minor isomer NMR Data C_6D_6 67

Table II.4 12-desmethyl-narbonolide (**2.29**) major isomer NMR Data C_6D_6 71

Table II.5 12-desmethyl-narbonolide (**2.29**) minor isomer NMR Data C_6D_6 72

Table II.6 12,14-didesmethyl-narbonolide (**2.30**) major isomer NMR Data C_6D_6 76

Table II.7 12,14-didesmethyl-narbonolide (**2.30**) minor isomer NMR Data C_6D_6 77

Table II.8 13-desethyl-narbonolide (**2.31**) major isomer NMR Data C_6D_6 81

Table II.9 Ester-narbonolide (**2.27**) 14-membered isomer NMR Data $CDCl_3$ 85

Table II.10 Ester-narbonolide (**2.33**) 6-membered isomer NMR Data $CDCl_3$ 86

Table II.11 14-desmethyl-narbomycin (**2.34**) NMR Data $CDCl_3$ 40 ºC 93

Table II.12 12-desmethyl-narbomycin (**2.35**) NMR Data $CDCl_3$ 40 ºC 97

Table III.1 ^{1}H NMR and ^{13}C NMR (600 MHz NMR, $CDCl_3$) 129

Table III.2 Site saturation mutagenesis NNK degenerate primers for the generation of Pik TE S148C enzyme variants. 132

Table III.3 epPCR primers for amplification between restriction sites NdeI and XhoI in the pET-28(a)+ Pik TE S148C vector. 135

Table III.4 Site directed mutagenesis primers for the generation of Pik TE S148C enzyme variants containing combinations of mutations from rounds of evolution. 137

Table III.5 Melting temperatures (Tm B) in quadruplicates for each protein and melting temperatures derivative (Tm D). 152

Table III.6 Melting temperatures (Tm B) mean values and corresponding standard (Std.) error values. 153

Table III.7 Crystallographic information and data collected. 164

Table III.8 Fractional Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Parameters ($Å^2 \times 10^3$) for **3.2**. U_{eq} is defined as 1/3 of the trace of the orthogonalized U_{ij} 167

Table III.9 Anisotropic Displacement Parameters ($\times 10^4$) for **3.2**. The anisotropic displacement factor exponent takes the form: $-2p^2[h^2a^{*2} \times U_{11} + ... + 2hka^* \times b^* \times U_{12}]$ 168

Table III.10 Bond Lengths in Å for **3.2**. 168

Table III.11 Bond Angles in ° for **3.2**. 169

Table III.12 Torsion Angles in ° for **3.2**. 170

Table III.13 Hydrogen Fractional Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Parameters ($Å^2 \times 10^3$) for **3.2**. U_{eq} is defined as 1/3 of the trace of the orthogonalized U_{ij}. 171

Table IV.1 Amide/ester pentaketide enzymatic reaction results for PikAIII-TE modules with engineered Pik TEs for the formation of 12-membered ring product. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data). 206

Table IV.2 Amide pentaketide enzymatic reaction results for PikAIII/AIV-TE modules with engineered Pik TEs for the formation of 6- and 14-membered ring products. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data). 207

Table IV.3 Ester pentaketide enzymatic reaction results for PikAIII/AIV-TE modules with engineered Pik TEs for the formation of 6- and 14-membered ring products. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data)..... 208

Table IV.4 Amide hexaketide enzymatic reaction results for PikAIV-TE modules with engineered Pik TEs for the formation of 6- and 14-membered ring products. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data)..... 209

List of Figures

Figure I.1 Examples of pharmaceutical NPs: aspirin, morphine, penicillin, taxol, lovastatin, doxorubicin and erythromycin. 2

Figure I.2 Chemoenzymatic methods combine both enzymes (biocatalysts) and synthetic strategies for the generation of complex molecules. 4

Figure I.3 PKS and NRPS domains including the cyclization and release domains present in both systems. 5

Figure I.4 Mechanism of the lactonization or hydrolysis of a CP-tethered peptide or polyketide intermediate. In this chapter, highlighted in purple is the bond formed via the described enzyme. 7

Figure I.5 Overall structure of PKS and NRPS thioesterases with the core α-helices depicted in teal, the lid α-helices depicted in yellow, β strands represented in purple. a. Topology diagram of TEs with the catalytic residues highlighted. b. Representation of Pik TE (PDB: 2H7X) highlighting the substrate channel typically seen in PKS TEs and a triketide substrate shown in pink. c. Representation of Vlm TE (PDB: 6ECE) depicting the buried, bowl shaped active site with the lid forming the top cap. 9

Figure I.6 Chemical structures of the lactones pikromycin, obafluorin, salinamide A and lysobactin. 11

Figure I.7 Chemical structures of the lactams sulfazecin, polymyxin B and fluvirucin B_1..... 13

Figure I.8 a. Chemical structures of cryptophycin analogs, teixobactin, zearalenone and radicicol. b. The *in-vitro* biocatalytic platform using the terminal two monomodules (PikAIII and PikAIV) of the Pik pathway. 16

Figure I.9 a. Engineering of the Pik TE for synthesis of novel macrolactones. b. Chemical structures of cilengitide and the tautomycetin polyketide natural products. 18

Figure I.10 a. Chemical structures of cereulide, valinomycin and elaiophylin. b. Proposed mechanism for TE oligomerization in conglobatin biosynthesis. 21

Figure I.11 a. The wA TE/CLC mechanism for the production of YWA1. b. PksA TE/CLC domain catalyzes the Claisen cyclization (C-C bond formation) of noranthrone. 23

Figure I.12 a. PhnA TE/CLC C1-C10 Claisen cyclization producing the naphthopyrone prephenalenone. b. Shunt products from the PhnA TE deletion (TE^0). 24

Figure II.1 Estimated deaths attributable to antibiotic resistance every year by 2050 (Figure from O'Neill, J., Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations, *The Review on Antimicrobial Resistance*, **2014**, 1-18, https://amr-review.org/ accessed on December 2023). 30

Figure II.2 Examples of clinically relevant macrolide antibiotics. 31

Figure II.3 Pikromycin biosynthetic pathway. 32

Figure II.4 Mechanism of PikAIII and PikAIV enzyme system for the generation of macrolactone rings. 34

Figure II.5 Scalable *in-vitro* chemoenzymatic synthesis of macrolactones using either PikAIII-TE or PikAIII/AIV, and whole cell bioconversion using engineered *S. venezuelae* to macrolides. Macrolactone components shown in green are introduced by the PKS modules as a result of chain elongation, processing and cyclization through one (10-DML), or two cycles (NBL). 35

Figure II.6 Evaluation of PikAIII-TE fusion protein against a panel of truncated (**2.9-2.12**) and stereoisomer (**2.13** and **2.14**) pentaketides. Highlighted in blue are the differences of these substrates in comparison to the Pik pentaketide (**2.8**)..... 36

Figure II.7 Native and epimeric hexaketides enzymatic reactions with Pik TE WT and Pik TE S148C. Epimeric pentaketide enzymatic reaction with reincorporation of Pik TE S148C generating PikAIII-TE S148C. 38

Figure II.8 New Pik pentaketide analogs. Highlighted in blue are the differences of these substrates in comparison to the Pik pentaketide (**2.8**)..... 39

Figure II.9 New pentaketides tested against PikAIII/AIV for the generation of 14-membered macrolactone rings. Highlighted in blue are the differences of these substrates in comparison to the Pik pentaketide (**2.8**). 42

Figure II.10 Proposed TE mechanism for the generation of both 6- and 14-membered rings present in the amide and ester enzymatic reactions with PikAIII/AIV. 43

Figure II.11 Total ion chromatogram (TIC) for amide pentaketide enzymatic reaction against PikAIII/AIV for 1 hour (blue), 2 hour (orange) and overnight (24 hours, green) time course to investigate formation of the 6-membered macrolactone ring. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run. 44

Figure II.12 Total ion chromatogram (TIC) for amide 14-membered macrolactone ring reaction against PikAIV module and standalone Pik TE WT and Pik TE S148C to test macrocycle opening and re-cyclization mediated by the TE. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run. 46

Figure II.13 Total ion chromatogram (TIC) for ester 14-membered macrolactone ring reaction against PikAIV module and standalone Pik TE WT and Pik TE S148C to test macrocycle opening

and re-cyclization mediated by the TE. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run. 47

Figure II.14 a. Scheme depicting the generation of narbonolide, narbomycin and pikromycin analogs from our strategy and inspired by the Pik biosynthesis. b. Glycosylated macrolides and corresponding oxidated macrolides showcasing the reactivity differences between PikC and JuvD P450s. 50

Figure II.15 Total ion chromatogram (TIC) for the truncated pentaketides reactions against PikAIII/AIV resulting in the formation of the desired 14-membered macrolactone rings. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run. 58

Figure II.16 Total ion chromatogram (TIC) resulting from tests with different reaction components/byproducts and thiols for the investigation of possibility of spontaneous amide macrolactone ring opening. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run..... 59

Figure II.17 Total ion chromatogram (TIC) resulting from reactions with amide and ester 14-membered ring macrolactones against Pik TE S148A. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. 60

Figure II.18 Total ion chromatogram (TIC) resulting from reactions with amide hexaketide reaction against PikAIV to test ring opening hypothesis and implications of the formation of the heptaketide acid intermediate. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. For synthetic procedure on amide hexaketide substrate, see Chapter 3 experimentals. 61

Figure III.1 Chemical synthesis of amide hexaketide (**3.1**). 106

Figure III.2 a. Amide hexaketide **3.1** reaction with diverse Pik TE variants to produce **3.2**. b. Pik TE WT with covalent pentaketide affinity label (PDB: 2HFJ) in yellow. Twenty-one residues identified within 6 Å of the substrate in cyan. Active site catalytic triad in purple (Ser148, Asp176, His268)..... 108

Figure III.3 a. Pik TE variants characterized from directed evolution. b. Pik TE WT crystal structure (PDB: 2HFJ) with mutations in cyan depicting Pik TE R2 using PyMOL mutagenesis tool. Pentaketide mimic affinity label in yellow and the TE catalytic triad in purple. c. Fold improvement for Pik TE mutants in terms of isolated yields (%) compared to the original Pik TE S148C (13%)..... 110

Figure III.4 a. Analysis for total turnover numbers (TTNs), melting temperatures (Tm, ^{o}C), initial rates (μM/min) substrate consumption and macrolactone formation for each Pik TE variant. b. Reaction time courses: substrate and products HPLC peaks shown as area under the curve (AUC) as a function of time (min). See experimental section for more details. 112

Figure III.5 a. LC-MS % yields and b. traces for Pik TE WT, Pik TE R1 C148S$_{reversion}$ and Pik TE R2 C148S$_{reversion}$ highlighting a gain of function for the WT enzyme. 113

Figure III.6 Amide hexaketide reaction with Pik TE WT or Pik TE S148C. Reaction conditions: water, sodium phosphate buffer (400 mM NaH_2PO_4•H_2O and Na_2HPO_4, 200 mM NaCl, pH 7.2), hexaketide (1 mM), Pik TE (10 μM), overnight, stationary, rt..... 128

Figure III.7 Example of extracted ion chromatogram (EIC) resulting from enzymatic reactions of the amide hexaketide with Pik TE WT and Pik TE S148C. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 10 minutes (min) total run..... 128

Figure III.8 Example of colonies screened after Gibson assembly method optimization: epPCR total mutations (silent + non silent) when varying concentrations of $MnCl_2$: colonies 1-10: 200 μM, colonies 11-20: 250 μM, colonies 21-28: 300 μM. 136

Figure III.9 Example of purified protein SDS-PAGE for Pik TE variants: Pik S148C, Pik TE R1 and Pik TE R2. NuPAGE® 4-12% Bis-Tris Mini Gel of purified Pik TE variants in MES SDS running buffer with BenchMark™ Pre-stained Protein Ladder. Lanes from left to right: Protein Ladder, Pellet, Supernatant, Flowthrough, Eluted Purified Pik TE variants (lanes repeat in the same order for S148C, R1 and R2). 141

Figure III.10 Amide hexaketide reaction with Pik TE variants from directed evolution campaign. Reaction conditions: water, sodium phosphate buffer (400 mM $NaH_2PO_4•H_2O$ and Na_2HPO_4, 200 mM NaCl, pH 7.2), hexaketide (0.25 μM), Pik TE variant (10 μM), overnight, stationary, rt.. 142

Figure III.11 Example of extracted ion chromatogram (EIC) resulting from enzymatic reactions amide hexaketide with Pik TE mutants. TOF-MS: 25 – 70% acetonitrile with 0.1% formic acid, 5 minutes (min) total run..... 143

Figure III.12 Pik TE directed evolution variants developed in rounds 1-3 and their corresponding LC-MS % yields for hybrid 12-membered macrolactone/lactam ring formation and depiction of the entire library screen..... 145

Figure III.13 Beneficial mutations found in rounds 1-2 that impacted selectivity for macrocyclization and their corresponding distances (Å) with the bound pentaketide affinity label in yellow in the Pik TE WT crystal structure (PDB: 2HFJ), catalytic triad in red. 145

Figure III.14 Standard curve for amide hexaketide (**3.1**) substrate..... 148

Figure III.15 Standard curve for hybrid 12-membered macrolactone/lactam ring (**3.2**) product. 149

Figure III.16 Total Turnover Numbers (TTNs) for Pik TE S148C, Pik TE R1 (S148C, A217T, G222V), Pik TE R2 (S148C, A217T, G222V, M271V, Y25C, L126V) and Pik TE R3 (S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W). 150

Figure III.17 a) Example of Total Turnover Numbers (TTNs) resulting from enzymatic reactions amide hexaketide with Pik TE mutants. Analytical HPLC: 25 – 70% acetonitrile, 17 minutes (min) total run. HPLC retention time drift is observed depending on different data collection days. b) Overlaid HPLC TTNs traces. Note: The peak observed to the left of the 12-ring corresponds to trace amounts (negligible) degradation of amide hexaketide (retention time: 7.60 min) due to exchange of thiophenol group with MeOH upon quenching the reactions. The peak observed to the left of hydrolysis product corresponds to spontaneous degradation of the hydrolysis molecule (retention time: 10.35 min) under the reaction conditions..... 151

Figure III.18 Melting curve plot and derivative plot for fluorescence emission readout for thermal shift assay Pik TE WT and variants (S148C, R1, R2, R3). Each curve represents the Tm average value for each protein. Also, melting curve plot on the complete Pik TE library and Tm values which followed same protocol as described before. 155

Figure III.19 Pik TE S148C reaction time course for substrate and macrolactone concentration (μM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error)..... 156

Figure III.20 Pik TE R1 reaction time course for substrate and macrolactone concentration (μM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error). 157

Figure III.21 Pik TE R2 reaction time course for substrate and macrolactone concentration (μM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error). 157

Figure III.22 Pik TE R3 reaction time course for substrate and macrolactone concentration (μM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error). 158

Figure III.23 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE S148C, with error bars representing standard deviation (error bars not visible, minimal error). 159

Figure III.24 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE R1 (S148C, A217T, G222V) with error bars representing standard deviation. 159

Figure III.25 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE R2 (S148C, A217T, G222V, M271V, Y25C, L126V) with error bars representing standard deviation (error bars not visible, minimal error). 160

Figure III.26 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE R3 (S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W) with error bars representing standard deviation (error bars not visible, minimal error). 160

Figure III.27 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE S148C, with error bars representing standard deviation. 161

Figure III.28 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE R1 (S148C, A217T, G222V) with error bars representing standard deviation. 161

Figure III.29 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE R2 (S148C, A217T, G222V, M271V, Y25C, L126V) with error bars representing standard deviation (error bars not visible, minimal error). 162

Figure III.30 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE R3 (S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W) with error bars representing standard deviation. 162

Figure III.31 12-membered macrolactone/lactam ring initial rates (μM/min) in terms of fold improvement for the engineered Pik TE variants. 163

Figure III.32 a) Example of the colorless block-shaped crystals obtained for **3.2**. b) X-ray crystal structure of **3.2**. 166

Figure III.33 Ellipsoid plot: X-ray crystal structure of hybrid 12-membered macrolactone/lactam ring (**3.2**). 173

Figure IV.1 Extracted ion chromatogram (EIC) resulting from enzymatic reaction of amide and ester pentaketides with PikAIII-TE and PikAIII-TE S148C. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. See also compound color scheme/compound legend accompanying the EIC. 190

Figure IV.2 Extracted ion chromatogram (EIC) resulting from enzymatic reaction of amide and ester pentaketides with PikAIII-TE S148A (TE inactive). TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. See also compound color scheme/compound legend accompanying the EIC in Figure **IV.1**. 191

Figure IV.3 Depiction of the SNAc byproducts offloaded from the ACP via the SNAc nucleophilic attack into thioester-ACP bond. 192

Figure IV.4 Total ion chromatogram (TIC) resulting from enzymatic reaction of amide pentaketide with PikAIII-TE S148C using MeMal-CoA as the extender unit. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. See also peak for electrospray ionization source (ESI) scan corresponding to the desired hybrid 12-membered macrolactone/lactam ring at the 4.755 min retention time. 193

Figure IV.5 Extracted ion chromatogram (EIC) resulting from thiophenol amide hexaketide transthioesterification reaction with SNAc, with subsequent TE mediated cyclization. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. 195

Figure IV.6 Electrospray ionization source (ESI) scan corresponding to Pik TE S148C and Pik TE S148S with SNAc added. 197

Figure IV.7 Electrospray ionization source (ESI) scan corresponding to Pik TE WT and Pik TE WT with SNAc added. 198

Figure IV.8 Electrospray ionization source (ESI) scan corresponding to Pik TE WT with SNAc added and Pik TE WT SNAc/DTT added. 199

Figure IV.9 Electrospray ionization source (ESI) scan corresponding to Pik TE S148C with SNAc added and Pik TE S148C SNAc/DTT added. 200

Figure IV.10 SNAc titration results: electrospray ionization source (ESI) scan corresponding to Pik TE S148C with SNAc at different concentrations ranging from 0.1 – 10.0 mM SNAc. 201

Figure IV.11 Extracted ion chromatogram (EIC) resulting from thiophenol methoxy protected Pik native hexaketide against Pik TE S148C. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. 203

Figure IV.12 Extracted ion chromatogram (EIC) resulting from enzymatic reaction of amide pentaketide with PikAIII-TEs. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13

minutes (min) total run. See also peak for electrospray ionization source (ESI) scan corresponding to the desired hybrid 12-membered macrolactone/lactam ring at the 4.756 min retention time. 205

Figure IV.13 Example on a control experiment for PCR Pik TE amplification and restriction digest of Pik TE amplicon and PKS module vector. 214

List of Abbreviations

α: alpha
Ac: acyl
Ac-CoA: acyl coenzyme A
β: beta
BCG: biosynthetic gene cluster
Bu: butyl
Bz: benzoyl
^{o}C: Celsius
CoA: coenzyme A
CYP: cytochrome P450
DCM: dichloromethane
DIBALH: diisobutylaluminium hydride
DIEA: *N,N*-diisopropylethylamine
DMAP: 4-dimethylaminopyridine
DMF: dimethylformamide
DMSO: dimethyl sulfoxide
dr: diastereomeric ratio
ee: enantiomeric excess
epPCR: error prone polymerase chain reaction
Et: ethyl
EtOAc: ethyl acetate
Et_2O: diethyl ether
Equiv.: equivalent
γ: gamma
gem: geminal
h: hour
HRMS: high resolution mass spectroscopy
hv: ultraviolet light
i: iso
i-Bu: isobutyl
i-Pr: isopropyl
IBX: 2-iodoxybenzoic acid
μ: micro
M: molarity
Me: methyl

MeMal-SNAc: methylmalonyl *N*-acetyl cysteamine
mg: milligram
n-Bu: butyl
n-Hex: hexyl
n-Pent: pentyl
n-Pr: propyl
$NADP^+$: nicotinamide adenine dinucleotide phosphate
NADPH: reduced form of $NADP^+$
NMR: nuclear magnetic resonance
NRPSs: nonribosomal peptide synthetases
NPs: natural products
PG: protecting group
Ph: phenyl
PhH: benzene
PhMe: toluene
PhSH: thiophenol
PKSs: polyketide synthases
Pr: propyl
psi: pounds per square inch
QTOF: quadruple time-of-flight mass spectrometry
R: variable molecular group unless specified
r.t.: room temperature
SAR: structure-activity relationship
sat.: saturated
SNAc: *N*-acetyl cysteamine
t: tertiary
t-Bu: *tert*-butyl
TBAF: tetrabutylammonium fluoride
TBS: *tert*-butyldimethyl silyl
TEA: triethylamine
TES: triethylsilyl
Tf: triflate
THF: tetrahydrofuran
TMS: trimethylsilyl
TOF-MS: Time-of-flight mass spectrometry

Abstract

Natural products (NPs) are a source of inspiration for the discovery and development of new pharmaceuticals, and they currently represent over one third of all FDA-approved small molecule drugs. These complex secondary metabolites and derivatives possess a wide range of biological activities and have been applied to treat human diseases, including cancer, parasitic infections, autoimmune disorders and bacterial pathogens. The isolation of these molecules from diverse organisms and chemical synthesis has effectively provided the desired products. However, these methods are often encountered with challenges and limitations such as low yields, unsuccessful synthesis and generation of waste and toxic materials. Chemoenzymatic approaches represent an alternative for chemical synthesis combined with biocatalysis or enzyme catalyzed reactions to obtain efficient production of these compounds. The synthesis of a variety of intermediates in conjunction with late stage biosynthetic machinery enables access to a repertoire of molecules, which could be otherwise challenging to produce utilizing synthetic methods alone. Type I modular polyketide synthases (PKSs) are a family of enzymes capable of synthesizing a diverse class of NPs. The biochemical features of these systems include the extension and processing of polyketide chains in a stepwise, stereospecific manner, organized by a series of modules divided into distinct catalytic domains. Taking advantage of these enzymes as biocatalysts would be a powerful addition to the synthetic chemistry toolbox that enables a more sustainable and environmentally responsible approach to molecular diversification. Despite the promise of these enzymes, deliberate engineering is required to expand their ability to process natural and unnatural polyketide substrates.

The work presented in this dissertation describes the chemoenzymatic synthesis of macrolactone rings using synthetic substrates and PKS modules from the antibiotic Pikromycin (Pik) biosynthetic pathway. In addition, the terminal cyclization domain, the Pik thioesterase (TE) was investigated employing protein engineering strategies via directed evolution. The Pik enzyme system PikAIII/AIV was explored for the generation of 14-membered macrolactones and macrolides utilizing unnatural pentaketide substrates. Interestingly, an amide and ester containing pentaketides showed efficient production of the 14-membered macrolactones using the PikAIII/AIV enzyme system, although the 12-membered macrolactones were not generated with the PikAIII-TE fusion protein, raising questions about the deciding factors for the formation of the macrocycles. Previous work postulated that the primary hurdle to applying PKS modules to create diverse macrolactones hinges on the selectivity of the TE domain. Thus, we decided to synthesize an unnatural amide hexaketide to study the standalone Pik TE against this unique substrate. We demonstrated the ability to form a hybrid 12-membered macrolactone/lactam ring in conjunction with a Pik TE S148C mutant. The low yields resulting from these Pik TE S148C reactions inspired us to embark on a directed evolution strategy to identify Pik TE variants with enhanced selectivity for the desired 12-membered rings. Our study provides critical insights into the amino acid residues that impart improved activity and yield with these substrates in the pursuit of combinatorial strategies to produce synthetically challenging molecules for drug discovery applications.

CHAPTER I

Introduction and Background

The majority of this introduction and background has been published as "Biosynthetic Cyclization Catalysts for the Assembly of Peptide and Polyketide Natural Products" and reproduced with permission from all authors Maria L. Adrover-Castellano, Jennifer J. Schmidt and David H. Sherman. *ChemCatChem*. **2021**, *13*, 2095-2116. https://doi.org/10.1002/cctc.202001886

I.1 Introduction

I.1.1 Natural Products in Medicine

NPs serve as the source of numerous medicines utilized today.[1, 2] Organisms across all kingdoms of life have the ability to produce these compounds as part of their secondary metabolism (molecules generally not essential for their growth, development and reproduction). NPs encompass a wide array of compounds with diverse biological properties, including antimicrobial, immunosuppressive, anticancer, and anti-inflammatory activities, among others, developed to treat human diseases.[3, 4] One of the earliest recordings of NPs usage dates back to the advanced medicinal system of ancient Mesopotamia, spanning from 2900 to 2600 BC. Examples of these include oils from *Cupressus sempervirens* (Cypress) and *Commiphora species* (myrrh), currently used to treat coughs, colds, and inflammation. Additional reports have emerged from Egypt (2900 BC), China (1100 BC and 100 BC), and Greece (300 BC) with documented medicinal herbs in various forms such as ointments, pills, gargles to infusions.[2] From the ancient historical evidence of NPs to the continuous scientific advancements over the years, many widely used drugs today have been discovered. Notable examples include aspirin (**1.1**) (anti-inflammatory drug and

blood thinner) derived from the bark of willow tree *Salix alba L.*, morphine (**1.2**) (analgesic) from the plant opium poppy *Papaver somniferum L.*, penicillin (**1.3**) (antibiotic) from the fungus *Penicillium notatum*, taxol (**1.4**) (anticancer) from the bark of pacific yew *Taxus brevifolia*, lovastatin (**1.5**) (cholesterol lowering) from the fungus *Aspergillus terreus*, doxorubicin (**1.6**) (anticancer) from the bacterium *Streptomyces peucetius var. caesius*, and erythromycin (**1.7**) (macrolide antibiotic) from the actinomycete bacteria *Saccharopolyspora erythraea* (**Figure I.1**).[2, 4] Currently, NPs account for over one third of all FDA-approved small molecule drugs, showcasing the significance of these pharmaceuticals in human health.[1]

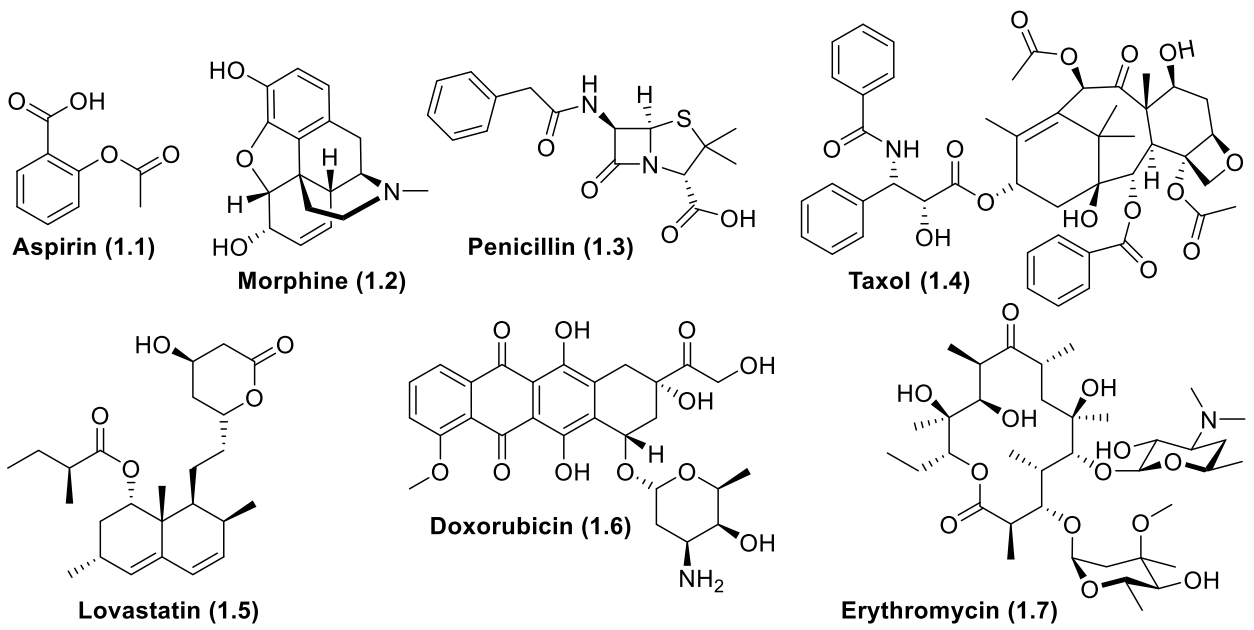


Figure I.1 Examples of pharmaceutical NPs: aspirin, morphine, penicillin, taxol, lovastatin, doxorubicin and erythromycin.

I.1.2 Chemical Synthesis of Natural Products and Related Analogs

NPs are characterized by their scaffold diversity and highly complex molecular structures. Although NPs are naturally derived, they are often produced in limited quantities, and challenges

associated with their isolation from the producing organisms make chemical synthesis an attractive technique to employ. The generation of NPs and related analogs allows for the study of new bioactivities as well as explorations to structure-activity relationships (SAR) to optimize compound leads.[5] Historically, total synthesis and synthetic organic chemistry have numerous successful examples of developing complex molecules, providing access to biologically active isolates and promising new candidates and analogs.[6] Despite the promise of these workable methodologies, total synthesis faces drawbacks and challenges related to stereochemical and structural diversity, often resulting in laborious, low yielding product synthesis that is insufficient for further diversification or medicinal assessment purposes.[4, 7] For instance, the first total synthesis of the antibiotic erythromycin[8-10] in 1981 required 50 synthetic steps and yielded less than 1%.

Conversely, semi-synthesis has been employed by performing structural and stereochemical modifications to existing NP isolates. While this approach reduces the synthetic step counts and improves the overall yields, modifications to the starting material scaffolds are often limited by reactive functional groups and difficulties in selectively altering diverse positions within these intricate structures. Although workable methodologies exist, they are often limited to extensively protected precursors, harsh reactions conditions (temperatures, reagents, solvents), and the formation of side products.[11-13] Chemoenzymatic methods provide an alternative for the chemical synthesis of intermediates from simpler building blocks, combined with enzyme catalyzed reactions to efficiently produce novel molecules and pharmaceuticals. By synthesizing a variety intermediates in conjunction with enzymes, access to a repertoire of molecules is facilitated, which could otherwise be challenging to produce using synthetic methods alone.[14-18] The use of enzymes as biocatalysts enables the generation of structurally distinct compounds while reducing waste,

utilizing water as the reaction solvent and avoiding rare earth metal catalyst or toxic conditions (**Figure I.2**).[17] The macrolide antibiotic erythromycin is naturally produced by enzymes called Polyketide Synthases (PKSs), similar to Nonribosomal Peptide Synthetases (NRPSs) found in its biosynthetic pathway, highlighting the great potential of studying these enzymes further to understand their activities and the mechanisms involved in generating such a complex structure.[19]

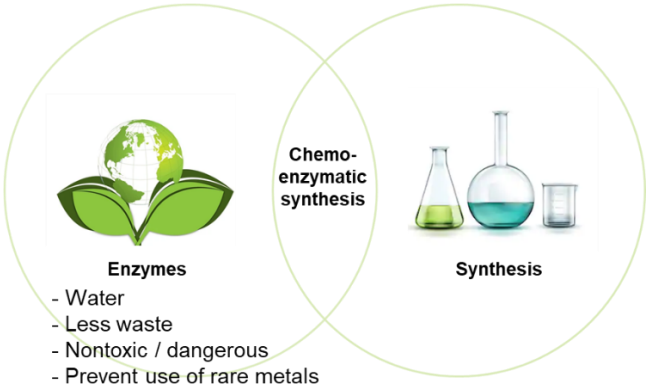


Figure I.2 Chemoenzymatic methods combine both enzymes (biocatalysts) and synthetic strategies for the generation of complex molecules.

I.1.3 Polyketide Synthases and Nonribosomal Peptide Synthetases

PKSs, NRPSs and hybrids (PKS/NRPS) thereof are responsible for the synthesis of numerous NPs in bacterial and fungal species.[20] Several of these compounds are clinically approved therapeutics to treat human diseases, including cancer (e.g. calicheamicin and bleomycin) autoimmune disorders (e.g. rapamycin), parasitic infections (e.g. ivermectin) and bacterial pathogens (e.g. erythromycin, tiacumicin B and vancomycin).[21] The multifunctional machinery of these megasynthases makes them particularly unique. Individual modules containing diverse domains incorporate specific substrates, as they work together to assemble complex nonribosomal peptides or polyketides in a linear manner to generate a final product. PKS and NRPS products possess widely varied structures with linear, cyclic and branched elements.[22]

The PKS and NRPS megasynthases contain modules that mediate substrate initiation, chain extension, processing and termination. A module contains a series of domains that catalyze diverse reactions on a peptide or polyketide chain. The NRPS core module contains an adenylation (A) domain that selects and activates amino acid monomers, a condensation (C) domain which catalyzes the formation of peptide bonds, and a thiolation domain (T) or peptidyl carrier protein (PCP) with a phosphopantetheine (Ppant) group for transferring substrates into the catalytic site. NRPS modules might also contain accessory domains including an epimerization (E) domain for the conversion of L-amino acids to D-amino acids. The PKS core module contains an acyltransferase (AT) domain, which selects and transfers an extender unit, an acyl carrier protein (ACP) also with a Ppant group (similarly to the PCP in NRPS) for extender unit loading, and a ketosynthase (KS) domain for decarboxylative Claisen condensation.[23] In the case of PKS, the elective processing domains include a ketoreductase (KR), dehydratase (DH) and enoylreductase (ER) (**Figure I.3**).[20]

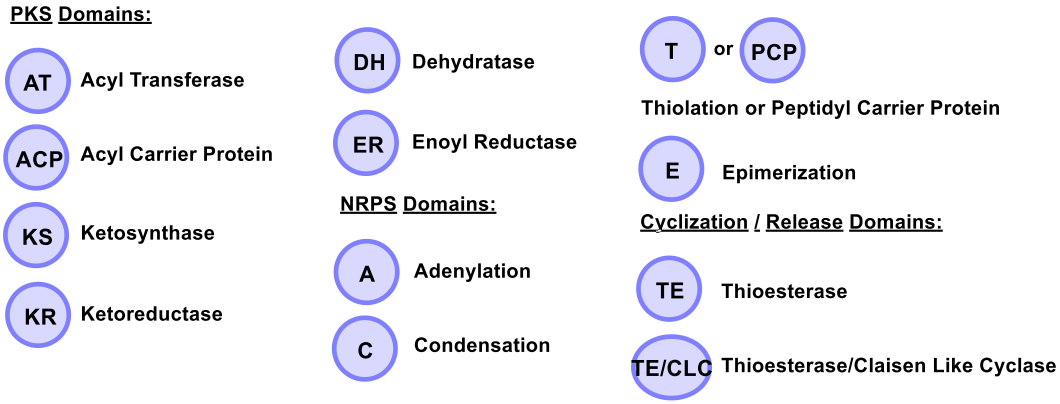


Figure I.3 PKS and NRPS domains including the cyclization and release domains present in both systems.

During the biosynthesis of a fully elongated and mature linear peptide or polyketide, a terminal domain, usually a thioesterase (TE) catalyzes the release of the peptide and polyketide as a linear

or cyclized product (**Figure I.3**).[23] In this portion of the thesis, the focus will be directed towards examples of biosynthetic cyclization catalysts found in PKS and NRPS systems. The general cyclization mechanism of the TE domain will be discussed including oligomerization and the fungal C-C bond forming Claisen-like cyclases (CLCs). Although various detailed examples on these topics have appeared in the literature previously,[20, 24-26] the main focus will be on the current literature relating to fascinating PKS and NRPS cyclization catalysts responsible for the assembly of particularly unique natural products and related compounds.

I.2 Thioesterases

As previously described, PKSs and NRPSs are organized into a series of modules that contain distinct catalytic domains. Utilizing peptidyl or acyl chain initiation, elongation and termination mechanisms, these enzymatic assembly lines are capable of generating a mature product. At the end of a biosynthetic pathway, a terminal TE domain usually located at the C-terminus of a PKS or the NRPS module is responsible for catalyzing the cyclization step during termination.[20, 27] Due to the importance and wide range of thioesterases found in nature, this topic has been reviewed with detailed discussions on PKS TEs, NRPS TEs, comparisons of their sequence as well as insights into their structure.[24, 28] In this portion, an overview of the cyclization mechanisms employed by TEs will be provided along with highlights of recent examples of the work in this area of natural product biosynthesis. Studies on the TE domain are particularly compelling due to their ability to catalyze diverse reactions, including cyclization via oligomerization, and through carbon-carbon bond formation, among others discussed below. Although an important feature of PKS and NRPS systems, the TEII class of thioesterases involved in editing of polyketide intermediates during assembly will not be discussed.[29]

I.2.1 Thioesterase Mechanistic Overview

TEs are prevalent in both PKS and NRPS pathways. These enzymes are broadly characterized as members of the α/β hydrolase superfamily, which also includes serine proteases. Macrocyclizing TEs, termed type I TEs are generally found at the end of the C-terminal module of PKS and NRPS pathways. In the active site, the TE features a conserved serine (Ser), histidine (His) and aspartate (Asp) catalytic triad. After the peptide or polyketide is fully extended and elongated, it is typically transferred from the upstream module PCP or ACP to the hydroxyl group of the Ser. The Ser residue acts as a catalytic nucleophile activated by the His and Asp dyad, which generates an acyl-O-TE covalent enzyme intermediate.[24, 30, 31] The Ppant group covalently attached to the ACP or PCP active site Ser residue mediates the approach and transfer of the thioester substrate to the TE. Then, an acyl hydroxyl or amino nucleophile adds into the carbonyl of the acyl-O-TE intermediate. An intermolecular nucleophile, generally water, leads to a linear hydrolyzed product, while an intramolecular nucleophile (amino or hydroxyl group) generates a lactam or lactone during cyclization, respectively (**Figure I.4**).[20, 24]

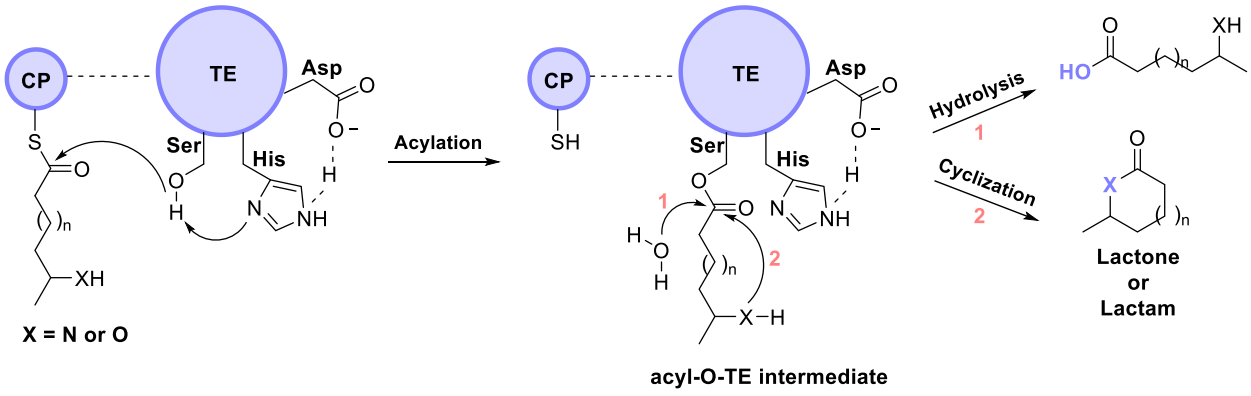


Figure I.4 Mechanism of the lactonization or hydrolysis of a CP-tethered peptide or polyketide intermediate. In this chapter, highlighted in purple is the bond formed via the described enzyme.

Multiple macrocyclizing TE X-ray crystal structures have been solved including for the erythromycin (**1.7**) and pikromycin (**1.8**) PKS TEs as well as tyrocidine[32] and fengycin[33] NRPS TEs. These structures provide the basis for information about the similarity of the protein fold, nature of the substrate channel and catalytic mechanism for ring closure.[34, 35] Despite nearly identical fold architecture (**Figure I.5a**), PKS TEs are dimers and contain dimerization helices at the N-terminus as well as a hydrophobic substrate channel through the protein (**Figure I.5b**). In contrast NRPS TEs are monomeric and do not contain these extra helices but sometimes a primary β-strand at the N-terminus. NRPS TEs also have a more bowl-shaped active site that is capped by a formal lid (**Figure I.5c**). These differences are born out in the current understanding of molecular mechanisms employed by TEs to effect macrocyclization. In 2006, the Fecik, Sherman and Smith groups captured a substrate mimic in the active site of the pikromycin (Pik) TE leading to a new hypothesis for macrocyclization.[36] The structure revealed a previously unrecognized water barrier in the enzyme that appears to induce a curled substrate conformation to direct macrolactone ring formation.[36] Similar studies on NRPS domains cited both key interactions in the active site (especially apparent in the fengycin TE) as well as inherent preorganization of the linear peptides as the driving force behind cyclization.[30, 33] Indeed, early studies on the tyrocidine TE demonstrated that it was capable of cyclizing substrates with widely differing amino acids compared to the native structure, lending biochemical evidence to preorganization of the linear intermediate as a driving force in the NRPS TE cyclization mechanism.[37] This new information provided insights for the design of enzymes capable of catalyzing regioselective macrocyclization of natural or synthetic substrates.[35]

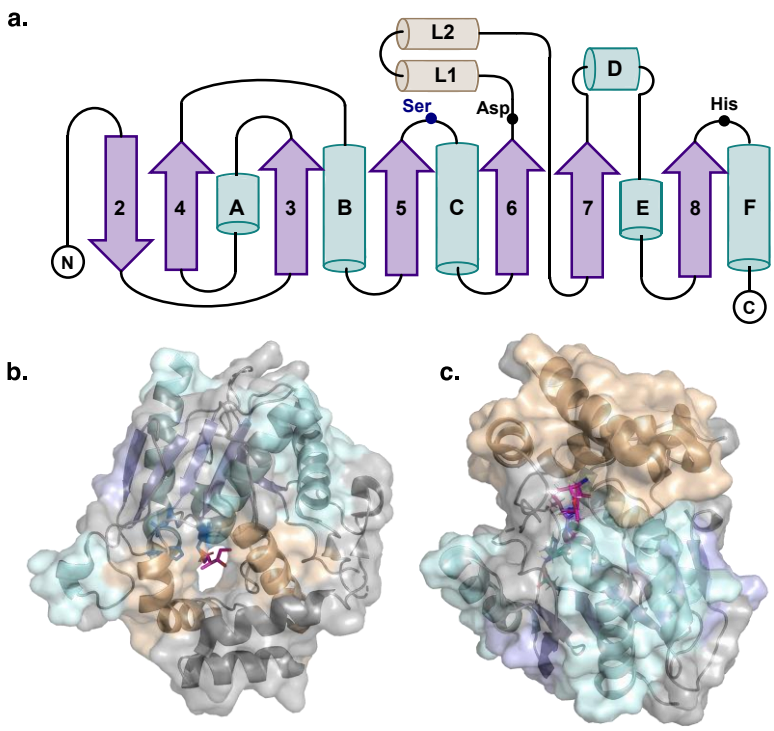


Figure I.5 Overall structure of PKS and NRPS thioesterases with the core α-helices depicted in teal, the lid α-helices depicted in yellow, β strands represented in purple. a. Topology diagram of TEs with the catalytic residues highlighted. b. Representation of Pik TE (PDB: 2H7X) highlighting the substrate channel typically seen in PKS TEs and a triketide substrate shown in pink. c. Representation of Vlm TE (PDB: 6ECE) depicting the buried, bowl shaped active site with the lid forming the top cap.

I.2.2 General Cyclization Reactions PKS/NRPS TEs (Applications for Lactones, Lactams, Chemoenzymatic Synthesis and Engineering in these Systems)

The canonical cyclization of PKSs and NRPSs refers to the TE mediated mechanism that includes the Ser-His-Asp catalytic triad and the acyl-O-TE intermediate, which facilitates the ring closing reaction.[38, 39] A number of examples of cyclic structures can be found including lactones and lactams that are formed via canonical and non-canonical mechanisms. The different types of mechanisms will be discussed below. Moreover, chemoenzymatic applications and engineering of

these systems for the synthesis of various natural product structural analogs will also be highlighted in the context of recent work involving TE domain mediated cyclization.

I.2.2.1 Lactones

Recently, the Wencewicz group studied the NRPS *cis*-monocyclic β-lactone antibiotic obafluorin (**1.9**) (**Figure I.6**).[40] Obafluorin (Obi) was isolated in 1984 from *Pseudomonas fluorescens* ATCC 39502.[41, 42] The biosynthetic gene cluster (BGC) was subsequently identified, and the ObiF TE domain shown to play a direct role in β-lactone formation. The authors demonstrated that the TE catalyzes cyclization of an active site-tethered β-hydroxythioester during product release to generate the strained ring system. Primary sequence analysis revealed that there is a cysteine active site residue (Cys1141) instead of the expected serine for the ObiF TE. This work has the potential for future NRPS engineering to produce targeted peptide β-lactones.

In 2016, the Moore group reported a study of salinamide A (**1.10**), which is a cyclic depsipeptide possessing antibacterial activity (**Figure I.6**).[43] Salinamides A and B are the first depsipeptides from this class isolated from marine actinomycete *Streptomyces* sp. CNB-091.[44] They found the *sln9* gene encodes a tetradomain NRPS module with a C-terminal TE. Biochemical characterization of the Sln9 TE showed that it catalyzes an intermolecular transesterification of a serine residue of desmethylsalinamide E with acylated glycyl thioesters to give desmethylsalinamide C in salinamide biosynthesis. Lysobactin (**1.11**), is another example of a cyclic depsipeptide with high antibacterial activity against human pathogens (**Figure I.6**). This molecule was originally isolated from *Lysobacter* sp. ATCC 53042 in 1988.[45, 46] In 2011, the Marahiel group identified and characterized the gene cluster responsible for the lysobactin biosynthesis.[47] The *lyb* gene cluster codes for two NRPSs, LybA and LybB. The latter was found to contain an unusual tandem TE architecture. The authors characterized *in-vitro* the two individual

terminal TEs (TE1, TE2) using a thiophenol activated linear lysobactin analog. Their results demonstrated that only the penultimate thioesterase domain mediates cyclization and simultaneous release of lysobactin. TE1 exclusively catalyzes the formation of the macrocyclic structure, whereas TE2 mediates hydrolytic cleavage of the synthetic substrate.

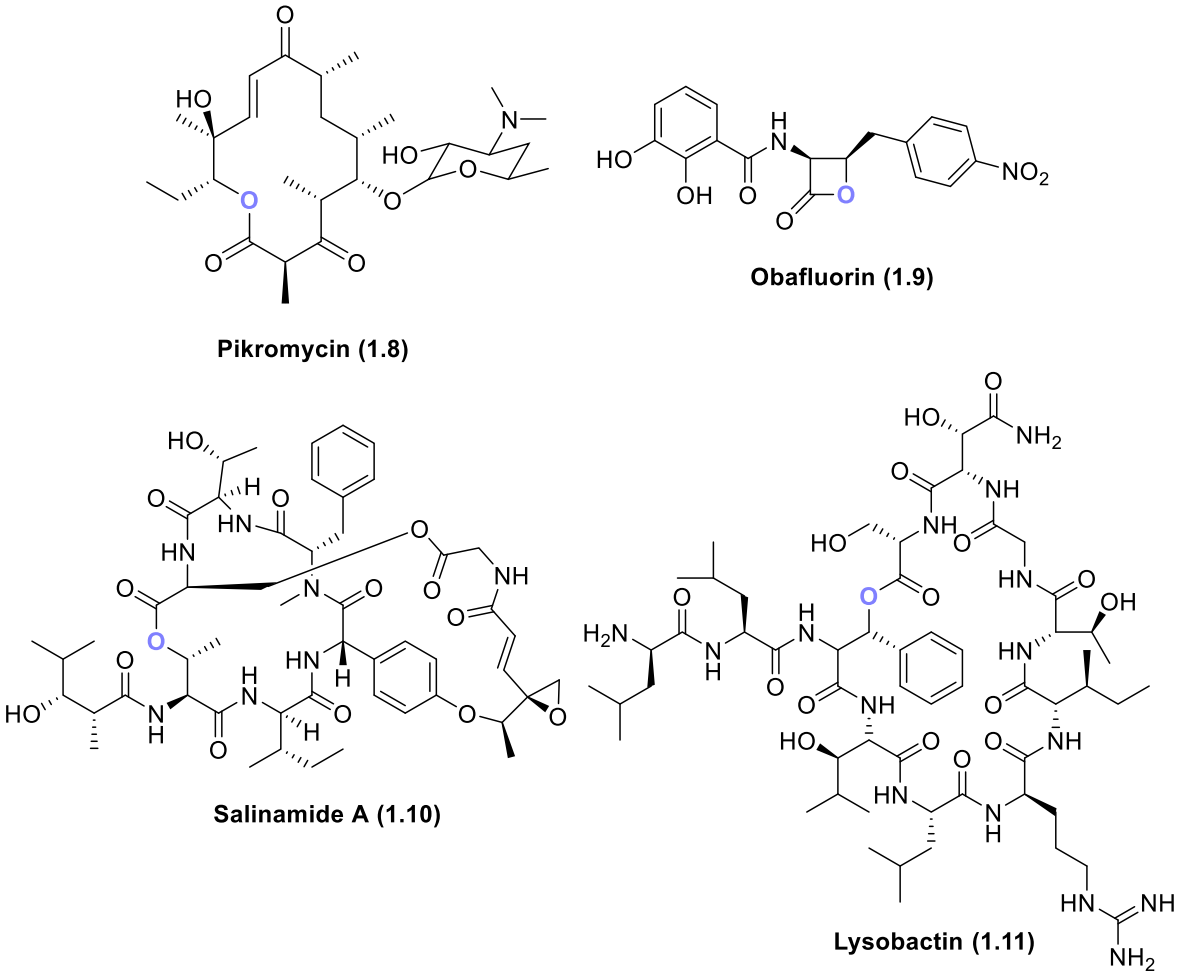


Figure I.6 Chemical structures of the lactones pikromycin, obafluorin, salinamide A and lysobactin.

I.2.2.2 Lactams

In 2018, the Townsend group revealed the biosynthetic routes for monobactam synthesis in sulfazecin (**1.12**).[48] This β-lactam antibiotic is produced by NRPS enzymes and was isolated from

Pseudomonas acidophila G-6302 in 1981 (**Figure I.7**).[49] In the gene cluster, they identified an unusual TE domain possessing a cysteine active site, and described possible routes for β-lactam formation. The proposed mechanism involves an unknown NRPS TE reaction type in which an assembled tripeptide is *N*-sulfonated *in-trans* before β-lactam ring synthesis in the TE domain, which is distinct compared to other systems**.**

In 2017, the Li and Velkov groups characterized the polymyxin B (**1.13**) TE to understand its mode of assembly (**Figure I.7**).[38] Polymyxins remain one of the few antibiotics available for treating antibiotic resistant bacteria. In their study they found that the TE contains a catalytic cysteine in the active site similarly to the TE involved in the synthesis of sulfazecin noted above. They also demonstrated that the polymyxin B pathway TE is able to cyclize a linear *N*-acetyl cysteamine (SNAc) polymyxin B analog. Additionally, homology model examinations of the TE substrate-binding cleft was found to be negatively charged, suggesting a potential mechanism for cyclization of the substrate. The authors also noted that charged interactions could play a role in the binding of the positively charged linear substrate to the negatively charged TE binding site facilitating cyclization.

The Schnarr group in 2013 identified and characterized the fluvirucin B_1 (**1.14**) polyketide synthase (**Figure I.7**).[50] This 14-membered macrolactam produced by *Actinomadura vulgaris* has shown antifungal and antiviral activity.[51-53] They found that the fluvirucin B_1 PKS consists of five extender modules, flanked by an initial N-terminal loading ACP that incorporates the initial β-alanine and a C-terminal TE, which catalyzes ring closure of the linear polyketide substrate. This macrolactam forming TE was also found to include a cysteine in the active site instead of a serine.

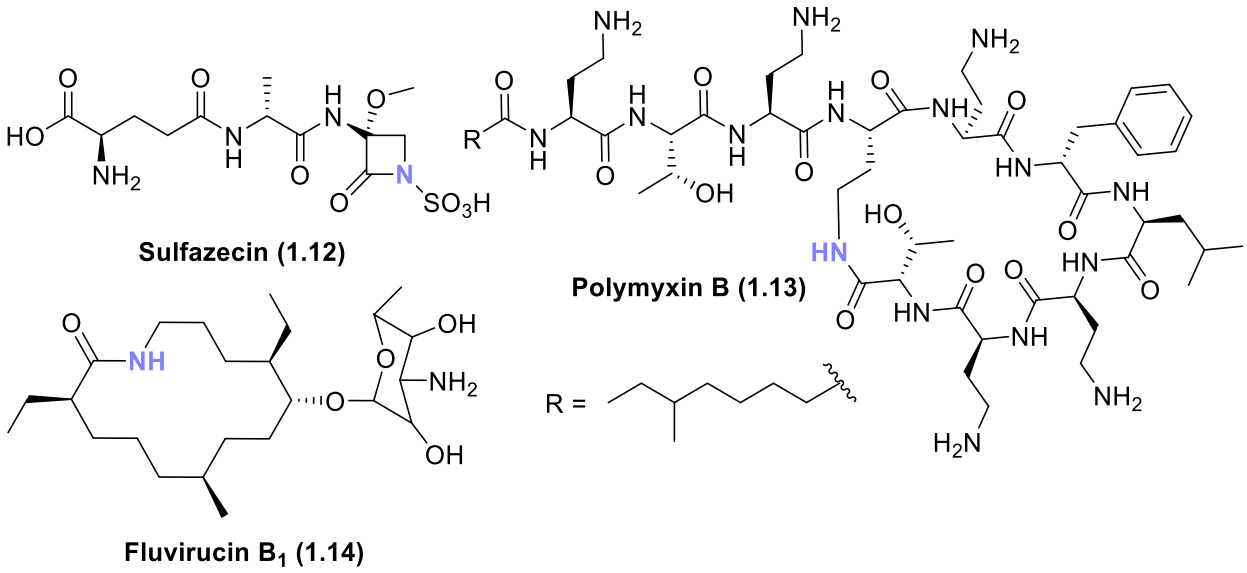


Figure I.7 Chemical structures of the lactams sulfazecin, polymyxin B and fluvirucin B_1.

I.2.2.3 *Chemoenzymatic Syntheses*

Recently, the Sherman group described the chemoenzymatic synthesis of anti-cancer cryptophycin analogs (**1.15**).[54] These macrocycles are a class of 16-membered ring depsipeptide natural products generated by hybrid PKS/NRPS system (**Figure I.8a**). Cryptophycins were first isolated in *Nostoc* sp. ATCC 53789, and showed antifungal activity.[55] These compounds were rediscovered in *Nostoc* sp. GSV 224, as a potent antiproliferative microtubule binding agent.[56] The penultimate step in cryptophycin biosynthesis involves a macrocyclization catalyzed by the cryptophycin thioesterase (Crp TE).[57] In their work, a series of novel cryptophycin chain elongation intermediates were synthesized and the Crp TE demonstrated its utility as a biocatalyst through the formation of a library of cryptophycin analogs. Additionally, biological activity studies identified a low picomolar potency analog containing a styrene functionality that overcomes the need for a β-epoxide group. This work shed light into the production of complex bioactive molecules for the potential treatment of malignant diseases and may represent a promising candidate as an antibody-drug conjugate payload.[58]

In 2018, the Zhang group studied the teixobactin (**1.16**) TE domains.[59] Teixobactin is a depsipeptide isolated from a β-proteobacterium *Eleftheria terrae*, with antimicrobial activity against many notorious Gram-positive pathogens (**Figure I.8a**).[60] The BGC encodes two large NRPSs, Txo1 and Txo2. In the Txo2 termination module two TE domains (TE1, TE2) were identified, similarly to the lysobactin biosynthetic system.[47] The Zhang group developed a TE based chemoenzymatic approach for the development of teixobactin analogs. Their work shows that the two TE domains operate together to generate a cyclized product and are functionally exchangeable. In contrast to lysobactin biosynthesis in which only the penultimate thioesterase domain (TE1) mediates cyclization while TE2 mediates hydrolytic cleavage, in teixobactin biosynthesis, the TE domains (TE1, TE2) appear to form a combined active site cavity that is essential for cyclization. Their work further demonstrates the impressive mechanistic diversity found in NRPS systems.

The Sherman group in 2013 reported an *in-vitro*, scalable biocatalytic platform utilizing the final two modules (PikAIII and PikAIV) of the Pik biosynthetic pathway for the generation of 12- and 14-membered ring macrolactones (**Figure I.8b**).[61] A thiophenol activated Pik pentaketide (**1.19**) substrate was synthesized, and loaded onto PikAIII-TE or PikAIII/AIV enzyme systems to afford 10-deoxymethynolide (**1.20**) and narbonolide (**1.21**), respectively. Furthermore, these macrolactones were converted to their macrolide counterparts *in-vivo* by direct appendage of D-desosamine and final P450-mediated C-H oxidation(s). In a follow-up study, a series of unnatural pentaketides were synthesized, which revealed the substrate tolerance of PikAIII-TE *in-vitro* for production of novel 12-membered ring macrolactones.[62] However, interrogation of pentaketides with epimerized chiral centers in the C9 hydroxyl group nucleophile revealed the inability of Pik TE to catalyze macrocycle formation. Instead, exclusive hydrolysis of the epimerized hexaketide

substrates was observed, suggesting that the TE domain is a catalytic bottleneck in the processing of these chain elongation intermediates.

The Boddy group in 2014 reported work on the TE domains from the zearalenone (**1.17**) and radicicol (**1.18**) biosynthetic pathways (**Figure I.8a**).[63] These natural products form part of the resorcylic acid lactone (RALs) class of macrocyclic fungal polyketides that contain a resorcylate (2,4-dihydroxybenzoate) within the 14-membered ring macrolactone. Interestingly, these compounds have opposite configuration in the lactone alcohol group, where zearalenone possesses an L (*S*) configuration, whereas radicicol contains a D (*R*) configuration. In their work, they show that the TE domains responsible for substrate cyclization are highly stereotolerant to D and L configurations contained in synthetic substrate analogs. Both zearalenone and radicicol TE have the ability to macrocyclize substrates with a different stereochemistry and low competing hydrolysis, in contrast to bacterial PKS TEs, which are highly stereoselective. They further demonstrated the ability of these RAL TEs in chemoenzymatic synthesis of 12- to 18-membered ring macrolactones and a 14-membered macrolactam.[64]

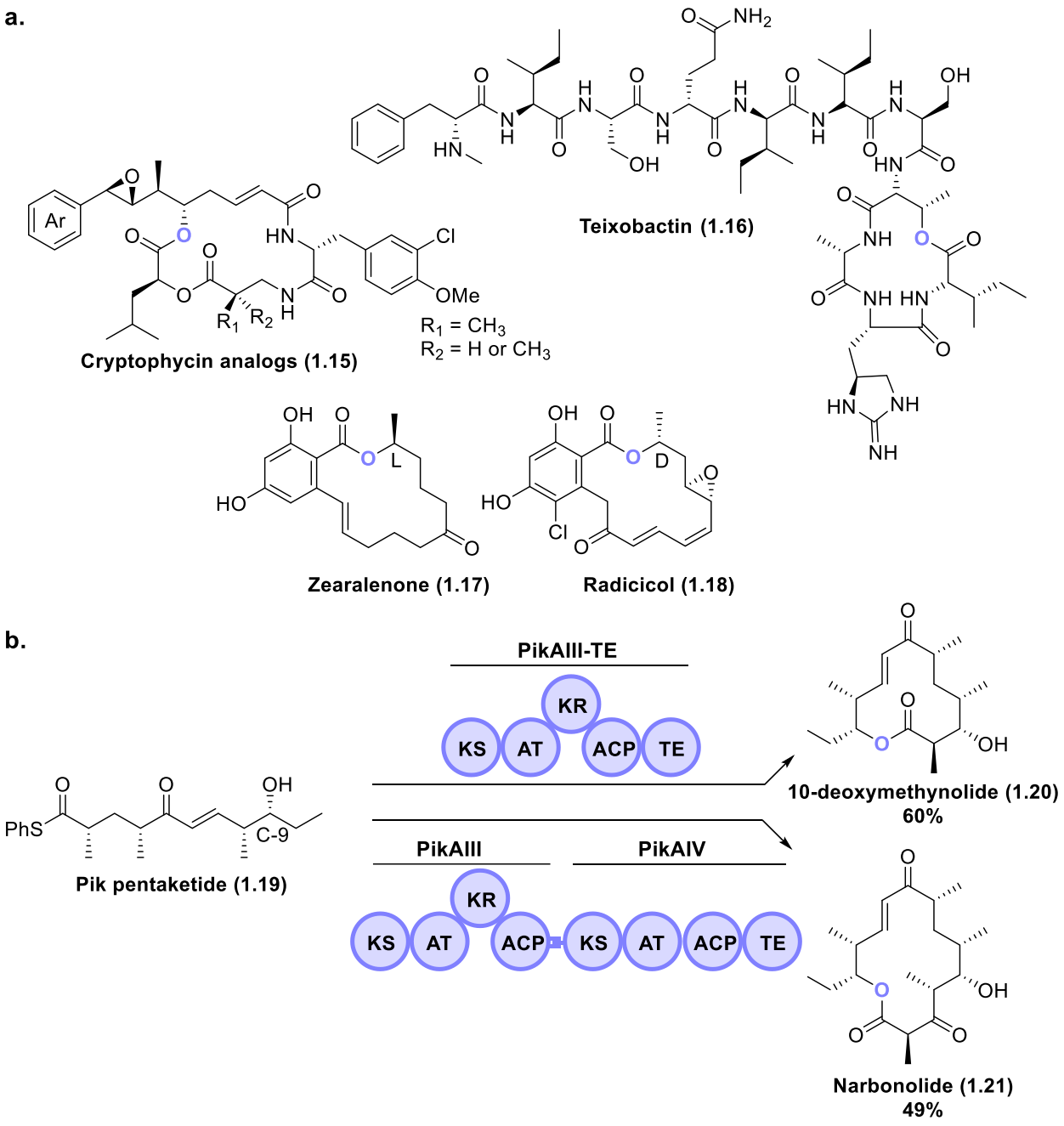


Figure I.8 a. Chemical structures of cryptophycin analogs, teixobactin, zearalenone and radicicol. b. The *in-vitro* biocatalytic platform using the terminal two monomodules (PikAIII and PikAIV) of the Pik pathway.

I.2.2.4 Engineering in these Systems

In 2017, the Sherman group reported the engineering of the Pik TE, where they mutated the serine 148 in the active site to a cysteine (**Figure I.9a**).[65] The application of the engineered Pik TE S148C into a hybrid protein PikAIII-TE S148C enabled processing of the C9 epimerized pentaketide (**1.22**) for the production of an epimeric 12-membered ring macrolactone (**1.23**). Comparison of quantum mechanics calculations for wild type Pik TE and the S148C mutant revealed that it significantly lowers the activation barriers to macrolactonization, resulting in a more efficient catalyst. In another recent study, the group created a series of hybrid PKS modules containing exchanged TE domains from heterologous pathways.[66] The modules selected were the terminal two modules of the pikromycin (Pik) and erythromycin (DEBS) pathways, and the penultimate module of the juvenimicin (Juv) system. These hybrid PKS modules were interrogated with native and non-native polyketide substrates. Their results showed that substrate TE pairing was a major driving factor in macrolactone production. These works highlight that the TE is often a critical catalytic gatekeeper for the processing of unnatural polyketide chain elongation intermediates. It also showed the substrate flexibility of PKS TE domains and their application in future engineering strategies for the production of polyketide natural product analogs.

The Zhu group in 2019, reported a chemoenzymatic method to synthesize the cyclic peptide cilengitide (**1.24**) based on the activity of the TE domain from *Microcystis aeruginosa* microcystin synthetase C (**Figure I.9b**).[67] Moreover, they generated a single S85C mutation in McyC TE, following the strategy used by the Sherman group for the Pik TE.[65] For their chemoenzymatic method using initially the wild type McyC TE domain, they synthesized a pentapeptide-BMT substrate. Enzymatic reactions showed that the McyC TE was able to recognize the linear pentapeptide chain and cyclized it into cilengitide. This was further improved when the McyC TE

S85C mutant was employed. Kinetic analysis revealed that the McyC TE S85C mutation resulted into a superior cyclization catalyst for the pentapeptide-BMT substrate for the production of cilengitide with a 12-fold increase in the K_{cat}/K_m for macrolactonization.

The Sherman and Kim groups in 2016 reported the polyketide tautomycetin (**1.25**) TE exchange with the Pik TE in *Streptomyces* sp. CK4412 (**Figure I.9b**).[68] In this work, the PKS cluster in *Streptomyces* sp. CK4412 that produces tautomycetin (TMC) was engineered by replacing the TMC TE domain with the Pik TE, in order to test the ability of the constructed TMC-Pik TE chimera to catalyze macrocyclization of the linear substrate. Their findings showed that indeed, the TMC-Pik TE chimera produced not only the expected TMC linear product, but also a cyclized form of TMC. This work shows the potential to engineer PKS strains for *in-vivo* production of new natural products by exchanging a hydrolyzing TE for a macrocyclizing TE.

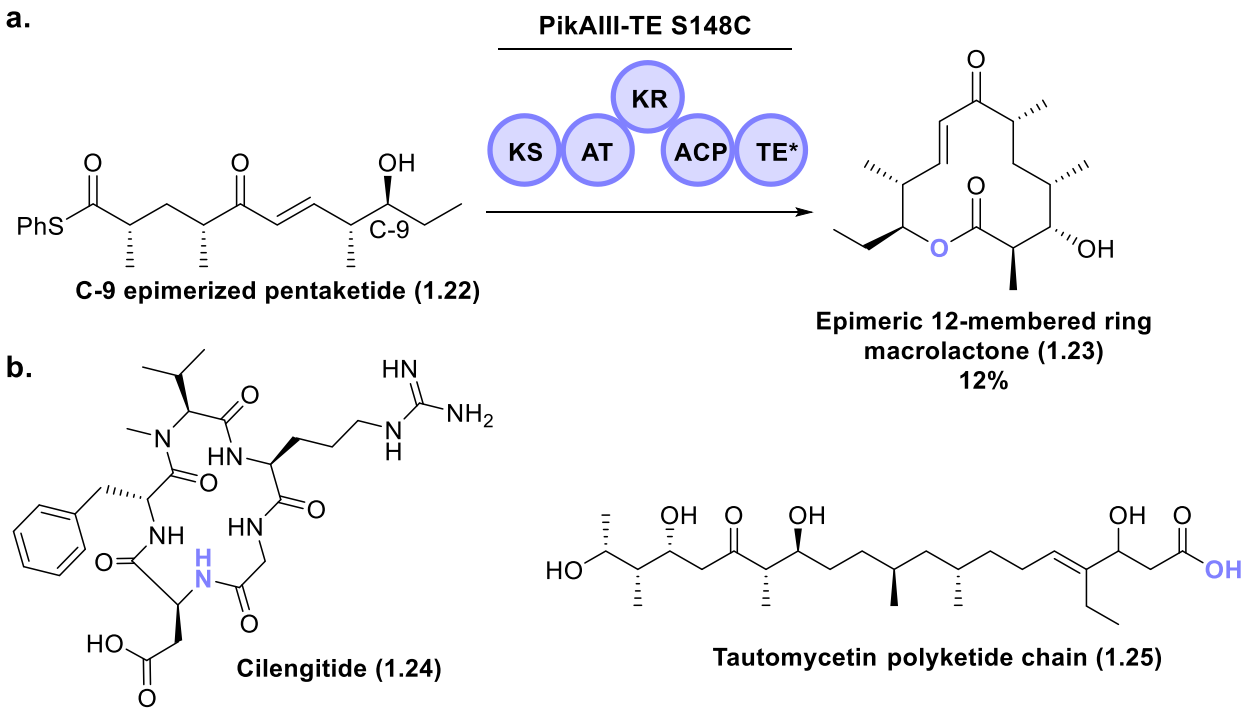


Figure I.9 a. Engineering of the Pik TE for synthesis of novel macrolactones. b. Chemical structures of cilengitide and the tautomycetin polyketide natural products.

I.3 Thioesterase Catalyzed Oligomerization

In addition to canonical head to tail macrocyclization, TE domains are also known to catalyze oligomerization of various natural products during biosynthesis. Cereulide (**1.26**), valinomycin (**1.27**), elaiophylin (**1.28**) and conglobatin (**1.29**) are representative examples of both PKS and NRPS derived compounds all formulated by an oligomerization mechanism (**Figure I.10**). Initially, one monomer is covalently attached to the active site serine of the thioesterase and the second to the carrier protein directly upstream (ACP or PCP). From here two modes of dimerization have been proposed including a "forward transfer" model (originally hypothesized for the enterobactin biosynthesis[69]) that involves nucleophilic attachment of the TE bound monomer on the carrier protein monomer, functionally building the oligomerized substrate on the TE. The second model, termed the "backward transfer" shifts the TE bound monomer back to a nucleophilic functional group on the carrier protein bound monomer, and building the oligomer on the carrier protein. This dimerized (or oligomerized when this cycle is repeated more than once) is then transferred back to the TE active site serine for final cyclization. Evidence for this backward mechanism initially comes from *in-vitro* work with the gramicidin S TE in which linear dimerized thioester intermediates were detected.[70] Recent structural work by the Schmeing and Chin groups on the valinomycin TE have further defined the mechanisms employed in this system.[71] Utilizing site-specific incorporation of a non-canonical amino acid 2,3-diaminopropionic acid (DAP) into recombinant TE, they demonstrated efficient capture of both the first and last acyl-enzyme intermediates. Their crystallographic data demonstrated that for this enzyme, the "backwards transfer" method is employed, further corroborating the gramicidin S TE data and indicating this is likely a global mechanism employed by this subset of TEs. Deeper investigations on the molecular mechanism of the cyclization function revealed a massive lid rearrangement associated

with the dodecadepsipeptidyl bound Vlm TE. The terminal domain reorients the substrate from its position during oligomerization placing it into a pocket that entropically controls cyclization. Overall, this work provided an innovative approach to characterize and capture diverse acyl-enzyme complexes of transiently acylated proteins.

The Boddy group has also investigated the TE from the highly homologous cereulide pathway, which is also produced by two NRPS, CesA and CesB.[72] Cereulide is a depsidodecapeptide isolated in 1994 from *Bacillus cereus*.[73, 74] Despite the similarities between the structures and BGCs of cereulide and the natural product valinomycin, it was previously suggested that cereulide was synthesized through a different mechanism.[75] However, the Boddy group demonstrated that cereulide is in fact produced in a similar manner to valinomycin where the Ces TE catalyzes the trimerization and macrocyclization of a tetradepsipeptide substrate. Interestingly, they also show the Ces TE can utilize related substrates like the valinomycin tetradepsipeptide and a cereulide-valinomycin hybrid tetradepsipeptide with the native substrate to generate natural product analogs. In 2015, the Leadlay group reported the C_2-symmetrical macrodiolide elaiophylin produced by PKS enzymes.[76] They investigated the elaiophylin (Ela) TE domain mechanism of the diolide formation *in-vitro* by using synthetic SNAc thioesters of tetraketide and pentaketide analogs of the natural octaketide monomers. It was determined that the Ela TE can catalyze homodimerization of the synthetic pentaketide to a novel 16-membered decaketide diolide. Iterative use of the Ela TE active site during ligation and cyclization is suggested since a linear dimeric thioester intermediate was identified in the ring formation. The authors also proposed that Ela TE must catalyze a total of two acylation and two deacylation reactions to form the diolide. Thus, in a similar manner to that determined for the NRPS dimerizing TEs, the Ela TE first ligates two monomers and then re-

loads the linear dimer in the active site for cyclization. However, further studies are required to firmly establish the Ela TE dimerization mechanism (**Figure I.10a**).

In the same year, the Leadlay group also studied conglobatin biosynthesis by a PKS/NRPS hybrid (**Figure I.10b**).[77] This compound is an unusual C_2-symmetrical macrodiolide with antitumor activity isolated from *Streptomyces conglobatus* ATCC 31005.[78] The group utilized a single step *in-vitro* cloning method to obtain the entire gene cluster that contains the five genes *congA-E* resulting in conglobatin production in a heterologous host. A model substrate was also utilized to mimic the conglobatin monomer, which revealed that the Cong TE acts iteratively. The Cong TE ligates two monomers and then re-binds the dimeric product and cyclizes it similarly to the Ela TE described above. Additional work showed that Cong TE is capable of combining two different monomers into hybrid linear dimers and trimers forming additional complex polyketides.

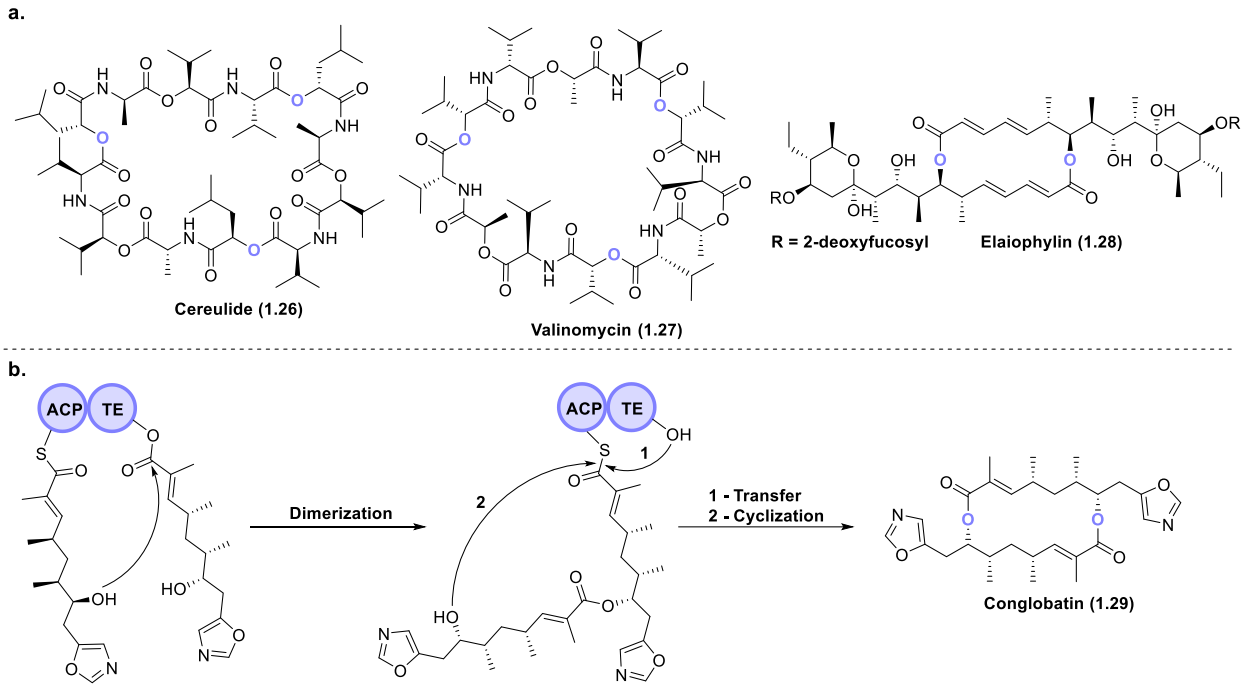


Figure I.10 a. Chemical structures of cereulide, valinomycin and elaiophylin. b. Proposed mechanism for TE oligomerization in conglobatin biosynthesis.

I.4 Fungal Thioesterase Domains: C-C bond forming Claisen-Like Cyclases

Fungal nonreducing PKS (NR-PKS) are megasynthases that include the SAT (starter unit: ACP transacylase), KS (β-ketosynthase), MAT (malonyl CoA: ACP transacylase), PT (product template), ACP (acyl carrier protein) and TE/CLC (thioesterase/Claisen like cyclase) domains.[79, 80] For the majority of fungal NR-PKS, the TE catalyzes the last step, which releases the product via a Claisen cyclization of a polyketide intermediate, leading to a C-C bond formation.[81] This process is accompanied by the cleavage of the thioester bond between the ACP and the polyketide to generate fused ring structures.[82, 83] Similarly to the TE domains from bacterial PKSs, the TE/CLC domain also contains a Ser, His and Asp catalytic triad.[34] In 2001, the Ebizuka group identified for the first time a TE with a Claisen cyclase function in the *wA* PKS gene cluster (**Figure I.11a**) from *Aspergillus nidulans*, which produces the naphthopyrone YWA1 (**1.30**).[82]

In 2008, the Kelleher and Townsend groups studied the NR-PKS PksA from the aflatoxin B_1 (**1.32**) biosynthesis in *Aspergillus parasiticus* (**Figure I.11a**).[79] With respect to the TE/CLC domain, their results suggested that it performs a Claisen cyclization on the intermediate from the PT domain (discussed further below) leading to the formation of **1.31**. In 2010, Townsend and Tsai groups reported a 1.7 Å crystal structure, which provided additional information related to the structure and function of the C-C bond forming TE/CLC domain.[39] Their study revealed that this domain conducts a Claisen cyclization to release the product from the 4'-phosphopantetheine (Ppant)-tethered ACP affording noranthrone (**1.31**). Further biosynthetic processing steps results in formation of aflatoxin B_1.

The Townsend group also reported studies on fungal PKS TE function involved in melanin biosynthesis.[84] In 1995, the Kubo group proposed that a polyketide synthase from the *PKS1* gene in *Colletotrichum lagenarium* was involved in production of this pigment.[85] During the

biosynthesis of these molecules, a 1,3,6,8-tetrahydroxynaphthalene (THN) intermediate undergoes two rounds of enzymatic reduction and dehydration forming 1,8- dihydroxynaphthalene (DHN), which ultimately affords melanin. THN is the major product from the NR-PKS Pks1 obtained from the heterologous expression in *Aspergillus oryzae*.[86, 87] In their work, they performed an *in-vitro* characterization of Pks1 from *Colletotrichum lagenarium*. The Townsend group showed that Pks1 TE is both a Claisen cyclase and deacetylase of an enzyme-bound monocyclic hexaketide precursor that generates THN. This rare dual function demonstrates the notable versatility of the α/β hydrolase fold family of enzymes.

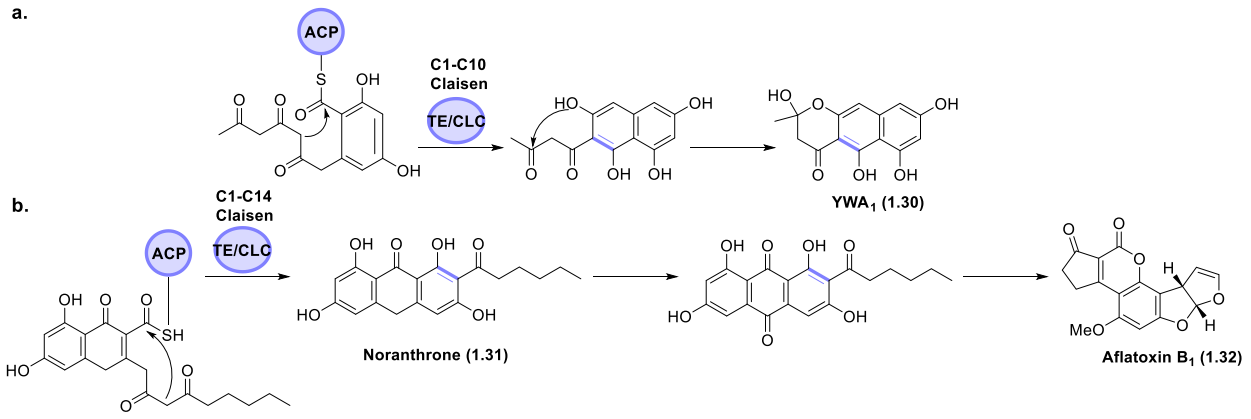


Figure I.11 a. The wA TE/CLC mechanism for the production of YWA1. b. PksA TE/CLC domain catalyzes the Claisen cyclization (C-C bond formation) of noranthrone.

The final example is the TE/CLC domain found in the herqueinone biosynthesis from *Penicillium herquei*. In 2016, the Houk and Tang groups studied the NR-PKS PhnA that is responsible for biosynthesis of a heptaketide backbone as well as for its cyclization into a hemiketal-containing naphtho-γ-pyrone prephenalenone (**1.33**).[80] They showed that the PhnA TE/CLC catalyzes a C1-C10 Claisen cyclization generating the corresponding naphthopyrone products (**Figure I.12a**). Remarkably, when the TE/CLC domain was deleted in PhnA (TE^0) shunt products **1.34** and **1.35** appeared, which are 8- and 10-membered ring macrolactones, respectively

(**Figure I.12b**). Their study revealed the basis for new and unexpected metabolites derived from polyketide cyclization.

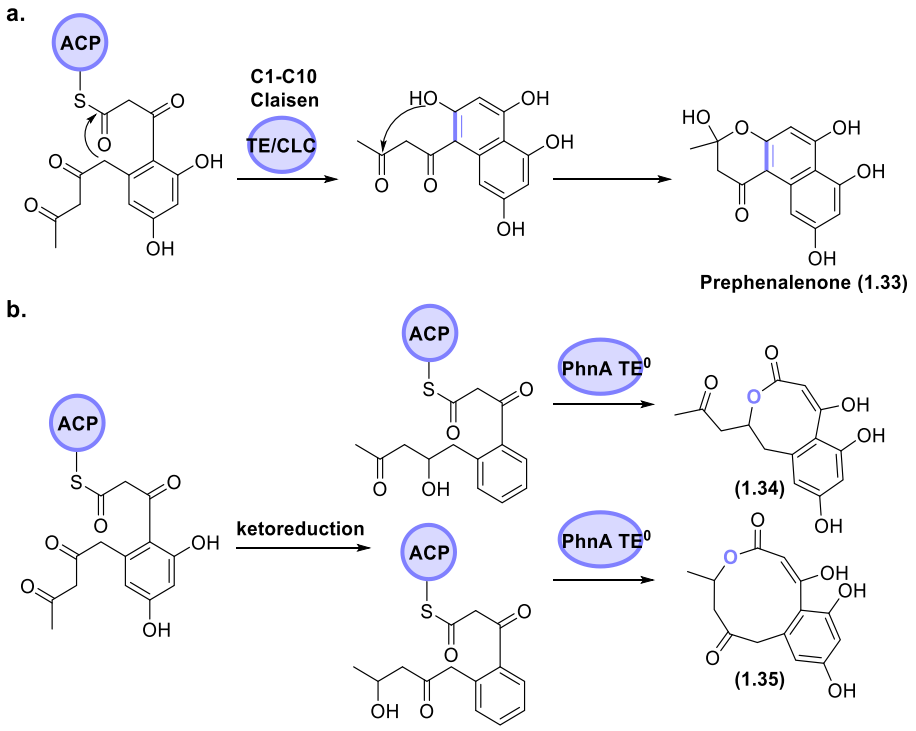


Figure I.12 a. PhnA TE/CLC C1-C10 Claisen cyclization producing the naphthopyrone prephenalenone. b. Shunt products from the PhnA TE deletion (TE^0).

I.5 Summary and Outlook on Thioesterases

In this section on TEs, fascinating work has been reported on peptide and polyketide biosynthetic cyclization catalysts over the past decade. Several examples of new cyclic structures were presented with interesting biological activities, and summarized the recent findings on the diversity of enzymes that catalyze cyclization of linear peptides or polyketide chains. These enzymes possess structural and mechanistic diversity. Even though TEs are relatively well understood in comparison with other cyclization catalysts, additional PKS and NRPS BGCs were recently characterized, showing the production of cyclic lactone and lactam structures mediated by a TE domain.[38, 40, 43, 47, 48, 50] Examples of these TE cyclases are derived from obafluorin,

salinamide A, lysobactin, sulfazecin, polymyxin B and fluvirucin B_1. In some cases, a cysteine instead of a serine active site residue was found[38, 40, 48, 50]; in addition to the unusual tandem TE architecture employed by lysobactin and teixobactin biosynthesis.[47, 59] Furthermore, chemoenzymatic syntheses[54, 59, 61-64] and pathway engineering in these systems[65-68] represent leading examples for the synthesis of novel compounds. TE domain oligomerization[71, 72, 76, 77, 88] and C-C bond forming Claisen-Like Cyclases[39, 79, 80, 84] (CLCs) were also described, which reflect the mechanistic diversity of these domains. The remarkable achievements toward understanding these systems show the potential of these cyclization biocatalysts. In the future, mechanistic, structural insights (X-ray crystallography and cryo-electron microscopy) and protein engineering (directed evolution) will guide the path for deeper exploration of the capabilities of the variant TEs for the synthesis of numerous natural products and bioactive, high value analogs. Recent crystal structures and mechanistic studies have demonstrated the impressive strategies resulting in cyclization of secondary metabolites.[80, 89-91] Future approaches using genome mining for the discovery of new biocatalysts, as well as protein engineering efforts on PKS and NRPS mega-enzyme systems are ripe with potential for the production of pharmaceutically relevant peptide and polyketide derived natural products.

I.6 Thesis Overview

My doctoral work focused on the chemical synthesis of polyketide substrates for the chemoenzymatic synthesis of macrolactones and macrolide antibiotics by application of modular PKSs, along with the enhancement of the thioesterase PKS domain macrocyclization selectivity through directed evolution. The first chapter of this dissertation covered the introduction and background on natural products as sources of medicines, PKSs, NRPSs, thioesterases and related enzymes for the generation of cyclic complex molecules.

The second chapter will cover the chemoenzymatic synthesis of macrolactone rings employing the last two PKSs modules within the Pik biosynthetic pathway. This portion covers the chemical synthesis of substrates with non-native functionalities for interrogation with the PikAIII-TE and PikAIII/AIV enzyme system. We successfully generated a subset of 14-membered macrolactone rings with different propensities for cyclization. This strategy coupled with the use of tailoring enzymes in the pathway and others, such as the use of glycosyltransferases and P450 oxidases allowed the generation of potentially novel macrolide antibiotics to be studied further.

The third chapter focuses on the directed evolution of the thioesterase Pik TE. Due to the discoveries that the TE appeared to be a catalytic bottleneck for macrocyclization, we generated an unnatural amide containing substrate for the study of the standalone enzyme. The TE reaction cyclization outcome, although low yielding for the amide substrate, allowed for the development of an engineering strategy aimed for yield improvement and ring selectivity. The Pik TE biocatalyst was engineered via site saturation and random mutagenesis, developing thousands of mutants (6,000+) to which 40+ have been efficiently characterized for the high yielding production of a hybrid macrolactone/lactam ring system.

The fourth chapter dives into the evaluation of the PikAIII-TE fusion protein and PikAIII/AIV for the generation of 12- and 14-membered macrolactone rings. These efforts started after the discovery that PikAIII-TE byproducts hindered the formation of the cyclized rings. We investigated potential reasons for the accumulation of byproducts, as well as fusion of the PikAIII module with other engineered Pik TEs from the directed evolution campaign. These resulted in the further investigation of PikAIII/AIV enzyme system and swapping of the wild type TEs with the engineered variants. The results indicated that depending on the TE variant incorporated to PikAIII-TE or PikAIV-TE (for the PikAIII/AIV system) diverse ratios of product formation were

obtained for the cyclized compounds. Although this is a work in progress, we envision that during the fifth and last chapter of this thesis, we will cover the knowledge gaps remaining for the generation of polyketide derived complex macrolactones and macrolides, in addition to the future work remaining for the generation of molecules with medicinal properties.

CHAPTER II

Chemoenzymatic Synthesis of Macrolactone Rings and Potential Macrolide Antibiotics Utilizing Polyketide Synthases (PKSs) and Post-PKS Tailoring Enzymes

The majority of this chapter is part of an ongoing manuscript "Structural Diversification of 14-Membered Macrolides by Chemoenzymatic Synthesis" and reproduced with permission from all authors Jennifer J. Schmidt, Maria L. Adrover-Castellano, Brian J. Curtis, Carolyn A. Glasser and David H. Sherman. *In progress*. **2023**.

