

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Genome Mining, Biosynthesis and Catalysis of a Novel Fungal Polyketide Synthase-Carnitine Acyltransferase Fusion Enzyme

**Permalink**

<https://escholarship.org/uc/item/6n99p1qs>

**Author**

Hang, Leibniz

**Publication Date**

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Genome Mining, Biosynthesis and Catalysis of a Novel Fungal Polyketide Synthase-Carnitine  
Acyltransferase Fusion Enzyme

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy in Chemistry

by

Leibniz Hang

2019

© Copyright by

Leibniz Hang

2019

## ABSTRACT OF THE DISSERTATION

Genome Mining, Biosynthesis and Catalysis of a Novel Fungal Polyketide Synthase-Carnitine

Acyltransferase Fusion Enzyme

By

Leibniz Hang

Doctor of Philosophy in Chemistry

University of California, Los Angeles, 2019

Professor Yi Tang, Chair

This dissertation describes the genome mining, biosynthesis, and catalysis of a novel fungal polyketide synthase-carnitine acyltransferase fusion enzyme. Fungal highly-reducing polyketide synthases (HRPKS) catalyze the biosynthesis of polyketide natural products such as the cholesterol-lowering drug, lovastatin. Investigation into unexplored families of HRPKS enzymes such as the HRPKS-cAT fusion enzymes provides opportunities to discover new drug leads. This research will also improve our understanding of HRPKS enzymatic catalysis for future engineering purposes.

Chapter 1 introduces the function, iterative catalysis, and biosynthetic framework of the fungal HRPKS enzymes. Partnering oxygenation enzymes which functionalize the polyketide scaffold are also assessed to emphasize the meticulous synchronization involved in HRPKS

biosynthesis. Several examples of HRPKS biosynthesis are provided to illustrate the coordination required for the formation of the final natural product.

Chapter 2 examines the genome mining of HRPKS-cAT enzymes and highlights their differences from canonical HRPKS catalysis. The HRPKS-cAT enzyme, Tv6-931, is reconstituted and the biosynthesis of polyketide-polyol conjugates bearing a rare *gem*-dimethyl moiety is discussed. Based on enzymatic assays using the Tv6-931 whole enzyme and dissected enzyme variants, an unusual polyketide “recapture” (oxyester-thioester transacylation) is proposed to explain the formation of the *gem*-dimethyl moiety.

Chapter 3 investigates the biosynthetic, polyketide “recapture” mechanism introduced in Chapter 2. Enzyme kinetic studies show that the cofactor-independent, polyketide recapture is necessary for the *gem*-dimethyl adduct to be the major product. Based on the oxyester-thioester transacylation mechanism (recapture), a chemoenzymatic strategy was developed to conjugate other nucleophiles to create a variety of polyketide oxyester, thioester and amide adducts. A one-pot, multi-enzyme approach was used to demonstrate polyketide functionalization using HRPKS-cAT enzymes for proof-of-principle.

This dissertation of Leibniz Hang is approved.

Steven G. Clarke

Neil Kamal Garg

Kendall N. Houk

Yi Tang, Committee Chair

University of California, Los Angeles

2019

## DEDICATION

This dissertation is dedicated to my family: Tracy Hang (wife), Tung Hang (father), Ha Hang (mother), Erin Hang (sister), Newton Hang (brother), and Judy Hang (sister-in-law). This dissertation would not have been possible without your support.

## TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION .....	ii-iii
COMMITTEE PAGE .....	iv
DEDICATION PAGE .....	v
TABLE OF CONTENTS.....	vi-vii
LIST OF FIGURES .....	viii-x
LIST OF TABLES .....	xi
LIST OF SCHEMES.....	xii
LIST OF ABBREVIATIONS.....	xii-xiii
ACKNOWLEDGEMENTS.....	xiv-xviii
BIOGRAPHICAL SKETCH .....	xix-xx
<b>Chapter 1</b> Introduction to Iterative Enzyme Catalysis in Fungal Polyketide Biosynthesis .....	1-33
Section 1.1    Iterative Enzyme Catalysis in Polyketide Biosynthesis.....	2-3
Section 1.2    Overview of Type I Iterative Fungal Highly-Reducing Polyketide Synthase (HRPKS) .....	3-4
Section 1.3    Overview of Post-PKS Iterative Oxygenases .....	5-8
Section 1.4    Lovastatin.....	8-12
Section 1.5    Chaetoglobosin A.....	12-16
Section 1.6    Cytochalasin E and Cytochalasin K.....	16-21
Section 1.7    Aurovertin E.....	21-24
Section 1.8    Conclusion .....	25
Section 1.9    References.....	26-33
<b>Chapter 2</b> Reversible Product Release and Recapture by a Fungal Polyketide Synthase Using a Carnitine Acyltransferase Domain.....	34-91
Section 2.1    Genome Mining of HRPKS-cAT Enzymes.....	35-37
Section 2.2    Heterologous Expression of Tv6-931 HRPKS-cAT.....	38-39
Section 2.3    Enzymatic Biosynthesis of Tv6-931 HRPKS-cAT.....	40-42



Section 2.4	Enzymatic Assays using <i>N</i> -acetylcysteamine Thioester Probes .....	42-44
Section 2.5	Polyketide Recapture and Post-Release Catalysis .....	45-47
Section 2.6	Supplementary Tables.....	48-56
Section 2.7	Supplementary Figures .....	57-70
Section 2.8	Experimental Methods .....	71-83
Section 2.9	References.....	84-91
<b>Chapter 3</b>	<b>Mechanistic Investigation of the Cofactor-Independent, Reversible Transacylation of Polyketide Oxyesters to Pantetheine Thioesters Catalyzed by a Carnitine Acyltransferase-Type Domain.....</b>	<b>92-191</b>
Section 3.1	Proposed Thioester Formation Catalyzed by a cAT-Type Domain in Polyketide Biosynthesis .....	93-96
Section 3.2	The Kinetic Competition Model of HRPKS and Potential Application to Tv6-931 Biosynthesis .....	96-98
Section 3.3	Kinetic Competition Between MT (Methylation) and cAT Domain (Esterification) .....	98-103
Section 3.4	Evaluation of cAT-catalyzed Oxyester-Thioester Transacylation (Polyketide “Recapture”) .....	103-105
Section 3.5	Chemoenzymatic Functionalization of Polyketides Using Tv6-931 Enzymes .....	105-111
Section 3.6	Coupling Tv6-931 HRPKS-cAT with a Downstream HRPKS for Polyketide Functionalization.....	111-113
Section 3.7	Discussion.....	114-116
Section 3.8	Supplementary Schemes .....	117-122
Section 3.9	Supplementary Tables.....	123-130
Section 3.10	Supplementary Figures .....	131-149
Section 3.11	Experimental Methods .....	150-182
Section 3.12	References.....	183-191

## LIST OF FIGURES

### Chapter 1

Figure 1.1	The domain architecture and the catalysis of the minimum fungal PKS domains and the fungal PKS tailoring domains in fungal polyketide synthases.....	5
Figure 1.2	The mechanism of P450 catalyzed, radical-mediated hydroxylation of a substrate .....	7
Figure 1.3	The mechanism of FMO mediated nucleophilic oxygenation .....	8
Figure 1.4	Biosynthesis of lovastatin .....	11
Figure 1.5	Biosynthesis of chaetoglobosin A.....	15
Figure 1.6	Biosynthesis of cytochalasin E and cytochalasin K.....	18
Figure 1.7	Proposed mechanism of the CcsB catalyzed conversion of ketocytochalasin to cytochalasin Z <sub>16</sub> .....	21
Figure 1.8	Proposed biosynthesis of aurovertin E.....	23

### Chapter 2

Figure 2.1	HRPKSs containing a C-terminal carnitine <i>O</i> -acyltransferase like domain.....	37
Figure 2.2	Functional characterization of Tv6-931 .....	39
Figure 2.3	Characterization of MT and cAT domains of Tv6-931 using SNAC substrates ...	43
Figure 2.4	Reversible product offloading by the cAT domain in Tv6-931.....	8
Figure S2.1	Fungal Phylogenetic Trees.....	57
Figure S2.2	SDS-PAGE Analysis of Purified Enzyme .....	59
Figure S2.3	List of Tv6-931 Offloading Substrates (Acyl-acceptors) .....	60
Figure S2.4	Enzymatic <sup>13</sup> C Isotope Labeling .....	61
Figure S2.5	HPLC Traces of Polyketide Products ( $\lambda = 236$ nm) .....	62
Figure S2.6	Enzymatic <sup>13</sup> C Isotope Labeling .....	63
Figure S2.7	Methylation Assays on Triketide SNACs.....	64
Figure S2.8	Titration of cAT in Enzymatic Synthesis.....	65
Figure S2.9	Sequence Alignment of the Tv6-931 cAT Domain and its Homologs.....	66
Figure S2.10	Conserved Residues between Authentic cAT and the cAT Domain .....	67

Figure S2.11	Comparison of the <sup>1</sup> H NMR spectra of compound <b>2</b> purified from the <i>in situ</i> assays and the feeding experiments.....	68
Figure S2.12	Verification of MT and cAT Activity of <i>apo</i> -Tv6-931 Enzyme .....	69
Figure S2.13	Abolishment of Biosynthesis in <i>apo</i> -Tv6-931 Enzyme.....	69
Figure S2.14	Methylation and Transacylation Assays with Tetraketide SNACs.....	70

### Chapter 3

Figure 3.1	Enzymatic synthesis of biological thioesters and their proposed relationship with polyketide biosynthesis.....	96
Figure 3.2	Kinetic experiments using SNAC probes with MT and cAT Domains.....	101
Figure 3.3	Kinetic transesterification experiments to estimate polyketide recapture .....	101
Figure S3.1	Mechanistic Comparison of the PKS TE Domain and Native cAT Enzyme .....	131
Figure S3.2	Fungal KS Phylogenetic Tree with Highlighted HRPKS-cAT Fusion Enzymes .....	132
Figure S3.3	Kinetic Competition Model of HRPKS .....	133
Figure S3.4	Determination of the Stereochemistry of the Distal, ε-Methyl Group.....	134
Figure S3.5	Heterologous Expression and Purification of Tv6-931 and LovB Enzyme Variants .....	135
Figure S3.6	Enzymatic Methylation Kinetic Assay .....	136
Figure S3.7	Enzymatic Release (Esterification) Kinetic Assay .....	137
Figure S3.8	Enzymatic Recapture (Oxyester-Thioester Transacylation) Kinetic Assay .....	138
Figure S3.9	Enzymatic Recapture (Transesterification) Kinetic Assay .....	140
Figure S3.10	Chemoenzymatic Functionalization of <b>1-PE</b> Using <i>holo</i> -ACP-cAT.....	141
Figure S3.11	Chemoenzymatic Functionalization of <b>2-PE</b> Using <i>holo</i> -ACP-cAT.....	142
Figure S3.12	Enzymatic Functionalization of <b>1-PE</b> and <b>2-PE</b> Using <i>apo</i> -ACP-cAT.....	143
Figure S3.13	Supplementation of Pantetheine Restores <i>apo</i> -ACP-cAT Transesterification Activity .....	144
Figure S3.14	Supplementation of Pantetheine Restores <i>apo</i> -ACP-cAT Transthoesterification Activity .....	145
Figure S3.15	Enzymatic Functionalization of <b>0-PE</b> Using <i>holo</i> -ACP.....	146

Figure S3.16	Coupling ACP-cAT with LovB enzyme with <b>1-PE</b> for chemoenzymatic synthesis of <b>1-Pante</b> , <b>3</b> , and <b>4</b> .....	147
Figure S3.17	Conversion of <b>1-Pante</b> by LovB enzyme for chemoenzymatic synthesis of <b>3</b> , and <b>4</b> .....	148
Figure S3.18	Coupling ACP-cAT with LovB enzyme with <b>2-PE</b> for chemoenzymatic synthesis of <b>2-Pante</b> , and <b>5</b> .....	149

## LIST OF TABLES

### Chapter 2

Table S2.1	DNA Primers Used .....	48
Table S2.2	DNA Plasmids Used .....	49
Table S2.3	Expression Strains Used .....	49
Table S2.4	NMR Data of Compound <b>1</b> .....	50
Table S2.5	NMR Data of Compound <b>2</b> .....	51
Table S2.6	NMR Data of Compound <b>5</b> .....	52
Table S2.7	NMR Data of Compound <b>10</b> .....	53
Table S2.8	HPLC-MS Mass Data of Compounds.....	54
Table S2.9	HR-MS Mass Data of Compounds .....	56

### Chapter 3

Table 3.1	Transacylation of <b>1-PE</b> by various nucleophiles .....	108
Table S3.1	List of SNAC Compounds .....	123
Table S3.2	DNA Primers Used .....	124
Table S3.3	DNA Plasmids Used .....	125
Table S3.4	Expression Strains Used .....	125
Table S3.5	Transacylation of <b>2-PE</b> by various nucleophiles .....	133
Table S3.6	HPLC-MS Mass Data of Compounds.....	127
Table S3.7	HR-MS Mass Data of Compounds .....	130

## LIST OF SCHEMES

### Chapter 3

Scheme 3.1	ACP-cAT-catalyzed transacylation of <b>1-PE</b> .....	106
Scheme 3.2	Coupling the ACP-cAT enzyme with another functionalization enzyme .....	112
Scheme S3.1	Synthetic scheme for the preparation of ( <i>R</i> )-2-methylbutanol .....	117
Scheme S3.2	Synthetic scheme for the preparation of <b>Tri<sub>0</sub>-SNAC</b> and <b>Tri<sub>1</sub>-SNAC</b> .....	117
Scheme S3.3	Synthetic scheme for the preparation of ( <i>R</i> )-2-methylbutanal.....	118
Scheme S3.4	Synthetic scheme for the preparation of <b>S-Tri<sub>0</sub>-SNAC</b> and <b>S-Tri<sub>1</sub>-SNAC</b> .....	118
Scheme S3.5	Synthetic scheme for the preparation of <b>0-SNAC</b> , <b>1-SNAC</b> and <b>2-SNAC</b> .....	119
Scheme S3.6	Synthetic scheme for the preparation of <b>S-0-SNAC</b> , <b>S-1-SNAC</b> and <b>S-2-SNAC</b> .....	120
Scheme S3.7	Mechanism of ACP-cAT Chemoenzymatic Transacylation Using <b>1-PE</b> .....	121
Scheme S3.8	Mechanism of ACP-cAT Chemoenzymatic Transacylation Using <b>2-PE</b> .....	122
Scheme S3.9	Enzymatic synthesis of <b>1-Pante</b> catalyzed by Tv6-931 <i>apo</i> -ACP-cAT enzyme .....	122

## LIST OF ABBREVIATIONS

HRPKS	highly-reducing polyketide synthase
cAT	carnitine acyltransferase; carnitine acyltransferase-type
PKS	polyketide synthase
FAS	fatty acid synthase
CoA	Coenzyme A
KS	ketosynthase
MAT	malonyl-CoA : ACP transacylase
ACP	acyl-carrier protein
MT	methyltransferase
KR	ketoreductase
DH	dehydratase
ER	enoylreductase
SAM	<i>S</i> -adenosylmethionine
NADPH	nicotinamide adenine dinucleotide phosphate hydride
P450	cytochrome P450 monooxygenase
CPR	cytochrome P450 reductase
FMO	flavin-containing monooxygenase
WT	wild-type
R	general alkyl group
THME	1,1,1-tris(hydroxymethyl)ethane (THME)
PE	pentaerythritol
Tris	tris(hydroxymethyl)aminomethane
<i>gem</i>	geminal
SNAC	<i>N</i> -acetyl cysteamine thioester

mal-CoA	malonyl-Coenzyme A
TCA	tricyclic acid
NAD <sup>+</sup>	nicotinamide adenine dinucleotide
ATP	adenosine triphosphate
NCL	native chemical ligation
LCMS	liquid chromatography - mass spectrometry
HPLC	high performance liquid chromatography
HR-MS	high resolution – mass spectrometry
MW	molecular weight
kDa	kilodaltons
ACN	acetonitrile
EtOAc	ethyl acetate
MeOH	methanol
NMR	nuclear magnetic resonance
RT	room temperature
UCLA	University of California, Los Angeles



## ACKNOWLEDGMENTS

When I first arrived at UCLA in July of 2014, I was both nervous, yet excited about the challenges that lied ahead. UCLA represented the continuation of my education and career. Paradoxically, it also signified a fresh start and a new opportunity to define myself as a scientist and an individual. Throughout this five-year journey, I have had varying levels of successes and struggles. Fortunately, I shared this road with many generous people who helped me along the way. My time at UCLA greatly shaped who I am today, and I owe my accomplishments and perseverance to the support of my mentors, colleagues, friends and family.

First, I would like to acknowledge my advisor, Professor Yi Tang, for his guidance, support and patience during my graduate studies. I doubt I will ever have another mentor-mentee relationship like I have with 老唐 (老大). Professor Tang helped reinforce the value of working hard, while also working smart. He also showed me what it takes to succeed in academia. Professor Tang possesses a peerless competitive drive in many things: academics, science, basketball, tennis, weightlifting, etc. He channels this competitiveness to intellectually, mentally and vigorously show students who is the Boss. I hope to emulate his intensity in the next stages of my career. On a personally level, few professors would embrace the playful personalities of myself and my colleagues like Professor Tang does.

Second, I would like to acknowledge to my committee members: Professor Kendall Houk, Professor Neil Garg, and Professor Steven G. Clarke. I have been very fortunate to have a great committee who has been invested in my education every step of the way. Thank you for taking the time to also serve on my oral committee and be my advisor on the NIH-CMB training

fellowship, respectively. I am indebted to each of you for your continued support even after I graduate from UCLA.

The Tang laboratory environment played an instrumental role in my growth both academically and personally. I am a proud member of a cohort of graduate students known as Fungal Journal Club (FJC). The other members include Dr. John Billingsley, Dr. Yan Yan and (soon Dr.) Nicholas Liu. I will always appreciate how we complement and support each other. John, we were always considered the opposite sides of the same coin. I learned a lot from you about science, politics and lifestyle. We always end up having friendly, philosophical debates. Nick, you were always considered the same side of the same coin. Regardless of the score, we will always have the Faculty Center—one of the greatest moments of our PhD, FJC, and possibly the Tang lab. Yan, you are a man of few words, but they always carry a lot of weight. FJC always had you to show us how to excel academically to compete with our friendly rivals, Yeast Book Club [and the post-docs].

To all of the Tang lab students that came before FJC, Dr. Ralph Cacho, Dr. Anthony DeNicola, Dr. Carly Bond, I appreciate your efforts to positively shape the graduate student culture. To all the student that came after FJC, Sunny Hung, Undramaa Bat-Erdene, Danielle Yee, Eun Bin Go, Josh Misa, Ike Okorafor, and Cooper Jamieson, your youth and positivity grew on me. You guys kept me young, and the lab is in good hands.

Several other members also played a vital role during my time in the Tang lab. Dr. Man-cheng Tang (TMC), you were instrumental to my growth. You were always very kind and funny, and I could always count on your advice and knowledge. Dr. Masao Ohashi and Dr. Zhang Zhuan, you two were great neighbors in the office and lab. Thank you for always taking the time

to listen to me and offer me sound advice. Dr. Yang Hai and Dr. Mengbin Chen, I appreciate you two always allowing me to bounce ideas off of you, and giving me insight in science and graduate school. I have learned a lot from the both of you.

I was very fortunate to also mentor talented undergraduate students who played critical roles in this dissertation. Claire Gwyneth Page, you provided a vital role in the success of my projects and I often leaned on you. Your contribution to my team will not be forgotten. Jessica Zhao, you came at a critical time in my PhD and your work was superb. I was lucky to have you join my team. I wish you both the best on your own PhD journey at Princeton University. Ariana Tamura, I recognize the effort that you put forth, and you have grown a lot since you first started.

Outside the lab, I found a group of special and loyal friends who were always there for me: Dr. Tim Chung, Dr. Janice Lin, Dr. Terri Lin, and Dr. Jonathan Lau. Tim, we first made it through UCSD and now UCLA together. Our food runs and tennis sessions are some of my best memories at UCLA. It was the closest thing we had to a work-life balance and normalcy. Janice, you were my first friend at UCLA, and the one with the strongest resolve. You were fiercely loyal to us and your beliefs, and I respect you for that. Terri, you are one of the most genuine people that I know, and the type of person we can always depend on. Jon, you were one of the most level-headed people that I met in graduate school. Even if you never complained, you always were willing to listen to the rest of us.

Most importantly, I want to thank my family: Tracy Hang (wife), Tung Hang (father), Ha Hang (mother), Erin Hang (sister), Newton Hang (brother), and Judy Hang (sister-in-law) for their continued support throughout my academic career.

Tracy, thank you for being there with me throughout this long road journey. I appreciate all the sacrifices that you made for me to be able to earn my PhD; it certainly was not easy. You dealt with your long commutes, my long lab hours, and even the wrong graduation gown rental, but you were always there to support me. Throughout it all, you always believed in me and that often made the difference. We finally made it!

Dad and Mom, you always pushed me to do my best and to work hard. It has now paid off. Even when things were hard, you prioritized my education and tried to help me push through it. I owe much of my success to the values that you instilled in me.

Erin, I am happy that you followed my footsteps at UCSD and maybe even one day at UCLA. I will always try to inspire you and be a good example. If someone like me can preserve through this challenge, I know you can do anything you put your mind to!

Newton and Judy, you two were always there for me. Despite my inconsistent schedule, you two always tried to make it work. You both took on more responsibilities so I could spend more time on my studies, and I will always be appreciative of your support. Newton, I always looked up to you, and I drew a lot of inspiration from you in school and academics.

I would also like to acknowledge the following funding agencies for their support: NIH-Cellular and Molecular Biology (CMB) Training Fellowship (T32GM007185), and UCLA Dissertation Year Fellowship.

Chapter 1 is an adapted version of the following manuscript reprinted with permission:  
**Hang, L.;** Liu, N.; Tang, Y., Coordinated and Iterative Enzyme Catalysis in Fungal Polyketide Biosynthesis. *Acs Catal.* **2016**, *6*, 5935-5945.

(doi/10.1021/acscatal.6b01559)

Copyright 2019 American Chemical Society

Chapter 2 is an adapted version of the following manuscript reprinted with permission:  
**Hang, L.;** Tang, M. C.; Harvey, C. J. B.; Page, C. G.; Li, J.; Hung, Y. S.; Liu, N.; Hillenmeyer, M. E.; Tang, Y., Reversible Product Release and Recapture by a Fungal Polyketide Synthase Using a Carnitine Acyltransferase Domain. *Angew. Chem. Int. Ed.* **2017**, *56* (32), 9556-9560.

(doi/10.1002/anie.201705237)

Copyright 2019 Wiley Company

Chapter 3 is an adapted version of the following *manuscript in preparation*:  
**Hang, L.;** Foerster, C.; Page, C.; Sanichar, R.; Zhao, Z-Y., Larade, S.; Antwi, I.; Vederas; Tang, Y., Mechanistic Investigation of the Cofactor-Independent, Reversible Transacylation of Polyketide Oxyesters to Pantetheine Thioesters Catalyzed by a Carnitine Acyltransferase-Type Domain. *Manuscript in Preparation*.