II.1 Abstract

Antibiotic resistance is one of the biggest threats to global health, and it is estimated to continue incrementally every year, with 10 million estimated deaths worldwide by 2050. For these reasons, we are motivated to utilize PKSs enzymes to facilitate the design and generation of novel antibiotics from the macrolides class to improve the development of new, effective therapeutics. PKSs are found in the Pik biosynthetic pathway which is responsible for the generation of both a 12-membered macrolactone, 10-DML and a 14-membered macrolactone, NBL. A scalable *in-vitro* biocatalytic platform utilizing the final two modules PikAIII/AIV as well as the PikAIII-TE fusion protein enabled the formation of 12- and 14-membered macrocycles respectively, from a thiophenol activated Pik pentaketide native intermediate. Following the successful *in-vitro* work, other unnatural substrates have been synthesized for the interrogation of the selectivity and flexibility of these PKS modules for the processing of non-native pentaketides. Herein, we show the efficient production of new 14-membered macrolactones from a panel of unnatural

pentaketides using the PikAIII/AIV enzyme system. In addition, we have carefully studied the generation of an amide and ester containing 14-membered rings along with unexpected 6-membered rings resulting from these enzymatic reactions. It was determined that the Pik TE plays a key catalytic role in opening the 14-membered macrolactones to generate exclusively the 6-membered rings over time. The newly generated macrocycles were converted to their corresponding macrolides via appendage of D-desosamine sugars and hydroxyl groups.

II.2 Introduction

II.2.1 Antibiotic Resistance and Macrolide Antibiotics

Antibiotic resistance is one of the biggest threats to global health according to the World Health Organization (WHO). The Centers for Disease Control and Prevention (CDC) showed that in the US alone, it causes more than 2 million infections and 23,000 deaths a year. These alarming numbers are estimated to continue incrementally every year, with 10 million estimated deaths worldwide by 2050 (**Figure II.1**). Emerging resistance to current, clinically used antibiotics represents a threat facing the long-term efficacy of macrolide antibiotics in the 21st century, and the prospect of a dwindling supply of effective therapeutics.[92]

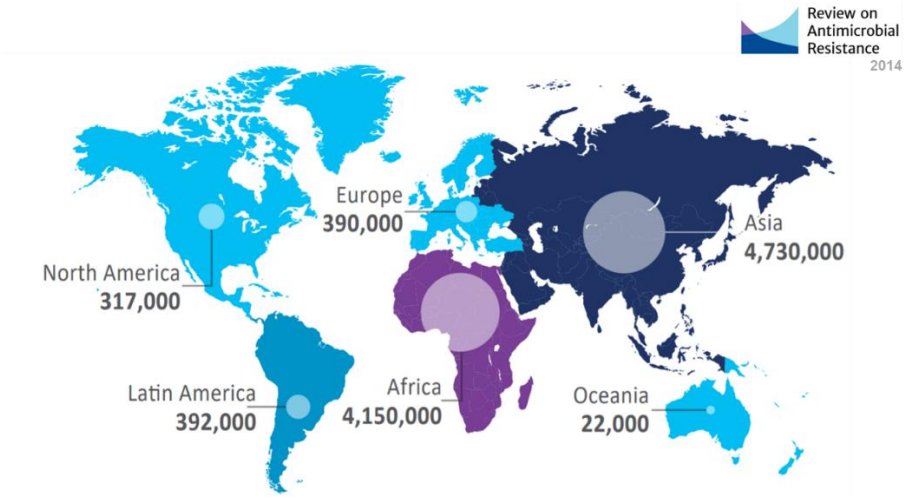


Figure II.1 Estimated deaths attributable to antibiotic resistance every year by 2050 (Figure from O'Neill, J., Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations, *The Review on Antimicrobial Resistance*, **2014**, 1-18, https://amr-review.org/ accessed on December 2023).

Macrolides are known to bind the 23S ribosomal RNA (rRNA) of the 50S ribosomal subunit blocking the peptide exit tunnel.[93, 94] As a result, translational elongation is inhibited, causing premature disassociation of the growing polypeptide chain from the ribosome, thus halting all protein synthesis and providing the basis for their antibacterial activity.[95, 96] However, bacterial resistance occurs through a variety of modifications, one of the most common of which is methylation of the rRNA portion of the 50S subunit, effectively blocking the macrolide binding site which confers resistance against the macrolides, lincosamides and streptogramins antibiotic classes (MLS).[97, 98] Semi-synthetic macrolides have been used since 1980 (**Figure II.2**) with the advent of clarithromycin (**2.2**) and later azithromycin (**2.3**) which showed better stability and improved pharmacokinetics compared to the first clinical macrolide erythromycin (**2.1**).[99] Recently, due to the threat of MLS resistance induced by the presence of the cladinose sugar on

many of the erythromycin derived antibiotics, the ketolide macrolides have been investigated with telithromycin (**2.4**) showing better efficacy in resistant strains.[92, 100]

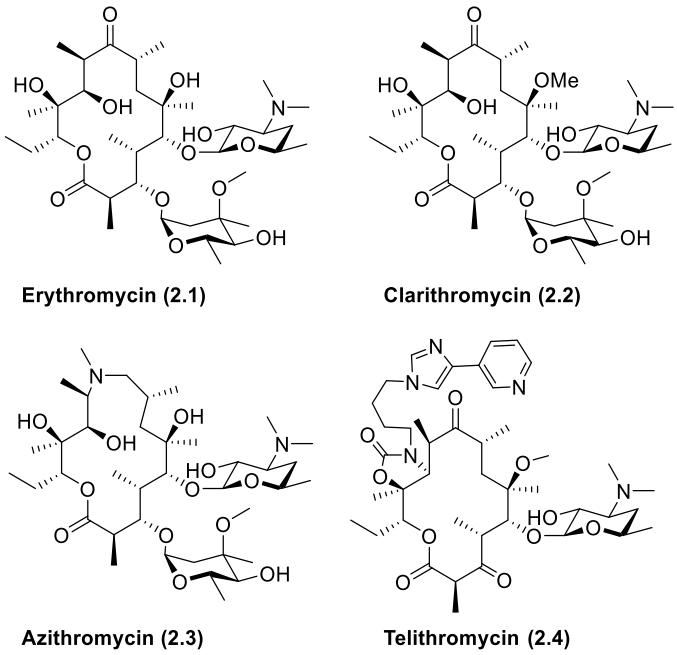


Figure II.2 Examples of clinically relevant macrolide antibiotics.

II.2.2 Macrolide Design and Applications of the Pik Biosynthetic Pathway

Current efforts are being directed towards the design and synthesis of novel macrolides to overcome bacterial resistance and instability shown with these antibiotics.[101-103] A subset of PKSs that generate these macrocycles, include pathways from the Pik[104], DEBS[105], Tyl[106] producing organisms. We have been focusing on the use of synthetic approaches to facilitate production of these compounds and their analogs using biocatalysis and enzyme engineering. The synthesis of diverse chain elongation intermediates in conjunction with late stage biosynthetic machinery (e.g. enzymes) facilitates access to a repertoire of molecules, which are otherwise challenging to generate using synthetic methods alone. Despite the success of these synthetic methods, they can generate waste and toxic materials having a negative impact on the environment and ultimately, human health.[15] As an alternative, we strive to develop PKS systems that will selectively perform

key transformations on polyketide chains to generate structurally distinct compounds while reducing waste, utilizing water as the reaction solvent and avoiding rare earth metal catalysts or toxic conditions.[17] PKSs offer an exciting avenue towards the production of diverse complex bioactive molecules as these enzymes are capable of the extension and processing of polyketide chains in a stepwise, stereospecific manner, organized by a series of catalytic domains that comprise distinct modules. [107, 108] At the end of the pathway, a terminal module, typically a TE domain catalyzes the formation of a cyclized product (10-DML (**2.5**) and NBL (**2.6**) for the case of Pik).[62, 109, 110] In Pik biosynthesis, the final bioactive compounds are obtained through further tailoring of the macrolactones by enzymes such as glycosyltransferases (e.g. DesVII) and P450 hydroxylases (e.g. PikC) obtaining the desired macrolides including pikromycin (**2.7**) (**Figure II.3**).[111]

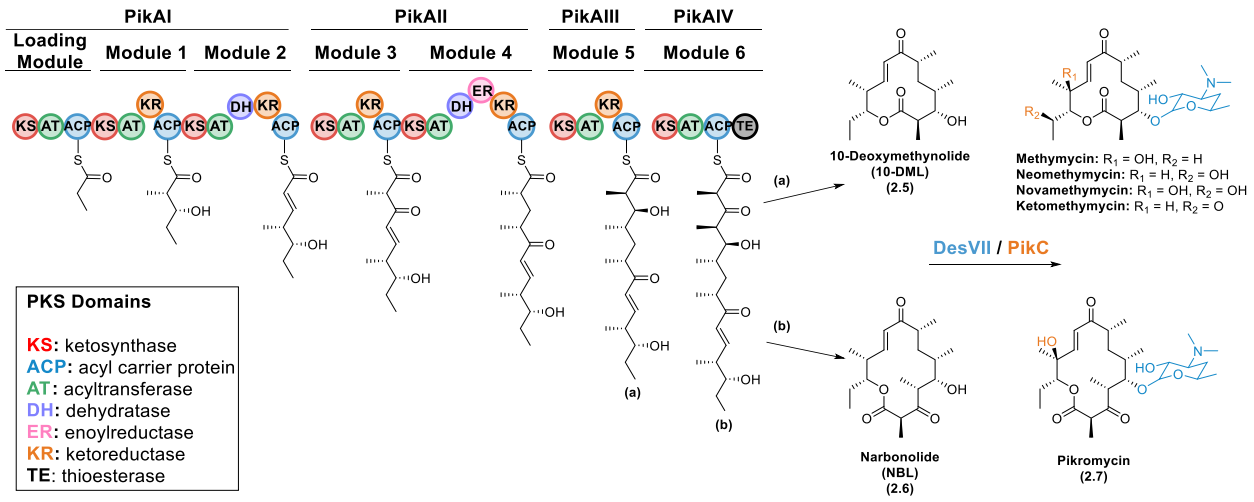


Figure II.3 Pikromycin biosynthetic pathway.

II.2.3 In-vitro Characterization of PikAIII-TE Fusion Protein and PikAIII/AIV System

The synthesis of unnatural chain elongation intermediates allows for the creation of structural diversity using the pikromycin and other PKS modules to produce new macrolactones.[61, 62, 112-115] Previous work by former graduate student in the Sherman group, Dr. Douglas A. Hansen leveraged

the Pik pathway utilizing synthetic chemistry in conjunction with either the unnatural fusion protein PikAIII-TE or PikAIII/AIV developing an *in-vitro* scalable biocatalytic platform for the production of 12- and 14-membered ring macrocycles (**Figure II.4**).[61] To accomplish this part, an efficient and scalable synthesis of the natural substrate, Pik pentaketide (**2.8**) activated as the thiophenol ester was achieved, as it was shown to be an effective substrate activation required for a Pik module.[112] For PikAIII-TE, the reaction proceeds by 1) the AT domain utilization and transfer to the ACP domain of a synthesized MeMal-SNAc extender derivative to incorporate 2-carbon units into the polyketide chain. 2) The KS domain accepts the pentaketide and proceeds with a decarboxylative Claisen condensation with the extender unit to the ACP. 3) The KR domain performs a reduction at the C3 position mediated by a recycling system composed of glucose 6-phosphate (G6P), glucose 6-phosphate dehydrogenase (G6PDH) and nicotinamide adenine dinucleotide phosphate ($NADP^+$). Lastly, the TE domain performs cyclization with the distal hydroxyl group existing in the polyketide. For the PikAIII/AIV system, it follows identical mechanism as described for PikAIII-TE with the difference of another round of extension with MeMal-SNAc and final TE domain cyclization (**Figure II.4**).[111, 116]

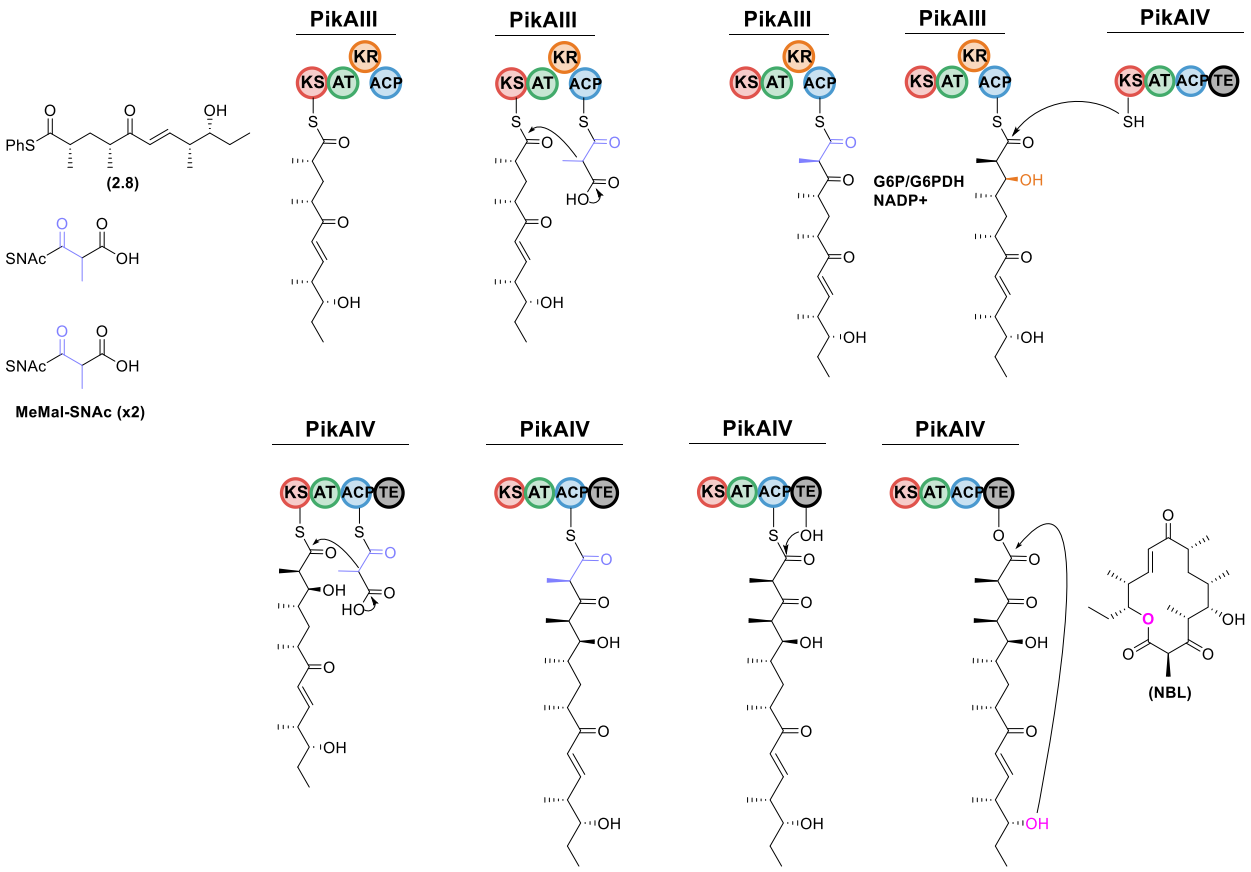


Figure II.4 Mechanism of PikAIII and PikAIV enzyme system for the generation of macrolactone rings.

For the Pik pentaketide (**2.8**), the conversion of 10-DML (**2.5**) with the engineered fusion protein PikAIII-TE occurred with 60% isolated yield and NBL (**2.6**) using PikAIII/AIV was obtained with 49% isolated yield. Furthermore, with engineered strains of *Streptomyces venezuelae* direct appendage of D-desosamine sugar and final P450-mediated C-H oxidation(s) steps enabled efficient synthesis of the macrolides (methymycins and pikromycin). The biotransformation to form a suite of mature macrolides was completed with ~70% isolated yields (**Figure II.5**).[61]

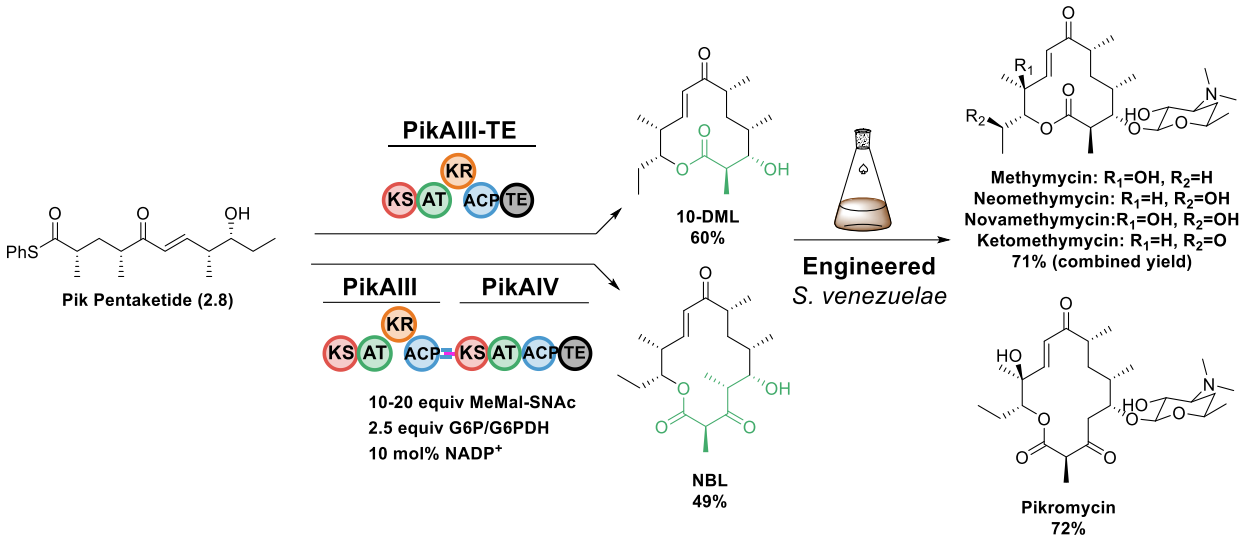


Figure II.5 Scalable *in-vitro* chemoenzymatic synthesis of macrolactones using either PikAIII-TE or PikAIII/AIV, and whole cell bioconversion using engineered *S. venezuelae* to macrolides. Macrolactone components shown in green are introduced by the PKS modules as a result of chain elongation, processing and cyclization through one (10-DML), or two cycles (NBL).

II.2.4 Synthesis of Unnatural Substrates to Interrogate the Flexibility of PikAIII-TE Fusion Protein for the Production of New Macrolactones

Furthermore, a library of unnatural pentaketides were synthesized mainly to interrogate the capabilities of PikAIII-TE fusion protein for the generation of 12-membered macrolactone rings applying the *in-vitro* platform shown above for the Pik system. These substrates were termed "truncated pentaketides" due to having truncations in the left half of the molecule (**2.9-2.12**) corresponding to the methyl group at C8 position in the native Pik pentaketide (**2.8**). Additionally, substrates bearing unnatural C8 and C9 stereocenters were designed (**2.13** and **2.14**). These set of substrates were then subjected to PikAIII-TE for macrocyclization. The corresponding macrolactone products **2.15-2.19** were reported, with isolated yields of 66% to 9%, showcasing the PikAIII-TE moderate flexibility and ability to process substrates with non-native

functionalities **(Figure II.6)**. Despite the success of the newly generated macrocycles, substrates with stereocenters **2.13** and **2.14** failed to be converted to the macrolactones **2.20** and **2.21**. Instead, the molecules identified corresponded to unidentifiable decomposition of the starting material or byproducts, along with trace amounts of extended products off-loaded as linear shunt products.[62]

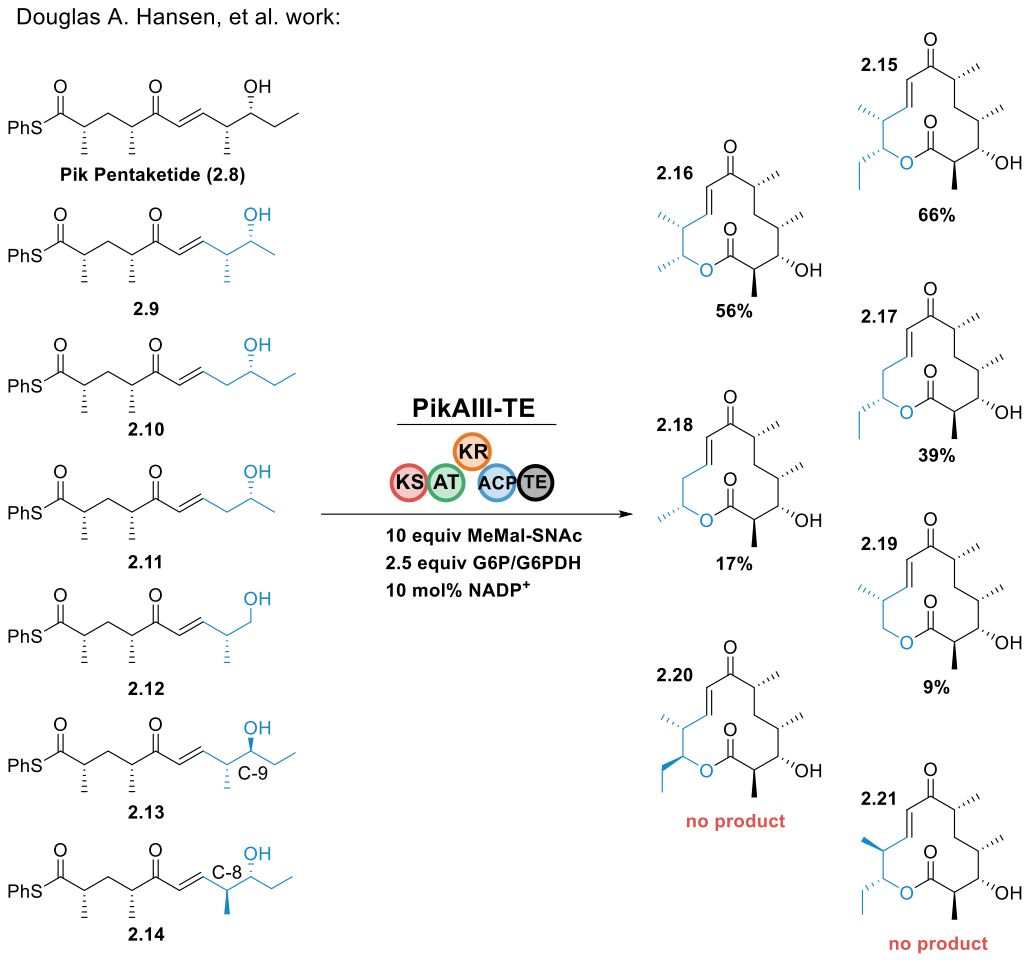


Figure II.6 Evaluation of PikAIII-TE fusion protein against a panel of truncated (**2.9-2.12**) and stereoisomer (**2.13** and **2.14**) pentaketides. Highlighted in blue are the differences of these substrates in comparison to the Pik pentaketide (**2.8**).

II.2.5 Synthesis of Unnatural Hexaketides to Investigate the TE as a Catalytic Gatekeeper

Considering macrolactonization in synthetic chemistry has been notoriously challenging[117, 118], further investigations to the TE domain as a catalytic gatekeeper for macrolactone formation was explored. Former graduate students in the Sherman group, Dr. Aaron A. Koch alongside Dr. Douglas A. Hansen identified the stereoisomer pentaketides (**2.13** and **2.14**) appeared to have been processed by the catalytic domains present in PikAIII-TE due the appearance of extended products prematurely off-loaded as linear shunt products, suggesting potential issues with the TE domain for cyclization.[62, 65] To investigate this hypothesis, a native methyl-protected Pik hexaketide (**2.22**) and a C11 epimerized nucleophilic hydroxyl group pentaketide (**2.23**) also methyl-protected were synthesized for direct probes for the Pik TE. These were activated as SNAc thioesters, as it was previously determined as the ideal substrate activation for TE loading.[112] The results showed the native hexaketide (**2.22**) to be converted by Pik TE WT in a 61% conversion, whereas the epimerized hexaketide (**2.23**) failed to undergo cyclization and only hydrolysis product was identified in a 63% conversion (**Figure II.7**). Since the hydrolysis product is TE mediated, it shows that cyclization was entirely blocked, suggesting the TE indeed is a key catalytic bottleneck worth studying further.[62, 65]

In an accompanying study, an engineered version of the Pik TE was achieved, by mutating the catalytic serine 148 into a more nucleophilic cysteine (Pik TE S148C).[65] The enzymatic reactions showed an improved 94% conversion of the native hexaketide (**2.22**) and the appearance of the cyclic product with 83% conversion of the epimeric hexaketide (**2.23**), highlighting the improvement of this mutation to the reaction outcome (**Figure II.7**). Additional computational analysis through molecular dynamics (MD) and quantum mechanics (QM) revealed an improvement in the reaction kinetics and a change in mechanism from a stepwise addition-

elimination to a concerted acyl substitution lower in energy resulting in a gain of function for the Pik TE. Interestingly, when the Pik TE S148C was reincorporated into the PikAIII-TE module, generating PikAIII-TE S148C, a lower 12% yield for the cyclized product was observed, alongside a C3-keto product also in a 12% yield (suggesting KR failing in the module) (**Figure II.7**). Nonetheless, even though the individual domains appear to readily accept a non-native substrate and the TE is shown to perform cyclization, it continues to suggest the TE is likely the catalytic gatekeeper for unproductive formation of the desired macrolactone products.[65]

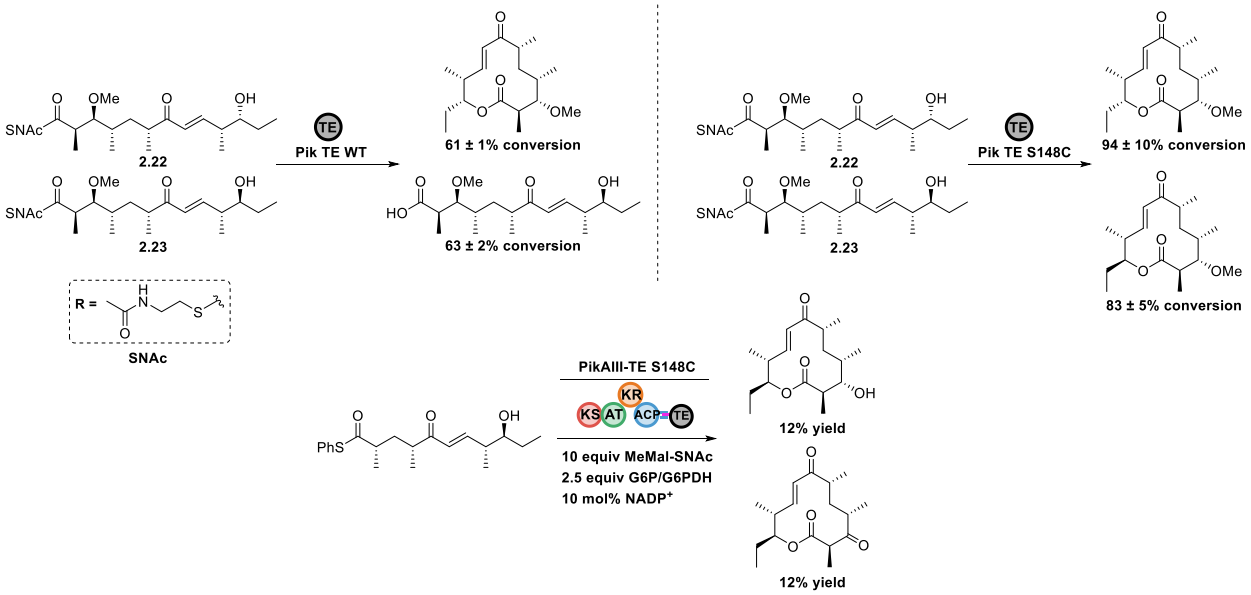


Figure II.7 Native and epimeric hexaketides enzymatic reactions with Pik TE WT and Pik TE S148C. Epimeric pentaketide enzymatic reaction with reincorporation of Pik TE S148C generating PikAIII-TE S148C.

II.3 Development of New Pik Pentaketide Analogs

Continuing the diversification and discovery of novel macrolactones and macrolides, a former graduate student also in the Sherman group, Dr. Jennifer J. Schmidt synthesized four additional unnatural pikromycin pentaketide analogs to probe even further the flexibility of PikAIII-TE and

PikAIII/AIV (unpublished work). These pentaketides contained functionalities not possibly found in the Pik system, including heteroatoms in place of carbons (amide **2.22** and ester **2.23**) and altered side chains (gem-dimethyl **2.24** and cyclohexyl **2.25**) for the use in PikAIII-TE fusion protein and PikAIII/AIV enzyme system (**Figure II.8**). In addition, Dr. Douglas A. Hansen truncated pentaketides[62] **2.9-2.12** were also investigated, this time around employing PikAIII/AIV for the potential generation of the corresponding 14-membered rings not studied previously, as only the PikAIII-TE enzyme was utilized to probe the formation of the 12-membered ring products. The goals of this study are to continue to shed light on the mechanistic deciding factors, including loading, extension, processing and cyclization of these intricate PKS modular enzymes and single domains, as well as generating new Pik macrolide analogs (by macrolactone appendage of D-desosamine sugar and final hydroxylations steps) to be tested for antibiotic activity.

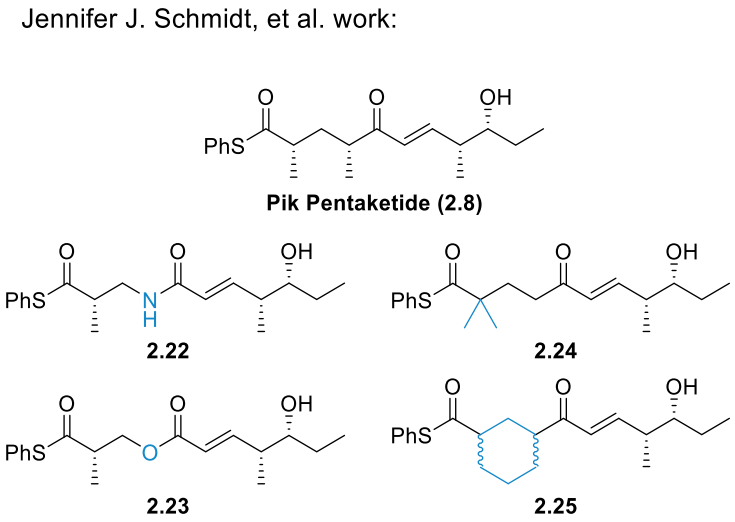


Figure II.8 New Pik pentaketide analogs. Highlighted in blue are the differences of these substrates in comparison to the Pik pentaketide (**2.8**).

II.4 Evaluation of New Pentaketides Against PikAIII-TE and PikAIII/AIV

II.4.1 PikAIII-TE Enzymatic Reactions

With the newly generated pentaketides in hand, we then proceeded to test them against the PikAIII-TE fusion protein to question the capabilities of the modular enzyme to generate the corresponding 12-membered macrolactone rings. Analytical scale enzymatic reactions of the four new pentaketides (amide, ester, gem-dimethyl and cyclohexyl) compared to no-enzyme controls showed no identifiable traces of 12-membered rings via liquid chromatography-mass spectrometry (LC-MS) analysis. Instead, extended and hydrolyzed byproducts for amide (**2.22**) and ester (**2.23**) pentaketides were identified, suggesting KS-loading, chain extension and β-processing of these analogs. Once again, these results suggest PKS domain flexibility to an extent, however the TE does not perform cyclization further aligning with the TE as a catalytic bottleneck investigations. Now, the gem-dimethyl (**2.24**) and cyclohexyl (**2.25**) pentaketides showed minimal substrate consumption, or no consumption at all, suggesting PKS flexibility has impediments for alkyl groups and bulkier substituents in close proximity to the KS-loading site. Subjecting these two pentaketides to a KS-AT didomain suggested via intact protein mass spectrometry that the pentaketides are in fact loaded to the KS, however no Claisen condensation occurs when utilizing MeMal-SNAc and the ACP domain *in trans*. Subsequently, due to the previous success of the truncated pentaketides **2.9-2.12** with PikAIII-TE for cyclization, and the gain of function for the PikAIII-TE S148C, we also tested the new pentaketides against the later protein. Similarly, the results showed no evidence of ring formation, and instead SNAc derived byproducts were identified.

II.4.2 PikAIII/AIV Enzymatic Reactions

Analogously, we tested the four new pentaketides as well as the four truncated pentaketides. For the truncated ones, we had to generate more of these substrates synthetically, as these were shown before by Dr. Douglas A. Hansen, and we only had limited quantities available for our purpose.[62] Having enough material of the eight pentaketides total allowed us to test first analytical scale enzymatic reactions with the PikAIII/AIV system. The resulting LC-MS data suggested formation of the 14-membered rings for the amide (**2.22**) and ester (**2.23**) and complete consumption of starting material. To our surprise, we identified two main peaks possessing the same mass to charge ratio (m/z) of the expected amide and ester ring products. The gem-dimethyl (**2.24**) and cyclohexyl (**2.25**) pentaketides showed once again minimal substrate consumption comparable to the analytical scale reactions performed with the PikAIII-TE enzyme. In the case of the four truncated pentaketides (**2.9-2.12**), these showed formation of the desired macrolactones and complete consumption of starting material (see experimental section at the end of the chapter).

Next, we performed preparative scales enzymatic reactions with ~10-20 mg of starting material to isolate and characterize the corresponding rings. The separation of fractions and analysis was conducted via high performance liquid chromatography (HPLC), and the products were confirmed by LC-MS and NMR. The amide and ester macrolactones were isolated as two compounds, one being the expected 14-membered ring macrolactones (**2.26**-**2.27**: 15-17% yields) and the others were 6-membered rings (**2.32**-**2.33**: 15% yields both). The resulting rings from the truncated pentaketides were isolated as a dynamic equilibrium of the hemiketal form between the C5 hydroxyl and the C9 α/β carbonyl for **2.29**-**2.30** and/or C2 isomerization caused by tautomerization for the case of **2.28**. Macrocycle **2.28** was isolated in a 71% yield of a 5:1 mixture with the tautomerization product. This was confirmed by NMR deuterium incorporation experiments,

showing a change from the doublet methyl group to a singlet. For **2.29** and **2.30**, these were isolated in 45% and 11% yields respectively. They both appeared as a dynamic equilibrium between the desired macrolactone and the hemiketalized product (8:1 for **2.29** and **2.30**) as confirmed by NMR. Lastly, macrolactone **2.31** was isolated in a 20% yield as a single isomer. Having 6/8 pentaketides processed and cyclized by the PikAIII/AIV module shows the flexibility of individual domains working together to assemble complex macrocyclic structures with functionalities not found in the native system for the production of narbonolide analogs (**Figure II.9**).

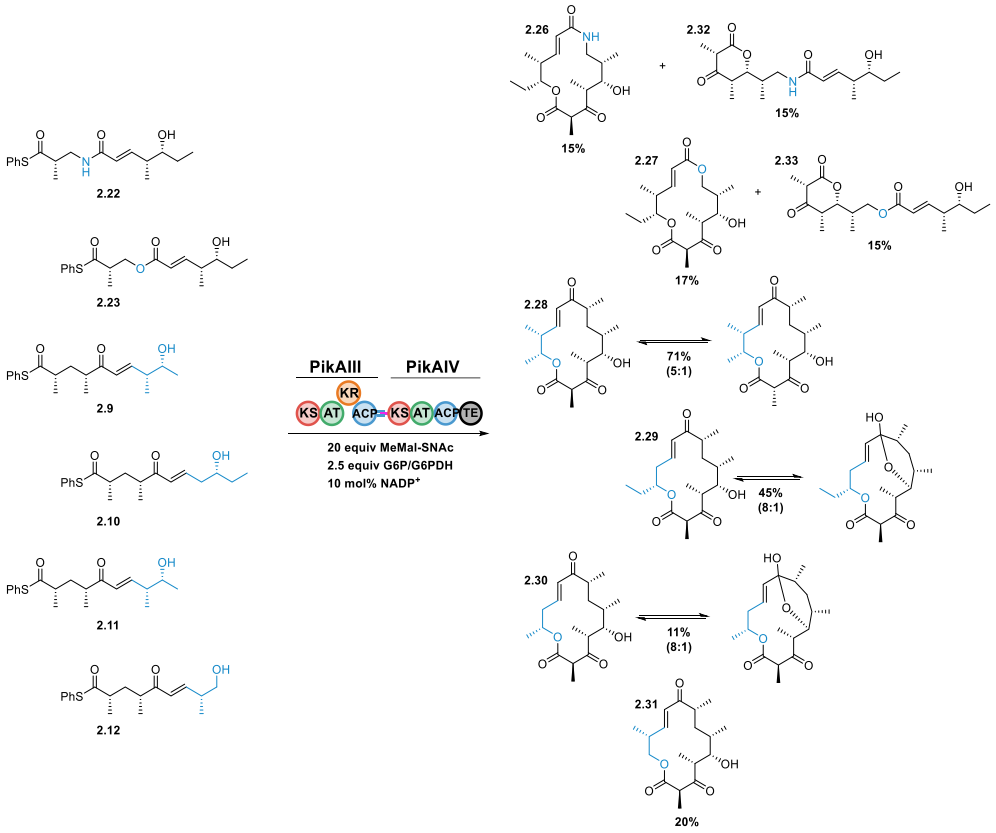


Figure II.9 New pentaketides tested against PikAIII/AIV for the generation of 14-membered macrolactone rings. Highlighted in blue are the differences of these substrates in comparison to the Pik pentaketide (**2.8**).

II.5 Identification of a Novel Macrolactone Ring Opening Activity for Pik TE

Based on the observation that the amide and ester pentaketide reactions against PikAIII/AIV generated 6-membered macrolactone rings alongside the desired 14-ring macrocycles, an ongoing hypothesis emerged, suggesting the TE performed cyclization using both hydroxyl groups present in the extended substrate. The first option was the TE induced nucleophilic addition of the C13 hydroxyl group onto the thioester bond generating the 14-ring. The second alternative related to the C5 hydroxyl group TE mediated nucleophilic addition to yield the 6-ring (**Figure II.10**). These results raised questions about the TE poor selectivity, low efficiency and yields for the isolation of the desired larger ring sized product.

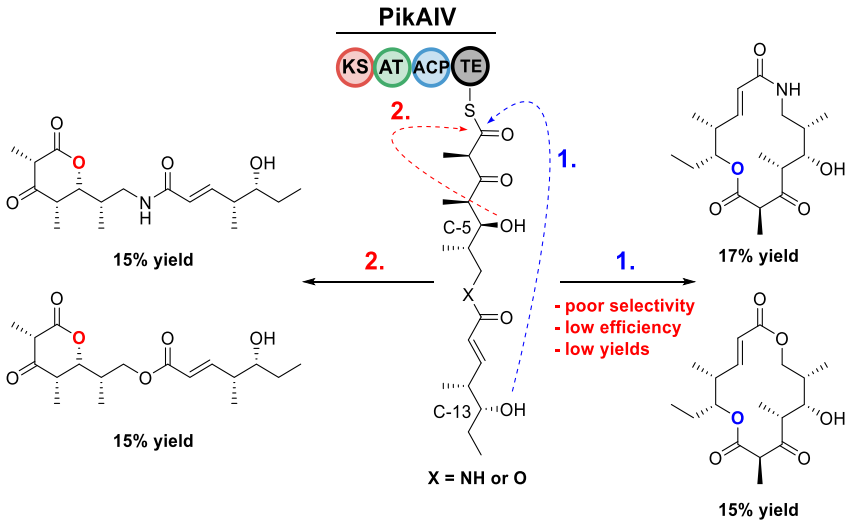


Figure II.10 Proposed TE mechanism for the generation of both 6- and 14-membered rings present in the amide and ester enzymatic reactions with PikAIII/AIV.

To investigate the hypothesis further, we decided to take a deeper look into the reaction outcome through LC-MS. Typically, the reactions both analytical and preparative scales were run at a 2 hour total time. As a starting point, we decided to utilize the amide pentaketide (**2.22**) against PikAIII/AIV taking aliquots at 1 hour, 2 hour and overnight (24 hours) time points. Surprisingly,

we determined that in the overnight reaction there is the complete disappearance of the 14-ring and only the appearance of even more 6-ring when compared to the 1 hour and 2 hour time points (**Figure II.11**). The reaction outcome suggested potential thermodynamically favored opening of the 14-membered ring into a linear intermediate that would then re-cyclize into the 6-ring. The question now was which module, catalytic domain and/or reaction component/byproduct is responsible for the hypothesized new mechanism?

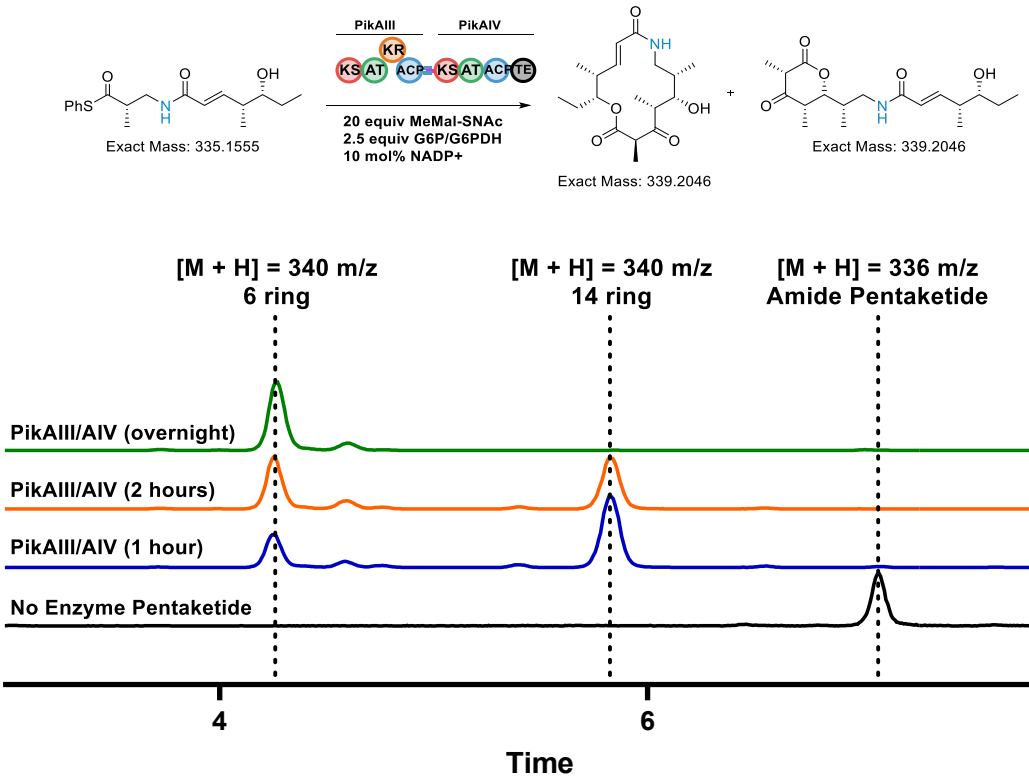


Figure II.11 Total ion chromatogram (TIC) for amide pentaketide enzymatic reaction against PikAIII/AIV for 1 hour (blue), 2 hour (orange) and overnight (24 hours, green) time course to investigate formation of the 6-membered macrolactone ring. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run.

Initially, we sought to investigate if any reaction components/byproducts or thiols present in the enzymatic reaction could be responsible for the ring opening mechanism. We evaluated the amide isolated 14-membered ring incubation with just reaction buffer, free SNAc, thiophenol, recycling system (G6P, G6PDH, $NADP^+$), recycling system and MeMal-SNAc, and 2-vinyl pyridine as a thiol-scavenger. None of these experiments suggested the presence of reaction components or free thiols in the reaction to be responsible for the 6-membered ring formation (see experimental section at the end of the chapter).

Interestingly, previous literature reports have demonstrated the TE from the epothilone biosynthetic pathway's capability to re-open the epothilone C macrolactone ring through a cyclo-release process. This TE-mediated mechanism involves water, leading to the hydrolyzed seco-epothilone C product.[119] With this information, we continued with examining the direct opening of both amide and ester rings by performing enzymatic reactions of the isolated macrolactones as the starting material and utilizing PikAIV and the standalone Pik TE WT and Pik TE S148C mutant. For PikAIV and Pik TE WT, the results showed the 14-membered ring is indeed being opened and re-cyclized to the 6-membered ring by the TE, as the 14-membered macrocycle is completely depleted from the overnight reaction as seen on LC-MS analysis (**Figure II.12, Figure II.13**). However, for the Pik TE S148C, it seems to be slower at degradation of the 14-ring as a 1:1 ratio of both ring sizes is still observed in the overnight reaction traces for amide and ester. It was also identified that the ester macrocycle is opened much faster by the Pik TE WT, as the 2 hour reaction trace shows only 6-ring formed. In the other hand, for the amide macrocycle, there is almost a 1:1 ratio of both 6- and 14-ring in the 2 hour reaction trace (**Figure II.12**). The same trend is observed also for amide and ester against PikAIV (**Figure II.13**). To further demonstrate the novel TE catalytic activity, we assessed an inactive form of the Pik TE through the mutation

of catalytic serine 148 to alanine (Pik TE S148A). Incubations of the amide and ester macrolactones with this inactive TE revealed no degradation of the 14-membered ring and no formation whatsoever of the 6-membered ring. This compelling evidence substantiates our findings, indicating that the TE serves as the catalyst responsible for the ring opening and closing of these macrolactone rings (see experimental section at the end of the chapter).

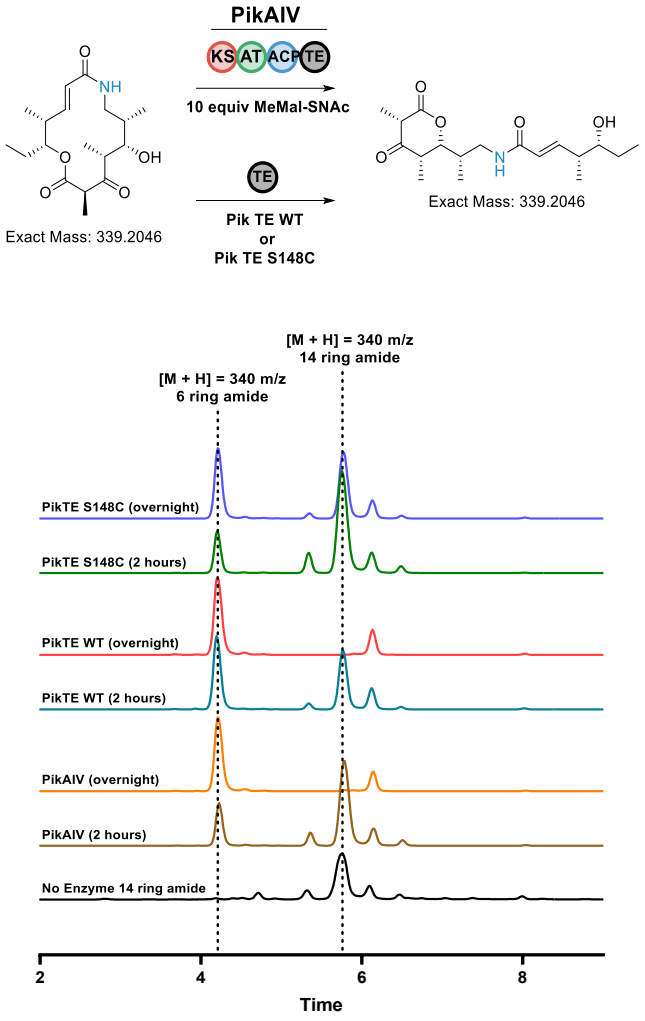


Figure II.12 Total ion chromatogram (TIC) for amide 14-membered macrolactone ring reaction against PikAIV module and standalone Pik TE WT and Pik TE S148C to test macrocycle opening

and re-cyclization mediated by the TE. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run.

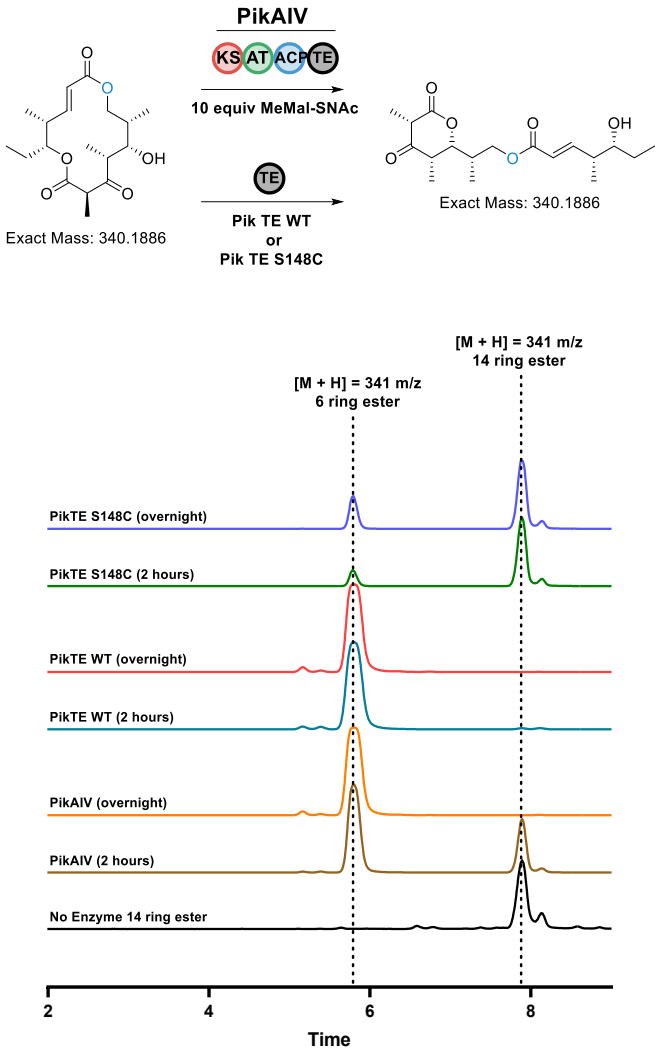


Figure II.13 Total ion chromatogram (TIC) for ester 14-membered macrolactone ring reaction against PikAIV module and standalone Pik TE WT and Pik TE S148C to test macrocycle opening and re-cyclization mediated by the TE. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run.

Another hypothesis remained relating to the amide and ester 6-membered ring formation. Our previous analysis with incubation of the macrolactones with PikAIV and the standalone TEs confirms that the TE has a catalytic role for the opening and re-cyclization of these rings. However, another possibility would be the Pik TE performing ring opening, then hydrolyzing of the heptaketide intermediate for both amide and ester leading to a spontaneous intramolecular cyclization of the free acid. Though, our current studies suggest this hypothesis is less likely, since reaction time courses showing consumption of the 14-membered ring and appearance of the 6-membered ring, the heptaketide acid identified in LC-MS traces does not decrease over time, in fact it increases (see experimental section at the end of the chapter).

II.6 *In-vivo* Glycosylation and Late-stage Macrolide C-H Oxidation (Post-PKS Tailoring Steps)

Having six new macrolactones from our preparative scale enzymatic reactions with PikAIII/AIV, we proceeded to generate the macrolide counterparts for future antimicrobial evaluation. For this purpose, utilizing an engineered strain of *Streptomyces venezuelae* (DHS316) that contains the PKS genes knocked out, as well as the P450 genes, we assessed the efficiency of the glycosyltransferases (PikBVII/PikBVIII) to append D-desosamine *in-vivo* into the macrocyclic scaffold.[120] The macrolactones (~10 mg scale) were spiked to growing cultures of the strain, along with acetyl-narbonolide activator and monitored for 8, 18 and 36 hours or until completion in the LC-MS instrument. Compounds **2.28**-**2.31** showed complete conversion to their macrolide counterparts (**2.34**-**2.37**). Further isolation in semi-prep HPLC furnished yields between 65-75%. This strain was also employed to assess amide **2.26** and ester **2.27** 14-membered macrolactones for its ability to produce the corresponding macrolides. Interestingly, neither biotransformation went to completion after monitoring for 72 hr. Further scale up and HPLC isolation and

characterization furnished 38% isolated yield of the amide containing macrolide **2.38** (with 35% recovered starting material) and a 33% isolated yield of the ester containing macrolide **2.39** (with 40% recovered starting material). Despite the overall lower isolation yields from these reactions, the recovery of intact macrolactone starting material allows for the increase in yields by successive biotransformation treatments for the production of narbomycin analogs.

Upon identification and characterization of the glycosylated macrolides, these were further functionalized via the P450 enzymes, both PikC from the pikromycin pathway and JuvD from the juvenimicin pathway to generate pikromycin analogs. PikC natively installs an allylic hydroxyl group at C12 in the macrolactone core. This enzyme has been studied previously, which lead to the development of an engineered form PikC D50N having improved oxidation turnover.[121-123] Analytical reactions indicated by mass in LC-MS of PikC WT-RhFRED and PikC D50N-RhFRED revealed that the variant formed performed better, and was the variant chosen for our next oxidative tailoring steps. An interesting finding from the initial reactions included the observed tendency of the methyl group at C12 position to be necessary for PikC to perform its function. This was evidenced from the reactivity of C12 methyl containing macrolides **2.34**, **2.37**, and **2.39**, however, no oxidation appeared from C12 methyl amide macrolide **2.38**. Whereas macrolides **2.35** and **2.36** both showed no evidence of oxidation presumably due to the lack of the C12 methyl group. Scale up, isolation and characterization revealed the PikC products to be the corresponding oxidized macrolides **2.40**, **2.42** and **2.44** isolated in 71% yields to trace amounts. Interestingly, the ester macrolide appeared to turnover to two products with the expected mass, whereas the amide macrolide showed no evidence of oxidation on an analytical scale. Moreover, we have previously shown that the JuvD P450 from the juvenimicin pathway catalyzes formation of a C15 primary oxidation product on unnatural macrolide substrates.[120] Analytical reaction with JuvD WT-

RhFRED revealed that a C15 methyl group could be necessary for activity, with **2.35** turning over to a product **2.41** of the expected mass. The ester and amide containing narbomycin analogs were also analyzed and although the ester **2.39** did not show any turnover to products **2.44**, the amide macrolide **2.38** was converted to multiple products **2.43** associated with a mass expected for a single oxidation (**Figure II.14**).

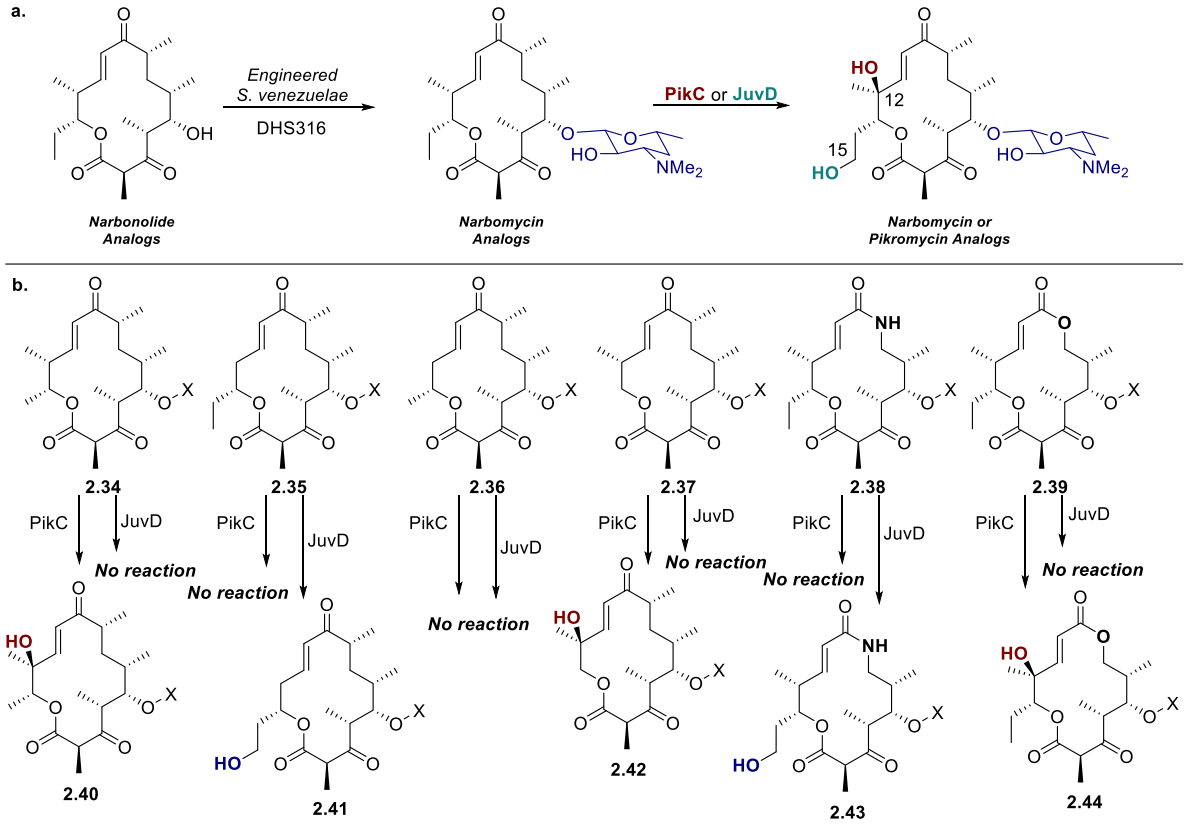


Figure II.14 a. Scheme depicting the generation of narbonolide, narbomycin and pikromycin analogs from our strategy and inspired by the Pik biosynthesis. b. Glycosylated macrolides and corresponding oxidated macrolides showcasing the reactivity differences between PikC and JuvD P450s.

II.7 Conclusion/Summary

In summary, we developed a platform that allowed for the development and exploration of novel substrates bearing non-native functionalities for the interrogation of PKS and post-PKS biocatalysts to generate macrolides. We started by synthesizing the corresponding substrates intended for reactions with PikAIII/AIV enzyme system to form macrolactone rings with yields between 12-71%. We also demonstrated the ability of Pik TE for the ring opening and closing of the amide/ester 14-membered macrolactones to the 6-membered ring products, which showcased a novel activity for the TE cyclization. Next, the macrolide counterparts were successfully produced utilizing an engineered strain of *Streptomyces venezuelae* (DHS316), and P450 oxidases PikC alongside JuvD with yields between 71% yields to trace amounts. The discovery of interesting reactivities, including the C12 methyl group for PikC catalysis, and the C15 primary oxidation by JuvD for the macrolides could shed light on the functionalities in the macrolactone scaffold required for the generation of diverse macrolides. Future studies with additional P450s, as well as the study of the macrolide bioactivities (antibiotic, anticancer, immunomodulators, among others) will be valuable for the continuous development of polyketide derived medicines such as the narbonolide, narbomycin and pikromycin analogs showed in this portion.

II.8 Experimental

Methods and Reagents

II.8.1 General Information, Materials and Instrumentation

Chemical reagents and solvents were purchased from EMD Millipore, Sigma Aldrich, Oakwood, Combi Blocks, Chem Impex, and Thermo-Fisher Scientific, unless otherwise noted. All reactions were performed in flame dried glass wear under an atmosphere of dry nitrogen, any changes will be stated. Reactions run at elevated temperatures were controlled by IKA RET control

visc (model RS 232 C), reactions run at room temperature (rt) was ~21-23 °C, reactions run cooler than room temperature were performed in an ice bath (0 °C), dry ice/acetone (-78 °C), or isopropanol/ThermoNESLAB (model CC100) for extended times/other temperatures. Commercially available starting materials and reagents were used as received unless otherwise noted. Dichloromethane was purchased as HPLC grade from Fisher and used directly. Tetrahydrofuran, *N,N* dimethylformamide, and pyridine were purchased as anhydrous and unstabilized (EMD Millipore DriSolv). Acetone was dried over Na_2SO_4 prior to use for reactions. Diisopropylamine and trimethylamine were distilled from calcium hydride directly prior to use.

^{1}H NMR spectra were recorded on either a Bruker 600 NMR System (600 MHz) or Varian 400-MR (400 MHz) spectrometer. Chemical shifts are reported in parts per million (ppm) using the solvent resonance as an internal standard ($CDCl_3$ 7.26 and CD_3OD 3.31 ppm). Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, h = hextet), coupling constant (Hz), and integration. Proton decoupled ^{13}C NMR spectra were recorded on Bruker 600 NMR System (150 MHz). Chemical shifts are reported in ppm using relative to residual solvent peaks ($CDCl_3$, 77.0 ppm or CD_3OD 49.0). Analytical thin layer chromatography was performed on EMD Millipore 0.25 mm silica gel F_{254} plates. Visualization was accomplished by a combination of 254 nm UV lamp and either potassium permanganate ($KMNO_4$), ceric ammonium molybdate (CAM), or p-anisaldehyde. Purifications were performed by forced air flash chromatography using EMD Millipore Silica Gel 60 (40 – 63 μm) or a Biotage Isolera one flash purification system. Columns used with Biotage Isolera one purification system include Biotage SNAP Ultra, SiliaSep, SiliaSepHP, and SiliaSep amine cartridges.

Deionized water was obtained from a Milli-Q system (EMD Milipore) using Q-Gard 2/Quantum Ex Ultrapure organex cartridges. Media components for *E. coli* growth were purchased

from EMD Millipore, Sigma Aldrich, and Thermo-Fisher Scientific. Glycerol was purchased from Thermo-Fisher Scientific. Lysogeny Broth (LB) broth and LB agar (Miller) were purchased in pre-made granulated form from EMD Millipore. Terrific Broth (TB) broth was made from individually purchased components and consisted of 4% v/v glycerol. Media and solutions were autoclaved or sterile filtered before use. For all solutions, pH was monitored using a VWR symphony SB70P pH meter calibrated according to the manufacturer's specifications. Kanamycin sulfate and Isopropyl β-D-1-thiogalactopyranoside (IPTG) were obtained from Gold Biotechnology. Lysozyme was purchased from DOT Scientific. Imidazole was purchased from AK Scientific. Amicon Ultra centrifugal filters and PD10 desalting columns used for protein concentration and buffer exchange were purchased from EMD Millipore, GE Healthcare or Cytiva.

PCR was completed using a Bio-Rad iCycler thermal cycler system. DNA oligonucleotides for cloning and mutagenesis were purchased from Integrated DNA Technologies (IDT). Restriction endonucleases and other associated molecular biology reagents were purchased from New England Biolabs (NEB). PCR cleanup/gel extraction and plasmid miniprep kits were purchased from QIAGEN. All DNA manipulations were accomplished following the manufacturer's protocols. DNA concentrations and protein purity indexes were measured using a NanoDrop One/One^{C} Microvolume UV-Vis Spectrophotometer. DNA sequencing was performed by GENEWIZ from Azenta Life Sciences for Sanger sequencing. Chemically competent *E. coli* cells were prepared using a Z-competent *E. coli* transformation buffer kit (Zymo Research). Optical density (OD_{600}) was measured using an Eppendorf BioPhotometer for cell growth.

High resolution mass spectra as well as analytical reaction analysis was performed on an Agilent Technologies 6230 TOF LC/MS equipped with an Agilent 1290 infinity LC system and Agilent Technologies 6545 Q-TOF LC/MS equipped also with an Agilent 1290 infinity LC system

at the Life Sciences Institute, University of Michigan. Column utilized for screening enzyme variants in 96-well plates was a Phenomenex Kinetex 1.7 µm Phenyl-Hexyl 100 Å LC Column 50 x 2.1 mm. Column for all other enzymatic reactions was a Phenomenex Synergi 4 µm Hydro-RP 80 Å, LC Column 100 x 2.0 mm, 4 micron, C18. LC/MS method was reversed phase, water and acetonitrile solvent system supplemented with formic acid (0.1% v/v). Preparative scale reactions were purified utilizing a Shimadzu LC 20AT prominence liquid chromatography system with a diode array detector, column was a Phenomenex Synergi 4 µm Hydro-RP 80 Å, HPLC Column 250 x 10.0 mm, 4 micron, C18. Analytical scale reactions (TTNs, reaction time courses including initial rates) were analyzed in analytical Shimadzu LC 20AD prominence liquid chromatography system with a diode array detector using 230 nm UV detection, column utilized was a Zorbax SB-Phenyl 4.6 mm x 15 cm PN 883975.912.

II.8.2 Synthetic Experimentals and Characterizations

Amide, Ester, Gem-dimethyl and Cyclohexyl Pentaketides Synthetic Procedures

The synthetic experimentals along with characterizations of these substrates will not be shown, as these were previously reported on Jennifer J. Schmidt PhD Dissertation, titled "Chemoenzymatic Assembly of Macrolactones: Merging Synthetic Chemistry with Biocatalysis to Afford Diverse Compound Libraries". University of Michigan, Ann Arbor. Rackham Graduate School. **2018**.

II.8.3 Protein Biochemistry

Expression and Purification of Pik PKS Enzymes

Cloning, expression and purification of PikAIII, PikAIV, PikAIII-TE and PikAIII-TE S148C have been described previously.[61] The desired plasmids were transformed into chemically

competent *E. coli* BAP1-pRARE cell line[124, 125], and grown overnight at 37 ºC in LB agar plates supplemented with kanamycin (50 µg/mL) and streptomycin (50 µg/mL). Single colonies were picked for overnight cultures in LB media with kanamycin (50 µg/mL) and streptomycin (50 µg/mL) for overnight incubation at 37 ºC, 180 rpm. Next, in 2.8 L baffled Fernbach flasks, 1L TB media containing TB salts (0.17 M KH_2PO_4 and 0.72 M K_2HPO_4) with kanamycin (50 µg/mL) and streptomycin (50 µg/mL) were inoculated using 5 mL of the overnight culture, and incubated at 37 ºC, 180 rpm. After the OD_{600} reached 1-1.5, the flasks containing the 1L growths were transferred to an ice bath until the internal temperature reached 18-20 ºC. These were then induced with 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated overnight at 18 ºC, 180 rpm. The following day, the mixtures were centrifuged at 6,000 x g for 30 minutes, the supernatant were discarded, and the cell pellets flash frozen in N_2 and stored in -80 ºC for future protein purification.

All subsequent steps were performed at 4 °C. Frozen cell pellets were resuspended in 5 mL of lysis buffer per grams of cells and stirred until thawed. The cells were treated with 1 mg/mL lysozyme, 0.1 mg/mL DNAse and then applied sonication (70% power / 2 mins total: 10 sec on, 50 sec off protocol). The cellular debris was pelleted by centrifugation at 50,000 x g, 30 mins, 4 ºC and the supernatant (lysate) filtered using 0.45 µm filters and added to a pre-equilibrated Ni-NTA resin column (equilibration: 10 CV wash buffer). After protein binding, the column was washed with 10 CV of wash buffer. Proteins were eluted with 5 CV of elution buffer and collected as 5 mL fractions. Protein fractions were analyzed for their absorption at 280 nm, combined and concentrated to 2.5 mL using 50 K MWCO Amicon Ultra-15 centrifugal filters and 3.5 mL of storage buffer exchanged using a pre-equilibrated PD10 desalting column (equilibration: 5 CV storage buffer). Upon completion of buffer exchange, the protein samples were analyzed once

again for their absorption at 280 nm to determine their concentration, frozen in liquid N_2 and stored at -80 °C. Samples were run on a NuPAGE 4-12% Bis-Tris protein gel to check for purity.

Buffers for protein purification:

Lysis: HEPES (50 mM), NaCl (300 mM), imidazole (10 mM), glycerol (10% v/v), pH 8.0

Wash: HEPES (50 mM), NaCl (300 mM), imidazole (30 mM), glycerol (10% v/v), pH 8.0

Elution: HEPES (50 mM), NaCl (300 mM), imidazole (300 mM), glycerol (10% v/v), pH 8.0

Storage: HEPES (50 mM), NaCl (150 mM), EDTA (1 mM), glycerol (20% v/v), pH 7.2

Pik PKS Enzymatic Reactions Analytical Scale and Preparative Scale Up

All analytical scale enzymatic reactions were performed at 50 µL. Reactions consisted of water, sodium phosphate reaction buffer (200 mM $NaH_2PO_4 \cdot H_2O$ and Na_2HPO_4, 100 mM NaCl, pH 7.2), nicotinamide adenine dinucleotide phosphate ($NADP^+$ at 0.5 mM), glucose-6-phosphate (G6P at 10 mM), glucose-6-phosphate dehydrogenase (G6PDH at 4 miliunits/uL), and allowed to incubate for 10 mins. To these were added methylmalonyl *N*-acetyl cysteamine (MeMal-NAc at 20 mM), PikAIII (3 µM), PikAIV (3 µM) and lastly pentaketide substrate (1 mM in DMSO). As negative controls, enzymes were omitted. All reactions were performed and analyzed in triplicates. The reactions were run for 2 hours, stationary, rt, then quenched by adding 3 volumes of MeOH (150 µL) and centrifuged at 13,000 x g, 10 mins, 4 °C for further LC-MS analysis. LC-MS conditions were as follows: mobile phase (A = deionized water + 0.1% formic acid, B = 95% acetonitrile/deionized water + 0.1% formic acid); 10% B for 1.0 min, 10% to 80% B over 7.0 min, 80% B for 2.0 min; 0.4 mL/min; injection volume 2 µL.

Preparative scale up enzymatic reactions (30-60 mL total volume) were performed to characterize and isolate the desired macrolactones. Reactions conditions were analogous to the analytical scale ones. Reactions consisted of water, sodium phosphate reaction buffer (200 mM $NaH_2PO_4 \bullet H_2O$ and Na_2HPO_4, 100 mM NaCl, pH 7.2), nicotinamide adenine dinucleotide phosphate ($NADP^+$ at 0.5 mM), glucose-6-phosphate (G6P at 10 mM), glucose-6-phosphate dehydrogenase (G6PDH at 4 miliunits/uL), and allowed to incubate for 10 mins. To these were added methylmalonyl *N*-acetyl cysteamine (MeMal-NAc at 20 mM), PikAIII (3 μM), PikAIV (3 μM) and lastly pentaketide substrate (~20 mg for 1 mM in DMSO). The reactions were run for 2 hours, stationary, rt. After completion, these were quenched with acetone (20-40 mL) and moved to -20 ^oC freezer for 30 mins to crash out protein. Next, these were centrifuged for 10 mins at 3,000 x g and the supernatant concentrated by rotary evaporation, extracted with EtOAc (3x) and dried over anhydrous Na_2SO_4. The crude product was concentrated by rotary evaporation, dissolved in methanol and purified by preparatory HPLC. Purification by HPLC monitoring at 210 nm and 254 nm was conducted as follows: mobile phase (A = deionized water + 0.1% formic acid, B = acetonitrile); 20% B for 2.0 min, 20% to 75% B over 38 min, 75% B for 5.0 min; 5 mL/min; each injection volume was 200 μL. HPLC fractions were collected, analyzed by mass in the LC-MS, concentrated and the products characterized by NMR affording macrolactone products.

Amide, Ester, Gem-dimethyl and Cyclohexyl Pentaketides

The analytical traces for amide, ester, gem-dimethyl and cyclohexyl pentaketides enzymatic reactions against PikAIII-TE, PikAIII-TE S148C and PikAIII/AIV will not be shown, as these were previously reported on Jennifer J. Schmidt PhD Dissertation, titled "Chemoenzymatic Assembly of Macrolactones: Merging Synthetic Chemistry with Biocatalysis to Afford Diverse Compound Libraries". University of Michigan, Ann Arbor. Rackham Graduate School. **2018**.

Truncated Pentaketides

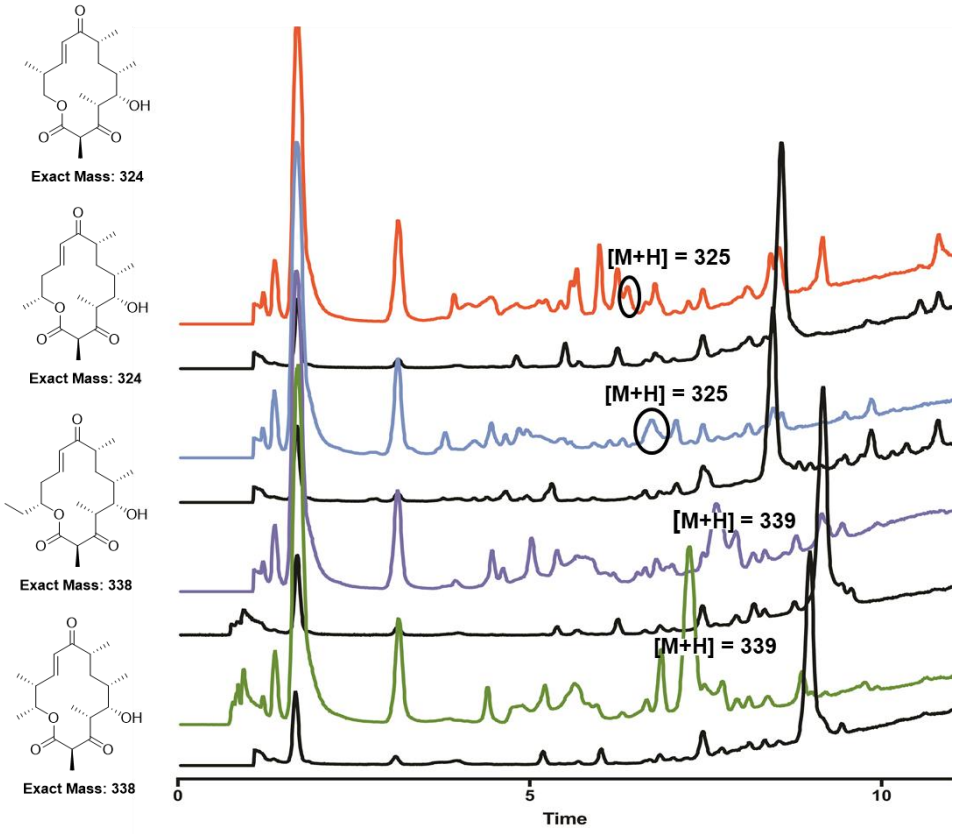


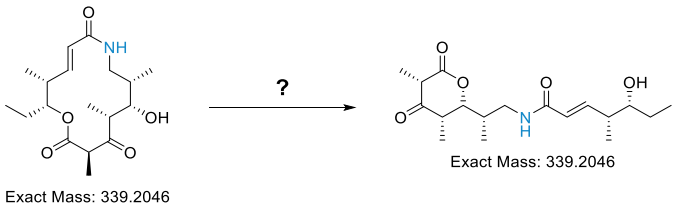
Figure II.15 Total ion chromatogram (TIC) for the truncated pentaketides reactions against PikAIII/AIV resulting in the formation of the desired 14-membered macrolactone rings. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 14 minutes (min) total run.

Amide and Ester 6-membered Ring Formation

The reactions shown below followed the same procedures as for "Pik PKS Enzymatic Reactions Analytical Scale" experimental section. The reactions times varied to 1 hour, 2 hours and overnight (24 hours). Expression and purification of PikTE WT and PikTE S148C has been reported previously.[62, 65]

Additional Data on Formation of Amide/Ester 6-membered Rings

Table II.1 Tests with different reaction conditions and thiols for investigation of spontaneous amide macrolactone ring opening.



Test	Description	Result
Buffer	14 amide ring	No 6 ring
SNAc: 	14 amide ring	No 6 ring
Thiophenol: 	14 amide ring	No 6 ring
Recycling System	14 amide ring	No 6 ring
Recycling System + MeMal-SNAc	14 amide ring	No 6 ring
2-vinyl pyridine: 	To enzymatic reaction (thiol scavenger)	6 ring forms

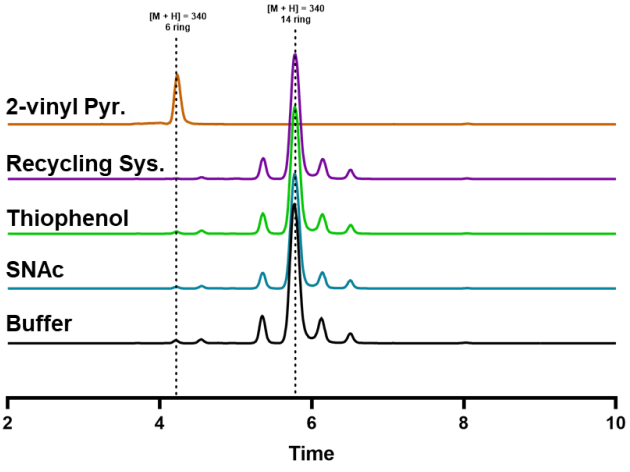


Figure II.16 Total ion chromatogram (TIC) resulting from tests with different reaction components/byproducts and thiols for the investigation of possibility of spontaneous amide

macrolactone ring opening. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run.

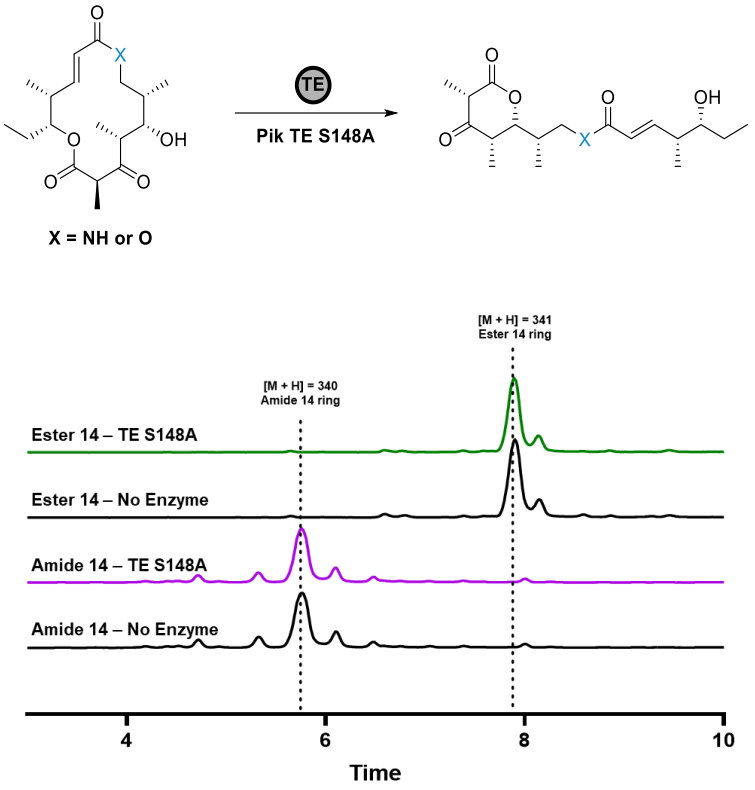


Figure II.17 Total ion chromatogram (TIC) resulting from reactions with amide and ester 14-membered ring macrolactones against Pik TE S148A. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run.

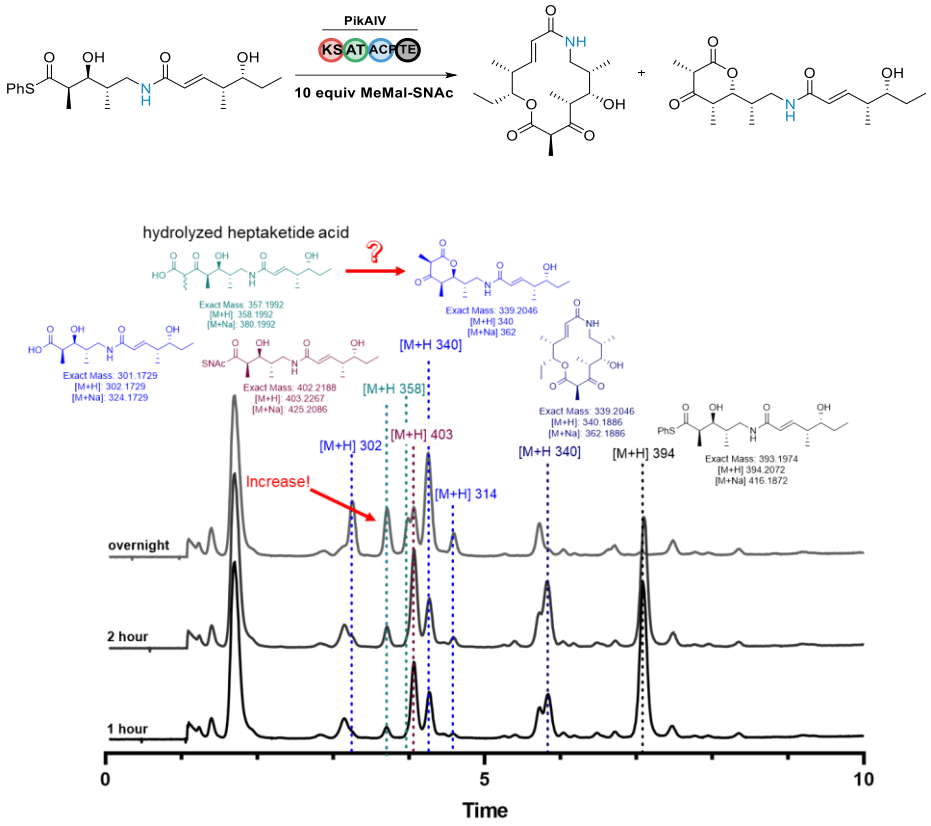


Figure II.18 Total ion chromatogram (TIC) resulting from reactions with amide hexaketide reaction against PikAIV to test ring opening hypothesis and implications of the formation of the heptaketide acid intermediate. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. For synthetic procedure on amide hexaketide substrate, see Chapter 3 experimentals.

Biotransformation with *Streptomyces venezuelae* Strain DHS316

The following *in-vivo* biotransformation reaction was carried out in an analogous manner to the published procedures with minor modifications.[120, 126] In 250 mL baffled Erlenmeyer flasks, 100 mL of SCM media were inoculated with 20 uL of *Streptomyces venezuelae* strain DHS316 spore stock and grown overnight. The next day, the OD_{600} was measured, followed by centrifugation of the cells for 10 mins at 3,000 x g. The resulting supernatant was removed until

~5 mL of slurry remained. Using the initial OD_{600}, the culture was used to inoculate another 100 mL biotransformation culture of SCM media to an adjusted OD_{600} = 0.1. The culture(s) were incubated for 1 hour at 28 ºC, 180 rpm. Then, ~10 mg of the desired narbonolide derivative (as a 50 mg/mL in DMSO solution) were added as well as 0.25 mg of acetyl-narbonolide activator (as a 20 mg/mL in DMSO solution). These were incubated for 18 hours, at 28 ºC, 180 rpm. After reaction completion (checked by LC-MS for product appearance), the culture(s) were centrifuged for 10 mins at 3,000 x g and the supernatant poured off and pH adjusted to 11 using 5 M NaOH. The solution was extracted with 3 x 50 mL of ethyl acetate, and the combined organic layers were dried over anhydrous Na_2SO_4. Solvent was removed under reduced pressure to yield crude product mixture. The reactions were purified by HPLC using a Synergi Hydro-RP C18 column, 250 x 10.0 mm, 4 micron, with 3 mL/min flow rate. A gradient of solvent A (Water + 0.1% NEt_3) and solvent B (Acetonitrile) was as follows: isocratic 50% B for 2 minutes, then a linear gradient from 50% - 95% B over 38 minutes, then isocratic 95% for 5 minutes.

Expression and Purification P450 PikC D50N-RhFRED and JuvD-RhFRED

Plasmid pET28b_PikC D50N-RhFRED previously generated[122] was transformed into *E. coli* BL21(DE3) cells. Single colonies were picked for overnight cultures in LB media with kanamycin (50 μg/mL) for overnight incubation at 37 ºC, 180 rpm. Next, in 2.8 L baffled Fernbach flasks, 1 L LB media + 0.4% glycerol containing kanamycin (50 μg/mL), thiamine (1 mM), and $FeCl_3$ (100 μM) were each inoculated with 15 mL of overnight seed culture and incubated at 37 ºC, 160 rpm. After the OD_{600} reached 0.6-1.0, the flasks containing the 1L growths were transferred to an ice bath until the internal temperature reached 18-20 ºC. These were then induced with 0.1 mM IPTG and 1.0 mM δ-aminolevulinic acid and incubated overnight at 18 ºC, 180 rpm. The following day,

the mixture was centrifuged at 6,000 x g for 30 minutes, the supernatant was discarded, and the cell pellet flash frozen and stored in -80 ºC for future protein purification.

All subsequent steps were performed at 4 °C. Frozen cell pellets were resuspended in 5 mL of lysis buffer per grams of cells and stirred until thawed. The cells were treated with 1 mg/mL lysozyme, 0.1 mg/mL DNAse and then applied sonication (70% power / 5 mins total: 5 sec on, 5 sec off protocol). The crude lysate was centrifuged at 50,000 x g, 30 mins, and the clarified lysate was filtered with 45 µm filters and added to a pre-equilibrated Ni-NTA resin column (equilibration: 10 CV wash buffer). After protein binding, the column was washed with 10 CV of wash buffer. Proteins were eluted with 5 CV of elution buffer and collected as 1.5 mL fractions. These were combined and concentrated to 2.5 mL using 10 K MWCO Amicon Ultra-15 centrifugal filters and 3.5 mL of storage buffer exchanged using a pre-equilibrated PD10 desalting column (equilibration: 5 CV storage buffer). Upon completion of buffer exchange, the protein samples were flash frozen in liquid N_2 stored at -80 ºC. The yield of functional P450 (as assessed by obtaining CO difference spectra according to the established protocol)[127] was ~8 mg per 1 L of overexpression culture.

Buffers for protein purification:

Lysis: Sodium phosphate (50 mM), NaCl (300 mM), imidazole (10 mM), glycerol (10% v/v), pH 8.0

Wash: Sodium phosphate (50 mM), NaCl (300 mM), imidazole (20 mM), glycerol (10% v/v), pH 8.0

Elution: Sodium phosphate (50 mM), NaCl (300 mM), imidazole (300 mM), glycerol (10% v/v), pH 8.0

Storage: Sodium phosphate (50 mM), EDTA (1 mM), DTT (0.2 mM), glycerol (10% v/v), pH 7.3

PikC D50N-RhFRED Analytical and Preparative Scale Up Enzymatic Reactions

Reactions with catalytically self-sufficient PikC D50N-RhFRED fusion enzymes were carried out as previously described.[120] For analytical scale 50 µL enzymatic reactions, the components were water, buffer (50 mM NaH_2PO_4, 1mM EDTA, 0.2 mM DTT, 10% v/v glycerol, pH 7.3), nicotinamide adenine dinucleotide phosphate ($NADP^+$ at 1 mM), glucose-6-phosphate (G6P at 5 mM), glucose-6-phosphate dehydrogenase (G6PDH at 1 miliunits/uL), and allowed to incubate for 10 mins. To this was added PikC D50N-RhFRED (5 µM) and lastly narbomycin derivative (0.5 mM in DMSO). As negative controls, enzymes were omitted. All reactions were performed and analyzed in triplicates. The reactions were run overnight, 30 ºC at 180 rpm, quenched by adding 3 volumes of MeOH (150 µL) and centrifuged at 13,000 x g, 10 mins, 4 ºC for further LC-MS analysis. LC-MS conditions were as follows: mobile phase (A = deionized water + 0.1% formic acid, B = 95% acetonitrile/deionized water + 0.1% formic acid); 10% B for 1.0 min, 10% to 80% B over 7.0 min, 80% B for 2.0 min; 0.4 mL/min; injection volume 2 µL.

The scale up reactions between PikC D50N-RhFRED and narbomycin derivatives (~5-10 mg for 0.5 mM in DMSO) were performed in 125 mL Erlenmeyer flasks (20 mL total reaction volume or more). These reactions were run analogously to the analytical PikC D50N-RhFRED reactions. Conversion after 12 hours as assessed by LC-MS. After completion, these were quenched with acetone and moved to -20 ºC freezer for 30 mins to crash out protein. Next, these were centrifuged for 10 mins at 3,000 x g and the supernatant concentrated by rotary evaporation, pH adjusted to 11, extracted with EtOAc (3x) and dried over anhydrous Na_2SO_4. The crude product was concentrated by rotary evaporation, dissolved in methanol and purified by preparatory HPLC. The reactions were purified by HPLC using a Synergi HydroRP C18 column, 250 x 10.0 mm, 4 micron,

with 3 mL/min flow rate. A gradient of solvent A (Water + 0.1% NEt_3) and solvent B (Acetonitrile) was as follows: isocratic 40% B for 2 minutes, then a linear gradient from 40% – 95% B over 38 minutes, then isocratic 95% for 5 minutes.

*JuvD-RhFRED followed analogous protocols to PikC D50N-RhFRED described herein.

II.9 NMR Spectra

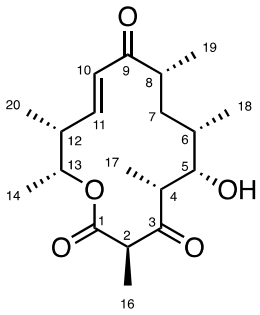


Table II.2 14-desmethyl-narbonolide (**2.28**) major isomer NMR Data C_6D_6

Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		169.8		
2	3.29 (q, J = 7.0)	50.5	16	1, 3, 16
3		208.0		
4	2.49 (qd, J = 5.3, 7.6)	49.7	5, 17	3, 5, 6, 17
5	3.89 (bt, J = 4.1)	72.7	4, 6	6
5-OH				
6	1.73 - 1.79 (m)	35.8	5, 7a, 7b, 18	5
7a	1.61 (ddd, J = 5.2, 8.1, 13.7)	36.8	6, 7b, 8	5, 6, 8, 9, 19
7b	1.38 (ddd, 5.9, 7.5, 13.8)		6, 7a, 8	5, 6, 8, 9, 19
8	2.83 (dq, J = 6.7, 13.1)	40.9	7a, 7b, 19	7, 9, 19
9		202.8		
10	6.01 (dd, J = 1.6, 16.1)	129.6	11, 12	9, 12
11	6.65 (dd, J = 6.1, 16.1)	146.4	10, 12	9, 12, 13, 20
12	2.09 (dddq, J = 1.5, 3.5, 6.9, 13.9)	39.5	10, 11, 13, 20	
13	4.93 (dq, J = 3.2, 6.6)	73.1	12, 14	1, 11
14	0.83 (d, J = 6.54)	15.5	13	12, 13
15				
16	1.25 (d, J = 7.0)	14.4	2	1, 2, 3
17	0.97 (d, J = 7.3)	11.6	4	3, 4, 5
18	0.88 (d, J = 7.2)	18.1	6	5, 6, 7
19	1.09 (d, J = 6.7)	17.6	8	7, 8, 9
20	0.65 (d, J = 7.0)	12.1	12	11, 12, 13

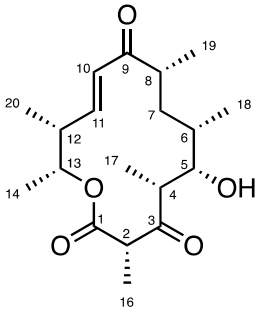
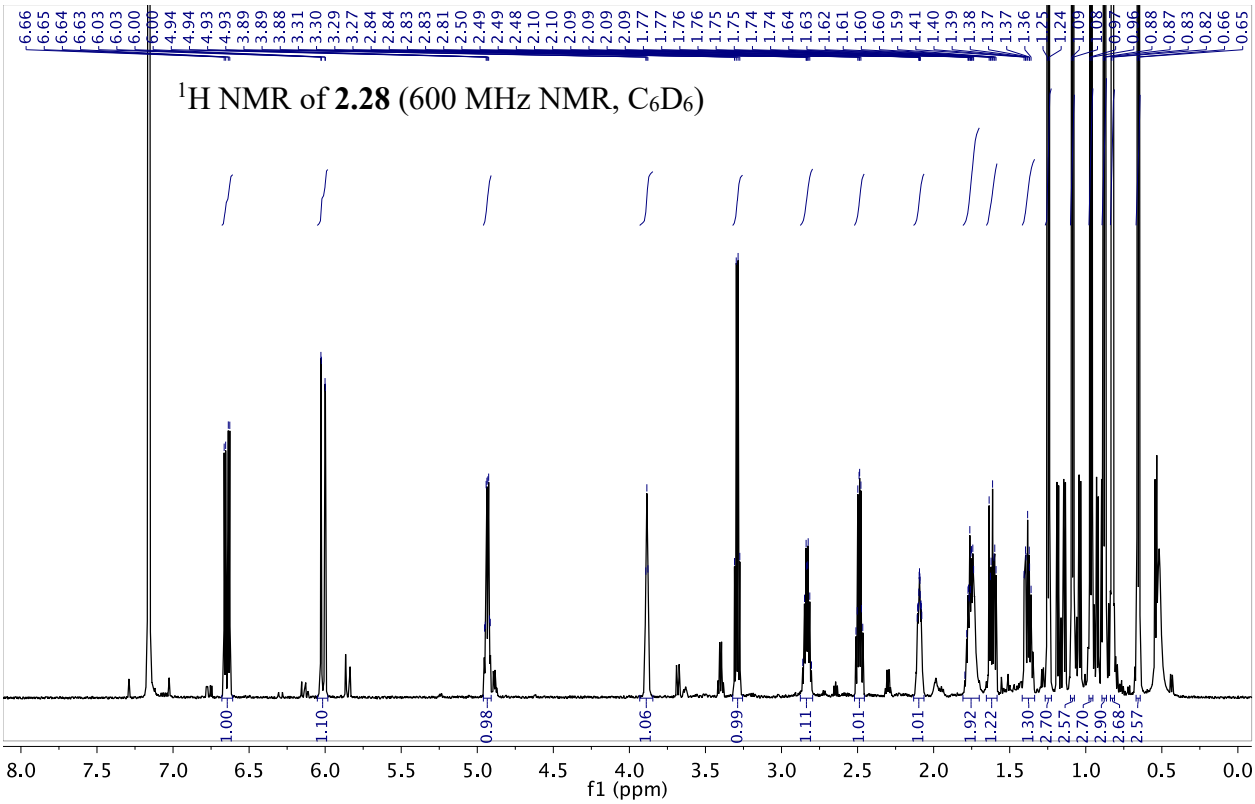
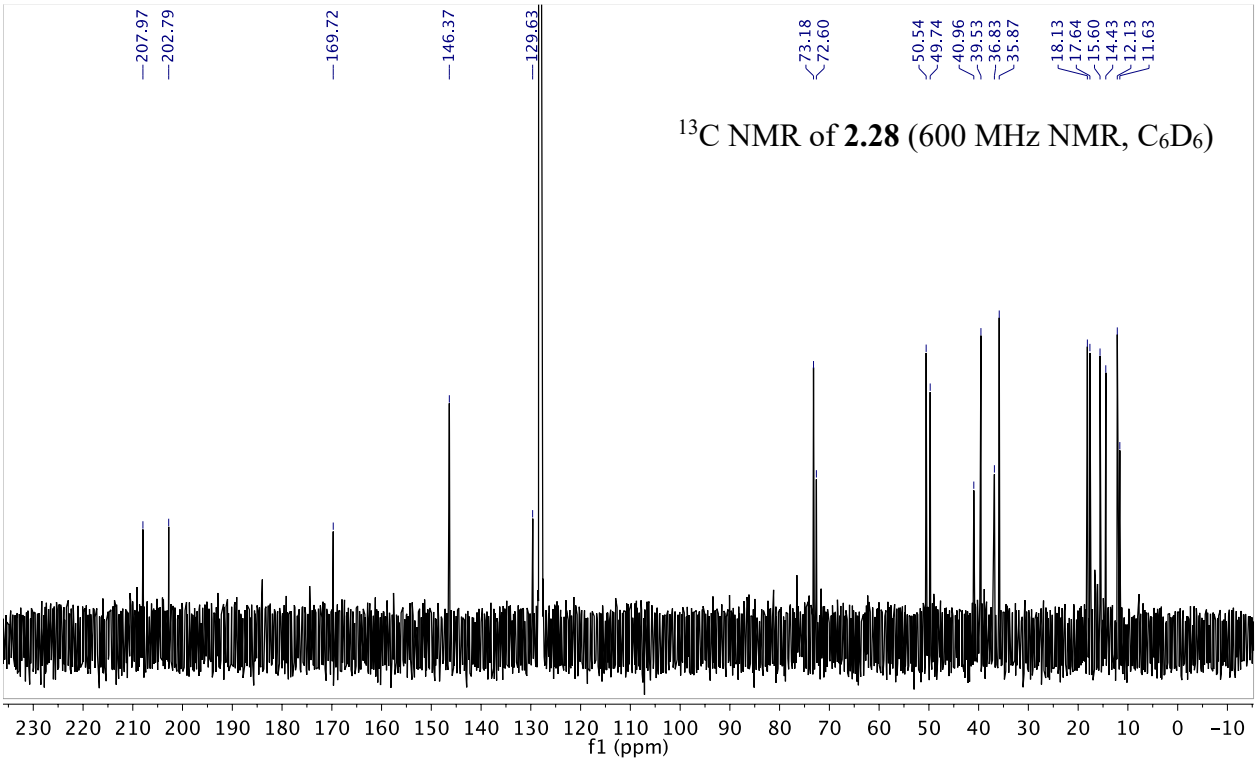


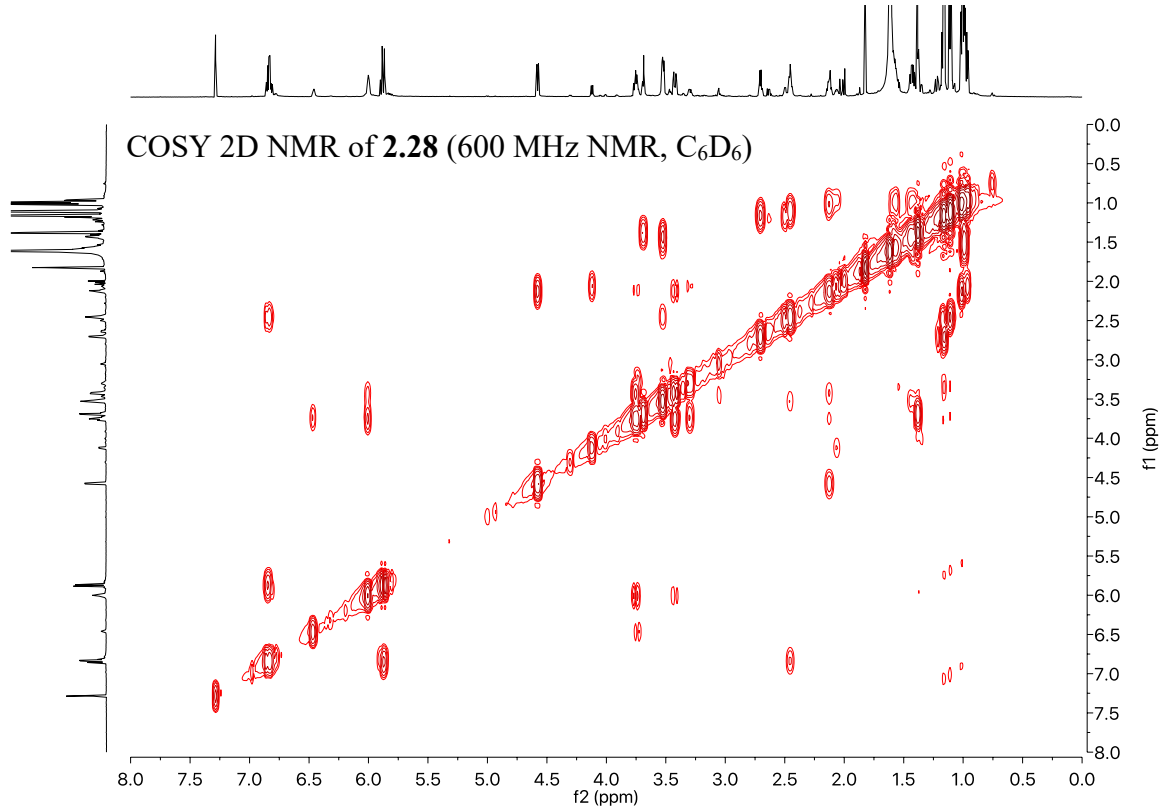
Table II.3 14-desmethyl-narbonolide (**2.28**) minor isomer NMR Data

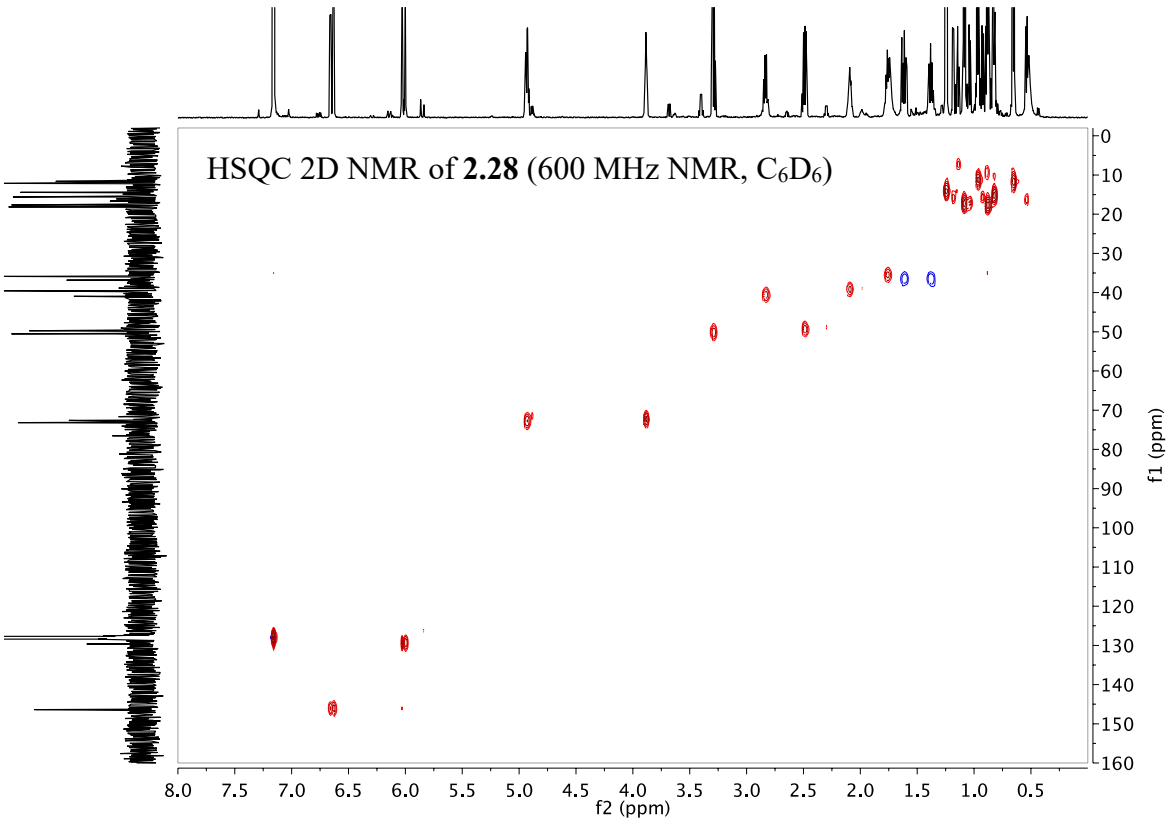
C_6D_6

Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		169.2		
2	3.38 (q, J = 7.2)	51.1	16	1, 3, 16
3		206.6		
4	2.63 (p, J = 6.8)	50.1	5, 17	3, 17
5	3.61 (bt, J = 6.7)	74.6	4, 6	7a, 17, 18
5-OH				
6	1.50 - 1.54 (m)	35.5	5, 7a, 7b, 18	
7a	1.72 (ddd, J = 2.7, 10.8, 13.5)	35.1	6, 7b, 8	5, 8, 18
7b	1.47 (ddd, J = 4.3, 9.3, 13.8)		6, 7a, 8	18
8	2.81 - 2.89 (m)	41.3	7a, 7b, 19	
9		203.1		
10	6.14 (dd, J = 1.7, 16.0)	128.6	11	9, 12
11	6.76 (dd, J = 5.7, 16.0)	147.2	10, 12	9, 12, 20
12	1.91 - 1.97 (m)	40.3	11, 13, 20	11, 13
13	4.92 (qd, J = 2.7, 6.7)	72.6	12, 14	11
14	0.84 (d, J = 6.5)	16.4	13	12, 13
15				
16	1.15 (d, J = 7.2)	14.7	2	1, 2, 3
17	0.94 (d, J = 7.0)	11.5	4, 6	3, 4, 5
18	0.86 (d, J = 6.5)	17.3	6	5, 6
19	1.06 (d, J = 6.8)	18.5	8	7, 8, 9
20	0.81 (d, J = 6.9)	10.7	12	11, 12, 13


^{1}H NMR of **2.28** (600 MHz NMR, C_6D_6)
f1 (ppm)


^{13}C NMR of **2.28** (600 MHz NMR, C_6D_6)
f1 (ppm)


COSY 2D NMR of **2.28** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)


HSQC 2D NMR of **2.28** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)

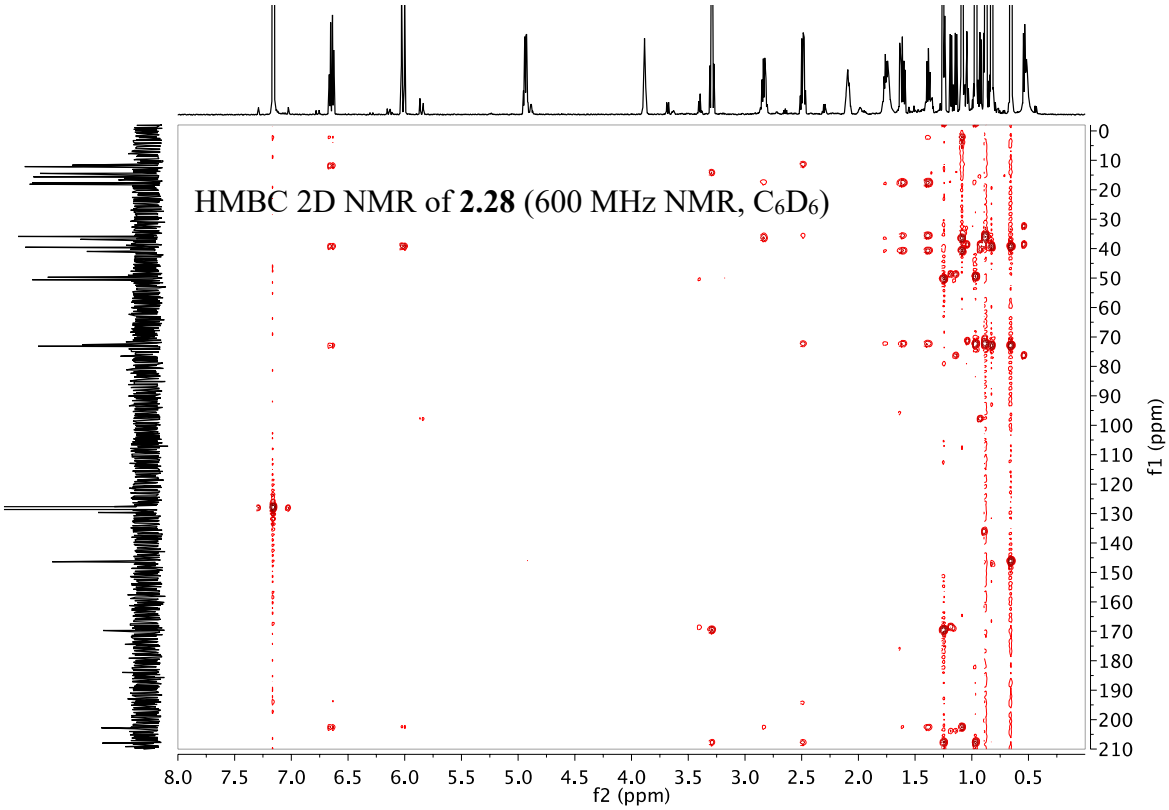
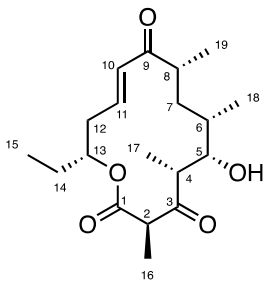

HMBC 2D NMR of **2.28** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)

Table II.4 12-desmethyl-narbonolide (**2.29**) major isomer NMR Data C_6D_6



Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		168.9		
2	3.53 (q, J = 7.0)	48.5	16	1, 3, 16
3		203.7		
4	2.23 (dq, J = 2.2, 6.4)	49.5	5, 17	3
5	3.76 (dd, J = 2.3, 10.2)	76.6	4, 6	17
5-OH				
6	1.39 - 1.47 (m)	32.6	5, 18	7
7a	1.31 - 1.39 (m)	38.7	7b	8
7b	0.90 - 0.99 (m)		7a, 8	6, 8, 18, 19
8	1.60 - 1.69 (m)	40.3	7b, 19	9, 10
9		98.1		
10	5.84 (d, J = 15.4)	131.0	11	9, 11, 12
11	5.90 (ddd, J = 3.7, 10.1, 15.7)	129.3	10, 12a, 12b	9
12a	2.06 (dt, J = 3.7, 12.7)	38.6	11, 12b, 13	11
12b	1.85 - 1.90 (m)		11, 12a, 13	11
13	4.72 - 4.79 (m)	73.2	12a, 12b, 14a, 14b	
14a	1.43 - 1.47 (m)	28.1	13, 14b, 15	
14b	1.26 - 1.30 (m)		12, 14a, 15	
15	0.76 (t, J = 7.4)	10.0	14a, 14b	13, 14
16	1.27 (d, J = 6.8)	17.3	2	1, 2, 3
17	1.15 (d, J = 6.9)	7.5	4	3, 4, 5
18	0.52 (d, J = 6.6)	16.5	6	5, 6, 7
19	0.90 (d, J = 6.8)	16.2	8	9
20				

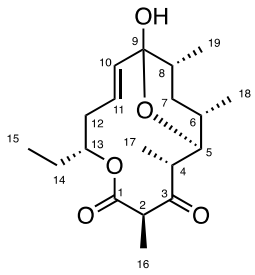
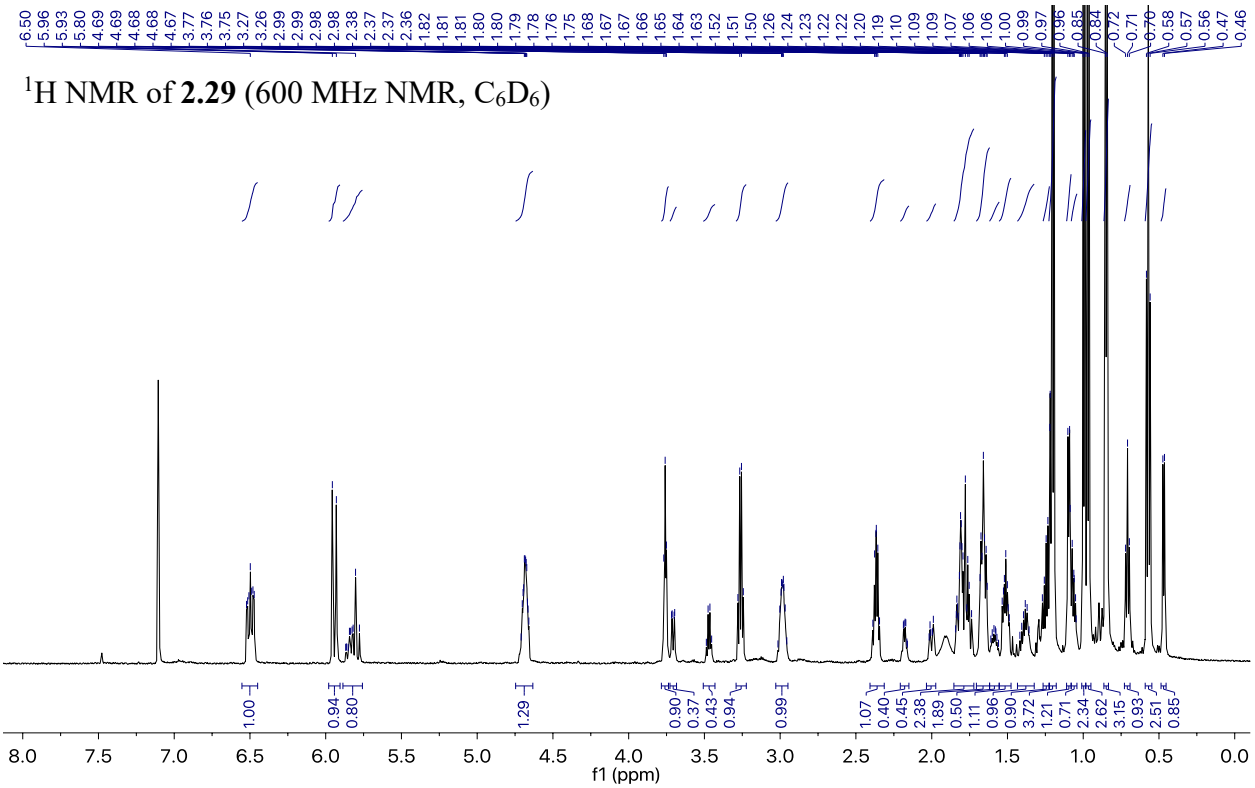
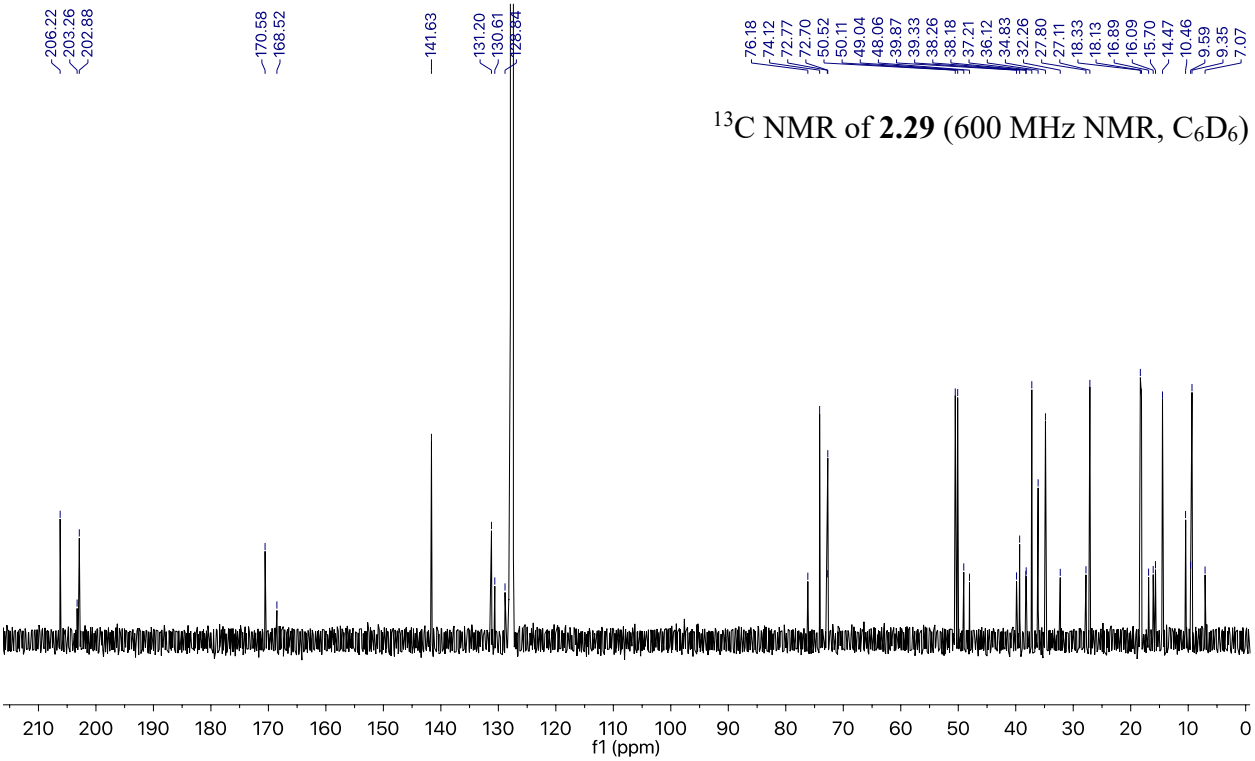


Table II.5 12-desmethyl-narbonolide (**2.29**) minor isomer NMR Data C_6D_6

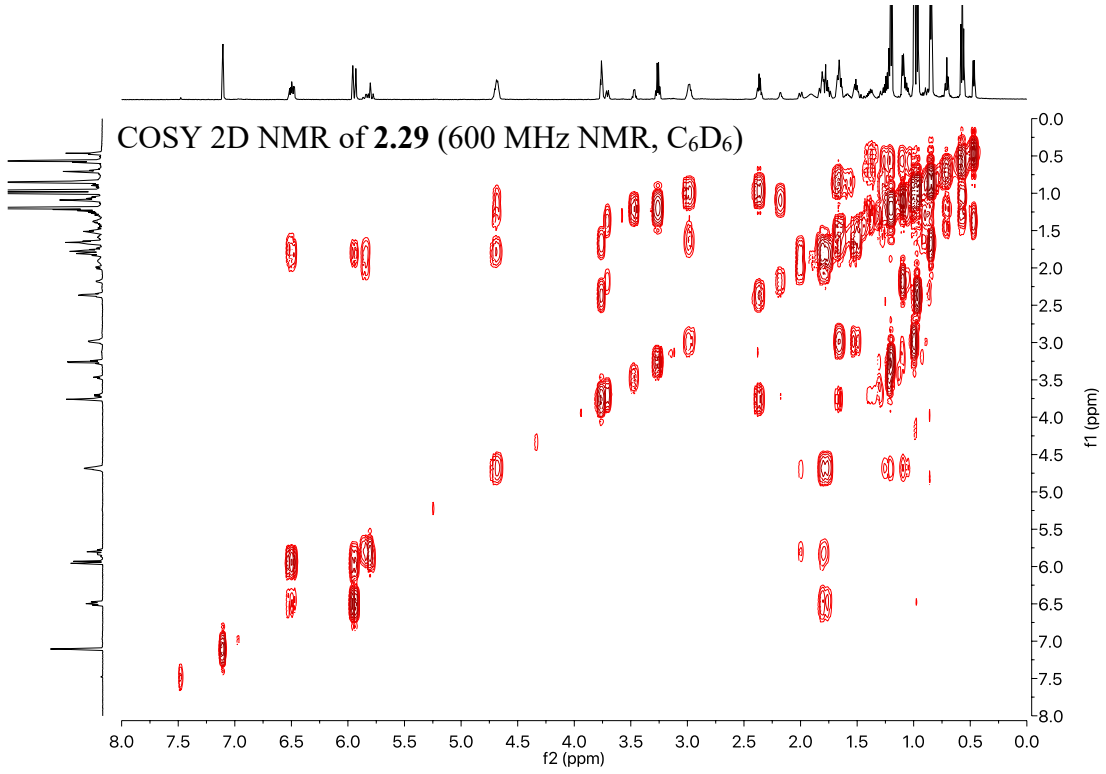
Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		168.9		
2	3.53 (q, J = 7.0)	48.5	16	1, 3, 16
3		203.7		
4	2.23 (qd, J = 6.4, 2.2)	49.5	5, 17	3
5	3.76 (dd, J = 10.2, 2.3)	76.6	4, 6	17
5-OH				
6	1.39 - 1.47 (m)	32.6	5, 18	7
7a	1.31 - 1.39 (m)	38.7	7b	8
7b	0.90 - 0.99 (m)		7a, 8	6, 8, 18, 19
8	1.60 - 1.69 (m)	40.3	7b, 19	9, 10
9		98.1		
10	5.84 (d, J = 15.4)	131.0	11	9, 11, 12
11	5.90 (ddd, J = 15.7, 10.1, 3.7)	129.3	10, 12a, 12b	9
12a	2.06 (dt, J = 12.7, 3.7)	38.6	11, 12b, 13	11
12b	1.85 - 1.90 (m)		11, 12a, 13	11
13	4.72 - 4.79 (m)	73.2	12a, 12b, 14a, 14b	
14a	1.43 - 1.47 (m)	28.1	13, 14b, 15	
14b	1.26 - 1.30 (m)		12, 14a, 15	
15	0.76 (t, J = 7.4)	10.0	14a, 14b	13, 14
16	1.27 (d, J = 6.4)	17.3	2	1, 2, 3
17	1.15 (d, J = 6.9)	7.5	4	3, 4, 5
18	0.52 (d, J = 6.6)	16.5	6	5, 6, 7
19	0.90 (d, J = 6.8)	16.2	8	9
20				

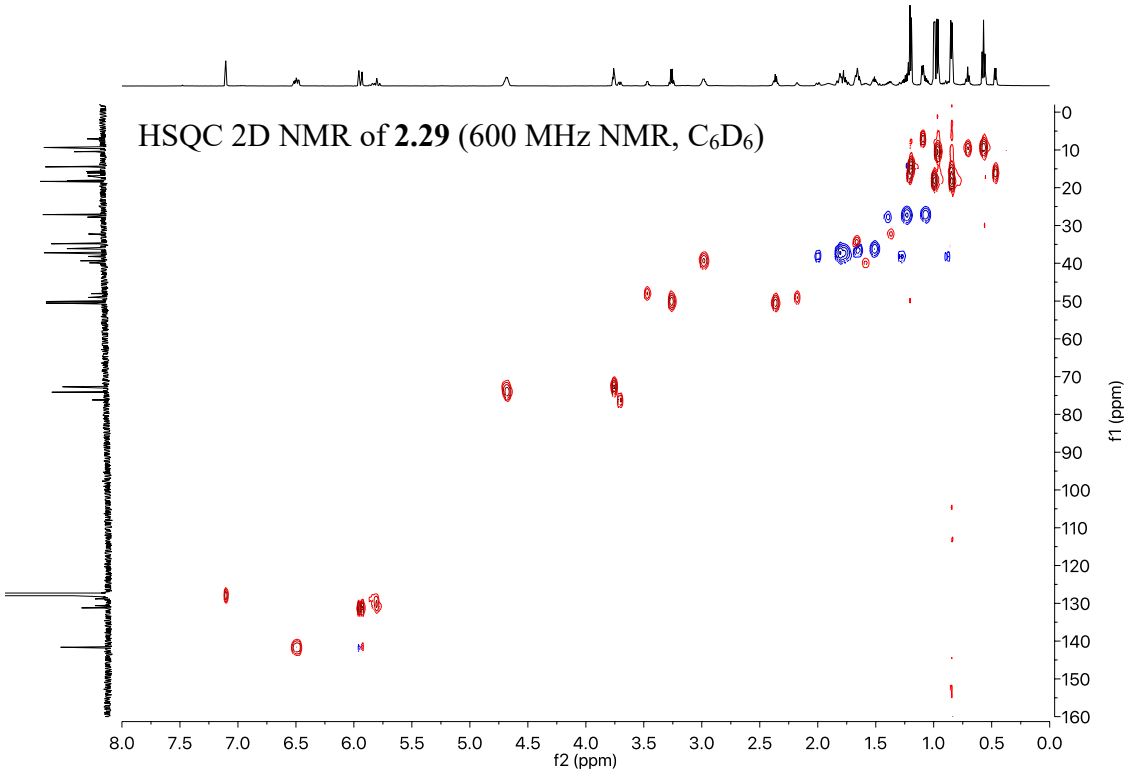


^{1}H NMR of **2.29** (600 MHz NMR, C_6D_6)



^{13}C NMR of **2.29** (600 MHz NMR, C_6D_6)


COSY 2D NMR of **2.29** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)


HSQC 2D NMR of **2.29** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)

HMBC 2D NMR of **2.29** (600 MHz NMR, C_6D_6)

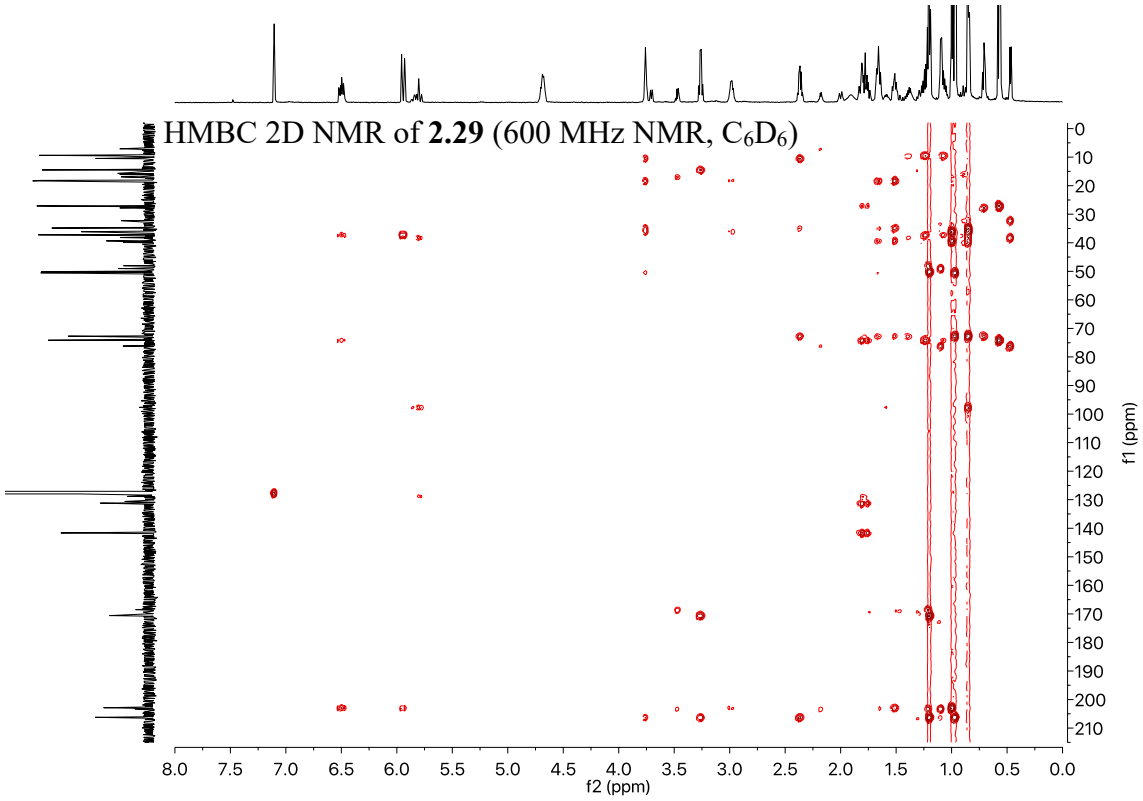
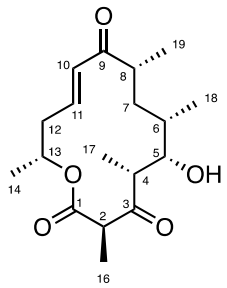


Table II.6 12,14-didesmethyl-narbonolide (**2.30**) major isomer NMR Data C_6D_6



Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		170.1		
2	3.29 (q, J = 7.0)	50.6	16	1, 3 16
3		207.2		
4	2.46 (p, J = 6.9)	50.5	5, 17	3, 5, 6, 17
5	3.78 (bt, J = 5.9)	73.3	4, 6	17
5-OH				
6	1.64 - 1.72 (m)	35.3	5, 7a, 7b, 18	
7a	1.61 (ddd, J = 14.0, 9.5, 4.1)	36.1	6, 8	5, 6, 8, 9, 18, 19
7b	1.48 (ddd, J = 14.0, 8.6, 5.3)		6, 8	5, 6, 8, 9, 18, 19
8	2.92 (dt, J = 15.5, 7.0)	40.8	7a, 7b, 19	7, 9, 19
9		202.6		
10	5.99 (d, J = 15.7)	130.8	11	9, 12
11	6.52 (dt, J = 15.2, 7.2)	141.6	10, 12	9, 12, 13
12	1.77 (t, J = 8.0)	39.2	11, 13	10, 11, 13, 14
13	4.78 (h, J = 6.2)	69.9	12, 14	1, 11, 12, 14
14	0.87 (d, J = 6.3)	20.0	13	12, 13
15				
16	1.24 (d, J = 6.9)	15.0	2	1, 2, 3
17	0.98 (d, J = 7.2)	11.5	4	3, 4, 5
18	0.92 (d, J = 6.9)	18.4	6	5, 6, 7
19	1.03 (d, J = 6.7)	18.1	8	7, 8, 9
20				

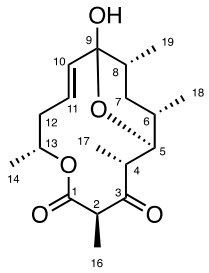
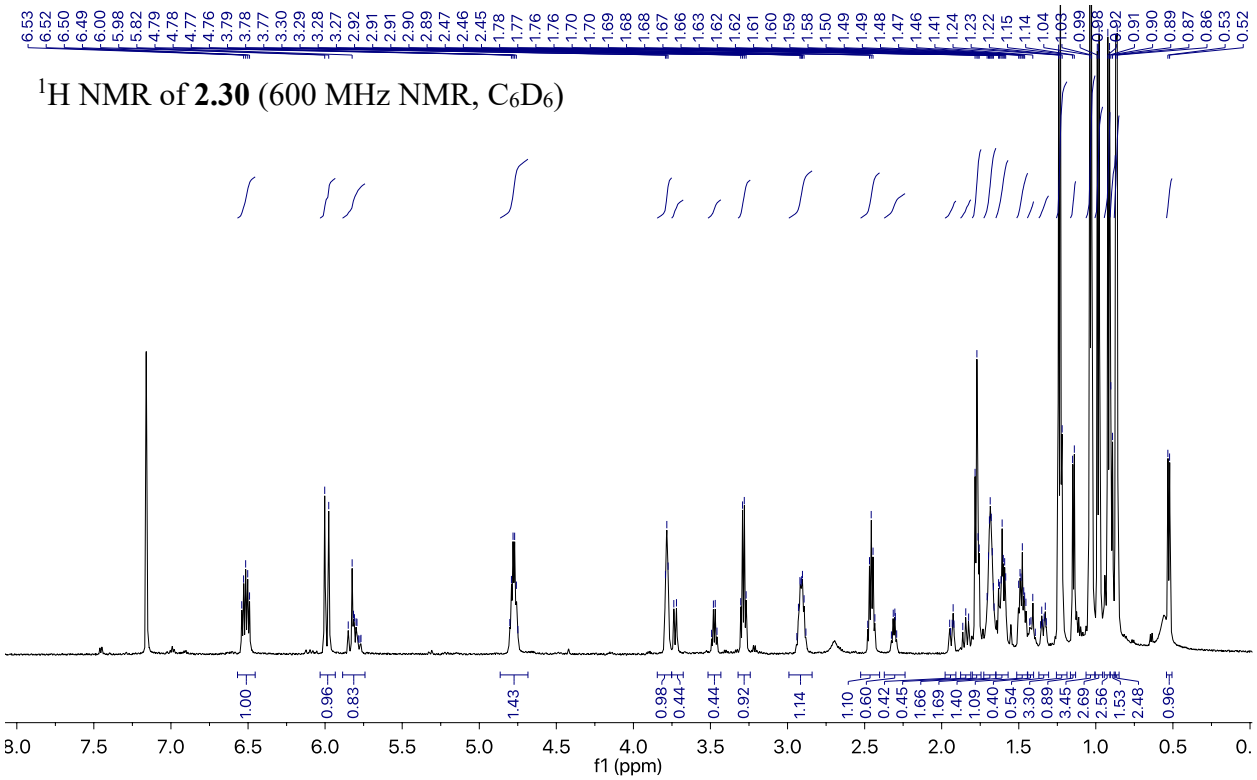
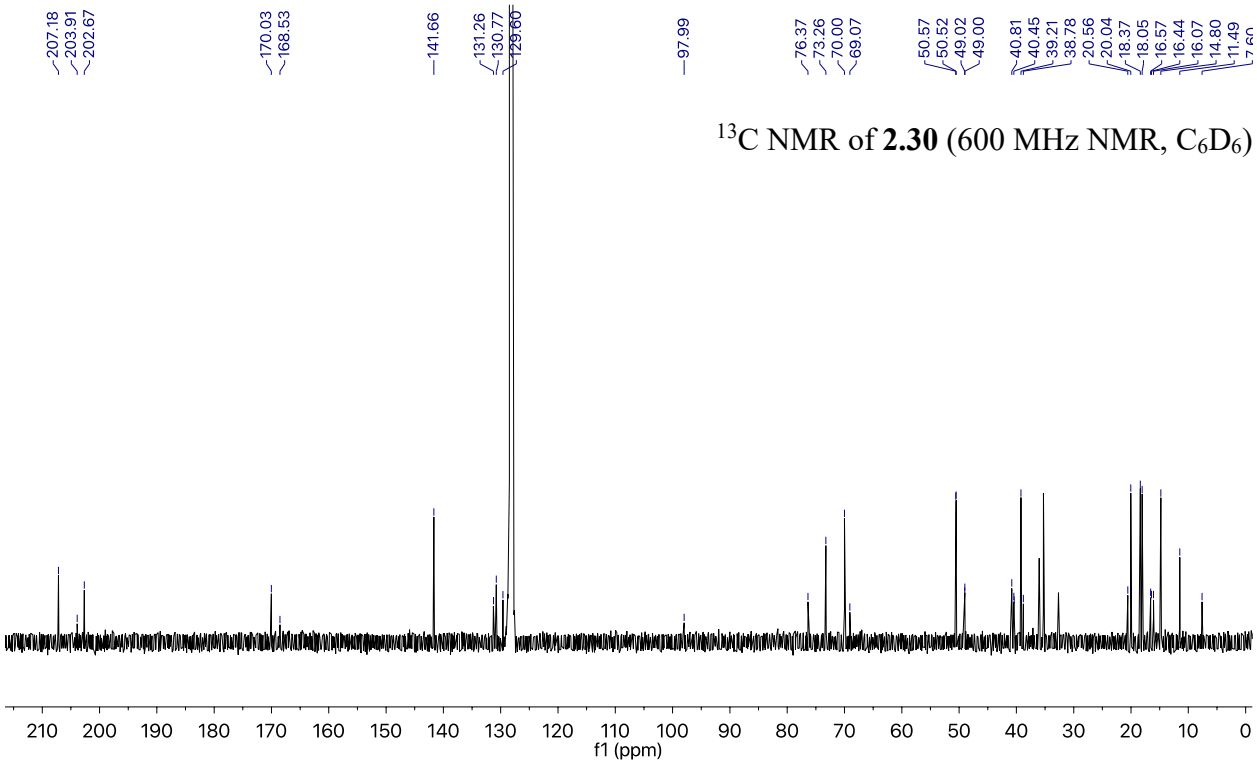


Table II.7 12,14-didesmethyl-narbonolide (**2.30**) minor isomer NMR Data C_6D_6

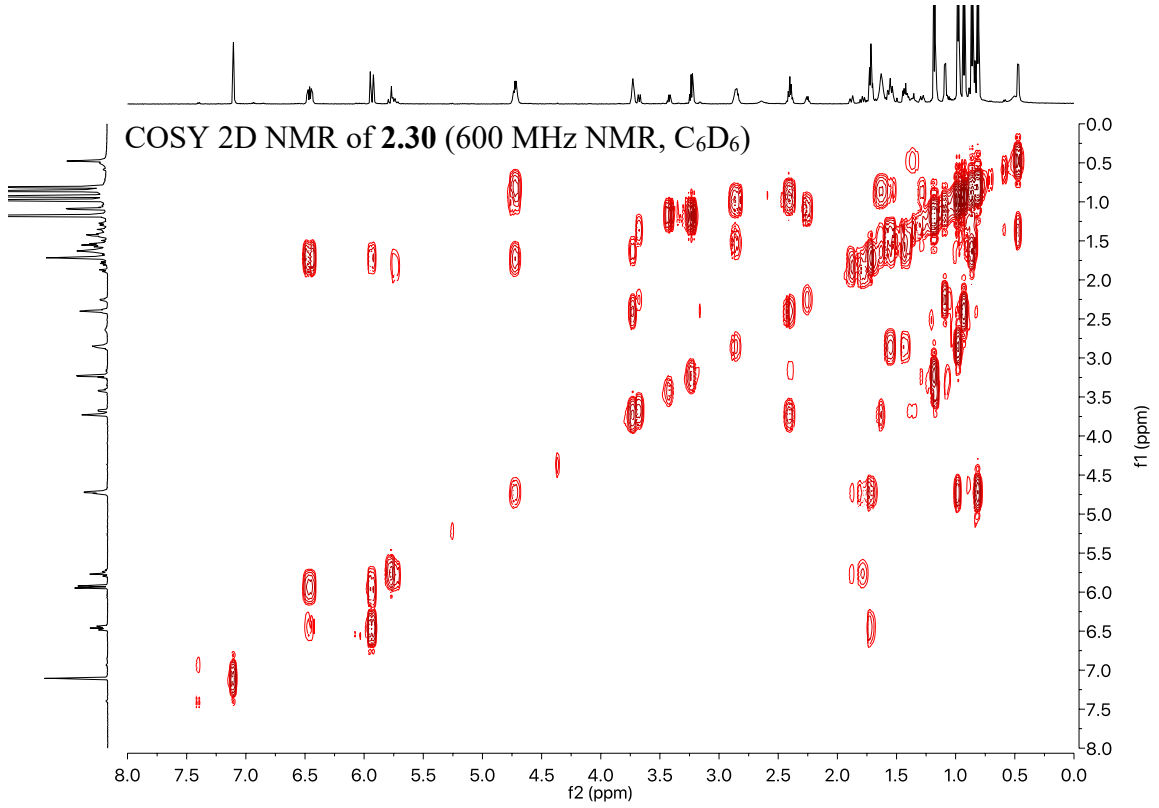
Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		168.5		
2	3.48 (q, J = 6.9)	49.1	16	1, 3, 16
3		203.9		
4	2.31 (q, J = 7.1)	49.1	5, 17	3, 5, 17
5	3.73 (d, J = 10.1)	76.4	4, 6	17
5-OH				
6	1.45 - 1.38 (m)	32.6	5, 7a, 18	
7a	1.34 (dt, J = 13.1, 4.0)	38.8	6, 7b	
7b	0.94 - 0.89 (m)		7a	
8	1.63 - 1.58 (m)	40.4		9
9		98.0		
10	5.84 (d, J = 15.4)	131.3	11	9, 11, 12
11	5.83 - 5.74 (m)	129.6	11, 12a, 12b	9, 10, 12
12a	1.94 (dt, J = 13.0, 3.5)	40.4	11, 13	
12b	1.89 - 1.81 (m)		11, 13	10, 11, 13
13	4.78 (h, J = 6.2)	69.1	12a, 12b, 14	
14	0.98 (d, J = 6.5)	20.6	13	12, 13
15				
16	1.23 (d, J = 7.1)	16.4	2	1, 2, 3
17	1.15 (d, J = 6.7)	7.6	4	3, 4, 5
18	0.53 (d, J = 6.6)	16.6	6	5, 6, 7
19	0.90 (d, J = 7.1)	16.0	8	7, 8, 9
20				

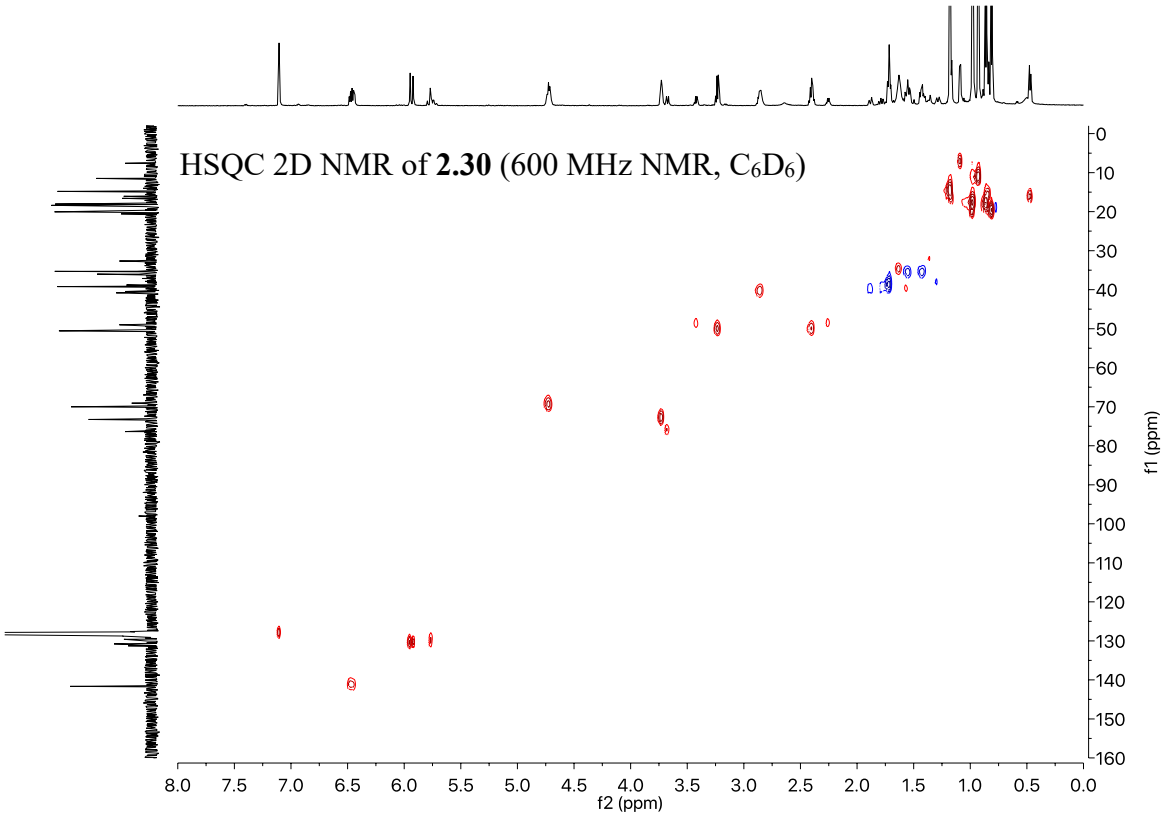


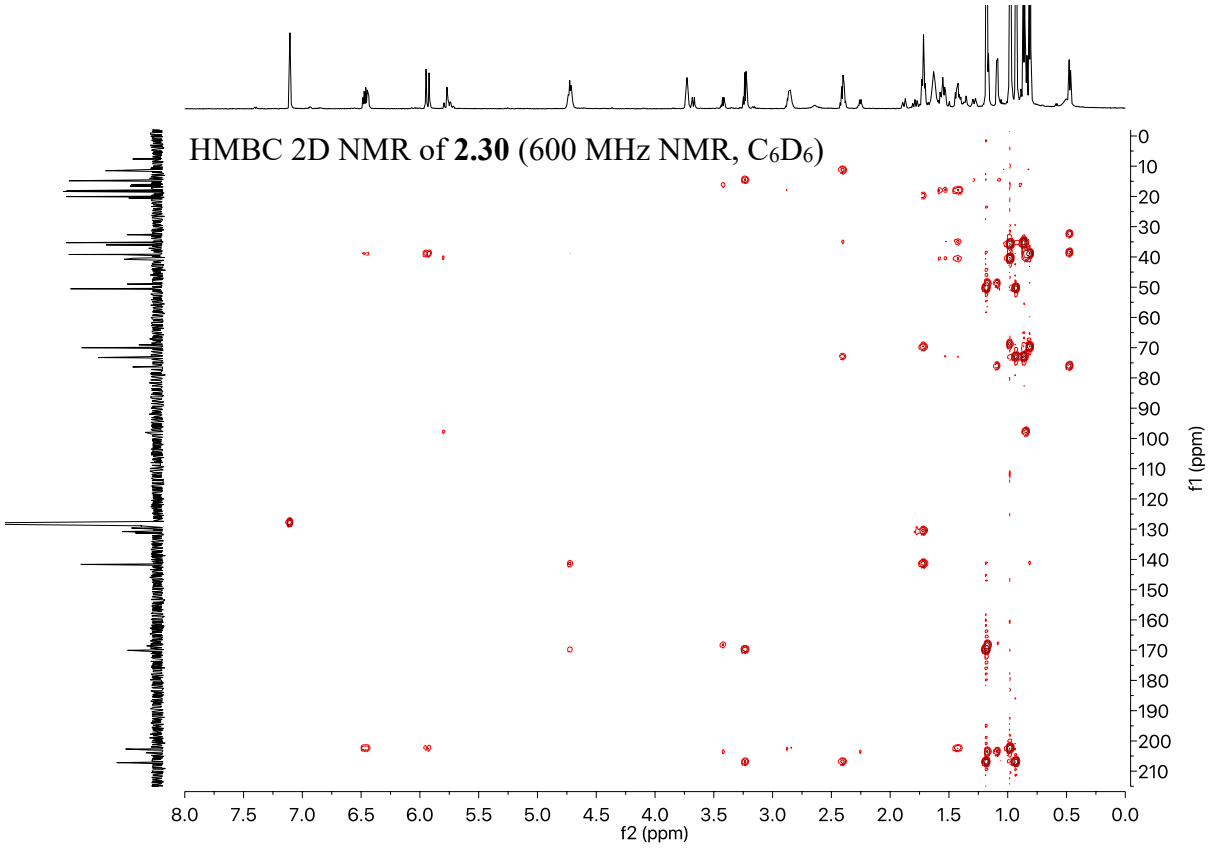
^{1}H NMR of **2.30** (600 MHz NMR, C_6D_6)



^{13}C NMR of **2.30** (600 MHz NMR, C_6D_6)


COSY 2D NMR of **2.30** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)


HSQC 2D NMR of **2.30** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)


HMBC 2D NMR of **2.30** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)

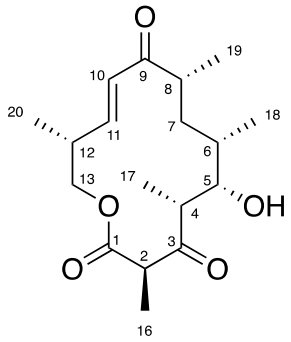
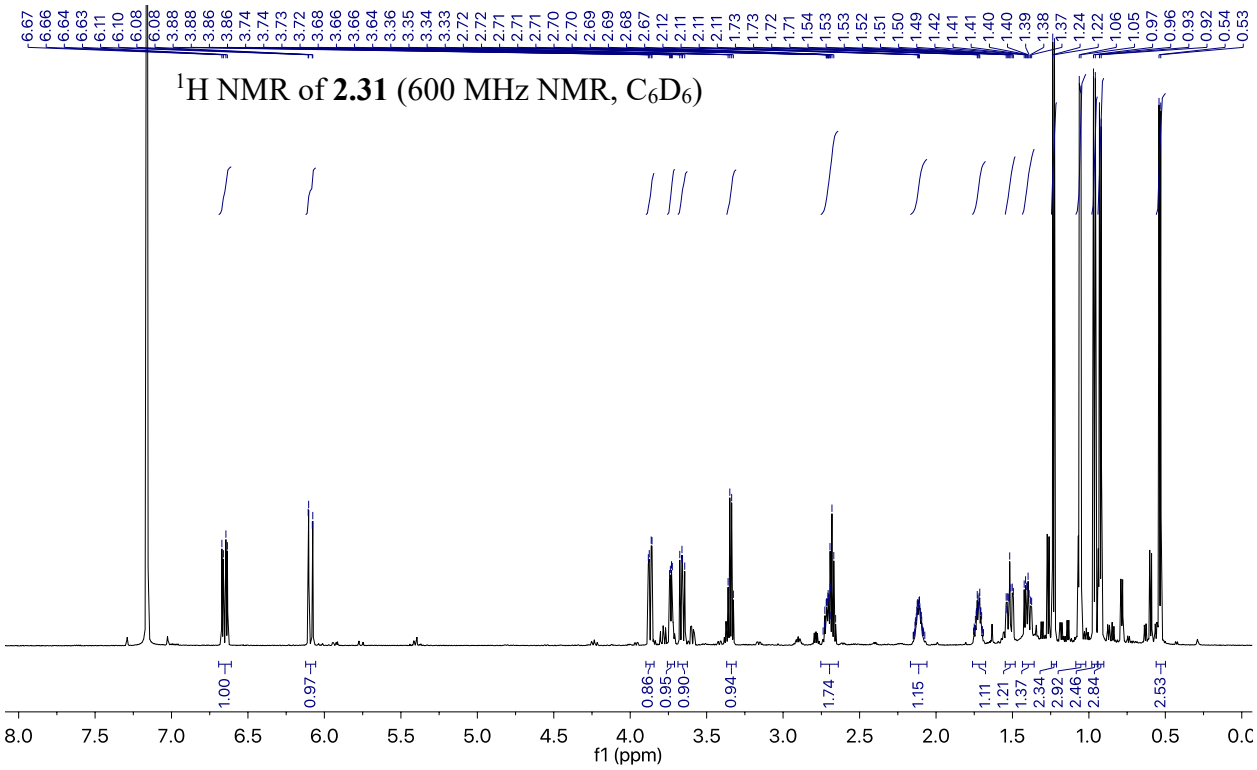
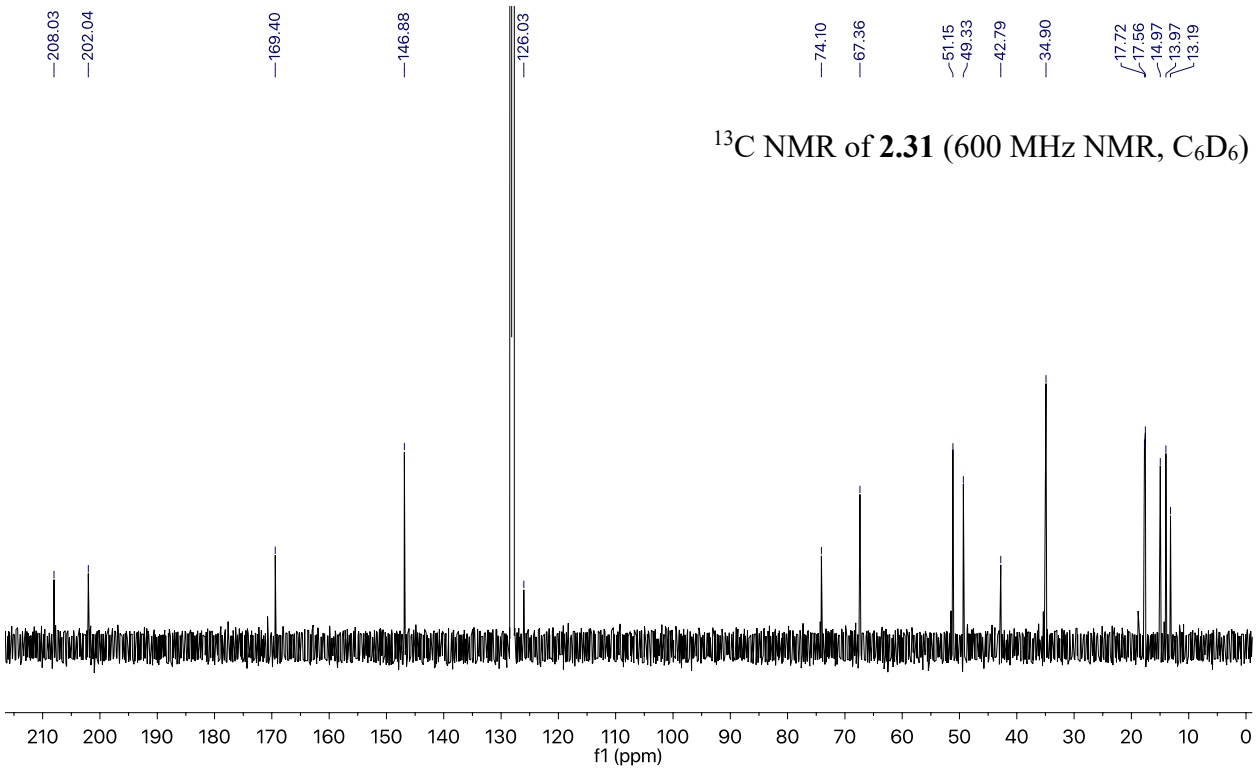


Table II.8 13-desethyl-narbonolide (**2.31**) major isomer NMR Data C_6D_6

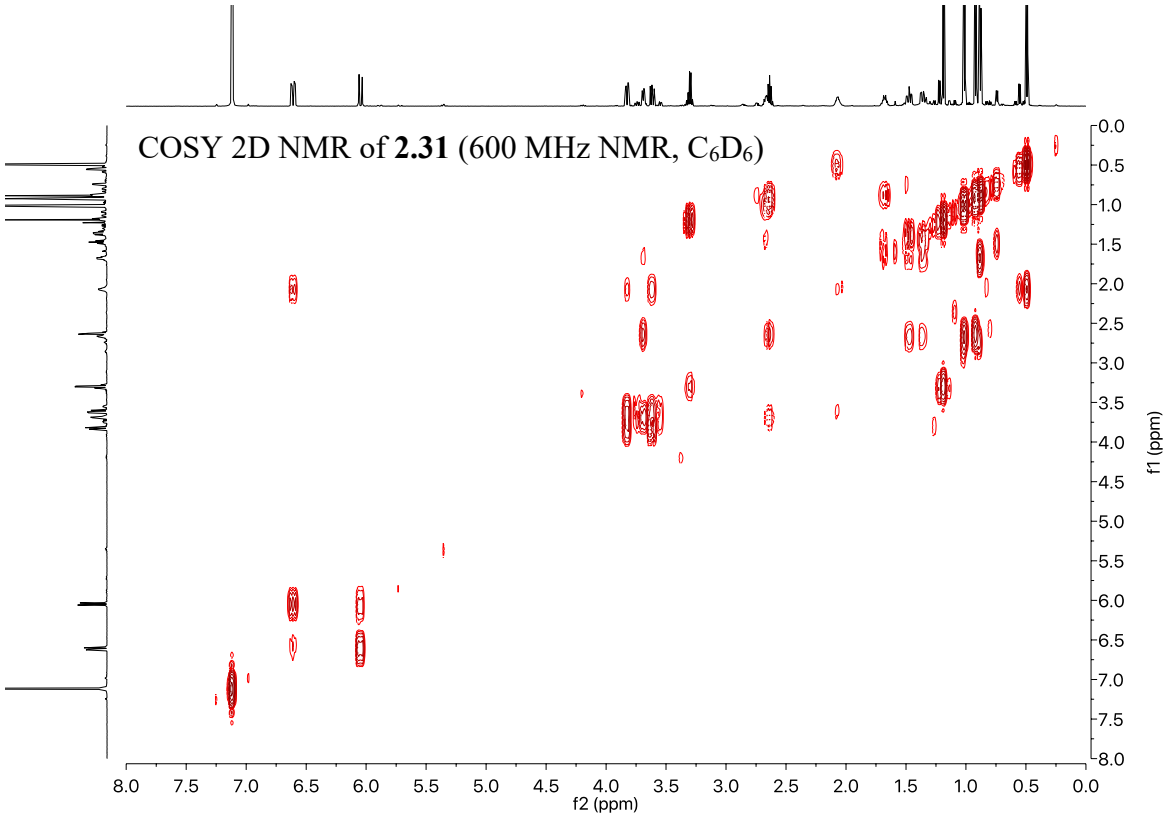
Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		169.4		
2	3.34 (q, J = 7.0)	51.2	16	1, 3, 16
3		208.0		
4	2.68 (p, J = 7.4)	49.3	5, 17	3, 5, 17
5	3.73 (dd, J = 7.5, 3.2)	74.1	4, 6	6, 18
5-OH				
6	1.73 (ddp, J = 13.3, 7.1, 3.1)	34.9	5, 7a, 7b, 18	5, 18
7a	1.52 (ddd, J = 14.1, 9.8, 4.4)	34.9	6, 7a, 8	8, 9, 18, 19
7b	1.40 (ddd, J = 14.0, 10.0, 5.3)		6, 7a, 8	8, 9, 18, 19
8	2.68 - 2.76 (m)	42.8	7a, 7b, 19	9, 19
9		202.0		
10	6.09 (dd, J = 16.0, 1.6))	126.0	11	9, 11, 12
11	6.65 (dd, J = 16.0, 5.7)	146.9	10, 12	9, 10, 12, 13, 20
12	2.07 - 2.17 (m)	34.9	11, 13a, 13b, 20	10, 11
13a	3.87 (dd, J = 11.0, 3.4)	67.4	12, 13b	1, 11, 20
13b	3.66 (dd, J = 11.0, 8.6)		12, 13a	1, 11, 12
14				
15				
16	1.23 (d, J= 7.0)	14.0	2	1, 2, 3
17	0.96 (d, J = 7.3)	13.2	4	3, 4, 5
18	0.93 (d, J = 7.0)	17.6	6	5, 6
19	1.06 (d, J = 6.8)	17.7	8	7, 8, 9
20	0.54 (d, J = 7.1)	15.0	12	11, 12, 13

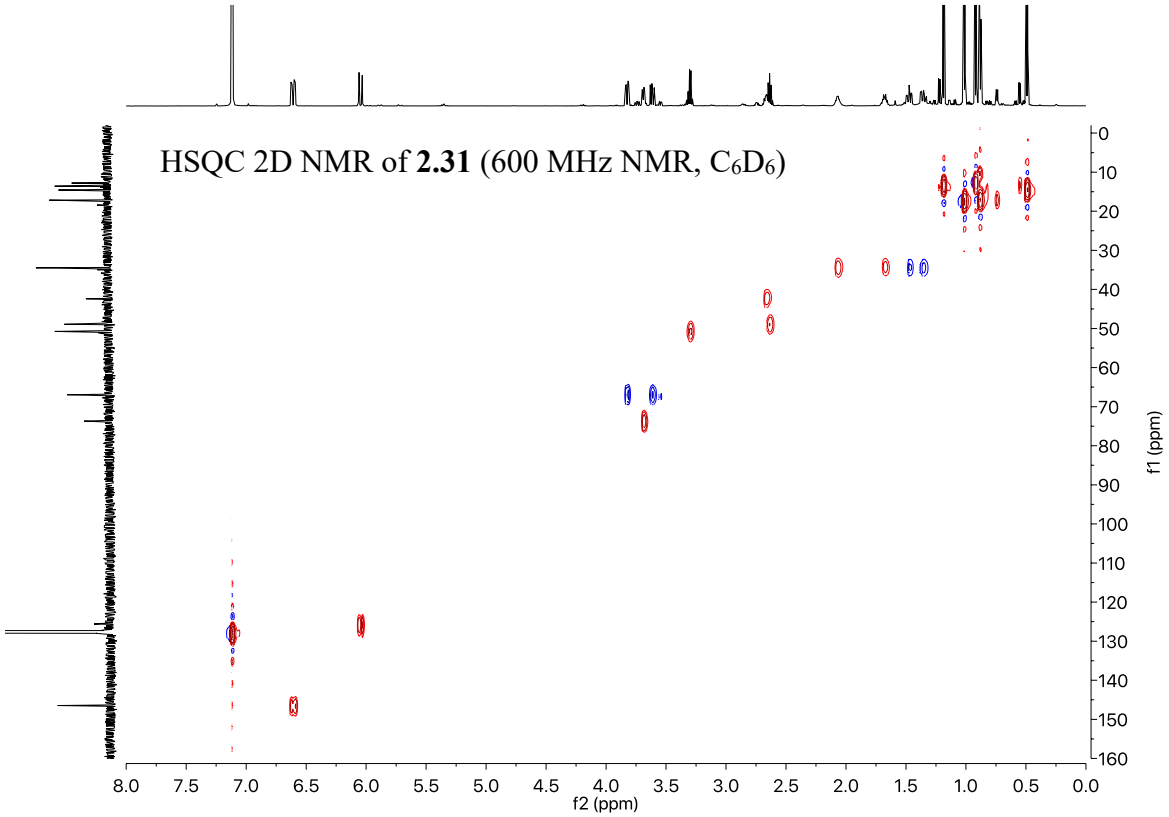


^{1}H NMR of **2.31** (600 MHz NMR, C_6D_6)



^{13}C NMR of **2.31** (600 MHz NMR, C_6D_6)


COSY 2D NMR of **2.31** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)


HSQC 2D NMR of **2.31** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)

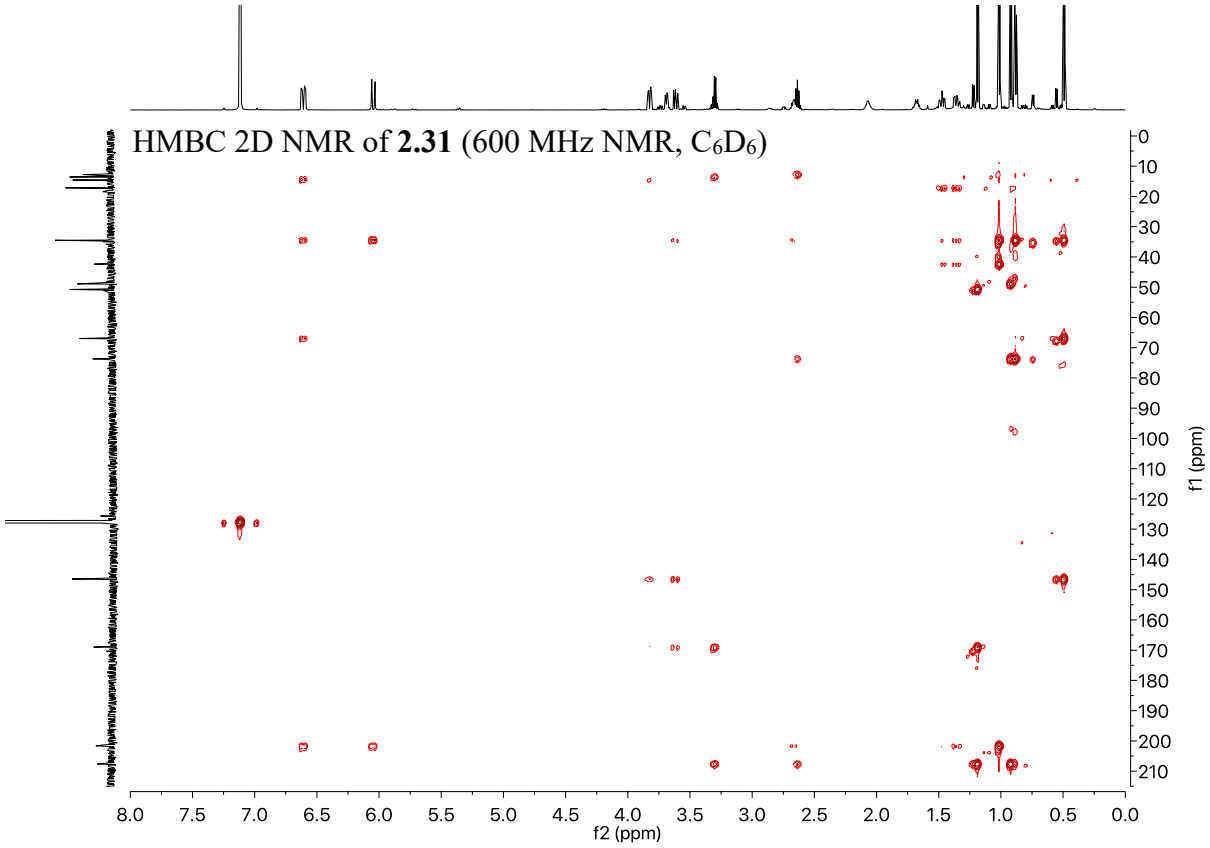
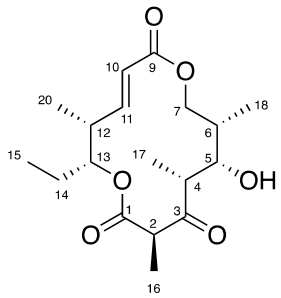

HMBC 2D NMR of **2.31** (600 MHz NMR, C_6D_6)
f2 (ppm)
f1 (ppm)

Table II.9 Ester-narbonolide (**2.27**) 14-membered isomer NMR Data $CDCl_3$



Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		171.9		
2	3.72 (q, J = 7.1)	50	16	1, 3, 16
3		206.5		
4	2.65 (m)	51.3	5, 17	3, 5, 6, 17
5	3.91 (m)	71.3	4, 6	
5-OH	2.26 (bd, J = 3.7)		5	
6	2.00 (m)	36.5	5, 7a, 7b, 18	18
7a	4.43 (dd, J = 3.4, 11.0)	65.7	6, 7b	5, 6, 9, 19
7b	4.14 (dd, J = 5.1, 11.0)		6, 7a	5, 6, 9, 19
8				
9		166.6		
10	5.82 (dd, J = 1.4, 15.9)	121.5	11	9, 12
11	6.92 (dd, J = 6.9, 15.9)	151.7	10, 12	9, 10, 12, 13, 20
12	2.62 (m)	39.1	11, 13, 20	10, 11, 20
13	5.12 (ddd, J = 2.7, 4.9, 9.3)	78.4	12, 14a, 14b	1, 15
14a	1.71 (ddq, J = 2.7, 7.4, 14.7)	25.1	13, 14b, 15	12, 13, 15
14b	1.58 (dqd, J = 4.4, 7.4, 14.7)		13, 14a, 15	12, 15
15	0.91 (t, J = 7.4)	10.3	14a, 14b	13, 14
16	1.38 (d, J = 7.12)	14.7	2	1, 2, 3
17	1.12 (d, J = 6.8)	10.3	4	3, 4, 5
18	1.07 (d, J = 7.2)	16.1	6	5, 6, 7
19				
20	1.12 (d, J = 6.8)	10.6	12	11, 12, 13

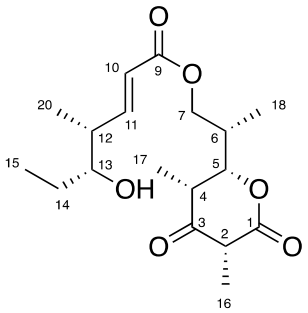
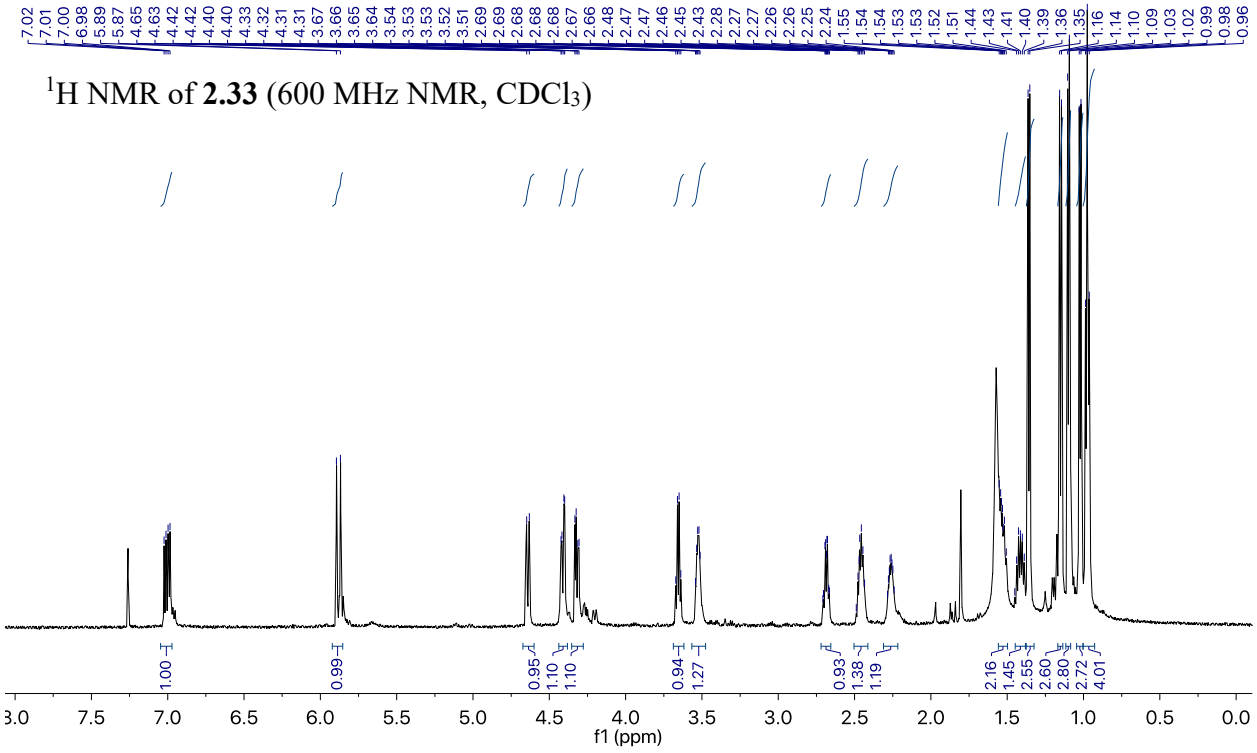


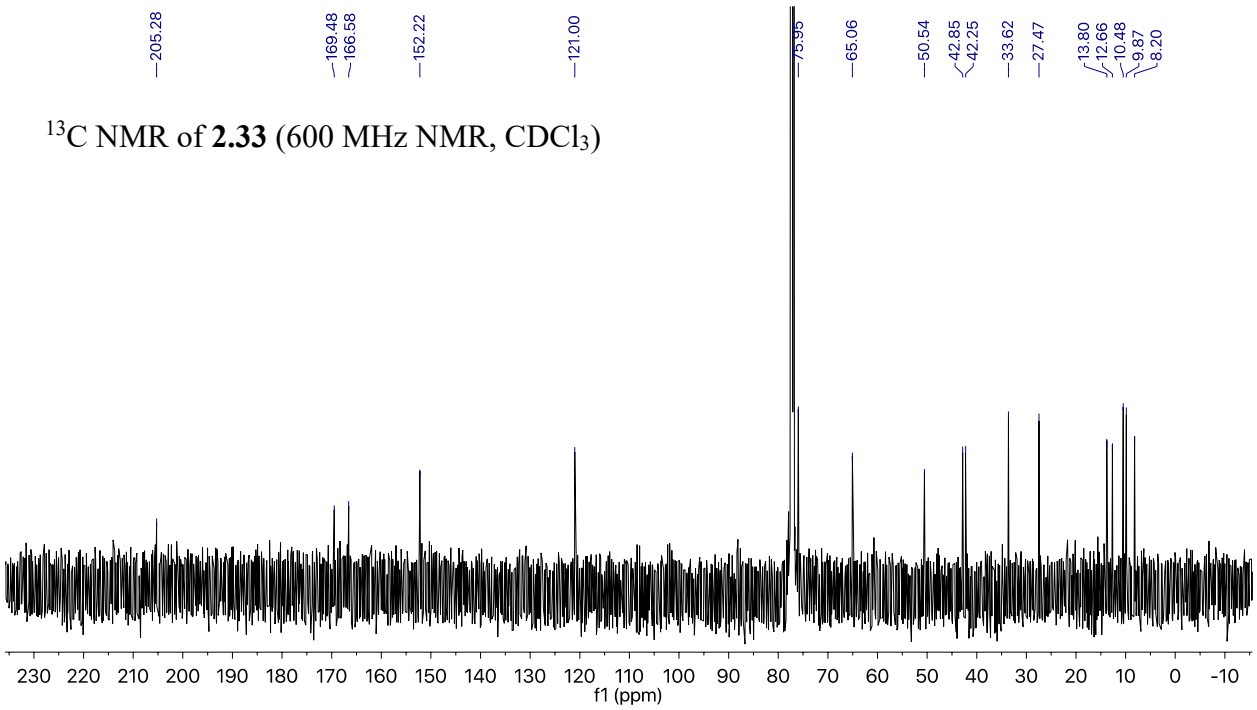
Table II.10 Ester-narbonolide (**2.33**) 6-membered isomer NMR Data $CDCl_3$

Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		169.5		
2	3.65 (q, J = 6.6)	50.5	16	1, 3, 16
3		205.3		
4	2.68 (qd, J = 7.4, 2.2)	42.9	5, 17	3, 17
5	4.64 (dd, J = 10.4, 2.5)	77.2	4, 6	6, 17
5-OH				
6	2.32 - 2.20 (m)	33.6	5, 7a, 7b, 18	
7a	4.41 (dd, J = 11.1, 3.5)	65.1	6, 7b	5, 6, 9, 18
7b	4.32 (dd, J = 11.2, 4.7)		6, 7a	5, 9, 18
8				
9		166.6		
10	5.88 (d, J = 15.8)	121.0	11	9, 12
11	7.00 (dd, J = 15.8, 7.7)	152.2	10, 12	9, 20
12	2.51 - 2.40 (m)	42.2	11, 13, 20	10, 11, 13, 14, 20
13	3.52 (dt, J = 12.2. 3.8)	76.0	12, 14a, 14b	
14a	1.57 - 1.49 (m)	27.5	13, 14b, 15	13, 15
14b	1.47 - 1.36 (m)		13, 14a, 15	
15	0.97 (t, J = 7.5)	10.5	14a, 14b	13, 14
16	1.36 (d, J = 6.6)	8.2	2	1, 2, 3
17	1.15 (d, J = 7.6)	9.9	4	3, 4, 5
18	1.02 (d, J = 7.0)	12.7	6	5, 6, 7
19				
20	1.10 (d, J = 6.8)	13.8	12	11, 12, 13

(6-membered ring)

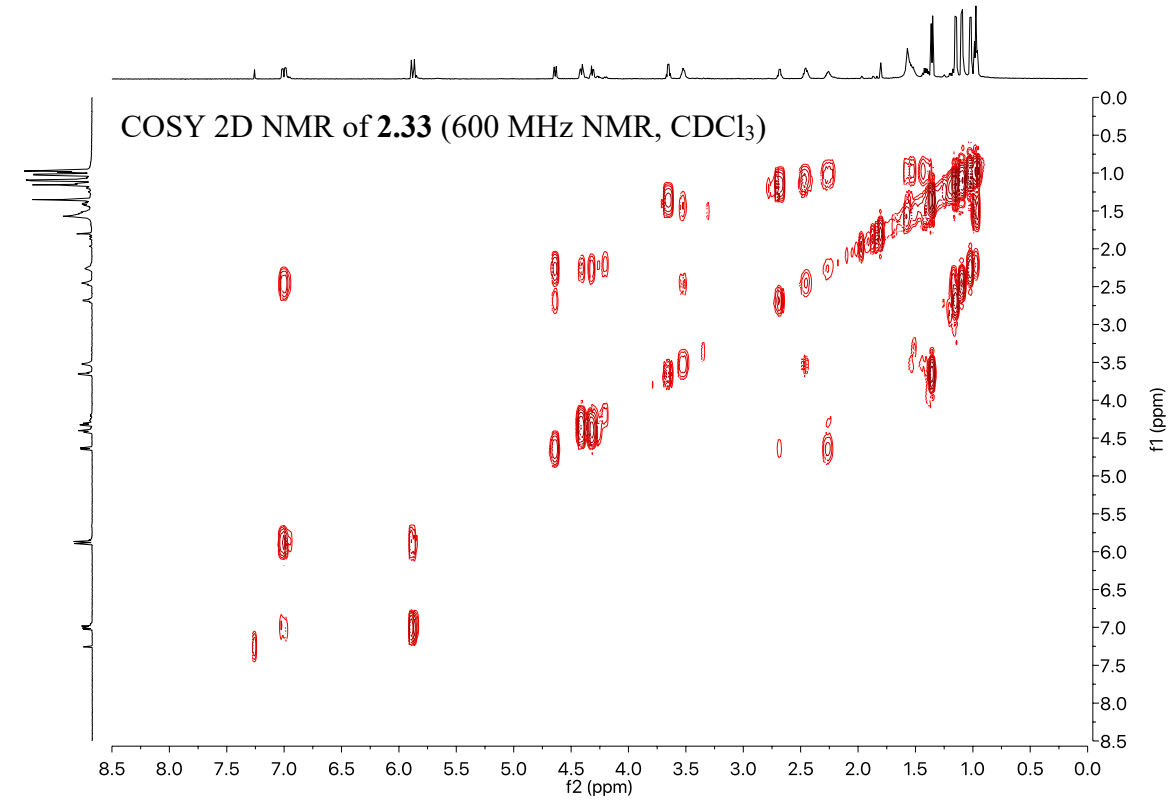


^{1}H NMR of **2.33** (600 MHz NMR, $CDCl_3$)