## BIOGRAPHICAL SKETCH

### EDUCATION:

**M.S. Organic Chemistry** December 2012  
University of California, San Diego

**B.S. Biochemistry/Cell Biology** March 2010  
University of California, San Diego  
Graduated with *Summa Cum Laude* Honors

### RESEARCH AND PROFESSIONAL EXPERIENCE:

**UCLA Department of Chemistry and Biochemistry** July 2014 – September 2019  
PhD Candidate *Advisor: Prof. Yi Tang*  
*Dissertation:* Genome Mining, Biosynthesis and Catalysis of Novel Fungal Polyketide Synthase-Carnitine Acyltransferase Fusion Enzymes

**UCSD Department of Biochemistry/Chemistry** January 2013 – June 2013  
Research Associate *Advisor: Prof. Jerry Yang*  
*Research Project:* Development of an Ion Channel-based Nano-sensors for Detection of Biomarkers  
Graduate Student Researcher  
*Thesis:* Development of a Polylactic Acid (PLA) Polymer with an Acid-Sensitive *N*-ethoxybenzylimidazole (NEBI) Crosslinker as a Drug Delivery System

**Bachem Americas, Inc** August 2013 – June 2014  
Synthetic Chemist  
*Production Project:* cGMP validation and synthesis of various APIs including medicinal oligopeptides, oligonucleotides derivatives and small molecules

### PUBLICATIONS:

- Li, L.; Tang, M. C.; Tang, S.; Gao, S.; Soliman, S.; **Hang, L.**; Xu, W.; Ye, T.; Watanabe, K.; Tang, Y., Genome Mining and Assembly-Line Biosynthesis of the UCS1025A Pyrrolizidinone Family of Fungal Alkaloids. *J. Am. Chem. Soc.* **2018**, *140* (6), 2067-2071.
- Kim, Y. H.; **Hang, L.**; Cifelli, J. L.; Sept, D.; Mayer, M.; Yang, J., Frequency-Based Analysis of Gramicidin A Nanopores Enabling Detection of Small Molecules with Picomolar Sensitivity. *Anal. Chem.* **2018**, *90* (3), 1635-1642.

6. **Hang, L.**; Tang, M. C.; Harvey, C. J. B.; Page, C. G.; Li, J.; Hung, Y. S.; Liu, N.; Hillenmeyer, M. E.; Tang, Y., Reversible Product Release and Recapture by a Fungal Polyketide Synthase Using a Carnitine Acyltransferase Domain. *Angew. Chem. Int. Ed.* **2017**, *56* (32), 9556-9560.
5. Liu, N.; Hung, Y. S.; Gao, S. S.; **Hang, L.**; Zou, Y.; Chooi, Y. H.; Tang, Y., Identification and Heterologous Production of a Benzoyl-Primed Tricarboxylic Acid Polyketide Intermediate from the Zaragozic Acid A Biosynthetic Pathway. *Org. Lett.* **2017**, *19* (13), 3560-3563.
4. Sato, M.; Dander, J. E.; Sato, C.; Hung, Y. S.; Gao, S. S.; Tang, M. C.; **Hang, L.**; Winter, J. M.; Garg, N. K.; Watanabe, K.; Tang, Y., Collaborative Biosynthesis of Maleimide- and Succinimide-Containing Natural Products by Fungal Polyketide Megasyntases. *J. Am. Chem. Soc.* **2017**, *139* (15), 5317-5320
3. **Hang, L.**; Liu, N.; Tang, Y., Coordinated and Iterative Enzyme Catalysis in Fungal Polyketide Biosynthesis. *Acs Catal.* **2016**, *6*, 5935-5945.
2. Yu, X.; Liu, F.; Zou, Y.; Tang, M. C.; **Hang, L.**; Houk, K. N.; Tang, Y., Biosynthesis of Strained Piperazine Alkaloids: Uncovering the Concise Pathway of Herquiline A. *J. Am. Chem. Soc.* **2016**, *138* (41), 13529-13532.
1. Gao, S. S.; Duan, A. B.; Xu, W.; Yu, P. Y.; **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*, 4249-4259.

## HONOR AND AWARDS

UCLA Thomas L. and Ruth F. Jacobs Award for Organic Graduates	2019
UCLA Dissertation Year Fellowship	2018 – 2019
UCLA Cellular and Molecular Biology Training Fellowship	2016 – 2018
UCLA Ernest F. Hare, Jr. Memorial Scholarship	2017
UCLA Scholarly Conference Travel Grant	2015
UCSD Provost Honors (Dean's List Equivalent)	2006 – 2010
UCSD Warren College Honors Program	2006 – 2010
UCSD Golden Key International Honor Society	2010

# **Chapter 1**

## **Introduction to Iterative Enzyme Catalysis in Fungal Polyketide**

### **Biosynthesis**

- 1.1 Iterative Enzyme Catalysis in Polyketide Biosynthesis**
- 1.2 Overview of Type I Iterative Fungal Highly-Reducing Polyketide Synthase (HRPKS)**
- 1.3 Overview of Post-PKS Iterative Oxygenases**
- 1.4 Lovastatin**
- 1.5 Chaetoglobosin A**
- 1.6 Cytochalasin E and Cytochalasin K**
- 1.7 Aurovertin E**
- 1.8 Conclusion**
- 1.9 References**



## Section 1.1 Iterative Enzymes Catalysis in Polyketide Biosynthesis

Polyketides play a vital role in modern medicine as some of the most important therapeutic drugs in the world.<sup>1</sup> The concise biosynthetic pathways of several fungal polyketide natural products feature the combination of iterative polyketide synthases (PKSs) and multifunctional and iterative oxygenases.<sup>2-4</sup> An iterative enzyme can be defined as a protein that reuses its catalytic unit to sequentially catalyze distinct modifications on a substrate. Typical examples of iterative enzymes include ubiquitous proteins such as nucleases, proteases and glycosidases. In the context of secondary metabolism, we define iterative enzymes as those that contain active sites that can catalyze multiple rounds of structural modifications during the biosynthesis of a natural product.

Fatty acid synthases (FASs) serve as a well-studied example of iterative enzymes.<sup>5</sup> FASs catalyze the homo-polymerization of malonyl-CoA and complete  $\beta$ -reduction to afford a saturated lipid chain. Although iterative PKSs share many similarities with FASs including domain architecture and catalytic mechanisms, they are programmed in a more sophisticated fashion to carry out combinatorial catalysis. The  $\beta$ -reduction domains are differentially used during each iteration of chain elongation to generate the structural diversity observed in many fungal natural products.<sup>2-3</sup> The diverse polyketide backbones are then further functionalized by tailoring enzymes such as oxygenases. Some of these tailoring oxygenases function iteratively and catalyze multiple rounds of oxidative modifications to generate the mature and bioactive polyketide natural product.

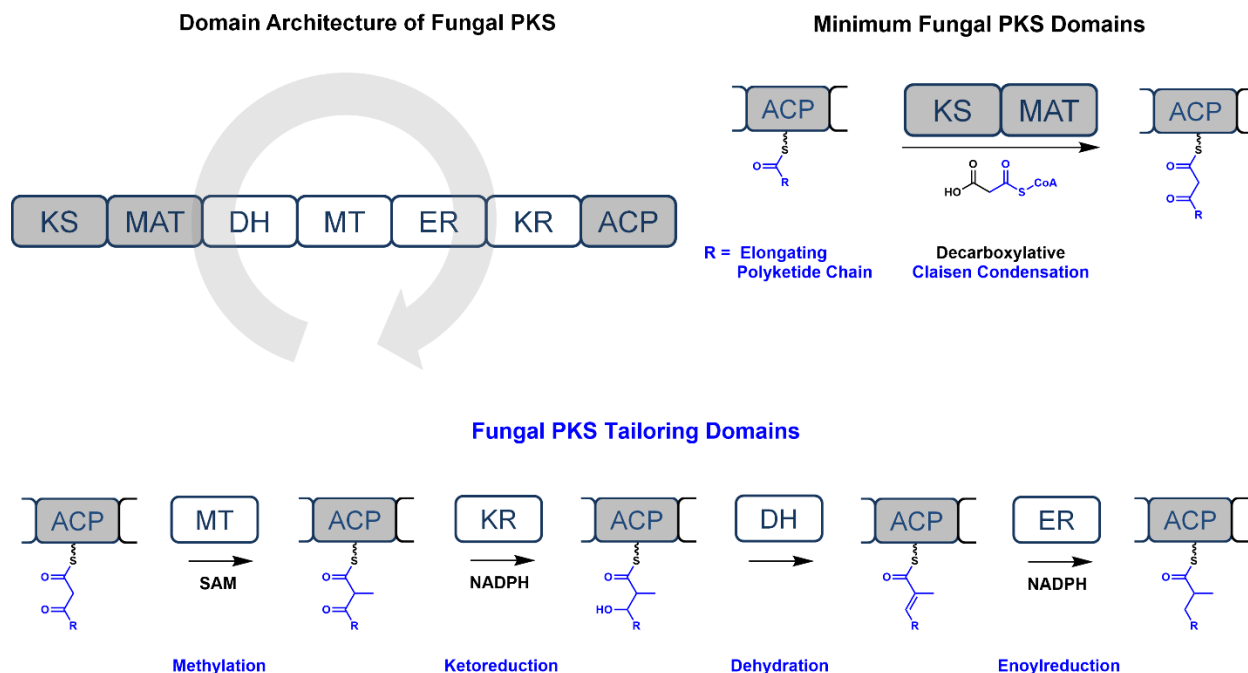
Together, the synergistic PKS and oxidative enzymes have enabled fungi to adeptly introduce chemical complexity to efficiently biosynthesize many chemically diverse and

biologically active molecules such as lovastatin (anti-hypercholesterolemia),<sup>6-7</sup> chaetoglobosin A (apoptosis-inducer),<sup>8</sup> cytochalasin E (anti-angiogenesis),<sup>9-10</sup> and aurovertins (ATP synthase inhibitor).<sup>11-12</sup> Inspired by their complex chemical structures and useful properties, these fungal polyketides are attractive targets for total chemical synthesis.<sup>13-16</sup> Some of these total syntheses draw strategies from Nature and employ biomimetic methodologies.<sup>16-17</sup> Yet, the elegance, brevity and synchronization of many enzymatic transformations often lack an equivalent chemical parallel. This chapter will highlight several succinct biosyntheses of fungal polyketides that are collaboratively forged by these iterative enzymes.

## **Section 1.2 Overview of Type I Iterative Fungal Highly-Reducing Polyketide Synthase (HRPKS)**

Type I PKSs are multi-domain megasynthases that possess the catalytic domains required for polyketide biosynthesis.<sup>18</sup> In most type I bacterial PKSs,<sup>19</sup> multiple sets of domains are typically compiled into modules and their biosynthesis proceeds in an assembly-line fashion. In contrast, type I fungal highly-reducing polyketide synthase (HRPKS) use a single set of domains in a highly programmed and permutative fashion.<sup>2-3</sup> The architecture of the fungal HRPKSs consist of the minimal fungal PKS components and the auxiliary tailoring domains (**Figure 1.1**). The  $\beta$ -ketoacyl synthase (KS),<sup>20</sup> malonyl-CoA: ACP transacylase (MAT) and acyl carrier protein (ACP)<sup>21</sup> form the minimal fungal PKS components—the basis for the chain-extending iterations through decarboxylative Claisen condensations (**Figure 1.1**). During each iteration, the minimal fungal PKS components catalyze the decarboxylative polymerization of malonyl-CoA to elongate the polyketide chain by a ketide (two carbons) unit.

Following each chain extension step, the ACP-bound,  $\beta$ -keto thioester intermediate may undergo a series of modifications from the tailoring domains such as  $\alpha$ -methylation by the methyltransferase (MT),  $\beta$ -ketoreduction by the ketoreductase (KR),<sup>22</sup> dehydration by the dehydratase (DH),<sup>23</sup> and enoylreduction by the enoylreductase (ER) domains (**Figure 1.1**).<sup>24</sup> The MT domain utilizes *S*-adenosylmethionine (SAM) as the methylating agent while the reductive domains use nicotinamide adenine dinucleotide phosphate hydride (NADPH) as the reducing agent. The  $\alpha$  and  $\beta$  position of each ketide unit will differ depending on the extent of methylation and reduction during each cycle. Through different permutative tailoring modifications following each chain extension, the same set of tailoring domains can install structural diversity into the  $\alpha$  and  $\beta$  positions of polyketide backbones.<sup>2-3</sup> The elongation-tailoring events proceed iteratively until the polyketide chain extension is terminated through product off-loading such as hydrolysis or reductive release. While these permutative programming steps may seem randomly coded at first glance, the examples shown in this chapter show that the complete polyketide products are exquisitely and precisely functionalized to facilitate the downstream modifications—many of which are oxidative. Currently, underlying programming rules for the iterative catalysis of both bacterial and fungal polyketide synthases remain an active area of research.



**Figure 1.1:** The domain architecture and the catalysis of the minimum fungal PKS domains and the fungal PKS tailoring domains in fungal polyketide synthases. In fungal PKSs, a single set of domains is used iteratively and permutatively to generate backbone structural diversity.

### Section 1.3 Overview of Post-PKS Iterative Oxygenases

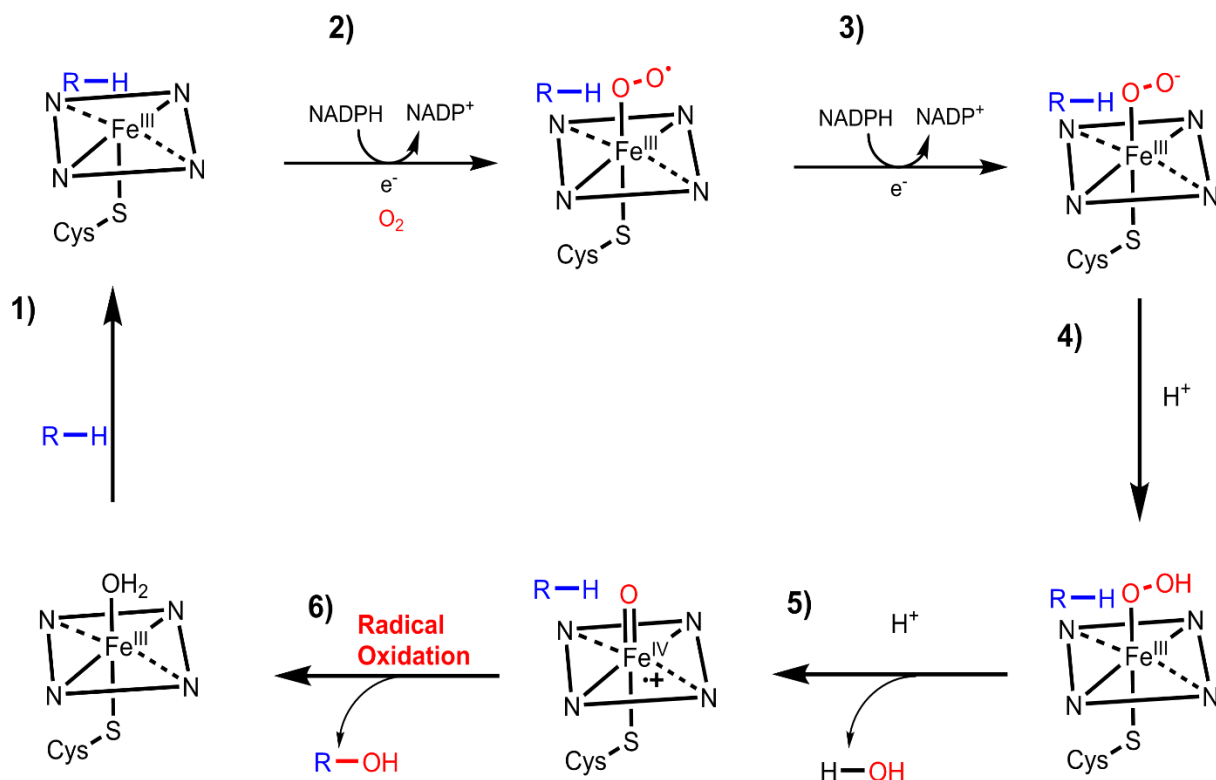
After the biosynthesis of the polyketide scaffold, further modifications are catalyzed by post-PKS tailoring enzymes. Oxygenases are common tailoring enzymes responsible for much of the chemical diversity of natural products.<sup>25</sup> These enzymes often carry out a diverse range of redox reactions including hydroxylations, epoxidations, dehydrogenations, cyclizations, and various rearrangements—often decreasing the lipophilicity of polyketide metabolites.<sup>26</sup> Recent discoveries have found several multifunctional oxygenases that can act iteratively on multiple sites of their substrates.<sup>4,27</sup> Thus, iterative oxygenases are enzymes that can introduce multiple oxygen atoms using molecular oxygen ( $O_2$ ) at different sites on a single substrate.

There are several major classes of oxygenases including cytochrome P450 monooxygenases (P450s), flavin-containing monooxygenases (FMOs), and non-heme, non-iron- $\alpha$ -ketoglutarate-dependent dioxygenases. Examples of iterative catalysis are found in each of these classes in fungal polyketide biosynthesis. Monooxygenases incorporate one oxygen atom from molecular oxygen while dioxygenases can incorporate both oxygen atoms. A notable example of an iterative, multifunctional  $\alpha$ -ketoglutarate-dependent dioxygenase is the AusE enzyme, which catalyzes several interesting oxidations on the terpene portion of the fungal polyketide hybrid molecule austinol.<sup>28</sup> However, this chapter will focus on examples which feature iterative oxidations from P450s and FMOs that work together with iterative PKSs.

#### *P450s are heme-dependent enzymes*

P450s are ubiquitous oxidative enzymes that span across all organisms, from bacteria to humans. P450s are heme-binding enzymes and share a highly conserved protein fold.<sup>29</sup> These oxygenases use heme complexes to oxidize a vast multitude of different substrates using molecular oxygen (**Figure 1.2**).<sup>30</sup> Despite their substantial range of substrate diversity, P450 oxidations can be highly stereoselective and regioselective. Furthermore, P450s can be organized into Class I and Class II subgroups, differing in their associated reductive partner. Fungal biosynthetic pathways typically use Class II P450s—microsomal transmembrane enzymes often associated with a transmembrane cytochrome P450 reductase (CPR) partnering enzymes which together moderate electron transfer from NADPH.<sup>29</sup> P450 enzymes in natural product biosynthesis use a single-electron manifold to produce radical intermediates. These reactive radical intermediates can lead to a variety of modifications including hydroxylation, epoxidation, dehydrogenation, radical homo-coupling, etc. A well-known example of an iterative P450-catalyzed reaction is the repeated oxidation of methyl groups to corresponding carboxylic acids

via alcohol and aldehyde intermediates.<sup>31-33</sup> This set of transformations is widely found in plant and fungal biosynthetic pathways and is also prominently featured in the conversion of lanosterol into cholesterol.<sup>34</sup>

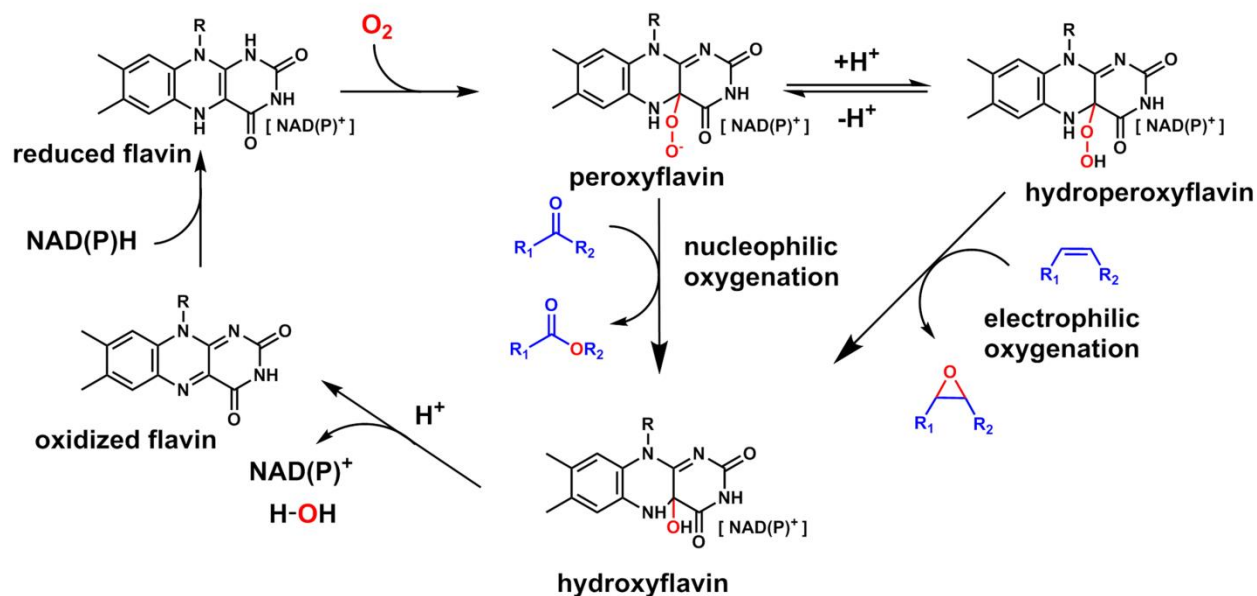


**Figure 1.2:** The mechanism of P450 catalyzed, radical-mediated hydroxylation of a substrate (**R**).

*FMOs are versatile oxidases and oxygenases*

FMOs are widespread enzymes that catalyze a large variety of substrate oxidations such as dehydrogenation, hydroxylations, epoxidations, Baeyer-Villiger oxidations, and sulfoxidations.<sup>35-36</sup> Unlike P450s which use a heme prosthetic group to activate molecular oxygen, FMOs instead use a flavin cofactor such as flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN), to generate reactive peroxy species that serve as nucleophiles (peroxyflavin,  $Fl-OO^-$ ) or electrophiles (hydroperoxyflavin,  $Fl-OOH$ ) (**Figure 1.3**).<sup>36</sup> The flavin

coenzymes are often tightly or covalently bound to the FMO.<sup>36</sup> After each round of catalysis, the flavin cofactor can be reduced in the presence of NAD(P)H to repeat the catalytic cycle.



**Figure 1.3:** The mechanism of FMO mediated nucleophilic oxygenation (see cytochalasin E/K, Section 1.6) and electrophilic oxygenation (see aurovertin E, Section 1.7).

## Section 1.4 Lovastatin

Lovastatin (**5**) is one of the most pharmaceutically relevant fungal natural products in use today. Lovastatin and its synthetic analogue (simvastatin) exhibit potent cholesterol-lowering activity by inhibiting the HMG-CoA reductase (HMGR) enzyme in cholesterol biosynthesis.<sup>37</sup> Lovastatin represents a classical example of a fungal natural product that is synthesized by a HRPKS. Produced by the filamentous fungus *Aspergillus terreus*, **5** is the esterified product of two highly-reduced polyketide chains catalyzed by LovB and LovF, respectively.