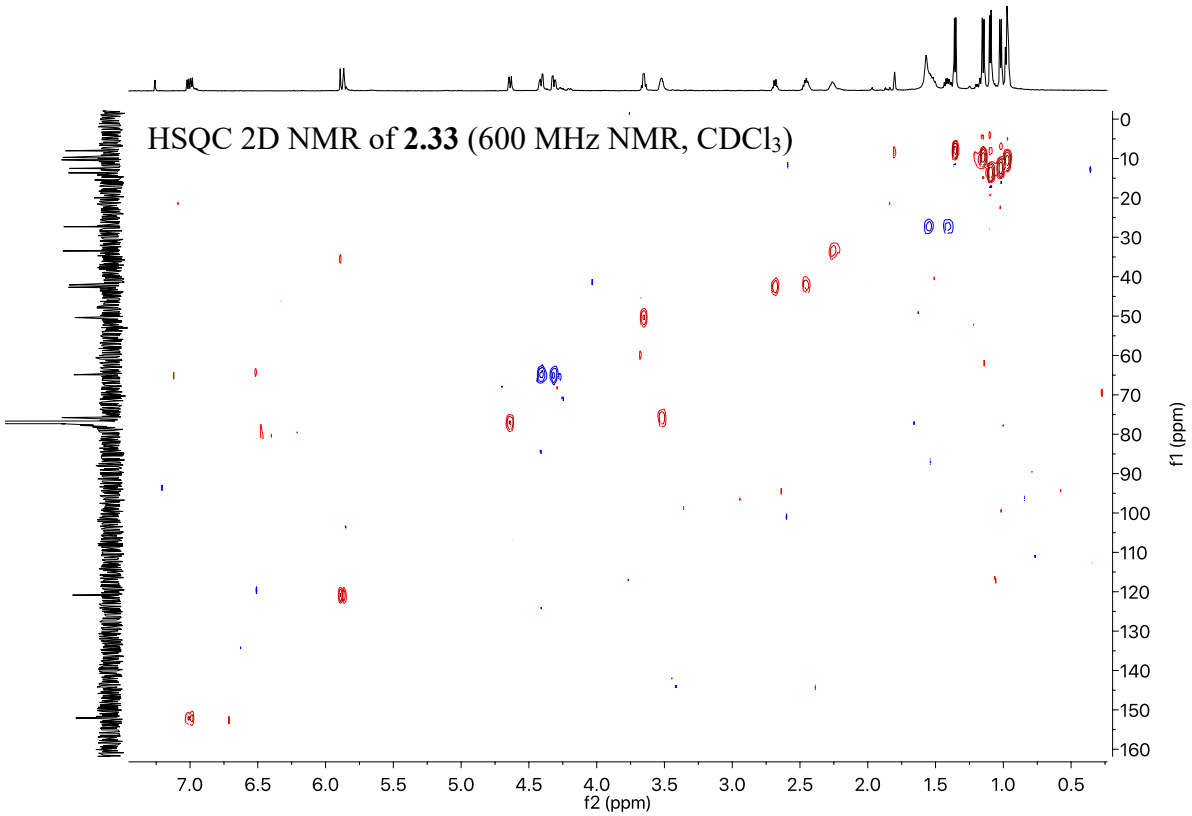


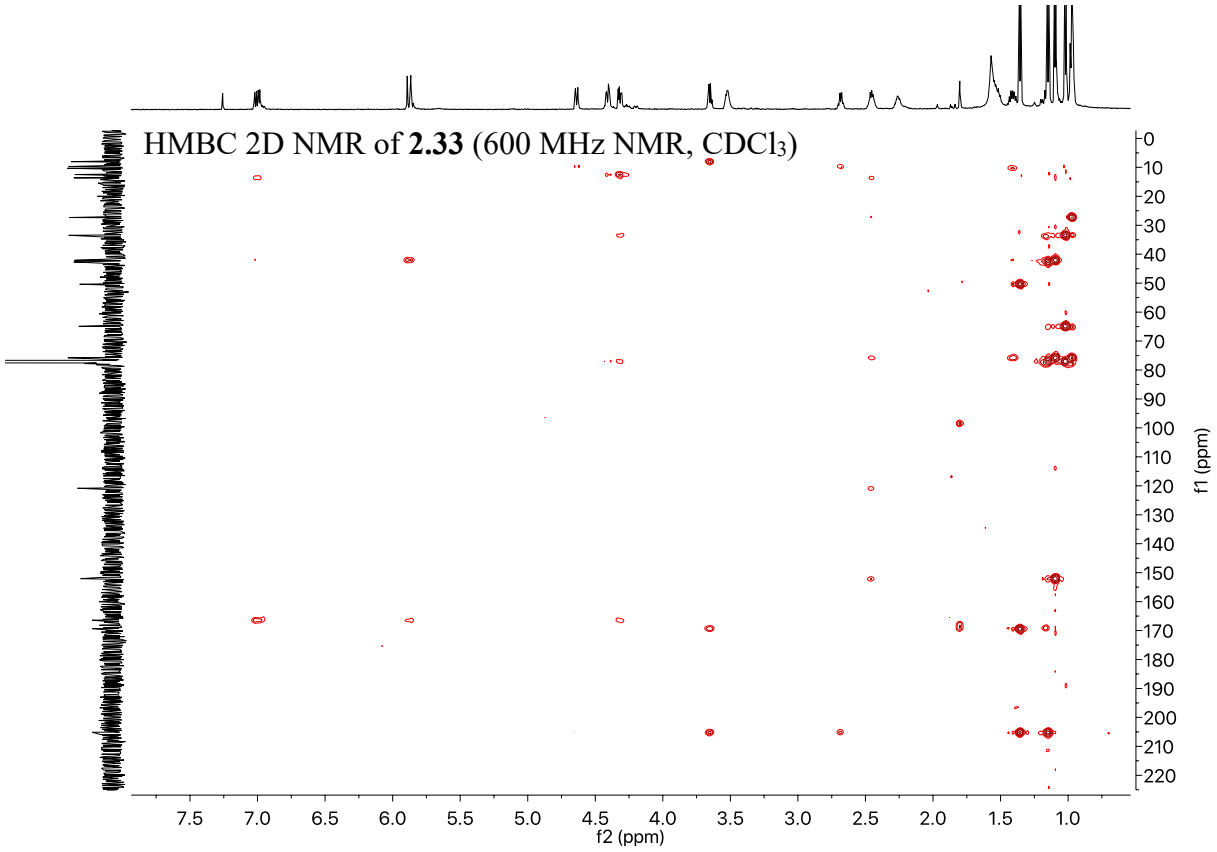
^{13}C NMR of **2.33** (600 MHz NMR, $CDCl_3$)

COSY 2D NMR of **2.33** (600 MHz NMR, $CDCl_3$)

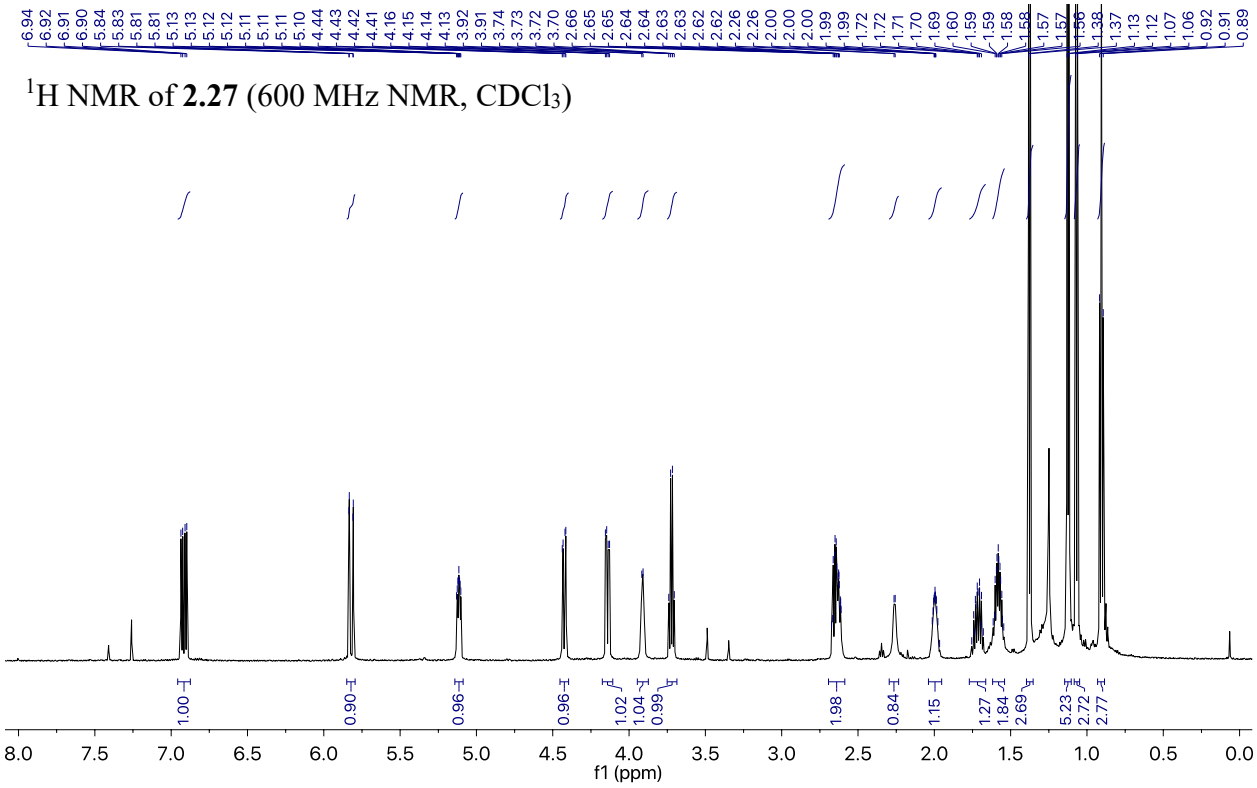


HSQC 2D NMR of **2.33** (600 MHz NMR, $CDCl_3$)

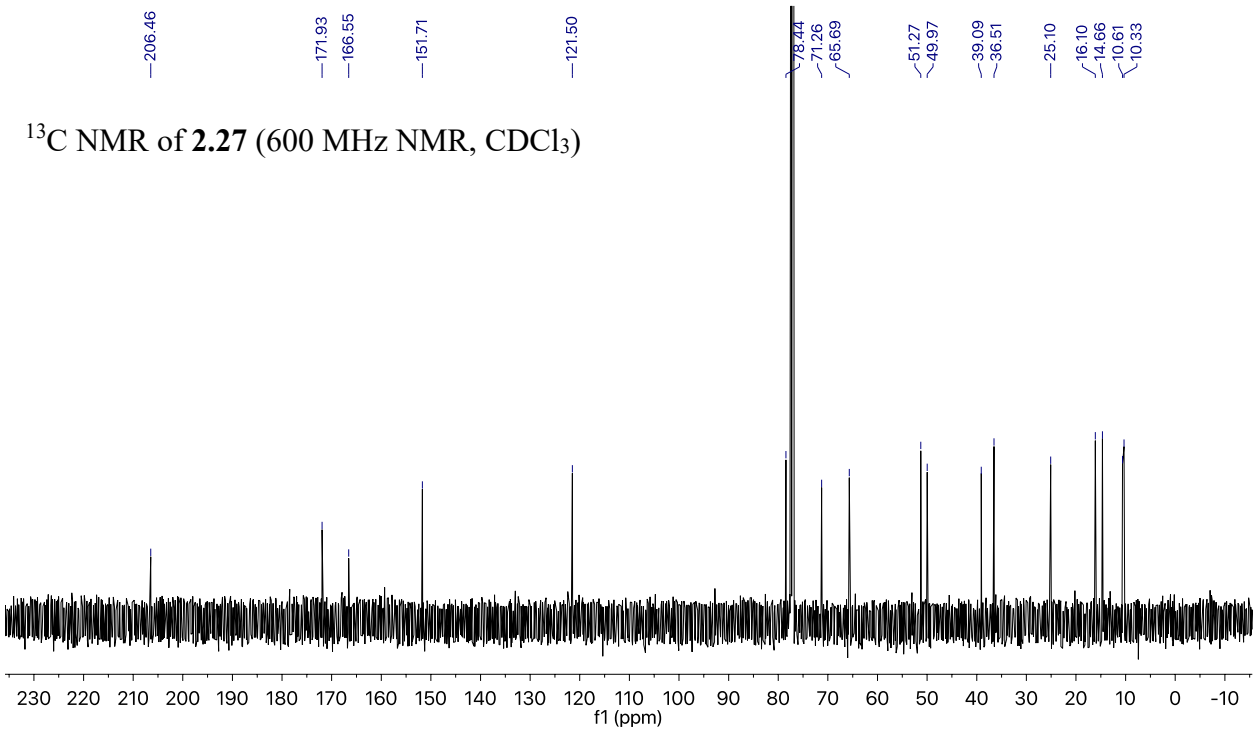



HMBC 2D NMR of **2.33** (600 MHz NMR, $CDCl_3$)
f2 (ppm)
f1 (ppm)

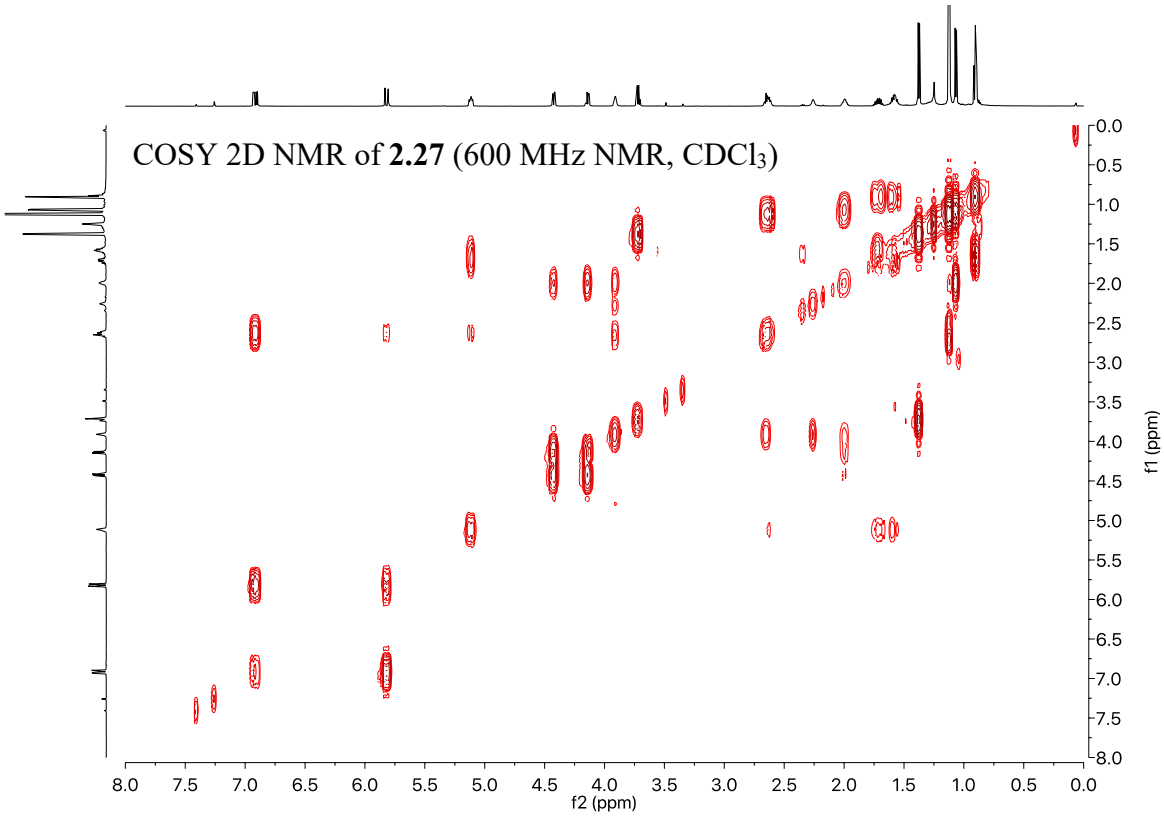
(14-membered ring)

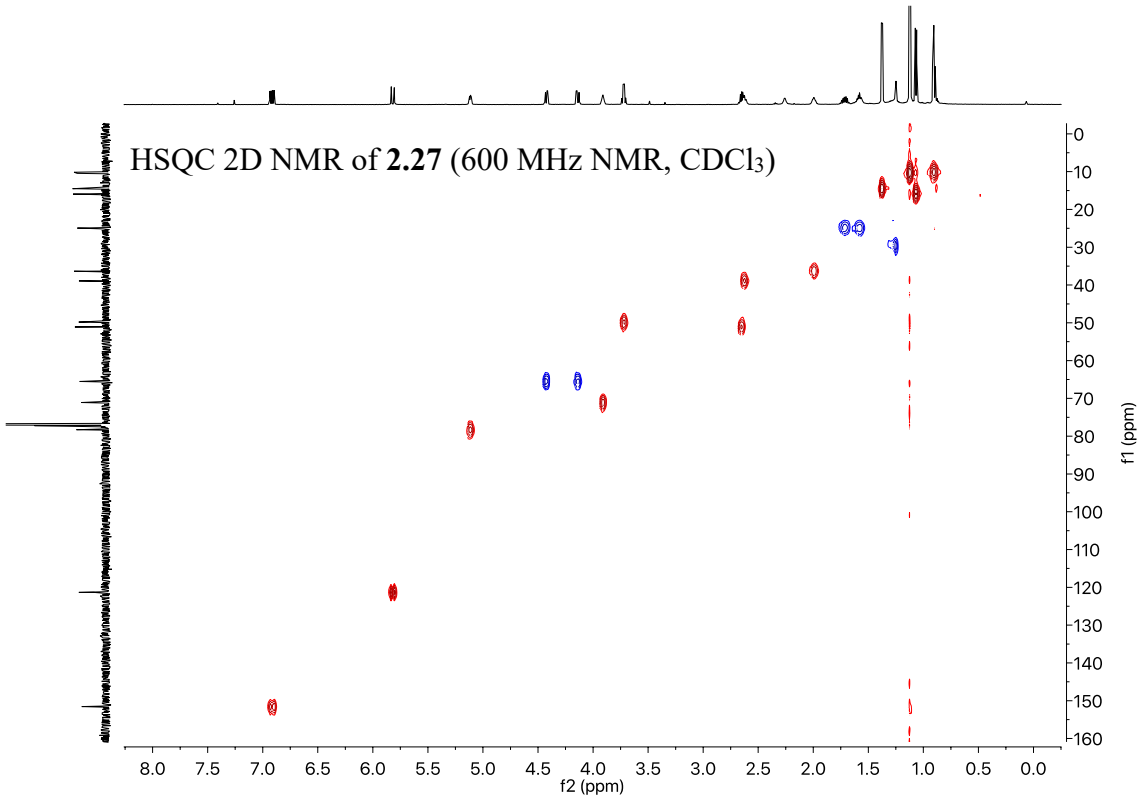


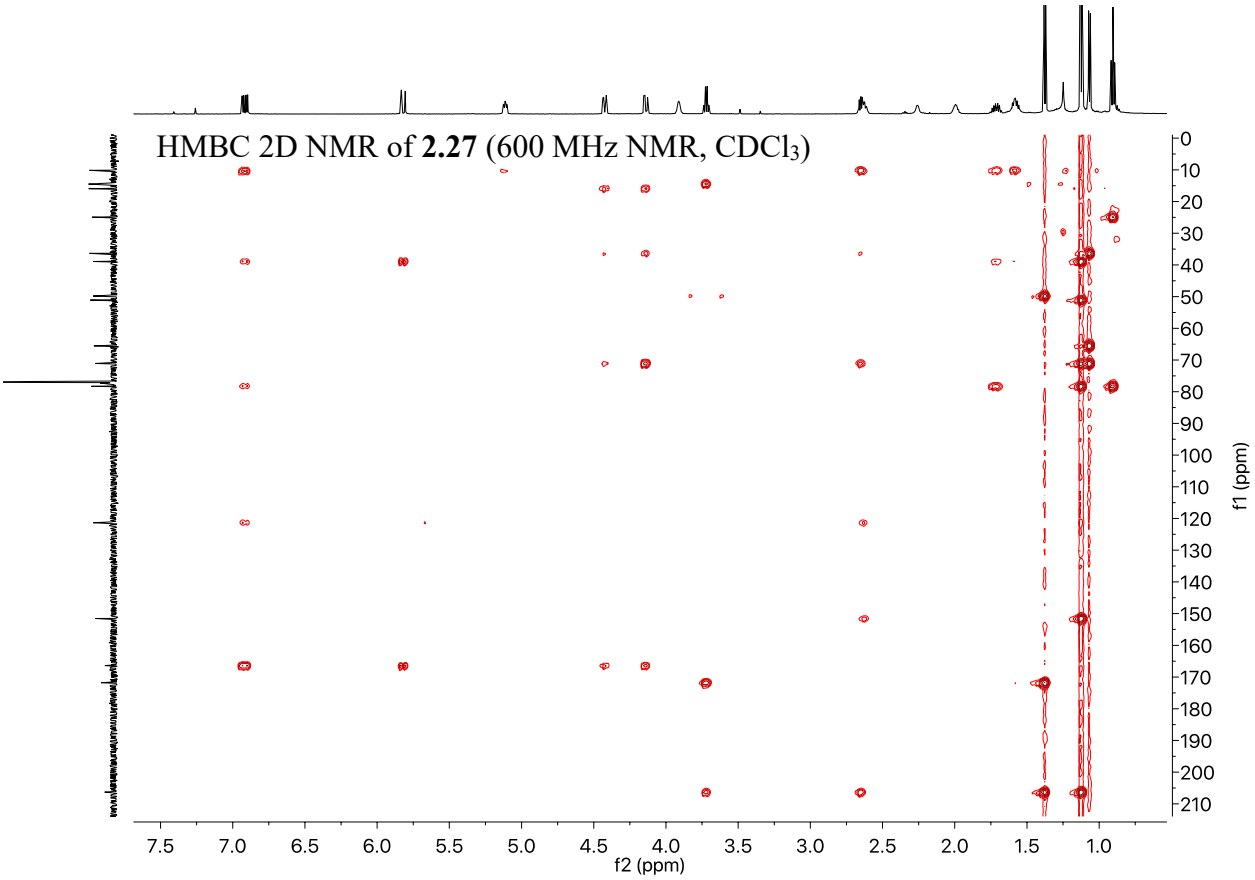
^{1}H NMR of **2.27** (600 MHz NMR, $CDCl_3$)



^{13}C NMR of **2.27** (600 MHz NMR, $CDCl_3$)


COSY 2D NMR of **2.27** (600 MHz NMR, $CDCl_3$)
f2 (ppm)
f1 (ppm)


HSQC 2D NMR of **2.27** (600 MHz NMR, $CDCl_3$)
f2 (ppm)
f1 (ppm)


HMBC 2D NMR of **2.27** (600 MHz NMR, $CDCl_3$)
f2 (ppm)
f1 (ppm)

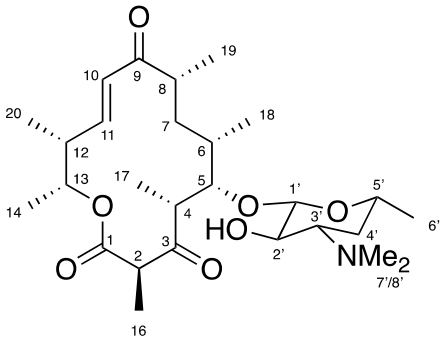
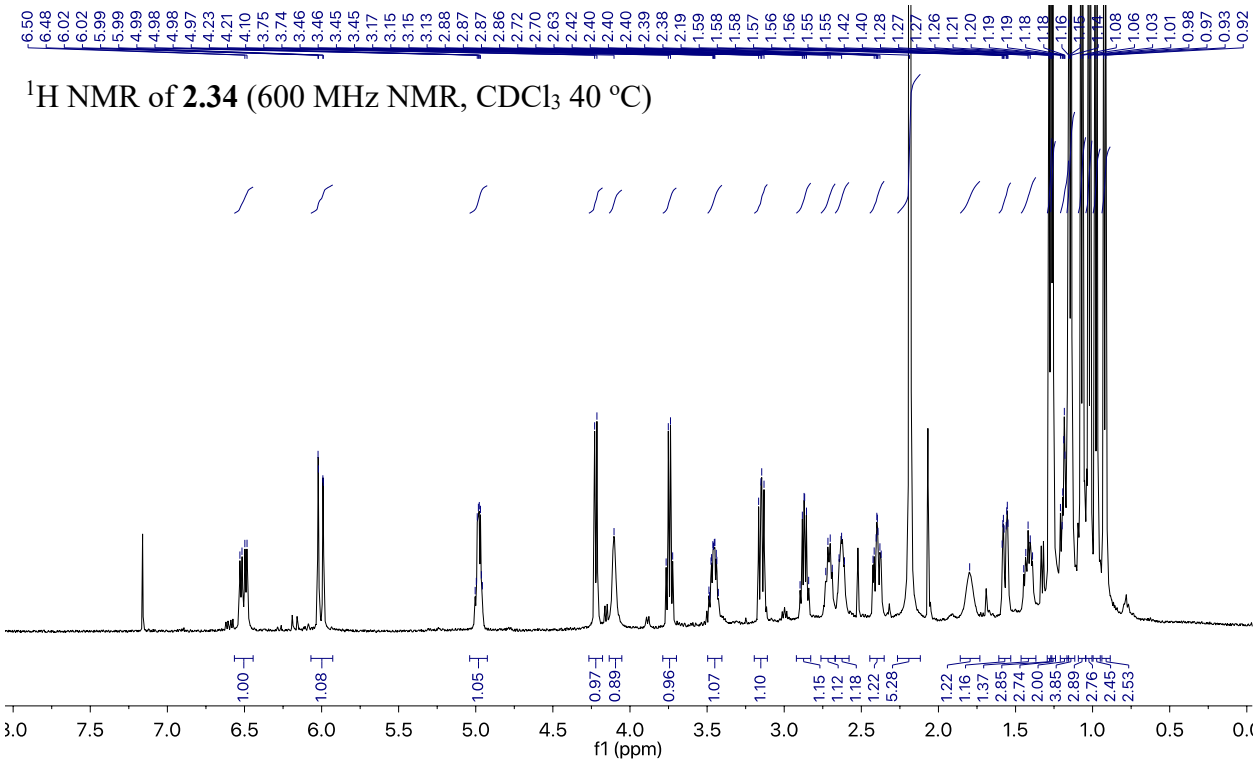
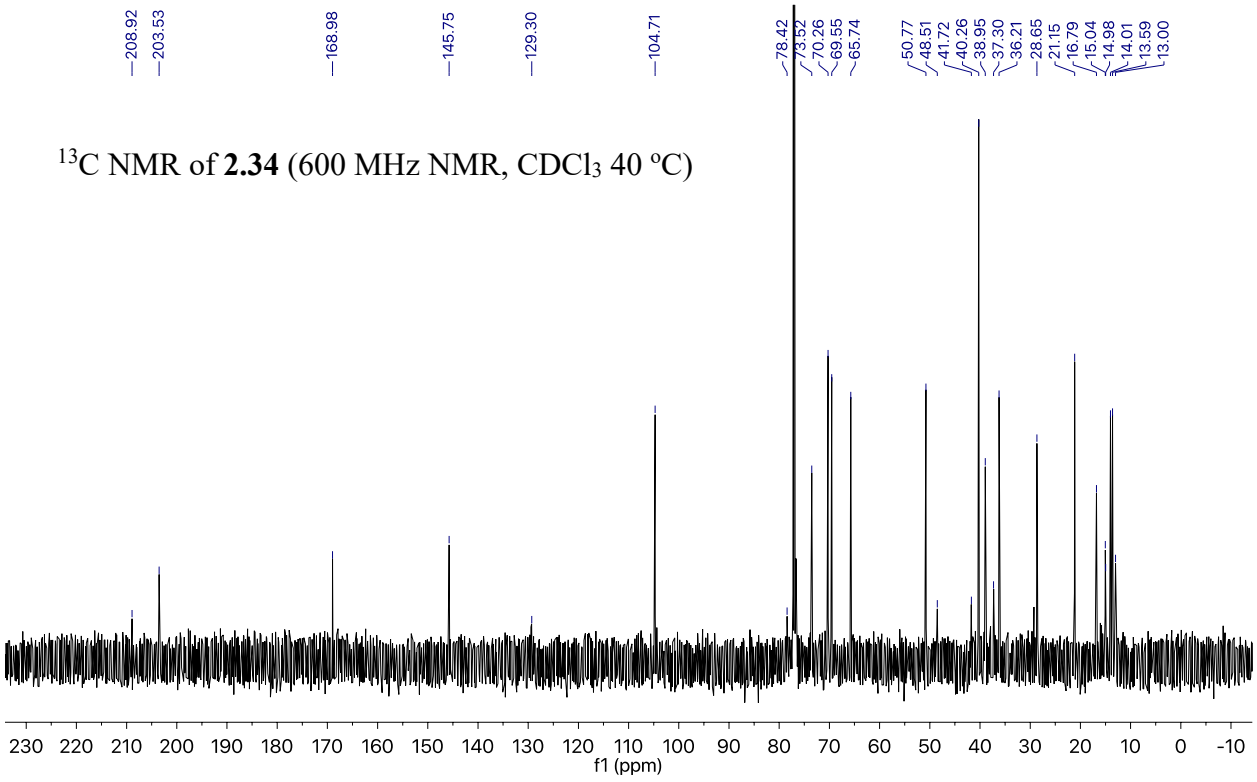


Table II.11 14-desmethyl-narbomycin (**2.34**) NMR Data $CDCl_3$ 40 ºC

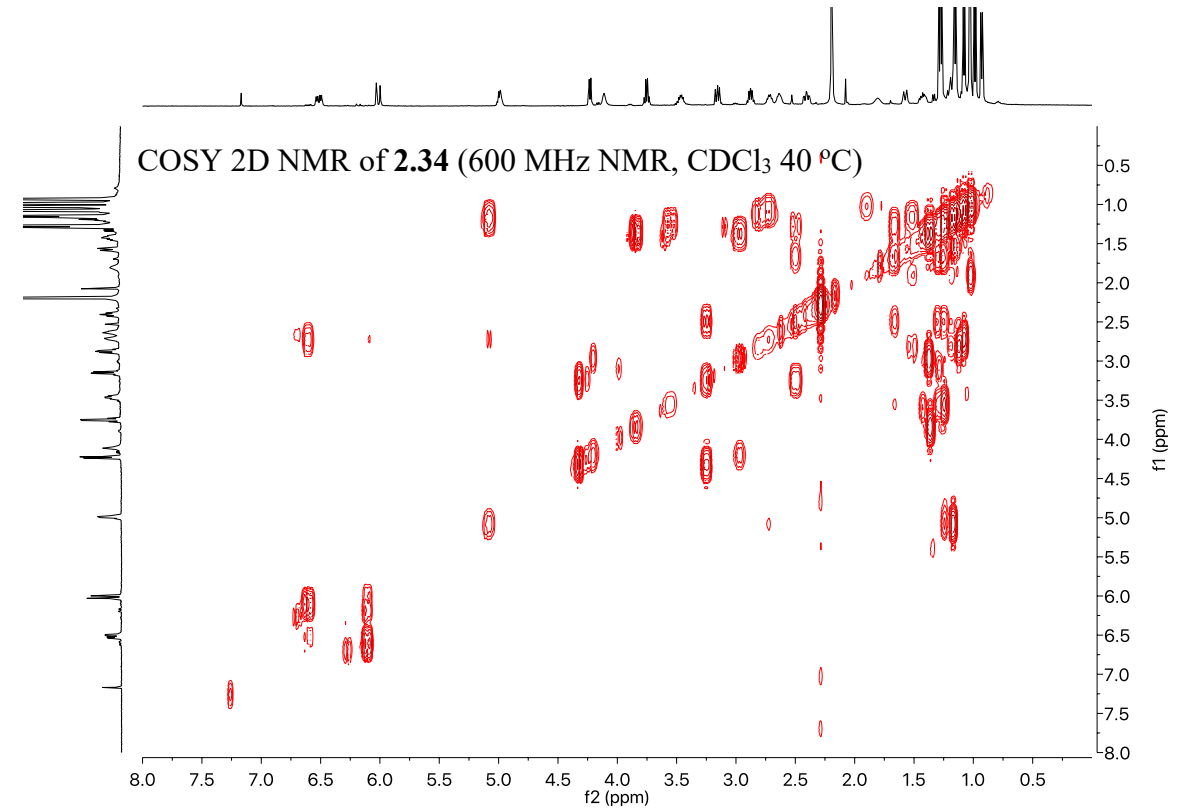
Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		169.0		
2	3.84 (q, J = 7.0)	50.8	16	1, 3, 16
3		208.9		
4	3.01 – 2.94 (m)	48.5	5, 17	3, 5, 6, 17
5	4.20 (bs)	78.4	4	
6	1.90 (bs)	36.2	18	
7a	1.52 (ddd, *J* = 14.1, 8.0, 6.1)	37.2	6, 7b, 8	5, 6, 8, 9, 18
7b	1.19 - 1.14 (m)		7a	5, 8, 9
8	2.86 – 2.76 (m)	41.7	7a, 19	7, 9
9		203.5		
10	6.11 (d, J = 16.0)	129.3	11	9, 12
11	6.61 (dd, J = 16.0, 6.8)	145.7	10, 12	9, 12, 13, 20
12	2.78 - 2.68 (m)	39.0	11, 13, 20	11, 13, 14, 20
13	5.08 (qd, J = 6.5, 3.5)	73.5	12, 14	1, 11
14	1.17 (d, J = 6.6)	15.1	13	12, 13
15				
16	1.36 (d, J = 6.8)	14.1	2	1, 2, 3
17	1.37 (d, J = 7.3)	13.6	4	3, 4
18	1.02 (d, J = 7.0)	16.8	6	5, 6, 7
19	1.12 (d, J = 6.6)	15.0	8	7, 8, 9
20	1.08 (d, J = 7.0)	13.0	12	11, 12, 13
1'	4.32 (d, J = 7.3)	104.7	2'	5, 3'
2'	3.25 (dd, J = 10.2, 7.3)	70.3	1', 3'	1', 3'
3'	2.49 (ddd, J = 12.0, 10.1, 4.0)	65.7	2', 4'a, 4'b	1', 2', 7'/8'
4'a	1.67 (ddd, J = 12.8, 4.0, 2.0)	28.6	3', 4'b	2', 3'
4'b	1.33 – 1.24 (m)		3', 4'a	5', 6'
5'	3.62 – 3.50 (m)	69.6	4'b, 6'	1', 3', 6'
6'	1.24 (d, J = 6.2)	21.2	5'	4', 5'
7/8'	2.29 (s)	40.3		3'

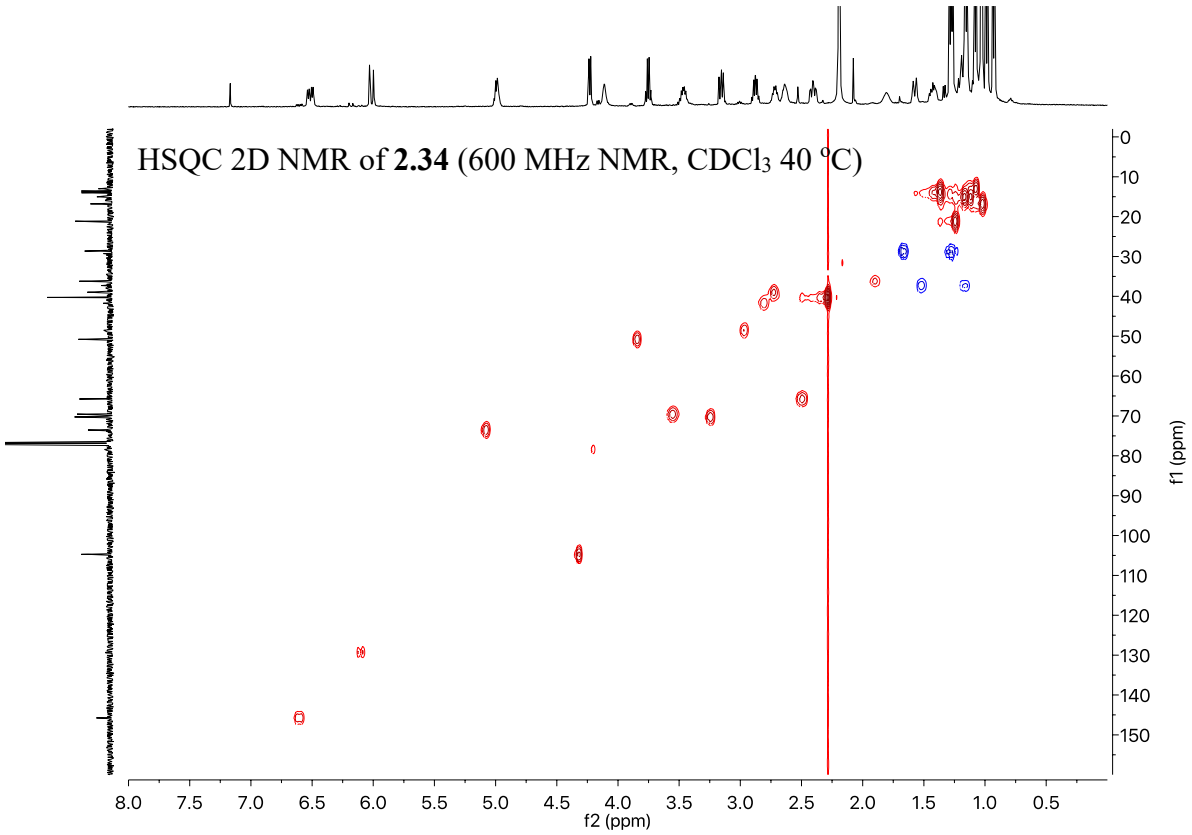


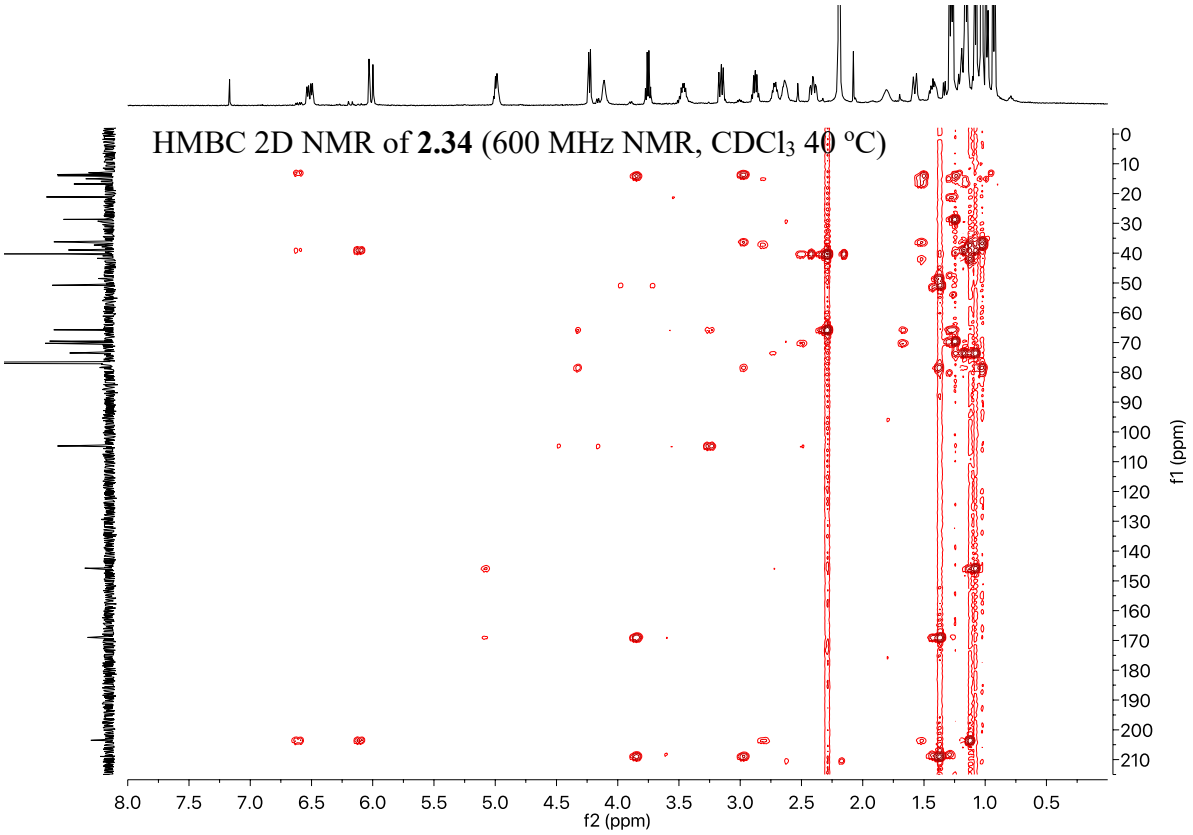
^{1}H NMR of **2.34** (600 MHz NMR, $CDCl_3$ 40 ºC)



^{13}C NMR of **2.34** (600 MHz NMR, $CDCl_3$ 40 ºC)


COSY 2D NMR of **2.34** (600 MHz NMR, $CDCl_3$ 40 ºC)
f2 (ppm)
f1 (ppm)


HSQC 2D NMR of **2.34** (600 MHz NMR, $CDCl_3$ 40 ºC)
f2 (ppm)
f1 (ppm)


HMBC 2D NMR of **2.34** (600 MHz NMR, $CDCl_3$ 40 ºC)
f2 (ppm)
f1 (ppm)

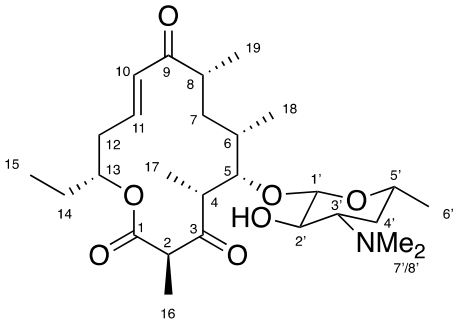
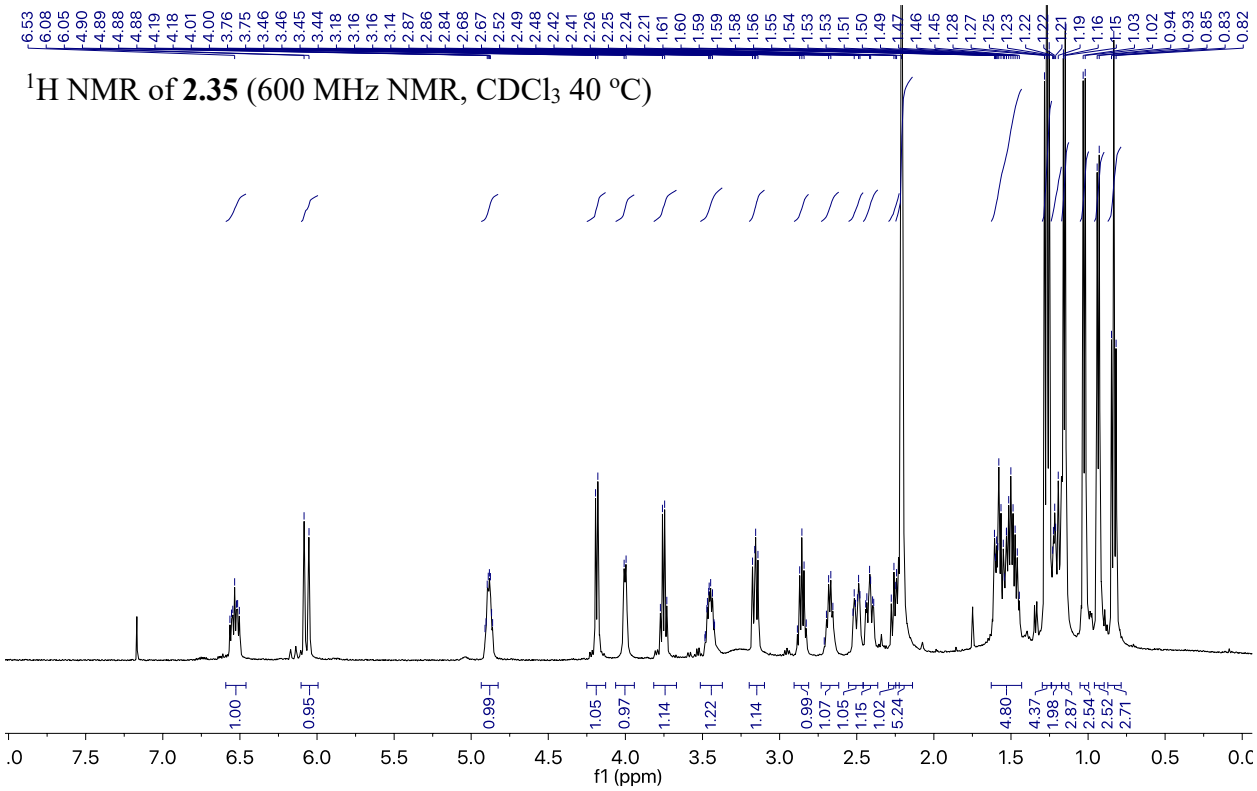
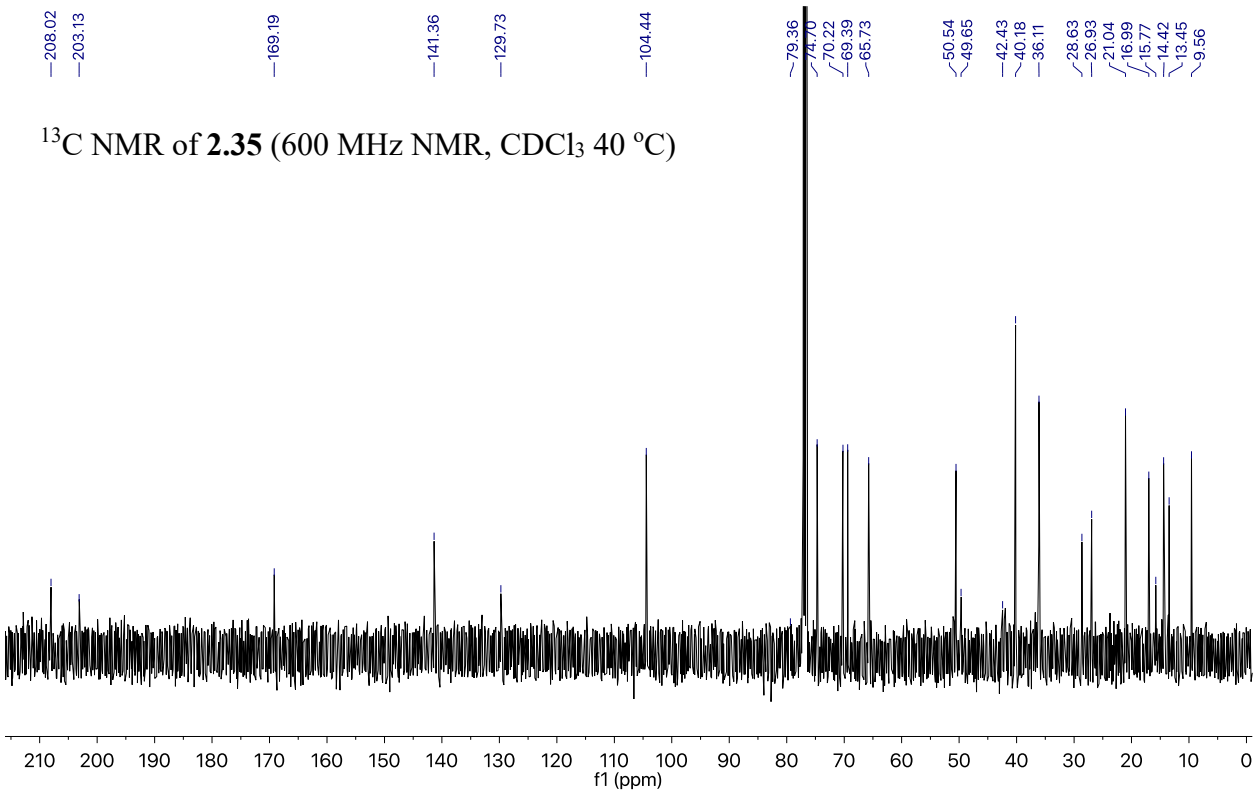


Table II.12 12-desmethyl-narbomycin (**2.35**) NMR Data $CDCl_3$ 40 ºC

Position	δ ^{1}H (mult, J Hz)	δ ^{13}C	COSY	HMBC
1		169.6		
2	3.84 (q, J = 6.9)	50.9	16	1, 3, 16
3		208.4		
4	2.95 (p, J = 7.2)	50.0	5, 17	3, 5, 6, 17
5	4.09 (d, J = 6.3)	79.7	4	1', 2, 6, 7, 18
6	1.72 - 1.52 (m)	36.4	18	7
7a	1.72 - 1.52 (m)	36.6	8	8, 18, 19
7b	1.33 - 1.29 (m)			18, 19
8	2.77 (dq, J = 13.3, 6.6)	130.0	7a, 19	7, 9, 19
9		203.5		
10	6.16 (d, J = 15.7)	130.0	11	9, 12
11	6.62 (ddd, J = 15.3, 9.4, 5.5)	141.7	10, 12a, 12b	9, 12
12a	2.59 (dt, 15.0, 4.2)	36.5	11, 12b, 13	
12b	2.38 - 2.31 (m)		11, 12a, 13	10, 11, 13
13	4.98 (qd, J = 7.1, 3.3)	75.0	12a, 12b, 14a, 14b	
14a	1.72 - 1.52 (m)	27.3	13, 15	12, 13, 15
14b	1.72 - 1.52 (m)		13, 15	12, 13, 15
15	0.92 (t, J = 7.4)	9.91	14a, 14b	13, 14
16	1.35 (d, J = 7.3)	14.7	2	1, 2, 3
17	1.37 (d, J = 7.7)	13.81	4	5
18	1.03 (d, J = 6.7)	17.3	6	5, 6, 7
19	1.12 (d, 6.6)	16.2	8	7, 8, 9
20				
1'	4.27 (d, J = 7.2)	104.8	2'	3', 5, 5'
2'	3.84 (dd, J =10.2, 7.3)	70.5	1', 3'	1', 3'
3'	2.51 (td, J = 11.6, 10.3, 3.8)	66.1	2', 4'a, 4'b	2'
4'a	1.72 - 1.52 (m)	29.0	3'	6'
4'b	1.27 - 1.23 (m)		3'	3', 6'
5'	3.54 (dq, J = 10.9, 5.7)	69.7	6'	
6'	1.24 (d, J = 6.2)	21.4	5'	4', 5'
7'/8'	2.30 (s)	42.3		3'

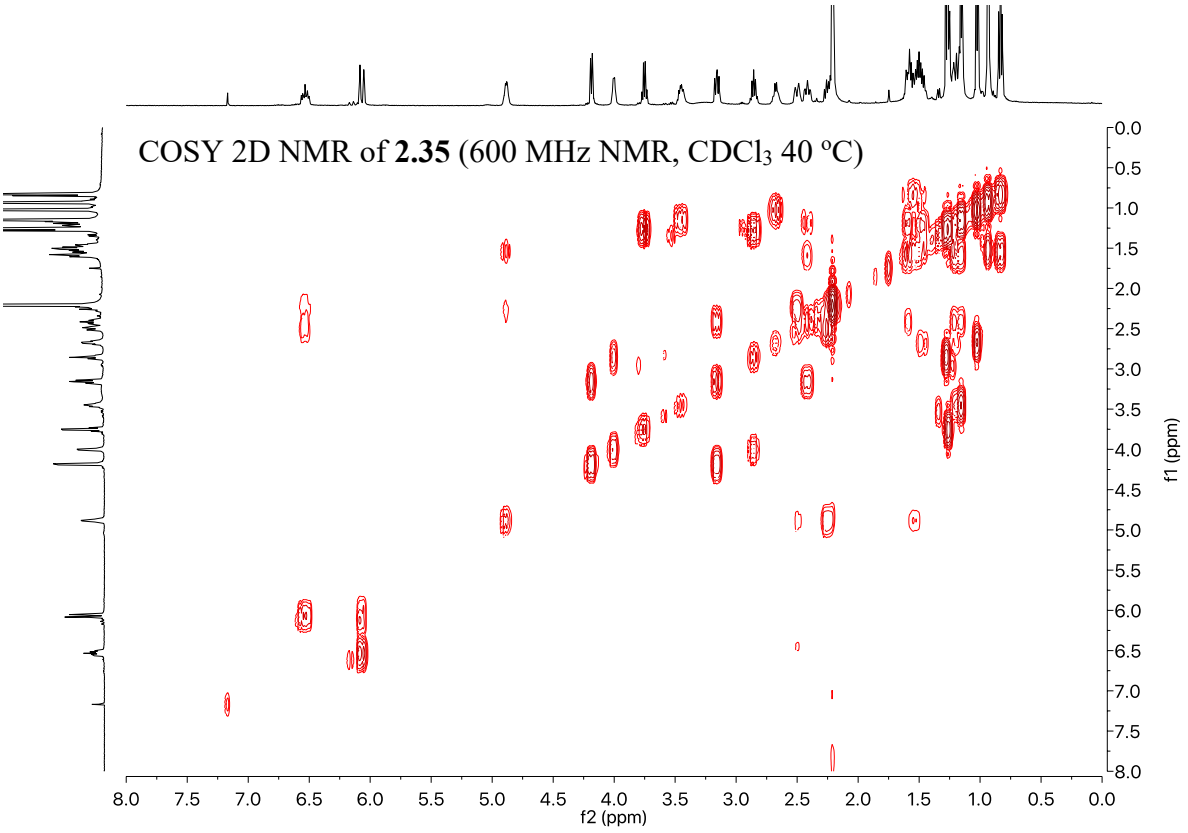


^{1}H NMR of **2.35** (600 MHz NMR, $CDCl_3$ 40 °C)

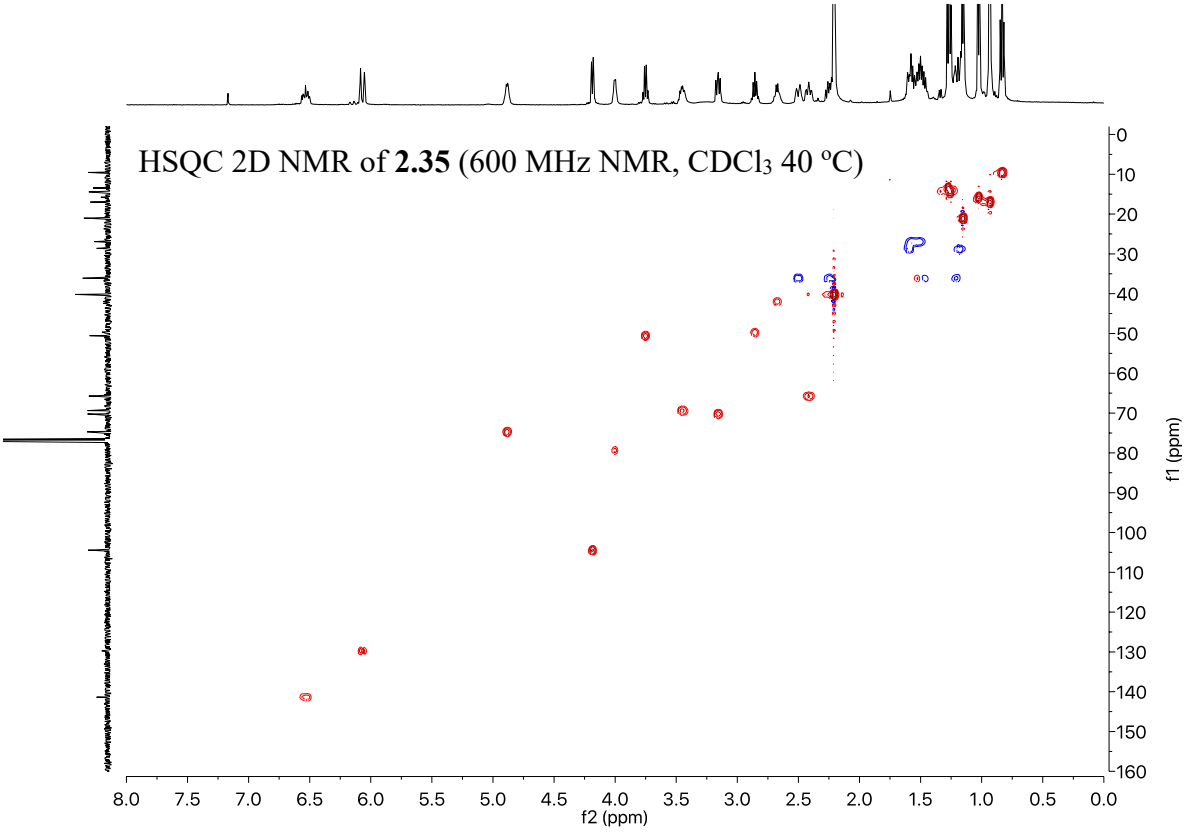


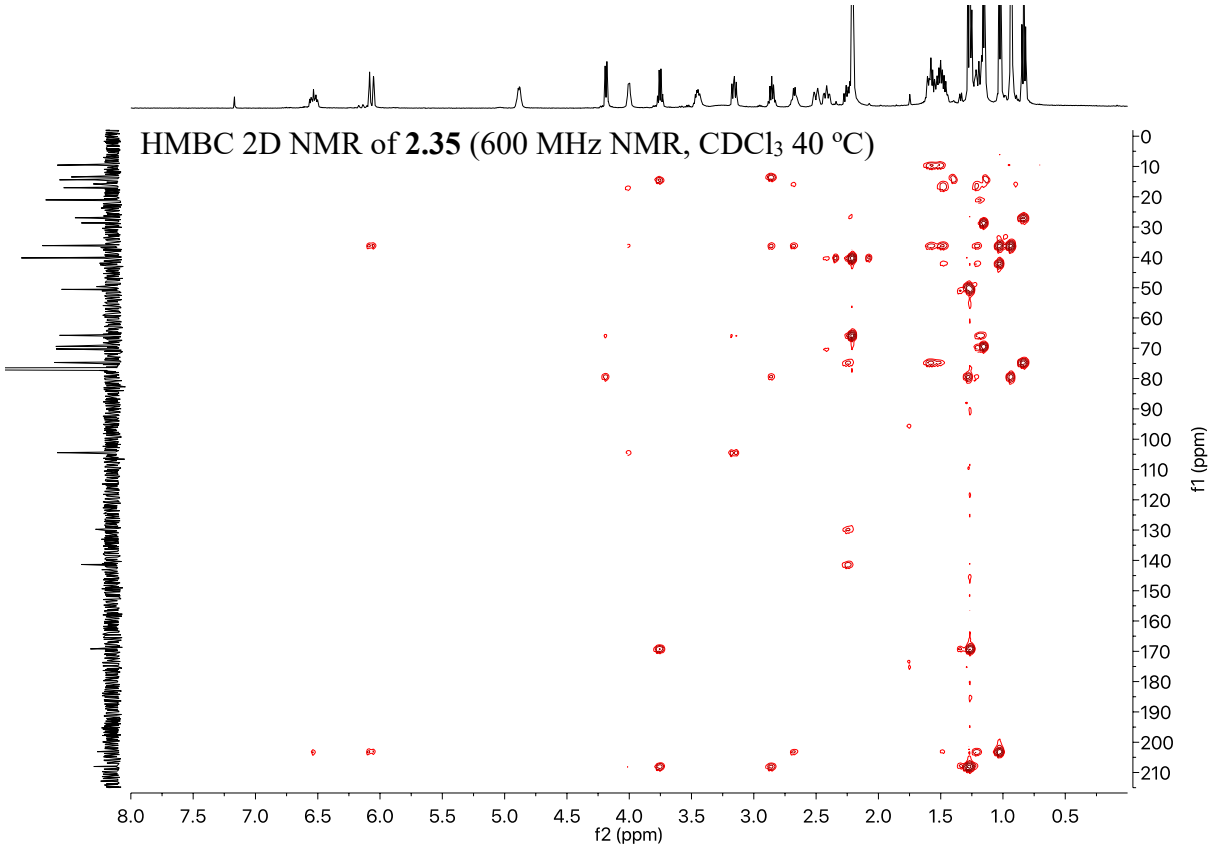
^{13}C NMR of **2.35** (600 MHz NMR, $CDCl_3$ 40 °C)

COSY 2D NMR of **2.35** (600 MHz NMR, $CDCl_3$ 40 ºC)



HSQC 2D NMR of **2.35** (600 MHz NMR, $CDCl_3$ 40 ºC)




HMBC 2D NMR of **2.35** (600 MHz NMR, $CDCl_3$ 40 °C)
f2 (ppm)
f1 (ppm)

CHAPTER III

Directed Evolution of the Pikromycin Thioesterase (Pik TE)

The majority of this chapter is part of an ongoing manuscript titled "Directed Evolution of a Modular Polyketide Synthase Thioesterase for Generation of a Hybrid Macrocyclic Ring System" and reproduced with permission from all authors Maria L. Adrover-Castellano, Jennifer J. Schmidt, Carolyn A. Glasser, Fengrui Qu and David H. Sherman. *In progress*. **2023**.

III.1 Abstract

Modular type I polyketide synthases (PKSs) comprise a family of enzymes capable of synthesizing a diverse class of natural products. The biochemical features of these systems include the extension and processing of polyketide chains in a stepwise, stereospecific manner, organized by a series of modules divided into distinct catalytic domains. Previous work has indicated that a primary hurdle for utilizing PKS modules to create diverse macrolactones hinges on the selectivity of the thioesterase (TE) domain. Herein, we generated a novel hybrid 12-membered macrolactone/lactam ring system employing an unnatural amide hexaketide intermediate in conjunction with an engineered TE S148C mutant from the pikromycin (Pik) biosynthetic pathway. This unnatural molecule was formed in attenuated yields compared to the native hexaketide substrate. Thus, we embarked on a directed evolution campaign to identify Pik TE variants with enhanced selectivity for the desired macrocyclic product. After three rounds of evolution, a series of mutant Pik TE proteins were identified, and further combinations of beneficial mutations produced a composite variant with 6-fold enhanced isolated yield compared to the parent enzyme. This study offers new insights into the range of amino acid residues, both

distal and proximal to the active site that impart improved selectivity and yield against the unnatural polyketide substrate. Overall, this work provides new knowledge about the modulation of selectivity in macrocyclization, and demonstrates an effective approach for overcoming a key PKS pathway gatekeeper.

III.2 Introduction

Macrocyclic scaffolds are found in many natural products from diverse sources. These commonly include twelve or more carbons atoms in their cyclic framework.[128] Synthetic strategies have been developed for macrolactonization, and macrolactamization, which include transition metal catalyzed coupling reactions, click chemistry and ring-closing metathesis.[128, 129] A wide variety of these molecules have been developed as pharmaceuticals with over one hundred FDA approved drugs carrying a macrocyclic core (e.g. macrolide and cyclic peptide antibiotics, anthelmintics, anticancer agents, immunomodulators); showing their current medicinal value and continued promise for future drug discovery and development.[130, 131] Macrolide antibiotics, including erythromycin and its semi-synthetic derivatives (clarithromycin, azithromycin, among others) act against bacterial pathogens as ribosome inhibitors and have synergistic immunomodulatory activity in the human host.[132, 133] The macrolide chemical structure contains a macrolactone ring possessing one or more sugar moieties (amino sugars, deoxy sugars or both) and other functionalities such as hydroxyl groups important for their mechanism of action in the bacterial ribosome peptidyl transferase center.[134] Frequently, the structural complexity of large-ring macrolactones is met with significant synthetic challenges. Options for generating these include total synthesis, semi-synthesis and chemoenzymatic strategies.[102, 117, 118, 120, 135-140] For example, the first total synthesis of erythromycin[8-10] in 1981 has over 50 synthetic steps and an overall yield of less than 1%. These large macrocycles possess stereochemical and structural

diversity, which often results in laborious, low yielding reaction schemes, possessing a barrier for further diversification or pursuing structure-activity relationship studies. Alternatively, semi-synthesis has been employed to generate widely used antibiotics such as the erythromycin derivatives clarithromycin, azithromycin, telithromycin, and others.[101, 102, 132, 137] Despite these approaches reducing the synthetic step counts and improving the overall yields, modifications to the starting material scaffolds are often limited by reactive functional groups and selectivity challenges involving diverse positions within these intricate ring structures.[101, 129, 139, 141-144]

Chemoenzymatic synthesis encompasses chemical synthesis of intermediates from simple building blocks, combined with enzyme catalyzed reactions to obtain efficient production of novel molecules and pharmaceuticals using aqueous reaction conditions, while avoiding toxic reagents, solvents and protecting groups.[111, 145-148] Although workable methodologies exist for macrolactonization, these are often met with unfavorable energetics related to a combination of entropic and enthalpic factors relating to intramolecular cyclization.[117, 118, 140, 149] To address the challenges of total and semi-synthesis, we have been investigating the pikromycin (Pik) biosynthetic pathway thioesterase (TE) domain as a key biocatalyst for macrolide antibiotic discovery.[61, 111] Previous studies revealed how the Pik TE binds its linear substrate and prepares it for cyclization reactions. Moreover, it can accommodate hexa- and heptaketides in its active site through a hydrophobic chamber that include substrate-protein anchoring hydrogen bonds. Remarkably, the TE induces a curled conformation of the linear polyketide through the hydrophobicity of the surrounding residues, coupled with a hydrophilic barrier at the exit site of the enzyme channel. The exit site barrier forces the substrate back into the active site, and positions the nucleophilic hydroxyl group for macrocyclization.[36, 150]

The Pik pathway is comprised of modular type I PKSs that selectively catalyzes key transformations on structurally distinct intermediates generating a 12-membered macrolactone ring, 10-deoxymethynolide (10-DML) and a 14-membered macrolactone ring, narbonolide (NBL). In general, PKSs are megasynthases responsible for the biosynthesis of numerous polyketide natural products including macrolides. PKS organization is comprised of modules divided into distinct catalytic domains with specific roles.[61, 107] Each round of polyketide elongation is performed by three main domains, the acyltransferase (AT) that selects an extender unit, an acyl carrier protein (ACP) and a ketosynthase (KS) that accepts the polyketide chain from an upstream ACP module and catalyzes a decarboxylative Claisen condensation reaction. Modules may contain combinations of other domains including ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) to transform the β-keto group into hydroxyl, alkene or alkane functionalities, respectively.[27, 111, 151] At the end of the pathway, a terminal module, typically a TE domain, catalyzes formation of the cyclized product.[151] Developing these enzymes as biocatalysts represents a complementary addition to synthetic chemistry and encompasses a sustainable and environmentally responsible approach to molecular diversification.[62, 65, 66] Thus, enhancing our understanding and improving Pik TE catalysis for macrocyclization represents a priority area to explore scalable applications of PKS systems.

Previously, we demonstrated the ability to form 12-membered macrolactones utilizing novel truncated pentaketides in conjunction with the PikAIII-TE fusion protein.[62] Depending on substrate structure, the cyclized products were obtained in 9%-66% yields, with hexaketide hydrolysis and truncated byproducts formed when using PikAIII-TE WT[62] and PikAIII-TE S148C mutant[65]. The results indicated that Pik TE has limited substrate flexibility and often functions as a gatekeeper in the processing of unnatural substrates.[62, 65, 66] Although diverse unnatural

macrolactones of different ring sizes have been generated, the introduction of functional groups not typically found in polyketide chain elongation intermediates often results in failed or inefficient cyclization. Thus, we were motivated to deliberately engineer and expand the ability of Pik TE to process structurally variant polyketide substrates. Protein engineering has been used widely to obtain enzymes with new and finely tuned properties[152-155] but has not been applied to polyketide macrocyclization catalysts. In this work, we pursued directed evolution utilizing an amide hexaketide (**3.1**) to create and assess a library of Pik TE variants leading to efficient formation of a hybrid macrolactone/lactam (**3.2**) derivative of 10-DML.[61, 112, 156, 157] We identified new variants with enhanced selectivity for macrocyclization, which was demonstrated by increased total turnover numbers (TTNs), initial reaction rates and isolated yields.

III.3 Chemical Synthesis of Amide Hexaketide and Site Saturation Mutagenesis (Round 1)

Since our previous studies had focused on simplifying the native hexaketide and altering the stereochemistry of nucleophilic hydroxyl group,[62, 65] we decided to explore the impact of replacing one of the central carbon atoms to NH, creating an amide functionality in the substrate. Within FDA approved pharmaceuticals, amides hold significant importance, with examples such as β-lactam antibiotics including penicillins and cephalosporins.[158] Amide functionality is prevalent in medicinal agents, accounting for over half of the reported compounds.[159-162] Various methods are employed for their synthesis, including dehydrative condensations between carboxyl and amino groups, utilization of condensation reagents, Beckmann rearrangements, and Schmidt reaction mechanisms.[163, 164] To further pursue diversification of polyketide systems, we decided to synthesize an amide-containing Pik hexaketide analog to explore TE-mediated cyclization to produce a hybrid ring system.

The target amide hexaketide (**3.1**) was synthesized starting with Evan's aldol condensation[165] between (*R*)-4-benzyl-3-propionyloxazolidin-2-one (**3.3**)[166, 167] and aldehyde **3.4** (generated from (*S*)-3-(Boc-amino)-2-methylpropionic acid) employing dibutylboron triflate to give **3.5**. Next, lithium hydroxide cleavage of the Evan's auxiliary[120, 168] afforded **3.6**, followed by trimethylsilyl (TMS) diazomethane esterification to make **3.7**. Subsequent deprotection of *N*-Boc enabled coupling with **3.8** to yield amide **3.9**. Hydrolysis with lithium hydroxide (**3.10**) followed by thioesterification and final cleavage of the silyl ether[120] with hydrofluoric acid (HF) generated the amide hexaketide (**3.1**) (**Figure III.1**).

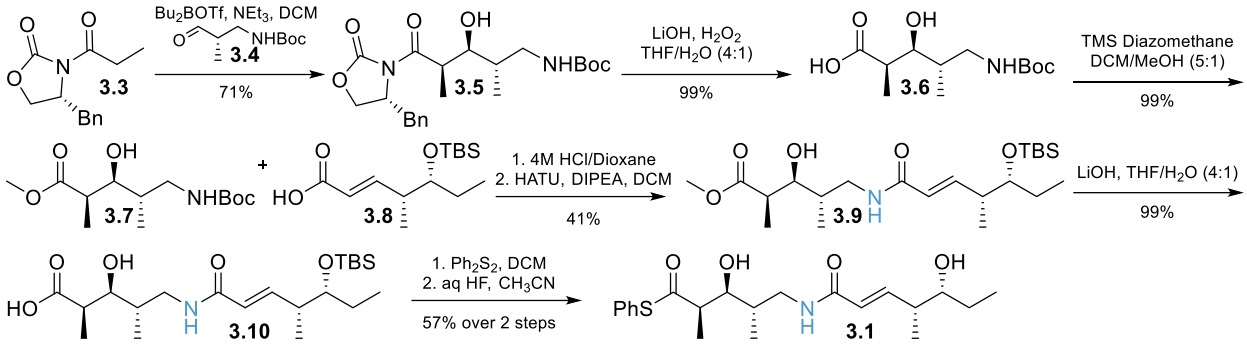


Figure III.1 Chemical synthesis of amide hexaketide (**3.1**).

III.3.1 Isolation of Novel 12-membered Macrolactone/Lactam Ring

With the desired substrate **3.1** in hand, we tested it against Pik TE WT and the Pik TE S148C active site mutant.[65] Pilot scale reactions indicated that Pik TE WT catalyzed formation of the amide hydrolysis product [M+H] = 302 m/z. Although reaction with Pik TE S148C also resulted in majority formation of the amide hydrolysis product, it was accompanied by the mass expected for the 12-membered macrolactone/lactam [M+H] = 284 m/z, suggesting potential formation of the desired macrocycle (**Figure III.7**). Subsequent scale-up, high-performance liquid chromatography (HPLC) purification and isolation provided a 13% isolated yield for the

macrocyclic compound **3.2** confirmed by NMR (**Table III.1**) and X-ray crystallography, which verified the molecular architecture and stereochemistry of the product (**Figure III.32**). Thus, the Pik TE S148C active site mutation was able to catalyze formation of the desired 12-membered ring, albeit in poor overall yield. As a next step, we decided to pursue directed evolution to create additional Pik TE S148C variants with enhanced selectivity for the desired macrocycle **3.2**.

III.3.2 Pik TE Directed Evolution Strategy

Our objective for reactions using Pik TE S148C with the amide hexaketide substrate **3.1** was to obtain incremental improvements in macrocycle formation with a corresponding decrease in hydrolysis products (**Figure III.2a**). To guide the first round of evolution, we examined the Pik TE WT X-ray crystal structure[36] with a covalent pentaketide affinity label (PDB: 2HFJ), and twenty-one active site resides were identified within 6 Å of the substrate (**Figure III.2b**). The crystal structure-substrate mimic complex likely represents a close approximation to the native hexaketide substrate covalently bound to Ser148 in the TE active site showing the Ser148, Asp176, His268 catalytic triad, and revealed the residues likely involved in binding and cyclization of **3.2**. Next, site saturation mutagenesis was employed to create thousands of active site mutants by changing each selected residue to all twenty canonical amino acids.[36, 169, 170] Screening 2,016 enzymatic reactions, we identified five single amino acid Pik TEs S148C variants (L29M, R160C, A217T, A217Y, G222V) with improved yields. These variants were further confirmed by scale-up protein production, purification and enzymatic reactions analyzed via liquid chromatography-mass spectrometry (LC-MS). The Pik TE S148C with additional single mutations L29M, A217T and A217Y were identified at the expected 6 Å distance. However, the corresponding G222V and R160C mutants were located 12 Å and 20 Å respectively, distal to the active site as unexpected non-silent mutations. Surprisingly, these amino acid changes resulted in enhanced ring formation.

Next, we explored combinations of beneficial mutations from round 1 (R1), and found that the Pik TE S148C, A217T, G222V variant showed 3-fold enhanced activity (39% yield based on LC-MS) compared to the original Pik TE S148C enzyme (**Figure III.3a**). This triple-mutant form of Pik TE was generated by combining S148C, with mutation A217T within 6 Å and the unexpected non-silent mutation G222V located 12 Å from the active site region.

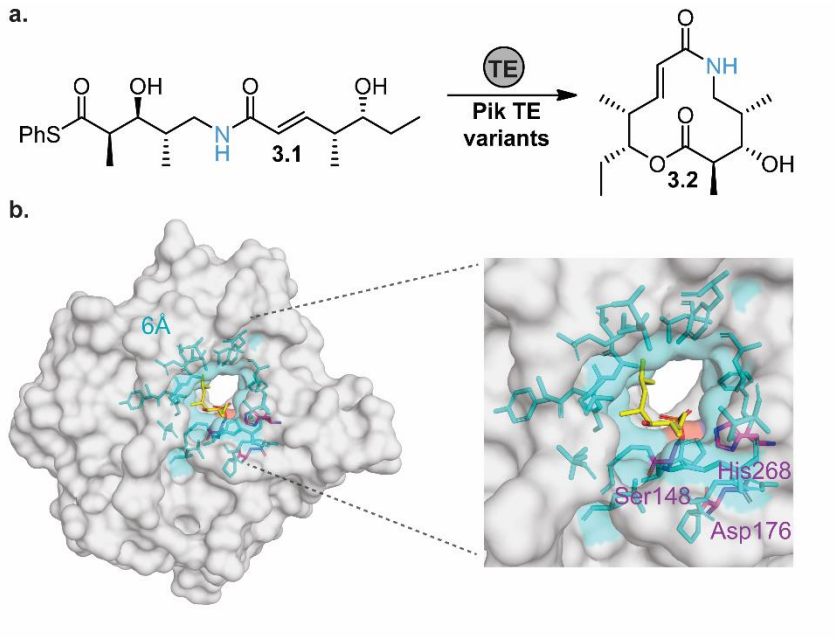


Figure III.2 a. Amide hexaketide **3.1** reaction with diverse Pik TE variants to produce **3.2**. b. Pik TE WT with covalent pentaketide affinity label (PDB: 2HFJ) in yellow. Twenty-one residues identified within 6 Å of the substrate in cyan. Active site catalytic triad in purple (Ser148, Asp176, His268).

III.4 Error Prone PCR (epPCR) Methodology (Rounds 2-3) and Generation of a TE Enzyme Mutant Library

Next, we conducted a second round of evolution focused on improvement of the optimal first round combined variant Pik TE S148C, A217T, G222V. Based on the beneficial G222V we were

motivated to pursue random mutagenesis in regions more distal to the active site. To this end, we employed an epPCR strategy to create random mutations across the *pikIV* TE gene.[170, 171] In round 2 (R2), we screened 2,880 enzymatic reactions and identified sixteen variants with improved production of the hybrid macrolactone/lactam **3.2**. Assessing the individual mutations revealed that the distances of mutated amino acid residues relative to the active site varied from 6 - 20 Å. This outcome demonstrated that critical residues impacting Pik TE selectivity, efficiency and turnover can reside at relatively remote distances from the active site. The identified variants with improved selectivity included single, double, and triple mutations added to the original Pik TE R1 variant. These were selected for enzyme production, purification, and confirmation of enhanced biocatalytic activity, similar to our approach during R1. For R2, we generated the combination TE variant comprised of six mutations (Pik TE S148C, A217T, G222V, M271V, Y25C, L126V) with a 5.5-fold improvement (71% yield based on LC-MS) (**Figure III.3b**).

We subsequently reiterated epPCR random mutagenesis utilizing the R2 variant for a third round of evolution. After screening 1,536 enzymatic reactions, six single amino acid variants (all ~20 Å from active site) were identified with potential improved ring formation. However, further scaleup revealed similar yields (71% yield based on LC-MS) compared to R2. Although two combination mutants were generated, including the combination mutant S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W from round 3 (R3) containing 12 mutations total by adding R2 and the six hits or mutations obtained in R3, none resulted in a yield increase (**Figure III.3a**).

To demonstrate efficiency of the engineered Pik TE variants at a larger scale, we conducted reactions utilizing ~10 mg of the amide hexaketide substrate **1**. HPLC purification and isolation of

the desired cyclic product afforded 42% (Pik TE R1), 79% (Pik TE R2) and 69% (Pik TE R3) isolated yields (**Figure III.3c**).

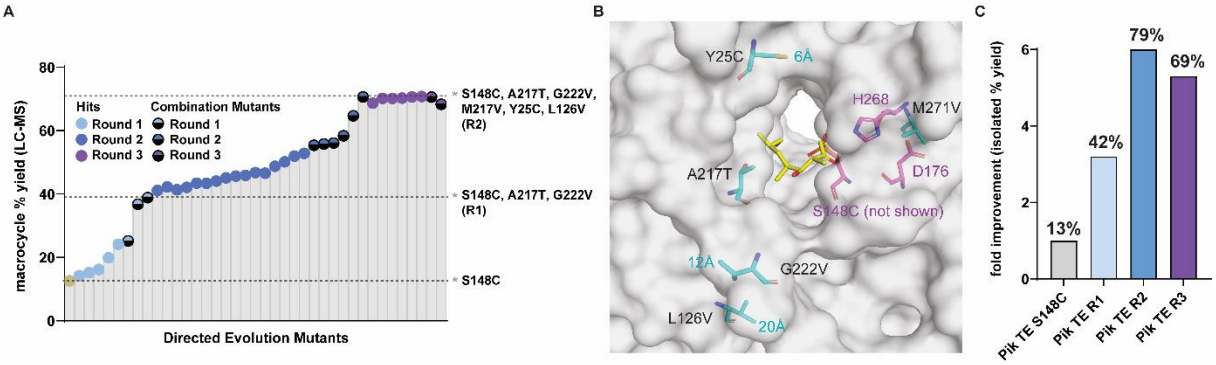


Figure III.3 a. Pik TE variants characterized from directed evolution. b. Pik TE WT crystal structure (PDB: 2HFJ) with mutations in cyan depicting Pik TE R2 using PyMOL mutagenesis tool. Pentaketide mimic affinity label in yellow and the TE catalytic triad in purple. c. Fold improvement for Pik TE mutants in terms of isolated yields (%) compared to the original Pik TE S148C (13%).

III.5 Total Turnover Numbers (TTNs), Thermal Shift Assay and Initial Rates for Substrate Consumption and Macrolactone Formation

Additional analysis of the Pik TE S148C, R1, R2 and R3 mutants, including TTNs, thermal shift assay melting temperatures (Tm's), and initial rates for the reactions, revealed individual variant's stability and catalytic efficiency.[172-174] Our results demonstrated an increase in TTN for cyclized product compared to Pik TE S148C (**Figure III.4a**). Notably, for the TTNs, both Pik TE R2 and R3 retained the highest amount of starting material among the four enzymes, suggesting potential enzyme destabilization[175] over the 18-hour reaction period and 0.2 mol % reduced biocatalyst loading (**Figure III.16**). Interestingly, Pik TE R1 and R3 exhibited highly similar TTN values, yet Pik TE R1 showed the lowest remaining substrate compared to R3. We then performed

a thermal shift assay using SYPRO orange dye to determine protein stability by thermal denaturation of the mutants.[173] The results indicated that both Pik TE WT and S148C had highly similar Tm values, whereas Pik TE R1 showed a Tm improvement of 0.6 ºC. Pik TE R2 and R3 showed a decrease in Tm by 5.0 ºC and 3.8 ºC respectively, suggesting enzyme denaturation at lower temperatures and subsequently lower protein stability compared to Pik TE WT, S148C and R1 (**Figure III.4a**). This data helps explain the results obtained for TTN values, since R1 as the more stable protein had the highest TTN value due to greater stability during the 18-hour TTN reaction time and 0.2 mol % reduced enzyme loading. Pik TE R2 and R3 TTN reactions had considerable amounts of starting material remaining, consistent with destabilization of the mutant proteins under the TTN experiment conditions (**Figure III.16**).

Additionally, a time course analysis was conducted with all four Pik TE mutants, illustrating the differences in selectivity for generating the hybrid macrolactone/lactam and hydrolysis products over three rounds of directed evolution (**Figure III.4b**). Initial rates for substrate consumption were determined for all Pik TEs, revealing similar values across the mutants, suggesting that higher concentrations of protein (1.0 mol %) and substrate have an impact on the reaction rate compared to TTN values. Likewise, initial rates for ring formation were determined, with increasing numbers for the evolved variants (**Figure III.4a**). Compared to the initial rate for Pik TE S148C, Pik TE R1 has a 3.5-fold improvement, followed by Pik TE R2 and Pik TE R3 with a ~14-fold enhancement on initial rates for cyclization. These results show the increase in selectivity for our evolved Pik TE variants for macrocyclization (**Figure III.4b**). The data also indicates that despite the TTN results and thermal shift assay showing Pik TE R2 and R3 variants becoming more unstable over time, addition of 1.0 mol % of the enzyme hastens the reaction

sufficiently to generate high quantities of the desired macrocyclic product without compromising the % yield.

a.

Variant	Total Turnover Number (TTN)	Melting Temperature (Tm, °C)	Initial Rate (μM/min) Substrate Consumption	Initial Rate (μM/min) Macrolactone Formation
Pik TE S148C	11.50 ± 2.70	43.79 ± 0.05	16.20 ± 0.11	0.41 ± 0.13
Pik TE R1	89.26 ± 0.43	44.46 ± 0.06	13.35 ± 0.32	1.42 ± 0.04
Pik TE R2	59.10 ± 6.60	38.85 ± 0.02	15.99 ± 0.29	5.92 ± 0.02
Pik TE R3	75.48 ± 8.61	40.06 ± 0.05	13.94 ± 0.27	5.85 ± 0.07

b.

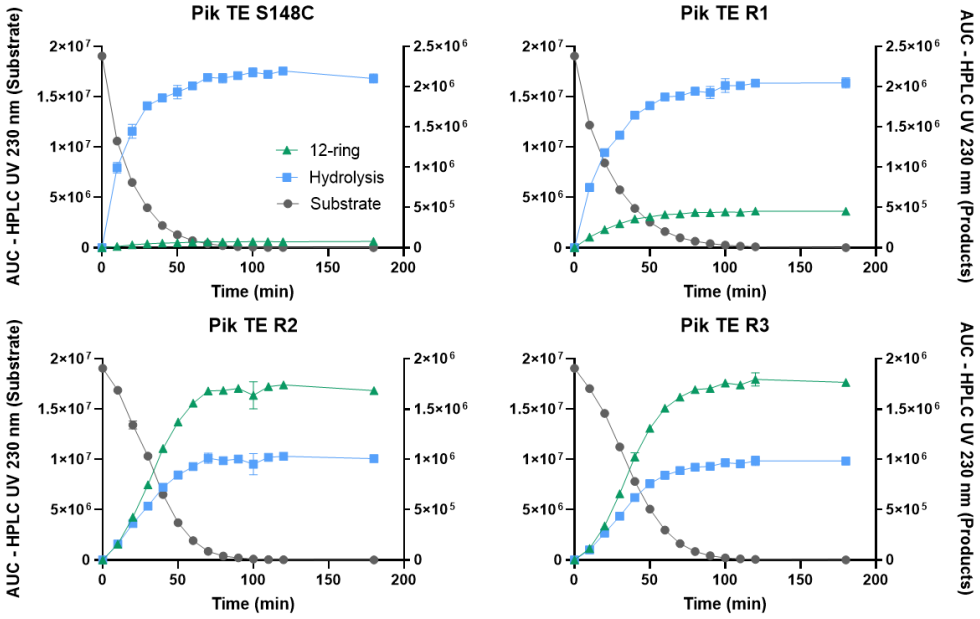


Figure III.4 a. Analysis for total turnover numbers (TTNs), melting temperatures (Tm, °C), initial rates (μM/min) substrate consumption and macrolactone formation for each Pik TE variant. b. Reaction time courses: substrate and products HPLC peaks shown as area under the curve (AUC) as a function of time (min). See experimental section for more details.

III.6 Pik TE Variants Cysteine 148 to Serine Reversions

Finally, to query the importance of the Ser→Cys active site mutation in the Pik TE R1 variant, it was reverted to the WT catalytic triad Ser148. The enzymatic reaction with the amide hexaketide

and Pik TE R1 C148S$_{reversion}$ showed a >10-fold decrease in ring formation (39% to 3% yield) by LC-MS analysis. This result confirms that the S148C mutation is a critical starting point for improved activity in the Pik TE R1 biocatalyst. Moreover, combining the S148C, A217T and G222V variants within the Pik TE R1 revealed synergistic effects of the combined individual mutations. Similarly, we reverted the Pik TE R2 variant back to the C148S WT residue, generating Pik TE R2 C148S$_{reversion}$. Testing the amide hexaketide against the Pik TE R2 reversion mutant resulted in a decrease from 71% to a 14% yield (LC-MS) (**Figure III.5a**). Despite the attenuated activity of the reversion mutants, which confirmed Pik TE S148C as foundational to success of the current strategy, these results demonstrate gain-of-function for the Pik TE WT, and could be a potential future avenue to independently generate the hybrid macrocyclic product **2** through further directed evolution (**Figure III.5b**).

a.

Variant	Ring Product (%)	Hydrolysis Product (%)
Pik TE WT	No product	100.00 ± 0.00
Pik TE R1 C148S$_{reversion}$	2.69 ± 0.01	97.31 ± 0.01
Pik TE R2 C148S$_{reversion}$	13.51 ± 0.15	86.49 ± 0.15

b.

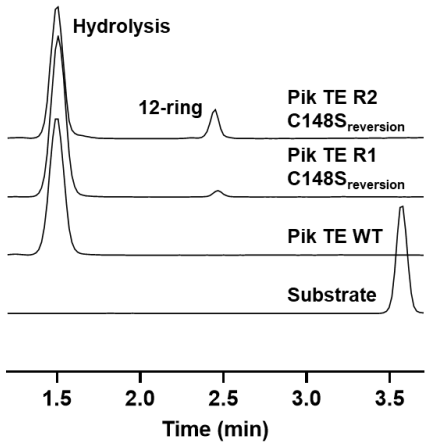


Figure III.5 a. LC-MS % yields and b. traces for Pik TE WT, Pik TE R1 C148S$_{reversion}$ and Pik TE R2 C148S$_{reversion}$ highlighting a gain of function for the WT enzyme.

III.7 Discussion

Macrolactonization represents a challenging transformation in organic synthesis, including unfavorable energetics of intramolecular reactions to form rings with more than six atoms. Although the generation of macrolactones and macrolides by total and semi-synthesis is feasible, limitations to both approaches including the large number of synthetic steps, low yields, difficulties for selective addition of diverse functional groups and the inherent reactivity of large macrocycles represents an area to explore with chemoenzymatic synthesis.[101, 102, 129, 131, 137, 141, 142] The discovery of modular type I PKSs that assemble complex natural products like erythromycin and pikromycin allowed us to identify enzymes within their pathways that catalyze macrolactonization.[61, 62, 65, 112, 113] In these biosynthetic systems, the PKS TE domain catalyzes formation of 12- and 14-membered ring structures.[151] In previous studies, the Pik TE was shown to be a catalytic bottleneck for the desired ring closing reaction and was identified as a gatekeeper for processing unnatural substrates, hence impeding structural diversification and production of new macrocycles.[62, 65, 66] This work demonstrates the value of TE directed evolution for macrocycle bioengineering to create new macrocyclic compounds.

Previously, the Pik TE S148C mutation revealed enhanced reaction kinetics and gain of function for the processing of an unnatural hexaketide diastereomer, which led to a shift in reaction mechanism between the WT and S148C variant as determined with quantum mechanical (QM) models.[65] In our case, the S148C mutation is foundational for the generation of the hybrid 12-membered macrolactone/lactam ring, whereas the Pik TE WT alone leads to exclusive formation of the seco acid amide hexaketide. With only two rounds of evolution, the Pik TE R2 provided a 6-fold improvement in isolated % yield compared to the Ser148Cys parent enzyme, enabling construction of a novel cyclic ring system with high efficiency and selectivity (**Figure III.3c**).

As a starting point for our work, we benefitted from the Pik TE/substrate mimic crystal structure complex,[36] which guided us to identify residues within the active site region, where site saturation mutagenesis was employed in a first round of evolution (**Figure III.2b**). Pik TE R1 contains the A217T mutation (6 Å from active site) that was shown in the Pik TE structure with two well-ordered water molecules coordinated between side chains Gln183 and backbone carbonyl of A217, and the C5-hydroxyl group of the pentaketide substrate mimic. This region was predicted to form a 'hydrophilic barrier' to direct the acyl chain back into the TE substrate channel's hydrophobic environment and guide the nucleophilic hydroxyl group toward active site the serine-ester bond.[36] We reason that the A217T could now have an additional coordination site with the threonine side chain via hydrogen bonding interactions with the nucleophilic hydroxyl group. Although the G222V mutation is 12 Å from the active site, increasing non-polar or hydrophobic effects of the more distal enzyme regions could induce conformational changes near the active site that promote cyclization and stability (**Figure III.3b**). Other studies have similarly shown that mutating residues relatively distant from the active site can improve enzymatic activity.[155, 176-178]

The epPCR random mutagenesis strategy enabled us to investigate residues from the entire Pik TE sequence. Over the course of this study, sixteen mutants were identified that improved product yields compared to variants generated during the first round of directed evolution. Additional synergistic effects were observed when combining three of the mutations M271V, Y25C, L126V with Pik TE R1 (**Figure III.3a**). Residue Y25 was shown in the Pik TE crystal structure to engage in direct interactions with the 12-membered cyclized product 10-DML. Therefore, we propose that Cys25 might promote polar contacts with the unnatural substrate that favor product formation. Mutation M271V is positioned behind the Asp176 and His268 catalytic triad residues 12 Å distal to the structure bound substrate, whereas L126V is 20 Å away, to the opposite side of the triad

(**Figure III.3b**). These valine mutations in Pik TE R2 could have hydrophobic interactions that create a more hospitable active site pocket and nearby regions of the catalytic site, allowing macrocyclization to occur more readily and excluding water that alleviates competing TE mediated hydrolysis.[179]

Regarding R3 mutations, all resided 20 Å distal to the active site located at the surface of the TE and failed to provide higher production of the desired macrocycle, thus representing an evolutionary plateau. Currently, our best yielding mutant Pik TE R2 (79% isolated yield), contains a total of three valines, two cysteines and one threonine (**Figure III.3b**). Enzyme performance limits with directed evolution have been reported previously.[180-182] Although predicting the potential benefits of additional mutations or rounds of evolution can be difficult, exploring existing mutations and understanding their impact on the enzyme's active site conformation, substrate binding, reaction mechanism, including the effects of individual mutations and combinations for overall protein stability could be informative for future studies.[183-185] In addition, complementary or parallel trajectories of individual variants can be pursued to maximally improve enzyme activities.[180, 181, 186, 187] Another option is ultrahigh-throughput droplet-based microfluidic screening to identify further optimized enzymes.[188]

TTN values for the Pik TE variants were experimentally shown to be Pik TE R1 > Pik TE R3 > Pik TE R2 > Pik TE S148C (**Figure III.4a**). We hypothesized that this trend may be attributed to enzymatic destabilization with increased ring selectivity for R3, resulting in elevated TTN values comparable to R1. Thus, although Pik TE R1 has a lower selectivity for macrocyclization, it remains active for a longer period.[189, 190] Although some degree of enzyme destabilization was observed in the thermal shift assay (**Figure III.4a**), it did not significantly impact the enzymatic reactions at both analytical and scale-up levels. This was mainly due to the higher enzyme

concentrations (1.0 mol %) and complete conversion of the substrate within one hour, overcoming possible destabilization challenges present in longer reaction times. In comparison, the initial rates for substrate consumption by the Pik TE S148C, R1, R2, and R3 mutants showed similar μM/min values, indicating an increase in protein concentration and substrate had an impact on the rate. Moreover, initial rates for compound **2** ring formation demonstrated an increase in enzymatic activity and shift in selectivity for the engineered TEs (**Figure III.4a**). In the reaction time courses, Pik TE R2 and R3 showed an initial increase in ring and hydrolysis products at the onset of the reaction and then ring formation surpassed seco-acid formation, which could relate to the enzyme mechanism parameters that influence hydrolysis or cyclization (**Figure III.4b**). In a previous study, we showed that Pik TE WT followed a stepwise addition-elimination mechanism with transient formation of a tetrahedral intermediate. By contrast, Pik TE S148C followed a concerted acyl substitution process, which is lower in energy.[65] The new mutations from our directed evolution campaign could have an impact on the biocatalyst energetic barriers in comparison to both Pik TE WT and S148C.[36, 150, 179, 191-197]

The Pik TE Ser→Cys catalytic triad mutation was also identified as fundamental in Pik TE R1 and Pik TE R2 variants as shown in reversion to the WT Ser148. Restoration of the Ser active site led to a >10-fold decrease in macrocycle formation (39% to 3% yield) in R1 and 5-fold decrease for R2 (71% to 14% yield) (**Figure III.5**). Overall, the current study represents a proof-of-concept approach for relieving the PKS TE bottleneck in the formation of new macrocycles. The results demonstrate that directed evolution improved selectivity and efficiency in the catalysis of an unnatural amide-containing hexaketide substrate, and offers significant promise for expanding to other substrates with diverse functionalities, aiming to generate novel macrolactones, macrolactams, depsipeptides, and other TE-mediated biosynthetic systems. For our case, future

computational analysis through QM, molecular dynamics (MD) calculations and machine learning (ML) approaches[198-203] could provide new mechanistic insights relating to mutations identified in our screen, as well as suggest additional amino acids that may modulate selectivity in TEs beyond the amide hexaketide substrate.

III.8 Conclusion/Summary

In summary, we have demonstrated the ability of the Pik TE to accept an unnatural amide hexaketide substrate for the generation of a hybrid 12-membered macrolactone/lactam ring. We employed directed evolution to engineer the Pik TE, starting from the previously reported Pik TE S148C mutant[65] for an increase in macrocyclization from 13% up to 79% isolated yields. With three rounds of evolution, we generated and screened >6,000 mutants through site saturation mutagenesis focused on the active site region (R1), and random mutagenesis covering the entire protein sequence (R2, R3). We initially identified residues of potential importance for cyclization using the previously reported Pik TE WT crystal structure with a covalently bound pentaketide mimic[36], which enabled target residues of interest to be identified for directed evolution. We expect that applying computational and ML-based approaches[65, 198-201] will deepen our mechanistic understanding of the cyclization process and the impact of select mutations on catalytic efficiency. Moreover, our Pik TE variant library will provide an ongoing resource to identify efficient biocatalysts for varied hexaketide/heptaketide substrates with the goal of obtaining novel bioactive macrolides and related molecules.

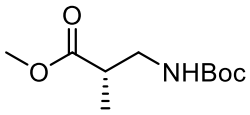
III.9 Experimental

III.9.1 General Information, Materials and Instrumentation

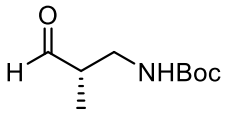
As described in Chapter 2.

Synthetic Experimentals and Characterizations

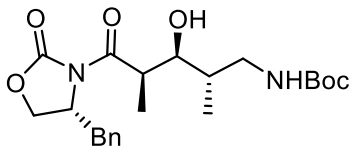
III.9.2 Amide Hexaketide Synthetic Procedure



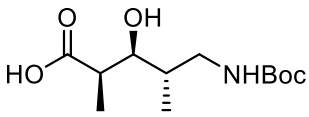
methyl (*S*)-3-((*tert*-butoxycarbonyl)amino)-2-methylpropanoate (3.0). To commercially available (*S*)-3-(Boc-amino)-2-methylpropionic acid (0.14 g, 0.67 mmol, 1 eq) suspended in 4:1 CH_2Cl_2/MeOH (20 mL) was added trimethylsilyl (TMS) diazomethane (2 M solution in diethyl ether, 0.15 mL, 1.5 eq) dropwise to form a yellow solution. The reaction was allowed to stir for 10 min at rt, quenched with acetic acid (0.05 mL) until the solution was again clear. The reaction was concentrated to give **3.0** (0.145 g, 99% yield) as a clear and colorless oil. All spectra were in accordance with published literature.[204]



***tert*-butyl (*S*)-(2-methyl-3-oxopropyl)carbamate (3.4)**. A clear and colorless stirring solution of **3.0** (0.40 g, 1.84 mmol, 1 eq) suspended in CH_2Cl_2 (20 mL) was cooled to -78 °C (dry ice/Acetone) and treated slowly with DIBAL (1 M in Heptane, 5.52 mL, 3 eq). The reaction was kept at -78 °C for 1 h, judged to be complete by TLC and quenched at -78 °C by the addition of 10 mL of MeOH. This was warmed to 0 °C and poured into an ice cold vigorously stirring solution of sat. sodium potassium tartrate (50 mL). This was allowed to stir warming to rt for 1 h. The aqueous layer was partitioned and extracted 3 x 40 mL with CH_2Cl_2. The organics were combined, dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified using automated chromatography (5% EtOAc in hexanes to 60% EtOAc in hexanes) to give **3.4** (265 mg, 75% yield) as a clear and colorless oil: R_f = 0.4 (20% EtOAc/hexanes); ^{1}H NMR (600 MHz, $CDCl_3$) δ 9.68 (s, 1H), 4.84 (s, 1H), 3.46 – 3.19 (m, 2H), 2.63 (q, *J* = 7.0 Hz, 1H), 1.43 (s, 9H), 1.15 (d, *J* = 7.4 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 204.14, 156.07, 79.65, 47.37, 40.76, 28.50, 11.44. HRMS (ES) calcd for $C_9H_{17}NO_3$ [M+H] 188.1281, found 188.1648.

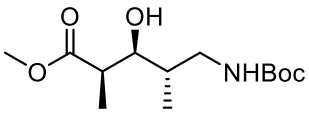


***tert*-butyl ((2*S*,3*S*,4*R*)-5-((*R*)-4-benzyl-2-oxooxazolidin-3-yl)-3-hydroxy-2,4-dimethyl-5-oxopentyl)carbamate** (**3.5**). To a clear and colorless solution of (*R*)-4-benzyl-3-propionyloxazolidin-2-one[166, 167] **3.3** (0.33 g, 1.41 mmol, 1 eq) in CH_2Cl_2 (10 mL) at -78 °C (dry ice/Acetone) was added Bu_2BOTf (1M in CH_2Cl_2, 1.56 mL, 1.1 eq) slowly. To this, triethyl amine was added and the reaction was stirred at -78 °C for 10 min, warmed to 0 °C, stirred for 30 min. The reaction was cooled to -78 °C prior to the addition of **3.4** aldehyde (0.26 g, 1.41 mmol, 1 eq) stirred for 0.5 h maintaining temperature. The reaction was warmed to 0 °C, stirred for 1 h and quenched by the addition of phosphate buffer (1 M, pH = 7, 10 mL), MeOH (5 mL), and 30% aq H_2O_2 (1 mL). The mixture was stirred for 1 h after addition, concentrated and the aqueous phase was extracted with CH_2Cl_2 (3 x 30 mL). The organics were combined, dried over Na_2SO_4 and concentrated. The residue was purified using automated chromatography (16% EtOAc in hexanes to 100% EtOAc in hexanes) to give **3.5** (420 mg, 71% yield) as a clear and colorless oil: R_f = 0.6 (50% EtOAc/hexanes); ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.36 (t, *J* = 7.3 Hz, 2H), 7.31 (d, *J* = 7.4 Hz, 1H), 7.23 (d, *J* = 6.8 Hz, 2H), 5.09 (s, 1H), 4.71 (ddt, *J* = 9.8, 6.1, 3.7 Hz, 1H), 4.22 – 4.19 (m, 2H), 3.92 (qd, *J* = 7.0, 2.4 Hz, 1H), 3.71 (dd, *J* = 9.4, 2.4 Hz, 1H), 3.47 (d, *J* = 14.0 Hz, 1H), 3.32 (dd, *J* = 13.4, 3.4 Hz, 1H), 3.09 (dd, *J* = 13.9, 5.1 Hz, 1H), 2.80 (dd, *J* = 13.4, 9.6 Hz, 1H), 1.84 – 1.73 (m, 1H), 1.45 (s, 9H), 1.26 (d, *J* = 7.2 Hz, 3H), 0.93 (d, *J* = 6.9 Hz, 3H). HRMS (ES) calcd for $C_{22}H_{32}N_2O_6$ [M+Na] 443.2153, found 443.2152.

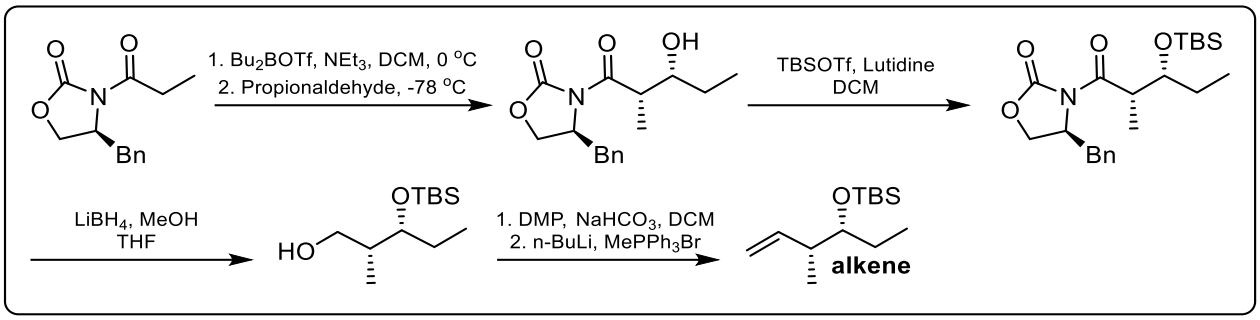


(2*R*,3*S*,4*S*)-5-((*tert*-butoxycarbonyl)amino)-3-hydroxy-2,4-dimethylpentanoic acid (3.6). To a stirred solution of **3.5** oxazolidinone (0.227 g, 0.54 mmol, 1 eq) in THF (4 mL) open to air was added LiOH monohydrate (0.036 g, 0.86 mmol, 1.6 eq) and 30% aq. H_2O_2 (67 µL, 4 eq) as a solution in H_2O (1 mL). The

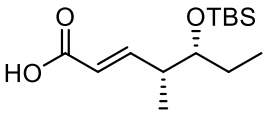
mixture was stirred at rt until consumption of starting material (as judged by TLC, 1 h). The reaction was quenched with saturated sodium sulfite (5 mL) and stirred at rt for 20 min. The organics were removed under reduced pressure and the resulting aqueous solution was acidified to pH = 1 by the addition of 1 N HCl and extracted with CH_2Cl_2 (3 x 20 mL). The organics were combined, dried over Na_2SO_4 and concentrated. The residue was purified using automated chromatography (0% MeOH in CH_2Cl_2+ 1% AcOH to 10% MeOH in CH_2Cl_2) to give **3.6** (140 mg, 99% yield) as a clear and colorless oil: R_f = 0.4 (5% MeOH/CH_2Cl_2); ^{1}H NMR (600 MHz, CD_3OD) δ 3.71 (dd, *J* = 8.1, 4.5 Hz, 1H), 3.22 (dd, *J* = 13.7, 4.2 Hz, 1H), 3.11 (dd, *J* = 13.7, 6.7 Hz, 1H), 2.62 (qd, *J* = 7.0, 4.6 Hz, 1H), 1.70 (ddt, *J* = 11.3, 7.0, 4.3 Hz, 1H), 1.44 (s, 9H), 1.13 (d, *J* = 7.0 Hz, 3H), 0.90 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR (150 MHz, CD_3OD) δ 179.07, 159.04, 80.02, 75.58, 43.86, 43.48, 38.36, 28.76, 15.13, 10.55. HRMS (ES) calcd for $C_{12}H_{23}NO_5$ [M+Na] 284.1468, found 284.1459.



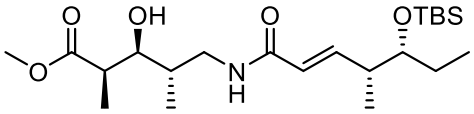
methyl (2*R*,3*S*,4*S*)-5-((*tert*-butoxycarbonyl)amino)-3-hydroxy-2,4-dimethyl pentanoate (3.7). To a stirred solution of **3.6** (0.140 g, 0.54 mmol, 1 eq) in CH_2Cl_2/MeOH (4:1, 8 mL) was added TMS diazomethane (2 M solution in diethyl ether, 0.32 mL, 1.2 eq) dropwise to form a yellow solution. The reaction was allowed to stir for 10 min at rt, quenched with acetic acid (0.05 mL) until the solution was again clear. The reaction was concentrated to give **3.7** (0.145 g, 99% yield) as a light yellow oil: R_f = 0.4 (33% EtOAc/Hexanes); ^{1}H NMR (600 MHz, CD_3OD) δ 4.85 (s, 3H), 3.69 (dd, *J* = 7.8, 5.0 Hz, 1H), 3.68 (s, 3H), 3.22 (dd, *J* = 13.7, 4.3 Hz, 1H), 3.08 (dd, *J* = 13.7, 6.7 Hz, 1H), 2.67 (qd, *J* = 7.0, 4.6 Hz, 1H), 1.71 – 1.61 (m, 1H), 1.44 (s, 9H), 1.13 (d, *J* = 7.0 Hz, 3H), 0.89 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR (150 MHz, CD_3OD) δ 177.37, 159.04, 80.02, 75.56, 52.19, 43.80, 43.68, 38.32, 28.76, 15.10, 10.35. HRMS (ES) calcd for $C_{13}H_{25}NO_5$ [M+Na] 298.1625, found 298.1644.



*Synthesis previously developed by Hansen, D. A.[61]

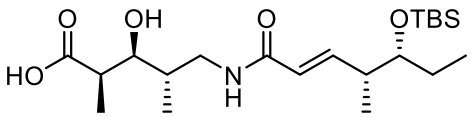


(4*R*,5*R*,*E*)-5-((*tert*-butyldimethylsilyl)oxy)-4-methylhept-2-enoic acid (3.8). To a long tube under N_2 was added Hoveyda Grubbs II (0.03 eq, 0.03 mmol, 0.019 g), 2-pentenoic acid (1 eq, 1.0 mmol, 0.10 g), and **alkene** (4 eq, 4.0 mmol, 0.913 g) in CH_2Cl_2 (2 mL, 0.5 M). This was allowed to reflux for 12 h, cooled, the solvent was removed under reduced pressure and purified directly by flash chromatography (1 – 25% EtOAc/Hexanes + 1% AcOH) to afford **3.8** (0.231 g, 88% yield) as a white solid: $R_f = 0.4$ (20% EtOAc/Hexanes); ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.12 (dd, J = 15.8, 7.3 Hz, 1H), 5.82 (d, J = 15.8, 1.2 Hz, 1H), 3.56 (q, J = 5.7 Hz, 1H), 2.50 (dq, 1H), 1.49 (ddd, J = 12.5, 7.6, 5.5 Hz, 1H), 1.39 (dp, J = 14.2, 7.2 Hz, 1H), 1.04 (d, J = 6.8, 1.0 Hz, 3H), 0.90 (s, 1.0 Hz, 9H), 0.87 (t, 3H), 0.04 (s, 4.1 Hz, 3H), 0.00 (s, 1.0 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 171.56, 155.23, 120.05, 76.40, 41.56, 27.01, 26.02, 18.27, 14.05, 9.82, -4.20, -4.38. HRMS (ES) calcd for $C_{14}H_{28}O_3Si$ [M-H] 271.1755, found 271.1759.



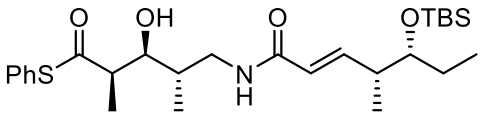
methyl-(2*R*,3*S*,4*S*)-5-((4*R*,5*R*,*E*)-5-((*tert* butyldimethylsilyl)oxy)-4-methylhept-2-enamido)-3-hydroxy-2,4-dimethylpentanoate (3.9). 3.7 (0.042 g, 0.2 mmol, 1 eq) was suspended in 4M HCl/dioxane (2 mL) and allowed to stir at rt for 30 min when judged to be complete by TLC. The reaction was concentrated and used directly. Acid fragment **3.8** (0.054 g, 0.2 mmol, 1 eq) was

suspended in CH_2Cl_2 (2 mL) and treated with HATU (0.083 g, 0.22 mmol, 1.1 eq) and DIPEA (0.086 mL, 0.5 mmol, 2.5 eq). To this was added the deprotected **3.7** in CH_2Cl_2 (1 mL) and the reaction was stirred at rt overnight. The reaction was diluted with water and CH_2Cl_2, the aqueous layer was partitioned, extracted with CH_2Cl_2 (2 x 20 mL), organics combined, dried over Na_2SO_4 and concentrated. The residue was purified using automated chromatography (0% EtOAc in hexanes to 100% EtOAc in hexanes) to give **3.9** (35.4 mg, 41% yield) as a clear and colorless oil: R_f = 0.5 (50% EtOAc/hexanes); ^{1}H NMR (600 MHz, $CDCl_3$) δ 6.82 (dd, *J* = 15.5, 7.5 Hz, 1H), 6.33 (s, 1H), 5.75 (dd, *J* = 15.5, 1.3 Hz, 1H), 3.84 (ddd, *J* = 13.3, 7.5, 3.5 Hz, 1H), 3.73 (dd, *J* = 9.4, 2.9 Hz, 1H), 3.71 (s, 3H), 3.53 (q, *J* = 5.5 Hz, 1H), 3.06 (ddd, *J* = 14.0, 6.3, 3.7 Hz, 1H), 2.63 (qd, *J* = 7.2, 2.9 Hz, 1H), 2.44 (h, *J* = 6.7 Hz, 1H), 1.82 – 1.71 (m, 1H), 1.52 – 1.35 (m, 2H), 1.18 (d, *J* = 7.2 Hz, 3H), 1.02 (d, *J* = 6.8 Hz, 3H), 0.89 (s, 9H), 0.89 (d, *J* = 6.3 Hz, 3H), 0.85 (t, *J* = 7.4 Hz, 3H), 0.03 (s, 6H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 176.99, 166.92, 147.76, 123.12, 76.50, 74.93, 52.11, 43.83, 41.50, 41.20, 36.16, 27.04, 26.06, 18.30, 15.07, 14.67, 9.58, -4.19, -4.29. HRMS (ES) calcd for $C_{22}H_{43}NO_5Si$ [M+H] 452.2803, found 452.2795.



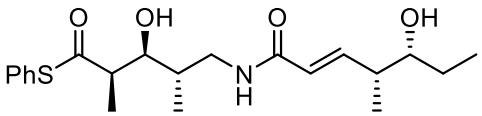
(2*R*,3*S*,4*S*)-5-((4*R*,5*R*,*E*)-5-((*tert*-butyldimethylsilyl)oxy)-4-methylhept-2-enamido)-3-hydroxy-2,4-dimethylpentanoic acid (3.10). To an open 4 d vial was added **3.9** (0.035 g, 0.08 mmol, 1 eq) in THF/H_2O (4:1, 1 mL) and treated with LiOH monohydrate (0.010 g, 0.24 mmol, 3 eq) and the reaction was stirred at rt for 1.5 h. The reaction was diluted with water (10 mL) and acidified to pH = 1 using 1 N HCl. The aqueous layer extracted with CH_2Cl_2 (2 x 20 mL), combined, dried over Na_2SO_4 and concentrated to give **3.10** (33 mg, 99% yield) as a clear and colorless oi: R_f = 0.1 (50% EtOAc/hexanes) ; ^{1}H NMR (600 MHz, $CDCl_3$) δ 6.91 (dd, *J* = 15.4, 7.3 Hz, 1H), 5.94 (t, *J* = 6.7 Hz, 1H), 5.77 (dd, *J* = 15.4, 1.4 Hz, 1H), 4.05 (ddd, *J* = 14.4, 8.3, 3.1

Hz, 1H), 3.59 – 3.54 (m, 2H), 2.94 (ddd, *J* = 14.4, 5.1, 3.6 Hz, 1H), 2.66 (qd, *J* = 7.3, 2.4 Hz, 1H), 2.46 (tdd, *J* = 6.8, 4.9, 1.5 Hz, 1H), 1.92 – 1.77 (m, 1H), 1.56 – 1.35 (m, 2H), 1.24 (d, *J* = 7.3 Hz, 3H), 1.03 (d, *J* = 6.8 Hz, 3H), 0.86 (t, *J* = 7.5 Hz, 3H), 0.04 (s, 3H), 0.02 (s, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 177.24, 168.50, 150.06, 121.81, 76.38, 71.52, 42.85, 41.23, 40.89, 36.02, 27.10, 26.05, 18.30, 14.63, 14.25, 9.95, 9.72, -4.18, -4.30. HRMS (ES) calcd for $C_{21}H_{41}NO_5Si$ [M+H] 416.2827, found 416.2923.



***S*-phenyl-(2*R*,3*S*,4*S*)-5-((4*R*,5*R*,*E*)-5-((*tert*-butyldimethylsilyl)oxy)-4-methylhept-2-enamido)-3-hydroxy-2,4-dimethylpentanethioate (3.10 thiophenol intermediate).** To a stirring solution of **3.10** (0.033 g, 0.08 mmol, 1 eq) and Ph_2S_2 (0.019 g 0.09 mmol, 1.1 eq) dissolved in CH_2Cl_2 (1 mL) cooled to 0 °C was added PBu_3 (22 μL, 0.019 mmol, 1,1 eq). After stirring at 0 °C for 0.5 h, the mixture was quenched by the addition of saturated $NaHCO_3$ (125 mL). The layers were separated, and the aqueous layer was extracted with CH_2Cl_2 (2 x 20 mL). The combined organic layers were dried with Na_2SO_4 and concentrated. The residue was purified using automated chromatography (20% EtOAc in hexanes to 70% EtOAc in hexanes) to give **3.10 thiophenol intermediate** (27.2 mg, 68% yield) as a clear and colorless oil: R_f = 0.3 (33% EtOAc/hexanes); ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.42 (d, *J* = 1.4 Hz, 5H), 6.82 (dd, *J* = 15.5, 7.5 Hz, 1H), 6.31 – 6.20 (m, 1H), 5.75 (dd, *J* = 15.5, 1.3 Hz, 1H), 3.81 (ddd, *J* = 12.2, 8.4, 3.3 Hz, 2H), 3.53 (q, *J* = 5.5 Hz, 1H), 3.11 (ddd, *J* = 14.0, 6.6, 4.5 Hz, 1H), 2.91 (qd, *J* = 7.1, 3.1 Hz, 1H), 2.49 – 2.36 (m, 1H), 1.80 (dddd, *J* = 10.3, 8.7, 6.8, 3.5 Hz, 1H), 1.52 – 1.34 (m, 2H), 1.30 (d, *J* = 7.1 Hz, 3H), 1.01 (d, *J* = 6.8 Hz, 3H), 0.93 (d, *J* = 6.9 Hz, 3H), 0.88 (s, 9H), 0.85 (t, *J* = 7.4 Hz, 3H), 0.03 (s, 6H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 202.65, 166.84, 147.74, 134.70, 129.70, 129.39, 127.33, 123.18,

76.50, 75.42, 50.24, 43.60, 41.19, 36.49, 27.04, 26.06, 18.29, 15.41, 14.71, 10.98, 9.60, -4.20, -4.27. HRMS (ES) calcd for $C_{27}H_{45}NO_4SSi$ [M+H] 530.2911, found 530.2913.



_S_-phenyl-(2_R_,3_S_,4_S_)-3-hydroxy-5-((4_R_,5_R_,_E_)-5-hydroxy-4-methylhept-2-enamido)-2,4-dimethylpentanethioate (3.1). To a stirring solution of **3.10 thiophenol intermediate** (0.022 g, 0.043 mmol) dissolved in CH_3CN (3 mL) in a polypropylene tube stirring at rt was added 48% HF in water (0.16 mL, 4.4mmol). The mixture was stirred for 2 h and poured into a chilled (ice bath) solution of 10% $NaHCO_3$ and stirred for 20 min. The mixture was extracted with CH_2Cl_2 (3 x 20 mL), dried with Na_2SO_4 and concentrated to a clear and colorless oil. The residue was purified using automated chromatography (20% EtOAc in hexanes to 70% EtOAc in hexanes) to give **3.1** (12.3 mg, 73% yield) as a clear and colorless oil: R_f = 0.3 (33% EtOAc/hexanes); ^{1}H NMR (600 MHz, $CDCl_3$) δδ 7.46 – 7.38 (m, 5H), 6.81 (dd, *J* = 15.4, 7.8 Hz, 1H), 6.33 (d, *J* = 7.2 Hz, 1H), 5.81 (dd, *J* = 15.3, 1.1 Hz, 1H), 3.85 – 3.75 (m, 2H), 3.48 (ddd, *J* = 8.9, 5.3, 3.8 Hz, 1H), 3.12 (ddd, *J* = 14.0, 6.6, 4.2 Hz, 1H), 2.91 (qd, *J* = 7.1, 3.0 Hz, 1H), 2.47 – 2.33 (m, 1H), 1.80 (tdd, *J* = 10.3, 8.7, 5.2 Hz, 1H), 1.55 (dqd, *J* = 14.9, 7.5, 3.8 Hz, 1H), 1.39 (ddq, *J* = 14.4, 8.9, 7.3 Hz, 1H), 1.31 (d, *J* = 7.1 Hz, 3H), 1.08 (d, *J* = 6.8 Hz, 3H), 0.96 (d, *J* = 7.1 Hz, 3H), 0.94 (t, *J* = 7.2 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 202.85, 166.51, 147.05, 134.70, 129.76, 129.42, 127.26, 123.97, 76.11, 75.49, 50.14, 43.78, 42.14, 36.38, 27.43, 15.39, 14.25, 10.92, 10.46. HRMS (ES) calcd for $C_{21}H_{31}NO_4S$ [M+H] 394.2047, found 394.2088.

Spectra (geometry: lock with width 20.01, height 17.45. proton from -0.9 to 12 carbon -10 to 218 ppm)

III.9.3 Protein Biochemistry

Protein Expression and Purification Pik TE WT and Pik TE S148C

Expression

Plasmid DNA[65, 112] for Pik TE WT or Pik TE S148C were transformed into chemically competent BL21(DE3) expression cells, plated on LB agar with kanamycin (50 µg/mL) plates and incubated overnight at 37 ºC to obtain colonies the next day. Single colonies were picked for overnight cultures in LB media with kanamycin (50 µg/mL) and grown overnight at 37 ºC, 180 rpm. Protein expression of enzymes were done by adding 5 mL of overnight cultures to 1 L of TB media, containing TB salts (0.17 M KH_2PO_4 and 0.72 M K_2HPO_4) with kanamycin (50 µg/mL) and incubated at 37 ºC, 180 rpm. After OD_{600} reached 1 – 1.5, these were transferred to an ice bath for 15 mins, induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated overnight at 18 ºC, 180 rpm. The following day, the mixture was centrifuged at 6,000 x g for 30 mins, the supernatant was discarded, and the cell pellet stored in -80 ºC for further protein purification.

Purification

The pelleted cells were resuspended in 5 mL of lysis buffer per gram of cells and vortexed or stirred at 4 ºC. Cell lysis was achieved by adding 1 mg/mL lysozyme, 0.1 mg/mL DNAse and then applied sonication (2 mins total: 10 sec on, 50 sec off protocol). The resulting cell lysate was pelleted by centrifugation at 50,000 x g, 30 mins, 4 ºC. The supernatant (lysate) was filtered using 0.45 µm filters and added to a pre-equilibrated Ni-NTA resin column (equilibration: 10 CV wash buffer). After protein binding, the column was washed with 10 CV of wash buffer. Proteins were eluted with 5 CV of elution buffer and collected as 1-1.5 mL fractions. The fractions were analyzed for their absorption at 280 nm, combined/concentrated to 2.5 mL using 10 K MWCO Amicon Ultra-15 centrifugal filters and 3.5 mL of storage buffer exchanged using a pre-equilibrated PD10

desalting column (equilibration: 5 CV storage buffer). Upon completion of buffer exchange, the protein samples were analyzed once again for their absorption at 280 nm to determine their concentration, frozen in liquid nitrogen and stored at -80 ºC. Samples were run on a NuPAGE 4-12% Bis-Tris protein gel to check for purity.

Buffers for protein purification:

Lysis: HEPES (50 mM), NaCl (300 mM), imidazole (10 mM), glycerol (10% v/v), pH 8.0

Wash: HEPES (50 mM), NaCl (300 mM), imidazole (30 mM), glycerol (10% v/v), pH 8.0

Elution: HEPES (50 mM), NaCl (300 mM), imidazole (300 mM), glycerol (10% v/v), pH 8.0

Storage: HEPES (50 mM), NaCl (150 mM), EDTA (1 mM), glycerol (20% v/v), pH 7.2

Enzymatic Evaluation Pik TE WT and Pik TE S148C with Hexaketide Substrate Analytical Enzymatic Reactions

Analytical scale 50 µL enzymatic reactions (**Figure III.6**) were done by adding water, sodium phosphate reaction buffer (400 mM NaH_2PO_4•H_2O and Na_2HPO_4, 200 mM NaCl, pH 7.2), 1 mM amide hexaketide (substrate) and lastly 10 µM of desired Pik TE WT or Pik TE S148C, overnight, stationary, rt. As negative controls, enzymes were omitted. All reactions were performed and analyzed in triplicates. Next day, the reactions were quenched by adding 3 volumes of MeOH (150 µL) and centrifuged at 13,000 x g, 10 mins, 4 ºC for further LC-MS analysis. LC-MS conditions were as follows: mobile phase (A = deionized water + 0.1% formic acid, B = 95% acetonitrile/deionized water + 0.1% formic acid); 10% B for 1.0 min, 10% to 80% B over 7.0 min, 80% B for 2.0 min; 0.4 mL/min; injection volume 2 µL, retention times and observed m/z values are seen below (**Figure III.6**).

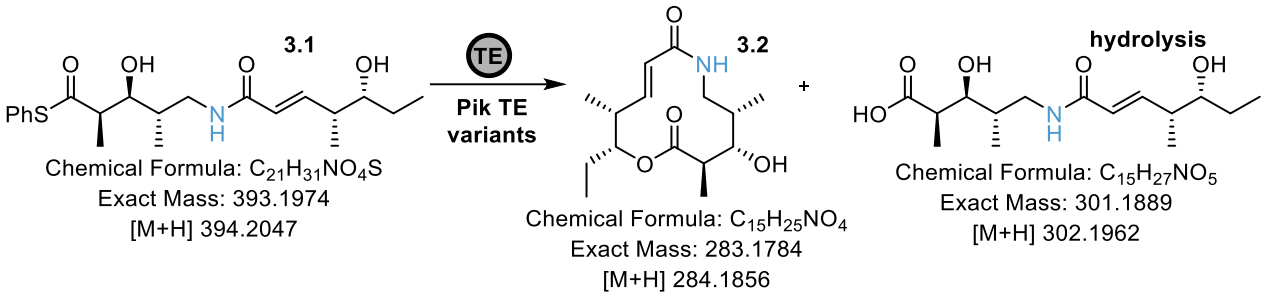


Figure III.6 Amide hexaketide reaction with Pik TE WT or Pik TE S148C. Reaction conditions: water, sodium phosphate buffer (400 mM $NaH_2PO_4 \cdot H_2O$ and Na_2HPO_4, 200 mM NaCl, pH 7.2), hexaketide (1 mM), Pik TE (10 µM), overnight, stationary, rt.

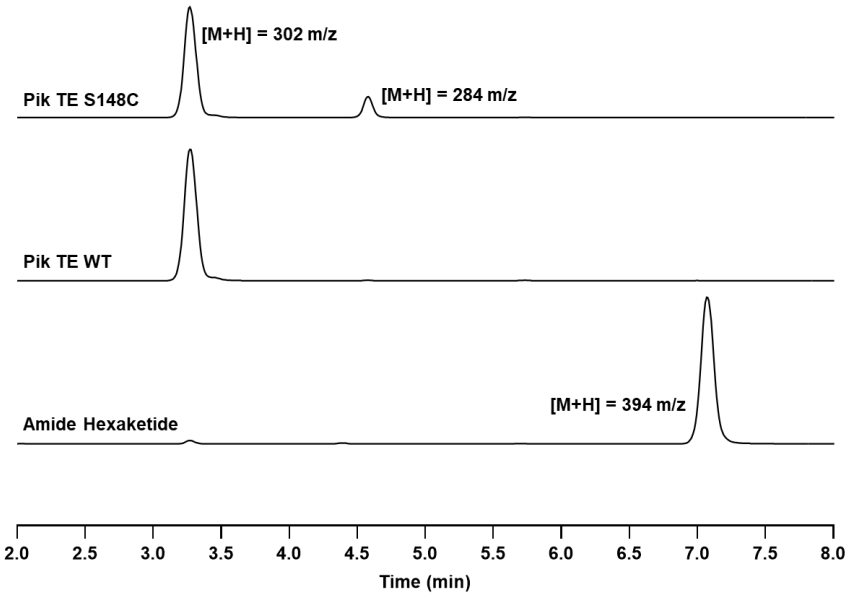
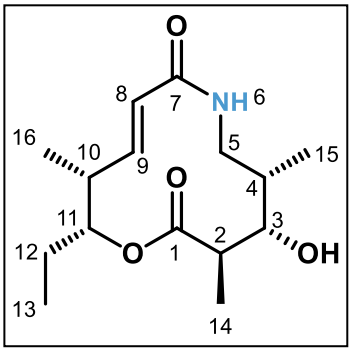


Figure III.7 Example of extracted ion chromatogram (EIC) resulting from enzymatic reactions of the amide hexaketide with Pik TE WT and Pik TE S148C. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 10 minutes (min) total run.

Preparative Enzymatic Reactions and Isolation of Macrolactone/Lactam

A preparative 15 mL enzymatic reaction was performed to characterize the desired macrolactone/lactam. Reaction conditions: water, sodium phosphate reaction buffer (400 mM $NaH_2PO_4 \bullet H_2O$ and Na_2HPO_4, 200 mM NaCl, pH 7.2), 1 mM amide hexaketide (~6 mg of substrate in DMSO) and lastly 10 µM of Pik TE WT or Pik TE S148C, 2 hours, stationary, rt. The reaction was quenched with acetone (15 mL) and moved to -20 ºC freezer for 30 mins to crash out protein. Next, filtered through a celite plug, concentrated by rotary evaporation, extracted with EtOAc (3x) and dried. The crude product was dissolved in methanol and purified by preparatory HPLC. Purification by HPLC monitoring at 210 nm and 254 nm was conducted as follows: mobile phase (A = deionized water + 0.1% formic acid, B = acetonitrile); 20% B for 2.0 min, 20% to 75% B over 38 min, 75% B for 5.0 min; 5 mL/min; each injection volume was 200 µL. HPLC fractions were collected, analyzed by mass in the LC-MS, concentrated and the products characterized by NMR affording macrolactone/lactam product (**Table III.1**).

Table III.1 ^{1}H NMR and ^{13}C NMR (600 MHz NMR, $CDCl_3$)



Position	^{1}H (mult, J Hz)	^{13}C	COSY	HMBC
1		175.10		
2	2.65 (m)	44.13	3	15, 3, 1
3	3.60 (dd, J = 10.1, 1.5 Hz)	77.46	2, 4	15, 4, 5a, 5b, 7
4	1.59 (m)	38.66	5a, 5b, 15	15, 5a, 5b
5a	3.56 (m)	43.68	6, 5b, 4	15
5b	2.77 (ddd, J = 15.1, 10.9, 4.6 Hz)	43.68	6, 5a, 4	15, 4, 7
6	5.90 (dd)		5a, 5b	
7		172.51		
8	6.20 (dd, J = 16.4, 1.4 Hz)	124.66	10, 9	10, 11, 7
9	6.38 (dd, J = 16.3, 5.5 Hz)	142.17	10	16, 10, 11, 7, 8
10	2.62 (m)	36.80	16, 9, 11	16, 8, 9
11	4.89 (ddd, J = 8.3, 6.5, 2.4 Hz)	75.15	12a, 10	16, 12a, 12b, 10, 9, 1
12a	1.74 (dt, J = 13.5, 7.5 Hz)	25.20	11, 13	16, 10, 11
12b	1.56 (m)	25.20	11	16, 10, 11
13	0.92 (t, J = 7.4 Hz)	10.35	12a, 12b	12a, 12b, 11
14	1.31 (d, J = 6.7 Hz)	16.36	2	2, 3, 1
15	1.08 (d, J = 6.7 Hz)	16.30	4	4, 5a, 5b, 3
16	1.13 (d, J = 6.9 Hz)	10.13	10	10, 11, 9
3'				

III.9.4 DNA Cloning and Mutagenesis

DNA Cloning and Sequence of Pik TE S148C

N-terminal histidine (His_6) tagged Pik TE S148C gene was codon optimized in a pET-28(a)+ vector between *NdeI* and *XhoI* restriction sites. The resulting plasmid was purchased from Twist Bioscience.

DNA sequence:

ATGTCAGGAGCTGATACAGGGGCGGGTGCAGGCATGTTTCGCGCTTTGTTCCGCCAA
GCTGTTGAAGATGATCGCTACGGCGAGTTCCTGGACGTGTTAGCGGAGGCATCTGCC
TTCCGTCCGCAGTTTGCGTCCCCGGAAGCATGTAGCGAACGCCTGGACCCGGTGTTG
CTGGCGGGCGGTCCGACCGACCGTGCGGAGGGACGTGCGGTGCTTGTCGGCTGCAC
GGGCACTGCGGCAAACGGTGGCCCGCATGAGTTCCTGCGTCTCTCCACCAGCTTTCA
GGAGGAGCGCGATTTTTTGGCTGTTCCGCTGCCGGGCTATGGTACGGGCACCGGTAC
CGGTACGGCTCTGCTGCCGGCTGACCTGGACACCGCGTTGGACGCGCAGGCGCGTG
CGATCCTGCGTGCGGCGGGGGACGCACCGGTTGTTTTGCTTGGTCATTGCGGTGGTG
CGTTGCTGGCTCACGAACTGGCGTTCCGCCTGGAAAGAGCCCATGGTGCACCGCCA

GCAGGCATCGTATTGGTTGATCCGTATCCGCCCGGTCACCAAGAACCGATTGAAGTT
TGGTCTCGTCAACTGGGTGAAGGCCTGTTCGCCGGCGAGCTGGAGCCTATGAGCGAT
GCGCGTTTGCTGGCTATGGGTCGTTACGCTCGCTTCCTCGCAGGTCCGCGTCCGGGT
CGTAGCAGCGCTCCGGTGTTACTGGTGCGTGCGTCGGAACCGCTGGGCGATTGGCAG
GAGGAGCGCGGTGACTGGCGTGCGCATTGGGATCTGCCACACACCGTCGCAGATGT
TCCGGGCGACCACTTTACCATGATGCGTGATCACGCCCCAGCGGTGGCGGAAGCCG
TGCTGAGCTGGCTGGACGCCATTGAGGGTATCGAGGGCGCGGGTAAA

Amino acid sequence:

MSGADTGAGAGMFRALFRQAVEDDRYGEFLDVLAEASAFRPQFASPEACSERLDPV
LLAGGPTDRAEGRAVLVGCTGTAANGGPHEFLRLSTSFQEERDFLAVPLPGYGTGTG
TGTALLPADLDTALDAQARAILRAAGDAPVVLLGHCGGALLAHELAFRLERAHGAPP
AGIVLVDPYPPGHQEPIEVWSRQLGEGLFAGELEPMSDARLLAMGRYARFLAGPRPG
RSSAPVLLVRASEPLGDWQEERGDWRAHWDLPHTVADVPGDHFTMMRDHAPAVAE
AVLSWLDAIEGIEGAGK

Site Saturation Mutagenesis Libraries

Method to generate single mutations around the active site of Pik TE S148C to cover all possible canonical amino acids at a given residue.[205] The Pik TE WT with a covalent pentaketide affinity label crystal structure[36] (PDB: 2HFJ) was suitable to determine twenty-one residues within 6 Å of the substrate for mutagenesis. A modified QuikChange method (Stratagene) PCR (50 µL total volume) contained distilled water, 4% DMSO, 0.2 µM degenerate NNK forward only primers (**Table III.2**), 100 ng Pik TE S148C plasmid, 1X *Pyrococcus furiosus* (*Pfu*) ultra II DNA polymerase reaction buffer, 0.2 mM dNTP mix (Promega: U151A) and 1 unit of *Pfu* ultra II fusion HotStart DNA polymerase (Agilent: 600672). PCR cycle conditions: 1) polymerase activation: 95 °C for 3 min, 2) denature: 95 °C for 30 sec, 3) annealing: 55 °C for 1 min (temperatures changed depending on Tm of primers), 4) extension: 65 °C for 12.4 min, 5) terminal extension: 65 °C for 10 min. Cycles run: 30 (for steps 2-4). Upon completion of

PCR amplification, 2 µL of DpnI (NEB: R0176S) were added and incubated at 37 ºC for 2 hours. After DpnI digestion, 5 µL from the reaction were taken directly for transformation in chemically competent BL21(DE3) expression cells, plated on LB agar with kanamycin (50 µg/mL) plates and incubated overnight at 37 ºC to obtain colonies the next day. Colonies were used for library screening in 96-well plates.

Table III.2 Site saturation mutagenesis NNK degenerate primers for the generation of Pik TE S148C enzyme variants.

Pik TE S148C residue	**NNK degenerate primer sequence 5' to 3'**
Tyr25	CGCCAAGCTGTTGAAGATGATCGCNNKGGCGAGTTCCTGGACGTGTTAGC
Leu29	GAAGATGATCGCTACGGCGAGTTCNNKGACGTGTTAGCGGAGGCATCTGC
Gly76	GTGCGGTGCTTGTCGGCTGCACGNNKACTGCGGCAAACGGTGGCCCGCATG
Thr77	GCGGTGCTTGTCGGCTGCACGGGCNNKGCGGCAAACGGTGGCCCGCATGAG
Ala78	GGTGCTTGTCGGCTGCACGGGCACTNNKGCAAACGGTGGCCCGCATGAGTTCCTGC
Ala79	GCTTGTCGGCTGCACGGGCACTGCGNNKAACGGTGGCCCGCATGAGTTCCTGCGTC

His147	GGGGACGCACCGGTTGTTTTGCTTGGTNNKTGCGG TGGTGCGTTGCTGGCTC
Gly149	GCACCGGTTGTTTTGCTTGGTCATTGCNNKGGTGC GTTGCTGGCTCACGAACTG
Gly150	CACCGGTTGTTTTGCTTGGTCATTGCGGTNNKGCGT TGCTGGCTCACGAACTGGC
Leu152	TTGCTTGGTCATTGCGGTGGTGCGNNKCTGGCTCA CGAACTGGCGTTCCG
Pro177	CGCCAGCAGGCATCGTATTGGTTGATNNKTATCCG CCCGGTCACCAAGAACC
Tyr178	GCCAGCAGGCATCGTATTGGTTGATCCGNNKCCGC CCGGTCACCAAGAACCGATTG
Gln183	GTTGATCCGTATCCGCCCGGTCACNNKGAACCGAT TGAAGTTTGGTCTCGTCAACTGGG
Ile186	GTATCCGCCCGGTCACCAAGAACCGNNKGAAGTTT GGTCTCGTCAACTGGGTG
Leu193	GTCACCAAGAACCGATTGAAGTTTGGTCTCGTCAA NNKGGTGAAGGCCTGTTCGCCGGCG
Leu197	GTTTGGTCTCGTCAACTGGGTGAAGGCNNKTTCGC CGGCGAGCTGGAGCCTATG
Met213	CTATGAGCGATGCGCGTTTGCTGGCTNNKGGTCGT TACGCTCGCTTCCTCGCAG

Tyr216	GCGCGTTTGCTGGCTATGGGTCGTNNKGCTCGCTT CCTCGCAGGTCCGCGTC
Ala217	GCGCGTTTGCTGGCTATGGGTCGTTACNNKCGCTT CCTCGCAGGTCCGCGTCC
Leu220	CTGGCTATGGGTCGTTACGCTCGCTTCNNKGCAGG TCCGCGTCCGGGTCGTAG
Phe269	GTCGCAGATGTTCCGGGCGACCACNNKACCATGAT GCGTGATCACGCCCC

Error Prone Polymerase Chain Reaction (epPCR) Libraries

To create our random mutagenesis Pik TE variant library, an epPCR methodology was utilized.[170, 171] The pET-28(a)+ vector containing the Pik TE S148C, A217T, G222V (for Round 2) or Pik TE S148C, A217T, G222V, M271V, Y25C, L126V (for Round 3) gene (also applies to other Pik TE gene variants generated) was digested with restriction enzymes *NdeI* (NEB: R0111S) and *XhoI* (NEB: R0146S) to generate an empty vector. In order to eliminate producing alterations elsewhere in the plasmid affecting important genes, we performed epPCR solely on the Pik TE S148C gene. epPCR (100 µL total volume) on the Pik TE variant gene was conducted by adding distilled water, 200 µM, 250 µM or 300 µM of $MnCl_2$•$4H_2O$ solution (stock: 1 mM in deionized water), 0.2 µM forward primer corresponding to *NdeI* restriction site of the pET-28(a)+ vector, 0.2 µM reverse primer corresponding to *XhoI* restriction site of the pET-28(a)+ vector (**Table III.3),** 200 ng of template DNA (Pik TE variant), 1X standard *Thermus aquaticus* (*Taq*) polymerase reaction buffer, 0.2 mM dNTP mix (Promega: U151A), 1 unit *Taq* DNA polymerase (NEB: M0273S). PCR cycle conditions: 1) polymerase activation: 95 °C for 40 sec, 2) denature: 95 °C

for 30 sec, 3) annealing: 68 °C for 30 sec, 4) extension: 72 °C for 56 sec, 5) terminal extension: 72 °C for 5 min. Cycles run: 30 (for steps 2-4). Upon completion of PCR amplification, 4 µL of DpnI were added and incubated at 37 ºC for 2 hours. After DpnI digestion, PCR cleanup was done using QIAquick PCR Purification Kit (QIAGEN: 28106). Next, we cloned via Gibson assembly the epPCR products into the empty pET-28(a)+ vector between restriction sites *NdeI* and *XhoI* to generate mutant TEs. Gibson assembly reaction set up: water, digested vector : epPCR products (2:1) and NEBuilder HiFi DNA Assembly Master Mix (NEB: E2621S) were mixed and incubated at 50 ºC for 1 hour. After 1 hour, the reactions were quickly put into ice, then 2 µL were taken from the mixture for transformation in chemically competent BL21(DE3) expression cells, plated on LB agar with kanamycin (50 µg/mL) plates and incubated overnight at 37 ºC to obtain colonies the next day (**Figure III.8**). Colonies were used for library screening in 96-well plates.

Table III.3 epPCR primers for amplification between restriction sites NdeI and XhoI in the pET-28(a)+ Pik TE S148C vector.

Restriction site	**epPCR primer sequence 5' to 3'**
NdeI	TGGTGCCGCGCGGCAGCCATATGTCA
XhoI	TCAGTGGTGGTGGTGGTGGTGCTCGA

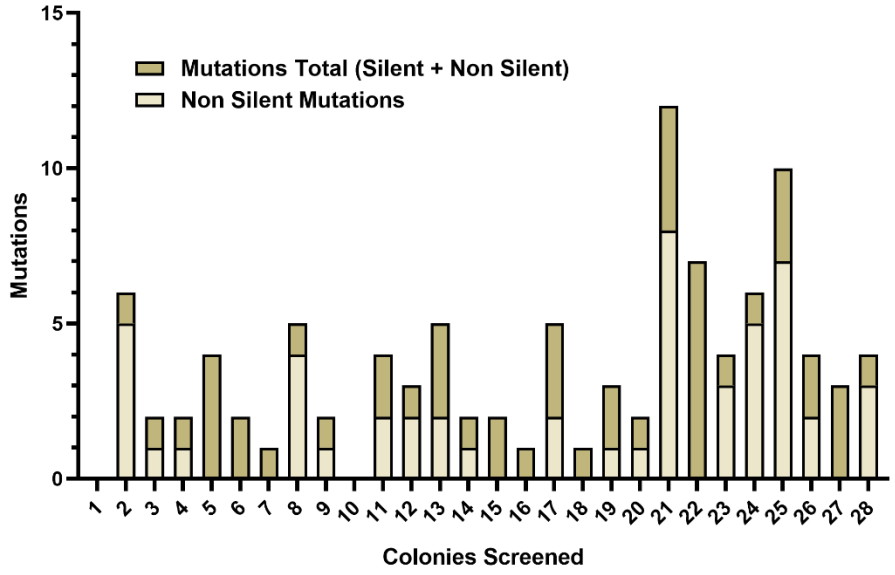


Figure III.8 Example of colonies screened after Gibson assembly method optimization: epPCR total mutations (silent + non silent) when varying concentrations of $MnCl_2$: colonies 1-10: 200 µM, colonies 11-20: 250 µM, colonies 21-28: 300 µM.

*200 µM and 250 µM resulted in similar amounts of mutations in the Pik TE variants. The chosen $MnCl_2$ concentrations were 200 µM and 300 µM for two diverse mutation rates for the enzyme libraries.

Site Directed Mutagenesis for Combination of Mutations

Combination of mutations resulting from directed evolution variants were achieved via site directed mutagenesis. A modified QuikChange method (Stratagene) PCR (50 µL total volume) contained distilled water, 4% DMSO, 0.2 µM forward only primers (**Table III.4**), 100 ng plasmid DNA template, 1X *Pyrococcus furiosus* (*Pfu*) ultra II DNA polymerase reaction buffer, 0.2 mM dNTP mix (Promega: U151A) and 1 unit of *Pfu* ultra II fusion HotStart DNA polymerase (Agilent: 600672). PCR cycle conditions: 1) polymerase activation: 95 °C for 3 min, 2) denature: 95 °C for

30 sec, 3) annealing: 55 °C for 1 min (temperatures changed depending on Tm of primers), 4) extension: 65 °C for 12.4 min, 5) terminal extension: 65 °C for 10 min. Cycles run: 30 (for steps 2-4). Upon completion of PCR amplification, 2 µL of DpnI (NEB: R0176S) were added and incubated at 37 ºC for 2 hours. After DpnI digestion, 5 µL from the reaction were taken directly for transformation in chemically competent BL21(DE3) expression cells, plated on LB agar with kanamycin (50 µg/mL) plates and incubated overnight at 37 ºC to obtain colonies the next day. Single colonies were picked for overnight cultures in LB media with kanamycin (50 µg/mL) and grown overnight at 37 ºC. Plasmid miniprep kit from QIAGEN was utilized to extract DNA for sequencing variants and confirm mutations.

Table III.4 Site directed mutagenesis primers for the generation of Pik TE S148C enzyme variants containing combinations of mutations from rounds of evolution.

Pik TE S148C variant plasmid DNA template	**Mutation change**	**Site directed mutagenesis primer sequence 5' to 3'**
S148C, L29M	A217T	CGCGTTTGCTGGCTATGGGTCGTTACACTCGC TTCCTCGCAGGTCCGCGTC
S148C, A217T	G222V	GGGTCGTTACACTCGCTTCCTCGCAGTGCCGC GTCCGGGTCGTAGCAGCG
S148C, A217T, G222V, L29M	R160C	GTTGCTGGCTCACGAACTGGCGTTCTGCCTGG AAAGAGCCCATGGTGCACCG
S148C, A217T, G222V	C148S	CGCACCGGTTGTTTTGCTTGGTCATTCCGGTG GTGCGTTGCTGGCTCACG
1) S148C, A217T, G222V, M271V *2)* S148C, A217T, G222V, M271V, Y25C	L126V	CTGCCGGCTGACCTGGACACCGCGGTGGACG CGCAGGCGCGTGCGATCCTG

S148C, A217T, G222V, M271V	Y25C	CGCCAAGCTGTTGAAGATGATCGCTGCGGCG AGTTCCTGGACGTGTTAGC
1) S148C, A217T, G222V, G112D, P141R, V188A *2)* S148C, A217T, G222V, Q94R, V142A, A289T *3)* S148C, A217T, G222V, H84R, Q192L	M271V	CAGATGTTCCGGGCGACCACTTTACCGTGATG CGTGATCACGCCCCAGCGG

III.9.5 High-throughput Screening Protein Expression and Enzymatic Reactions

Protein Expression and Generation of Crude Cell Lysates in 96-well Plates

Transformed Pik TE variants in chemically competent BL21(DE3) expression cells, plated on LB agar with kanamycin (50 µg/mL) plates and incubated overnight at 37 ºC to obtain colonies the next day. Seed cultures were prepared using LB media with kanamycin (50 µg/mL). Then, picked single colonies with sterilized toothpicks and plated in sterilized deep 96-well plate for incubation overnight at 37 ºC, 180 rpm. To a clean deep 96-well plate, 1 mL of overnight express instant TB medium (Sigma Aldrich: 71491) containing TB salts (0.17 M KH_2PO_4 and 0.72 M K_2HPO_4) and kanamycin (50 µg/mL) were added to each well, followed by the addition of 50 µL of the overnight seed culture. The plates were sealed with adhesive gas permeable seals to ensure proper *E. coli* growth (Thermo-Fisher Scientific: AB-0718). Glycerol stocks were prepared with 50 µL of overnight culture added to 50 µL of a sterile solution 40% glycerol in water to sterilized 96-well PCR plates (DOT Scientific: 370LP - PCR), sealed with foil sealing film (CellTreat: 229133) and stored at -80 ºC for future hits identification. Plates were grown for 5 hours at 37 ºC, 180 rpm. After OD_{600} reached 1 – 1.5, these were incubated overnight at 18 ºC, 180 rpm. Next

day, centrifuged the plates at 3,000 rpm for 30 minutes. The resulting supernatant was removed, and the deep 96-well plate was stored in -80 ºC for efficient future cell lysis. The following day, the cell pellet was resuspended in 600 µL of sodium phosphate reaction buffer (400 mM $NaH_2PO_4•H_2O$ and Na_2HPO_4, 200 mM NaCl, pH 7.2) containing lysozyme (0.25 mg/mL). After resuspension, these were incubated for 1 hour at -80 ºC. The cell pellets were thawed in a water bath at 37 ºC for 15 minutes, shaken at 200 rpm for 15 mins. The cell pellet became viscous. To each well, 6 µL of DNAse was added for a final concentration of 10 µg/mL (stock: 2mg/mL in water) followed by 12 µL of $MgSO_4$ for a final concentration of 20 mM (stock: 1M in deionized water). Plates were shaken at 200 rpm for 15 mins. At this point the cell pellets became non-viscous (tested for this to ensure proper lysis). Lastly, centrifuged plates at 3,000 rpm for 30 mins. The recovered supernatant corresponds to the crude cell lysate ready for enzymatic reactions.

Enzymatic Reactions in 96-well Plates

To a clean conical V-shape 96-well plate (Corning: 3897), 1.25 µL of amide hexaketide was added to each well (stock: 1 mM in DMSO) for a 25 µM final concentration in 50 µL total volume enzymatic reactions. Then, 48.75 µL of protein (crude cell lysate from 96-well plate expression) were added. These were run overnight at room temperature in the 96-well plates closed with plates lids. The next day, 150 µL of methanol were added to each well to quench the reactions. The plates were sealed with foil sealing film (CellTreat: 229133) to avoid methanol vaporization. The well plates were put into the -20 ºC fridge for 15 mins to help crash out protein to the bottom of the conical V-shape plates. These were spun down in the centrifuge at 2,000 rpm for 20 mins. After centrifugation, 100 µL of reserpine internal standard (stock: 600 nM in water) for a 300 nM concentration in sample was added to the plates, followed by addition of 100 µL of reaction mixture to a clean 96-well plate (Axygen: P-96-450V-C-S) for hit identification via LC-MS

analysis. LC-MS conditions were as follows: mobile phase (A = deionized water + 0.1% formic acid, B = 95% acetonitrile/deionized water + 0.1% formic acid); 10% B for 0.2 min, 10% to 100% B over 2.6 min, 100% B for 0.2 min; 0.6 mL/min; injection volume 12 µL.

III.9.6 Enzyme Variants Identification, Expression, Purifications and Enzymatic Reactions

Hits Identification of New Variants

After identification of potential hits, we went back to the glycerol stocks corresponding to the desired wells and plates for DNA isolation. The positive wells were picked with sterilized toothpicks, transferred to 5 mL LB media with kanamycin (50 µg/mL), and grown overnight at 37 ºC. The overnight cultures were lysed, and plasmid DNA samples extracted using QIAprep Spin Miniprep Kit (250) (QIAGEN: 27106). DNA samples were sent for sequencing to determine mutations resulting from rounds of evolution.

Protein Expression and Purification of New Variants

Once the variants were identified, these extracted plasmids were transformed into chemically competent BL21(DE3) expression cells, plated on LB agar with kanamycin (50 µg/mL) plates and incubated overnight at 37 ºC to obtain colonies the next day. Single colonies were picked for overnight cultures in LB media with kanamycin (50 µg/mL) and grown overnight at 37 ºC. Protein expression of mutants were done by adding 250 µL of overnight cultures to 50 mL of TB media, containing TB salts (0.17 M KH_2PO_4 and 0.72 M K_2HPO_4) with kanamycin (50 µg/mL). After OD_{600} reached 1 – 1.5, these were transferred to an ice bath for 15 mins, induced with 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated overnight at 18 ºC, 180 rpm. The following day, the mixture was centrifuged at 3,000 rpm for 30 mins, the supernatant was discarded, and the cell pellet stored in -80 ºC for further protein purification. The purification of

the variants was accomplished in the same manner as described herein for Pik TE WT and Pik TE S148C.

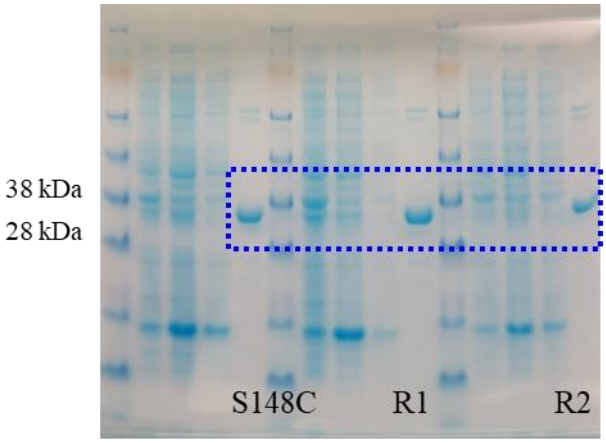


Figure III.9 Example of purified protein SDS-PAGE for Pik TE variants: Pik S148C, Pik TE R1 and Pik TE R2. NuPAGE® 4-12% Bis-Tris Mini Gel of purified Pik TE variants in MES SDS running buffer with BenchMark™ Pre-stained Protein Ladder. Lanes from left to right: Protein Ladder, Pellet, Supernatant, Flowthrough, Eluted Purified Pik TE variants (lanes repeat in the same order for S148C, R1 and R2).

Enzymatic Reactions of New Purified Variants

Analytical scale 50 µL enzymatic reactions (**Figure III.9**) were done by adding water, sodium phosphate reaction buffer (400 mM NaH_2PO_4•H_2O and Na_2HPO_4, 200 mM NaCl, pH 7.2), 25 µM amide hexaketide (substrate) and lastly 10 µM of desired Pik TE variant, overnight, stationary, rt. As negative controls, enzymes were omitted. All reactions were performed and analyzed in triplicates. Next day, the reactions were quenched by adding 3 volumes of MeOH (150 µL) and centrifuged at 13,000 x g, 10 mins, 4 ºC for further LC-MS analysis. LC-MS conditions were as follows: mobile phase (A = deionized water + 0.1% formic acid, B = 95% acetonitrile/deionized water + 0.1% formic acid); 25% B for 0.5 min, 25% to 70% B over 4.0 min, 70% B for 0.5 min;

0.5 mL/min; injection volume 10 µL, retention times and observed m/z values are seen below (**Figure III.10**).

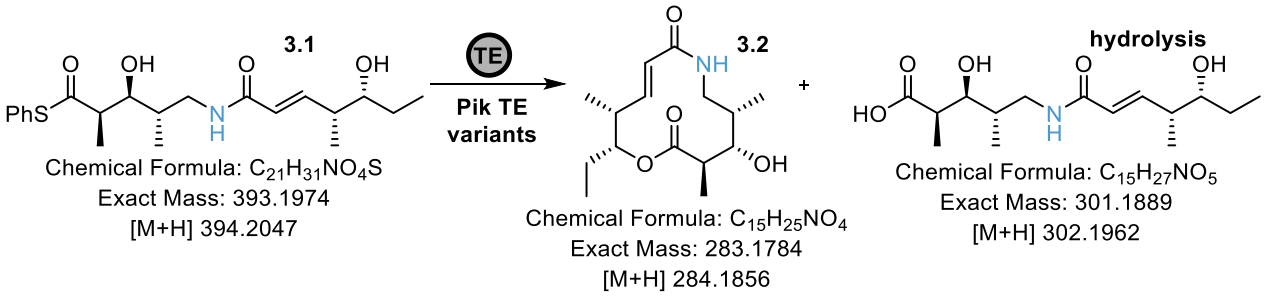


Figure III.10 Amide hexaketide reaction with Pik TE variants from directed evolution campaign. Reaction conditions: water, sodium phosphate buffer (400 mM $NaH_2PO_4 \cdot H_2O$ and Na_2HPO_4, 200 mM NaCl, pH 7.2), hexaketide (0.25 µM), Pik TE variant (10 µM), overnight, stationary, rt.

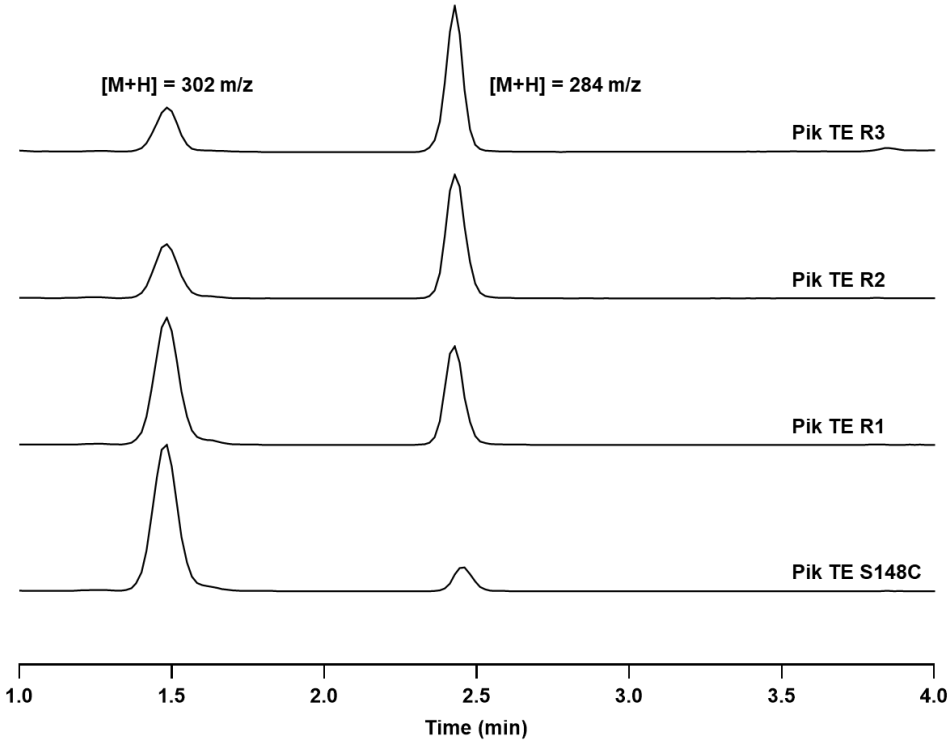
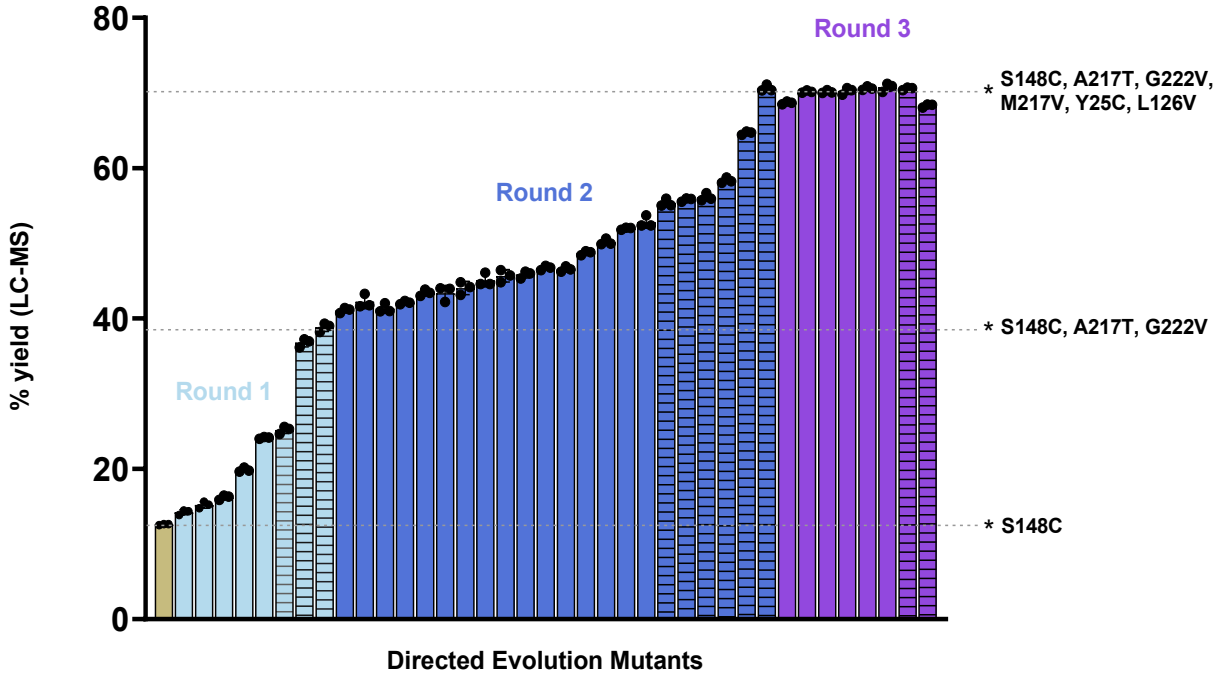


Figure III.11 Example of extracted ion chromatogram (EIC) resulting from enzymatic reactions amide hexaketide with Pik TE mutants. TOF-MS: 25 – 70% acetonitrile with 0.1% formic acid, 5 minutes (min) total run.


80
60
40
20
0
% yield (LC-MS)
Round 1
Round 2
Round 3
* S148C, A217T, G222V, M217V, Y25C, L126V
* S148C, A217T, G222V
* S148C
Directed Evolution Mutants

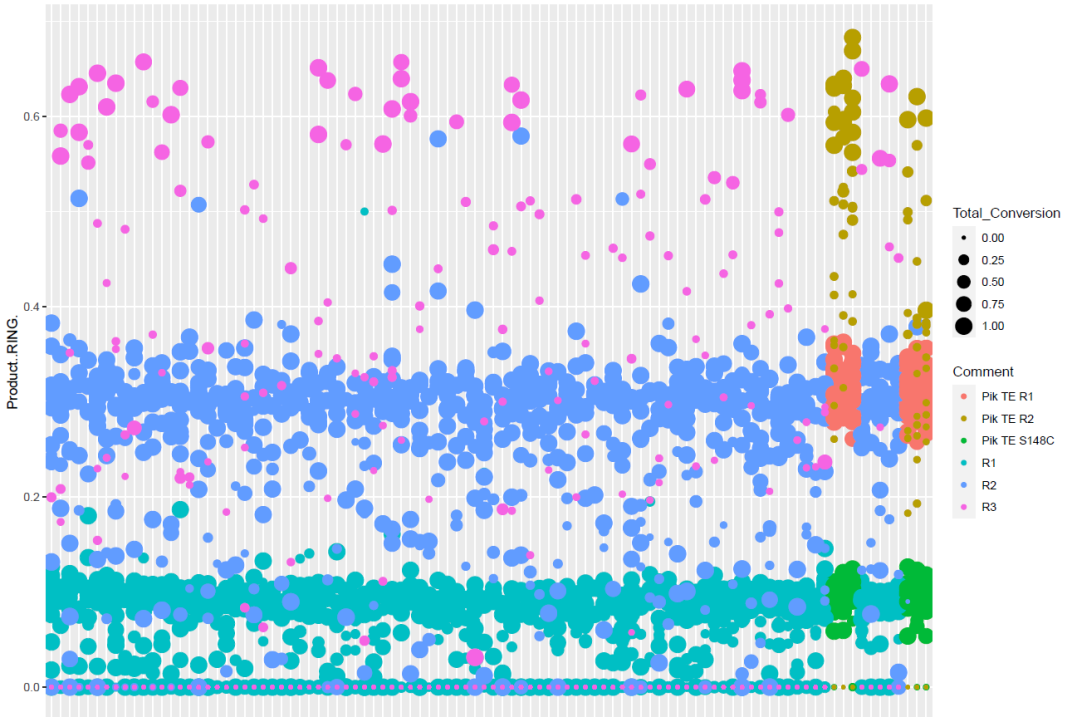

0.6
0.4
0.2
0.0
Product.RING.
Total_Conversion
0.00
0.25
0.50
0.75
1.00
Comment
Pik TE R1
Pik TE R2
Pik TE S148C
R1
R2
R3

Figure III.12 Pik TE directed evolution variants developed in rounds 1-3 and their corresponding LC-MS % yields for hybrid 12-membered macrolactone/lactam ring formation and depiction of the entire library screen.

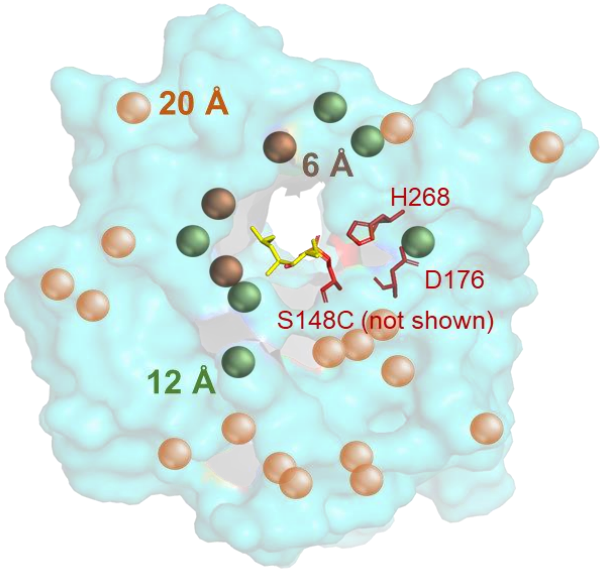


Figure III.13 Beneficial mutations found in rounds 1-2 that impacted selectivity for macrocyclization and their corresponding distances (Å) with the bound pentaketide affinity label in yellow in the Pik TE WT crystal structure (PDB: 2HFJ), catalytic triad in red.

Sequences of Pik TE Variants from Directed Evolution Rounds

Pik TE S148C, A217T, G222V (Round 1)

DNA sequence:

ATGTCAGGAGCTGATACAGGGGCGGGTGCAGGCATGTTTCGCGCTTTGTTCCGCCAA
GCTGTTGAAGATGATCGCTACGGCGAGTTCCTGGACGTGTTAGCGGAGGCATCTGCC
TTCCGTCCGCAGTTTGCGTCCCCGGAAGCATGTAGCGAACGCCTGGACCCGGTGTTG
CTGGCGGGCGGTCCGACCGACCGTGCGGAGGGACGTGCGGTGCTTGTCGGCTGCAC
GGGCACTGCGGCAAACGGTGGCCCGCATGAGTTCCTGCGTCTCTCCACCAGCTTTCA

GGAGGAGCGCGATTTTTTGGCTGTTCCGCTGCCGGGCTATGGTACGGGCACCGGTAC
CGGTACGGCTCTGCTGCCGGCTGACCTGGACACCGCGTTGGACGCGCAGGCGCGTG
CGATCCTGCGTGCGGCGGGGGACGCACCGGTTGTTTTGCTTGGTCATTGCGGTGGTG
CGTTGCTGGCTCACGAACTGGCGTTCCGCCTGGAAAGAGCCCATGGTGCACCGCCA
GCAGGCATCGTATTGGTTGATCCGTATCCGCCCGGTCACCAAGAACCGATTGAAGTT
TGGTCTCGTCAACTGGGTGAAGGCCTGTTCGCCGGCGAGCTGGAGCCTATGAGCGAT
GCGCGTTTGCTGGCTATGGGTCGTTACACTCGCTTCCTCGCAGTGCCGCGTCCGGGT
CGTAGCAGCGCTCCGGTGTTACTGGTGCGTGCGTCGGAACCGCTGGGCGATTGGCAG
GAGGAGCGCGGTGACTGGCGTGCGCATTGGGATCTGCCACACACCGTCGCAGATGT
TCCGGGCGACCACTTTACCATGATGCGTGATCACGCCCCAGCGGTGGCGGAAGCCG
TGCTGAGCTGGCTGGACGCCATTGAGGGTATCGAGGGCGCGGGTAAA

Amino acid sequence:

MSGADTGAGAGMFRALFRQAVEDDRYGEFLDVLAEASAFRPQFASPEACSERLDPVLL
AGGPTDRAEGRAVLVGCTGTAANGGPHEFLRLSTSFQEERDFLAVPLPGYGTGTGTGTA
LLPADLDTALDAQARAILRAAGDAPVVLLGHCGGALLAHELAFRLERAHGAPPAGIVL
VDPYPPGHQEPIEVWSRQLGEGLFAGELEPMSDARLLAMGRYTRFLAVPRPGRSSAPVL
LVRASEPLGDWQEERGDWRAHWDLPHTVADVPGDHFTMMRDHAPAVAEAVLSWLD
AIEGIEGAGK

Pik TE S148C, A217T, G222V, M271V, Y25C, L126V (Round 2)

DNA sequence:

ATGTCAGGAGCTGATACAGGGGCGGGTGCAGGCATGTTTCGCGCTTTGTTCCGCCAA
GCTGTTGAAGATGATCGCTGCGGCGAGTTCCTGGACGTGTTAGCGGAGGCATCTGCC
TTCCGTCCGCAGTTTGCGTCCCCGGAAGCATGTAGCGAACGCCTGGACCCGGTGTTG
CTGGCGGGCGGTCCGACCGACCGTGCGGAGGGACGTGCGGTGCTTGTCGGCTGCAC
GGGCACTGCGGCAAACGGTGGCCCGCATGAGTTCCTGCGTCTCTCCACCAGCTTTCA
GGAGGAGCGCGATTTTTTGGCTGTTCCGCTGCCGGGCTATGGTACGGGCACCGGTAC
CGGTACGGCTCTGCTGCCGGCTGACCTGGACACCGCGGTGGACGCGCAGGCGCGTG
CGATCCTGCGTGCGGCGGGGGACGCACCGGTTGTTTTGCTTGGTCATTGCGGTGGTG

CGTTGCTGGCTCACGAACTGGCGTTCCGCCTGGAAAGAGCCCATGGTGCACCGCCA
GCAGGCATCGTATTGGTTGATCCGTATCCGCCCGGTCACCAAGAACCGATTGAAGTT
TGGTCTCGTCAACTGGGTGAAGGCCTGTTCGCCGGCGAGCTGGAGCCTATGAGCGAT
GCGCGTTTGCTGGCTATGGGTCGTTACACTCGCTTCCTCGCAGTGCCGCGTCCGGGT
CGTAGCAGCGCTCCGGTGTTACTGGTGCGTGCGTCGGAACCGCTGGGCGATTGGCAG
GAGGAGCGCGGTGACTGGCGTGCGCATTGGGATCTGCCACACACCGTCGCAGATGT
TCCGGGCGACCACTTTACCGTGATGCGTGATCACGCCCCAGCGGTGGCGGAAGCCG
TGCTGAGCTGGCTGGACGCCATTGAGGGTATCGAGGGCGCGGGTAAA

Amino acid sequence:

MSGADTGAGAGMFRALFRQAVEDDRCGEFLDVLAEASAFRPQFASPEACSERLDPVLL
AGGPTDRAEGRAVLVGCTGTAANGGPHEFLRLSTSFQEERDFLAVPLPGYGTGTGTGTA
LLPADLDTAVDAQARAILRAAGDAPVVLLGHCGGALLAHELAFRLERAHGAPPAGIVL
VDPYPPGHQEPIEVWSRQLGEGLFAGELEPMSDARLLAMGRYTRFLAVPRPGRSSAPVL
LVRASEPLGDWQEERGDWRAHWDLPHTVADVPGDHFTVMRDHAPAVAEAVLSWLDA
IEGIEGAGK

Pik TE S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W (Round 3)

DNA sequence:

ATGTCAGGAGCTGATACAGGGGCGGGTGGAGGCATGTTTCGCGCTTTGTTCCGCAAA
GCTGTTGAAGATGATCGCTGCGGCGAGTTCCTGGACGTGTTAGCGGAGGCATCTGCC
TTCCGTCCGCAGTTTGCGTCCCCGGAAGCATGTAGCGAACGCCTGGACCCGGTGTTG
CTGGCGGGCGGTCCGACCGACCGTGCGGAGGGACGTGCGGTGCTTGTCGGCTGCAC
GGGCACTGCGGCAAACGGTGGCCCGCATGAGTTCCAGCGTCTCTCCACCAGCTTTCA
GGAGGAGCGCGATTTTTTGGCTGTTCCGCTGCCGGGCTATGGTACGGGCACCGGTAC
CGGTACGGCTCTGCTGCCGGTTGACCTGGACACCGCGGTGGACGCGCAGGCGCGTG
CGATCCTGCGTGCGGCGTGGGACGCACCGGTTGTTTTGCTTGGTCATTGCGGTGGTG
CGTTGCTGGCTCACGAACTGGCGTTCCGCCTGGAAAGAGCCCATGGTGCACCGCCA
GCAGGCATCGTATTGGTTGATCCGTATCCGCCCGGTCACCAAGAACCGATTGAAGTT

TGGTCTCGTCAACTGGGTGAAGGCCTGTTCGCCGGCGAGCTGGAGCCTATGAGCGAT
GCGCGTTTGCTGGCTATGGGTCGTTACACTCGCTTCCTCGCAGTGCCGCGTCCGGGT
CGTAGCAGCGCTCCGGTGTTACTGGTGCGTGCGTCGGAACCGCTGGGCGATTGGCAG
GAGGAGCGCGGTGACTGGCGTGCGCATTGGGATCTGCCACACACCGTCGCAGATGT
TCCGGGCGACCACTTTACCGTGATGCGTGATCACGCCCCAGCGGTGGCGGAAGCCG
TGCTGAGCTGGCTGGACGCCATTGAGGGTATCGGGGGCGCGGGTAAA

Amino acid sequence:

MSGADTGAGGGMFRALFRKAVEDDRCGEFLDVLAEASAFRPQFASPEACSERLDPVLL
AGGPTDRAEGRAVLVGCTGTAANGGPHEFQRLSTSFQEERDFLAVPLPGYGTGTGTGT
ALLPVDLDTAVDAQARAILRAAWDAPVVLLGHCGGALLAHELAFRLERAHGAPPAGIV
LVDPYPPGHQEPIEVWSRQLGEGLFAGELEPMSDARLLAMGRYTRFLAVPRPGRSSAPV
LLVRASEPLGDWQEERGDWRAHWDLPHTVADVPGDHFTVMRDHAPAVAEAVLSWLD
AIEGIGGAGK

III.9.7 Total Turnover Numbers (TTNs), Thermal Shift Assay and Initial Rates for Substrate Consumption and Macrolactone Formation

Examples of Calibration Curves for Amide Hexaketide (Substrate) and Hybrid 12-membered Macrolactone/Lactam Ring (Product)

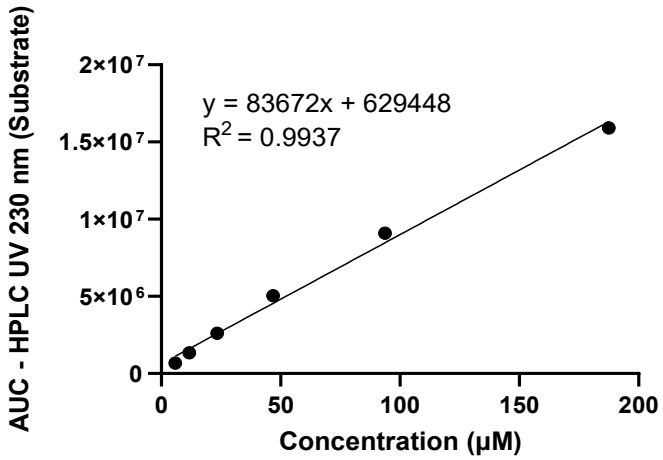


Figure III.14 Standard curve for amide hexaketide (**3.1**) substrate.

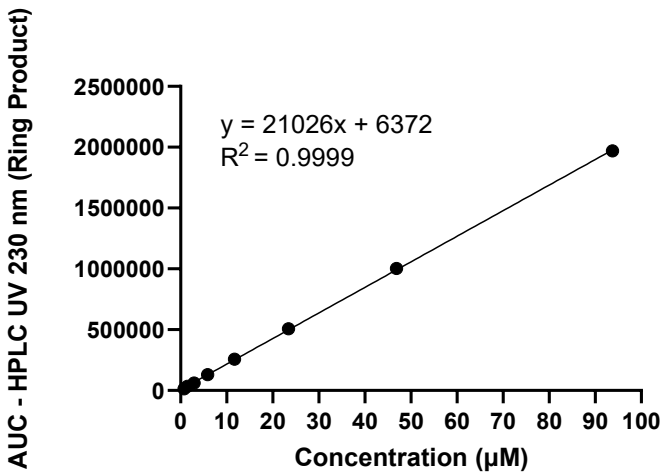


Figure III.15 Standard curve for hybrid 12-membered macrolactone/lactam ring (**3.2**) product.