### Polyketide Core Assembly (LovB/LovC and LovF)

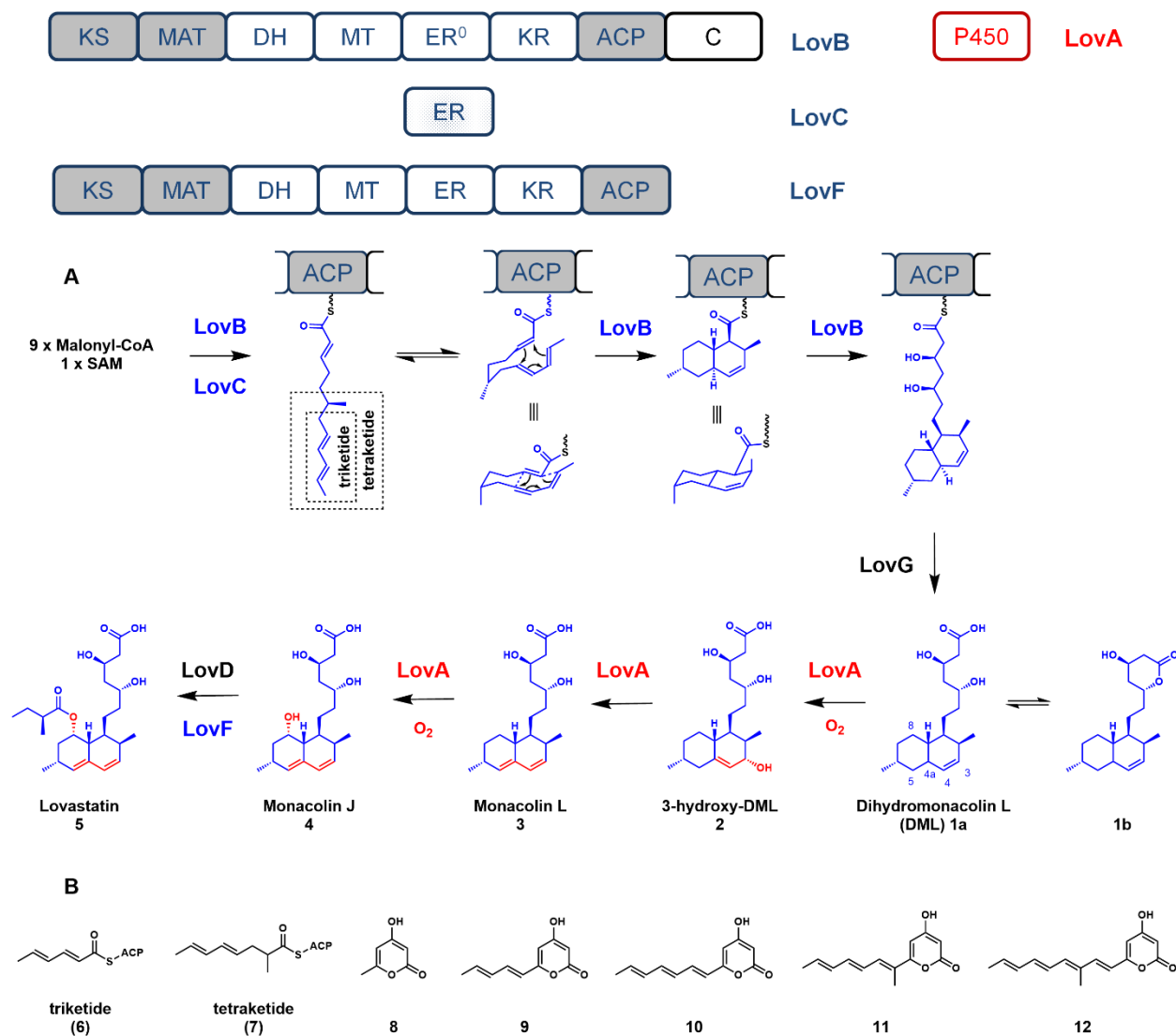
LovB (nonaketide synthase) and LovC (*trans*-acting ER partner) catalyze ~35 highly programmed reactions to produce the ACP-bound dihydromonacolin L (DML, **1a**), the nonaketide precursor to **5**.<sup>38</sup> The multiple reductions are catalyzed by the KR and DH domains within LovB in concert with LovC. LovB also catalyzes a Diels-Alder reaction to form the decalin core, a yet unresolved step proposed to occur at the hexaketide stage. The polyketide chain is then offloaded from LovB by the thioesterase LovG to yield **1a**, which can undergo iterative, oxidative tailoring around the decalin core to yield monacolin J (**4**). The acyltransferase LovD can then transesterify an  $\alpha$ -methylbutyrate diketide from LovF onto the newly installed 8-hydroxyl group of **4** to afford **5** (**Figure 1.4A**).

LovB exhibits extraordinary specificity and precision in catalyzing the formation of **1a**. Each structural portion of lovastatin is important for its activity as a HMGR inhibitor. The 3,5-dihydroxy acid portion is the warhead that binds to the active site of HMGR, while the decalin ring provides the hydrophobic anchor for high-affinity binding. To accomplish this, LovB shows an impressive ability to control the timing of reductions and methylation steps during polyketide chain elongation, with an ability to correct its mistakes.<sup>38</sup> In the first two rounds of polyketide chain elongation, only the KR and DH tailoring domains are active to afford the triketide (**6**). In the next round, the MT of LovB and the LovC will also function to produce the methylated tetraketide (**7**). As a method of proofreading, aberrantly tailored acyl intermediates are offloaded through  $\alpha$ -pyrone formation. *In vitro* studies with LovB showed that the absence of the reductive NADPH cofactor led to polyketide chain offloading by lactonization to produce the triketide  $\alpha$ -pyrone (**8**).<sup>38</sup> In the absence of either of SAM or LovC, offloading would also proceed to form other  $\alpha$ -pyrones (**Figure 1.4B**). These findings illustrate that if the PKS cannot correctly



proceed to later stages of biosynthesis, the PKS is programmed to catalyze two additional condensation cycles (KS/MAT) to offload the shunt polyketide products.<sup>38</sup> This proofreading mechanism prevents stalling of the incorrectly modified acyl chain on LovB.

Furthermore, mechanistic studies of LovB provided valuable insight into the programming of iterative PKSs. For example at the tetraketide stage, the methylation step is a prerequisite for ER activity as LovC fails to recognize substrates that are not methylated.<sup>38</sup> Recent kinetic studies also compared the substrate specificity of the KR and MT domains.<sup>39</sup> While the KR domain displays broader substrate promiscuity, the MT domain exhibits high specificity only for the tetraketide substrate. The experiments also illustrated that the kinetic competition between the KR and MT domains prevents methylation at incorrect positions. These experiments suggest that the substrate specificity of the catalytic domains may decide the sequence of reactions in the polyketide biosynthesis. Studies on LovB have unveiled key features of PKSs, including the strict checks and balances of individual tailoring domains on the fidelity of previous iterations.<sup>38-39</sup> Subtle mistakes are recognized and rejected in the later iterations. Collectively, LovB is an incredibly accurate enzyme. This accuracy is clearly exemplified in the linear hexaketide triene intermediate. The presence of the three double bonds is a result of precise tailoring to set up the intramolecular Diels-Alder reaction that forges the decalin core (**Figure 1.4A**). Clearly, any mistakes prior to that would not lead to formation of the decalin core.



**Figure 1.4:** Biosynthesis of lovastatin. A) The polyketide products produced from an iterative PKS are highlighted in blue and oxidative tailoring produced from an iterative oxygenase are highlighted in red. The portions of the polyketide chain reflected by intermediates **6** and **7** are shown in the dashed boxes. B) The natural triketide and tetraketide intermediates and the offloaded polyketide pyrone shunt products.

Additional functionalization by post-PKS iterative oxygenase (LovA)

After formation of **1a**, several oxidation steps are required to generate the penultimate intermediate **4**, which is acylated at the C8-hydroxyl group by the LovF product to yield **5**. Although the lovastatin biosynthetic gene cluster encodes two P450 genes (LovA and Orf17), gene disruption of LovA in *Aspergillus terreus* confirmed its role as the only oxidative enzyme

required in lovastatin biosynthesis.<sup>7</sup> Furthermore, heterologous reconstitution of LovA and its cytochrome P450 oxidoreductase (CPR) partner in *Saccharomyces cerevisiae* confirmed the iterative oxygenase activity of LovA.<sup>40</sup> It was demonstrated that LovA first converted **1a** to monacolin L (**3**), and then to **4** via sequential radical hydroxylation reactions (**Figure 1.4A**). The conversion of **1a** to 3-hydroxymonacolin L (**2**) proceeds through allylic hydrogen abstraction followed by an oxygen rebound to introduce the hydroxyl group at the C3 $\alpha$  position (**Figure 1.2** and **Figure 1.4A**). Dehydration of the hydroxyl group yields the diene in the decalin core of **3**. LovA then performs the second oxidation on the C8 $\alpha$  position of the decalin to introduce the hydroxyl group.

Intriguingly, the second olefin in the decalin core of **3** cannot be produced by a PKS alone. However, LovB shrewdly placed the first unsaturation between C3-C4 in DML to activate the decalin ring for an allylic hydroxylation-elimination sequence to yield the second unsaturation between C4a and C5. The C3-C4  $\pi$ -bond has been experimentally demonstrated to be vital for substrate recognition by LovA since the saturated DML derivative cannot be taken as a substrate.<sup>40</sup> Furthermore, the lack of hydroxylation at the C8 $\alpha$  position prior to the second unsaturation possibly also suggests the order of oxidation events catalyzed by LovA is exact (see chaetoglobosin P450, Section 1.5, for contrast). It is possible that the conformational change (chair to half-chair) upon diene formation in the decalin substrate may be required for the second hydroxylation by LovA to occur.

## Section 1.5 Chaetoglobosin A

Cytochalasans are a family of fungal polyketide-amino acid hybrid natural products that have interesting biological activities.<sup>41</sup> Many cytochalasans exhibit antibiotic, anti-inflammatory,

anti-angiogenic and anti-cytokinesis activity.<sup>41</sup> Furthermore, the representative members of the cytochalasan family all share an isoindolone moiety fused to a macrocycle in its tricyclic core. Among the cytochalasans, the highly oxygenated chaetoglobosin A (**20**) inhibits actin polymerization and was the first member to have its biosynthetic gene cluster elucidated.<sup>42</sup>

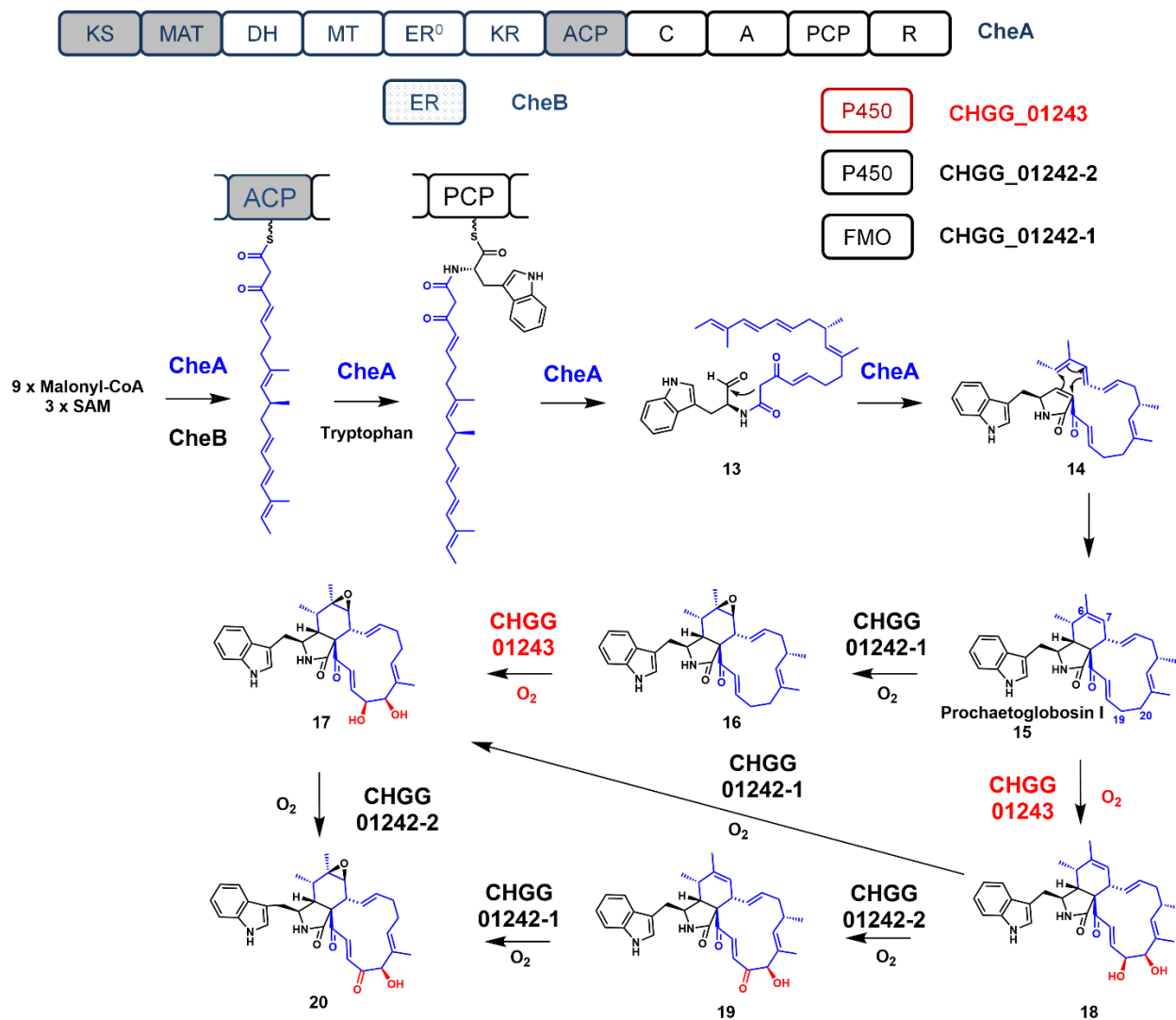
Polyketide core assembly (CheA/CheB and CHGG 01239/CHGG 01240)

The backbone of chaetoglobosin A is assembled through the polyketide synthase-non-ribosomal synthetase (PKS-NRPS) hybrid enzyme, CheA (**Figure 1.5**). This finding was first confirmed by RNA-mediated gene silencing in *Penicillium expansum*.<sup>42</sup> A functionally equivalent CheA homolog was also discovered from another chaetoglobosin A producer, *Chaetomium globosum*.<sup>43</sup> Similar to the LovB PKS, the PKS portion of CheA also lacks a functional ER domain (**Figure 1.5**), and is complemented with a *trans*-acting ER partnering enzyme (CheB) in the construction of the nonaketide portion of the molecule. Differing from the lovastatin nonaketide precursor, the CheA PKS product is then condensed with a tryptophanyl thioester attached to the fused NRPS module of CheA. The nonaketide-tryptophanyl thioester intermediate **13** is then reductively offloaded from CheA by the reduction (R) domain on the NRPS module (**Figure 1.5**).<sup>42-43</sup>

An intramolecular Knoevenagel condensation between the  $\beta$ -ketoamide and the aldehyde produces the pyrrolinone **14** that can serve as a dienophile in a subsequent intramolecular Diels-Alder reaction. The resulting Diels-Alder reaction precisely and stereoselectively forges the perhydro-isoindolone core of prochaetoglobosin I **15**.<sup>44</sup> The cyclized **15** requires a series of downstream oxidations to produce the final chaetoglobosin A product.

Comparing the HRPKS domain architecture of LovB and CheA, the main difference lies in the complete NRPS module appended at the C-terminus of CheA (**Figure 1.4-1.5**).

Condensation of PKS product with an aminoacyl building block serves several purposes: 1) this represents one method of chain termination which is likely controlled through the substrate specificity of the corresponding C domain in the NRPS module; 2) the amino acid unit introduces structural variation, especially a basic nitrogen that is absent from the PKS alone—providing a convenient entry point for equipping polyketide products with the versatile nitrogen atom; and 3) formation of the pyrrolinone through Knoevenagel condensation generates a dienophile for a Diels-Alder reaction. Again, the diene is formed from the permutative activities of the PKS module. Interestingly, the diene is part of a triene moiety and the Diels-Alderase is able to regioselectively and stereoselectively perform the cycloaddition to yield **15**. The functional differences between LovB and CheA therefore result in vastly different polyketide structures and biological activities. The contrast highlights the importance of understanding the programming rules of PKSs from seemingly similar PKS architecture.



**Figure 1.5:** Biosynthetic pathway of chaetoglobosin A. The polyketide products produced from an iterative PKS are highlighted in blue and oxidative tailoring produced from an iterative oxygenase are highlighted in red.

Additional functionalization by post-PKS iterative oxygenase (CHGG\_01243)

After assembly of the tricyclic core of **15**, downstream oxidative tailoring by three oxidative enzymes CHGG\_01242-2 (FMO), CHGG\_01242-1 (P450), and CHGG\_01243 (P450) decorate the polyketide portion of the molecule with different oxygenated functional groups (alcohol, ketone and epoxide) to form **20**.<sup>43</sup> Different combinations of targeted gene deletions of these three oxygenases were constructed to generate single and double gene deletion mutants.

Product analysis from these mutants indicated that P450 CHGG\_01243 performs the iterative hydroxylations on C19 and C20; both are allylic carbons within the macrocycle scaffold (**Figure 1.5**). Subsequently, the FMO CHGG\_01242-2 catalyzes oxidation of the newly introduced C19 hydroxyl to a ketone. Finally, the P450 CHGG\_01242-1 performs the epoxidation across the C6-C7  $\pi$ -bond to yield **20**.

When analyzing the product profile from the various mutants, it was discovered that the order of the downstream tailoring oxidations is not strictly linear (**Figure 1.5**). The ketone formation must occur after CHGG\_01243 dihydroxylation, but the epoxidation by CHGG\_01242-1 can occur either before or after these two reactions. This shows a relatively broad substrate specificity of the iterative oxygenase CHGG\_01243 towards the remote segment of the molecule, as it can utilize either **15** or **16** as substrates. This contrasts with the highly selective nature of the LovA iterative oxygenase in the lovastatin biosynthetic pathway.

## **Section 1.6 Cytochalasin E and Cytochalasin K**

Despite the similar name, the cytochalasins are actually specific metabolites of the general cytochalasan family.<sup>41</sup> The cytochalasins are fungal polyketide-phenylalanine hybrids<sup>45-46</sup> that inhibit several cellular processes such as the polymerization of actin and cell division.<sup>47-48</sup> For example, cytochalasin E (**26**) has garnered significant attention due to its strong anti-angiogenic properties.<sup>49</sup> The *ccs* gene cluster in *Aspergillus clavatus* responsible for cytochalasin E and cytochalasin K (cytochalasin E/K, **26** and **27**) biosynthesis was discovered for members of this family.<sup>9</sup>

*Polyketide core assembly (CcsA/CcsC)*

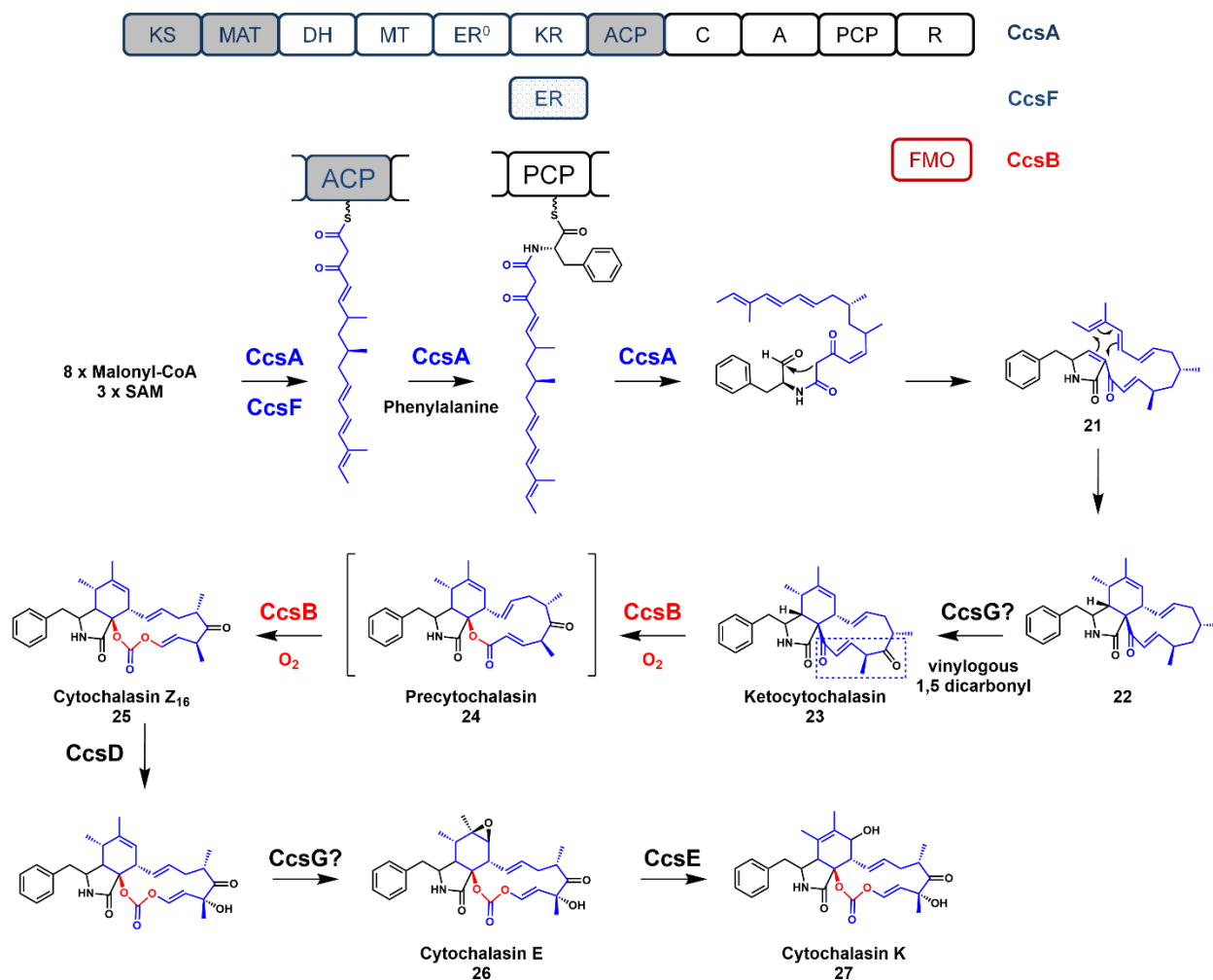
Considering the structural similarities between all cytochalasans, the HRPKSs responsible for the biosynthesis of chaetoglobosin A and cytochalasin E/K expectedly share similar domain architectures. When comparing the HRPKSs for chaetoglobosin A (CheA/B) and cytochalasin E/K (CcsA/C), both multidomain enzymes possess the same set of tailoring domains including the MT, all of the reductive domains, a partnering *trans*-acting ER, and a fused NRPS module.<sup>9</sup> Furthermore, the biosyntheses of their linear polyketide precursor are strikingly similar—differing only by a double bond and an ethylene unit (**Figure 1.5-1.6**).

The distinguishing feature of the cytochalasins from other members of the cytochalasan family is the incorporated amino acid, phenylalanine.<sup>46</sup> The ACP-bound linear polyketide precursor is condensed with the phenylalanyl thioester by the NRPS module to produce the  $\beta$ -ketoamide that can be cyclized by a Knoevenagel condensation to yield the pyrrolinone **21**. Operating under similar programming rules as CheA, the essential structural features required for the cycloaddition are found in the acyclic precursor, such as the diene in the triene tail and the dienophile. An intramolecular Diels-Alder reaction similarly furnishes the perhydro-isoindolone core and the large carbocycle **22**<sup>50</sup>—the hallmark of the cytochalasan family (**Figure 1.6**).<sup>41</sup> The Diels-Alder adduct **22** then undergoes an iterative oxidation at C17 to produce the essential vinylogous, 1,5-dicarbonyl system in ketocytochalasin **23** (**Figure 1.6-1.7**).

The subtle differences between the iterative HRPKSs for chaetoglobosin A and cytochalasin E/K further highlights the precise calibration of iterative catalysis by the HRPKS. In the chaetoglobosin A biosynthesis, the absent ethylene unit is flanked by  $\pi$  bonds which forms the vicinal allylic carbons. These activated carbons are exactly the targets of the iterative allylic hydroxylation by the P450 enzyme, CHGG\_01243, in the *cheA* pathway.<sup>43</sup> Although cytochalasin E/K lacks the diol derived from the vicinal allylic carbons, the vinylogous 1,5-



dicarbonyl sets up a most fascinating set of iterative oxygenation to yield the carbonate group found in the final natural products.



**Figure 1.6:** Biosynthesis of cytochalasin E and cytochalasin K. The polyketide products produced from an iterative PKS are highlighted in blue and oxidative tailoring produced from an iterative oxygenase are highlighted in red.

Additional functionalization by post-PKS iterative oxygenase (CcsB)

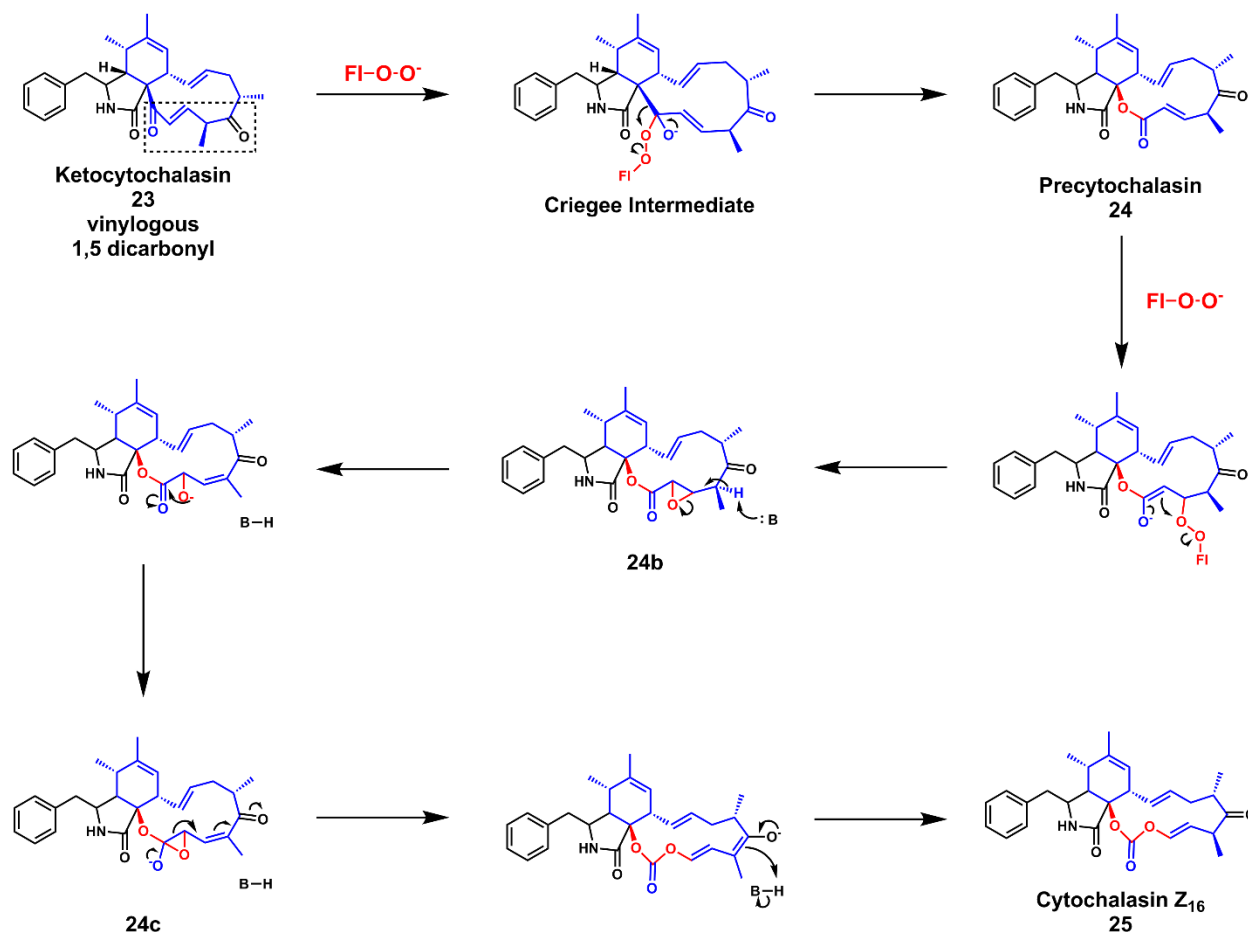
Among the cytochalasans, cytochalasin E/K are particularly interesting due to the unique vinyl carbonate moiety incorporated in the macrocyclic ring.<sup>51</sup> The exceptionally rare carbonate group in natural products is essential for the cytotoxic properties of cytochalasin E/K, as the

corresponding ester is ~500 fold lower in activity.<sup>52-53</sup> Isotopic labeling experiments suggested that a Baeyer-Villiger monooxygenase (BVMO)<sup>54-55</sup> can convert deoxaphomin (carbocycle) to cytochalasin B (lactone) via an oxygen derived from molecular oxygen.<sup>10, 56-57</sup> Consequently, the carbonate moiety in cytochalasin E/K has been postulated to result from two consecutive oxygen insertions via a single BVMO enzyme. This is an unprecedented reaction in both synthetic and biosynthetic chemistry.

Deletion of the BVMO, CcsB, in the *A. clavatus* strain abolished the production of cytochalasin E/K and resulted in the accumulation of **23**, the ketone containing substrate.<sup>10</sup> *In vitro* assays with purified CcsB enzyme demonstrated the conversion of **23** to the carbonate cytochalasin Z<sub>16</sub> **25** via two oxygen atom insertions and a vinyl ketone ester shunt product (not shown).<sup>10</sup> The shunt product is isomerized from the ester product precytochalasin **24**, which was not isolated under assay conditions. A point mutation on the FMO enzyme's catalytic arginine to alanine also abolished the formation of **23** and the lactone shunt product (**Figure 1.7**). Together, these findings illustrate the first known example of an FMO enzyme iteratively catalyzing two oxygen insertions to convert a ketone functional group to a carbonate moiety.

While the first oxygen insertion to form **24** is expected to follow the classical BVMO mechanism with regards to migratory aptitude, the unusual second oxygen insertion is expected to be inserted via a completely different mechanism. Based on experimental data and survey of all known cytochalasins, it appears that the vinylogous-1,5-dicarbonyl system is a prerequisite for carbonate formation.<sup>10</sup> The energetically-favored mechanism for the second oxygen insertion could proceed by an initial Michael addition **24** by the peroxyflavin anion to furnish the  $\alpha$ - $\beta$  epoxide on the lactone (**24b**) (**Figure 1.7**). A subsequent ring-opening elimination can then produce an epoxy-alkoxide product (**24c**), which upon on ring opening across the enone can then

yield the carbonate moiety in cytochalasin Z<sub>16</sub> (**25**). Formation of the carbonate product requires cleavage of the C-C bond in the epoxide and the remote vinyl ketone serving as an electron sink. Based on this mechanism, the iterative HRPKS, CcsA, precisely orchestrates the construction of the polyketide precursor that can be oxidized to yield the essential vinylogous-1,5-dicarbonyl **23**. CcsB can utilize the unique electronic properties of the polyketide system to convert a ketone into the carbonate found in cytochalasin E/K. While sequence alignment comparisons suggest that CcsB is a typical BVMO, its pairing with a unique substrate yields a remarkably rare functional group among natural products.



**Figure 1.7:** Proposed mechanism of the CcsB catalyzed conversion of ketocytochalasin to cytochalasin Z<sub>16</sub>. Alternatively, a direct attack of precytochalasin (**24**), followed by a vinyl migration on the Criegee intermediate can also form the same epoxy-alkoxide system (**24c**) that can collapse to form **25**; this would make CcsB a true, iterative BVMO.

## Section 1.7: Aurovertin E

Aurovertin E (**34**) represents the biosynthetic precursor and the representative structure of the fungal polyketide family of aurovertins. The acetylated derivative, aurovertin B, exhibits potent and non-competitive inhibition of F<sub>1</sub> ATPase, as well as induction of apoptosis and cell arrest in breast cancer cells.<sup>12, 58</sup> The defining structural hallmark of the aurovertins includes a methylated  $\alpha$ -pyrone attached to a 2,6-dioxabicyclo[3.2.1]octane (DBO) ring system via a triene linker. Unlike the lovastatin, chaetoglobosin A and cytochalasin E examples, the heterocyclic

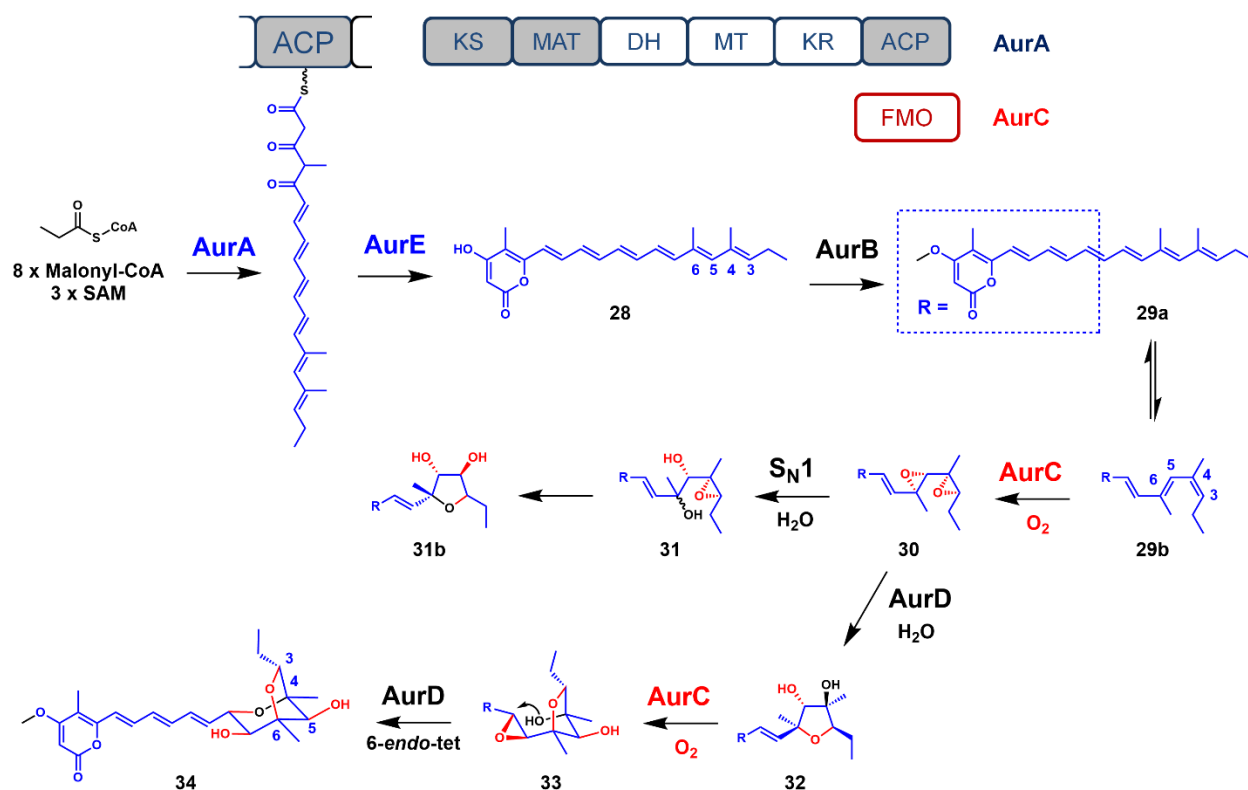
portion of aurovertin E (**34**) is not derived from a Diels-Alder reaction, but rather through the iterative epoxidation of a polyene precursor and a cascade of concerted epoxide openings that result in bridged, cyclic ether formations.

#### Polyketide core assembly (AurA)

Another key differentiating feature of aurovertin E is the conjugated polyene structure deficient in fully reduced carbons. The only methylene unit within the entire structure has been shown to be derived from a propionate starter unit through isotope labeling.<sup>59</sup> This structural feature is manifested in the domain architecture of the PKS. The AurA enzyme from *Calcarisporium arbuscula* is a HRPKS that possesses the minimal PKS unit along with the MT, KR and DH tailoring domains. However, it lacks a *cis*-acting or *trans*-acting ER domain. Consequently, the PKS can only reduce each  $\beta$ -carbonyl to the enoyl functional group, consistent with the observed polyolefin structure. The programming rules of AurA are remarkably precise, as the PKS employs all of its reductive domains for six iterations to generate a hexa-ene thioester intermediate. However, in the final two iterations, the PKS is able to shut down the reductive KR domain to generate a tricarbonyl tail that is required for subsequent chain release steps. Furthermore, the PKS precisely uses its MT domain to introduce two  $\alpha$ -methyl groups at C4 and C6 only. These methyl groups may be later involved in the post-PKS oxidation that forges the DBO ring system.

The tricarbonyl hexaene linear precursor is offloaded by enzymatic or spontaneous intramolecular cyclization (**Figure 1.8**). The offloading proceeds by enolization of the  $\delta$ -carbonyl, followed by an additional-elimination on the thioester carbonyl to produce the  $\alpha$ -pyrone-polyene conjugated intermediate **28**. Curiously, this  $\alpha$ -pyrone-polyene motif is also

produced by LovB in the absence of its ER partner (LovC) (**Figure 1.4**). This scaffold is also present in other related fungal pyrone-polyene metabolites such as citreoviridin,<sup>60</sup> asteltoxin,<sup>61</sup> and asteltoxin B.<sup>62</sup> Pausing of the KR activity in the final two extension cycles is essential for the enol-induced lactonization to produce the  $\alpha$ -pyrone substructure in **28**.



**Figure 1.8:** Proposed biosynthesis of aurovertin E. The polyketide products produced from an iterative PKS are highlighted in blue and oxidative tailoring produced from an iterative oxygenase are highlighted in red.

Additional functionalization by post-PKS iterative oxygenases (AurC)

Isotope  $^{18}\text{O}$ -labeling studies shed valuable insight on the origin of the oxygen atoms in the DBO ring system. Using  $^{18}\text{O}_2$ , the oxygen atom within the C3-O-C6 ether linkage as well as the equatorial hydroxyl groups on C5 and C7 were shown to be derived from molecular oxygen. The lack of labeling of the oxygen within the C4-O-C8 ether linkage implies that the oxygen atom is most certainly derived from water.<sup>63</sup> Despite having three oxygen atoms derived from  $\text{O}_2$

in the DBO ring system, the identified *aur* gene cluster revealed only a single FMO, AurC. Taken together, these observations hinted the utilization of an iterative FMO enzyme in selective epoxidations of the double bonds in **28**.<sup>64</sup>

Considering that the pyrone-polyene intermediate **28** possesses six consecutive alkenes, an oxygenase must exhibit high regioselectivity to oxidize the distal alkenes but not the proximal triene linker. Interestingly, these distal olefins are substituted with methyl groups installed by the MT domain. Density functional theory (DFT) calculations on the pyrone-polyene intermediate **28** revealed the non-planarity on C3-C4-C5-C6 in its preferred conformation. Presumably, this conformation avoids the possible A<sup>1,3</sup>-allylic strain between the  $\alpha$ -methyl groups. The induced asymmetry may therefore play a role in the specific epoxidation by the FMO enzyme.

Based on the absolute stereochemistry, isotope labeling experiments and detailed biochemical studies, formation of the DBO system was shown to be constructed using an epoxide-opening, cyclic ether ring-closing cascade.<sup>63-64</sup> Gene inactivation of the AurC oxygenase enzyme in *Calcarisporium arbuscula* abolished the production of aurovertin E and led to the accumulation of a methylated pyrone-polyene intermediate **29a**. The reactive **29b** pyrone-polyene resulted from the trans-cis isomerization of **29a**.

Furthermore, coexpression of the AurA (PKS), AurC (FMO) and AurB (*O*-MT) in yeast led to the production of tetrahydrofuranyl isomers **32** and **31b**. Chemical complementation with **32** to the  $\Delta$ *aurA* (PKS-knockout) strain restored the production of aurovertin E. The formation of **32** is consistent with the biomimetic total synthesis<sup>16</sup> and the structures of related natural products.<sup>60-62</sup> An epoxide hydrolase, AurD, was identified to control the regioselectivity of the epoxide openings (**Figure 1.8**). Overall, the distal olefins were converted to the DBO

substructure by an iterative FMO enzyme and an assisting epoxide hydrolase. The MT domain may have subtly introduced an asymmetry in the polyene for FMO substrate recognition. The collaboration between two iterative enzymes produced the complex framework of aurovertin E.

## **Section 1.8: Conclusion**

As shown in the four HRPKS examples, the synchronization between an iterative PKS and an iterative oxygenase can succinctly construct complex fungal polyketide natural products. The examples discussed in this perspective underscore how the iterative HRPKS produces precisely chiseled intermediates specifically for the iterative oxygenase partner. Many hypotheses may emerge to account for the evolutionary origin to develop these highly efficient tandem enzymatic systems. Naturally, one also may speculate on the evolutionary advantages of iterative enzymes as a whole. Considering the ubiquitous nature of many iterative enzymes such as non-specific proteases, a rational assumption could be the metabolic efficiency of multifunctional catalysis. Recognizing that polyketides constitute the most abundant secondary metabolites in fungi,<sup>65</sup> it is also tempting to consider the selective advantages that these iterative enzymes partners offer. Fungi are well known as rich producers of secondary metabolites, yet they are constrained by their microbial genome size. Iterative enzymes in natural product biosynthesis likely have evolved to help circumvent these intrinsic limitations.



## Section 1.9: References

1. Weissman, K. J., Polyketide biosynthesis: understanding and exploiting modularity. *Philos. T. Roy. Soc. A* **2004**, *362* (1825), 2671-2690.
2. Chooi, Y. H.; Tang, Y., Navigating the Fungal Polyketide Chemical Space: From Genes to Molecules. *J. Org. Chem.* **2012**, *77* (22), 9933-9953.
3. Cox, R. J.; Simpson, T. J., Fungal Type I Polyketide Synthases. *Method. Enzymol.* **2009**, *459*, 49-78.
4. Cochrane, R. V.; Vederas, J. C., Highly selective but multifunctional oxygenases in secondary metabolism. *Acc. Chem. Res.* **2014**, *47* (10), 3148-61.
5. Jenni, S.; Leibundgut, M.; Boehringer, D.; Frick, C.; Mikolasek, B.; Ban, N., Structure of fungal fatty acid synthase and implications for iterative substrate shuttling. *Science* **2007**, *316* (5822), 254-261.
6. Hendrickson, L.; Ray Davis, C.; Roach, C.; Kim Nguyen, D.; Aldrich, T.; McAda, P. C.; Reeves, C. D., Lovastatin biosynthesis in *Aspergillus terreus*: Characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chemistry & Biology* **1999**, *6* (7), 429-439.
7. Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Hutchinson, C. R., Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **1999**, *284* (5418), 1368-1372.
8. Oikawa, H.; Murakami, Y.; Ichihara, A., Biosynthetic Study of Chaetoglobosin-a - Origins of the Oxygen and Hydrogen-Atoms, and Indirect Evidence for Biological Diels-Alder Reaction. *J. Chem. Soc. Perk. T. I* **1992**, (21), 2955-2959.

9. Qiao, K. J.; Chooi, Y. H.; Tang, Y., Identification and engineering of the cytochalasin gene cluster from *Aspergillus clavatus* NRRL 1. *Metab. Eng.* **2011**, *13* (6), 723-732.
10. Hu, Y. C.; Dietrich, D.; Xu, W.; Patel, A.; Thuss, J. A. J.; Wang, J. J.; Yin, W. B.; Qiao, K. J.; Houk, K. N.; Vederas, J. C.; Tang, Y., A carbonate-forming Baeyer-Villiger monooxygenase. *Nat. Chem. Biol.* **2014**, *10* (7), 552-554.
11. Wang, F.; Luo, D. Q.; Liu, J. K., Aurovertin E, a new polyene pyrone from the basidiomycete *Albatrellus confluens*. *J. Antibiot.* **2005**, *58* (6), 412-415.
12. vanRaaij, M. J.; Abrahams, J. P.; Leslie, A. G. W.; Walker, J. E., The structure of bovine F-1-ATPase complexed with the antibiotic inhibitor aurovertin B. *P. Natl. Acad. Sci. U. S. A.* **1996**, *93* (14), 6913-6917.
13. Hirama, M.; Iwashita, M., Total Synthesis of (+)-Monacolin-K (Mevinolin). *Tetrahedron Lett.* **1983**, *24* (17), 1811-1812.
14. Haidle, A. M.; Myers, A. G., An enantioselective, modular, and general route to the cytochalasins: Synthesis of L-696,474 and cytochalasin B. *P. Natl. Acad. Sci. U. S. A.* **2004**, *101* (33), 12048-12053.
15. Nishiyama, S.; Toshima, H.; Kanai, H.; Yamamura, S., Total Synthesis and the Absolute-Configuration of Aurovertin-B, an Acute Neurotoxic Metabolite. *Tetrahedron* **1988**, *44* (20), 6315-6324.
16. Forbes, J. E.; Pattenden, G., Total Synthesis of Preaurovertin, Putative Biogenetic Precursor of Aurovertin - Biosynthetic Interrelationships between the Aurovertins, Citreoviridinols and Asteltoxin. *J. Chem. Soc. Perk. T. 1* **1991**, (8), 1959-1966.
17. Stork, G.; Nakahara, Y.; Nakahara, Y.; Greenlee, W. J., Total Synthesis of Cytochalasin-B. *J. Am. Chem. Soc.* **1978**, *100* (24), 7775-7777.