Total Turnover Numbers (TTNs)

For determination of TTNs, analytical scale 50 μL enzymatic reactions were done by adding water, sodium phosphate reaction buffer (400 mM NaH_2PO_4•H_2O and Na_2HPO_4, 200 mM NaCl, pH 7.2), 500 μM amide hexaketide (substrate) and lastly 1 μM of desired Pik TE variant, 18 hours, stationary, 25 ^oC temperature controlled in thermocycler. All reactions were performed and analyzed in triplicates. After 18 hours, the reactions were quenched by adding 3 volumes of MeOH (150 μL) and centrifuged at 13,000 x g, 10 mins, 4 ^oC for further HPLC analysis using 230 nm UV detection. Analytical HPLC conditions were as follows: mobile phase (A = deionized water + 0.1% formic acid, B = 100% acetonitrile); 25% B for 1.0 min, 25% to 70% B over 14.0 min, 70% B for 2.0 min; 0.4 mL/min; injection volume 50 μL, retention times values and observed spectra are shown below. TTN values were determined by evaluating the areas of HPLC peaks corresponding to macrolactone product formation and evaluating the μM values with standard curves (run simultaneously and fitted to a linear curve). TTN values were calculated using the following formula TTN = moles of macrolactone generated (μM) ÷ moles of TE utilized (μM).

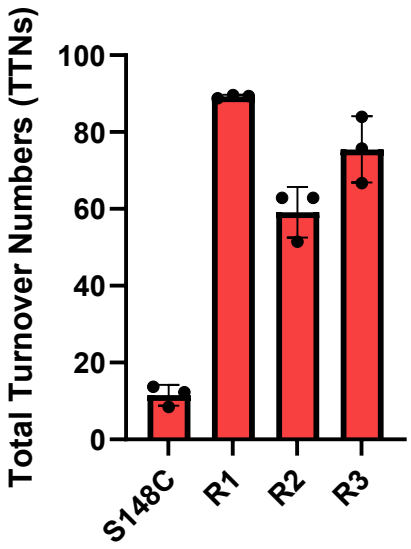
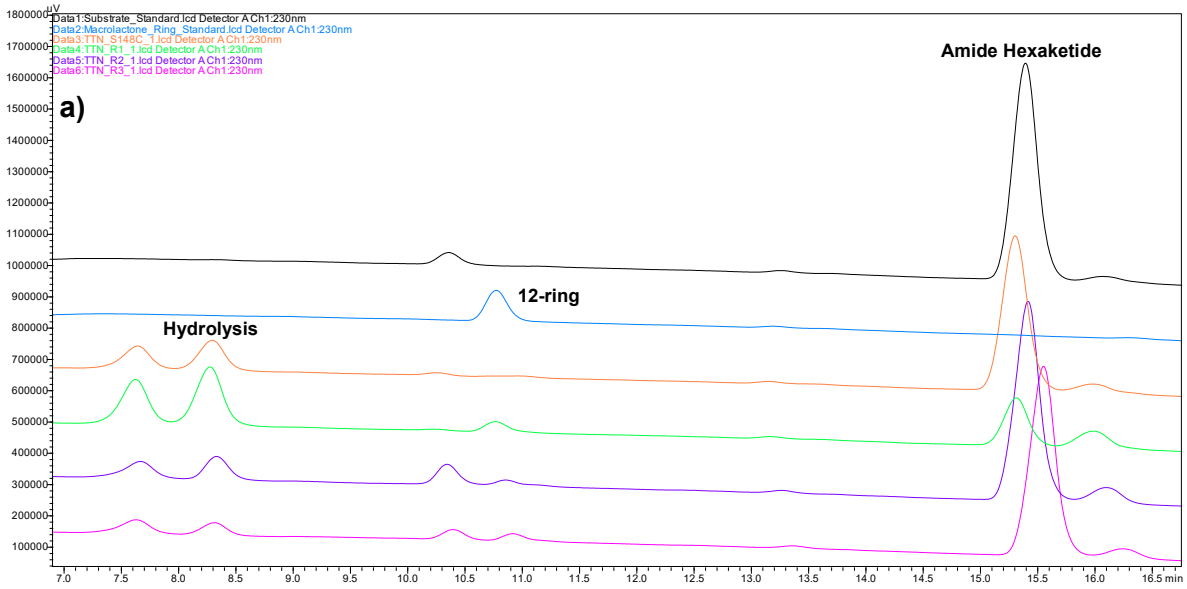


Figure III.16 Total Turnover Numbers (TTNs) for Pik TE S148C, Pik TE R1 (S148C, A217T, G222V), Pik TE R2 (S148C, A217T, G222V, M271V, Y25C, L126V) and Pik TE R3 (S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W).



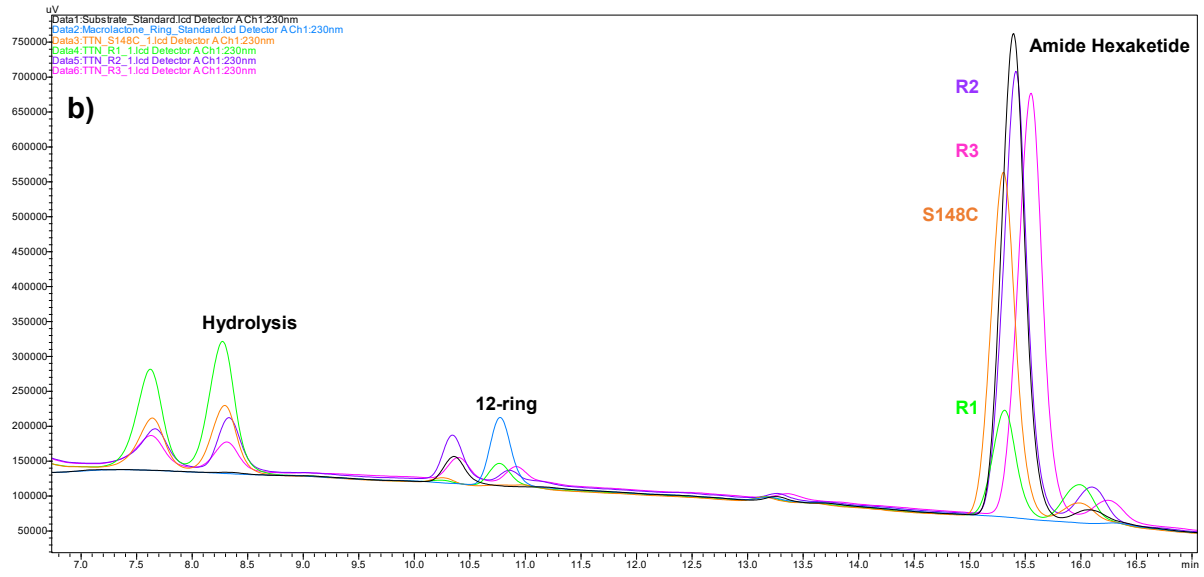


Figure III.17 a) Example of Total Turnover Numbers (TTNs) resulting from enzymatic reactions amide hexaketide with Pik TE mutants. Analytical HPLC: 25 – 70% acetonitrile, 17 minutes (min) total run. HPLC retention time drift is observed depending on different data collection days. b) Overlaid HPLC TTNs traces. Note: The peak observed to the left of the 12-ring corresponds to trace amounts (negligible) degradation of amide hexaketide (retention time: 7.60 min) due to exchange of thiophenol group with MeOH upon quenching the reactions. The peak observed to the left of hydrolysis product corresponds to spontaneous degradation of the hydrolysis molecule (retention time: 10.35 min) under the reaction conditions.

Thermal Shift Assay

Protein stability of the Pik TE mutants was determined by thermal shift assay using SYPRO orange dye (Invitrogen: S6650) to monitor enzyme thermal denaturation with sensitive fluorescence detection. The methodology for this assay followed previously described protocols with minor adjustments.[173] For each variant (Pik TE WT, S148C, R1, R2, R3), 0.2 mg/mL of protein in storage buffer (HEPES (50 mM), NaCl (150 mM), EDTA (1 mM), glycerol (20% v/v),

pH 7.2) was utilized alongside 10X SYPRO in storage buffer, in a 384-well plate format, 10 uL total volume in each well. Each protein was prepared and analyzed in quadruplicates, using QuantStudio 7 real-time PCR instrument (Center for Structural Biology at UM). The melt curve stage started at 25 ºC for 15 secs then ramped from 25 ºC to 95 ºC for 15 secs with recordings at 0.03 ºC/s.

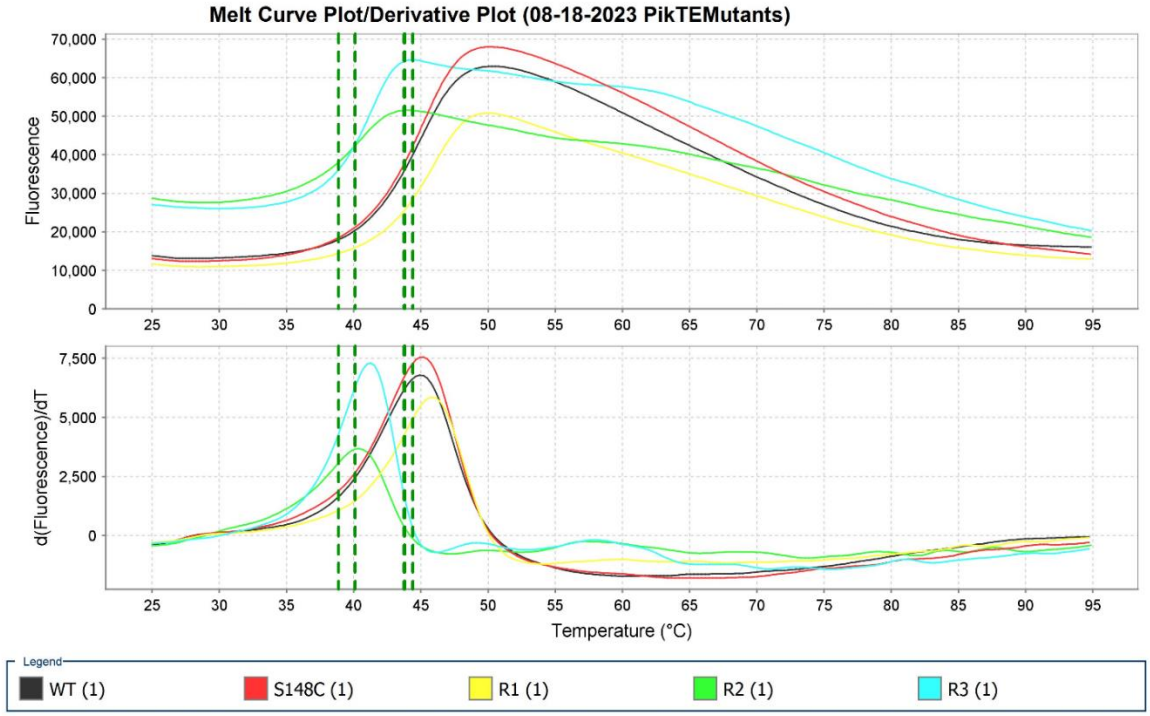
Table III.5 Melting temperatures (Tm B) in quadruplicates for each protein and melting temperatures derivative (Tm D).

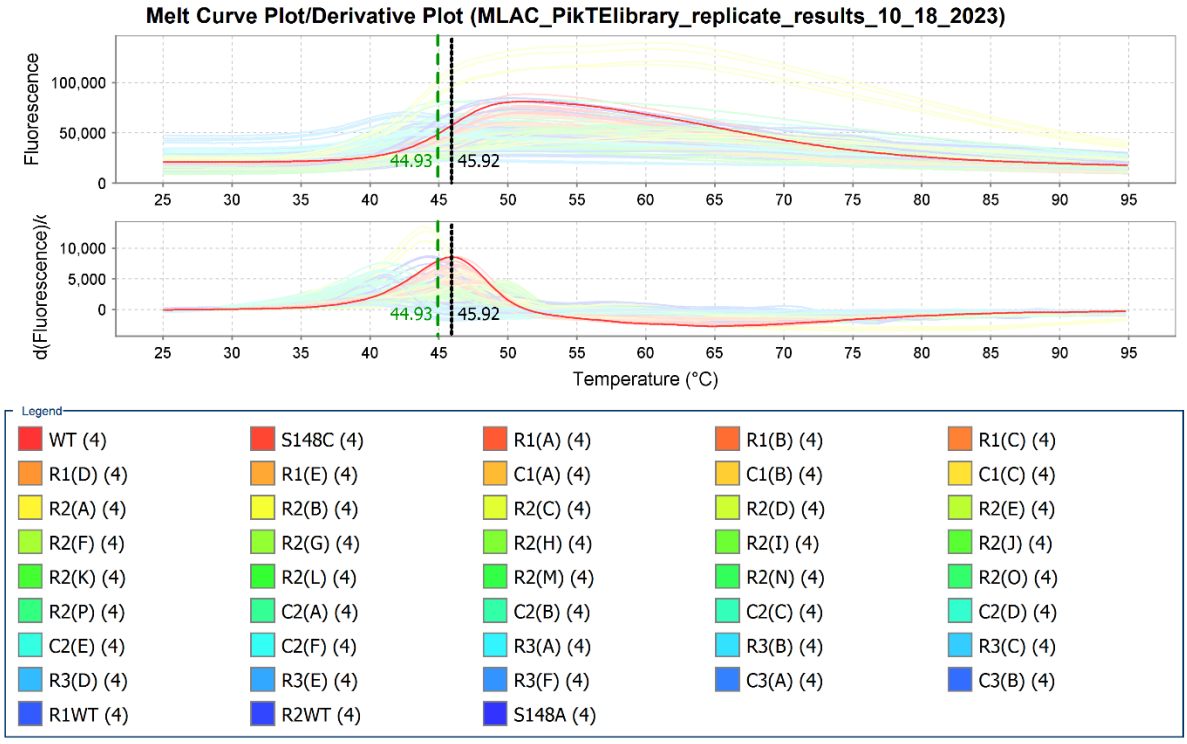
Protein (0.2 mg/mL)	**SYPRO**	**Tm B**	**B Fit**	**Tm B Start**	**Tm B End**	**Tm D**
WT	10X	43.72432	1.355092	28.9373	51.15715	45.04542
WT	10X	43.93996	1.386368	29.43245	51.5904	45.18533
WT	10X	43.80145	1.404628	29.12298	51.5904	44.90551
WT	10X	44.07802	1.376236	29.18488	51.5904	45.18533
S148C	10X	43.88963	1.372849	28.9373	51.71419	45.18533
S148C	10X	43.86144	1.366409	29.12298	51.5904	45.18533
S148C	10X	43.66988	1.365381	27.82322	51.71419	45.04542
S148C	10X	43.74246	1.357278	28.87541	51.65229	45.04542
R1	10X	44.39959	1.363241	29.49435	51.5904	45.88489
R1	10X	44.32479	1.382902	29.49435	51.65229	45.74498
R1	10X	44.57628	1.359779	30.23707	51.83798	45.88489
R1	10X	44.52873	1.377424	29.49435	51.83798	45.88489
R2	10X	38.83318	1.2829	28.19458	46.14375	40.42832

R2	10X	38.88606	1.318599	29.12298	46.14375	40.42832
R2	10X	38.81594	1.286083	28.44215	45.95807	40.28841
R2	10X	38.86703	1.289069	28.81352	45.83429	40.42832
R3	10X	39.97553	1.404569	28.38026	47.13405	41.12788
R3	10X	40.11462	1.360404	29.30867	47.25784	41.2678
R3	10X	39.97742	1.421399	27.88511	47.6292	41.12788
R3	10X	40.16014	1.368638	28.25647	47.50542	41.2678

Table III.6 Melting temperatures (Tm B) mean values and corresponding standard (Std.) error values.

Protein	**Tm B Mean**	**Tm B Lower 95%**	**Tm B Upper 95%**	**Tm B Median**	**Tm B Std. Error**	**Tm B Min**	**Tm B Max**
WT	43.88594	43.63759	44.13428	43.87071	0.078036	43.72432	44.07802
S148C	43.79085	43.62724	43.95446	43.80195	0.05141	43.66988	43.88963
R1	44.45735	44.27327	44.64143	44.46416	0.057841	44.32479	44.57628
R2	38.85055	38.79997	38.90114	38.85011	0.015896	38.81594	38.88606
R3	40.05693	39.90617	40.20768	40.04602	0.047372	39.97553	40.16014


Melt Curve Plot/Derivative Plot (08-18-2023 PikTEMutants)
Fluorescence
d(Fluorescence)/dT
Temperature (°C)
Legend
WT (1)
S148C (1)
R1 (1)
R2 (1)
R3 (1)


Melt Curve Plot/Derivative Plot (MLAC_PikTElibrary_replicate_results_10_18_2023)
Fluorescence
d(Fluorescence)/
44.93
45.92
Temperature (°C)
Legend
WT (4)
S148C (4)
R1(A) (4)
R1(B) (4)
R1(C) (4)
R1(D) (4)
R1(E) (4)
C1(A) (4)
C1(B) (4)
C1(C) (4)
R2(A) (4)
R2(B) (4)
R2(C) (4)
R2(D) (4)
R2(E) (4)
R2(F) (4)
R2(G) (4)
R2(H) (4)
R2(I) (4)
R2(J) (4)
R2(K) (4)
R2(L) (4)
R2(M) (4)
R2(N) (4)
R2(O) (4)
R2(P) (4)
C2(A) (4)
C2(B) (4)
C2(C) (4)
C2(D) (4)
C2(E) (4)
C2(F) (4)
R3(A) (4)
R3(B) (4)
R3(C) (4)
R3(D) (4)
R3(E) (4)
R3(F) (4)
C3(A) (4)
C3(B) (4)
R1WT (4)
R2WT (4)
S148A (4)

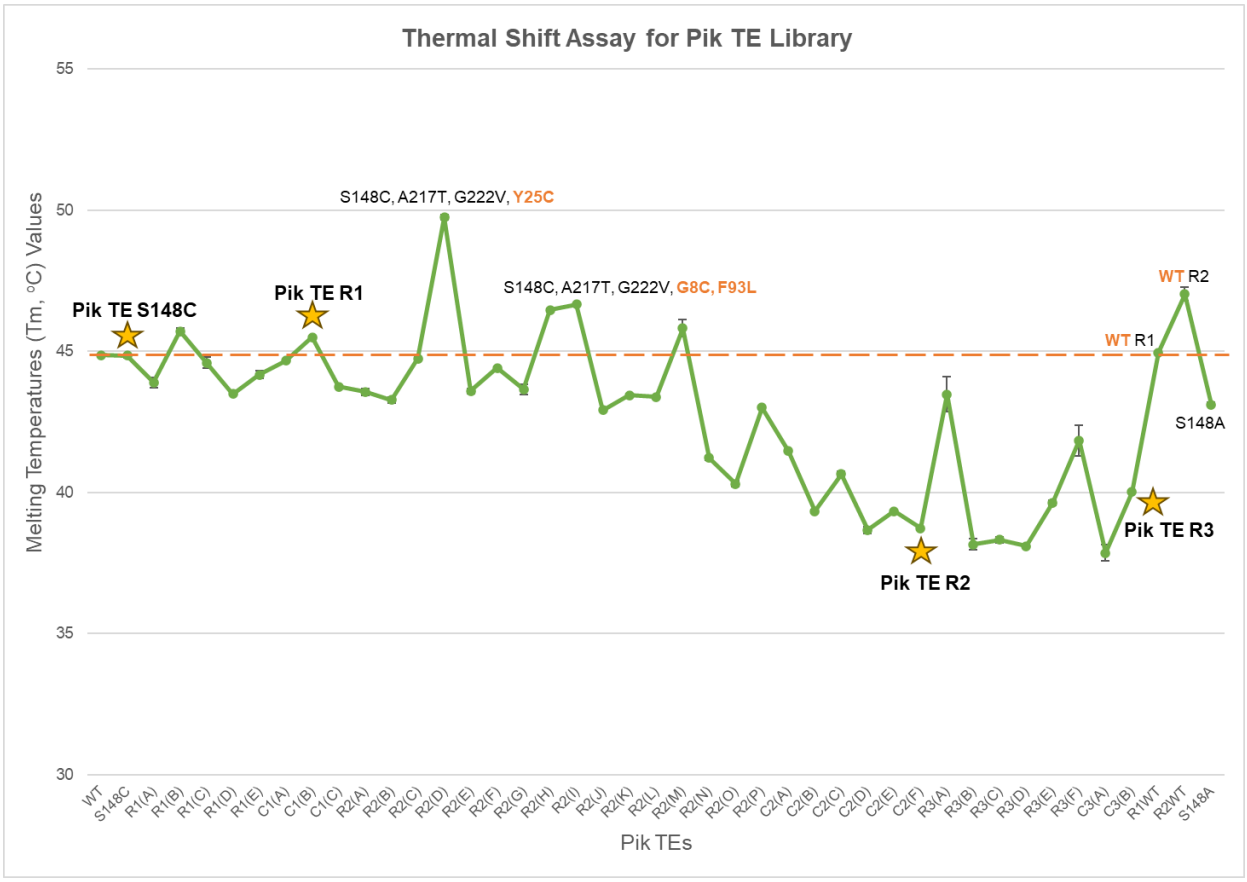


Figure III.18 Melting curve plot and derivative plot for fluorescence emission readout for thermal shift assay Pik TE WT and variants (S148C, R1, R2, R3). Each curve represents the Tm average value for each protein. Also, melting curve plot on the complete Pik TE library and Tm values which followed same protocol as described before.

Reaction Time Courses: Substrate and Hybrid 12-membered Macrolactone/Lactam Ring Concentrations (µM) as a Function of Time (min)

For the reaction time courses, analytical scale 300 µL enzymatic reactions were done by adding water, sodium phosphate reaction buffer (400 mM $NaH_2PO_4•H_2O$ and Na_2HPO_4, 200 mM NaCl, pH 7.2), 1 mM amide hexaketide (substrate) and lastly 10 µM of desired Pik TE variant, 3 hours (taking 20 µL aliquots every 10 minutes), stationary, rt. All reactions were performed and analyzed

in triplicates. After taking the 10 mins aliquots, the reactions were quenched by adding 3 volumes of MeOH (60 µL) and centrifuged at 13,000 x g, 10 mins, 4 ºC for further HPLC analysis using 230 nm UV detection. Analytical HPLC conditions were as follows: mobile phase (A = deionized water + 0.1% formic acid, B = 100% acetonitrile); 25% B for 1.0 min, 25% to 70% B over 14.0 min, 70% B for 2.0 min; 0.4 mL/min; injection volume 50 µL. Note: Only the substrate and macrolactone product µM values were efficiently determined, due to the hydrolysis product being unstable under the reaction conditions resulting on undesired degradation products (see **Figure III.16**).

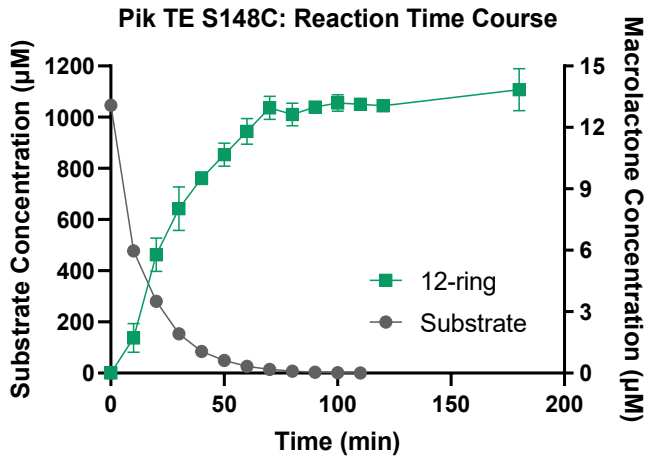


Figure III.19 Pik TE S148C reaction time course for substrate and macrolactone concentration (µM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error).

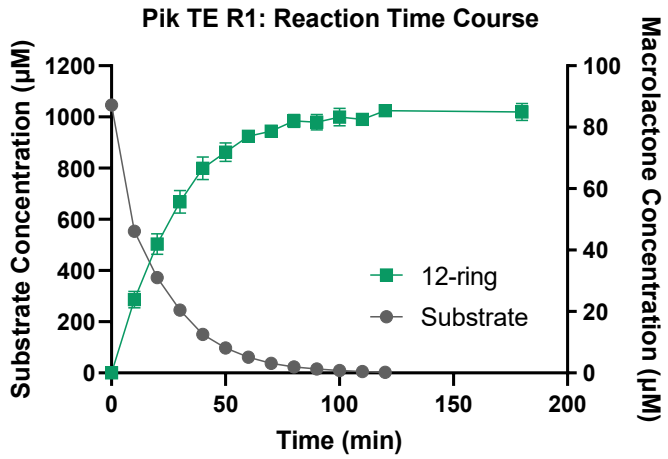


Figure III.20 Pik TE R1 reaction time course for substrate and macrolactone concentration (µM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error).

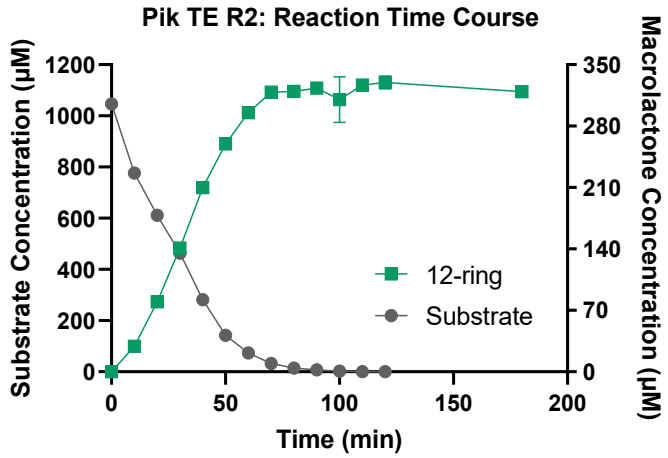


Figure III.21 Pik TE R2 reaction time course for substrate and macrolactone concentration (µM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error).

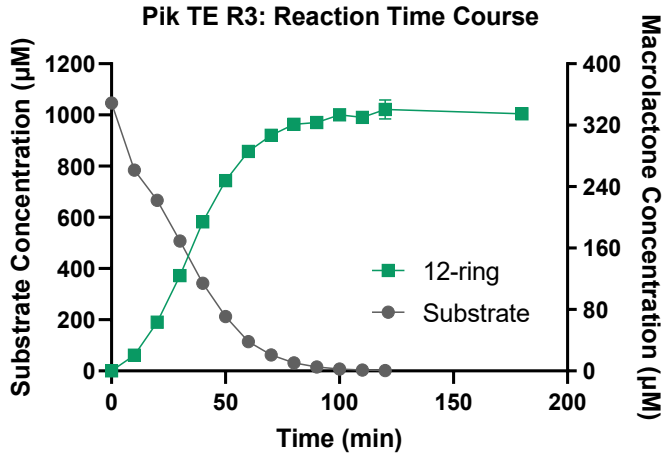


Figure III.22 Pik TE R3 reaction time course for substrate and macrolactone concentration (µM): Time points recorded every 10 minutes, with error bars representing standard deviation (error bars not visible, minimal error).

Initial Rates (µM/min) for Substrate Consumption and Macrolactone Formation

The initial rates were determined from the reaction time courses: substrate and hybrid 12-membered macrolactone/lactam ring concentrations (µM) as a function of time (min) shown above. The plots for initial rates (µM/min) of substrate consumption and macrolactone formation for the Pik TE variants was determined from the concentration (µM) over time (min) in which the rates remained linear, with error bars representing standard deviation. The linear equations ($y = mx + b$) and R^2 values for the plots were determined, where the slope (m) = the initial rate in µM/min.

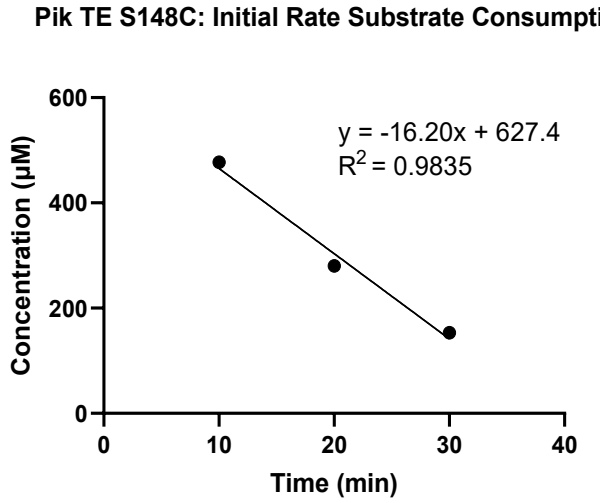


Figure III.23 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE S148C, with error bars representing standard deviation (error bars not visible, minimal error).

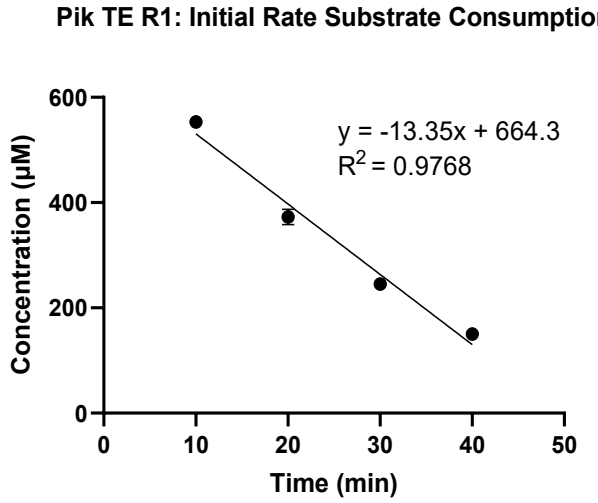


Figure III.24 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE R1 (S148C, A217T, G222V) with error bars representing standard deviation.

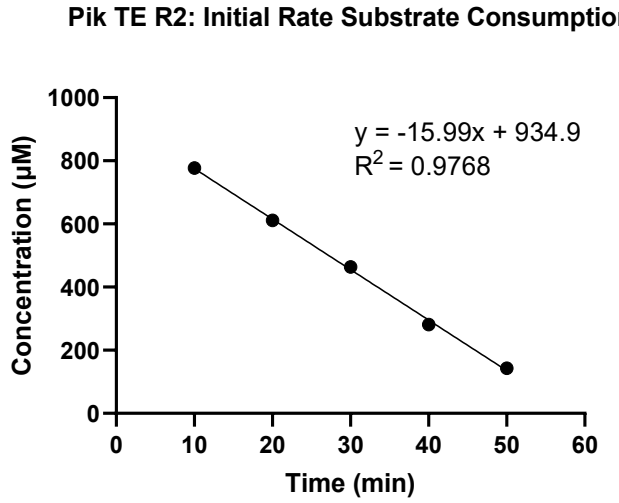


Figure III.25 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE R2 (S148C, A217T, G222V, M271V, Y25C, L126V) with error bars representing standard deviation (error bars not visible, minimal error).

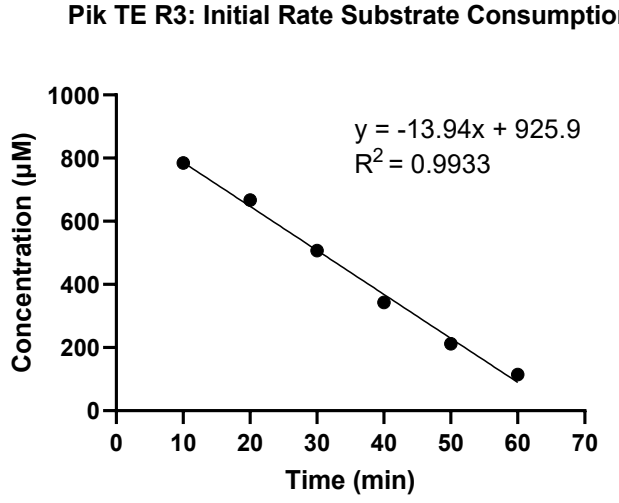


Figure III.26 Plot of amide hexaketide substrate concentration over time in reactions with Pik TE R3 (S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W) with error bars representing standard deviation (error bars not visible, minimal error).

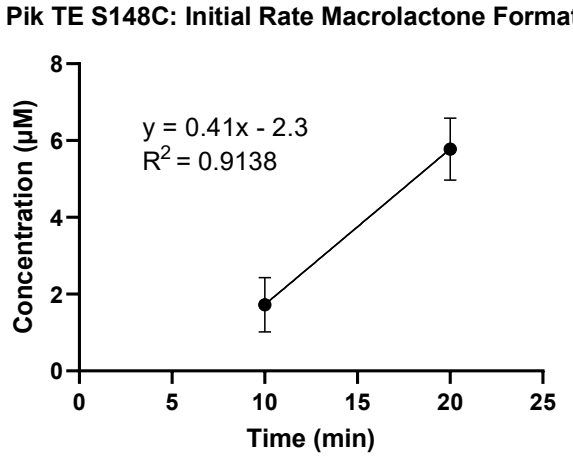


Figure III.27 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE S148C, with error bars representing standard deviation.

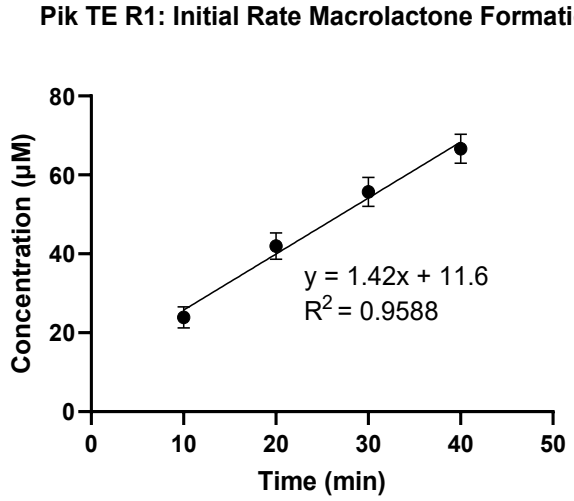


Figure III.28 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE R1 (S148C, A217T, G222V) with error bars representing standard deviation.

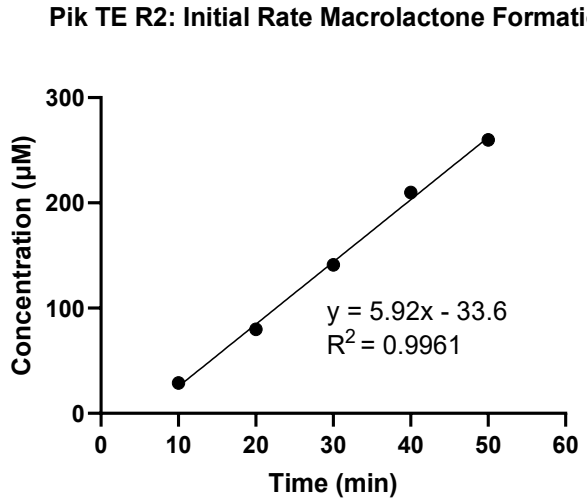


Figure III.29 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE R2 (S148C, A217T, G222V, M271V, Y25C, L126V) with error bars representing standard deviation (error bars not visible, minimal error).

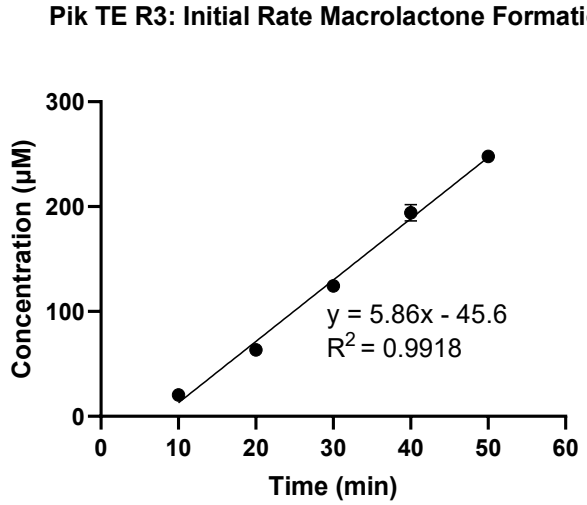


Figure III.30 Plot of hybrid 12-membered macrolactone/lactam ring product concentration over time in reactions with Pik TE R3 (S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W) with error bars representing standard deviation.

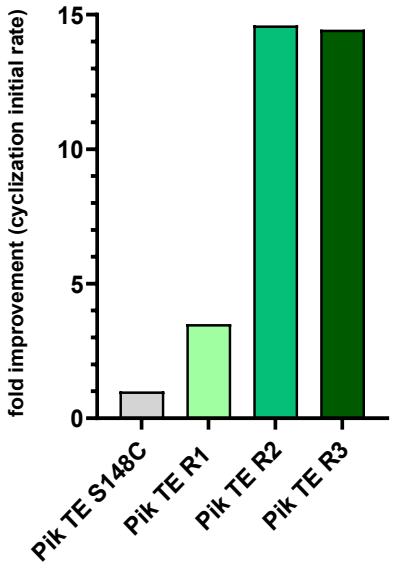


Figure III.31 12-membered macrolactone/lactam ring initial rates (μM/min) in terms of fold improvement for the engineered Pik TE variants.

III.9.8 X-ray Crystallography for Hybrid 12-membered Macrolactone/Lactam Ring Molecule

Experimental for crystal formation and x-ray structure determination

The crystals were grown through vapor diffusion method. In a 1 dram vial, the macrolactone/lactam (**3.2**) was dissolved utilizing a minimum volume of CH_2Cl_2, then placed inside of a 4 dram vial containing Et_2O (~5X v/v). These grew at room temperature (25 ºC) for 3.5 weeks. Single colorless block-shaped crystals of **3.2** were used as supplied. A suitable crystal with dimensions 0.17 × 0.17 × 0.14 mm^3 was selected and mounted on a dtrek-CrysAlisPro-abstract goniometer imported rigaku-d*trek images diffractometer. The crystal was kept at a steady T = 85 K during data collection. The structure was solved with the ShelXT 2018/2 solution program[206] using Intrinsic Phasing methods and by using Olex2 1.5-alpha as the graphical interface.[207] The

model was refined with ShelXL 2018/3 using full matrix least squares minimization[208] on F^2. Cambridge Crystallographic Data Centre (CCDC) X-Ray **Deposition Number 2294580**.

Table III.7 Crystallographic information and data collected.

Compound	**hybrid 12-membered macrolactone/lactam ring**
Formula	$C_{15}H_{25}NO_4$
$D_{calc.}$/ g cm^{-3}	1.237
μ/mm^{-1}	0.724
Formula Weight	283.36
Color	colorless
Shape	block-shaped
Size/mm^3	0.17×0.17×0.14
T/K	85
Crystal System	orthorhombic
Flack Parameter	0.00(10)
Hooft Parameter	0.09(4)
Space Group	$P2_12_12_1$
a/Å	8.10460(10)

b/Å	8.8948(2)
c/Å	21.1134(4)
α/°	90
β/°	90
γ/°	90
V/Å^3	1522.04(5)
Z	4
Z'	1
Wavelength/Å	1.54184
Radiation type	Cu K_α
Θ_{min}/°	4.188
Θ_{max}/°	69.448
Measured Refl's.	22570
Indep't Refl's	2836
Refl's I≥2 σ(I)	2824
R_{int}	0.0441
Parameters	191

Restraints	0
Largest Peak	0.185
Deepest Hole	-0.206
GooF	1.117
wR_2 (all data)	0.0782
wR_2	0.0781
R_1 (all data)	0.0315
R_1	0.0314

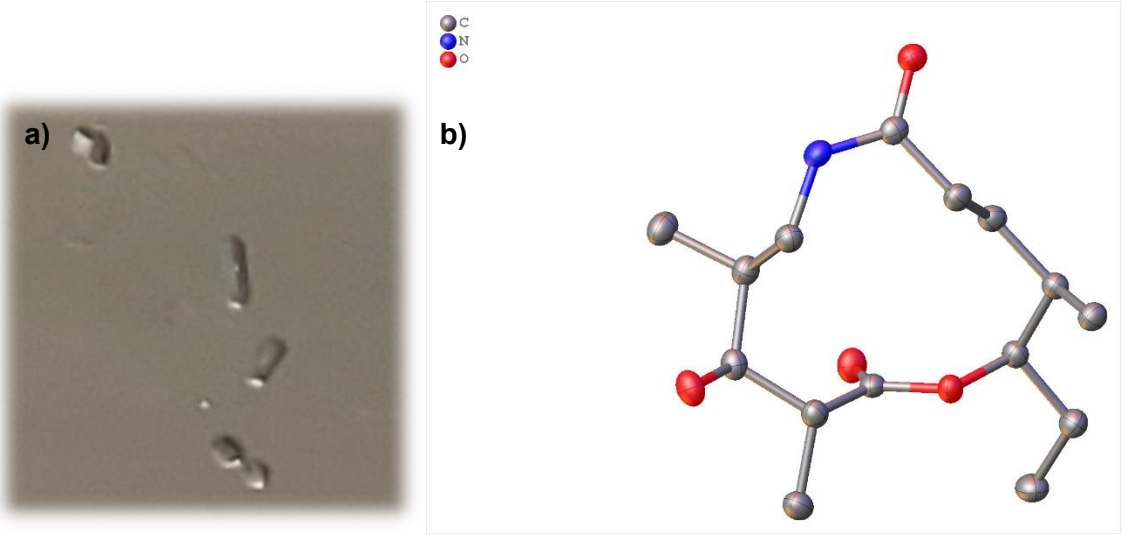


Figure III.32 a) Example of the colorless block-shaped crystals obtained for **3.2**. b) X-ray crystal structure of **3.2**.

Table III.8 Fractional Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Parameters ($Å^2 \times 10^3$) for **3.2**. U_{eq} is defined as 1/3 of the trace of the orthogonalized U_{ij}.

Atom	x	y	z	U_{eq}
O1	6849.3(16)	4035.3(14)	3417.8(6)	22.8(3)
O2	4762.4(17)	3134.1(16)	2819.1(6)	27.7(3)
O3	3578.6(19)	695.4(15)	4527.2(6)	29.0(3)
O4	2083.6(16)	7878.9(15)	4402.2(6)	26.4(3)
N1	2176(2)	5403.2(19)	4637.0(8)	24.9(4)
C1	2863(2)	6686(2)	4430.0(8)	22.1(4)
C2	2950(2)	3917(2)	4625.8(9)	24.2(4)
C3	2474(2)	2976(2)	4044.2(9)	24.4(4)
C4	3598(2)	1596(2)	3964.0(9)	24.0(4)
C5	5414(2)	1967(2)	3822.1(9)	22.7(4)
C6	5593(2)	3099(2)	3294.1(8)	21.1(4)
C7	7103(2)	5336(2)	3007.6(8)	22.4(4)
C8	6937(2)	6721(2)	3437.8(9)	22.5(4)
C9	5166(2)	6806(2)	3662.1(9)	23.4(4)
C10	4655(2)	6613(2)	4251.5(9)	23.5(4)
C11	669(3)	2480(3)	4073.7(10)	32.1(5)
C12	6378(3)	552(2)	3634.3(10)	28.8(4)
C13	8774(2)	5174(2)	2686.9(9)	26.8(4)
C14	8861(3)	3821(2)	2246.0(10)	31.6(5)
C15	8196(2)	6755(2)	3974.1(9)	25.3(4)

Table III.9 Anisotropic Displacement Parameters ($\times 10^4$) for **3.2**. The anisotropic displacement factor exponent takes the form: $-2p^2[h^2a^{*2} \times U_{11} + ... + 2hka^* \times b^* \times U_{12}]$

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	20.0(6)	20.8(6)	27.6(7)	3.0(5)	-2.8(5)	-1.6(5)
O2	27.9(7)	30.8(7)	24.3(6)	0.5(6)	-4.8(6)	-5.2(6)
O3	36.5(8)	24.5(7)	26.1(7)	4.1(5)	-1.2(6)	-5.8(6)
O4	22.5(7)	24.6(7)	32.0(7)	0.2(5)	0.0(5)	0.9(5)
N1	19.3(8)	26.4(8)	28.9(8)	-1.2(7)	4.5(7)	0.5(7)
C1	21.0(9)	25.8(9)	19.6(8)	-2.0(7)	-2.8(7)	0.1(8)
C2	22.3(9)	23.5(9)	26.8(9)	1.4(8)	-0.2(7)	0.5(8)
C3	21.7(9)	26.9(10)	24.8(9)	1.0(7)	-0.5(7)	-1.7(7)
C4	25.9(9)	23.4(9)	22.8(9)	2.1(7)	-0.6(7)	-3.5(8)
C5	21.6(9)	21.8(10)	24.8(9)	0.5(7)	-2.5(7)	-1.0(7)
C6	18.5(8)	19.0(9)	25.8(9)	-2.6(7)	2.2(7)	1.5(7)
C7	21.9(9)	20.3(9)	25.0(9)	2.7(7)	-0.5(7)	-0.2(8)
C8	21.9(9)	19.9(8)	25.7(9)	1.2(7)	0.0(7)	0.5(7)
C9	21.2(9)	22.2(9)	26.7(9)	0.0(7)	-2.8(7)	2.2(8)
C10	19.5(8)	23.9(9)	27.2(9)	-0.1(8)	-3.0(7)	0.3(7)
C11	24.1(10)	36.7(12)	35.4(11)	-5.4(9)	0.1(8)	-3.4(8)
C12	26.7(10)	23.2(10)	36.5(10)	0.2(8)	-2.6(8)	1.5(8)
C13	24.7(9)	25.3(10)	30.4(10)	0.5(8)	3.6(8)	-0.3(8)
C14	30.5(10)	27.7(10)	36.5(11)	-3.2(9)	6.3(9)	3.9(8)
C15	21.6(9)	26.1(9)	28.4(9)	-1.2(8)	-0.7(7)	0.3(8)

Table III.10 Bond Lengths in Å for **3.2**.

Atom	Atom	Length/Å
O1	C6	1.341(2)
O1	C7	1.460(2)
O2	C6	1.208(2)
O3	C4	1.434(2)
O4	C1	1.236(2)
N1	C1	1.343(3)
N1	C2	1.463(3)
C1	C10	1.502(2)
C2	C3	1.535(3)
C3	C4	1.539(3)

Atom	Atom	Length/Å
C3	C11	1.529(3)
C4	C5	1.538(3)
C5	C6	1.509(2)
C5	C12	1.534(3)
C7	C8	1.537(3)
C7	C13	1.521(3)
C8	C9	1.513(2)
C8	C15	1.524(3)
C9	C10	1.323(3)
C13	C14	1.523(3)

Table III.11 Bond Angles in ° for **3.2**.

Atom	Atom	Atom	Angle/°
C6	O1	C7	118.92(14)
C1	N1	C2	125.75(15)
O4	C1	N1	122.16(16)
O4	C1	C10	121.30(18)
N1	C1	C10	116.51(17)
N1	C2	C3	113.42(15)
C2	C3	C4	111.98(15)
C11	C3	C2	111.42(16)
C11	C3	C4	109.90(16)
O3	C4	C3	110.37(14)
O3	C4	C5	106.97(15)
C5	C4	C3	114.63(15)
C6	C5	C4	112.29(15)
C6	C5	C12	107.91(15)
C12	C5	C4	111.19(16)
O1	C6	C5	110.09(15)
O2	C6	O1	124.67(17)
O2	C6	C5	125.21(17)
O1	C7	C8	105.82(13)
O1	C7	C13	108.32(15)
C13	C7	C8	114.64(16)
C9	C8	C7	107.93(15)
C9	C8	C15	113.64(16)
C15	C8	C7	113.36(15)
C10	C9	C8	125.81(17)
C9	C10	C1	122.24(17)
C7	C13	C14	112.86(16)

Table III.12 Torsion Angles in ° for **3.2**.

Atom	Atom	Atom	Atom	Angle/°
O1	C7	C8	C9	-66.61(17)
O1	C7	C8	C15	60.19(19)
O1	C7	C13	C14	64.3(2)
O3	C4	C5	C6	-171.66(14)
O3	C4	C5	C12	67.31(18)
O4	C1	C10	C9	-67.8(3)
N1	C1	C10	C9	114.0(2)
N1	C2	C3	C4	166.66(15)
N1	C2	C3	C11	-69.8(2)
C1	N1	C2	C3	-95.8(2)
C2	N1	C1	O4	173.45(17)
C2	N1	C1	C10	-8.4(3)
C2	C3	C4	O3	55.97(19)
C2	C3	C4	C5	-64.9(2)
C3	C4	C5	C6	-49.0(2)
C3	C4	C5	C12	-169.99(16)
C4	C5	C6	O1	142.75(15)
C4	C5	C6	O2	-39.3(2)
C6	O1	C7	C8	119.53(16)
C6	O1	C7	C13	-117.09(17)
C7	O1	C6	O2	9.9(3)
C7	O1	C6	C5	-172.21(14)
C7	C8	C9	C10	112.8(2)
C8	C7	C13	C14	-177.80(16)
C8	C9	C10	C1	-178.69(18)
C11	C3	C4	O3	-68.43(19)
C11	C3	C4	C5	170.72(16)
C12	C5	C6	O1	-94.35(17)
C12	C5	C6	O2	83.6(2)
C13	C7	C8	C9	174.10(16)
C13	C7	C8	C15	-59.1(2)
C15	C8	C9	C10	-13.8(3)

Table III.13 Hydrogen Fractional Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Parameters ($Å^2 \times 10^3$) for **3.2**. U_{eq} is defined as 1/3 of the trace of the orthogonalized U_{ij}.

Atom	x	y	z	U_{eq}
H3	3117.5	-130.54	4450.42	44
H1	1160(30)	5480(30)	4744(10)	22(5)
H2A	2628.75	3363.08	5013.34	29
H2B	4163.28	4042.9	4633.09	29
H3A	2610.25	3624.64	3660.76	29
H4	3158.33	974.9	3606.97	29
H5	5923.86	2396.02	4213.74	27
H7	6218.02	5362.06	2677.71	27
H8	7142.67	7627.01	3168.43	27
H9	4349.12	7017.38	3352.01	28
H10	5448.34	6422.14	4572.8	28
H11A	352.25	2027.72	3668.24	48
H11B	-33	3354.19	4158.51	48
H11C	530.52	1739.35	4413.28	48
H12A	5956.54	167.96	3230.4	43
H12B	6242.93	-217.49	3962.57	43
H12C	7550.28	800.28	3589.97	43
H13A	9637.69	5075.38	3015.83	32
H13B	9008.32	6096.99	2440.93	32
H14A	8019.59	3917.71	1915.33	47
H14B	8664.73	2899.22	2488.99	47
H14C	9956.4	3774.56	2050.02	47
H15A	8006.95	7645.63	4237.64	38
H15B	9311.61	6793.56	3796.21	38
H15C	8077.15	5847.61	4233.43	38

Crystal Data

$C_{15}H_{25}NO_4$, M_r = 283.36, orthorhombic, $P2_12_12_1$ (No. 19), a = 8.10460(10) Å, b = 8.8948(2) Å, c = 21.1134(4) Å, $\alpha = \beta = \gamma = 90^\circ$, V = 1522.04(5) $Å^3$, T = 85 K, Z = 4, Z' = 1, μ(Cu K_α) = 0.724, 22570 reflections measured, 2836 unique (R_{int} = 0.0441) which were used in all calculations. The final wR_2 was 0.0782 (all data) and R_1 was 0.0314 (I≥2 σ(I)).

Structure Quality Indicators

Reflections:	d min (Cu\a) 2Θ=138.9°	0.82	I/σ(I)	49.4	Rint m=7.99	4.41%	Full 135.4° 99.9
Refinement:	Shift	0.000	Max Peak	0.2	Min Peak	-0.2	GooF 1.117	Hooft .09(4)

A colorless block-shaped-shaped crystal with dimensions 0.17 × 0.17 × 0.14 mm^3 was mounted. Data were collected using a dtrek-CrysAlisPro-abstract goniometer imported rigaku-d*trek images diffractometer operating at T = 85 K.

Data were measured using ω scans with Cu K_α radiation. The diffraction pattern was indexed and the total number of runs and images was based on the strategy calculation from the program CrysAlisPro[209] 1.171.42.93a. The maximum resolution that was achieved was Θ = 69.448° (0.82 Å).

The unit cell was refined using CrysAlisPro 1.171.42.93a on 18014 reflections, 80% of the observed reflections.[209]

Data reduction, scaling and absorption corrections were performed using CrysAlisPro[209] 1.171.42.93a The final completeness is 99.90 % out to 69.448° in Θ. A multi-scan absorption correction was performed using CrysAlisPro 1.171.42.93a. Empirical absorption correction using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm. The absorption coefficient μ of this material is 0.724 mm^{-1} at this wavelength (λ = 1.54184Å) and the minimum and maximum transmissions are 0.939 and 1.000.

The structure was solved and the space group $P2_12_12_1$ (No. 19) determined by the ShelXT 2018/2 structure solution program[206] using Intrinsic Phasing methods and refined by full matrix least squares minimization on F^2 using version 2018/3 of ShelXL 2018/3. All non-hydrogen atoms were

refined anisotropically. Hydrogen atom positions were calculated geometrically and refined using the riding model. Most hydrogen atom positions were calculated geometrically and refined using the riding model, but some hydrogen atoms were refined freely.

_exptl_absorpt_process_details: CrysAlisPro 1.171.42.93a using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm.[209]

There is a single formula unit in the asymmetric unit, which is represented by the reported sum formula. In other words: Z is 4 and Z' is 1. The moiety formula is C15 H25 N O4.

The Flack parameter was refined to 0.00(10). Determination of absolute structure using Bayesian statistics on Bijvoet differences using the Olex2 results in 0.09(4). The chiral atoms in this structure are: C3(S), C4(S), C5(R), C7(R), C8(R). Note: The Flack parameter is used to determine chirality of the crystal studied, the value should be near 0, a value of 1 means that the stereochemistry is wrong, and the model should be inverted. A value of 0.5 means that the crystal consists of a racemic mixture of the two enantiomers.

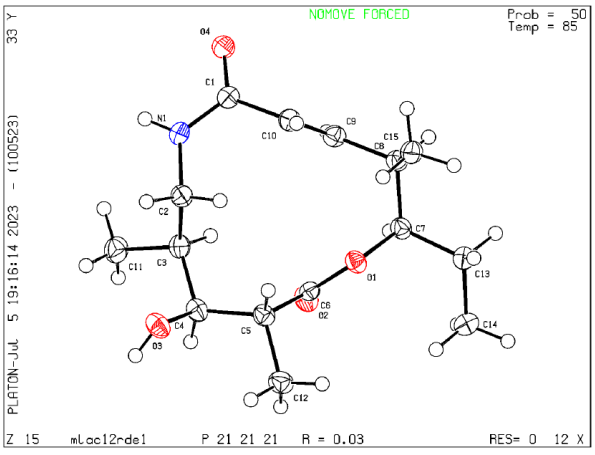
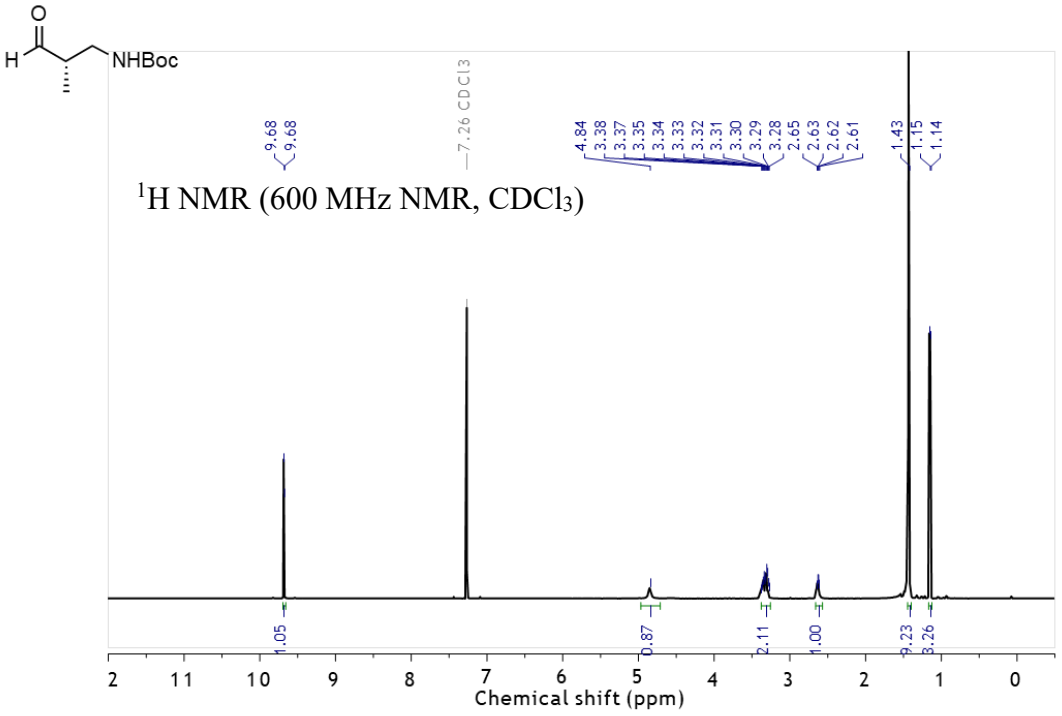
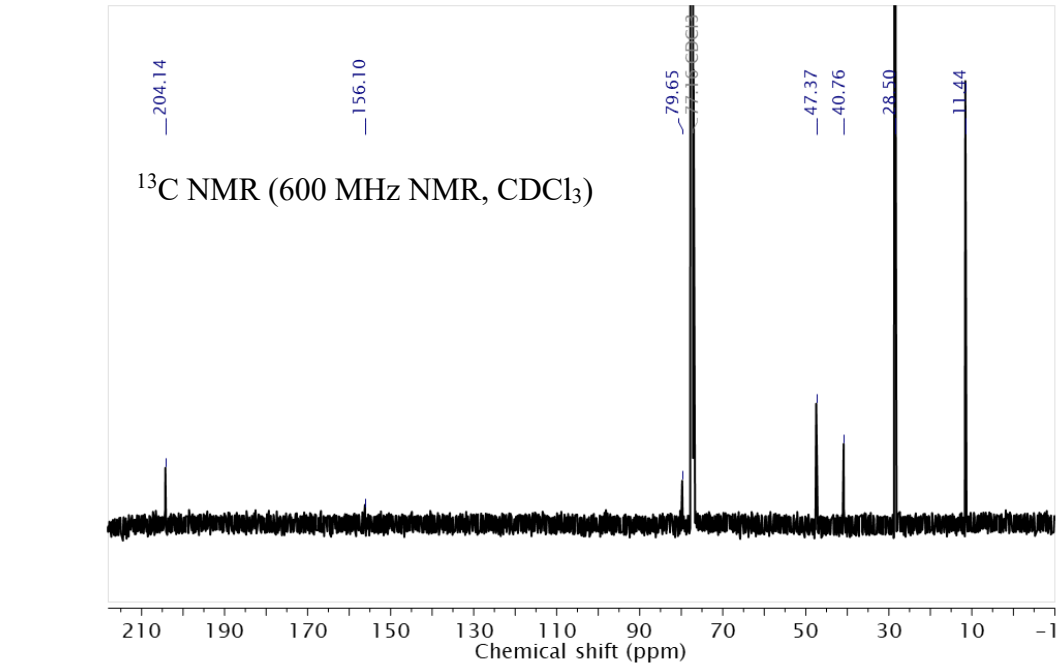


Figure III.33 Ellipsoid plot: X-ray crystal structure of hybrid 12-membered macrolactone/lactam ring (**3.2**).

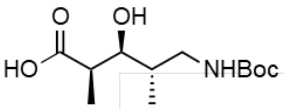
III.10 NMR Spectra



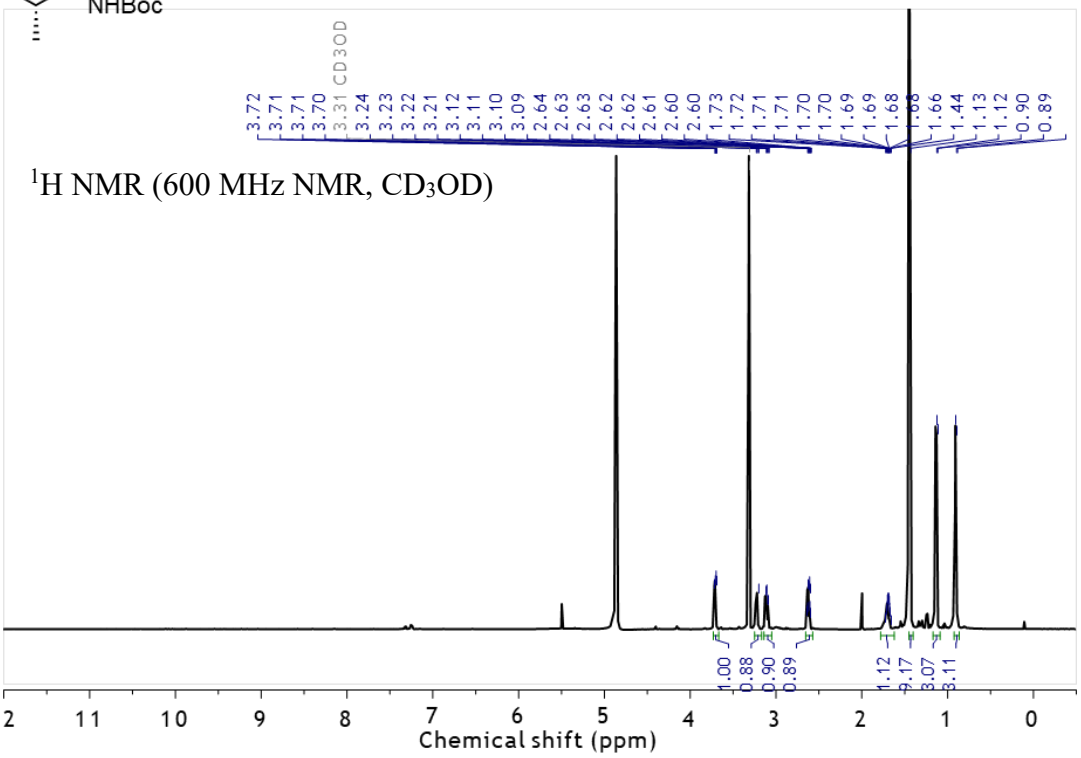
^{1}H NMR (600 MHz NMR, $CDCl_3$)



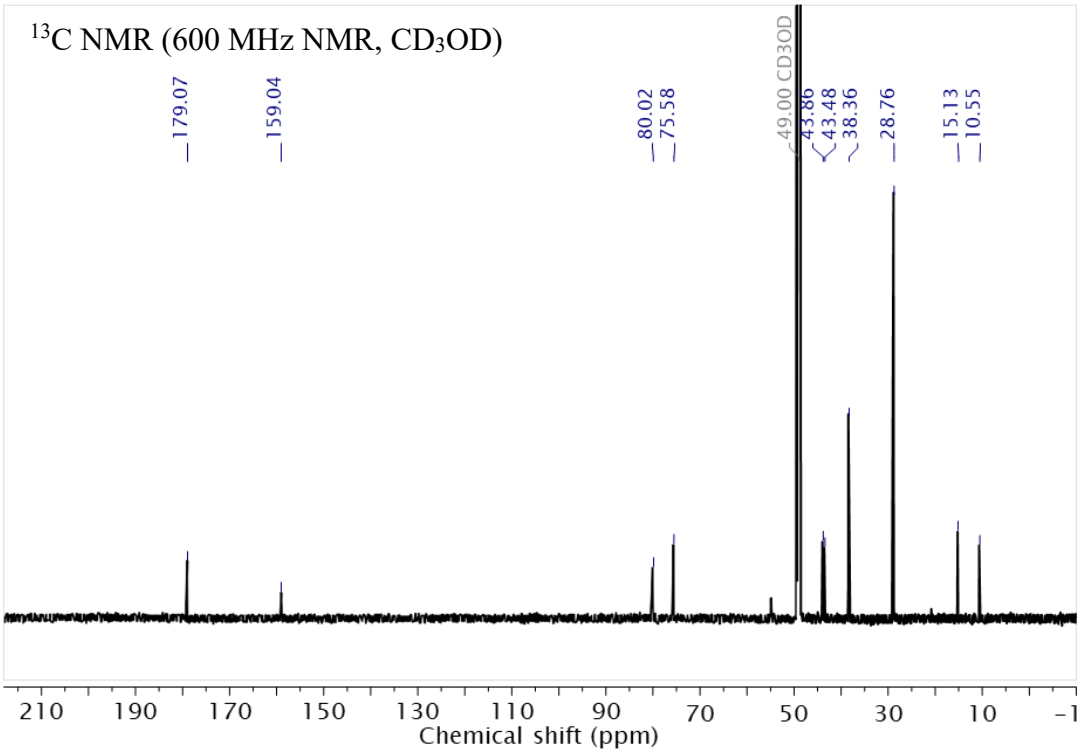
^{13}C NMR (600 MHz NMR, $CDCl_3$)

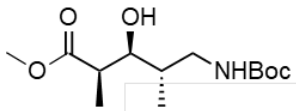


^{1}H NMR (600 MHz NMR, CD_3OD)

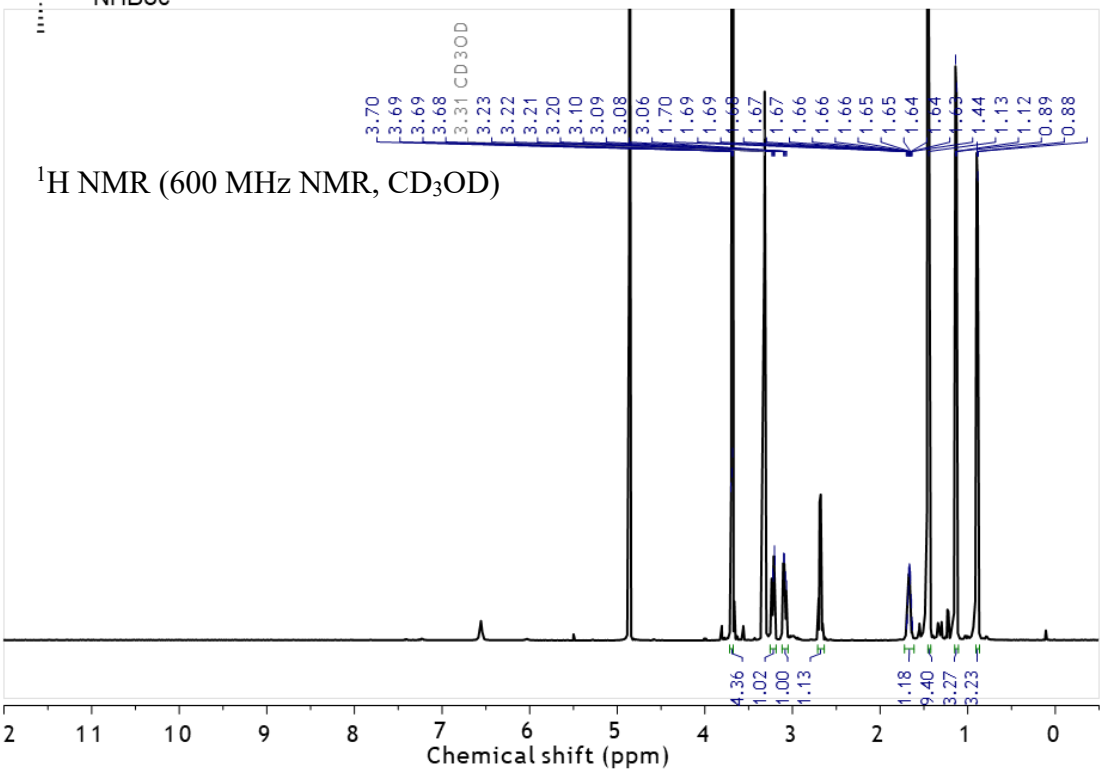


^{13}C NMR (600 MHz NMR, CD_3OD)

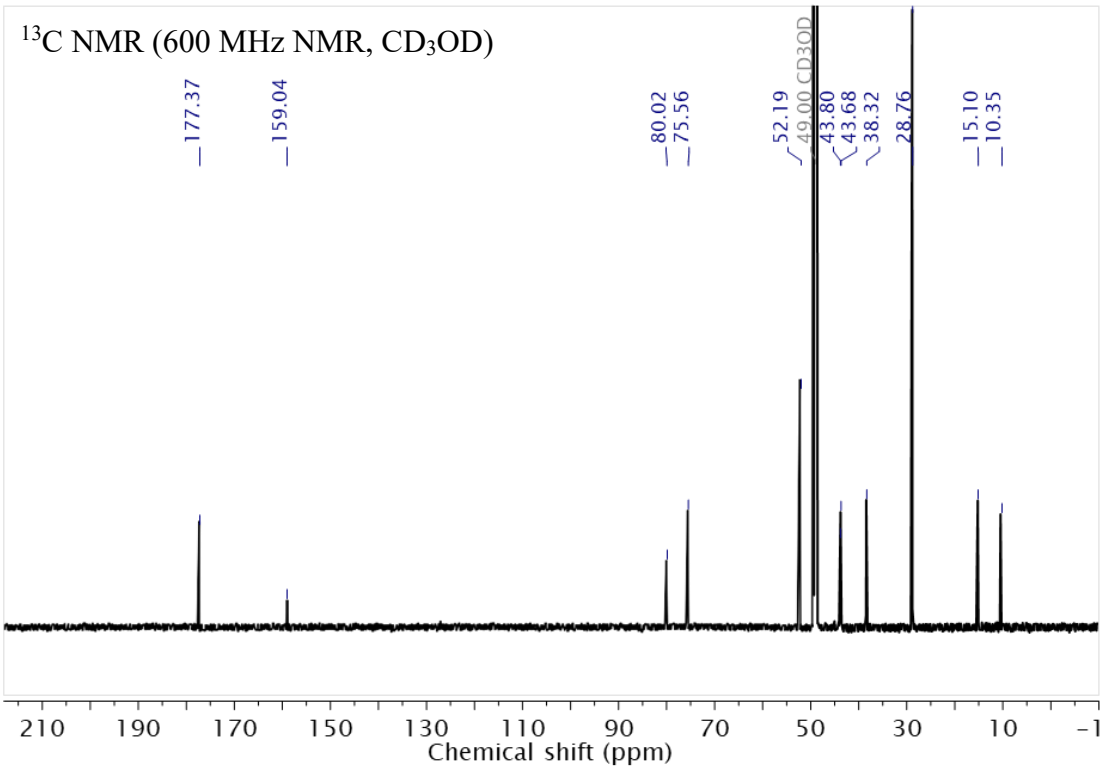


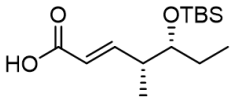


^{1}H NMR (600 MHz NMR, CD_3OD)



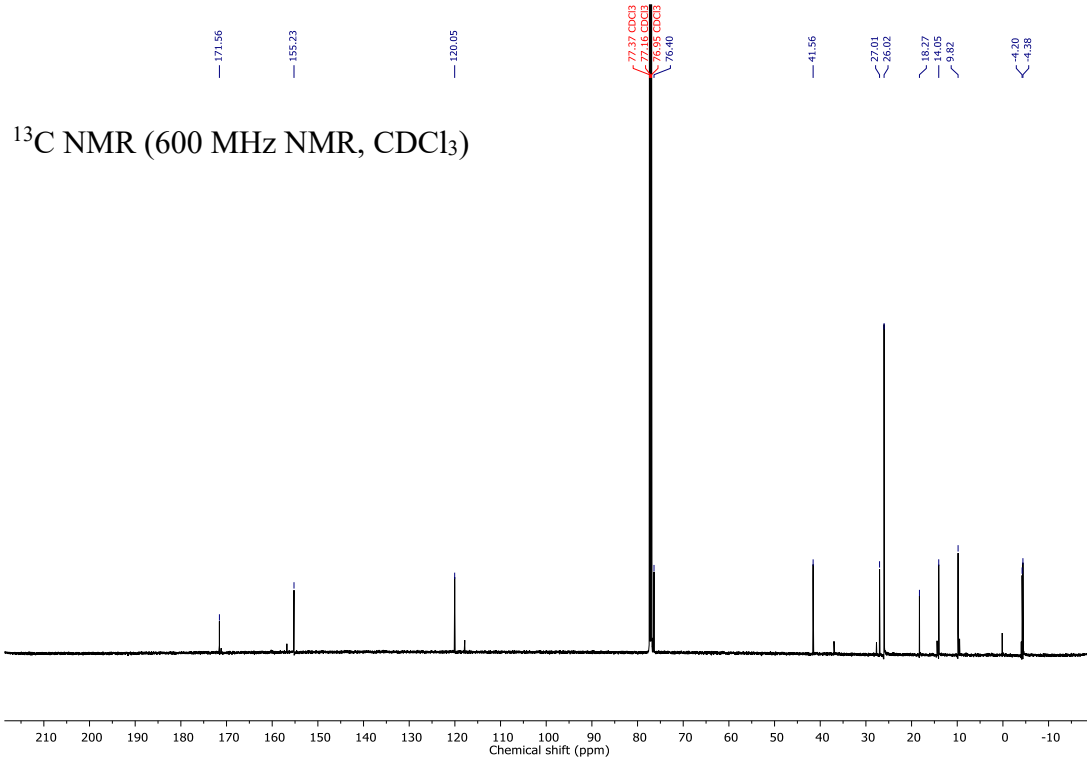
^{13}C NMR (600 MHz NMR, CD_3OD)

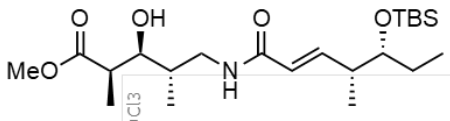


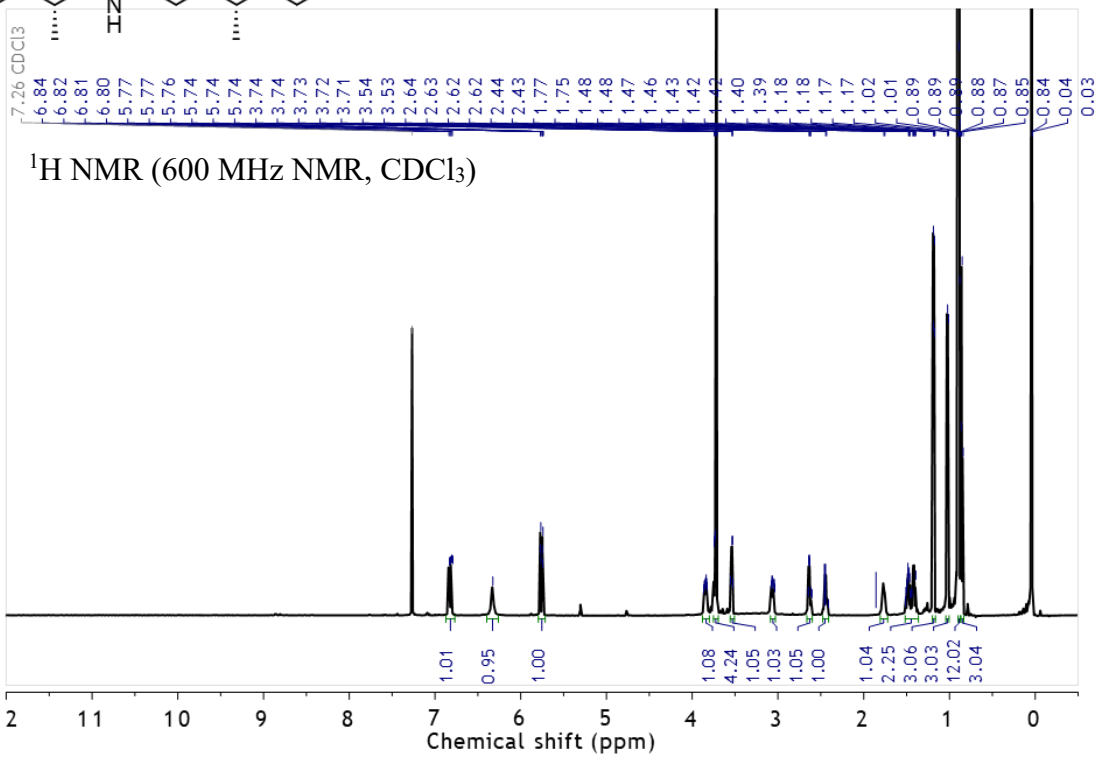


^{1}H NMR (600 MHz NMR, $CDCl_3$)

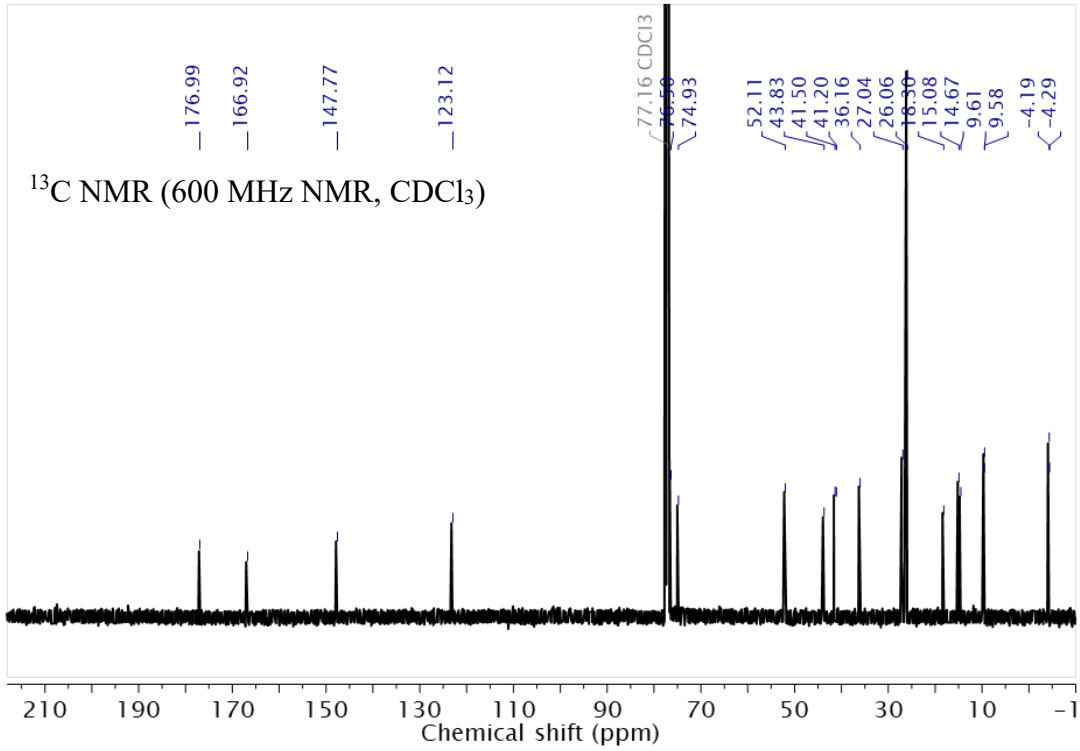
^{13}C NMR (600 MHz NMR, $CDCl_3$)



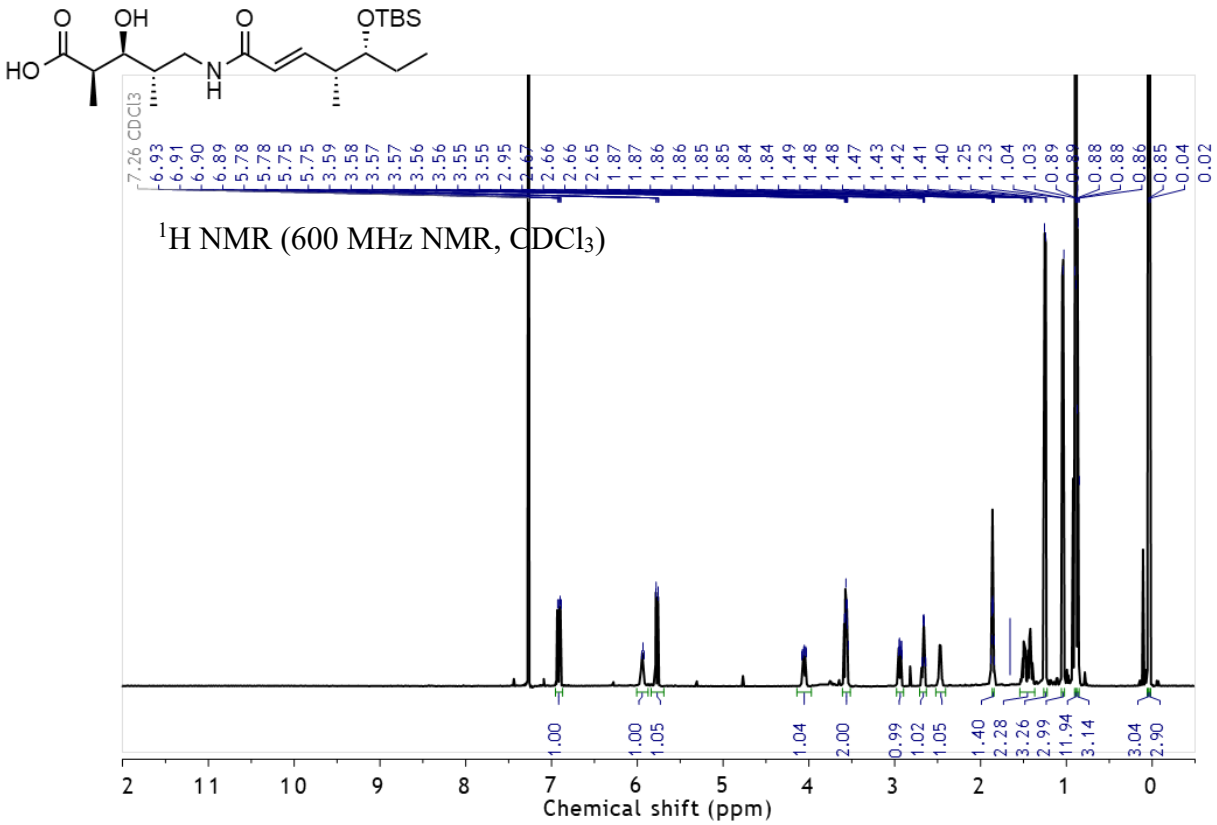




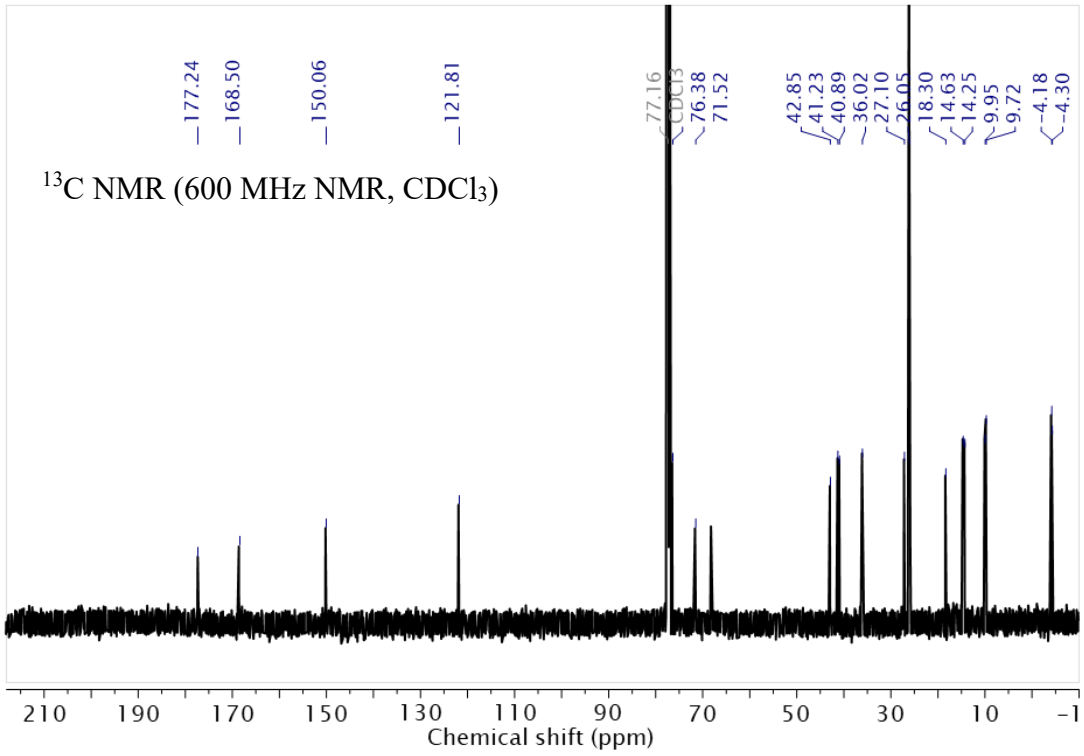
^{1}H NMR (600 MHz NMR, $CDCl_3$)



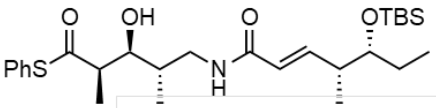
^{13}C NMR (600 MHz NMR, $CDCl_3$)

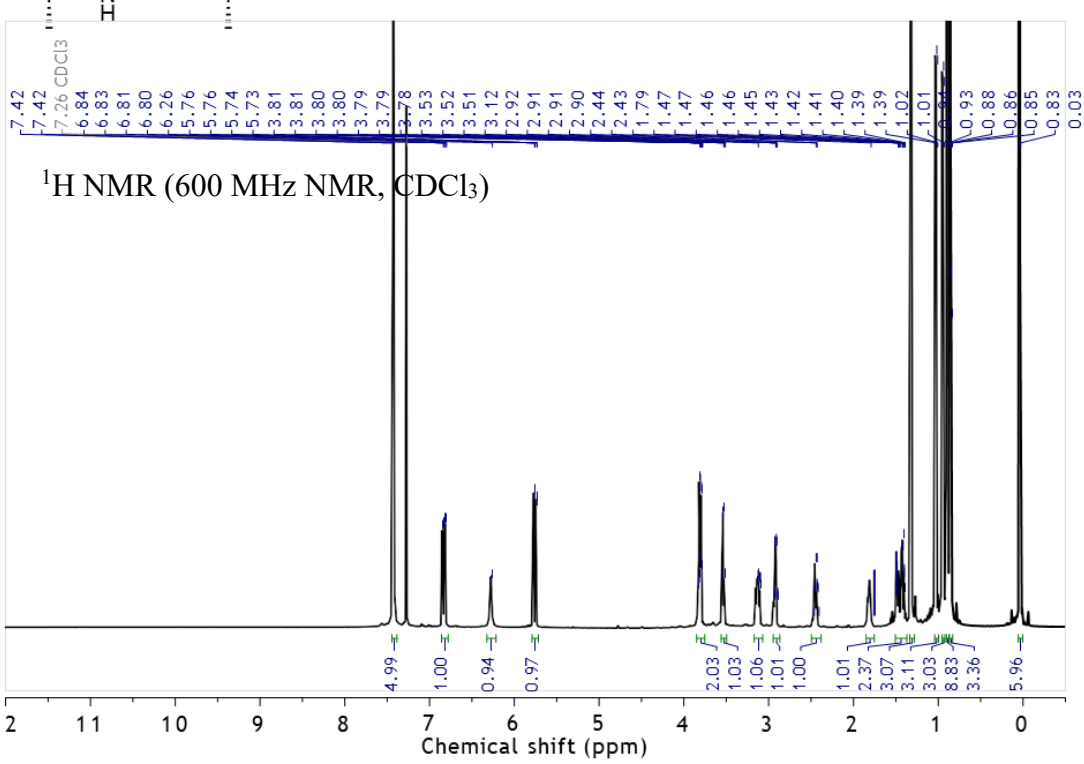


^{1}H NMR (600 MHz NMR, $CDCl_3$)