18. Staunton, J.; Weissman, K. J., Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **2001**, *18* (4), 380-416.
19. Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R. H.; Hakansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G., Structure of a modular polyketide synthase. *Nature* **2014**, *510* (7506), 512-+.
20. Tang, Y. Y.; Kim, C. Y.; Mathews, I. I.; Cane, D. E.; Khosla, C., The 2.7-angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *P. Natl. Acad. Sci. U. S. A.* **2006**, *103* (30), 11124-11129.
21. Lim, J.; Kong, R.; Murugan, E.; Ho, C. L.; Liang, Z. X.; Yang, D., Solution Structures of the Acyl Carrier Protein Domain from the Highly Reducing Type I Iterative Polyketide Synthase CalE8. *Plos One* **2011**, *6* (6).
22. Zheng, J. T.; Keatinge-Clay, A. T., The status of type I polyketide synthase ketoreductases. *Medchemcomm* **2013**, *4* (1), 34-40.
23. Akey, D. L.; Razelun, J. R.; Tehranisa, J.; Sherman, D. H.; Gerwick, W. H.; Smith, J. L., Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway. *Structure* **2010**, *18* (1), 94-105.
24. Ames, B. D.; Nguyen, C.; Bruegger, J.; Smith, P.; Xu, W.; Ma, S.; Wong, E.; Wong, S.; Xie, X. K.; Li, J. W. H.; Vederas, J. C.; Tang, Y.; Tsai, S. C., Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *P. Natl. Acad. Sci. U. S. A.* **2012**, *109* (28), 11144-11149.
25. Walsh, C. T., A chemocentric view of the natural product inventory. *Nat. Chem. Biol.* **2015**, *11* (9), 620-624.

26. Cashman, J. R., Some distinctions between flavin-containing and cytochrome P450 monooxygenases. *Biochem. Biophys. Res. Commun.* **2005**, *338* (1), 599-604.
27. Wu, L. F.; Meng, S.; Tang, G. L., Ferrous iron and alpha-ketoglutarate-dependent dioxygenases in the biosynthesis of microbial natural products. *BBA-Proteins Proteom.* **2016**, *1864* (5), 453-470.
28. Matsuda, Y.; Awakawa, T.; Wakimoto, T.; Abe, I., Spiro-Ring Formation is Catalyzed by a Multifunctional Dioxygenase in Austinol Biosynthesis. *J. Am. Chem. Soc.* **2013**, *135* (30), 10962-10965.
29. Dawson, J., Cytochrome P450. Structure, mechanism, and biochemistry, 2nd edition - DeMontellano, P. *Science* **1996**, *271* (5255), 1507-1508.
30. Coon, M. J., Cytochrome P450: Nature's most versatile biological catalyst. *Annu. Rev. Pharmacol.* **2005**, *45*, 1-25.
31. Abe, S.; Sado, A.; Tanaka, K.; Kisugi, T.; Asami, K.; Ota, S.; Il Kim, H.; Yoneyama, K.; Xie, X.; Ohnishi, T.; Seto, Y.; Yamaguchi, S.; Akiyama, K.; Yoneyama, K.; Nomura, T., Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro. *P. Natl. Acad. Sci. U. S. A.* **2014**, *111* (50), 18084-18089.
32. Helliwell, C. A.; Chandler, P. M.; Poole, A.; Dennis, E. S.; Peacock, W. J., The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *P. Natl. Acad. Sci. U. S. A.* **2001**, *98* (4), 2065-2070.
33. Ro, D. K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; Chang, M. C. Y.; Withers, S. T.; Shiba, Y.; Sarpong,

- R.; Keasling, J. D., Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **2006**, *440* (7086), 940-943.
34. Risley, J. M., Cholesterol biosynthesis: Lanosterol to cholesterol. *J. Chem. Educ.* **2002**, *79* (3), 377-384.
35. van Berkel, W. J.; Kamerbeek, N. M.; Fraaije, M. W., Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.* **2006**, *124* (4), 670-89.
36. Walsh, C. T.; Wencewicz, T. A., Flavoenzymes: versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.* **2013**, *30* (1), 175-200.
37. Tobert, J. A., Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. *Nat. Rev. Drug Discov.* **2003**, *2* (7), 517-526.
38. Ma, S. M.; Li, J. W. H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. A.; Xie, X. K.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y., Complete Reconstitution of a Highly Reducing Iterative Polyketide Synthase. *Science* **2009**, *326* (5952), 589-592.
39. Cacho, R. A.; Thuss, J.; Xu, W.; Sanichar, R.; Gao, Z. Z.; Nguyen, A.; Vederas, J. C.; Tang, Y., Understanding Programming of Fungal Iterative Polyketide Synthases: The Biochemical Basis for Regioselectivity by the Methyltransferase Domain in the Lovastatin Megasyntase. *J. Am. Chem. Soc.* **2015**, *137* (50), 15688-15691.
40. Barriuso, J.; Nguyen, D. T.; Li, J. W. H.; Roberts, J. N.; MacNevin, G.; Chaytor, J. L.; Marcus, S. L.; Vederas, J. C.; Ro, D.-K., Double Oxidation of the Cyclic Nonaketide Dihydromonacolin L to Monacolin J by a Single Cytochrome P450 Monooxygenase, LovA. *J. Am. Chem. Soc.* **2011**, *133* (21), 8078-8081.
41. Scherlach, K.; Boettger, D.; Remme, N.; Hertweck, C., The chemistry and biology of cytochalasans. *Nat. Prod. Rep.* **2010**, *27* (6), 869-886.

42. Schumann, J.; Hertweck, C., Molecular basis of cytochalasan biosynthesis in fungi: Gene cluster analysis and evidence for the involvement of a PKS-NRPS hybrid synthase by RNA silencing. *J. Am. Chem. Soc.* **2007**, *129* (31), 9564-+.
43. Ishiuchi, K.; Nakazawa, T.; Yagishita, F.; Mino, T.; Noguchi, H.; Hotta, K.; Watanabe, K., Combinatorial Generation of Complexity by Redox Enzymes in the Chaetoglobosin A Biosynthesis. *J. Am. Chem. Soc.* **2013**, *135* (19), 7371-7377.
44. Sato, M.; Yagishita, F.; Mino, T.; Uchiyama, N.; Patel, A.; Chooi, Y. H.; Goda, Y.; Xu, W.; Noguchi, H.; Yamamoto, T.; Hotta, K.; Houk, K. N.; Tang, Y.; Watanabe, K., Involvement of Lipocalin-like CghA in Decalin-Forming Stereoselective Intramolecular [4+2] Cycloaddition. *Chembiochem* **2015**, *16* (16), 2294-2298.
45. Binder, M.; Kiechel, J. R.; Tamm, C., Zur Biogenese des Antibioticums Phomin. 1. Teil: Die Grundbausteine. *Helvetica Chimica Acta* **1970**, *53* (7), 1797-1812.
46. Vederas, J. C.; Tamm, C., Biosynthesis of Cytochalasans .6. Mode of Incorporation of Phenylalanine into Cytochalasin-D. *Helvetica Chimica Acta* **1976**, *59* (2), 558-566.
47. Cooper, J. A., Effects of Cytochalasin and Phalloidin on Actin. *J. Cell. Biol.* **1987**, *105* (4), 1473-1478.
48. Sampath, P.; Pollard, T. D., Effects of Cytochalasin, Phalloidin, and Ph on the Elongation of Actin-Filaments. *Biochemistry* **1991**, *30* (7), 1973-1980.
49. Udagawa, T.; Yuan, J.; Panigrahy, D.; Chang, Y. H.; Shah, J.; D'Amato, R. J., Cytochalasin E, an epoxide containing Aspergillus-derived fungal metabolite, inhibits angiogenesis and tumor growth. *J. Pharmacol. Exp. Ther.* **2000**, *294* (2), 421-427.
50. Vedejs, E.; Campbell, J. B.; Gadwood, R. C.; Rodgers, J. D.; Spear, K. L.; Watanabe, Y., Synthesis of the Cytochalasin D Isoindolone Unit - Solutions to the Problem of

- Regiochemistry in N-Benzoylpyrrolinone Diels-Alder Reactions. *J. Org. Chem.* **1982**, *47* (8), 1534-1546.
51. Buchi, G.; Kitaura, Y.; Yuan, S. S.; Wright, H. E.; Clardy, J.; Demain, A. L.; Glinsuko, T.; Hunt, N.; Wogan, G. N., Structure of Cytochalasin-E, a Toxic Metabolite of *Aspergillus Clavatus*. *J. Am. Chem. Soc.* **1973**, *95* (16), 5423-5425.
52. Zhang, H.; Liu, H. B.; Yue, J. M., Organic Carbonates from Natural Sources. *Chem. Rev.* **2014**, *114* (1), 883-898.
53. Wang, F. Z.; Wei, H. J.; Zhu, T. J.; Li, D. H.; Lin, Z. J.; Gu, Q. Q., Three New Cytochalasins from the Marine-Derived Fungus *Spicaria elegans* KLA03 by Supplementing the Cultures with L- and D-Tryptophan. *Chem. Biodivers.* **2011**, *8* (5), 887-894.
54. de Gonzalo, G.; Mihovilovic, M. D.; Fraaije, M. W., Recent Developments in the Application of Baeyer-Villiger Monooxygenases as Biocatalysts. *Chembiochem* **2010**, *11* (16), 2208-2231.
55. Leisch, H.; Morley, K.; Lau, P. C. K., Baeyer-Villiger Monooxygenases: More Than Just Green Chemistry. *Chem. Rev.* **2011**, *111* (7), 4165-4222.
56. Robert, J. L.; Tamm, C., Biosynthesis of Cytochalasins .5. Incorporation of Deoxaphomin into Cytochalasin B (Phomin). *Helvetica Chimica Acta* **1975**, *58* (8), 2501-2504.
57. Vederas, J. C.; Nakashima, T. T.; Diakur, J., Detection of O-18-Label in Cytochalasin-B by C-13 Nmr. *Planta Med.* **1980**, *39* (3), 201-202.
58. Huang, T. C.; Chang, H. Y.; Hsu, C. H.; Kuo, W. H.; Chang, K. J.; Juan, H. F., Targeting therapy for breast carcinoma by ATP synthase inhibitor aurovertin B. *J. Proteome. Res.* **2008**, *7* (4), 1433-1444.

59. Steyn, P. S.; Vlegaar, R.; Wessels, P. L., Biosynthesis of the Aurovertin-B and Aurovertin-D - the Role of Methionine and Propionate in the Simultaneous Operation of 2 Independent Biosynthetic Pathways. *J. Chem. Soc. Perk. T. I* **1981**, (4), 1298-1308.
60. Ueno, Y.; Ueno, I., Isolation and Acute Toxicity of Citreoviridin, a Neurotoxic Mycotoxin of *Penicillium-Citreo-Viride* Biourge. *Jpn J. Exp. Med.* **1972**, 42 (2), 91-&.
61. Kawai, K.; Fukushima, H.; Nozawa, Y., Inhibition of Mitochondrial Respiration by Asteltoxin, a Respiratory Toxin from *Emericella-Variecolor*. *Toxicol. Lett.* **1985**, 28 (2-3), 73-77.
62. Adachi, H.; Doi, H.; Kasahara, Y.; Sawa, R.; Nakajima, K.; Kubota, Y.; Hosokawa, N.; Tateishi, K.; Nomoto, A., Asteltoxins from the Entomopathogenic Fungus *Pochonia bulbillosa* 8-H-28. *J. Nat. Prod.* **2015**, 78 (7), 1730-1734.
63. Steyn, P. S.; Vlegaar, R., Mechanistic Studies on the Biosynthesis of the Aurovertins Using O-18-Labeled Precursors. *Chem. Commun.* **1985**, (24), 1796-1798.
64. Mao, X. M.; Zhan, Z. J.; Grayson, M. N.; Tang, M. C.; Xu, W.; Li, Y. Q.; Yin, W. B.; Lin, H. C.; Chooi, Y. H.; Houk, K. N.; Tang, Y., Efficient Biosynthesis of Fungal Polyketides Containing the Dioxabicyclo-octane Ring System. *J. Am. Chem. Soc.* **2015**, 137 (37), 11904-11907.
65. Keller, N. P., Translating biosynthetic gene clusters into fungal armor and weaponry. *Nat. Chem. Biol.* **2015**, 11 (9), 671-677.



# Chapter 2

## Reversible Product Release and Recapture by a Fungal Polyketide Synthase Using a Carnitine Acyltransferase Domain

- 2.1 Genome Mining of HRPKS-cAT Enzymes
- 2.2 Heterologous Expression of Tv6-931 HRPKS-cAT
- 2.3 Enzymatic Biosynthesis of Tv6-931 HRPKS-cAT
- 2.4 Enzymatic Assays using *N*-acetylcysteamine Thioester Probes
- 2.5 Polyketide Recapture and Post-Release Catalysis
- 2.6 Supplementary Tables
- 2.7 Supplementary Figures
- 2.8 Experimental Methods
- 2.9 References

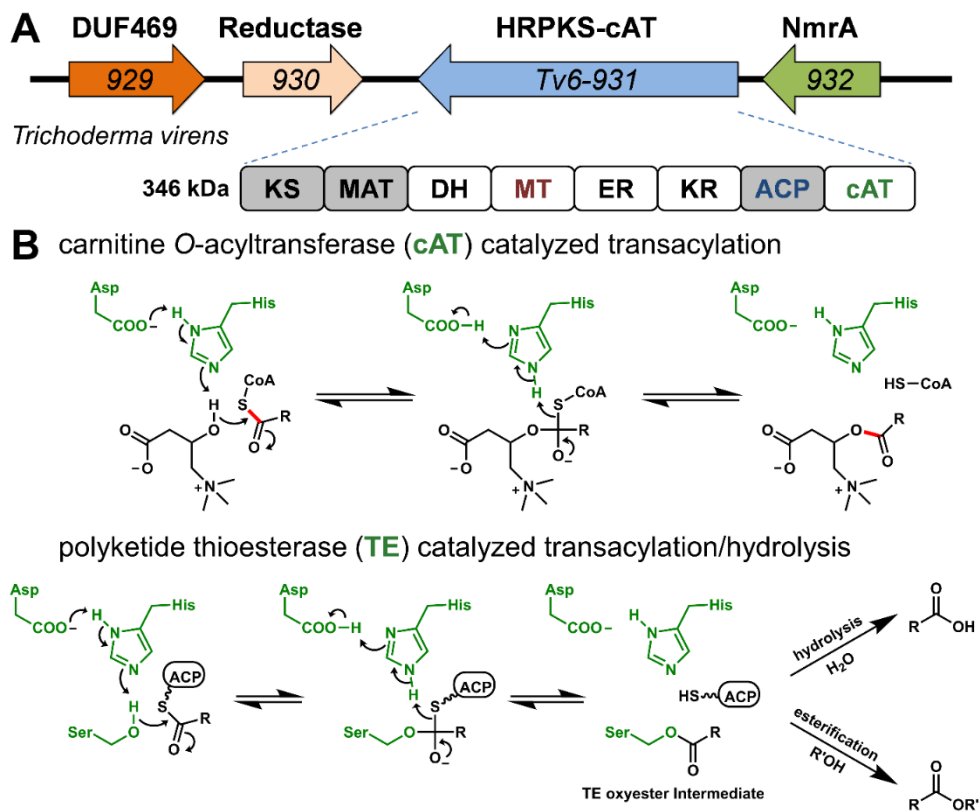
## Section 2.1 Genome Mining of HRPKS-cAT Enzymes

Fungal polyketides constitutes an important family of natural products that have significant pharmaceutical value.<sup>1</sup> The biosynthesis of reduced fungal polyketides, such as lovastatin, by the iterative highly reducing polyketide synthase (HRPKS) is a highly programmed and enigmatic process.<sup>2-3</sup> This is mainly due to the iterative and permutative use of a single set of elongation and tailoring domains to synthesize a precisely modified polyketide chain. Recent studies showed that tailoring domains, which typically include methyltransferase (MT), ketoreductase (KR), dehydratase (DH) and enoylreductase (ER), have well-defined substrate preferences towards growing polyketide intermediates, and are kinetically synchronized during each iteration.<sup>4-6</sup> Furthermore, offloading of the completed polyketide chain, which represents the termination step in the iterative process, is also well coordinated with the rest of the domains to ensure a product with the correct length and chemical features is released.<sup>7-8</sup> Notwithstanding these new insights, prediction of product structures from the large number of HRPKS enzymes identified from fungal genome sequences remains unfeasible at the current time. Therefore, discovery and understanding of new HRPKS programming features remain important objectives towards genome mining and biosynthetic engineering.

We searched through 581 sequenced ascomycetes and basidiomycetes genomes and identified ~4984 genes that are annotated as containing ketosynthase (KS) domains.<sup>9</sup> Based on phylogenetic classification (**Figure S2.1A**), we identified an uncharacterized clade of approximately ~100 members that contains a C-terminal domain with strong sequence homology to carnitine *O*-acyltransferase (cAT, **Figure 2.1A**). cAT is important in the shuttling of short to long-chain fatty acid molecules to the mitochondrial matrix by catalyzing acylation with carnitine (**Figure 1B**).<sup>10</sup> The fusion of cAT with HRPKS therefore represents a potentially new

mode of product release. Fungal HRPKS product release, either hydrolysis or transesterification, is canonically catalyzed by a separate, *in trans* thioesterase (TE) domains that uses an active site nucleophilic serine and proceeds through an TE-bound oxyester intermediate (**Figure 2.1B**).<sup>11-12</sup> In contrast, cAT domains use an active site histidine as a general base to deprotonate the carnitine hydroxyl group, followed by direct transesterification.<sup>13-14</sup> This reaction is readily reversible to yield the CoA thioester with an equilibrium constant of ~1.<sup>14-16</sup>

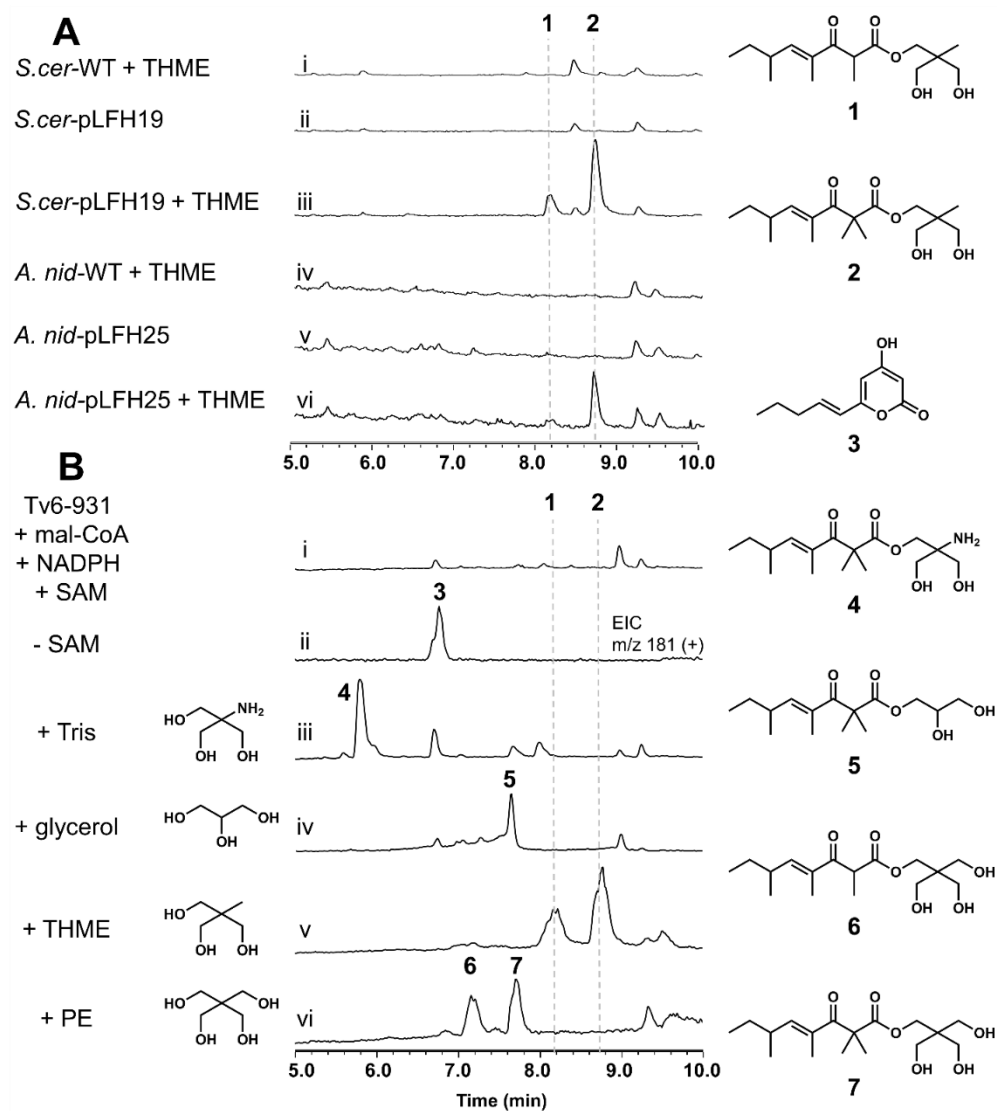
A phylogenetic tree of the ~100 members of cAT containing HRPKSs is shown in **Figure S2.1B**, among which two members are found in gene clusters that have been associated with the biosynthesis of AF-toxin and sordarin.<sup>17-18</sup> The role of the C-terminal cAT-like domain in this large group of uncharacterized HRPKSs is therefore intriguing, especially its functional relationship to the programming rules of the upstream HRPKS.



**Figure 2.1:** HRPKSs containing a C-terminal carnitine *O*-acyltransferase like domain. (A) Tv6-931 is a HRPKS-cAT; (B) Mechanistic comparison of reactions catalyzed by cAT and the canonical thioesterase (TE).

## Section 2.2 Heterologous Expression of Tv6-931 HRPKS-cAT

One member of this family of HRPKS is Tv6-931 encoded in the genome of *Trichoderma virens* (**Figure 2.1A**), and is well-conserved across several *Trichoderma* species. The HRPKS contains all of the reductive domains, as well as a MT domain. Genes surrounding the HRPKSs do not provide additional insights into a plausible product structure. Due to difficulties associated with working with *T. virens*, we used heterologous expression to investigate the potential product. The continuous coding sequence was synthesized, assembled, and cloned into either a yeast 2 $\mu$  vector (pLFH19) or an *Aspergillus* vector (pLFH25) (**Table S2.2**). Transformation of the vectors into the corresponding *Saccharomyces cerevisiae* (BJ5464-NpgA)<sup>19</sup> or *Aspergillus nidulans* A1145, however, did not yield any new products (**Figure 2.2A, ii and v**). Addition of genes surrounding Tv6-931 into yeast also did not lead to new metabolites.



**Figure 2.2:** Functional characterization of Tv6-931 (Total Ion Count). (A) Heterologous expression in *S. cerevisiae* and *A. nidulans* lead to **1** and **2** in the presence of THME. (B) Biochemical assays using purified enzyme revealed the polyketide chain structure and list of offloading nucleophiles.

## Section 2.3 Enzymatic Biosynthesis Using Tv6-931 HRPKS-cAT

We then turned to enzymatic assays using Tv6-931 purified by Ni-NTA chromatography to determine its function (**Figure S2.2**; See Section 2.6 Experimental Methods). Assay of the enzyme in the presence of only malonyl-CoA yielded triacetic acid lactone, indicating the KS, malonyl-CoA:ACP transacylase (MAT) and acyl carrier protein (ACP) domains are all functional. Addition of NADPH led to synthesis of the conjugated  $\alpha$ -pyrone **3** (**Figure 2.2B**, ii),<sup>8, 19</sup> indicating the reductive and dehydrating domains (KR, DH and ER) are active. Unexpectedly, addition of SAM led to no detectable polyketide products (**Figure 2.2B**, i). The absence of any  $\alpha$ -pyrone product such as **3** suggested that the MT domain may be active, but the offloading mechanism was not fully reconstituted. Addition of carnitine and related compounds did not lead to formation of esterified products (**Figure S2.3**). Serendipitously, when repeating the same assay (+NADPH, +SAM) in Tris buffer, or PBS buffer containing glycerol, we detected formation of two new products, **4** (MW 315) and **5** (MW 286), respectively (**Figure 2.2B**, iii and iv, and **Figure S2.5**). Subtracting the mass of either Tris or glycerol from the molecular weights of **4** and **5**, respectively, yielded the same residual molecular weight of 195 (C<sub>12</sub>O<sub>2</sub>H<sub>19</sub>), hinting that the same polyketide fragment is offloaded by Tris or glycerol. To gain insights into polyketide length, we performed the assay in the presence of [2-<sup>13</sup>C]-malonate and MatB to generate labeled [2-<sup>13</sup>C]-malonyl-CoA *in situ*.<sup>20</sup> A mass increase of 4 mu in **4** and **5** was observed (**Figure S2.4**), indicating the product is a tetraketide (C<sub>8</sub>). The remaining four carbon atoms in the molecule must therefore be derived from SAM, yet only three possible C $\alpha$ -methylation positions are present in a tetraketide.

We assayed additional amino and hydroxyl containing substrates as offloading nucleophiles using the *in vitro* assay, including linear and branched alcohols (**Figure S2.3**). While several substrates supported the formation of the polyketide adduct to different degrees, two polyol compounds: 1,1,1-tris(hydroxymethyl)ethane (THME) and pentaerythritol (PE) led to significantly improved product turnover in **2** and **7**, respectively (**Figure 2.2B**, iv and v). In each of these assays, we also observed a second product that is 14 mu less, **1** and **6**. <sup>13</sup>C-malonate assay revealed that both **1** and **6** contain one less SAM-derived methyl group. Based on product yield, the order of releasing nucleophile preference by Tv6-931 is THME~PE>Tris>glycerol (**Figure S2.3**).

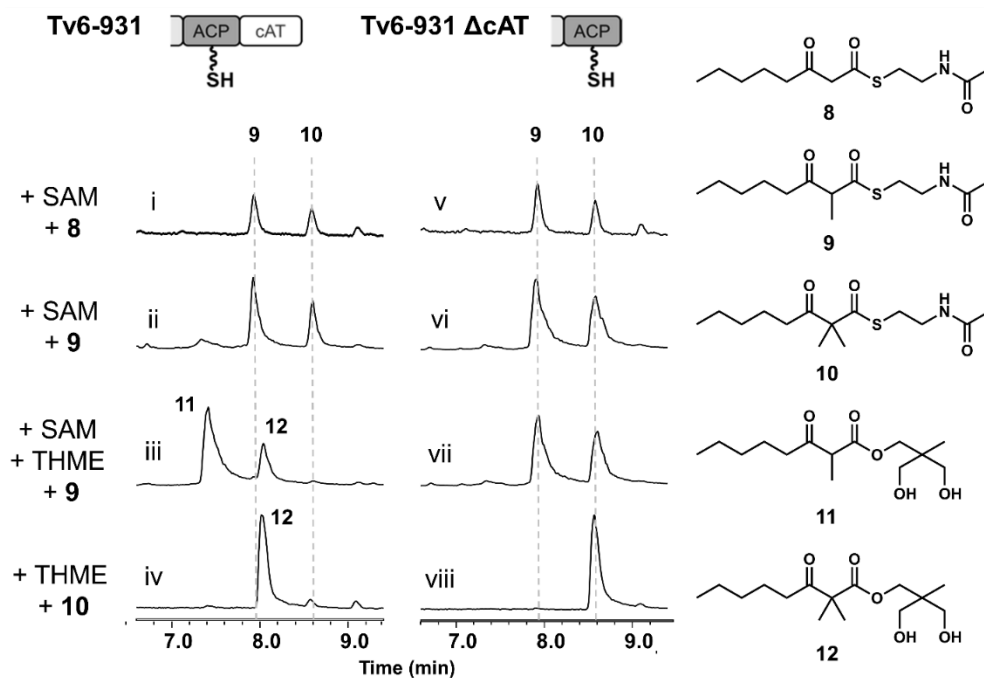
We repeated the heterologous expression experiment with the addition of 1% releasing substrate. Both yeast and *A. nidulans* produced predominantly **2** and minor **1** upon addition of THME (**Figure 2.2A**, iii and vi). Similarly, addition of glycerol or PE yielded **5** or **6** and **7**, respectively (**Figure S2.6**). We purified **1**, **2**, and **5** from yeast and elucidated the structures with NMR (See Section 2.7). As expected, both **1** and **2** are tetraketides esterified with THME, while **5** is esterified through a primary hydroxyl group of glycerol. The polyketide portion of **1** is 2,4,6-trimethyl-3-oxooct-4-enoate, with three SAM-derived methyl groups. The polyketide portion of **2** and **5**, however, contains a *gem*-dimethyl at the last  $\alpha$ -carbon. This is consistent with the suggestion of four SAM-derived methyl group from <sup>13</sup>C labeling studies described above. The *gem*-dimethyl group in polyketides are relatively rare, most examples are from bacterial polyketides,<sup>21-29</sup> including the anticancer drugs epothilone and precursor of epoxyketone.<sup>30-31</sup> Keasling and coworkers showed that MT domains in these PKSs can *gem*-dimethylate either the  $\beta$ -keto polyketide intermediate, or malonyl-ACP prior to KS-catalyzed condensation.<sup>28</sup> Keatinge-Clay and coworkers showed the GphH MT domain from the gephyronic acid biosynthetic



pathway can *gem*-dimethylate acetoacetyl-*S-N*-acetylcysteamine (~1 turnover in 72 hours).<sup>29</sup> Fungal polyketides containing *gem*-dimethyl groups are rarer,<sup>32-33</sup> with no responsible MT identified so far.

## Section 2.4 Enzymatic Assays using *N*-acetylcysteamine Thioester Probes

To investigate the activity of the Tv6-931 MT domain, we used simplified triketide (3-oxohexanoyl-SNAC) and tetraketide (3-oxooctanoyl-SNAC, **8**) to perform *in vitro* methylation assays. Both the full length Tv6-931 and a truncated version lacking the cAT domain (Tv6-931  $\Delta$ cAT) were used to probe any potential role of the cAT domain in the successive C $_{\alpha}$ -methylation reactions. In the presence of the triketide substrate, Tv6-931 catalyzed only one methylation step to yield 2-methyl-3-oxohexanoyl-SNAC (**Figure S2.7**). In contrast, when incubated with **8**, we detected the conversion to both the 2-methyl-3-oxooctanoyl-SNAC **9** (matched to standard), and a new compound with mass corresponding to 2,2-dimethyl-3-oxooctanoyl-SNAC **10** (**Figure 2.3**, i). Direct assay of Tv6-931 with **9** also led to formation of **10** (**Figure 2.3**, ii). The identity of **10** was confirmed by NMR characterization (**Table S2.7**). The truncated Tv6-931  $\Delta$ cAT also catalyzed the methylation reactions (**Figure 2.3**, v and vi). These results therefore confirmed that the MT domain is specifically programmed to differentiate between triketide and tetraketide substrates. Independent of the cAT domain, MT can indeed catalyze two successive C $_{\alpha}$ -methylation reactions on the tetraketide substrate **8** to **10**. The catalytic efficiency of ~1 turnover per hour, albeit still low as expected with SNAC substrates, is significantly higher than that reported for GphH.<sup>13</sup>



**Figure 2.3:** Characterization of MT and cAT domains of Tv6-931 using SNAC substrates (Total Ion Count). Reaction conditions: 15  $\mu$ M enzyme (Tv6-931 or Tv6-931  $\Delta$ cAT), 5 mM SAM, 2 mM SNAC substrate, 5 mM offloading nucleophile when added, PBS buffer, pH 7.4, 16 hours of reaction.

We also recapitulated the activities of the cAT domain using the SNAC substrates. When THME, SAM and **9** were added to Tv6-931, we observed a majority of **9** readily converted to the ester **11** with a small amount to the 2,2-dimethyl-3-oxooctanoyl-THME ester **12** (**Figure 2.3**, iii). When **10** was directly added to Tv6-931 with THME, nearly complete conversion to **12** was seen (**Figure 2.3**, iv). In contrast, the truncated Tv6-931  $\Delta$ cAT enzyme was unable to transesterify the products as THME esters (**Figure 2.3**, vii and viii). Addition of the standalone recombinant cAT domain *in trans* at increasing concentrations progressively restored the release of ester products (**Figure S2.9**). Collectively, these experiments confirmed the role of the cAT domain in Tv6-931 is responsible for product offloading. Comparison of the active site residues of Tv6-931 cAT domain with authentic carnitine-*O*-acyltransferases revealed conservation of the catalytic histidine and residues that coordinate to a water molecule in the active site (**Figure S2.9-S2.10**). A serine residue proposed to stabilize the oxyanion is also conserved. Residues that interact with carnitine are not conserved, expected since carnitine was not accepted as an offloading nucleophile.

Despite confirming the activity of the cAT domain in Tv6-931, it remains unclear what is the chemical strategy behind Nature's use of the *in cis* cAT domain as an offloading catalyst, instead of the more canonical TE domain. A particular important feature of the acyl transfer reaction between acyl CoA-thioesters and carnitine catalyzed by carnitine-*O*-acyltransferase is that the reaction is readily reversible, which allows shuttling of acyl-CoA across the mitochondria membrane.<sup>14-16</sup> We therefore hypothesized that the reversibility feature could be advantageous to the programming rules of Tv6-931, especially with respect to the *gem*-dimethylation activities of the MT domain.

## Section 2.5 Polyketide Recapture and Post-Release Catalysis

To test the reversibility of the Tv6-931 cAT domain, we performed a post-release transesterification assay in which the THME esters **1** or **2** is mixed with an equal molar amount of the free alcohol PE, a kinetically comparable substrate to cAT as THME. Interestingly, we observed the transesterification readily occurs with the final distributions between THME (**1** or **2**) and PE esters (**6** or **7**) to be close to 1:1 (**Figure 2.4**, ii and iii). In contrast, the reaction does not occur when Tv6-931  $\Delta$ cAT was used (**Figure 2.4**, v and vi). However, the standalone cAT does not catalyze the interconversion, suggesting the reversible reaction is dependent on the HRPKS, namely the thiol-terminating phosphosphopantetheinyl arm of the ACP domain. This was readily confirmed by expressing the *apo* form of Tv6-931 using a yeast strain without an integrated phosphopantetheinyl transferase NpgA. No transesterification can be detected using the *apo* enzyme (**Figure 2.4**, viii and ix).

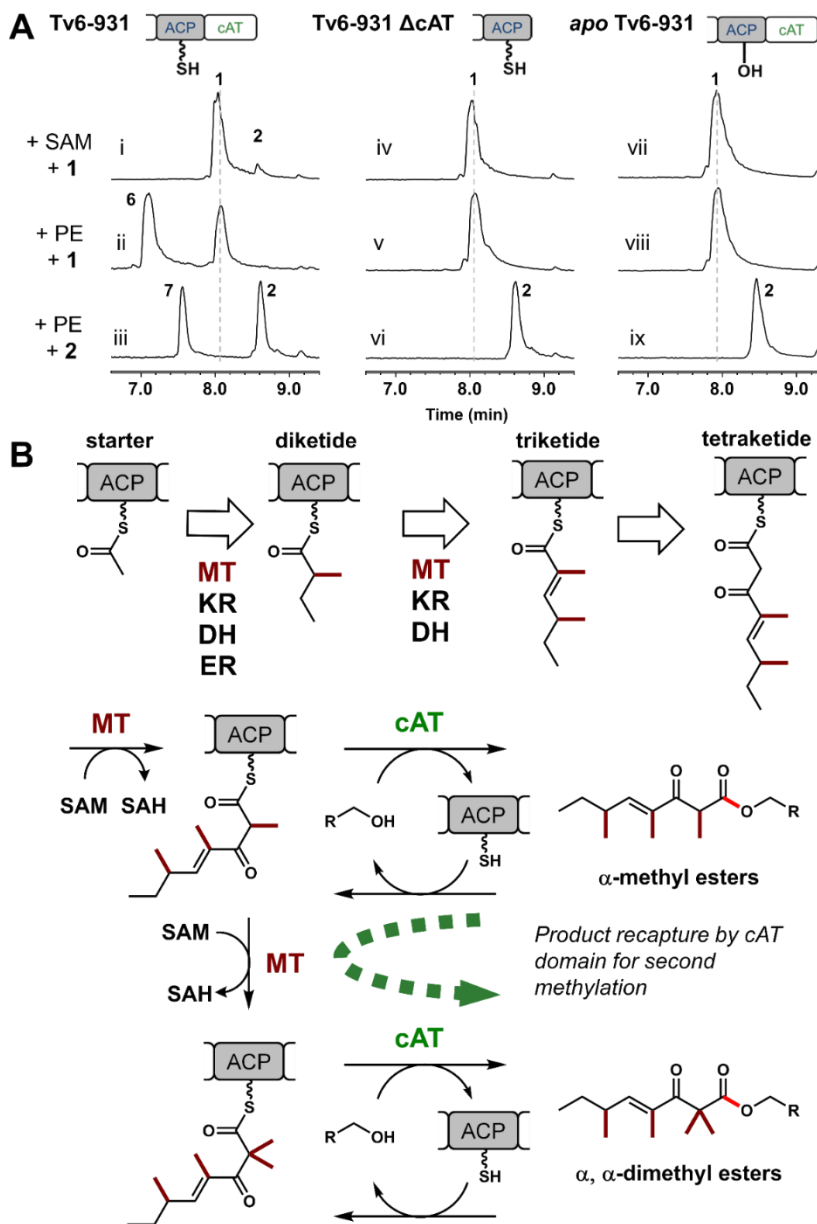
Our results demonstrate the reversible release and recapture of the polyketide product back on to the ACP domain, catalyzed by the appended cAT domain (**Figure 2.4B**). The involvement of ACP-SH to form the reloaded acyl thioester is consistent with the reaction mechanism of canonical carnitine-*O*-acyltransferase. While this may appear to be a futile cycle in terms of polyketide biosynthesis, we propose that assembly line reentry may be a strategy to recycle prematurely offloaded 2-monomethyl products back on to the ACP and to allow the second methylation to take place. One observation from *in vitro* assays suggests such a kinetic competition between offloading by cAT and the second methylation step by MT (**Figure 2.2B**, iii-vi): 2-monomethyl products such as **1** and **6** are more abundant with better offloading nucleophiles such as PE and THME; while 2-dimethyl products are major products for slower substrates such as Tris and glycerol. While we do not yet know the natural offloading substrate

of Tv6-931, it is likely to be a kinetically competent one that will compete with the second methylation step. To test the product recapture hypothesis shown in **Figure 2.4B**, we directly added **1** to different versions of Tv6-931 in the presence of SAM. We were able to observe the conversion of **1** to **2** (**Figure 2.4A**, i) only in the presence of *holo* Tv6-931. The identity of the product was verified using crude <sup>1</sup>H-NMR and compared to purified **2** (**Figure S2.11**). In contrast, the  $\Delta$ cAT and *apo* versions did not support any conversion of **1** to **2**. We verified the MT and cAT domains of the *apo* form were active using **9** and **10** (**Figure S2.12**), thereby confirming the lack of activity in the *apo* protein is solely due to the absence of the phosphopantetheinyl prosthetic group (**Figure S2.13**).

The post-release methylation of **1** to **2** via cAT and ACP provides a complementary route to the direct iterative methylation to produce the *gem*-dimethyl polyketide **2** (**Figure 2.4B**). Given that the methylation steps are essentially irreversible, the recycling mechanism facilitated should lead to the *gem*-dimethyl products as the final products. This mechanism may explain that under *in vivo* conditions in *A. nidulans* (**Figure 2.2A**, vi), the predominant product was **2**, while under *in vitro* conditions where equilibrium favors the offloaded product, a significant fraction of products remained the monomethylated **1**.

In conclusion, the cAT domain studied here can catalyze reversible release and recapturing of the polyketide product, consistent with the mechanism of carnitine-*O*-acyltransferase in primary metabolism. A remaining unanswered question is the natural nucleophile of the cAT domain, as the substrates we discovered here are not expected to be present in *T. virens*. It is evident that the natural substrate should be unique to *T. virens*, as it is absent in both *S. cerevisiae* and *A. nidulans*. Active metabolomics work is ongoing to discover the authentic natural product of Tv6-931. Hence, our work illustrates a potential bottleneck of

genome mining efforts when a required substrate present in the natural host is unavailable in model heterologous organisms.



**Figure 2.4:** Reversible product offloading by the cAT domain in Tv6-931. (A) Assay demonstrating post-release methylation; (B) Proposed mechanism and interplay between the cAT and MT domains.

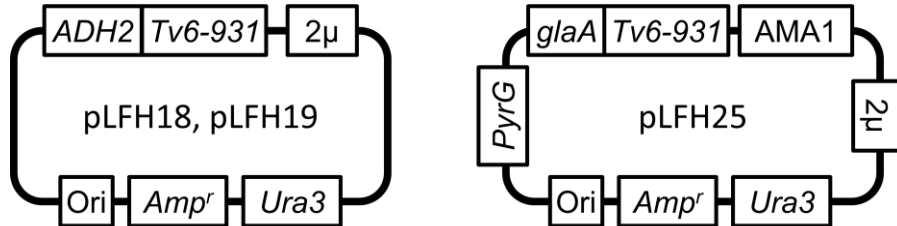
## Section 2.6: Supplementary Tables

**Table S2.1: DNA Primers Used**

<b>Primer Name</b>	<b>Sequence of Primer (5' to 3')</b>
pXW55-N-Term-Ura-F	AGCAGCACGTTCCCTTATATG
pXW55-N-8xHis-R	ACTAGTGTGATGGTGTGATGGTGTGATGGTGTGATGGCTAGCCATATGGT ATTACG
pXW55-C-Term-F	TAATTTAAATGACAAATTTGTGCG
pXW55-C-Term-Ura-R	ATGTGCGAAAGCTACATATAAG
Tv6-931-N-His- HRPKS-P1-F	AATACCATATGGCTAGCCATCACCATCACCATCACCATCACACTA GTATGTCTCCACAAG
Tv6-931-HRPKS-P1-R	AGTCAACATATTGGAAAGATATG
Tv6-931-HRPKS-P2-F	CGAAGTAATCATATCTTTCCAA
Tv6-931-HRPKS-P2-R	AAATCACCGAAAATCCTTTTCA
Tv6-931-HRPKS-P3-F	GAGATTGTTGAAAAGGATTTTCG
Tv6-931-HRPKS-cAT- Native-R	TGAAGGCATCGGTCCGCACAAATTTGTCATTTAAATTAACAACA TTGATCTCAATGATA
Tv6-931-HRPKS- standalone-R	CGCACAAATTTGTCATTTAAATTAGTGATGGTGTGATGGTGTGATGCAC GGTAACGAAGGCAGC
Tv6-931-cAT- standalone-F	TATATAGCTAGCACCAAGTCTTCTAATTTGGTC
Tv6-931-cAT- standalone-R	TATATAGCGGCCCGCTAAACAACATTGATCTCAATGA
HRPKS-cAT-gDNA- pYTU-P1-F	ATCCCCAGCATCATTACACCTCAGCATTAAATTAAGGCGGCCGCAT GTCACCCCAAGAGCC
HRPKS-cAT-gDNA- pYTU-P1-R	AGCATCTGAGAACCGAACTG
HRPKS-cAT-gDNA- pYTU-P2-F	AGACAACCTGGTATCCCCAG
HRPKS-cAT-gDNA- pYTU-P2-R	CGTAGAACAGATGAACGCTG
HRPKS-cAT-gDNA- pYTU-P3-F	TCAAATTCTTCAGGGAATCAC
HRPKS-cAT-gDNA- pYTU-P3-R	CTAGTGCTTGCTCATTATAATAG
HRPKS-cAT-gDNA- pYTU-P4-F	AGTCCAGCAGGACTATTATAATGAGCAAGCACTAGATA
HRPKS-cAT-gDNA- pYTU-P4-R	AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCCGCCTTAGATGGGT AATATAATAAGTTGG

**Table S2.2: DNA Plasmids Used**

Plasmid Name	Plasmid Backbone	Description of Plasmid
pTMC-1-106	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -FLAG-Tv6-931-C-6xHis gene under the <i>ADH2</i> promoter
pLFH17	pET-28a	pET <i>E. coli</i> plasmid with <i>N</i> -6xHis-Tv6-931 cAT gene under the T7 promoter
pLFH18	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -8xHis-Tv6-931 $\Delta$ cAT gene under the <i>ADH2</i> promoter
pLFH19	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -8xHis-Tv6-931 gene under the <i>ADH2</i> promoter
pLFH25	pYTU	AMA1 <i>A. nid.</i> expression plasmid with Tv6-931 gene under the <i>glaA</i> promoter



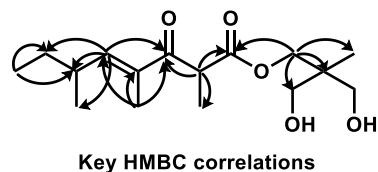
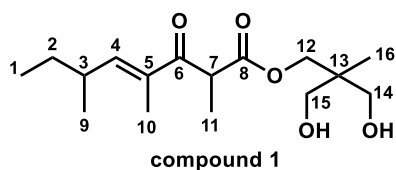
**Table S2.3: Expression Strains Used**

Strain Name	Organism	Description of Expression Strain
<i>E.coli</i> -pLFH17	<i>E. coli</i>	<i>E. coli</i> SoluBL21™ strain expressing the pLFH17 plasmid
<i>S.cer</i> -pLFH18	<i>S. cer.</i>	<i>S. cer.</i> BJ5464-npgA strain expressing the pLFH18 plasmid
<i>S.cer</i> -pLFH19	<i>S. cer.</i>	<i>S. cer.</i> BJ5464-npgA strain expressing the pLFH19 plasmid
<i>S.cer</i> -pLFH19- <i>apo</i>	<i>S. cer.</i>	<i>S. cer.</i> BJ5464- <i>apo</i> strain expressing the pLFH19 plasmid
<i>A.nid</i> -pLFH25	<i>A. nid</i>	<i>A. nid.</i> A1145 strain expressing the pLFH25 plasmid



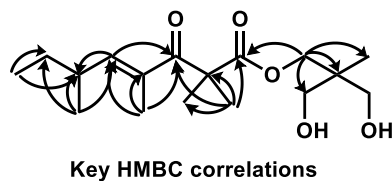
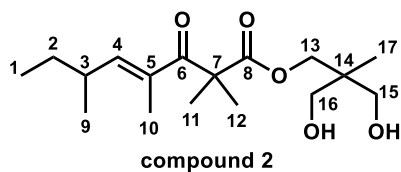
**Table S2.4: NMR Data of Compound 1**<sup>1</sup>H NMR spectrum (500 MHz), <sup>13</sup>C NMR spectrum (125 MHz), CDCl<sub>3</sub>

Position	$\delta_{\text{H}}$ , mult ( <i>J</i> in Hz)	$\delta_{\text{C}}$	COSY
1	0.87, t (7.5)	12.03/12.07	H2
2	1.36, m, overlap 1.47, m	29.74/29.79	H1
3	2.52, m	35.65/35.66	H4, H9
4	6.43, d (9.6)	150.57/150.59	H3
5		135.09/135.12	
6		199.21/199.34	
7	4.24, q (7.1)	47.00	H11
8		172.36/172.39	
9	1.04, d (6.7)	19.72	
10	1.81, s	11.98/12.00	
11	1.39, d (7.1)	14.64/14.72	H7
12	4.12, d (12.6) 4.30, d (11.3)	67.43	
13		41.04/41.07	
14	3.53, m, overlap	67.71/67.73 <sup>a</sup>	
14-OH	2.74, m, overlap		H14
15	3.53, m, overlap	67.97/67.99 <sup>a</sup>	
15-OH	2.74, m, overlap		H15
16	0.79, s	16.94/16.96	

Note: <sup>a</sup> Assignments denoted by superscript letters are interchangeable.