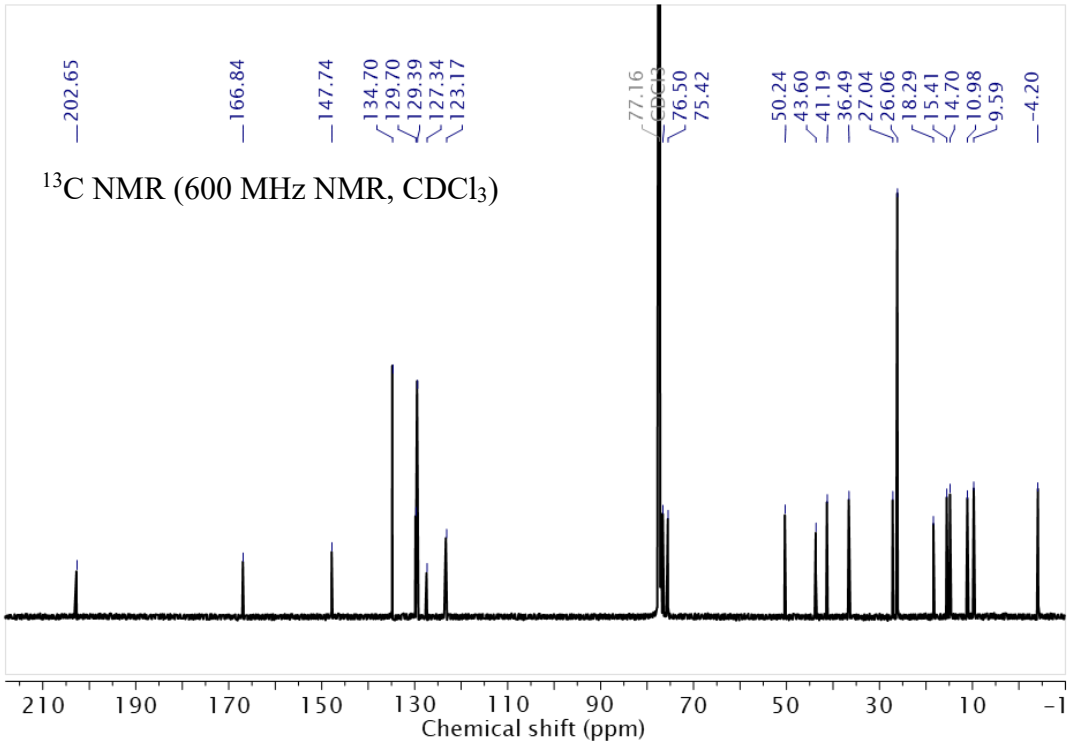


^{13}C NMR (600 MHz NMR, $CDCl_3$)

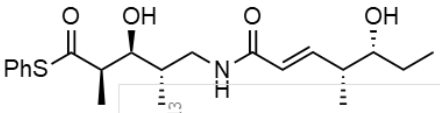


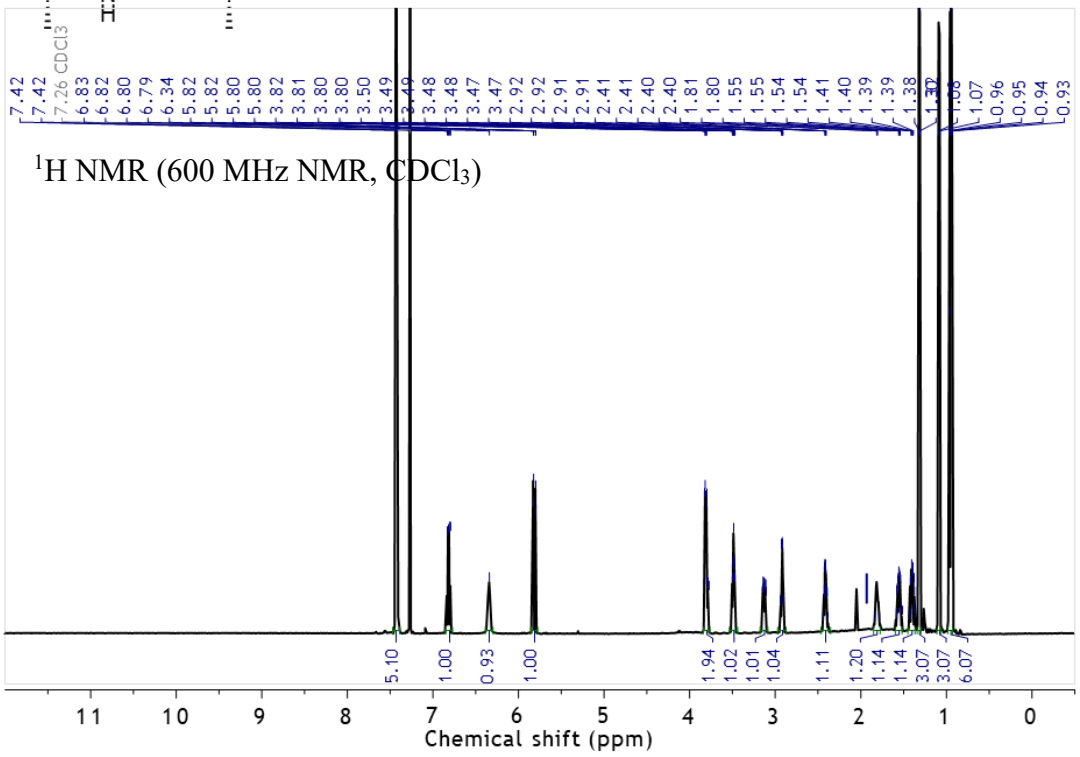


^{1}H NMR (600 MHz NMR, $CDCl_3$)

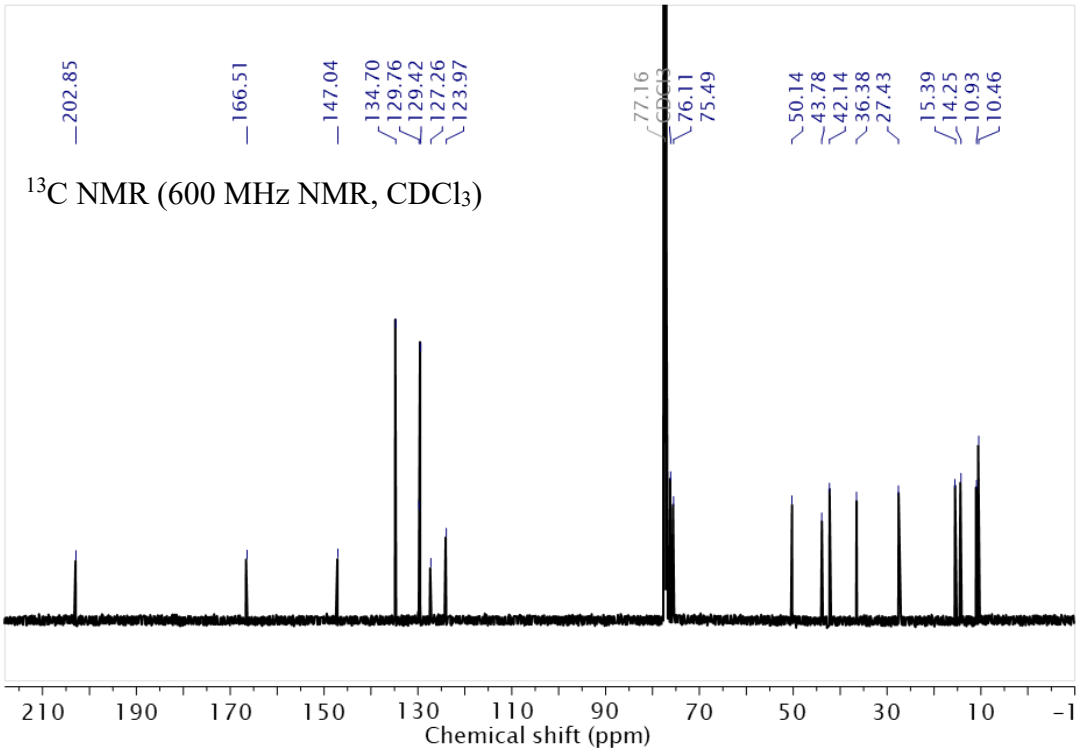


^{13}C NMR (600 MHz NMR, $CDCl_3$)

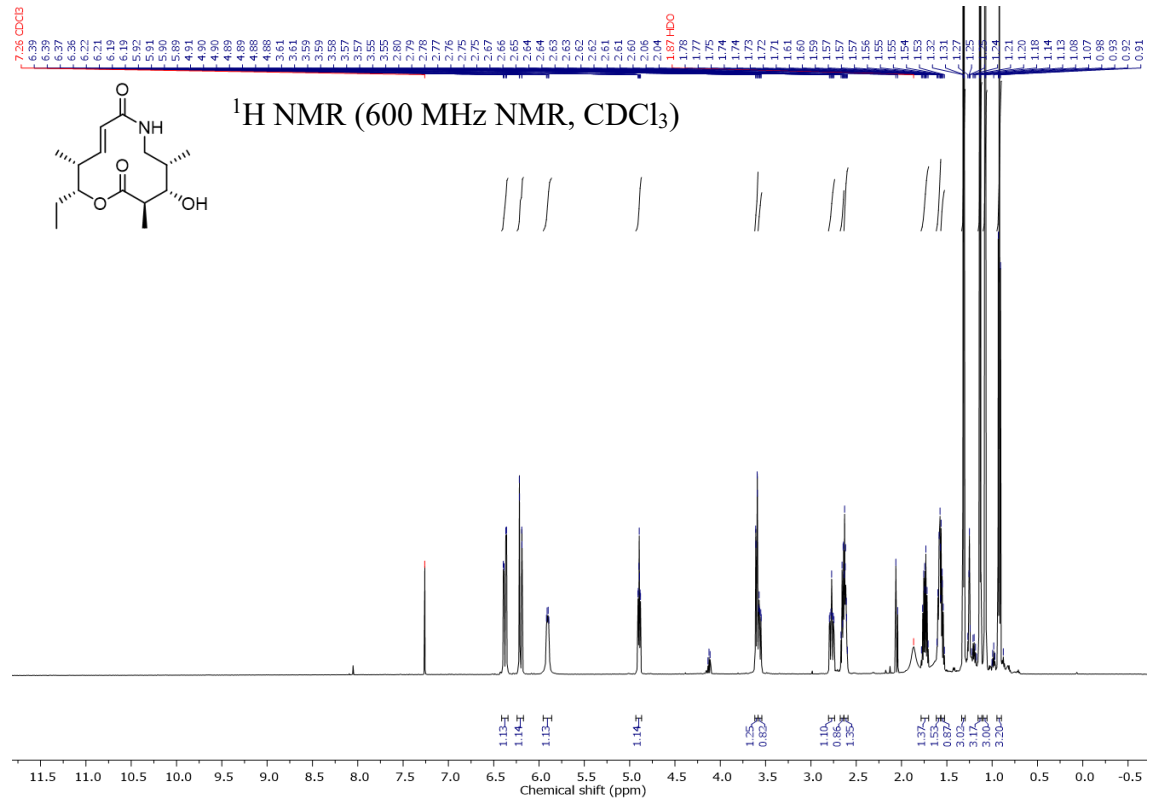


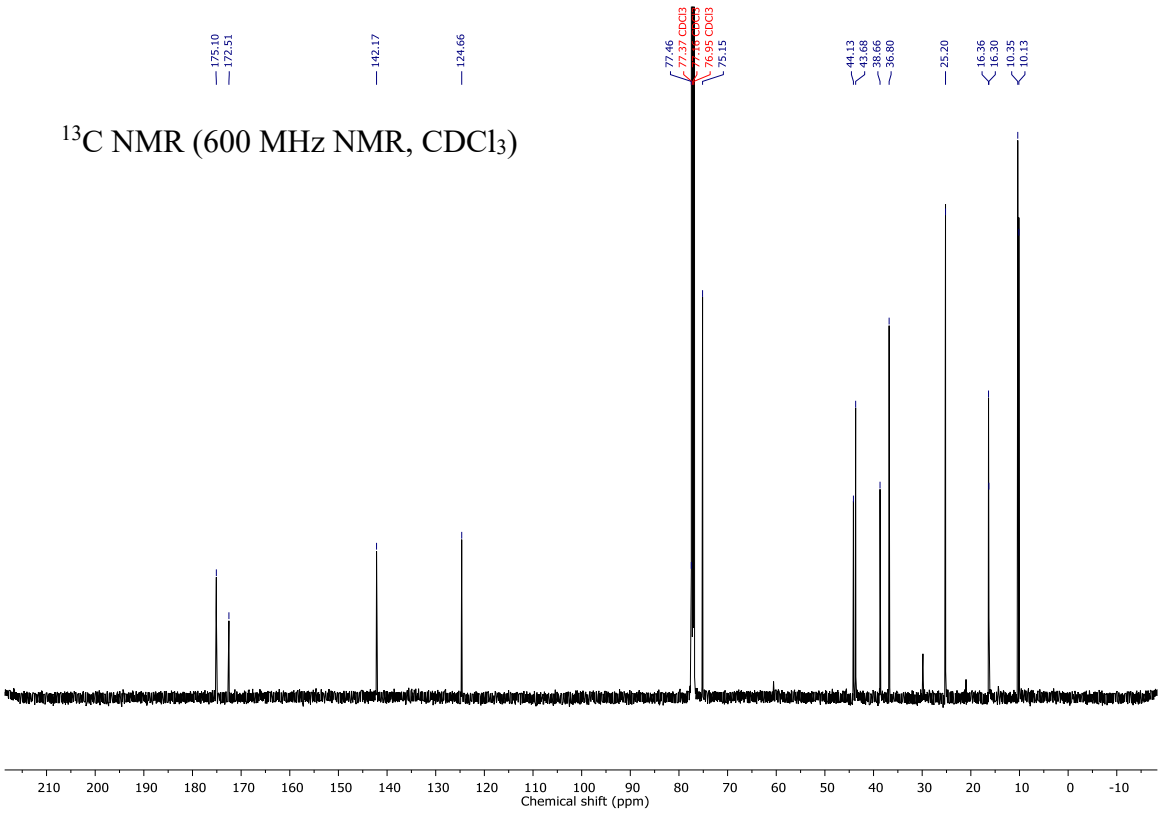


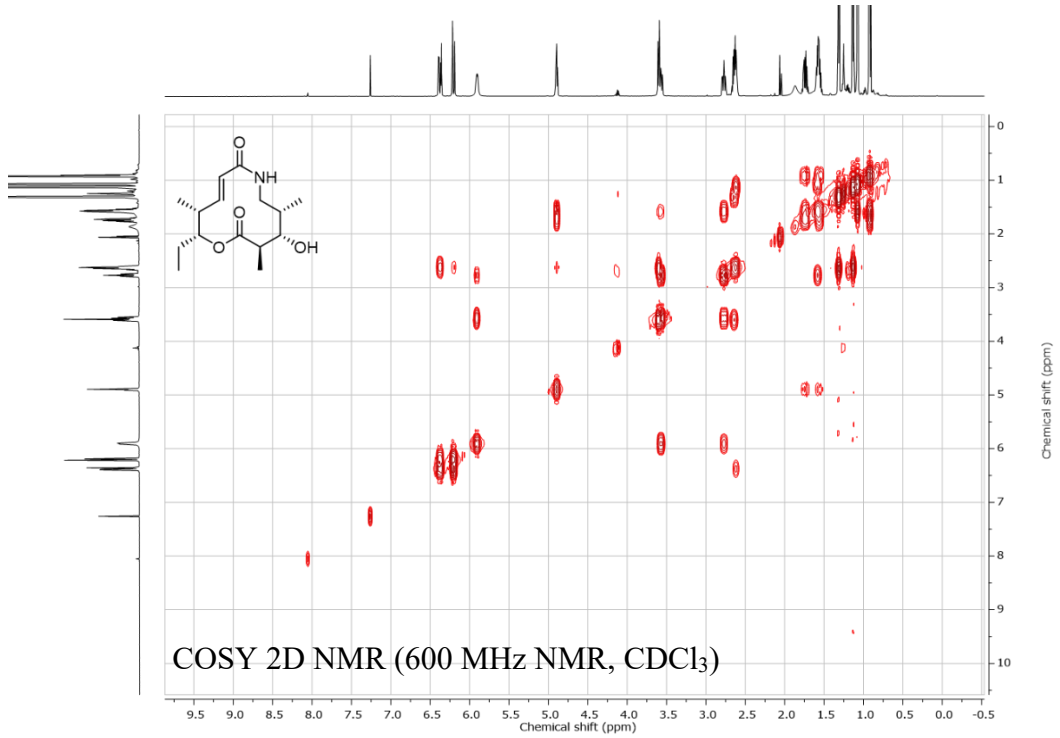
^{1}H NMR (600 MHz NMR, $CDCl_3$)



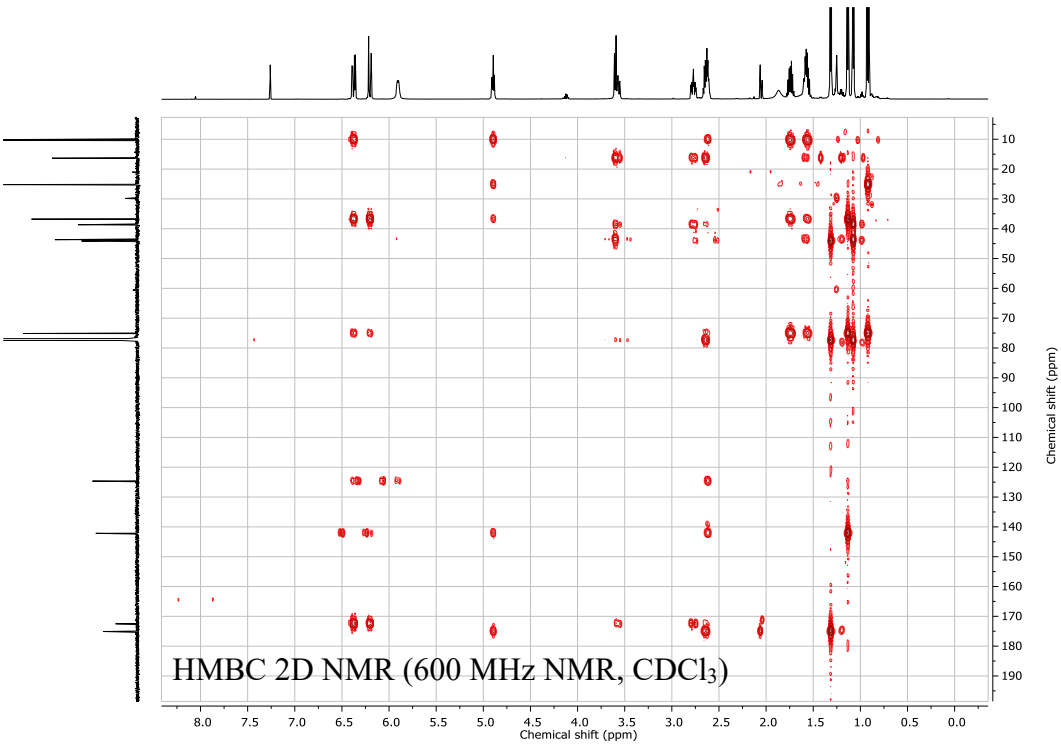
^{13}C NMR (600 MHz NMR, $CDCl_3$)


^{1}H NMR (600 MHz NMR, $CDCl_3$)
NH
O
O
O
OH
Chemical shift (ppm)

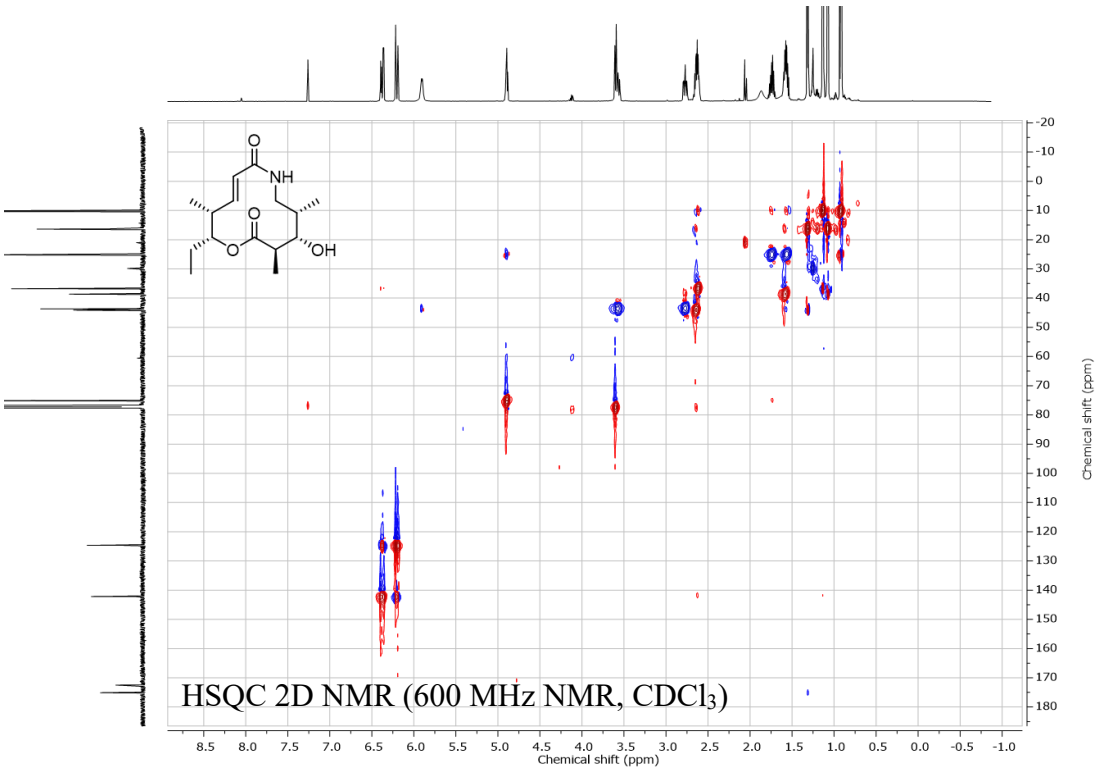

^{13}C NMR (600 MHz NMR, $CDCl_3$)
Chemical shift (ppm)



COSY 2D NMR (600 MHz NMR, $CDCl_3$)



HMBC 2D NMR (600 MHz NMR, $CDCl_3$)


NH
O
O
O
OH
HSQC 2D NMR (600 MHz NMR, CDCl3)
Chemical shift (ppm)
Chemical shift (ppm)

CHAPTER IV

Evaluation of PKS Modules PikAIII-TE and PikAIII/AIV for the Production of 12- and 14-membered Macrolactone Rings

IV.1 Abstract

The fusion protein PikAIII-TE is investigated in detail, answering the questions as to whether it can generate the desired amide/ester 12-membered macrolactones and what are the modular enzyme challenges. It was speculated that similar to the result obtained using Pik TE S148C which forms the hybrid 12-membered macrolactone/lactam ring, the enzyme should also be capable of cyclization in the module. However, formation of these rings was not identified in the module containing Pik TE S148C. Analytical scale enzymatic reactions indicated by mass for both amide and ester pentaketides, with modules PikAIII-TE WT and PikAIII-TE S148C show different masses for byproduct formation. These byproducts include hydrolyzed pentaketides, hydrolyzed hexaketides, SNAc pentaketides (resulting from transthioesterification of thiophenol pentaketides with SNAc following MeMal-SNAc extender unit consumption) and SNAc hexaketides (resulting from transthioesterification in the enzyme with SNAc). Notably, the SNAc hexaketides are the most predominant species formed after enzymatic reactions with the module PikAIII-TE S148C. Since in the reaction there are only trace amounts of the corresponding hydrolyzed amide and ester hexaketides, it suggests SNAc hexaketide formation is the main issue preventing cyclization from occurring. Additional experiments failed to eliminate SNAc from the enzymatic reactions to prevent the formation of these byproducts. Intact protein mass-spec suggests potential SNAc

binding or tight non-covalent interaction at the enzyme active site and more studies are underway. We also decided to utilize the native extender unit methylmalonyl-coenzyme A (MeMal-CoA) to avoid having free SNAc in solution. The hybrid 12-membered macrolactone/lactam ring appeared for the first time in trace amounts using MeMal-CoA with module PikAIII-TE S148C and matching our product standard retention time, however no product with PikAIII-TE WT. In terms of the ester pentaketide for both PikAIII-TE WT and PikAIII-TE S148C revealed only the hydrolyzed hexaketide and no evidence for the ester 12-membered macrolactone ring. Even though the amide macrocycle was generated in trace amounts, it still showed limitations worth studying further when using the modular enzyme PikAIII-TE, as the hydrolyzed hexaketides appear as major products from the reaction. Lastly, PikAIII-TE and PikAIV cloning and fusion with our new Pik TE variants from directed evolution showed interesting enzymatic reaction results for ring product formation at different ratios. In the future, additional efforts such cryo-EM and cryo-ET full module 3D structures could shed light into understanding how PKSs work together to assemble these complex macrocyclic structures.

IV.2 Introduction

In Chapter 2, we have demonstrated the ability to form 14-membered macrolactone rings utilizing a novel amide and ester pentaketide in conjunction with the PikAIII/AIV enzyme system. By contrast, the corresponding hexaketide hydrolysis product and other byproducts were formed exclusively when using the fusion protein PikAIII-TE WT and PikAIII-TE S148C mutant, raising questions about the enzyme selectivity and flexibility for macrocyclization. In this Chapter 4, we investigate PikAIII-TE and PikAIII-TE S148C to answer the question as to whether it can generate the desired amide/ester 12-membered macrolactones rings and what are the challenges when using the modular enzyme. In previous work, it was shown that Pik TE S148C worked as a standalone

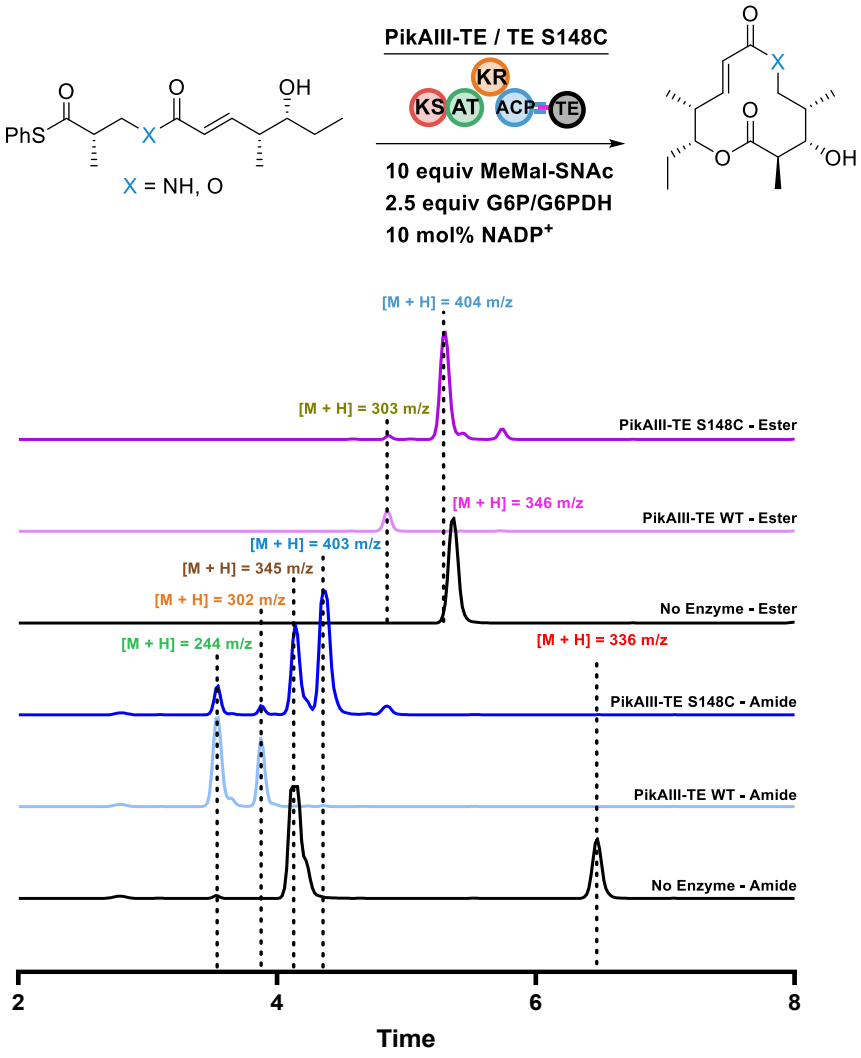
enzyme to generate a desired macrocycle to an outstanding 94% conversion from a C11 epimeric Pik hexaketide substrate (shown in Chapter 2). However, it was found that when Pik TE S148S was reincorporated into the fusion protein PikAIII-TE forming PikAIII-TE S148C, the ring formation lowered from a 94% to a 12%, with appearance of a C3-keto product also in a 12% yield (suggesting KR domain skipping or failure). These findings motivated a deeper examination of domain constraints in modular PKSs. Analogously, we utilized variants from Chapter 3 Pik TE directed evolution campaign, to ligate/fuse additional Pik TEs into PikAIII-TE and PikAIII/AIV modular system to understand the TE effects for not just cyclization, but the overall inter-domain interactions within the PKS modules that influence product formation.

IV.3 Analysis of Amide and Ester Pentaketides in the Modules PikAIII-TE WT and PikAIII-TE S148C

During the assessment of the analytical scale in 2 hours and overnight reactions of both amide and ester pentaketides with module PikAIII-TE WT and PikAIII-TE S148C different masses for product formation were identified except for the desired 12-membered macrolactone rings (**Figure IV.1**). For PikAIII-TE WT the masses obtained corresponded to amide: hydrolyzed pentaketide [M+H] = 244 m/z, hydrolyzed hexaketide [M+H] = 302 m/z, for ester: hydrolyzed hexaketide [M+H] = 303 m/z. Now, for PikAIII-TE S148C the masses are amide: hydrolyzed pentaketide [M+H] = 244 m/z, hydrolyzed hexaketide [M+H] = 302 m/z, transthioesterification of thiophenol pentaketide with SNAc from MeMal-SNAc generated SNAc pentaketide [M+H] = 345 m/z, transthioesterification of enzyme with SNAc produced SNAc hexaketide [M+H] = 403 m/z, in the case of ester: only SNAc hexaketide [M+H] = 404 m/z (**Figure IV.1**). Interestingly, the hydrolyzed hexaketides are mostly found in the PikAIII-TE WT reaction, suggesting that the Pik TE WT is hydrolyzing the intermediate instead of performing cyclization. This is in accordance with the

results when using the synthesized amide hexaketide in the Pik TE WT enzymatic reaction which does not generate the desired 12-membered ring, instead it only performs hydrolysis of the hexaketide from the serine in the Pik TE WT active site.

Notably, the SNAc hexaketides are the most predominant species formed after enzymatic reactions of amide and ester pentaketides with the PikAIII-TE S148C (**Figure IV.1**). It is suspected that similar to the result using Pik TE S148C which forms the hybrid 12-membered macrolactone/lactam ring, the Pik TE S148C should also be capable of cyclization in the module. However, formation of the desired amide and ester macrocyclic rings are not identified. Since in the reaction there are only traces amounts of the corresponding hydrolyzed amide and ester hexaketides, it suggests that the SNAc hexaketides formation are the main issue preventing cyclization from occurring. Another challenge present is the transthioesterification that generates the amide SNAc pentaketide, which consumes thiophenol starting material. The order of addition of the substrate was changed to last in the reaction, which still failed to reduce the amide SNAc pentaketide product suggesting the SNAc transthioesterification is faster and preferred when starting material is present. The ester SNAc pentaketide product was not identified, indicating a substrate preference during the modular PikAIII-TE WT and PikAIII-TE S148C enzymatic reaction. To prevent the production of the undesired SNAc hexaketides, 20 mM 2-vinylpyridine as a thiol scavenger[62] was added resulting in a slight reduction of SNAc pentaketides and SNAc hexaketides failing to drastically reduce or eliminate these undesired products.


PikAIII-TE / TE S148C
KR
KS AT ACP TE
10 equiv MeMal-SNAc
2.5 equiv G6P/G6PDH
10 mol% NADP+
PhS
X = NH, O
OH
[M + H] = 404 m/z
[M + H] = 303 m/z
PikAIII-TE S148C - Ester
[M + H] = 346 m/z
PikAIII-TE WT - Ester
[M + H] = 403 m/z
[M + H] = 345 m/z
[M + H] = 302 m/z
No Enzyme - Ester
[M + H] = 244 m/z
[M + H] = 336 m/z
PikAIII-TE S148C - Amide
PikAIII-TE WT - Amide
No Enzyme - Amide
2
4
6
8
Time

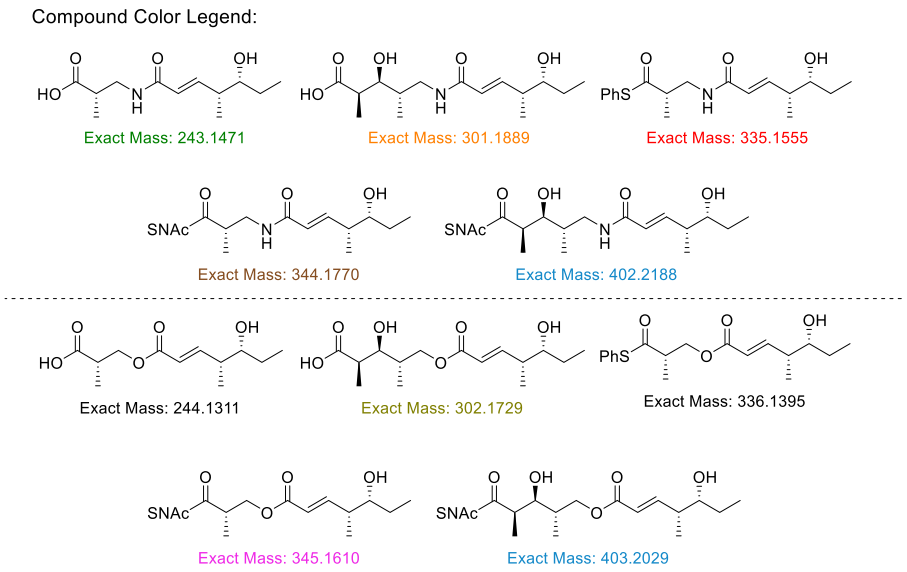


Figure IV.1 Extracted ion chromatogram (EIC) resulting from enzymatic reaction of amide and ester pentaketides with PikAIII-TE and PikAIII-TE S148C. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. See also compound color scheme/compound legend accompanying the EIC.

IV.4 Where are these SNAc Byproducts Being Off-Loaded From? The ACP or the TE?

In order to investigate where the amide and ester SNAc hexaketides are being off-loaded in the PikAIII-TE WT and PikAIII-TE S148C modules, we designed the PikAIII-TE S148A inactive TE mutant. The reaction with the amide pentaketide generated mainly amide SNAc pentaketide [M+H] = 345 m/z and to our surprise, two peaks corresponding to the C2 epimerized species for the C3-keto amide SNAc hexaketide [M+H] = 401 m/z. The resulting species suggest the unsuccessful reduction by the KR in the C3 position of the growing polyketide chain. This process leads to epimerization at C2 position-methyl group due to the acidity of the α-proton next to the C1 carbonyl, resulting in exchange of the proton in water causing modification of the configuration at the C2 chiral center (**Figure IV.2**). For the ester pentaketide, the corresponding ester SNAc hexaketide was found [M+H] = 404 m/z and also the species corresponding to the C3-keto ester

SNAc hexaketide [M+H] = 402 m/z, which the epimerization products are likely unresolved in the single peak (**Figure IV.2**). The results suggests the amide and ester SNAc hexaketide byproducts are likely off-loaded from the ACP and not the TE (**Figure IV.3**). An interesting finding is the S148A change in the TE, thus inactivating the enzyme, seems to cause the KR to be skipped during the reaction. Examples of these have been found in the literature where the KR is bypassed in PKS modular reactions when the TE is engineered or swapped between systems which is a finding worth investigating further.[65, 66, 210]

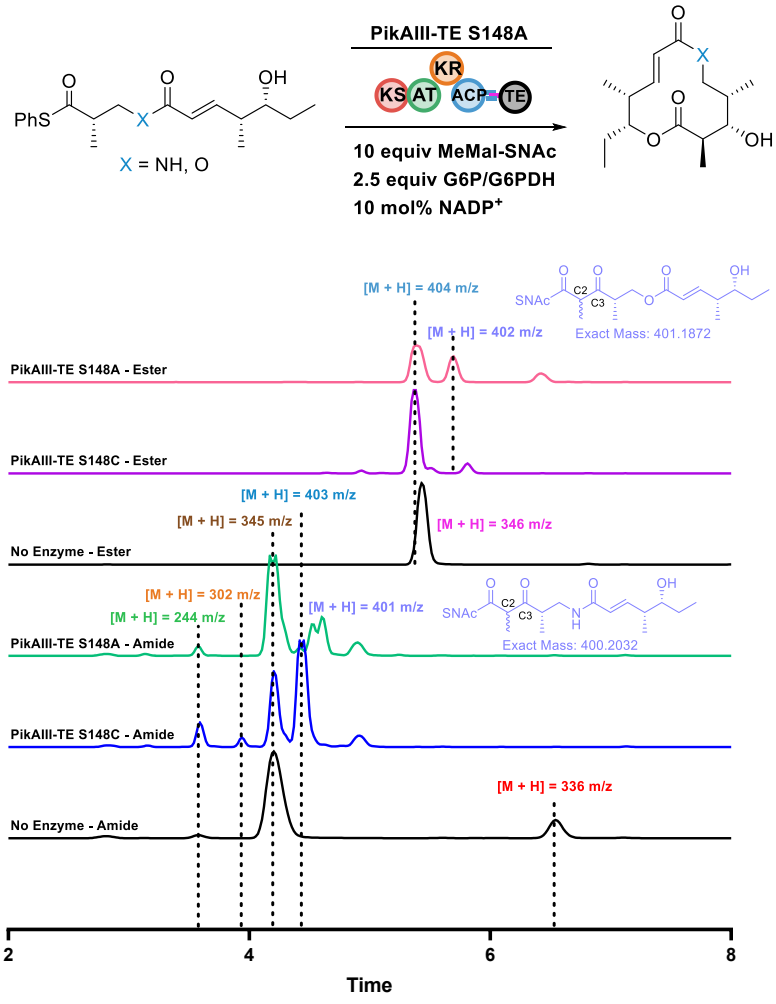


Figure IV.2 Extracted ion chromatogram (EIC) resulting from enzymatic reaction of amide and ester pentaketides with PikAIII-TE S148A (TE inactive). TOF-MS: 20 – 100% acetonitrile/water

with 0.1% formic acid, 13 minutes (min) total run. See also compound color scheme/compound legend accompanying the EIC in Figure **IV.1**.

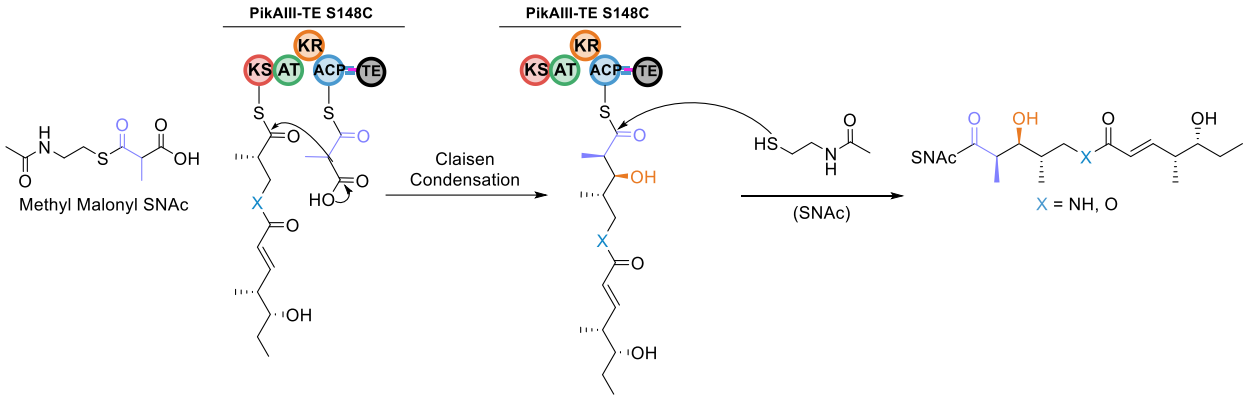


Figure IV.3 Depiction of the SNAc byproducts offloaded from the ACP via the SNAc nucleophilic attack into thioester-ACP bond.

IV.5 PikAIII-TE WT and PikAIII-TE S148C Enzymatic Reactions using MeMal-CoA as the Extender Unit

Having failed to generate the amide and ester 12-membered macrolactone rings using MM-SNAc extender unit, we decided to utilize the native extender unit methylmalonyl-coenzyme A (MeMal-CoA) to avoid having free SNAc in solution. Analytical scale reactions indicated by mass showed the enzymatic reaction with PikAIII-TE WT using the amide pentaketide resulted mainly on the formation of the hydrolyzed pentaketide [M+H] = 244 m/z and hydrolyzed hexaketide [M+H] = 302 m/z products. Now, when utilizing PikAIII-TE S148C the reaction also gave hydrolyzed pentaketide [M+H] = 244 m/z and hydrolyzed hexaketide [M+H] = 302 m/z, however a peak corresponding to the hybrid 12-membered macrolactone/lactam ring [M+H] = 284 m/z, appeared for the first time in trace amounts using the module PikAIII-TE S148C and matching our product standard retention time (**Figure IV.4**). In terms of the ester pentaketide for both PikAIII-TE WT and PikAIII-TE S148C showed only the hydrolyzed hexaketide [M+H] = 303 m/z and no

evidence for the ester 12-membered macrolactone ring. MeMal-CoA only appears to generate the desired 12-membered macrolactone ring for the amide product and not the ester. Even though the hybrid 12-membered macrolactone/lactam ring was generated in trace amounts, it still showed limitations when using the modular enzyme PikAIII-TE, as the hydrolyzed hexaketide products appear as major products from the reaction. This suggests that the Pik TE WT and Pik TE S148C even when utilizing MeMal-CoA, preferably generates the hydrolysis product instead of the cyclized macrolactone, suggesting once more that the Pik TE is a catalytic hurdle for the enzymatic reaction (**Figure IV.4**).

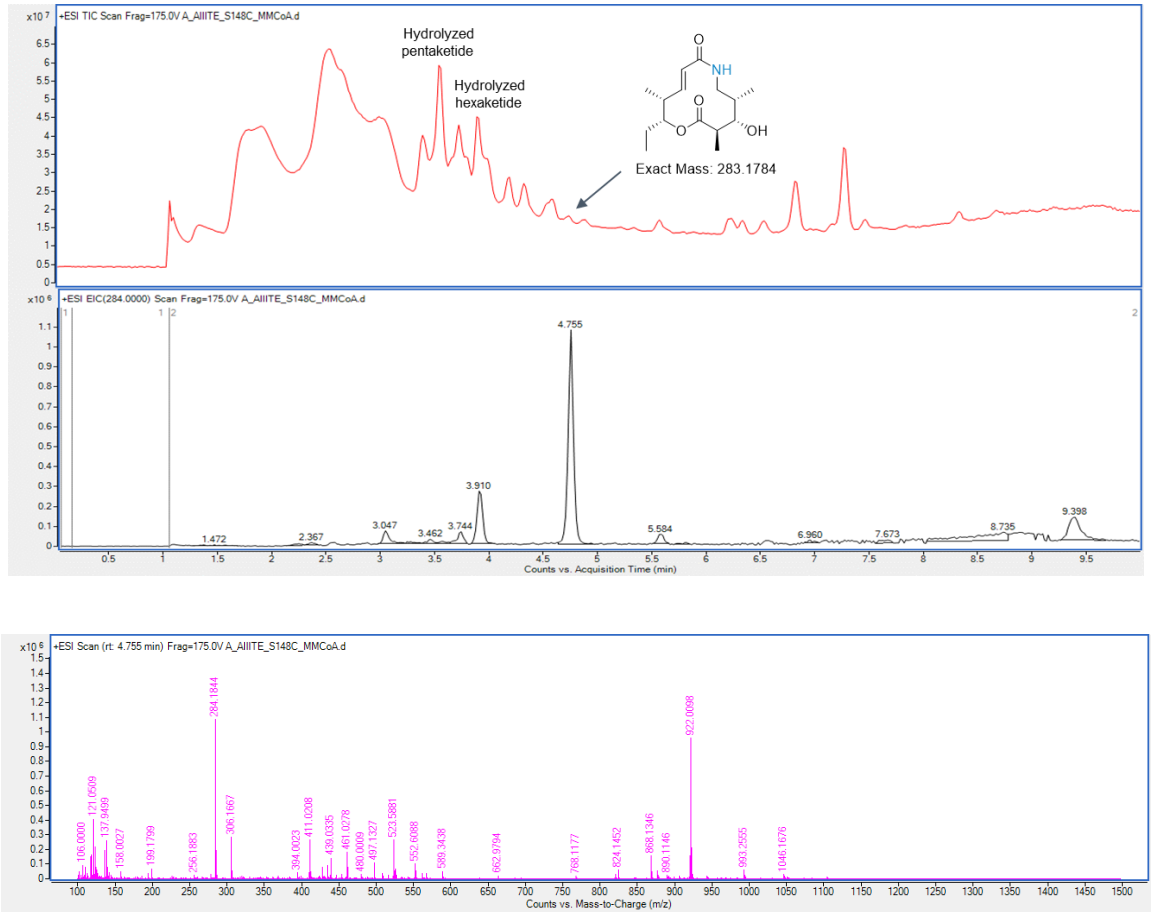


Figure IV.4 Total ion chromatogram (TIC) resulting from enzymatic reaction of amide pentaketide with PikAIII-TE S148C using MeMal-CoA as the extender unit. TOF-MS: 20 – 100%

acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. See also peak for electrospray ionization source (ESI) scan corresponding to the desired hybrid 12-membered macrolactone/lactam ring at the 4.755 min retention time.

IV.6 SNAc Transthioesterification Reactions to Test Substrate Activating Group Reactivity Differences Against Pik TE

The results obtained from the modular reactions using PikAIII-TE WT/S148C inspired the investigation of the substrate activating group reactivity difference against the standalone Pik TE. From previous studies, it was found that Pik TE prefers SNAc containing substrates, rather than thiophenolated ones.[112] We decided to utilize the readily available synthesized thiophenol containing amide hexaketide, to perform a transthioesterification reaction using SNAc, to then utilize it for a TE mediated reaction. The goal was to identify if the standalone TE could take on the SNAc containing hexaketide, and cyclize it to generate the desired hybrid 12-membered macrolactone/lactam ring (**Figure IV.5**). For the reaction, we first started by incubation of the substrate with 10 mM SNAc and subsequent addition of the TE. First, we tried Pik TE WT which resulted in formation of the hydrolyzed hexaketide [M+H] = 302 m/z, and no ring formation, as expected form the thiophenol reactions done previously. To our surprise, when using Pik TE S148C, only the thiophenol substrate [M+H] = 394 m/z was identified, meaning no SNAc transthioesterification occurred. We also found in the total ion chromatogram, the presence of the dimeric form of SNAc suggesting that it needed to be reduced in order to be a monomer, and possible reason why the SNAc transthioesterification reaction was unsuccessful. Next, 1 mM of TCEP was added as a reducing agent to prevent SNAc from being in its oxidized form. The TE reactions were performed analogously, where Pik TE WT again resulted in the formation of the hydrolyzed hexaketide [M+H] = 302 m/z, and no ring formation. For Pik TE S148C, only the

SNAc amide hexaketide [M+H] = 403 m/z was found, no hydrolysis or cyclization, potentially showcasing that the SNAc activating group is not tolerated as well as the thiophenol one. However, since the Pik TE WT appeared to not discriminate between thiophenol or SNAc amide hexaketide, we hypothesized that the key difference between the two enzymes, one being Ser148 and the other Cys148 mutation, that perhaps SNAc is generating a disulfide bond with the catalytic cysteine, avoiding cyclization from occurring (**Figure IV.5**).

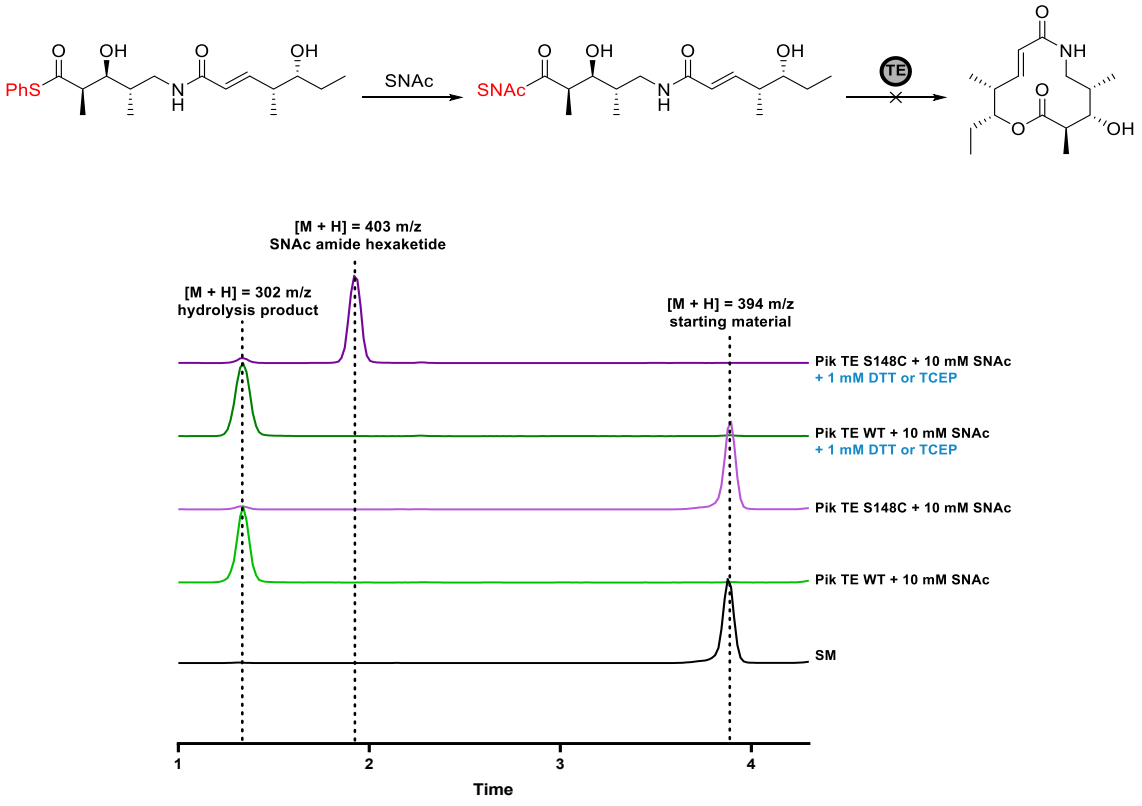


Figure IV.5 Extracted ion chromatogram (EIC) resulting from thiophenol amide hexaketide transthioesterification reaction with SNAc, with subsequent TE mediated cyclization. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run.

IV.7 Mass Spectrometry Analysis of SNAc in the Pik TE S148C Active Site

To investigate the hypothesis of SNAc generating a disulfide bond with Pik TE S148C active site cysteine, we decided to perform a LC-MS analysis using intact protein mass spectrometry. We first evaluated the Pik TE S148C sequence, where we found 3 possible cysteines for disulfide bond formation (including Cys148 in the active site). We then compared the deconvolution results obtained for the Pik TE S148C alone control which was 33,631.54 Da to the Pik TE S148C + 10 mM SNAc which was 33,983.58 Da **(Figure IV.6**). The mass difference between the two is + 352.04 Da, which suggests potentially the possible disulfide bond formation or a tight non-covalent interaction of SNAc with all 3 cysteines in the proteins (352.04 ÷ 3 = 117.34 Da), with ~0.8 Da mass error expected for the molecular weight of SNAc (118.17 Da) and the LC-MS error. As a control, we used Pik TE WT, revealing 33,615.70 Da and 33,850.43 Da after adding 10 mM SNAc. Again, this mass difference suggests the two cysteines in Pik TE WT being compromised by SNAc (**Figure IV.7**).

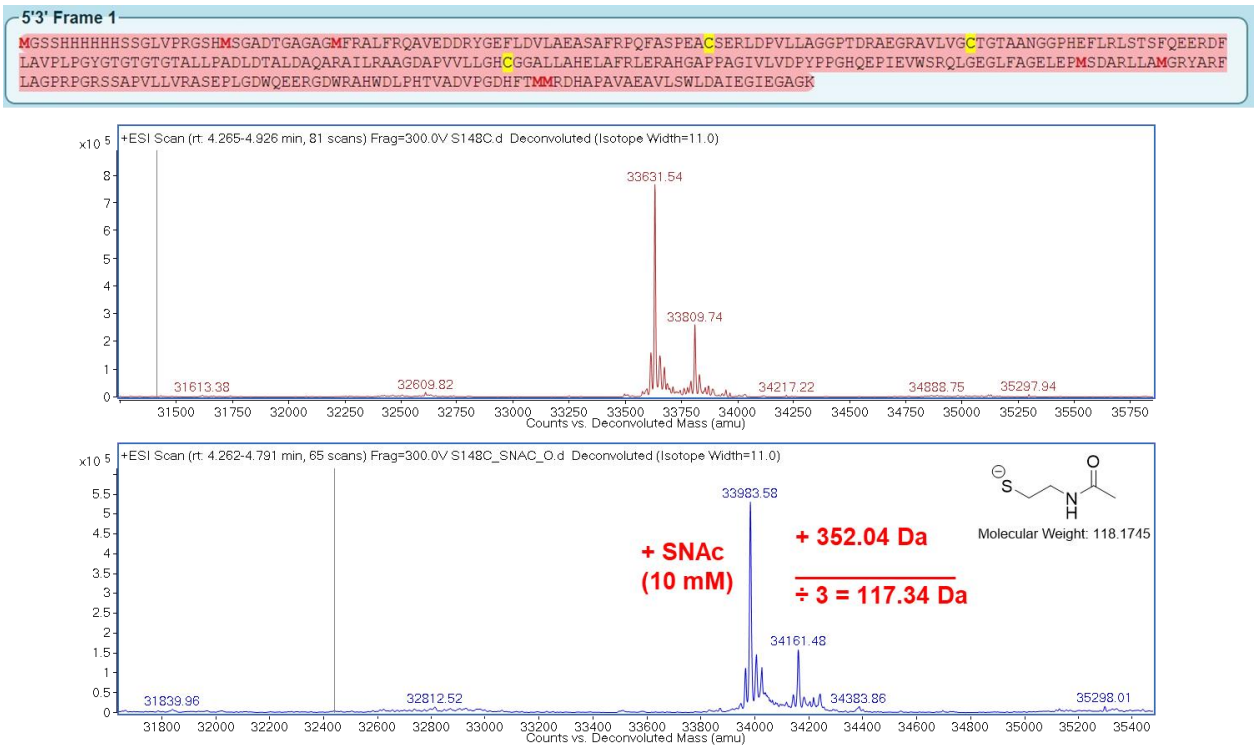


Figure IV.6 Electrospray ionization source (ESI) scan corresponding to Pik TE S148C and Pik TE S148S with SNAc added.

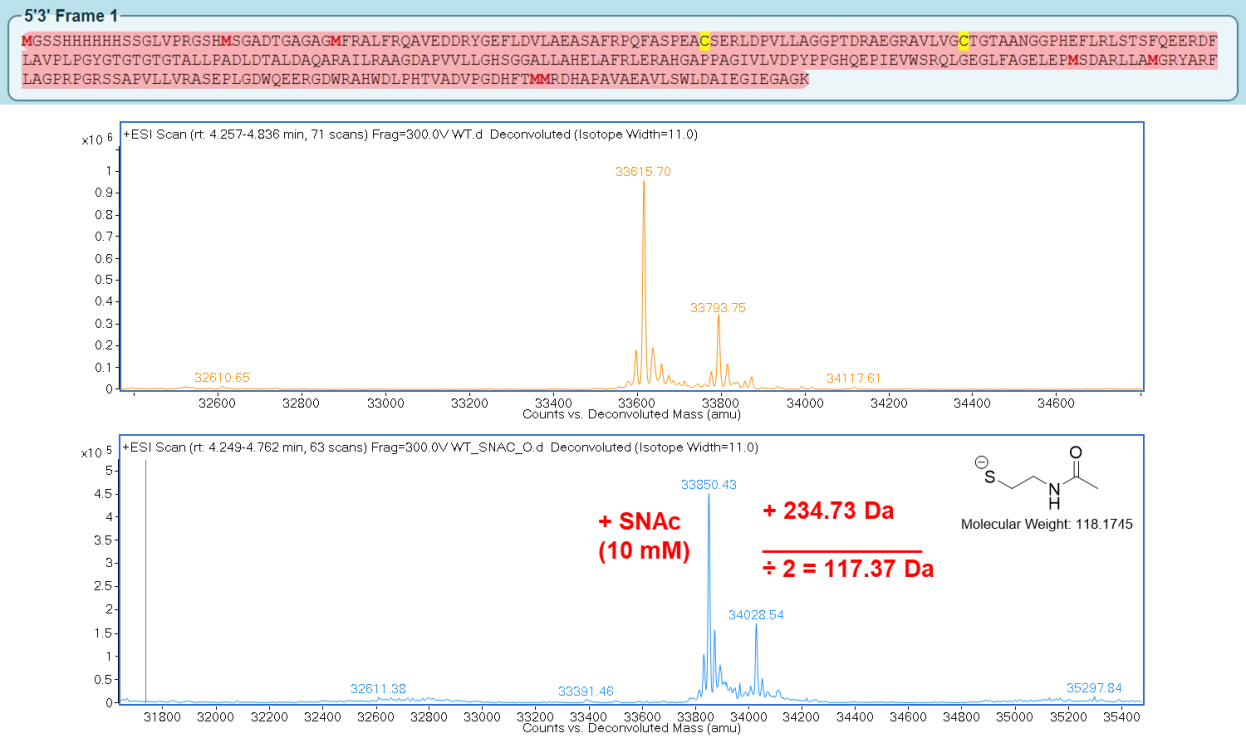


Figure IV.7 Electrospray ionization source (ESI) scan corresponding to Pik TE WT and Pik TE WT with SNAc added.

As a next step, we utilized DTT as a reducing agent, to avoid having SNAc potentially being covalently bound to the protein. To our advantage, we were able to show that adding 1 mM DTT efficiently removed SNAc from the two cysteines in Pik TE WT, as seen on the intact protein mass results, where the mass is kept intact to 33,615.71 Da (**Figure IV.8**).

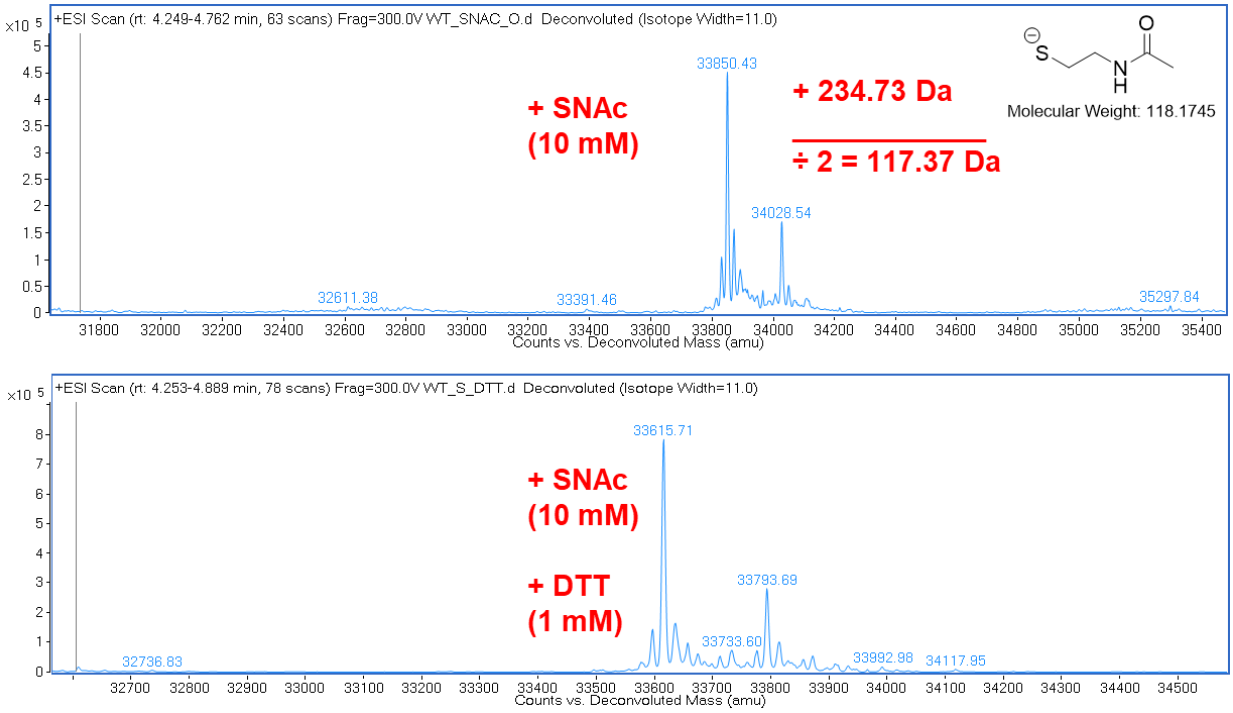


Figure IV.8 Electrospray ionization source (ESI) scan corresponding to Pik TE WT with SNAc added and Pik TE WT SNAc/DTT added.

Similarly, we tested Pik TE S148C with 1 mM DTT and the results again showed reversion to Pik TE S148C mass to 33,631.50 Da, and appearance of a peak with 33,747.43 Da (**Figure IV.9**). We reasoned this could still be 1 cysteine being compromised with SNAc, however the mass difference corresponds to ~2 Da, which is not under the mass difference error of the LC-MS. This could perhaps represent an ionization of the protein that we might not be aware of, or a deprotonation event. We also obtained similar results when changing the reducing agent from DTT to TCEP. In addition, we added any of the aforementioned reducing agents to an enzymatic reaction (PikAIII-TE WT/S148C and Pik TE WT/S148C) where the desired hybrid 12-membered macrolactone/lactam ring is not formed, and it showed no formation of the desired macrocycle.

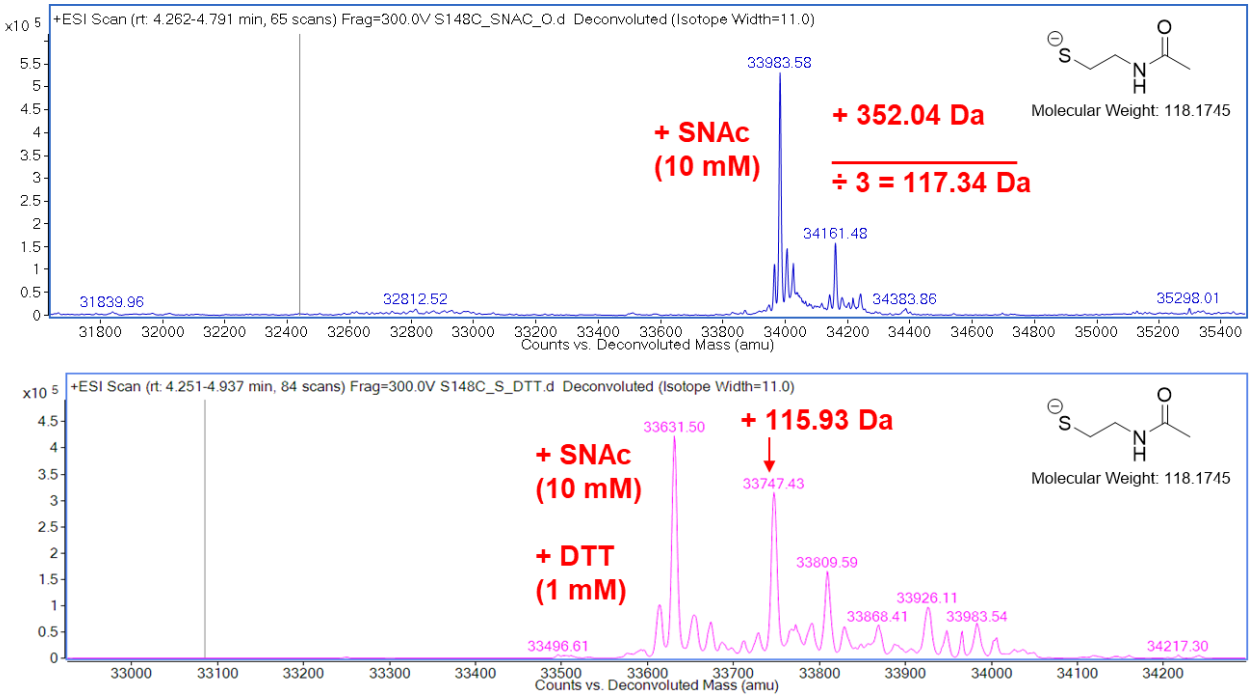


Figure IV.9 Electrospray ionization source (ESI) scan corresponding to Pik TE S148C with SNAc added and Pik TE S148C SNAc/DTT added.

To further demonstrate the potential detrimental effects of SNAc in the standalone Pik TE S148C, we performed a titration starting with 0.1 mM SNAc to 10.0 mM SNAc, as those concentrations are to be expected as byproducts from the PKS reactions and resulting SNAc in solution (**Figure IV.10**). The results revealed at the onset of the titration with 0.1 mM SNAc, the Pik TE S148C molecular weight is intact. Now 0.5 mM-5.0 mM SNAc resulted in gradual disappearance of the intact protein mass and a shift to the addition of ~116 Da, similar to the results obtained when adding DTT or TCEP as reducing agent. Lastly, 7.0 mM-10.0 mM SNAc resulted in appearance of the main molecular weight addition of ~352 Da, corresponding to the previous observed mass added when having all cysteines with SNAc potentially covalently bound. Overall, these results suggest that the addition of ~116 Da that we were not able to eliminate even by adding a reducing agent might correspond to the catalytic Cys148 disulfide bonded with SNAc, and might

be the reason why the reactions of SNAc amide hexaketide and Pik TE S148C were unsuccessful. Despite the fact that these reactions only showed SNAc starting material, we noticed that removal of SNAc from the substrate was highly difficult, as LC-MS results showed appearance of free SNAc from the transthioesterification reactions. In the future, a similar titration using Pik TE WT might shed light on the hypothesis of SNAc covalently bound to cysteines in these proteins. In addition, synthesis of the SNAc amide hexaketide rather than transthioesterification of the available thiophenol amide hexaketide might result in a cleaner sample free of remaining SNAc, as well as potentially using MeMal-thiophenol or other activating groups for the extender unit.

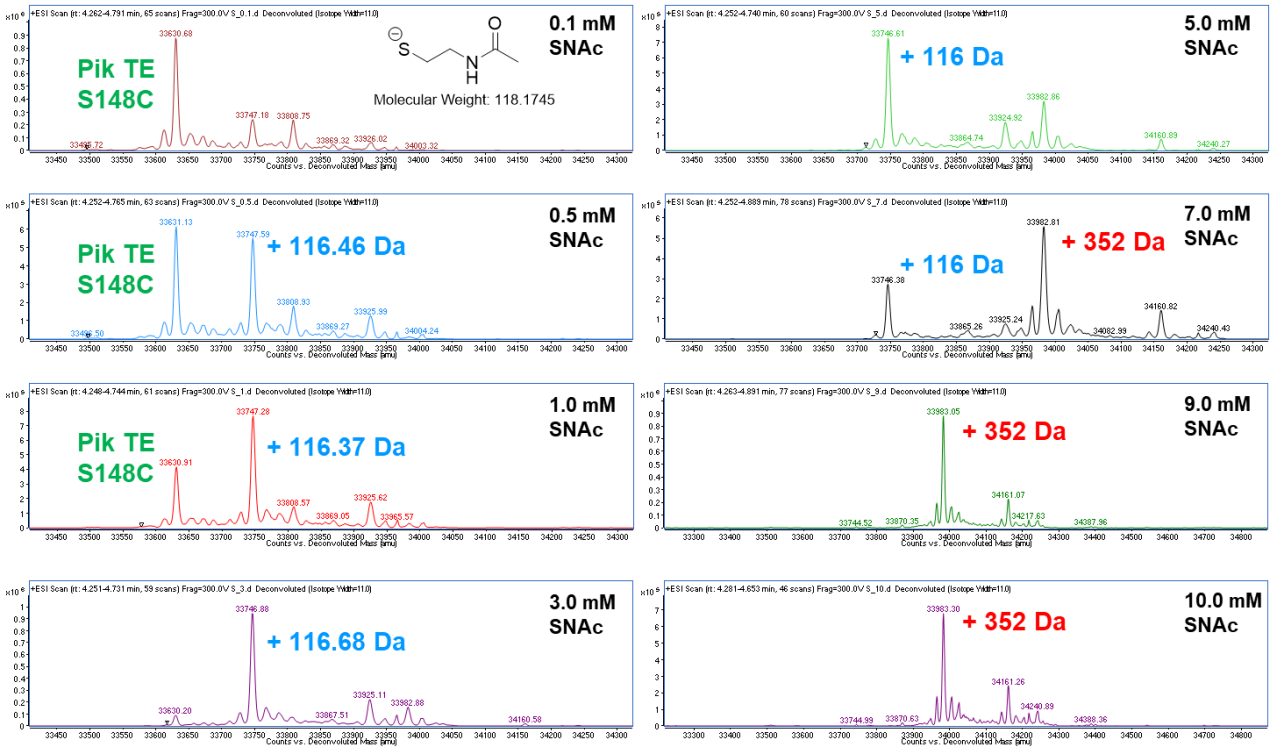


Figure IV.10 SNAc titration results: electrospray ionization source (ESI) scan corresponding to Pik TE S148C with SNAc at different concentrations ranging from 0.1 – 10.0 mM SNAc.

From our studies using intact protein mass spectrometry and the amide hexaketide, we decided to test our hypothesis with a different substrate. We tested the previously generated thiophenol

methoxy protected Pik native hexaketide against Pik TE S148C and WT (**Figure IV.11**).[62, 65] The enzymatic reaction with Pik TE S184C resulted in the expected formation of the 12-membered ring mass [M+H] = 311 m/z, with some starting material remaining [M+H] = 421 m/z. Adding 10 mM of SNAc to the reaction completely abolished the formation of the ring at [M+H] = 311 m/z, where only starting material was found, meaning the enzyme was deactivated by SNAc. To our surprise, adding 1 mM DTT to the reaction then generated the desired macrocycle at [M+H] = 311 m/z and the formation of the SNAc Pik hexaketide due to transthioesterification of the activating groups. The macrolactone formation is hence decreased by the formation of the SNAc byproduct. Previous data has shown that the SNAc activated Pik hexaketide[62, 65] reaction against Pik TE S148C results in 94% conversion, meaning that the amount of SNAc byproduct resulting from the reaction appears to not impact greatly the macrocyclization. However, considering higher amounts of SNAc byproducts derived from MeMal-SNAc does seem to impact cyclization of pentaketides when testing these as substrates for PikAIII-TE or PikAIII/AIV enzyme systems.

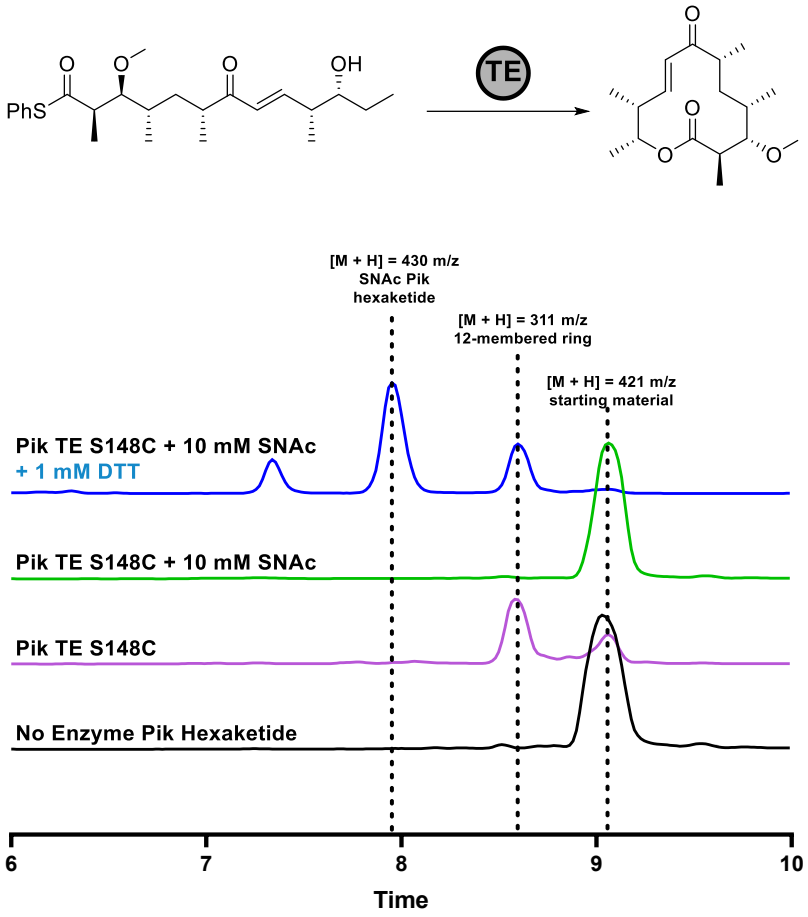


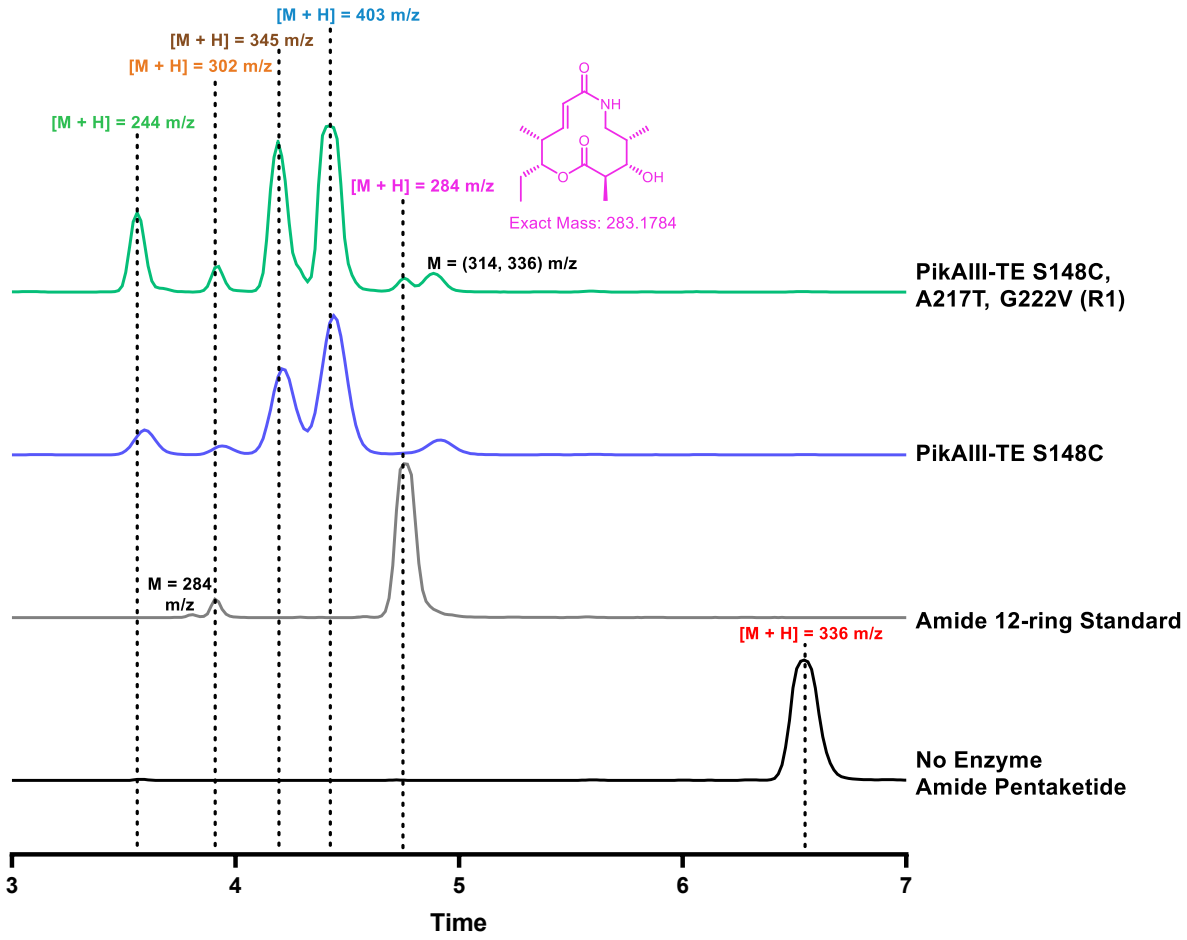
Figure IV.11 Extracted ion chromatogram (EIC) resulting from thiophenol methoxy protected Pik native hexaketide against Pik TE S148C. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run.

IV.8 Ligations of Pik TE Mutants from Directed Evolution into PKS Modules

IV.8.1 Enzymatic Reactions of Amide Pentaketide Against PikAIII-TE S148C, A217T, G222V (R1) Mutant and PikAIII-TE S148C, A217T, G222V, M271V, Y25C, L126V (R2)

After identification of the standalone Pik TE R1 mutant as a more efficient and selective macrocyclization catalyst. We decided to perform a ligation of this standalone domain into the full fusion enzyme PikAIII-TE to generate PikAIII-TE S148C, A217T, G222V (R1). We performed

this ligation by cloning the digested TE R1 gene (with XhoI, HindIII restriction enzymes) into the empty vector containing PikAIII via T4 ligase reaction overnight at 18 ºC.[66] Upon confirmation of the desired fusion protein via DNA sequencing, the protein was expressed into *E. coli* and purified. The resulting pure enzyme was used in analytical reactions against the amide pentaketide and analyzed in LC-MS traces (**Figure IV.12**). The data showed similarities with PikAIII-TE S148C, as the main products from the reaction corresponded to hydrolyzed pentaketide [M+H] = 244 m/z, hydrolyzed hexaketide [M+H] = 302 m/z, SNAc pentaketide [M+H] = 345 m/z and SNAc hexaketide [M+H] = 403 m/z. To our surprise, we identified a new peak corresponding to the expected mass of the 12-membered macrolactone/lactam ring [M+H] = 284 m/z. Even though the results suggest a lower yield than desired, we were able to identify the formation of the ring for the very first time using the PikAIII-TE fusion protein in the presence of MeMal-SNAc (**Figure IV.12**). Subsequently, we tried the Pik TE R2 mutant, and generated PikAIII-TE S148C, A217T, G222V, M271V, Y25C, L126V (R2). The results showed the exact same LC-MS trace as shown for PikAIII-TE R1, meaning that more mutations in the TE did not have an effect on the ring product formation. As for ester pentaketide reactions, these were not as successful as the amide pentaketide ones, with no ring formation observed (**Table IV.1**). We still need to investigate what are the constraints when using the fusion module and others and how these individual domains work together to extend and elongate the polyketide substrates to generate the desired macrolactone rings.



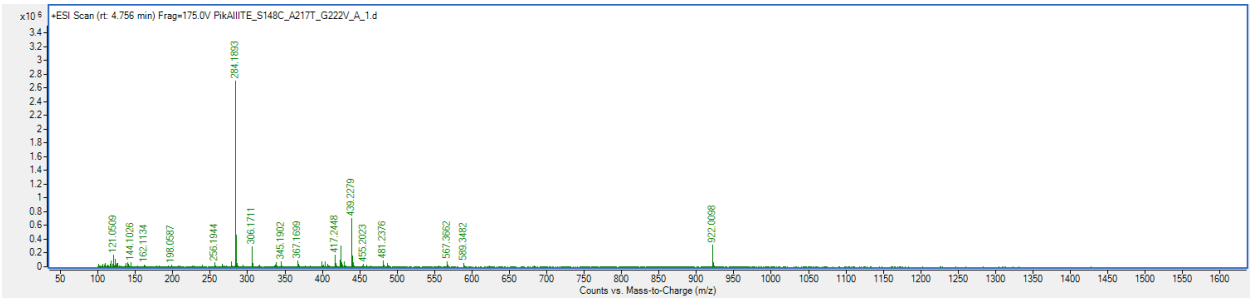


Figure IV.12 Extracted ion chromatogram (EIC) resulting from enzymatic reaction of amide pentaketide with PikAIII-TEs. TOF-MS: 20 – 100% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run. See also peak for electrospray ionization source (ESI) scan corresponding to the desired hybrid 12-membered macrolactone/lactam ring at the 4.756 min retention time.

Table IV.1 Amide/ester pentaketide enzymatic reaction results for PikAIII-TE modules with engineered Pik TEs for the formation of 12-membered ring product. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data).

entry	substrate	Pik modules	time	product (12 ring)
1	Amide Pentaketide	AIII-TE WT or S148C	overnight	-
2	Amide Pentaketide	AIII-TE S148C, A217T, G222V (**R1**)	overnight	12 ring trace
3	Amide Pentaketide	AIII-TE S148C, A217T, G222V, M271V, Y25C, L126V (**R2**)	overnight	12 ring trace
4	Ester Pentaketide	AIII-TE WT or S148C	overnight	-
5	Ester Pentaketide	AIII-TE S148C, A217T, G222V (**R1**)	overnight	-
6	Ester Pentaketide	AIII-TE S148C, A217T, G222V, M271V, Y25C, L126V (**R2**)	overnight	-

IV.8.2 Enzymatic Reactions of Amide/Ester Pentaketide Against PikAIII/AIV-TE S148C, A217T, G222V (R1) Mutant and PikAIII/AIV-TE S148C, A217T, G222V, M271V, Y25C, L126V (R2)

We then followed a similar strategy as described for PikAIII-TE, with the fusion of the engineered TE variants.[66] We ligated Pik TE R1 into PikAIV, generating PikAIV-TE S148C, A217T, G222V (R1), as well as incorporation of Pik TE R2, to produce PikAIV-TE S148C, A217T, G222V, M271V, Y25C, L126V (R2). First, the resulting purified modules were used in analytical reactions against the amide pentaketide and analyzed in LC-MS traces. The results highlighted on **Table IV.2** show the comparison with AIV-TE WT module, time of the reaction and the ratio of 6:14 membered ring products. To our surprise, when using PikAIII/AIV-TE S148C, the ratio of 6:14 rings remained 1:1 throughout the course of the reaction, suggesting that the TE S148C mutation blocks the generation of additional 14-ring opening to the 6-ring when compared to the TE S148C WT in the AIV module, where contrarily overnight reactions show

only 6-ring production. When utilizing PikAIII/AIV-TE R1, the results overtime only show the generation of 6-membered ring, suggesting a higher selectivity for the Pik TE R1 for the smaller ring sized product. Then, for PikAIII/AIV-TE R2, it predominantly generated the 6-ring, however, it overtime shifts the reaction thermodynamics to the formation of 14-ring via ring opening/closure.

Table IV.2 Amide pentaketide enzymatic reaction results for PikAIII/AIV-TE modules with engineered Pik TEs for the formation of 6- and 14-membered ring products. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data).

entry	substrate	Pik modules	time	products (6:14 ring)
1	Amide Pentaketide	AIII / AIV TE WT	30 mins	1:3
2	Amide Pentaketide	AIII / AIV TE S148C	30 mins	1:1
3	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V **(R1)**	30 mins	6 ring:trace
4	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V **(R2)**	30 mins	6:1
5	Amide Pentaketide	AIII / AIV TE WT	1h	1:2
6	Amide Pentaketide	AIII / AIV TE S148C	1h	1:1
7	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V **(R1)**	1h	6 ring:trace
8	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V **(R2)**	1h	6:1
9	Amide Pentaketide	AIII / AIV TE WT	2h	1:1
10	Amide Pentaketide	AIII / AIV TE S148C	2h	1:1
11	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V **(R1)**	2h	6 ring:trace
12	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V **(R2)**	2h	5:1
13	Amide Pentaketide	AIII / AIV TE WT	overnight	6 ring
14	Amide Pentaketide	AIII / AIV TE S148C	overnight	1:1
15	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V **(R1)**	overnight	6 ring:trace
16	Amide Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V **(R2)**	overnight	3:1

Second, we tested the ester pentaketide against the newly generated PikAIV-TEs. Again, these reactions followed the same analysis as described above for the amide pentaketide substrate. The results revealed in **Table IV.3** are much more different than those obtained for the amide pentaketide case. PikAIII/AIV-TE S148C showed the ratio of 6:14 ring to be consistent over time to 1:2, suggesting the primary selectivity for the 14-ring. Now, PikAIII/AIV-TE R1 reactions

showed 2:1 formation of products, unchanged during the time course. And lastly, PikAIII/AIV-TE R2 remained 1:1 of ring formation during the reaction time. These results suggest the potential use of the PikAIII/AIV-TE variants for amide/ester ring formation changing the selectivity for the ring sized macrolactone being formed.

Table IV.3 Ester pentaketide enzymatic reaction results for PikAIII/AIV-TE modules with engineered Pik TEs for the formation of 6- and 14-membered ring products. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data).

entry	substrate	Pik modules	time	products (6:14 ring)
1	Ester Pentaketide	AIII / AIV TE WT	30 mins	1:2
2	Ester Pentaketide	AIII / AIV TE S148C	30 mins	1:2
3	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V (**R1**)	30 mins	2:1
4	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V (**R2**)	30 mins	**1:1**
5	Ester Pentaketide	AIII / AIV TE WT	1h	1:1
6	Ester Pentaketide	AIII / AIV TE S148C	1h	**1:2**
7	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V (**R1**)	1h	2:1
8	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V (**R2**)	1h	**1:1**
9	Ester Pentaketide	AIII / AIV TE WT	2h	2:1
10	Ester Pentaketide	AIII / AIV TE S148C	2h	**1:2**
11	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V (**R1**)	2h	2:1
12	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V (**R2**)	2h	**1:1**
13	Ester Pentaketide	AIII / AIV TE WT	overnight	6 ring
14	Ester Pentaketide	AIII / AIV TE S148C	overnight	**1:2**
15	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V (**R1**)	overnight	2:1
16	Ester Pentaketide	AIII / AIV TE S148C, A217T, G222V, M271V, Y25C, L126V (**R2**)	overnight	**1:1**

Lastly, since we had the synthesized amide hexaketide substrate, we also decided to explore the formation of the corresponding 6:14 rings, using just PikAIV-TE variants. The results, as shown below in **Table IV.4** demonstrated the similarities when using the amide pentaketide and the PikAIII/AIV system. Furthermore, the reaction of the amide hexaketide against PikAIII/AIV-TE R2, to our surprise, produced a mass identified for the hybrid 12-membered macrolactone/lactam ring. The data suggests the Pik TE R2 variant has an apparent direct

selectivity for the 12-ring generation, without blocking the formation of the 6- and 14-cyclized products, a finding worth studying further.

Table IV.4 Amide hexaketide enzymatic reaction results for PikAIV-TE modules with engineered Pik TEs for the formation of 6- and 14-membered ring products. TOF-MS: 10 – 80% acetonitrile/water with 0.1% formic acid, 13 minutes (min) total run (TIC and EIC data).

entry	substrate	Pik modules	time	products (6:14 ring)
1	Amide Hexaketide	AIV TE WT	2h	1:1
2	Amide Hexaketide	AIV TE S148C	30 mins, 1h, 2h	**1:1**
3	Amide Hexaketide	AIV TE S148C, A217T, G222V (**R1**)	30 mins, 1h, 2h	**6 ring:trace**
4	Amide Hexaketide	AIV TE S148C, A217T, G222V, M271V, Y25C, L126V (**R2**)	30 mins, 1h, 2h	**4:1 *Formation of 12 ring**
5	Amide Hexaketide	AIV TE WT	overnight	6 ring:trace
6	Amide Hexaketide	AIV TE S148C	overnight	**1:1**
7	Amide Hexaketide	AIV TE S148C, A217T, G222V (**R1**)	overnight	**6 ring:trace**

IV.9 Conclusion/Summary

In summary, we studied in detail the PKS PikAIII-TE fusion protein capabilities and constrains applying the amide/ester containing pentaketide substrates. When subjecting these substrates to PikAIII-TE WT and PikAIII-TE S148C, several byproducts were generated, none of which are cyclized macrolactones. One major tendency observed is the appearance of SNAc hexaketides only when the TE is S148C in PikAIII-TE S148C, and lower quantities of hydrolyzed penta- and hexaketides are found. In the contrary, for when the TE is WT in PikAIII-TE WT, the only byproducts observed correspond to hydrolyzed penta- and hexaketides. The formation of these offloaded SNAc byproducts potentially blocked cyclization from occurring in the TE domain.

Elimination of SNAc from the reactions by using thiol scavengers resulted in only a slight decrease in these byproducts, without any ring formation. When utilizing MeMal-CoA and PikAIII-TE S148C module, the corresponding hybrid 12-membered macrolactone/lactam ring appeared although in small quantities, and for ester no ring was identified. Intact protein MS on the Pik TEs allowed the potential discovery that SNAc is forming a disulfide bond in the protein or presents tight non-covalent interactions, suggesting that it could be the main impediment as to why the TE is not capable of cyclization. We also ligated the engineered TEs from the directed evolution campaign to the modules PikAIII-TE and PikAIV-TE with the goal of understanding the effects on swapping or changing the TEs in the module for polyketide processing and also terminal cyclization. We discovered the formation of ring products at different ratios, depending on the substrate utilized and the engineered module, and on occasions, no ring production whatsoever. These modular enzymes, featuring altered functions, underscore the significance of directed evolution applied to individual domains, such as the Pik TE, in order to modify reaction outcomes and the resulting products. Taking a broader perspective, additional domains like KS, AT, ACP, and others may benefit from protein engineering in the future. It will involve leveraging site saturation mutagenesis for creating targeted mutant libraries and random mutagenesis to explore a broader spectrum of mutant forms. These approaches are crucial for testing PKS enzymes functioning together at the modular level. Such endeavors align with our overarching goals of comprehending PKS modules and mastering the art of engineering them effectively to achieve desired outcomes. Future work on additional engineered modules and PKS module 3D structures using cryo-EM tools could allow us to understand inter-domain interactions and how these PKSs work together for the processing, elongation and cyclization of polyketides into macrolactones and macrolides.

IV.10 Experimental

IV.10.1 General Information, Materials and Instrumentation

As described in Chapters 2 and 3.

IV.10.2 Synthetic Experimentals and Characterizations

As described in Chapters 2 and 3.

IV.10.3 Protein Biochemistry

As described in Chapters 2 and 3.

Intact Protein MS

Protein samples were analyzed in the QTOF-MS instrument, C4 column and prepared as follows: Pik TE WT alone: water = 41.9 µL, sodium phosphate buffer = 3.1 µL (50 mM final from 800 mM buffer stock), Pik TE WT = 5 µL (20 mM final from 200 µM dilution) / 50 µL total volume, for QTOF sample, diluted 1:1 with water, spun down → 30 µL to LC-MS vial.

Pik TE S148C alone: water = 41.9 µL, sodium phosphate buffer = 3.1 µL (50 mM final from 800 mM buffer stock), Pik TE S148C = 5 µL (20 mM final from 200 µM dilution) / 50 µL total volume, for QTOF sample, diluted 1:1 with water, spun down → 30 µL to LC-MS vial.

Pik TE WT + SNAc: water = 40.2 µL, sodium phosphate buffer = 3.1 µL (50 mM final from 800 mM buffer stock), SNAc = 1.7 µL (10 mM final from a 300 mM stock), Pik TE WT = 5 µL (20 mM final from 200 µM dilution) / 50 µL total volume, for QTOF sample, diluted 1:1 with water, spun down → 30 µL to LC-MS vial.

Pik TE S148C + SNAc: water = 40.2 µL, sodium phosphate buffer = 3.1 µL (50 mM final from 800 mM buffer stock), SNAc = 1.7 µL (10 mM final from a 300 mM stock), Pik TE S148C = 5 µL (20 mM final from 200 µM dilution) / 50 µL total volume, for QTOF sample, diluted 1:1 with water, spun down → 30 µL to LC-MS vial.

Pik TE S148C + SNAc + DTT: water = 36.1 µL, sodium phosphate buffer = 3.1 µL (50 mM final from 800 mM buffer stock), SNAc = 3.3 µL (20 mM final from a 300 mM stock), Pik TE S148C = 5 µL (20 mM final from 200 µM dilution), DTT = 2.5 µL (1 mM final from 20 mM stock) / 50 µL total volume, for QTOF sample, diluted 1:1 with water, spun down → 30 µL to LC-MS vial.

Pik TE WT + SNAc + DTT: water = 36.1 µL, sodium phosphate buffer = 3.1 µL (50 mM final from 800 mM buffer stock), SNAc = 3.3 µL (20 mM final from a 300 mM stock), Pik TE WT = 5 µL (20 mM final from 200 µM dilution), DTT = 2.5 µL (1 mM final from 20 mM stock) / 50 µL total volume, for QTOF sample, diluted 1:1 with water, spun down → 30 µL to LC-MS vial.

*sodium phosphate buffer = 800 mM $NaH_2PO_4 \cdot H_2O$ and Na_2HPO_4, 400 mM NaCl, pH 7.2

IV.10.4 DNA Cloning and Mutagenesis

Pik TE Variants PCR Amplifications and Ligations into PikAIII and PikAIV Modules

PCR amplification for Pik TE variants genes: The HindIII and XhoI restriction sites were added by PCR amplifications with DNA primers for HindIII (forward primer) and XhoI (reverse primer). 50 µL reaction: A modified QuikChange method (Stratagene) PCR contained distilled water, 4% DMSO, 0.2 µM both forward/reverse primers, 100 ng plasmid DNA template, 1X

Pyrococcus furiosus (*Pfu*) ultra II DNA polymerase reaction buffer, 0.2 mM dNTP mix (Promega: U151A) and 1 unit of *Pfu* ultra II fusion HotStart DNA polymerase (Agilent: 600672). PCR cycle conditions: 1) polymerase activation: 95 °C for 3 min, 2) denature: 95 °C for 30 sec, 3) annealing: 68 °C for 30 sec (temperatures changed depending on Tm of primers), 4) extension: 65 °C for 1.5 min, 5) terminal extension: 65 °C for 10 min. Cycles run: 30 (for steps 2-4). Upon completion of PCR amplification, 2 µL of DpnI (NEB: R0176S) were added and incubated at 37 ºC for 2 hours. After DpnI digestion, the DNA was purified with QIAGEN PCR clean up kit.

Restriction digest Pik TE amplicons and module vectors: The restriction enzymes utilized for this purpose included HindIII and XhoI located at both start and end of the Pik TE gene in the PikAIII-TE and PikAIV-TE vectors. 55 µL reaction: DNA (~100 ng either PCR amplicon product or vector), cut-smart 10X buffer, and water (if needed) were added to PCR tubes. Both restriction enzymes (120-140 ng) were added to the tubes, and they were incubated at 37 ºC for 1.5 hours at which time a 5 µL aliquot of the two reactions was taken and diluted with 5 µL water and 2 µL Purple DNA loading dye. Also, prepared a sample of the uncut plasmid as a control: 2 µL of plasmid, 8 µL water, 2 µL dye. These were run on 1% agarose DNA gel, at 135 volts for 30 min to determine if the reactions were complete. Once reaction completion was determined, 5 µL of DNA loading dye was added to the reaction of the digested vector and it was run on a 1% agarose DNA gel fitted with large comb at 114 volts for 60 min. The band was extracted under UV light and subjected to gel purification kit. Run alongside the gel purification was a PCR clean up kit on the digested amplicon DNA for ligation.

Ligation Pik TEs and modules: Water, T4 DNA ligase, T4 DNA ligase buffer, vector DNA, and insert DNA (ratio of insert:vector = 5:1) were added to a PCR tube and T4 DNA ligase was lastly added. The 20 µL reaction was mixed gently by up and down pipetting. The reaction was

incubated in a PCR thermocycler overnight at 16 ºC. The reaction was chilled on ice, 1 aliquot of chemically BL21(DE3) competent cells was defrosted on ice, and 5 µL of DNA was added to the tube. Heat shock: 5 mins on ice, 47 ºC warm bath for 45 sec. These were allowed to sit on ice for 5 min, 400 µL of SOC media was added, and the reactions were shaken at 37 ºC for 1 hour. The reactions were plated on LB kanamycin (50 µg/mL) plates using glass beads and grown overnight in the 37 ºC incubator. The plasmid miniprep kit from QIAGEN was utilized to extract DNA for sequencing plasmids and confirm successful Pik TE WT or variant ligation to the desired Pik module for further protein expression and purification.

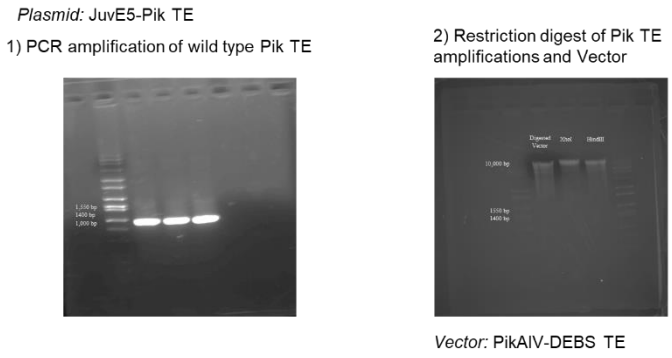


Figure IV.13 Example on a control experiment for PCR Pik TE amplification and restriction digest of Pik TE amplicon and PKS module vector.

CHAPTER V

Summary, Discussion and Future Directions

PKSs have been responsible for the wealth of numerous clinically relevant molecules to treat human diseases including cancer, autoimmune disorders, parasitic infections, bacterial pathogens, among others.[3, 4] My dissertation described the chemoenzymatic synthesis of macrolactone rings using synthetic substrates and PKS modules (PikAIII-TE fusion and PikAIII/AIV) from the antibiotic pikromycin or Pik biosynthetic pathway together with post-PKS tailoring enzymes (glycosyltransferases, P450 oxidases), and individual domains such as the Pik TE. The work herein strives to continue with the development of compounds with medicinal properties such as antibiotics, utilizing PKS megasynthases and the thioesterase cyclization domain. We investigated the capabilities and limitations of modules in the Pik system as well as engineered the Pik TE for higher yielding production of macrolactones. This chapter will summarize the findings of the three main projects I worked on for my dissertation with the goal of connecting the main discoveries and present the future work desired to understand how PKSs work together to assemble complex NPs and analogs, with the ultimate goal of generating improved biocatalysts for the production of novel molecules with biological activities.

V.1 Chemoenzymatic Synthesis of Macrolactone Rings and Potential Macrolide Antibiotics Utilizing PKSs and Post-PKS Tailoring Enzymes

V.1.1 Summary and Discussion of Work Completed

Previously, it was shown that the Pik fusion protein PikAIII-TE was capable of accepting unnatural substrates with truncations in the left half of the molecule corresponding to the methyl

group at C8 position in the native Pik pentaketide. Additionally, substrates bearing unnatural C8 and C9 stereocenters were incorporated. Both of these types of molecules were studied for the flexibility and restrictions of the modules and individual catalytic domains for the production of 12-membered macrolactone rings.[62, 65] For Chapter 2, we then synthesized new substrates having heteroatoms in place of carbons, such as the amide/ester pentaketides, in addition to substrates with altered side chains like the gem-dimethyl and cyclohexyl containing pentaketides. We decided to continue probing the diversification of macrolactones and macrolides using PikAIII-TE for the generation of 12-membered macrocycles and PikAIII/AIV for the first time to generate larger 14-membered macrocycles. With four newly pentaketides in hand, as well as the previously generated four truncated pentaketides[62, 65], we proceeded to perform the enzymatic reactions with the PKS modules to generate narbonolide, narbomycin and pikromycin analogs inspired by the Pik pathway.

The design and synthesis of pentaketide intermediates for the fusion protein PikAIII-TE and the last two modules in the Pik pathway (PikAIII/AIV) allowed the finding that these enzymes display flexibility to processing and generating the desired macrocyclic products when exposed to substrates with non-native functionalities. First, the analytical scale reactions analyzed in LC-MS for the amide, ester, gem-dimethyl and cyclohexyl pentaketides utilizing PikAIII-TE failed to generate the cyclic products. Instead, only amide and ester were accepted and extended by the catalytic domains, however only hydrolyzed byproducts and others were identified, suggesting PKS domain flexibility to an extent. Once again, the TE was unable to cyclize the corresponding intermediates, showing the TE as a catalytic bottleneck, supporting previous findings in the Sherman lab.[62, 65] The gem-dimethyl and cyclohexyl had minimal substate consumption, meaning that the PKS domains were not flexible enough to accept intermediates bearing bulkier altered side

chains. Subsequently, we expanded the analytical LC-MS analysis with PikAIII/AIV and successfully identified the masses corresponding to the 14-membered rings for the truncated pentaketides previously generated. We also identified masses for the amide and ester rings, and minimal substrate consumption for gem-dimethyl and cyclohexyl consistent with the PikAIII-TE findings. Further HPLC preparatory scale reactions with ~10 mg of each substrate allowed the isolation and NMR characterization of the 14-membered rings with yields between 71% to 11%.

Another important aspect of our study included the discovery of a novel macrolactone ring opening activity for the Pik TE cyclase. This finding relates to the appearance of an amide/ester 6-membered ring when performing the PKS reactions with PikAIII/AIV system. We hypothesized that Pik TE had a role for the smaller ring sized formation due to reaction time courses performed at 1 hour, 2 hour and overnight (24 hours) by taking reaction aliquots at each point. The time courses showed the complete overnight disappearance of the 14-ring for a shift into formation of the 6-ring. This outcome suggested the reaction is thermodynamically favored for the 6-ring. To further prove our hypothesis, we used the isolated 14-membered macrocycles with PikAIV and standalone Pik TEs (WT and S148C) and we identified that indeed these rings are being opened by the TE to then recyclize these generating the smaller ones. As a final proof of concept, we also tested Pik TE S148A bearing a catalytic Ser148 inactivation to an unreactive Ala, showing that certainly the active TE is required for the formation of the 6-membered ring.

After the generation of the Pik narbonolide analogs, we continued with the tailoring steps with the engineered strain of *Streptomyces venezuelae* (DHS316) for incorporating desosamine sugars and the utilization of P450s including PikC D50N-RhFRED and JuvD WT-RhFRED to generate the macrolide counterparts. Although the tailoring steps were met with success for the glycosyltransferases with yields between 33% to 75%, the P450 reactions failed to generate a

complete set of macrolides for the study of their potential medicinal properties with unreacted narbomycin analogs for the reactions, depending on whether the PikC or JuvD P450s were utilized. The yields for the macrolides or pikromycin analogs were first observed in analytical reaction LC-MS traces, and then HPLC scaled up reactions encompassed isolated 71% yield to trace amounts of product, suggesting more work to be done to understand P450 trends for oxidation of macrolides.

V.1.2 Future Directions

Chapter 2 shows the potential of using Pik PKS modules and their abilities to produce macrocycles with structural modifications when compared to the ring products formed in the Pik biosynthetic pathway. Although this work is near completion, follow up studies could potentially shed light on the module capabilities at large, but also investigations on single domains such as the KS and TE.

Findings that still need further assessments are the unreactive gem-dimethyl and cyclohexyl pentaketide substrates. We found that although these seem to barely react with the modular enzymes (PikAIII-TE and PikAIII/AIV), the pentaketides are in fact loaded to the KS, however no Claisen condensation occurs when utilizing MeMal-SNAc and the ACP domain *in trans* when subjecting them with KS-AT didomains and analyzing the intact MS data. In the future, these projects could benefit from a detailed KS study, including KS protein engineering strategies such as directed evolution.

Similarly, the findings of the PikAIII-TE not being able to generate amide and ester 12-membered rings, but being able to accept these and offload instead hydrolytic byproducts suggest yet again the Pik TE as a bottleneck for cyclization. Later, studies on modular PKS function and structural analysis via cryo-EM could shed light on the modular inter-domain communications and

protein-protein interactions that allow the diversification of macrolactones/macrolides. We also hoped that additional TE protein engineering efforts, besides the ones presented in Chapter 3 herein could potentially overcome hurdles when the TE is present in the PKS modules.

The last aspect that relates to single domain examinations include the ring opening-reclosing activity not known before for Pik TE. It would be intriguing to pursue the identification of other TEs that have evolved to perform ring opening with amide/ester 14-rings and other macrocycles. From a synthetic chemistry standpoint, finding additional biocatalysts capable of generating rings of diverse sizes with alternative selectivities could be applicable to chemoenzymatic strategies for total synthesis in addition to the study of the effect of ring size for biological activities in diverse compound libraries.

Lastly, the post-PKS tailoring enzymes showed limitations mainly when using the P450s. Although reactivity patterns were found for the apparent requirement of the C12 methyl group for PikC catalysis, and the C15 primary oxidation by JuvD for the macrolides, the addition of more P450s from diverse organisms and biosynthetic pathways, in addition to testing additional macrolides for oxidation could be supportive to determine functional groups essential for the P450s. These efforts will continue to be of high interest to perform in the near future analysis of the macrolides generated and others for their biological activities as antibiotics, anticancer agents and immunomodulators.

V.2 Directed Evolution of the Pik TE

V.2.1 Summary and Discussion of Work Completed

Chapter 3 revealed a protein engineering strategy with exciting results towards the study and development of improved TEs for the higher yielding cyclization of macrolactones and later

macrolides. We expanded the study of the pentaketide library presented in Chapter 2 by taking a closer look into the amide substrate.

Since the hybrid 12-membered macrolactone/lactam ring was not observed when utilizing PikAIII-TE fusion protein, we decided to synthesize an amide hexaketide intermediate to investigate if the Pik TE standalone is capable of cyclization. After completion of the synthetic intermediate, we performed enzymatic reactions against both the Pik TE WT and the Pik TE S148C engineered mutant. To our surprise, only the Pik TE S148C was capable of ring formation as shown in LC-MS traces for the analytical reactions. However, although the reactions went to completion, the competing reaction for the TE was the enzyme mediated hydrolysis observed as the higher yielding species. Further HPLC scale up and isolation revealed that indeed the 12-ring was formed in a 13% yield as verified by NMR and X-ray crystallography. Based on the low yielding formation of the macrocycle when employing the Pik TE S148C we decided to embark on a directed evolution strategy aimed at the generation of a Pik TE mutant library to expand the use of the TE PKS domains for ring formation and decrease on the undesired hydrolysis product.

The first round (R1) of evolution consisted of looking into the active site region of the Pik TE X-ray crystal structure with a covalent pentaketide affinity label (PDB: 2HFJ). Twenty-one residues were identified within 6 Å of the substrate. The crystal structure-substrate mimic complex likely represented a close approximation to the native hexaketide substrate covalently bound to Ser148 in the TE active site showing the Ser148, Asp176, His268 catalytic triad, and revealed the residues likely involved in binding and cyclization. This mutagenesis strategy allowed the generation of thousands of active site mutants by changing each residue to all possible canonical amino acids. We identified five hits with improved LC-MS yields which were carefully characterized by DNA sequencing, protein expression/purification and subsequent enzymatic

reaction LC-MS yield confirmation. We found a curious result where two of the hits were actually located within 12 Å away of the substrate in the Pik TE X-ray. These two hits resulted as non-silent mutations during the mutagenesis PCR reactions. Next, combinations of beneficial mutations produced a triple-mutant Pik TE S148C, A217T, G222V (Pik TE R1) by combining mutation S148C with A217T within 6 Å and G222V within 12 Å from the active site region, with a 3-fold increase on ring from 13% to 39%.

The second (R2) and third (R3) rounds of evolution were inspired by the finding that mutations distal from the active site appeared to be significant for cyclization. For this purpose, we pursued random mutagenesis via epPCR strategies. Screening again thousands of enzymatic reactions, we came across sixteen hits for round 2 and six hits for round 3. The mutations found varied from 6 - 20 Å, showcasing that indeed variations to other residues in the entire protein sequence improves the selectivity for macrocyclization rather than hydrolysis. In R2 the combination mutant generated was Pik TE S148C, A217T, G222V, M271V, Y25C, L126 (Pik TE R2) with a 5.5-fold improvement on cyclization from 13% to 71%. Meanwhile, R3 remained on a plateau for yield improvement, where we could not improve the 71% yield any higher with combination mutant Pik TE S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V, E294G, G138W (Pik TE R3). Although initially the R3 hits seemed to be promising, protein scale up and yield confirmation stalled at 71%. Lastly, scale up and HPLC isolation for the reactions using ~10 mg of amide hexaketide showed the desired cyclic product to be afforded in 42% (Pik TE R1), 79% (Pik TE R2) and 69% (Pik TE R3) isolated yields for the best yielding mutants at every round of evolution.

We were also eager to identify and understand the effects on the mutations for Pik TE catalysis by performing TTNs, thermal shift assay to measure stability and initial rates for substrate

consumption/macrolactone formation. The TTN results suggested possible enzyme destabilization over an 18-hour reaction period and lower catalyst loading or 0.2 mol %. For that reason, we performed the thermal shift assay where the TEs were challenged with increased temperatures to measure protein denaturation and overall stability. The data suggested that Pik TE WT and S148C retained similar melting temperature or Tm values. Now, R1 is slightly more stable as shown with the Tm's in comparison with Pik TE WT. In contrast, R2 and R3 showed a decrease in protein denaturation, where R2 was shown to be the least stable when compared again the Tm's to Pik TE WT. This suggested that although the mutations acquired during the directed evolution strategy improved the yield for ring formation, it had tradeoffs with the protein stability especially at R2 and R3. However, although the stability was decreased, the initial rates for substrate consumption and cyclization suggested that when the reactions are subjected to five times more TE or 1.0 mol %, these are complete in less than 1 hour and stability issues are unnoticed due to faster reaction rates with higher catalyst loading.

Another fascinating result found was the importance of the S148C mutation for Pik TE catalysis. Reversions to the Ser148 proved to be detrimental to the formation of hybrid 12-membered macrolactone/lactam ring. The enzymatic reaction with the amide hexaketide and Pik TE R1 C148S$_{reversion}$ showed a >10-fold decrease in ring formation (39% to 3% yield) by LC-MS analysis and similarly Pik TE R2 reversion mutant resulted in a decrease from 71% to a 14% yield (LC-MS). In our case, the Cys148 residue was found to be foundational for macrocyclization.

V.2.2 Future Directions

The work completed in Chapter 3 was met with success for the development of a Pik TE library with over 40+ variants purified and characterized for their properties and 6,000+ additional mutants counting the entire library screened. Future investigations for Pik TE and others will be

advantageous for our knowledge of single PKS domains and how to use TEs for the generation of macrolactones and macrolides.

One of the ongoing work to be finished for the Pik TE mutant library generated involves a collaboration with the Sigman Group at the University of Utah, where we have been working together for almost a year now towards the goal of understanding the effects of the mutations identified in the directed evolution screen. We provided them with the entire library of 40+ Pik TEs data, including yields, TTN values, stability screen and initial rates. Our goal is to perform MD and ML approaches to deepen our mechanistic knowledge of the cyclization process, the impact of the identified mutations on catalysis as well as predicting new mutations/variants. We hope to answer questions as to why the third round of evolution reached a plateau, and how to overcome it. Also, the predictions of other substrates in the active sites of these variants will be used to generate a potential library of hexa- and heptaketide substrates for the production of macrolactones in PKS systems. We hope to continue this collaboration to improve TEs in general for macrocyclization aided by computational methods.

Currently, we are also testing other substrates (hexa- and heptaketides) for the Pik TE WT, S148C, R1, R2 and R3 variants, to identify how the cyclization yields might differ compared to the amide hexaketide. More work is required to test additional substrates with diverse structures and functional groups, which will need more synthetic work to make these, and also biocatalysis efforts for testing the reaction outcomes and possible trends.

Another piece of information not covered in Chapter 3 was the synthesis of an ester hexaketide with a methoxy group at the C3 position. This substrate was synthesized by Carolyn A. Glasser, a former undergraduate student working with Jennifer J. Schmidt and later with Brian J. Curtis for the last couple of steps of the synthesis. Carolyn's independent research led to a senior thesis which

covered the outcomes for the ester methoxy substrate. This is a work still in progress, where we want to identify if we can analyze via MD simulations how the substrate binds and how the methoxy group might be impacting cyclization in Pik TE. Formation of the ester 12-membered ring has not been identified yet, and reactions with the methoxy substrate led to unreacted starting material and mostly hydrolysis products depending on the TE variant utilized. Ongoing efforts include generating an ester hexaketide bearing a hydroxyl group at C3 for comparison purposes and also synthesizing the corresponding methoxy amide hexaketide.

Lastly, we would like to continue with testing TEs from other organisms and biosynthetic pathways. The Sherman lab has a library of TEs (unpublished data) that can be studied further by incorporating mutations found in Chapter 3 for the case of Pik TE. We hope to verify the applicability of the residues identified for the cyclization selectivity and efficiency in TEs beyond the Pik pathway.

V.3 Evaluation of PKS Module PikAIII-TE and PikAIII/AIV for the Production of 12- and 14-membered Macrolactone Rings

V.3.1 Summary and Discussion of Work Completed

Chapter 4 pursued a thorough analysis of PikAIII-TE fusion protein and PikAIII/IV system towards the goal of shedding light on the enzymatic capabilities and limitations of the last two modules in the Pik pathway.

Initial LC-MS analysis of the amide/ester pentaketides against both PikAIII-TE WT and PikAIII-TE S148C revealed the potential issues as to why the module is unable of producing the desired 12-membered macrolactone rings. A trend was observed when the TE is WT in PikAIII-TE, where the only byproducts identified correspond to hydrolysis of the penta- and hexaketides. In contrast, when the TE is mutated to S148C in PikAIII-TE, the predominant byproducts found

are the SNAc containing penta- and hexaketides and to a lesser extent, the hydrolysis products found for PikAIII-TE WT. We hypothesized that the formation of these offloaded SNAc byproducts potentially blocked cyclization from occurring in the TE domain. We then identified that potentially, the SNAc byproducts are being offloaded by the ACP domain. This is due to running reactions against PikAIII-TE S148A inactive TE, where we observed the same SNAc byproducts as before including SNAc hexaketide byproducts that had a ketone instead of a hydroxyl group at C3 position due to KR skipping. We also avoided the use of the extender unit MeMal-SNAc in the reactions due to generating free SNAc in solution by performing enzymatic reactions with the native PKS extender unit MeMal-CoA. The outcome of the reactions proved that the desired 12-membered macrolactone/lactam ring was produced although in trace amounts and the ester 12-ring was not generated. Even though the hybrid 12-membered macrolactone/lactam ring was identified, the major product from the reaction was still the hydrolysis of the hexaketide intermediate, suggesting that the modular enzyme still has limitations when working together with the TE for cyclization.

As a last study of the effects of SNAc in the reactions, we performed intact protein MS and identified potential issues. We hypothesized that SNAc could be inducing a disulfide bond with available cysteines with the TEs, especially with the catalytic Cys148 in the mutant form of Pik TE avoiding cyclization in some cases. However, further studies need to be done to confirm or disprove this finding.

Finally, we swapped the Pik TEs in PikAIII-TE and PikAIV with the variants R1, R2 and R3 from Chapter 3 to investigate the effects on full module processing and cyclization when exchanging these TEs. We identified via LC-MS the formation of ring products at different ratios

when using amide/ester substrates, depending on the substrate and the TE mutant utilized and in some cases no ring formation was observed.

V.3.2 Future Directions

Despite Chapter 4 being a shorter study in this dissertation, it opened doors for other projects to emerge in the future. We still would like to compare and continue probing the trends found for PikAIII-TE WT against PikAIII-TE S148C. There was a trend present when the TE is WT, only hydrolysis byproducts are formed, but when the TE is S148C in the PikAIII-TE module, then the hydrolysis byproducts decrease considerably and SNAc byproducts appear in higher amounts. The fact that no SNAc byproducts are found when the TE WT is present, suggested issues at the protein level that we are not yet aware of. We performed investigations to try to decipher if these SNAc byproducts prevent cyclization in the TE, since we know from Chapter 3 that the standalone Pik TE can in fact generate the hybrid 12-membered macrolactone/lactam ring, then why is it not formed in the module? Future studies should focus on the effects of free SNAc in solution, since our work with intact protein MS might suggest potential SNAc interactions or binding in the TEs hypothetically preventing ring formation. Follow up studies can also incorporate the use of MeMal-CoA in PKS reactions as a proof of concept in comparison with MeMal-SNAc.

An additional domain to investigate resulting from this work is the KR domain. We identified a trend when the TEs are swapped or inactivated, the KR seems to stop working as a reductase enzyme as C3-keto products are mainly observed resulting from enzymatic reactions. There have been other examples of this in the Sherman lab[66], and there is more to be done in that realm.

We also tried changing thiophenol activating groups to SNAc groups in the substrates, including the amide hexaketide, which the results met with some failure. It would be beneficial to study the effect of activating groups for KS and TE loading. Although the Sherman lab has done

work on this before[112], other activating groups can be synthesized and tested. We are currently collaborating with the Healy Group at the University of New York in Abu Dhabi, where we have tested the effect on activating groups with innovative shorter length substrates, which they have synthesized for KS loading and PKS module processing. These work will also be important to connect the dots on PKS catalysis, since we have observed similar trends as seen before with our substrates.

To finish, future work will need to be done to continue our efforts to generate variant forms of the Pik PKS modules PikAIII-TE and PikAIIII/AIV to gain knowledge on the TE requirements for cyclization in the module and how these can play a role in inter-domain connections and protein-protein interactions. We wish to continue with other PKS modules from different pathways, and swap the corresponding TEs with the mutants generated in Chapter 3, in addition to TEs from other pathways[66], and diverse library of substrates. With this, we hope to see how PKS modules work together to generate complex molecules as well as later structural studies using Cryo-EM.

References

1. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020,** *83* (3), 770-803.
2. Dias, D. A.; Urban, S.; Roessner, U., A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2012,** *2* (2), 303-336.
3. Bernardini, S.; Tiezzi, A.; Laghezza Masci, V.; Ovidi, E., Natural Products for Human Health: An Historical Overview of the Drug Discovery Approaches. *Nat. Prod. Res.* **2018,** *32* (16), 1926-1950.
4. Pham, J. V.; Yilma, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C.; Park, S. R.; Yoon, Y. J., A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front. Microbiol.* **2019,** *10*.
5. Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; International Natural Product Sciences, T.; Supuran, C. T., Natural Products in Drug Discovery: Advances and Opportunities. *Nat. Rev. Drug Discov.* **2021,** *20* (3), 200-216.
6. Baran, P. S., Natural Product Total Synthesis: As Exciting as Ever and Here To Stay. *J. Am. Chem. Soc.* **2018,** *140* (14), 4751-4755.
7. Hong, J. Y., Natural Product Synthesis at the Interface of Chemistry and Biology. *Chem. Eur. J.* **2014,** *20* (33), 10204-10212.
8. Woodward, R. B.; Logusch, E.; Nambiar, K. P.; Sakan, K.; Ward, D. E.; Auyeung, B. W.; Balaram, P.; Browne, L. J.; Card, P. J.; Chen, C. H.; Chenevert, R. B.; Fliri, A.; Frobel, K.; Gais, H. J.; Garratt, D. G.; Hayakawa, K.; Heggie, W.; Hesson, D. P.; Hoppe, D.; Hoppe, I.; Hyatt, J. A.; Ikeda, D.; Jacobi, P. A.; Kim, K. S.; Kobuke, Y.; Kojima, K.; Krowicki, K.; Lee, V. J.; Leutert, T.; Malchenko, S.; Martens, J.; Matthews, R. S.; Ong, B. S.; Press, J. B.; Rajanbabu, T. V.; Rousseau, G.; Sauter, H. M.; Suzuki, M.; Tatsuta, K.; Tolbert, L. M.; Truesdale, E. A.; Uchida, I.; Ueda, Y.; Uyehara, T.; Vasella, A. T.; Vladuchick, W. C.; Wade, P. A.; Williams, R. M.; Wong, H. N. C., Asymmetric Total Synthesis of Erythromycin. 1. Synthesis of an Erythronolide a Seco Acid-Derivative Via Asymmetric Induction. *J. Am. Chem. Soc.* **1981,** *103* (11), 3210-3213.
9. Woodward, R. B.; Logusch, E.; Nambiar, K. P.; Sakan, K.; Ward, D. E.; Auyeung, B. W.; Balaram, P.; Browne, L. J.; Card, P. J.; Chen, C. H.; Chenevert, R. B.; Fliri, A.; Frobel, K.; Gais, H. J.; Garratt, D. G.; Hayakawa, K.; Heggie, W.; Hesson, D. P.; Hoppe, D.; Hoppe, I.; Hyatt, J. A.; Ikeda, D.; Jacobi, P. A.; Kim, K. S.; Kobuke, Y.; Kojima, K.; Krowicki, K.; Lee, V. J.; Leutert, T.; Malchenko, S.; Martens, J.; Matthews, R. S.; Ong, B. S.; Press, J. B.; Rajanbabu, T. V.; Rousseau, G.; Sauter, H. M.; Suzuki, M.; Tatsuta, K.; Tolbert, L. M.; Truesdale, E. A.; Uchida, I.; Ueda, Y.; Uyehara, T.; Vasella, A. T.; Vladuchick, W. C.; Wade, P. A.; Williams, R. M.; Wong, H. N. C., Asymmetric Total Synthesis of Erythromycin. 2. Synthesis of an Erythronolide a Lactone System. *J. Am. Chem. Soc.* **1981,** *103* (11), 3213-3215.

10. Woodward, R. B.; Logusch, E.; Nambiar, K. P.; Sakan, K.; Ward, D. E.; Auyeung, B. W.; Balaram, P.; Browne, L. J.; Card, P. J.; Chen, C. H.; Chenevert, R. B.; Fliri, A.; Frobel, K.; Gais, H. J.; Garratt, D. G.; Hayakawa, K.; Heggie, W.; Hesson, D. P.; Hoppe, D.; Hoppe, I.; Hyatt, J. A.; Ikeda, D.; Jacobi, P. A.; Kim, K. S.; Kobuke, Y.; Kojima, K.; Krowicki, K.; Lee, V. J.; Leutert, T.; Malchenko, S.; Martens, J.; Matthews, R. S.; Ong, B. S.; Press, J. B.; Rajanbabu, T. V.; Rousseau, G.; Sauter, H. M.; Suzuki, M.; Tatsuta, K.; Tolbert, L. M.; Truesdale, E. A.; Uchida, I.; Ueda, Y.; Uyehara, T.; Vasella, A. T.; Vladuchick, W. C.; Wade, P. A.; Williams, R. M.; Wong, H. N. C., Asymmetric Total Synthesis of Erythromycin. 3. Total Synthesis of Erythromycin. *J. Am. Chem. Soc.* **1981,** *103* (11), 3215-3217.
11. Wang, Z.; Hui, C. G., Contemporary Advancements in the Semi-Synthesis of Bioactive Terpenoids and Steroids. *Org. Biomol. Chem.* **2021,** *19* (17), 3791-3812.
12. Majhi, S.; Das, D., Chemical Derivatization of Natural Products: Semisynthesis and Pharmacological Aspects-A Decade Update. *Tetrahedron* **2021,** *78*.
13. Vollmann, D. J.; Winand, L.; Nett, M., Emerging Concepts in the Semisynthetic and Mutasynthetic Production of Natural Products. *Curr. Opin. Biotech.* **2022,** *77*, 102761.
14. Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K., Engineering the Third Wave of Biocatalysis. *Nature* **2012,** *485* (7397), 185-194.
15. Sheldon, R. A.; Brady, D., Broadening the Scope of Biocatalysis in Sustainable Organic Synthesis. *ChemSusChem* **2019,** *12* (13), 2859-2881.
16. Wu, S.; Snajdrova, R.; Moore, J. C.; Baldenius, K.; Bornscheuer, U. T., Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chem. Int. Ed.* **2021,** *60* (1), 88-119.
17. Patel, R. N., Biocatalysis for Synthesis of Pharmaceuticals. *Bioorg. Med. Chem.* **2018,** *26* (7), 1252-1274.
18. Sun, H.; Zhang, H.; Ang, E. L.; Zhao, H., Biocatalysis for the Synthesis of Pharmaceuticals and Pharmaceutical Intermediates. *Bioorg. Med. Chem.* **2018,** *26* (7), 1275-1284.
19. Zhang, H. R.; Wang, Y.; Wu, J. Q.; Skalina, K.; Pfeifer, B. A., Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. coli as a Heterologous Host. *Chem. Biol.* **2010,** *17* (11), 1232-1240.
20. Du, L.; Lou, L., PKS and NRPS Release Mechanisms. *Nat. Prod. Rep.* **2010,** *27* (2), 255-278.
21. Williams, G. J., Engineering Polyketide Synthases and Nonribosomal Peptide Synthetases. *Curr. Opin. Struct. Biol.* **2013,** *23* (4), 603-612.
22. Hwang, S.; Lee, N.; Cho, S.; Palsson, B.; Cho, B. K., Repurposing Modular Polyketide Synthases and Non-Ribosomal Peptide Synthetases for Novel Chemical Biosynthesis. *Front. Mol. Biosci.* **2020,** *7* (87).
23. Ansari, M. Z.; Yadav, G.; Gokhale, R. S.; Mohanty, D., NRPS-PKS: A Knowledge-Based Resource for Analysis of NRPS/PKS Megasynthases. *Nucleic Acids Res.* **2004,** *32* (Web Server issue), W405-W413.
24. Horsman, M. E.; Hari, T. P. A.; Boddy, C. N., Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016,** *33* (2), 183-202.
25. Mullowney, M. W.; McClure, R. A.; Robey, M. T.; Kelleher, N. L.; Thomson, R. J., Natural Products from Thioester Reductase Containing Biosynthetic Pathways. *Nat. Prod. Rep.* **2018,** *35* (9), 847-878.
26. Tang, M. C.; Zou, Y.; Watanabe, K.; Walsh, C. T.; Tang, Y., Oxidative Cyclization in Natural Product Biosynthesis. *Chem. Rev.* **2017,** *117* (8), 5226-5333.

27. Weissman, K. J., Genetic Engineering of Modular PKSs: From Combinatorial Biosynthesis to Synthetic Biology. *Nat. Prod. Rep.* **2016,** *33* (2), 203-230.
28. Cantu, D. C.; Chen, Y.; Reilly, P. J., Thioesterases: A New Perspective Based on Their Primary and Tertiary Structures. *Protein Sci.* **2010,** *19* (7), 1281-1295.
29. Kotowska, M.; Pawlik, K., Roles of Type II Thioesterases and their Application for Secondary Metabolite Yield Improvement. *Appl. Microbiol. Biotechnol.* **2014,** *98* (18), 7735-7746.
30. Bruner, S. D.; Weber, T.; Kohli, R. M.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T.; Stubbs, M. T., Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE. *Structure* **2002,** *10* (3), 301-310.
31. Miller, B. R.; Gulick, A. M., In *Nonribosomal Peptide and Polyketide Biosynthesis: Methods in Molecular Biology*, Evans, B., Ed. Humana Press, New York: 2016; Vol. 1401, pp 3-29.
32. Loll, P. J.; Upton, E. C.; Nahoum, V.; Economou, N. J.; Cocklin, S., The High Resolution Structure of Tyrocidine A Reveals an Amphipathic Dimer. *Biochim. Biophys. Acta* **2014,** *1838* (5), 1199-1207.
33. Samel, S. A.; Wagner, B.; Marahiel, M. A.; Essen, L. O., The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-Ribosomal Lipopeptide. *J. Mol. Biol.* **2006,** *359* (4), 876-889.
34. Tsai, S. C.; Miercke, L. J.; Krucinski, J.; Gokhale, R.; Chen, J. C.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M., Crystal Structure of the Macrocycle-Forming Thioesterase Domain of the Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** *98* (26), 14808-14813.
35. Tsai, S. C.; Lu, H.; Cane, D. E.; Khosla, C.; Stroud, R. M., Insights into Channel Architecture and Substrate Specificity from Crystal Structures of Two Macrocycle-Forming Thioesterases of Modular Polyketide Synthases. *Biochemistry* **2002,** *41* (42), 12598-12606.
36. Akey, D. L.; Kittendorf, J. D.; Giraldes, J. W.; Fecik, R. A.; Sherman, D. H.; Smith, J. L., Structural Basis for Macrolactonization by the Pikromycin Thioesterase. *Nat. Chem. Biol.* **2006,** *2* (10), 537-542.
37. Trauger, J. W.; Kohli, R. M.; Walsh, C. T., Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry* **2001,** *40* (24), 7092-7098.
38. Galea, C. A.; Roberts, K. D.; Zhu, Y.; Thompson, P. E.; Li, J.; Velkov, T., Functional Characterization of the Unique Terminal Thioesterase Domain from Polymyxin Synthetase. *Biochemistry* **2017,** *56* (4), 657-668.
39. Korman, T. P.; Crawford, J. M.; Labonte, J. W.; Newman, A. G.; Wong, J.; Townsend, C. A.; Tsai, S. C., Structure and Function of an Iterative Polyketide Synthase Thioesterase Domain Catalyzing Claisen Cyclization in Aflatoxin Biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (14), 6246-6251.
40. Schaffer, J. E.; Reck, M. R.; Prasad, N. K.; Wencewicz, T. A., β-Lactone Formation During Product Release from a Nonribosomal Peptide Synthetase. *Nat. Chem. Biol.* **2017,** *13* (7), 737-744.
41. Wells, J. S.; Trejo, W. H.; Principe, P. A.; Sykes, R. B., Obafluorin, a Novel Beta-Lactone Produced by Pseudomonas fluorescens - Taxonomy, Fermentation and Biological Properties. *J. Antibiot. (Tokyo)* **1984,** *37* (7), 802-803.

42. Tymiak, A. A.; Culver, C. A.; Malley, M. F.; Gougoutas, J. Z., Structure of Obafluorin - An Antibacterial Beta-Lactone from Pseudomonas fluorescens. *J. Org. Chem.* **1985,** *50* (26), 5491-5495.
43. Ray, L.; Yamanaka, K.; Moore, B. S., A Peptidyl-Transesterifying Type I Thioesterase in Salinamide Biosynthesis. *Angew. Chem. Int. Ed.* **2016,** *55* (1), 364-367.
44. Trischman, J. A.; Tapiolas, D. M.; Jensen, P. R.; Dwight, R.; Fenical, W.; Mckee, T. C.; Ireland, C. M.; Stout, T. J.; Clardy, J., Salinamides A and B: Anti-Inflammatory Depsipeptides from a Marine Streptomycete. *J. Am. Chem. Soc.* **1994,** *116* (2), 757-758.
45. Bonner, D. P.; Osullivan, J.; Tanaka, S. K.; Clark, J. M.; Whitney, R. R., Lysobactin, a Novel Antibacterial Agent Produced by Lysobacter Sp. 2. Biological Properties. *J. Antibiot. (Tokyo)* **1988,** *41* (12), 1745-1751.
46. Osullivan, J.; Mccullough, J. E.; Tymiak, A. A.; Kirsch, D. R.; Trejo, W. H.; Principe, P. A., Lysobactin, a Novel Antibacterial Agent Produced by Lysobacter Sp. 1. Taxonomy, Isolation and Partial Characterization. *J. Antibiot. (Tokyo)* **1988,** *41* (12), 1740-1744.
47. Hou, J.; Robbel, L.; Marahiel, M. A., Identification and Characterization of the Lysobactin Biosynthetic Gene Cluster Reveals Mechanistic Insights into an Unusual Termination Module Architecture. *Chem. Biol.* **2011,** *18* (5), 655-664.
48. Oliver, R. A.; Li, R. F.; Townsend, C. A., Monobactam Formation in Sulfazecin by a Nonribosomal Peptide Synthetase Thioesterase. *Nat. Chem. Biol.* **2018,** *14* (1), 5-7.
49. Asai, M.; Haibara, K.; Muroi, M.; Kintaka, K.; Kishi, T., Sulfazecin, a Novel Beta-Lactam Antibiotic of Bacterial Origin - Isolation and Chemical Characterization. *J. Antibiot. (Tokyo)* **1981,** *34* (6), 621-627.
50. Lin, T. Y.; Borketey, L. S.; Prasad, G.; Waters, S. A.; Schnarr, N. A., Sequence, Cloning, and Analysis of the Fluvirucin B1 Polyketide Synthase from Actinomadura vulgaris. *ACS Synth. Biol.* **2013,** *2* (11), 635-642.
51. Naruse, N.; Konishi, M.; Oki, T.; Inouye, Y.; Kakisawa, H., Fluvirucin-A1, Fluvirucin-A2, Fluvirucin-B1, Fluvirucin-B2, Fluvirucin-B3, Fluvirucin-B4 and Fluvirucin-B5, New Antibiotics Active against Influenza-a Virus .3. The Stereochemistry and Absolute-Configuration of Fluvirucin-A1. *J. Antibiot. (Tokyo)* **1991,** *44* (7), 756-761.
52. Naruse, N.; Tenmyo, O.; Kawano, K.; Tomita, K.; Ohgusa, N.; Miyaki, T.; Konishi, M.; Oki, T., Fluvirucin-A1, Fluvirucin-A2, Fluvirucin-B1, Fluvirucin-B2, Fluvirucin-B3, Fluvirucin-B4, and Fluvirucin-B5, New Antibiotics Active against Influenza-a Virus .1. Production, Isolation, Chemical-Properties and Biological-Activities. *J. Antibiot. (Tokyo)* **1991,** *44* (7), 733-740.
53. Naruse, N.; Tsuno, T.; Sawada, Y.; Konishi, M.; Oki, T., Fluvirucin-A1, Fluvirucin-A2, Fluvirucin-B1, Fluvirucin-B2, Fluvirucin-B3, Fluvirucin-B4, and Fluvirucin-B5, New Antibiotics Active against Influenza-a Virus .2. Structure Determination. *J. Antibiot. (Tokyo)* **1991,** *44* (7), 741-755.
54. Schmidt, J. J.; Khatri, Y.; Brody, S. I.; Zhu, C.; Pietraszkiewicz, H.; Valeriote, F. A.; Sherman, D. H., A Versatile Chemoenzymatic Synthesis for the Discovery of Potent Cryptophycin Analogs. *ACS Chem. Biol.* **2020,** *15* (2), 524-532.
55. Schwartz, R. E.; Hirsch, C. F.; Sesin, D. F.; Flor, J. E.; Chartrain, M.; Fromtling, R. E.; Harris, G. H.; Salvatore, M. J.; Liesch, J. M.; Yudin, K., Pharmaceuticals from Cultured Algae. *J. Ind. Microbiol.* **1990,** *5* (2-3), 113-123.

56. Smith, C. D.; Zhang, X. Q.; Mooberry, S. L.; Patterson, G. M. L.; Moore, R. E., Cryptophycin - A New Antimicrotubule Agent Active Against Drug-Resistant Cells. *Cancer Res.* **1994,** *54* (14), 3779-3784.
57. Magarvey, N. A.; Beck, Z. Q.; Golakoti, T.; Ding, Y.; Huber, U.; Hemscheidt, T. K.; Abelson, D.; Moore, R. E.; Sherman, D. H., Biosynthetic Characterization and Chemoenzymatic Assembly of the Cryptophycins. Potent anticancer Agents from Cyanobionts. *ACS Chem. Biol.* **2006,** *1* (12), 766-779.
58. Verma, V. A.; Pillow, T. H.; DePalatis, L.; Li, G.; Phillips, G. L.; Polson, A. G.; Raab, H. E.; Spencer, S.; Zheng, B., The Cryptophycins as Potent Payloads for Antibody Drug Conjugates. *Bioorg. Med. Chem. Lett.* **2015,** *25* (4), 864-868.
59. Mandalapu, D.; Ji, X.; Chen, J.; Guo, C.; Liu, W. Q.; Ding, W.; Zhou, J.; Zhang, Q., Thioesterase-Mediated Synthesis of Teixobactin Analogues: Mechanism and Substrate Specificity. *J. Org. Chem.* **2018,** *83* (13), 7271-7275.
60. Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Schaberle, T. F.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. A.; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K., A New Antibiotic Kills Pathogens Without Detectable Resistance. *Nature* **2015,** *517* (7535), 455-459.
61. Hansen, D. A.; Rath, C. M.; Eisman, E. B.; Narayan, A. R.; Kittendorf, J. D.; Mortison, J. D.; Yoon, Y. J.; Sherman, D. H., Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin. *J. Am. Chem. Soc.* **2013,** *135* (30), 11232-11238.
62. Hansen, D. A.; Koch, A. A.; Sherman, D. H., Identification of a Thioesterase Bottleneck in the Pikromycin Pathway through Full-Module Processing of Unnatural Pentaketides. *J. Am. Chem. Soc.* **2017,** *139* (38), 13450-13455.
63. Heberlig, G. W.; Wirz, M.; Wang, M.; Boddy, C. N., Resorcylic Acid Lactone Biosynthesis Relies on a Stereotolerant Macrocyclizing Thioesterase. *Org. Lett.* **2014,** *16* (22), 5858-5861.
64. Heberlig, G. W.; Brown, J. T. C.; Simard, R. D.; Wirz, M.; Zhang, W.; Wang, M.; Susser, L. I.; Horsman, M. E.; Boddy, C. N., Chemoenzymatic Macrocycle Synthesis Using Resorcylic Acid Lactone Thioesterase Domains. *Org. Biomol. Chem.* **2018,** *16* (32), 5771-5779.
65. Koch, A. A.; Hansen, D. A.; Shende, V. V.; Furan, L. R.; Houk, K. N.; Jimenez-Oses, G.; Sherman, D. H., A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst. *J. Am. Chem. Soc.* **2017,** *139* (38), 13456-13465.
66. Koch, A. A.; Schmidt, J. J.; Lowell, A. N.; Hansen, D. A.; Coburn, K. M.; Chemler, J. A.; Sherman, D. H., Probing Selectivity and Creating Structural Diversity Through Hybrid Polyketide Synthases. *Angew. Chem. Int. Ed.* **2020,** *59* (32), 13575-13580.
67. Qiao, L.; Fang, J.; Zhu, P.; Huang, H.; Dang, C.; Pang, J.; Gao, W.; Qiu, X.; Huang, L.; Li, Y., A Novel Chemoenzymatic Approach to Produce Cilengitide Using the Thioesterase Domain from Microcystis aeruginosa Microcystin Synthetase C. *Protein J.* **2019,** *38* (6), 658-666.
68. Tripathi, A.; Choi, S. S.; Sherman, D. H.; Kim, E. S., Thioesterase Domain Swapping of a Linear Polyketide Tautomycetin with a Macrocyclic Polyketide Pikromycin in Streptomyces sp. CK4412. *J. Ind. Microbiol.* **2016,** *43* (8), 1189-1193.
69. Gehring, A. M.; Mori, I.; Walsh, C. T., Reconstitution and Characterization of the Escherichia coli Enterobactin Synthetase from EntB, EntE, and EntF. *Biochemistry* **1998,** *37* (8), 2648-2659.

70. Hoyer, K. M.; Mahlert, C.; Marahiel, M. A., The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chem. Biol.* **2007,** *14* (1), 13-22.
71. Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dornan, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W., Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature* **2019,** *565* (7737), 112-117.
72. Heberlig, G. W.; Boddy, C. N., Thioesterase from Cereulide Biosynthesis Is Responsible for Oligomerization and Macrocyclization of a Linear Tetradepsipeptide. *J. Nat. Prod.* **2020,** *83* (6), 1990-1997.
73. Agata, N.; Mori, M.; Ohta, M.; Suwan, S.; Ohtani, I.; Isobe, M., A Novel Dodecadepsipeptide, Cereulide, Isolated from Bacillus cereus Causes Vacuole Formation in Hep-2 Cells. *FEMS Microbiol. Lett.* **1994,** *121* (1), 31-34.
74. Agata, N.; Ohta, M.; Mori, M.; Isobe, M., A Novel Dodecadepsipeptide, Cereulide, Is an Emetic Toxin of Bacillus cereus. *FEMS Microbiol. Lett.* **1995,** *129* (1), 17-19.
75. Marxen, S.; Stark, T. D.; Rutschle, A.; Lucking, G.; Frenzel, E.; Scherer, S.; Ehling-Schulz, M.; Hofmann, T., Depsipeptide Intermediates Interrogate Proposed Biosynthesis of Cereulide, the Emetic Toxin of Bacillus cereus. *Sci. Rep.* **2015,** *5* (10637).
76. Zhou, Y. J.; Prediger, P.; Dias, L. C.; Murphy, A. C.; Leadlay, P. F., Macrodiolide Formation by the Thioesterase of a Modular Polyketide Synthase. *Angew. Chem. Int. Ed.* **2015,** *54* (17), 5232-5235.
77. Zhou, Y. J.; Murphy, A. C.; Samborskyy, M.; Prediger, P.; Dias, L. C.; Leadlay, P. F., Iterative Mechanism of Macrodiolide Formation in the Anticancer Compound Conglobatin. *Chem. Biol.* **2015,** *22* (6), 745-754.
78. Westley, J. W.; Liu, C. M.; Evans, R. H.; Blount, J. F., Conglobatin, a Novel Macrolide Dilactone from Streptomyces conglobatus ATCC 31005. *J. Antibiot. (Tokyo)* **1979,** *32* (9), 874-877.
79. Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A., Deconstruction of Iterative Multidomain Polyketide Synthase Function. *Science* **2008,** *320* (5873), 243-246.
80. Gao, S. S.; Duan, A.; Xu, W.; Yu, P.; Hang, L.; Houk, K. N.; Tang, Y., Phenalenone Polyketide Cyclization Catalyzed by Fungal Polyketide Synthase and Flavin-Dependent Monooxygenase. *J. Am. Chem. Soc.* **2016,** *138* (12), 4249-4259.
81. Zhou, H.; Li, Y.; Tang, Y., Cyclization of Aromatic Polyketides from Bacteria and Fungi. *Nat. Prod. Rep.* **2010,** *27* (6), 839-868.
82. Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y., Identification of Claisen Cyclase Domain in Fungal Polyketide Synthase WA, a Naphthopyrone Synthase of Aspergillus nidulans. *Chem. Biol.* **2001,** *8* (2), 189-197.
83. Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S., Physically Discrete Beta-Lactamase-Type Thioesterase Catalyzes Product Release in Atrochrysone Synthesis by Iterative Type I Polyketide Synthase. *Chem. Biol.* **2009,** *16* (6), 613-623.
84. Vagstad, A. L.; Hill, E. A.; Labonte, J. W.; Townsend, C. A., Characterization of a Fungal Thioesterase Having Claisen Cyclase and Deacetylase Activities in Melanin Biosynthesis. *Chem. Biol.* **2012,** *19* (12), 1525-1534.

85. Takano, Y.; Kubo, Y.; Shimizu, K.; Mise, K.; Okuno, T.; Furusawa, I., Structural-Analysis of Pks1, a Polyketide Synthase Gene Involved in Melanin Biosynthesis in Colletotrichum lagenarium. *Mol. Gen. Genet.* **1995,** *249* (2), 162-167.
86. Fujii, I.; Mori, Y.; Watanabe, A.; Kubo, Y.; Tsuji, G.; Ebizuka, Y., Enzymatic Synthesis of 1,3,6,8-Tetrahydroxynaphthalene Solely from Malonyl Coenzyme A by a Fungal Iterative Type I Polyketide Synthase PKS1. *Biochemistry* **2000,** *39* (30), 8853-8858.
87. Fujii, I.; Mori, Y.; Watanabe, A.; Kubo, Y.; Tsuji, G.; Ebizuka, Y., Heterologous Expression and Product Identification of Colletotrichum lagenarium Polyketide Synthase Encoded by the PKS1 Gene Involved in Melanin Biosynthesis. *Biosci. Biotechnol. Biochem.* **1999,** *63* (8), 1445-1452.
88. Kudo, F.; Asou, Y.; Watanabe, M.; Kitayama, T.; Eguchi, T., Potent Oligomerization and Macrocyclization Activity of the Thioesterase Domain of Vicenistatin Polyketide Synthase. *Synlett.* **2012,** (12), 1843-1846.
89. Matsuda, K.; Zhai, R.; Mori, T.; Kobayashi, M.; Sano, A.; Abe, I.; Wakimoto, T., Heterochiral Coupling in Non-Ribosomal Peptide Macrolactamization. *Nat. Catal.* **2020**, 507-515.
90. Crawford, J. M.; Korman, T. P.; Labonte, J. W.; Vagstad, A. L.; Hill, E. A.; Kamari-Bidkorpeh, O.; Tsai, S. C.; Townsend, C. A., Structural Basis for Biosynthetic Programming of Fungal Aromatic Polyketide Cyclization. *Nature* **2009,** *461* (7267), 1139-1143.
91. Barajas, J. F.; Shakya, G.; Moreno, G.; Rivera, H.; Jackson, D. R.; Topper, C. L.; Vagstad, A. L.; La Clair, J. J.; Townsend, C. A.; Burkart, M. D.; Tsai, S. C., Polyketide Mimetics Yield Structural and Mechanistic Insights into Product Template Domain Function in Nonreducing Polyketide Synthases. *Proc. Natl. Acad. Sci. U.S.A.* **2017,** *114* (21), E4142-E4148.
92. Fair, R. J.; Tor, Y., Antibiotics and Bacterial Resistance in the 21st Century. *Perspect. Medicin. Chem.* **2014,** *6*, 25-64.
93. Tenson, T.; Lovmar, M.; Ehrenberg, M., The Mechanism of Action of Macrolides, Lincosamides and Streptogramin B Reveals the Nascent Peptide Exit Path in the Ribosome. *J. Mol. Biol.* **2003,** *330* (5), 1005-1014.
94. Golkar, T.; Zielinski, M.; Berghuis, A. M., Look and Outlook on Enzyme-Mediated Macrolide Resistance. *Front. Microbiol.* **2018,** *9* (1942).
95. Katz, L.; Ashley, G. W., Translation and Protein Synthesis: Macrolides. *Chem. Rev.* **2005,** *105* (2), 499-528.
96. Witzky, A.; Tollerson, R., 2nd; Ibba, M., Translational Control of Antibiotic Resistance. *Open Biol.* **2019,** *9* (7), 190051.
97. Liu, M.; Douthwaite, S., Activity of the Ketolide Telithromycin is Refractory to Erm Monomethylation of Bacterial rRNA. *Antimicrob. Agents Chemother.* **2002,** *46* (6), 1629-1633.
98. Depardieu, F.; Courvalin, P., Mutation in 23S rRNA Responsible for Resistance to 16-Membered Macrolides and Streptogramins in Streptococcus pneumoniae. *Antimicrob. Agents Chemother.* **2001,** *45* (1), 319-323.
99. Lebel, M., Pharmacokinetic Properties of Clarithromycin: A Comparison with Erythromycin and Azithromycin. *Can. J. Infect. Dis.* **1993,** *4* (3), 148-152.
100. Echols, R. M., Understanding the Regulatory Hurdles for Antibacterial Drug Development in the Post-Ketek World. *Ann. N. Y. Acad. Sci.* **2011,** *1241*, 153-161.
101. Seiple, I. B.; Zhang, Z.; Jakubec, P.; Langlois-Mercier, A.; Wright, P. M.; Hog, D. T.; Yabu, K.; Allu, S. R.; Fukuzaki, T.; Carlsen, P. N.; Kitamura, Y.; Zhou, X.; Condakes, M.

L.; Szczypinski, F. T.; Green, W. D.; Myers, A. G., A Platform for the Discovery of New Macrolide Antibiotics. *Nature* **2016,** *533* (7603), 338-345.
102. Fernandes, P.; Martens, E.; Pereira, D., Nature Nurtures the Design of New Semi-Synthetic Macrolide Antibiotics. *J. Antibiot. (Tokyo)* **2017,** *70* (5), 527-533.
103. Rossiter, S. E.; Fletcher, M. H.; Wuest, W. M., Natural Products as Platforms To Overcome Antibiotic Resistance. *Chem. Rev.* **2017,** *117* (19), 12415-12474.
104. Xue, Y.; Zhao, L.; Liu, H. W.; Sherman, D. H., A Gene Cluster for Macrolide Antibiotic Biosynthesis in Streptomyces venezuelae: Architecture of Metabolic Diversity. *Proc. Natl. Acad. Sci. U.S.A* **1998,** *95* (21), 12111-12116.
105. Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L., Modular Organization of Genes Required for Complex Polyketide Biosynthesis. *Science* **1991,** *252* (5006), 675-679.
106. Jung, W. S.; Lee, S. K.; Hong, J. S.; Park, S. R.; Jeong, S. J.; Han, A. R.; Sohng, J. K.; Kim, B. G.; Choi, C. Y.; Sherman, D. H.; Yoon, Y. J., Heterologous Expression of Tylosin Polyketide Synthase and Production of a Hybrid Bioactive Macrolide in Streptomyces venezuelae. *Appl. Microbiol. Biotechnol.* **2006,** *72* (4), 763-769.
107. Weissman, K. J., Uncovering the Structures of Modular Polyketide Synthases. *Nat. Prod. Rep.* **2015,** *32* (3), 436-453.
108. Weissman, K. J., The Structural Biology of Biosynthetic Megaenzymes. *Nat. Chem. Biol.* **2015,** *11* (9), 660-670.
109. Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R.; Hakansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G., Structure of a Modular Polyketide Synthase. *Nature* **2014,** *510* (7506), 512-517.
110. Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R.; Hakansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G., Structural Rearrangements of a Polyketide Synthase Module During its Catalytic Cycle. *Nature* **2014,** *510* (7506), 560-564.
111. Mortison, J. D.; Sherman, D. H., Frontiers and Opportunities in Chemoenzymatic Synthesis. *J. Org. Chem.* **2010,** *75* (21), 7041-7051.
112. Hansen, D. A.; Koch, A. A.; Sherman, D. H., Substrate Controlled Divergence in Polyketide Synthase Catalysis. *J. Am. Chem. Soc.* **2015,** *137* (11), 3735-3738.
113. Mortison, J. D.; Kittendorf, J. D.; Sherman, D. H., Synthesis and Biochemical Analysis of Complex Chain-Elongation Intermediates for Interrogation of Molecular Specificity in the Erythromycin and Pikromycin Polyketide Synthases. *J. Am. Chem. Soc.* **2009,** *131* (43), 15784-15793.
114. Aldrich, C. C.; Beck, B. J.; Fecik, R. A.; Sherman, D. H., Biochemical Investigation of Pikromycin Biosynthesis Employing Native Penta- and Hexaketide Chain Elongation Intermediates. *J. Am. Chem. Soc.* **2005,** *127* (23), 8441-8452.
115. Aldrich, C. C.; Venkatraman, L.; Sherman, D. H.; Fecik, R. A., Chemoenzymatic Synthesis of the Polyketide Macrolactone 10-Deoxymethynolide. *J. Am. Chem. Soc.* **2005,** *127* (25), 8910-8911.
116. Weissman, K. J., Polyketide Biosynthesis: Understanding and Exploiting Modularity. *Philos. Trans. A Math. Phys. Eng. Sci.* **2004,** *362* (1825), 2671-2690.
117. Parenty, A.; Moreau, X.; Niel, G.; Campagne, J. M., Update 1 of: Macrolactonizations in the Total Synthesis of Natural Products. *Chem. Rev.* **2013,** *113* (1), PR1-40.

118. Furstner, A., Lessons from Natural Product Total Synthesis: Macrocyclization and Postcyclization Strategies. *Acc. Chem. Res.* **2021,** *54* (4), 861-874.
119. Boddy, C. N.; Schneider, T. L.; Hotta, K.; Walsh, C. T.; Khosla, C., Epothilone C Macrolactonization and Hydrolysis are Catalyzed by the Isolated Thioesterase Domain of Epothilone Polyketide Synthase. *J. Am. Chem. Soc.* **2003,** *125* (12), 3428-3429.
120. Lowell, A. N.; DeMars, M. D.; Slocum, S. T.; Yu, F. A.; Anand, K.; Chemler, J. A.; Korakavi, N.; Priessnitz, J. K.; Park, S. R.; Koch, A. A.; Schultz, P. J.; Sherman, D. H., Chemoenzymatic Total Synthesis and Structural Diversification of Tylactone-Based Macrolide Antibiotics through Late-Stage Polyketide Assembly, Tailoring, and C-H Functionalization. *J. Am. Chem. Soc.* **2017,** *139* (23), 7913-7920.
121. Negretti, S.; Narayan, A. R.; Chiou, K. C.; Kells, P. M.; Stachowski, J. L.; Hansen, D. A.; Podust, L. M.; Montgomery, J.; Sherman, D. H., Directing Group-Controlled Regioselectivity in an Enzymatic C-H Bond Oxygenation. *J. Am. Chem. Soc.* **2014,** *136* (13), 4901-4904.
122. Li, S. Y.; Chaulagain, M. R.; Knauff, A. R.; Podust, L. M.; Montgomery, J.; Sherman, D. H., Selective Oxidation of Carbolide C-H Bonds by an Engineered Macrolide P450 Mono-Oxygenase. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (44), 18463-18468.
123. Narayan, A. R.; Jimenez-Oses, G.; Liu, P.; Negretti, S.; Zhao, W.; Gilbert, M. M.; Ramabhadran, R. O.; Yang, Y. F.; Furan, L. R.; Li, Z.; Podust, L. M.; Montgomery, J.; Houk, K. N.; Sherman, D. H., Enzymatic Hydroxylation of an Unactivated Methylene C-H Bond Guided by Molecular Dynamics Simulations. *Nat. Chem.* **2015,** *7* (8), 653-660.
124. Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C., Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E. coli. *Science* **2001,** *291* (5509), 1790-1792.
125. Chemler, J. A.; Tripathi, A.; Hansen, D. A.; O'Neil-Johnson, M.; Williams, R. B.; Starks, C.; Park, S. R.; Sherman, D. H., Evolution of Efficient Modular Polyketide Synthases by Homologous Recombination. *J. Am. Chem. Soc.* **2015,** *137* (33), 10603-10609.
126. DeMars, M. D.; Sheng, F.; Park, S. R.; Lowell, A. N.; Podust, L. M.; Sherman, D. H., Biochemical and Structural Characterization of MycCI, a Versatile P450 Biocatalyst from the Mycinamicin Biosynthetic Pathway. *ACS Chem. Biol.* **2016,** *11* (9), 2642-2654.
127. Omura, T.; Sato, R., The Carbon Monoxide-Binding Pigment of Liver Microsomes. I. Evidence for Its Hemoprotein Nature. *J. Biol. Chem.* **1964,** *239*, 2370-2378.
128. Saridakis, I.; Kaiser, D.; Maulide, N., Unconventional Macrocyclizations in Natural Product Synthesis. *ACS Cent. Sci.* **2020,** *6* (11), 1869-1889.
129. Yu, X.; Sun, D., Macrocyclic Drugs and Synthetic Methodologies Toward Macrocycles. *Molecules* **2013,** *18* (6), 6230-6268.
130. Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K., The Exploration of Macrocycles for Drug Discovery - An Underexploited Structural Class. *Nat. Rev. Drug Discov.* **2008,** *7* (7), 608-624.
131. Cummings, M. D.; Sekharan, S., Structure-Based Macrocycle Design in Small-Molecule Drug Discovery and Simple Metrics To Identify Opportunities for Macrocyclization of Small-Molecule Ligands. *J. Med. Chem.* **2019,** *62* (15), 6843-6853.
132. Kirst, H. A., Semi-Synthetic Derivatives of Erythromycin. *Prog. Med. Chem.* **1993,** *30*, 57-88.

133. Parnham, M. J.; Haber, V. E.; Giamarellos-Bourboulis, E. J.; Perletti, G.; Verleden, G. M.; Vos, R., Azithromycin: Mechanisms of Action and their Relevance for Clinical Applications. *Pharmacol. Therapeut.* **2014,** *143* (2), 225-245.
134. Breiner-Goldstein, E.; Eyal, Z.; Matzov, D.; Halfon, Y.; Cimicata, G.; Baum, M.; Rokney, A.; Ezernitchi, A. V.; Lowell, A. N.; Schmidt, J. J.; Rozenberg, H.; Zimmerman, E.; Bashan, A.; Valinsky, L.; Anzai, Y.; Sherman, D. H.; Yonath, A., Ribosome-Binding and Anti-Microbial Studies of the Mycinamicins, 16-Membered Macrolide Antibiotics from Micromonospora griseorubida. *Nucleic Acids Res.* **2021,** *49* (16), 9560-9573.
135. Kumar, C. N. S. S. P., Total Synthesis of Macrolides. In *Organic Synthesis - A Nascent Relook*, Nandeshwarappa, B. P., Ed. 2019.
136. Shinde, P. B.; Oh, H. S.; Choi, H.; Rathwell, K.; Ban, Y. H.; Kim, E. J.; Yang, I.; Lee, D. G.; Sherman, D. H.; Kang, H. Y.; Yoon, Y. J., Chemoenzymatic Synthesis of Glycosylated Macrolactam Analogues of the Macrolide Antibiotic YC-17. *Adv. Synth. Catal.* **2015,** *357* (12), 2697-2711.
137. Jelić, D.; Antolović, R., From Erythromycin to Azithromycin and New Potential Ribosome-Binding Antimicrobials. *Antibiotics (Basel)* **2016,** *5* (3), 29.
138. Tatsuta, K., Total Synthesis of the Big Four Antibiotics and Related Antibiotics. *J. Antibiot. (Tokyo)* **2013,** *66* (3), 107-129.
139. Hogan, P. C.; Chen, C. L.; Mulvihill, K. M.; Lawrence, J. F.; Moorhead, E.; Rickmeier, J.; Myers, A. G., Large-Scale Preparation of Key Building Blocks for the Manufacture of Fully Synthetic Macrolide Antibiotics. *J. Antibiot. (Tokyo)* **2018,** *71* (2), 318-325.
140. Marti-Centelles, V.; Pandey, M. D.; Burguete, M. I.; Luis, S. V., Macrocyclization Reactions: The Importance of Conformational, Configurational, and Template-Induced Preorganization. *Chem. Rev.* **2015,** *115* (16), 8736-834.
141. Force, G.; Perfetto, A.; Mayer, R. J.; Ciofini, I.; Leboeuf, D., Macrolactonization Reactions Driven by a Pentafluorobenzoyl Group. *Angew. Chem. Int. Ed.* **2021,** *60* (36), 19843-19851.
142. Marsault, E.; Peterson, M. L., Macrocycles are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* **2011,** *54* (7), 1961-2004.
143. Yang, M.; Wang, X. W.; Zhao, J. F., Ynamide-Mediated Macrolactonization. *ACS Catal.* **2020,** *10* (9), 5230-5235.
144. Sengupta, S.; Mehta, G., Macrocyclization via C-H Functionalization: A New Paradigm in Macrocycle Synthesis. *Org. Biomol. Chem.* **2020,** *18* (10), 1851-1876.
145. Li, J.; Amatuni, A.; Renata, H., Recent Advances in the Chemoenzymatic Synthesis of Bioactive Natural Products. *Curr. Opin. Chem. Biol.* **2020,** *55*, 111-118.
146. Stout, C. N.; Renata, H., Reinvigorating the Chiral Pool: Chemoenzymatic Approaches to Complex Peptides and Terpenoids. *Acc. Chem. Res.* **2021,** *54* (5), 1143-1156.
147. Yang, L. C.; Deng, H. P.; Renata, H., Recent Progress and Developments in Chemoenzymatic and Biocatalytic Dynamic Kinetic Resolution. *Org. Process Res. Dev.* **2022,** *26* (7), 1925-1943.
148. Vanable, E. P.; Habgood, L. G.; Patrone, J. D., Current Progress in the Chemoenzymatic Synthesis of Natural Products. *Molecules* **2022,** *27* (19), 6373.
149. Sisco, S. W.; Larson, B. M.; Moore, J. S., Relaxing Conformational Constraints in Dynamic Macrocycle Synthesis. *Macromolecules* **2014,** *47* (12), 3829-3836.

150. Giraldes, J. W.; Akey, D. L.; Kittendorf, J. D.; Sherman, D. H.; Smith, J. L.; Fecik, R. A., Structural and Mechanistic Insights into Polyketide Macrolactonization from Polyketide-Based Affinity Labels. *Nat. Chem. Biol.* **2006,** *2* (10), 531-536.
151. Adrover-Castellano, M. L.; Schmidt, J. J.; Sherman, D. H., Biosynthetic Cyclization Catalysts for the Assembly of Peptide and Polyketide Natural Products. *ChemCatChem* **2021,** *13* (9), 2095-2116.
152. Kaur, J.; Sharma, R., Directed Evolution: An Approach to Engineer Enzymes. *Crit. Rev. Biotechnol.* **2006,** *26* (3), 165-199.
153. Nannemann, D. P.; Birmingham, W. R.; Scism, R. A.; Bachmann, B. O., Assessing Directed Evolution Methods for the Generation of Biosynthetic Enzymes with Potential in Drug Biosynthesis. *Future Med. Chem.* **2011,** *3* (7), 809-819.
154. Dalby, P. A., Strategy and Success for the Directed Evolution of Enzymes. *Curr. Opin. Struct. Biol.* **2011,** *21* (4), 473-480.
155. Arnold, F. H., Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem. Int. Ed.* **2018,** *57* (16), 4143-4148.
156. Xiao, H.; Bao, Z.; Zhao, H., High Throughput Screening and Selection Methods for Directed Enzyme Evolution. *Ind. Eng. Chem. Res.* **2015,** *54* (16), 4011-4020.
157. Ye, L.; Yang, C.; Yu, H., From Molecular Engineering to Process Engineering: Development of High-Throughput Screening Methods in Enzyme Directed Evolution. *Appl. Microbiol. Biotechnol.* **2018,** *102* (2), 559-567.
158. Tiz, D. B.; Bagnoli, L.; Rosati, O.; Marini, F.; Santi, C.; Sancineto, L., FDA-Approved Small Molecules in 2022: Clinical Uses and Their Synthesis. *Pharmaceutics* **2022,** *14* (11), 2538.
159. Lubberink, M.; Finnigan, W.; Schnepel, C.; Baldwin, C. R.; Turner, N. J.; Flitsch, S. L., One-Step Biocatalytic Synthesis of Sustainable Surfactants by Selective Amide Bond Formation. *Angew. Chem. Int. Ed.* **2022,** *61* (30), e202205054.
160. Brown, D. G.; Bostrom, J., Analysis of Past and Present Synthetic Methodologies on Medicinal Chemistry: Where Have all the New Reactions Gone? *J. Med. Chem.* **2016,** *59* (10), 4443-4458.
161. Massolo, E.; Pirola, M.; Benaglia, M., Amide Bond Formation Strategies: Latest Advances on a Dateless Transformation. *Eur. J. Org. Chem.* **2020,** (30), 4641-4651.
162. Boström, J.; Brown, D. G.; Young, R. J.; Keseru, G. M., Expanding the Medicinal Chemistry Synthetic Toolbox. *Nat. Rev. Drug Discov.* **2018,** *17* (12), 709-727.
163. Muramatsu, W.; Hattori, T.; Yamamoto, H., Amide Bond Formation: Beyond the Dilemma Between Activation and Racemisation. *Chem. Commun.* **2021,** *57* (52), 6346-6359.
164. Bering, L.; Craven, E. J.; Thomas, S. A. S.; Shepherd, S. A.; Micklefield, J., Merging Enzymes with Chemocatalysis for Amide Bond Synthesis. *Nat. Commun.* **2022,** *13* (1), 380.
165. Evans, D. A.; Bartroli, J.; Shih, T. L., Enantioselective Aldol Condensations. 2. Erythro-Selective Chiral Alol Condensations via Boron Enolates. *J. Am. Chem. Soc.* **1981,** *103* (8), 2127-2129.
166. Gage, J. R.; Evans, D. A., Diastereoselective Aldol Condensation Using a Chiral Oxazolidinone Auxiliary - (2S*, 3S*)-3-Hydroxy-3-Phenyl-2-Methylpropanoic Acid. *Org. Synth.* **1990,** *68*, 83-91.
167. Willwacher, J.; Kausch-Busies, N.; Furstner, A., Divergent Total Synthesis of the Antimitotic Agent Leiodermatolide. *Angew. Chem. Int. Ed.* **2012,** *51* (48), 12041-12046.

168. Beutner, G. L.; Cohen, B. M.; DelMonte, A. J.; Dixon, D. D.; Fraunhoffer, K. J.; Glace, A. W.; Lo, E.; Stevens, J. M.; Vanyo, D.; Wilbert, C., Revisiting the Cleavage of Evans Oxazolidinones with LiOH/H2O2. *Org. Process Res. Dev.* **2019,** *23* (7), 1378-1385.
169. Brandenberg, O. F.; Chen, K.; Arnold, F. H., Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds. *J. Am. Chem. Soc.* **2019,** *141* (22), 8989-8995.
170. Romney, D. K.; Murciano-Calles, J.; Wehrmuller, J. E.; Arnold, F. H., Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. *J. Am. Chem. Soc.* **2017,** *139* (31), 10769-10776.
171. McIsaac, R. S.; Engqvist, M. K.; Wannier, T.; Rosenthal, A. Z.; Herwig, L.; Flytzanis, N. C.; Imasheva, E. S.; Lanyi, J. K.; Balashov, S. P.; Gradinaru, V.; Arnold, F. H., Directed Evolution of a Far-Red Fluorescent Rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.* **2014,** *111* (36), 13034-13039.
172. Rogers, T. A.; Bommarius, A. S., Utilizing Simple Biochemical Measurements to Predict Lifetime Output of Biocatalysts in Continuous Isothermal Processes. *Chem. Eng. Sci.* **2010,** *65* (6), 2118-2124.
173. Huynh, K.; Partch, C. L., Analysis of Protein Stability and Ligand Interactions by Thermal Shift Assay. *Curr. Protoc. Protein Sci.* **2015,** *79*, 28.9.1-28.9.14.
174. Gomes, M. D.; Woodley, J. M., Considerations when Measuring Biocatalyst Performance. *Molecules* **2019,** *24* (19), 3573.
175. Bommarius, A. S., Total Turnover Number - A Key Criterion for Process Evaluation. *Chem. Ing. Tech.* **2023,** *95* (4), 491-497.
176. Payne, J. T.; Poor, C. B.; Lewis, J. C., Directed Evolution of RebH for Site-Selective Halogenation of Large Biologically Active Molecules. *Angew. Chem. Int. Ed.* **2015,** *54* (14), 4226-4230.
177. Mendonca, L. M.; Marana, S. R., Single Mutations Outside the Active Site Affect the Substrate Specificity in a β-Glycosidase. *Biochim. Biophys. Acta* **2011,** *1814* (12), 1616-1623.
178. Rix, G.; Watkins-Dulaney, E. J.; Almhjell, P. J.; Boville, C. E.; Arnold, F. H.; Liu, C. C., Scalable Continuous Evolution for the Generation of Diverse Enzyme Variants Encompassing Promiscuous Activities. *Nat. Commun.* **2020,** *11* (1), 5644.
179. Liu, L.; Tao, W. T.; Bai, L. Q.; Kim, E. S.; Zhao, Y. L.; Shi, T., Why Does Tautomycetin Thioesterase Prefer Hydrolysis to Macrocyclization? Theoretical Study on its Catalytic Mechanism. *Catal. Sci. Technol.* **2019,** *9* (22), 6391-6403.
180. Goldsmith, M.; Aggarwal, N.; Ashani, Y.; Jubran, H.; Greisen, P. J.; Ovchinnikov, S.; Leader, H.; Baker, D.; Sussman, J. L.; Goldenzweig, A.; Fleishman, S. J.; Tawfik, D. S., Overcoming an Optimization Plateau in the Directed Evolution of Highly Efficient Nerve Agent Bioscavengers. *Protein Eng. Des. Sel.* **2017,** *30* (4), 333-345.
181. Baier, F.; Hong, N.; Yang, G.; Pabis, A.; Miton, C. M.; Barrozo, A.; Carr, P. D.; Kamerlin, S. C. L.; Jackson, C. J.; Tokuriki, N., Cryptic Genetic Variation Shapes the Adaptive Evolutionary Potential of Enzymes. *Elife* **2019,** *8*, e40789.
182. Li, H.; Bao, Q. Q.; Zhao, J. F.; Xu, Y. B.; Yang, S. Y.; Xue, W. S.; Sun, Y.; Liu, Y. P., Directed Evolution Engineering to Improve Activity of Glucose Dehydrogenase by Increasing Pocket Hydrophobicity. *Front. Microbiol.* **2022,** *13*, 1044226.
183. Tokuriki, N.; Tawfik, D. S., Stability Effects of Mutations and Protein Evolvability. *Curr. Opin. Struct. Biol.* **2009,** *19* (5), 596-604.

184. Studer, R. A.; Christin, P. A.; Williams, M. A.; Orengo, C. A., Stability-Activity Tradeoffs Constrain the Adaptive Evolution of RubisCO. *Proc. Natl. Acad. Sci. U.S.A.* **2014,** *111* (6), 2223-2228.
185. Romero, P. A.; Arnold, F. H., Exploring Protein Fitness Landscapes by Directed Evolution. *Nat. Rev. Mol. Cell. Biol.* **2009,** *10* (12), 866-876.
186. Yang, G. R.; Miton, C. M.; Tokuriki, N., A Mechanistic View of Enzyme Evolution. *Protein Sci.* **2020,** *29* (8), 1724-1747.
187. Pinto, G. P.; Corbella, M.; Demkiv, A. O.; Kamerlin, S. C. L., Exploiting Enzyme Evolution for Computational Protein Design. *Trends Biochem. Sci.* **2022,** *47* (5), 375-389.
188. Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R.; Baker, D.; Griffiths, A. D.; Hilvert, D., Emergence of a Catalytic Tetrad During Evolution of a Highly Active Artificial Aldolase. *Nat. Chem.* **2017,** *9* (1), 50-56.
189. Stimple, S. D.; Smith, M. D.; Tessier, P. M., Directed Evolution Methods for Overcoming Trade-Offs Between Protein Activity and Stability. *AIChE J.* **2020,** *66* (3), e16814.
190. Yang, T.; Ye, Z. X.; Lynch, M. D., "Multiagent" Screening Improves Directed Enzyme Evolution by Identifying Epistatic Mutations. *ACS Synth. Biol.* **2022,** *11* (5), 1971-1983.
191. Peracchi, A., Enzyme Catalysis: Removing Chemically 'Essential' Residues by Site-Directed Mutagenesis. *Trends Biochem. Sci.* **2001,** *26* (8), 497-503.
192. Chen, X. P.; Shi, T.; Wang, X. L.; Wang, J. T.; Chen, Q. H.; Bai, L. Q.; Zhao, Y. L., Theoretical Studies on the Mechanism of Thioesterase-Catalyzed Macrocyclization in Erythromycin Biosynthesis. *ACS Catal.* **2016,** *6* (7), 4369-4378.
193. Kokkonen, P.; Slanska, M.; Dockalova, V.; Pinto, G. P.; Sanchez-Carnerero, E. M.; Damborsky, J.; Klan, P.; Prokop, Z.; Bednar, D., The Impact of Tunnel Mutations on Enzymatic Catalysis Depends on the Tunnel-Substrate Complementarity and the Rate-Limiting Step. *Comput. Struct. Biotec.* **2020,** *18*, 805-813.
194. Schafer, J. W.; Zoi, I.; Antoniou, D.; Schwartz, S. D., Optimization of the Turnover in Artificial Enzymes via Directed Evolution Results in the Coupling of Protein Dynamics to Chemistry. *J. Am. Chem. Soc.* **2019,** *141* (26), 10431-10439.
195. Patel, K. D.; d'Andrea, F. B.; Gaudelli, N. M.; Buller, A. R.; Townsend, C. A.; Gulick, A. M., Structure of a Bound Peptide Phosphonate Reveals the Mechanism of Nocardicin Bifunctional Thioesterase Epimerase-Hydrolase Half-Reactions. *Nat. Commun.* **2019,** *10*, 3868.
196. Shi, T.; Liu, L. X.; Tao, W. T.; Luo, S. G.; Fan, S. B.; Wang, X. L.; Bai, L. Q.; Zhao, Y. L., Theoretical Studies on the Catalytic Mechanism and Substrate Diversity for Macrocyclization of Pikromycin Thioesterase. *ACS Catal.* **2018,** *8* (5), 4323-4332.
197. Zhou, Y. C.; Tao, W. T.; Qi, Z.; Wei, J. H.; Shi, T.; Kang, Q. J.; Zheng, J. T.; Zhao, Y. L.; Bai, L. Q., Structural and Mechanistic Insights into Chain Release of the Polyene PKS Thioesterase Domain. *ACS Catal.* **2022,** *12* (1), 762-776.
198. Kelly, S. P.; Shende, V. V.; Flynn, A. R.; Dan, Q.; Ye, Y.; Smith, J. L.; Tsukamoto, S.; Sigman, M. S.; Sherman, D. H., Data Science-Driven Analysis of Substrate-Permissive Diketopiperazine Reverse Prenyltransferase NotF: Applications in Protein Engineering and Cascade Biocatalytic Synthesis of (-)-Eurotiumin A. *J. Am. Chem. Soc.* **2022,** *144* (42), 19326-19336.
199. Yang, S.; DeMars, M. D.; Grandner, J. M.; Olson, N. M.; Anzai, Y.; Sherman, D. H.; Houk, K. N., Computational-Based Mechanistic Study and Engineering of Cytochrome P450 MycG for Selective Oxidation of 16-Membered Macrolide Antibiotics. *J. Am. Chem. Soc.* **2020,** *142* (42), 17981-17988.

200. Shende, V. V.; Harris, N. R.; Sanders, J. N.; Newmister, S. A.; Khatri, Y.; Movassaghi, M.; Houk, K. N.; Sherman, D. H., Molecular Dynamics Simulations Guide Chimeragenesis and Engineered Control of Chemoselectivity in Diketopiperazine Dimerases. *Angew. Chem. Int. Ed.* **2023,** *62*, e2022102.
201. Clements, H. D.; Flynn, A. R.; Nicholls, B. T.; Grosheva, D.; Lefave, S. J.; Merriman, M. T.; Hyster, T. K.; Sigman, M. S., Using Data Science for Mechanistic Insights and Selectivity Predictions in a Non-Natural Biocatalytic Reaction. *J. Am. Chem. Soc.* **2023,** *145* (32), 17656-17664.
202. Andorfer, M. C.; Evans, D.; Yang, S.; He, C. Q.; Girlich, A. M.; Vergara-Coll, J.; Sukumar, N.; Houk, K. N.; Lewis, J. C., Analysis of Laboratory-Evolved Flavin-Dependent Halogenases Affords a Computational Model for Predicting Halogenase Site Selectivity. *Chem. Catalysis* **2022,** *2* (10), 2658-2674.
203. Wittmann, B. J.; Johnston, K. E.; Wu, Z.; Arnold, F. H., Advances in Machine Learning for Directed Evolution. *Curr. Opin. Struct. Biol.* **2021,** *69*, 11-18.
204. Lühr, S.; Holz, J.; Zayas, O.; Seidelmann, O.; Domke, L.; Borner, A., Synthesis of Enantiopure β2-Homoalanine Derivatives via Rhodium Catalyzed Asymmetric Hydrogenation. *Tetrahedron Asymmetry* **2013,** *24* (7), 395-401.
205. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z. G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P., Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol.* **2013,** *2* (2), 83-92.
206. Sheldrick, G. M., SHELXT - Integrated Space-Group and Crystal-Structure Determination. *Acta Crystallogr. A* **2015,** *A71*, 3-8.
207. Dolomanov, O. V.; Bourhis, L. J.; Gildea, R. J.; Howard, J. A. K.; Puschmann, H., OLEX2: A Complete Structure Solution, Refinement and Analysis Program. *J. Appl. Crystallogr.* **2009,** *42*, 339-341.
208. Sheldrick, G. M., Crystal Structure Refinement with SHELXL. *Acta Crystallogr. A* **2015,** *C71*, 3-8.
209. Rigaku, O. D., *CrysAlisPro Software System, Rigaku Oxford Diffraction*. Yarnton, UK: 2023.
210. Kalkreuter, E.; CroweTipton, J. M.; Lowell, A. N.; Sherman, D. H.; Williams, G. J., Engineering the Substrate Specificity of a Modular Polyketide Synthase for Installation of Consecutive Non-Natural Extender Units. *J. Am. Chem. Soc.* **2019,** *141* (5), 1961-1969.