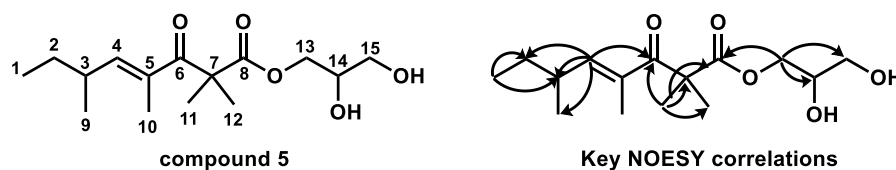
**Table S2.5: NMR Data of Compound 2**<sup>1</sup>H NMR spectrum (500 MHz), <sup>13</sup>C NMR spectrum (125 MHz), CDCl<sub>3</sub>

Position	$\delta_{\text{H}}$ , mult ( <i>J</i> in Hz)	$\delta_{\text{C}}$	COSY
1	0.85, t (7.4)	12.74	H2
2	1.30, m 1.41, m	29.69	
3	2.45, m	35.47	H4, H9
4	6.12, d (9.5)	148.50	H3
5		133.42	
6		199.38	
7		53.10	
8		176.49	
9	0.99, d (6.6)	19.73	H3
10	1.79, s	12.02	
11	1.44 <sup>a</sup> , s	24.92 <sup>b</sup>	
12	1.45 <sup>a</sup> , s	24.87 <sup>b</sup>	
13	4.14, m	67.52	
14		40.85	
15	3.52, m, overlap	67.85 <sup>c</sup>	
15-OH	2.56, t (5.9)		H15
16	3.52, m, overlap	67.87 <sup>c</sup>	
16-OH	2.56, t (5.9)		H16
17	0.79, s	16.91	

Note: <sup>a-c</sup> Assignments denoted by superscript letters are interchangeable.

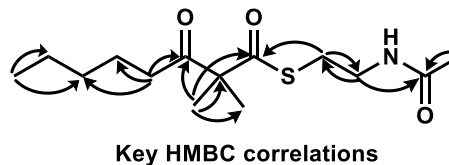
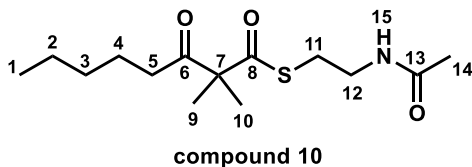
**Table S2.6: NMR Data of Compound 5**<sup>1</sup>H NMR spectrum (500 MHz), <sup>13</sup>C NMR spectrum (125 MHz), CDCl<sub>3</sub>

Position	$\delta_{\text{H}}$ , mult ( <i>J</i> in Hz)	$\delta_{\text{C}}$	COSY
1	0.84, t (7.5)	12.77	H2
2	1.30, m 1.45, m, overlap	29.69	H1, H3
3	2.47, m, overlap	35.48	H2, H4, H9
4	6.12, d (9.6)	148.58	H3
5		133.42	
6		199.49	
7		53.00	
8		175.87	
9	0.98, d (6.7)	19.70/19.72	H3
10	1.79, s	12.00	
11	1.44 <sup>a</sup> , s	24.80 <sup>b</sup>	
12	1.45 <sup>a</sup> , s	24.83 <sup>b</sup>	
13	4.14, m	66.14/66.22	H14
14	3.88, brs	70.04	H13, H15
14-OH	2.02 <sup>c</sup>		
15	3.55, m 3.65, m	63.40	H14
15-OH	2.47 <sup>c</sup>		

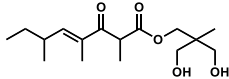
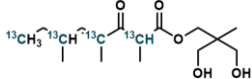
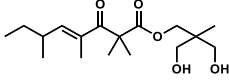
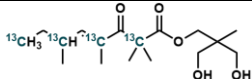
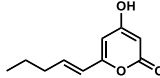
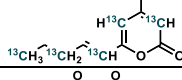
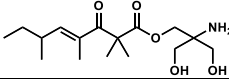
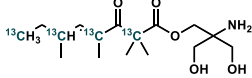
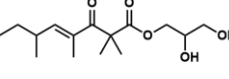
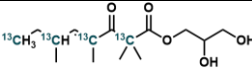
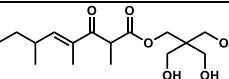
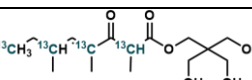
Note: <sup>a-c</sup> Assignments denoted by superscript letters are interchangeable.

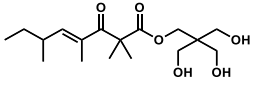
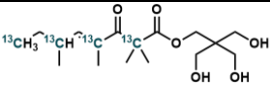
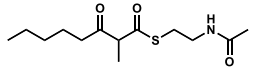
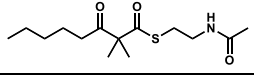
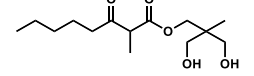
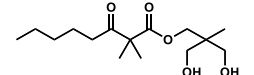
**Table S2.7: NMR Data of Compound 10**<sup>1</sup>H NMR spectrum (500 MHz), <sup>13</sup>C NMR spectrum (125 MHz), CDCl<sub>3</sub>

Position	$\delta_{\text{H}}$ , mult ( <i>J</i> in Hz)	$\delta_{\text{C}}$	COSY
1	0.88, t (7.2)	14.07	H2
2	1.27, m, overlap	22.62	
3	1.27, m, overlap	31.40	
4	1.57, m	23.65	H3, H5
5	2.45, t (7.3)	38.18	H4
6		207.90	
7		63.34	
8		202.03	
9	1.43, s, overlap	22.43	
10	1.43, s, overlap	22.43	
11	3.07, t (6.4)	28.78	H12
12	3.44, q (6.1)	39.55	H11, H15
13		170.37	
14	1.97, s	23.37	
15-NH	5.78, brs		H12

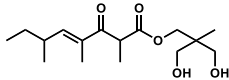
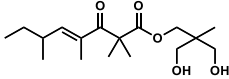
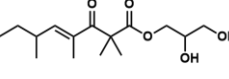
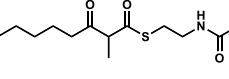
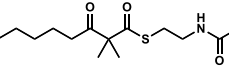


**Table S2.8: HPLC-MS Mass Data of Compounds**

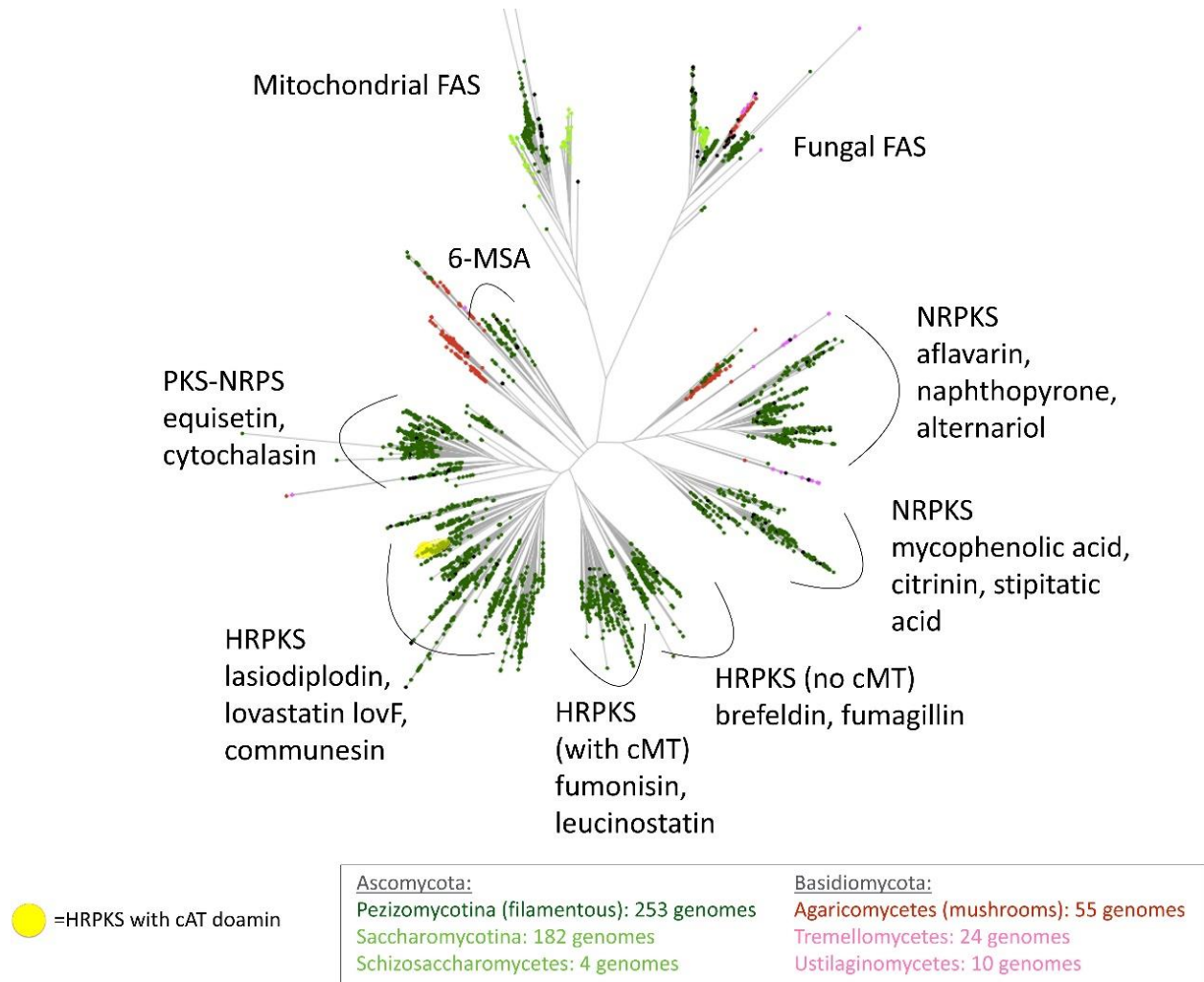
Molecule	Compound Structure	Chemical Formula	Adducts Found	Expected Mass (m/z)	Mass Found (m/z)
<b>1</b>		C <sub>16</sub> H <sub>28</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	301	301
			[M+Na] <sup>+</sup>	323	323
			[M+H-H <sub>2</sub> O] <sup>+</sup>	283	283
			[M-H] <sup>-</sup>	299	299
<sup>13</sup> C-labeled <b>1</b>		<sup>13</sup> C <sub>4</sub> C <sub>12</sub> H <sub>28</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	305	305
			[M+Na] <sup>+</sup>	327	327
			[M+H-H <sub>2</sub> O] <sup>+</sup>	287	287
<b>2</b>		C <sub>17</sub> H <sub>30</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	315	315
			[M+Na] <sup>+</sup>	337	337
			[M+H-H <sub>2</sub> O] <sup>+</sup>	297	297
<sup>13</sup> C-labeled <b>2</b>		<sup>13</sup> C <sub>4</sub> C <sub>13</sub> H <sub>30</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	319	319
			[M+Na] <sup>+</sup>	341	341
			[M+H-H <sub>2</sub> O] <sup>+</sup>	301	301
			[M-H] <sup>-</sup>	317	317
<b>3</b>		C <sub>10</sub> H <sub>12</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	181	181
			[M-H] <sup>-</sup>	179	179
<sup>13</sup> C-labeled <b>3</b>		<sup>13</sup> C <sub>5</sub> C <sub>5</sub> H <sub>12</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	186	186
			[M-H] <sup>-</sup>	184	184
<b>4</b>		C <sub>16</sub> H <sub>29</sub> NO <sub>5</sub>	[M+H] <sup>+</sup>	316	316
			[M+Na] <sup>+</sup>	338	338
<sup>13</sup> C-labeled <b>4</b>		<sup>13</sup> C <sub>4</sub> C <sub>12</sub> H <sub>29</sub> NO <sub>5</sub>	[M+H] <sup>+</sup>	320	320
			[M+Na] <sup>+</sup>	342	342
<b>5</b>		C <sub>15</sub> H <sub>26</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	287	287
			[M+Na] <sup>+</sup>	309	309
			[M+H-H <sub>2</sub> O] <sup>+</sup>	269	269
<sup>13</sup> C-labeled <b>5</b>		<sup>13</sup> C <sub>4</sub> C <sub>11</sub> H <sub>26</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	291	291
			[M+Na] <sup>+</sup>	313	313
			[M+H-H <sub>2</sub> O] <sup>+</sup>	273	273
<b>6</b>		C <sub>16</sub> H <sub>28</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	317	317
			[M+Na] <sup>+</sup>	339	339
			[M+H-H <sub>2</sub> O] <sup>+</sup>	299	299
<sup>13</sup> C-labeled <b>6</b>		<sup>13</sup> C <sub>4</sub> C <sub>12</sub> H <sub>28</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	321	321
			[M+Na] <sup>+</sup>	343	343
			[M+H-H <sub>2</sub> O] <sup>+</sup>	303	303

<b>7</b>		$C_{17}H_{30}O_6$	$[M+H]^+$	331	331
			$[M+Na]^+$	353	353
			$[M+H-H_2O]^+$	313	313
<sup>13</sup> C-labeled <b>7</b>		$^{13}C_4C_{13}H_{30}O_6$	$[M+H]^+$	335	335
			$[M+Na]^+$	357	357
			$[M+H-H_2O]^+$	317	317
<b>9</b>		$C_{13}H_{23}NO_3S$	$[M+H]^+$	274	274
			$[M+Na]^+$	296	296
			$[M-H]^-$	272	272
<b>10</b>		$C_{14}H_{25}NO_3S$	$[M+H]^+$	288	288
			$[M+Na]^+$	310	310
<b>11</b>		$C_{14}H_{26}O_5$	$[M+H]^+$	275	275
			$[M+Na]^+$	297	297
			$[M+H-H_2O]^+$	257	257
<b>12</b>		$C_{15}H_{28}O_5$	$[M+H]^+$	289	289
			$[M+Na]^+$	311	311
			$[M+H-H_2O]^+$	271	271

**Table S9: HR-MS Mass Data of Compounds**

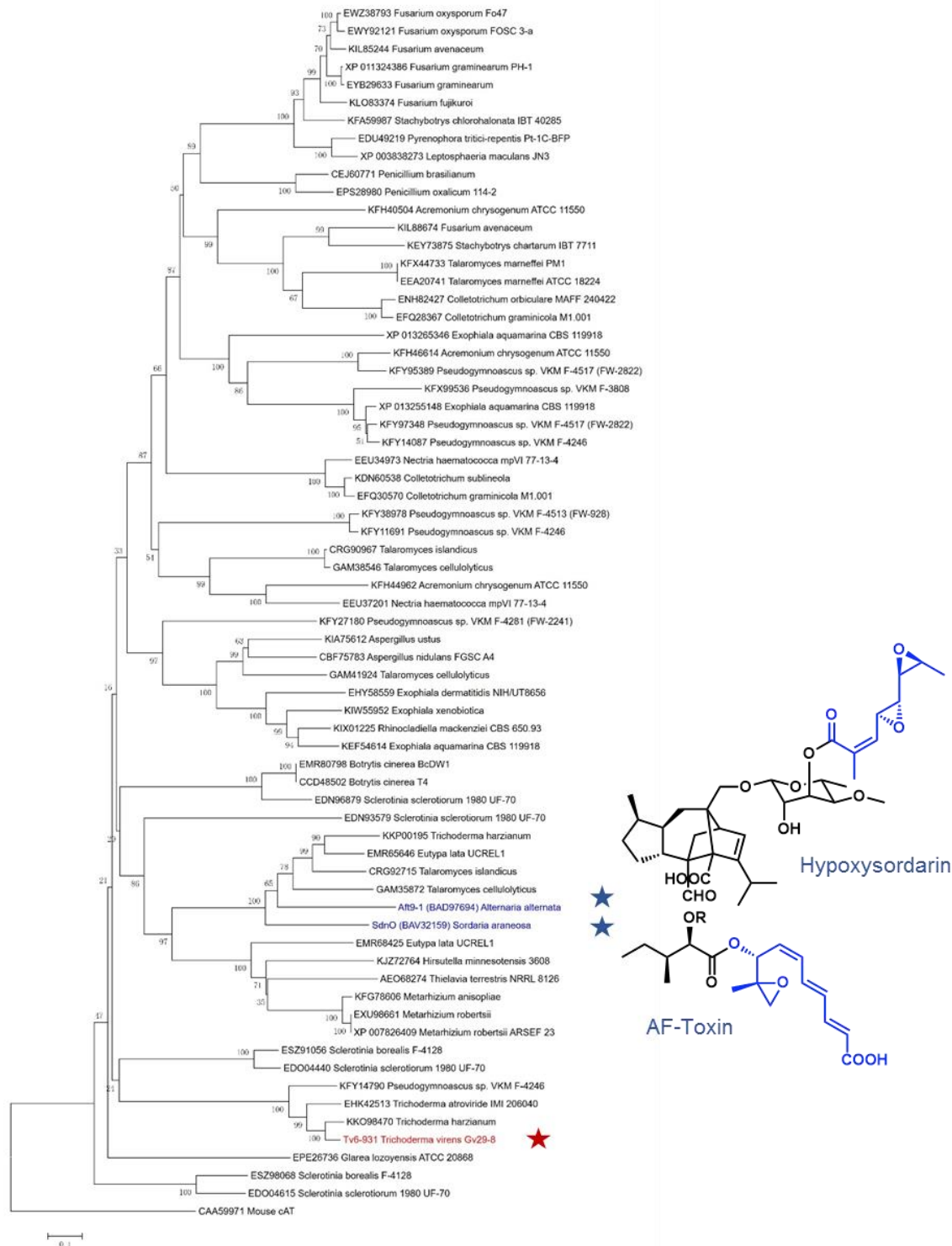
Molecule	Compound Structure	Chemical Formula	Adducts Found	Expected Mass (m/z)	Mass Found (m/z)
<b>1</b>		C <sub>16</sub> H <sub>28</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	301.20150	301.19953
			[M+NH <sub>4</sub> ] <sup>+</sup>	318.22805	318.22589
			[M+H-H <sub>2</sub> O] <sup>+</sup>	283.19093	283.18879
			[M-H] <sup>-</sup>	299.18585	299.18523
<b>2</b>		C <sub>17</sub> H <sub>30</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	315.21715	315.21494
			[M+NH <sub>4</sub> ] <sup>+</sup>	332.24370	332.24145
			[M+H-H <sub>2</sub> O] <sup>+</sup>	297.20658	297.20428
			[M-H] <sup>-</sup>	313.20150	313.20099
<b>5</b>		C <sub>15</sub> H <sub>26</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	287.18585	287.18488
			[M+NH <sub>4</sub> ] <sup>+</sup>	304.21240	304.21152
			[M+H-H <sub>2</sub> O] <sup>+</sup>	269.17528	269.17399
<b>9</b>		C <sub>13</sub> H <sub>23</sub> NO <sub>3</sub> S	[M+H] <sup>+</sup>	274.14769	274.14785
			[M+NH <sub>4</sub> ] <sup>+</sup>	291.17424	291.17454
			[M-H] <sup>-</sup>	272.13204	272.13326
<b>10</b>		C <sub>14</sub> H <sub>25</sub> NO <sub>3</sub> S	[M+H] <sup>+</sup>	288.16334	288.16364
			[M+NH <sub>4</sub> ] <sup>+</sup>	305.18989	305.19035
			[M-H] <sup>-</sup>	286.14769	286.14898

## Section 2.7: Supplementary Figures

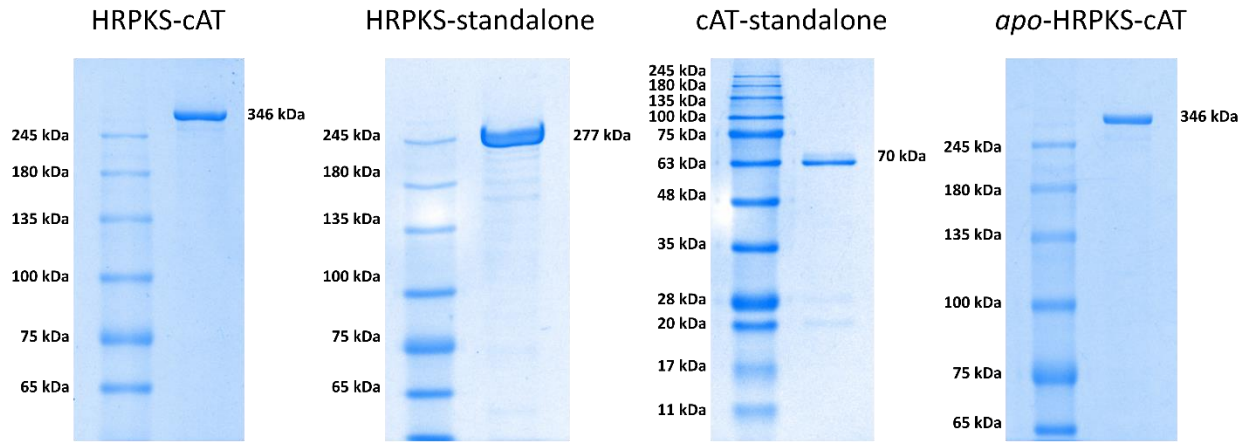


**Figure S2.1 Fungal Phylogenetic Trees A) Fungal KS Phylogenetic Tree:** Phylogenetic tree constructed using the KS domains of the PKS core enzymes identified by homology in sequenced fungal genomes from GenBank.<sup>9</sup> All the HRPKSs with cAT domains (highlight in yellow color in the figure) fall into a subgroup within one of the HRPKS clades.



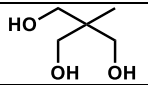
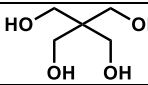
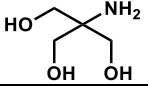
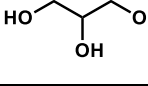
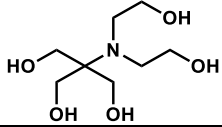
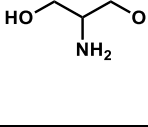
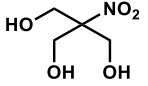
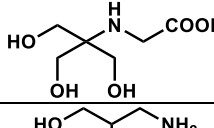
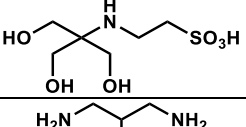
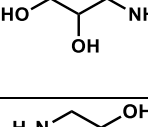
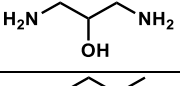
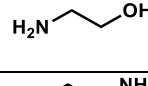
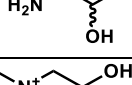
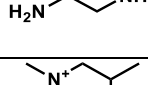
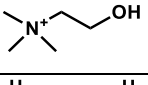
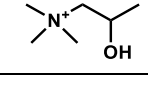
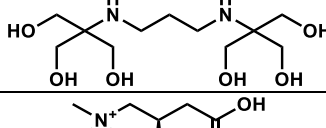
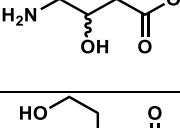
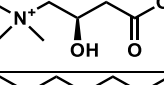
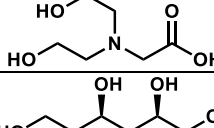
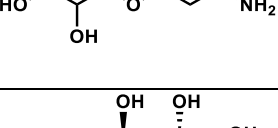
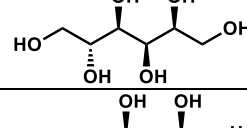
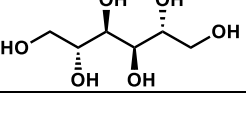
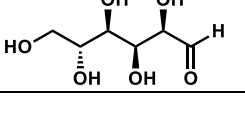


**Figure S2.1 (continued) B) Fungal cAT Phylogenetic Tree:** Phylogenetic tree of the cAT domains of the HRPKS-cAT enzymes identified from Figure S2.1A. The mouse cAT was used as an out-group to root the tree. Putative polyketides derived from HRPKS-cAT are denoted in blue (hypoxysordarin and AF-Toxin)



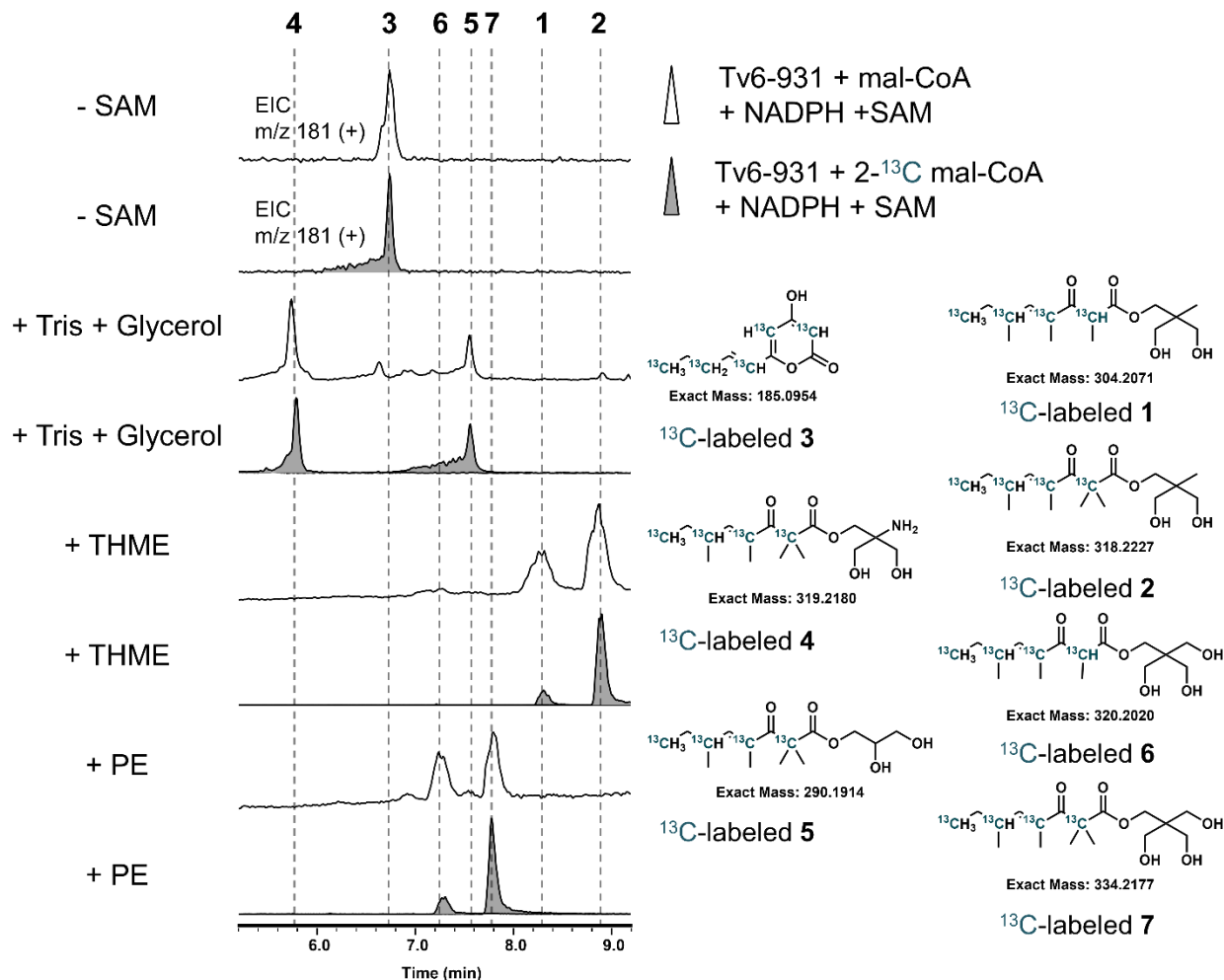
**Figure S2.2: SDS-PAGE Analysis of Purified Enzymes**

The Tv6-931 HRPKS-cAT, HRPKS-standalone protein, cAT-standalone protein and *apo*-HRPKS-cAT were all purified at >20 mg/L and then analyzed for purity using SDS-PAGE (6% and 12% acrylamide gels)

Tested Substrate	Relative Catalysis	Tested Substrate	Relative Catalysis
	+++++		+++++
	+++		+
	+		Trace
	Trace		Trace
	Trace		Product Not Detected
	Product Not Detected		Product Not Detected
	Product Not Detected		Product Not Detected
	Product Not Detected		Product Not Detected
	Product Not Detected		Product Not Detected
	Product Not Detected		Product Not Detected
	Product Not Detected		Product Not Detected
	Product Not Detected		Product Not Detected

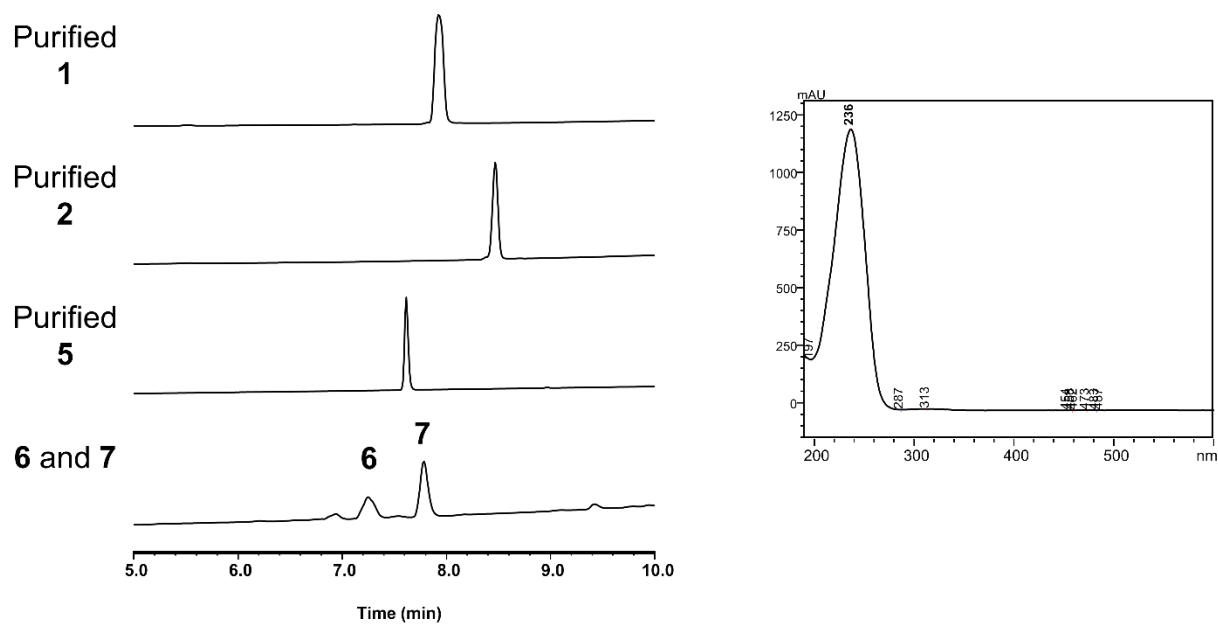
**Figure S2.3: List of Tv6-931 Offloading Substrates (Acyl-acceptors)**

A list of tested substrates as accepted nucleophiles for the Tv6-931 enzyme. The better substrates tend to be polyols. The presence of a nitrogen atom is not critical for catalysis. Many compounds including carnitine and choline show no catalysis.



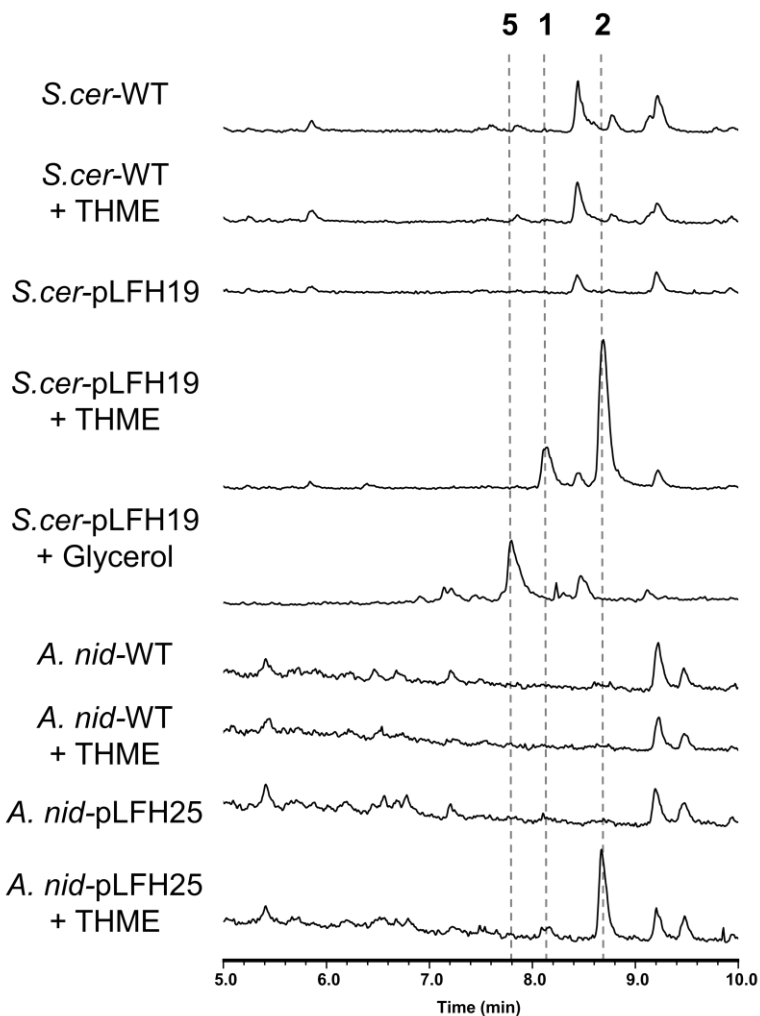
**Figure S2.4: Enzymatic <sup>13</sup>C Isotope Labeling**

LC-MS Mass Analysis (see **Table S8** for full data). Enzymatic synthesis assays with endogenous mal-CoA and <sup>13</sup>C-labeled mal-CoA (*in situ*), Tv6-931 enzyme, NADPH, SAM and various offloading nucleophiles (Total Ion Count). The <sup>13</sup>C-labeled  $\alpha$ -pyrone shows an increase in mass of 5 mu. The <sup>13</sup>C-labeled polyketide-polyol conjugates show an increase in mass of 4 mu.



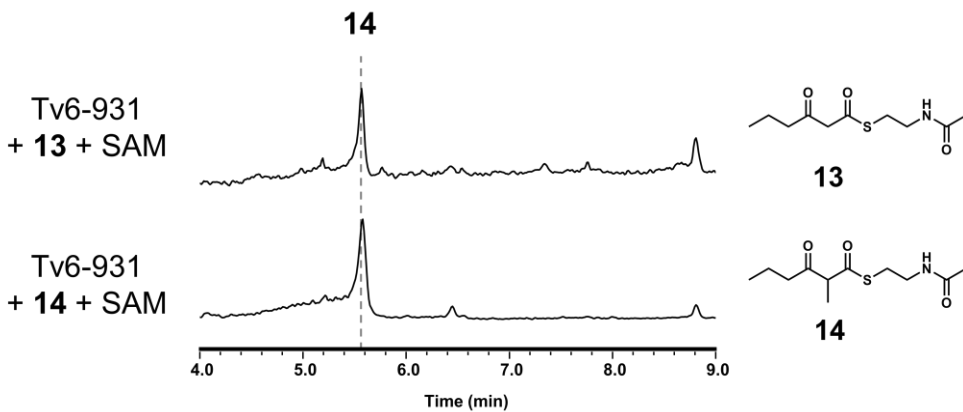
**Figure S2.5: HPLC Traces of Polyketide Products ( $\lambda = 236$  nm)**

The polyketide-polyol conjugates have a  $\lambda_{\text{max}}$  at 236 nm, indicative of a short  $\pi$  conjugation. HPLC traces of polyketide-polyol conjugates at  $\lambda = 236$  nm reflect the absorption [and purity].



**Figure S2.6: Heterologous Biosynthesis**

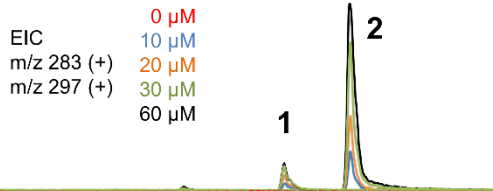
The biosynthesis of polyketide-polyol conjugates in heterologous hosts, *S. cer.* and *A. nid* (Total Ion Count). Biosynthesis is observed only if both the Tv6-931 HRPKS-cAT gene is introduced and an accepted polyol substrate is supplemented to the media.



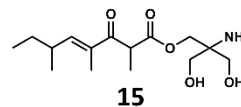
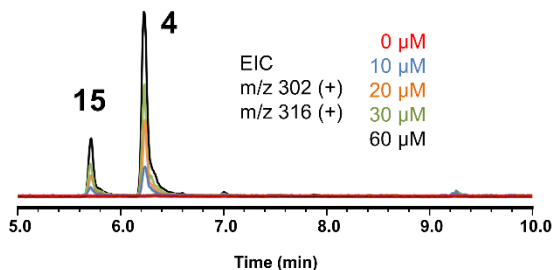
**Figure S2.7: Methylation Assays on Triketide SNACs**

Methylation assays with triketide SNACs and Tv6-931 HRPKS-cAT enzyme only result in the monomethyl SNAC-adduct. These results differ from the results from the tetraketide SNACs (Total Ion Count).

Tv6-931  $\Delta$ cAT  
+ Cofactors + THME  
+ Tv6-931 cAT  
(Titration)



Tv6-931  $\Delta$ cAT  
+ Cofactors + Tris  
+ Tv6-931 cAT  
(Titration)



**Figure S2.8: Titration of cAT in Enzymatic Synthesis**

Tv6-931  $\Delta$ cAT enzyme is insufficient for the enzymatic synthesis of polyketide-polyol conjugates when incubated with its cofactors and an acceptable substrate. Titration of the Tv6-931 cAT enzyme to the reaction mixture progressively restores the enzymatic synthesis of the polyketide-polyol conjugates (Total Ion Count).



```

cAT    L P R L P V P P L Q Q S L D Y Y L K A L Q P I V - S E E E W A H T K Q L V D E F Q T S G G V G E R L Q K G L E R R A - - K K M E N W L S E W L K T A Y I Q L P D K P - - V V I Y S S P G V I L P K Q D F V D L
M-cPT-1 L P K L P V P R V S A T I Q R Y L E S V R P L L - D D E E Y Y R M E L L A K E F Q - - D K T A P R L Q W Y L V L K S - W W A S N Y V S D W W E E Y I Y L R G R S P L M V N S N Y Y V M D L V L I K - - - - N T
L-cPT-1 L P R L P V P A V K D T V N R Y L Q S V R P L M - K E E D F K R M T A L Q A D F A - - V G L S P R L Q W Y L K L K S - W W A T N Y V S D W W E E Y I Y L R G R G P L M V N S N Y Y A M D L L Y L I - - - P T
cPT-II  L P R L P I P K L E D T M K R Y L N A Q P L L - D S Q F R R T E A L C K N F E - - T G V S K E L H A H L A Q D K Q N K H T S Y I S G P A F D - M Y L T A R D S I V L N F N P F V A - - F N P D P K S E Y
cOT     L P S L P V P S L E E S L K K Y L E S V K P F A - N Q E E Y K K T E E I V Q K F Q - - S G I G E K L H Q K L L E R A - - K G K R N W L E E W L N V A Y I D V R I P S Q L N V N F A G P A A H F E H Y W P P K E
PhS-cAT L P K F P L Y S L D T T M Q V F M E S I G H L G - N E Q E L Q K T R E A V S A L L S P G M G T R L Q A R L E I L N N D P K V D N W L I D I Y N R A W W S A R D M G P R G H N F F G T H E L S - - - - K S P
SS-cAT  L P K Q P L P D L D N S L S Q F L S D L L P V L - T S E E Y A R Y E N Y V T E F L K P E G F G R L Q A R L S K V A Q D P K V E N W L G E F Y A T S M F I K G R R S L V P W L N F F A T H F M S - - - - P V P
PeS-cAT L P K Q P L M D L D E A L G M I S G S R H F A - T E E E Y T A L C Q A I E T F G Q P G S T G R O L Y E K L V T M A S D P K V D N M M V D F L D A T F I R K R F P L V P L S S F F A T H H D A - - - - A V P
FF-cAT  L P R Y P L V D L D E A I T D L L N S V G H F A H T R E E Y E E L S R K A H V L A A P D S L S R R Y N R L R V K A D D P N V E S W I A D L L L K A L H I K R R Y P L A P F G N F L G T H F D S - - - - P T V
Tv6-931-cAT L P R Y P L Q T L E M T L D I F L D S V G H I A - N E E E L K H T R E A I S A F Q T P D G L S Q O L Q A R L A K L S G E Q G E N N E V V D M Y V R N K W I R G R D W R P R L R N F F A T L P G Q - - - - D T A R

cAT    Q G Q L R F A A K I E G V L D F K S M I D - N E T L P - - - - - V E F L G - G O P L C M N Q Y Y Q I L S S C R E P G P K D S V N V F L K S K - - - R P P T H
M-cPT-1 D V Q A A R L G N I I H A M I M Y R R K L D R E E I K P - - - - - V M A L G - I V P M C S Y Q M E R M F N T T R I P G K D T D V L Q H L - - - - - S D S R H
L-cPT-1 H I Q A A R A G N A I H A I L L Y R R K L D R E E I K P - - - - - I R L L G S T I P L C S A Q W E R M F N T S R I P G E E T D T I Q H M - - - - - R D S K H
cPT-II  N D Q L T R A T N L T V S A V R F L K T L R A G L L E P E V F H L N P S K S D T D A F K R L I R F V P P S L S W Y G A Y L V N A Y P L D M S Q Y F R L F N S T R I P R P N R D E L F T D - - - - - T K A R H
cOT     G T Q L E R G S I T L W H N L N Y V Q L L R - K E K L P - - - - - V H K V G - N T P L D M N Q F R M L F S T C K V P G I T R D S I M N Y F R T E S E G R S P N H
PhS-cAT Q T Q A E C A A L I S L A A F E Y K L S L D A G T V K Q - - - - - D F R N - E Q P L C M E T V H W L F N S N R V P V L G D R V D R W - - - - - P G N D Y
SS-cAT  H T A A E R A A I I S N A A F R F R K R L E R G E V G Y - - - - - G Y Y N - E Q P I S K D A Y Q W F N A T R E P S N K P D F M R R Y - - - - - P G N D Y
PeS-cAT H P Q A E R A A I A T T A F R F K E A V D A G T L E P - - - - - H W Y F - H I P S C M D S W Q W L F N T T R E P Q L E V D V M R Q Y - - - - - P G N D Y
FF-cAT  H T Q A E R A A I L T T A L Y E F K A D R D A G R L E P - - - - - D F L G - T R A N C G H S L S W L F N A V R E P N V D C K M M K Y - - - - - P D N E Y
Tv6-931-cAT Q P Q A D A Q A L T L G A Y G Y K L A D A G T V Q Q - - - - - D Y Y N - E Q A L D M A T V Y W L F S S N R T P A L G C D G Y D R Y - - - - - P E S D Y

cAT    I T V V H N Y Q F F E L D V Y N S D G T P L T S D Q I F V Q L E K I W N - S S L - Q S N K E P V G I L T S N H R N T W A K A Y N N L I K D K - V N R E S V N S I Q K S I F T V C L D K Q V P R V S D D V Y R - -
M-cPT-1 V A Y Y H K G R F F K L W L Y E G - A R L L K P O D L E M Q F Q R L D D P S P P Q P G E E K L A A L T A G G R V E W A Q A Q A F F S S G - K N K A A L E A I E R A A F F V A L D E E S Y D P E D - E A S
L-cPT-1 I V Y Y H R G R Y F K W L Y H D - G R L L K P R E M E Q Q M Q R L D N T S E P Q G E A R L A A L T A G D R V P N A R C R Q A Y F G R G - K N K Q S L D A V E K A A F F T V L D E T E E G Y R S E D P D T S
cPT-II  L L I L R K G H Y F V F D V L D Q D G N I V N P S E I Q A H L K Y I L S D S S P - - V P E F P V A Y L T S E N R D V W A E L R Q K L I F D G - - N E E T L K K V D S A V F C L C L D D F P - - - - - M K D
cOT     I V V L C R G R A F V F D V I H E - G C L V T P P E L L R Q L T Y I H K - K C H S E P D G P I A A L T S E E R T W A K A R E Y L I G L D P E N L A L L E K I Q S S L L Y V S M E D S S P H V T P E D Y S - -
PhS-cAT L V A M R H G H Y K V P L M L A D G Q K I S H L K L K A I F Q A I Q L - - - A P E E V N W A S V L T T G N R D I W A K T R D E V I A A S S I N E E F I T T V K E S L F A V C L E D S S P A N G H E R A - - -
SS-cAT  L I A F R R G H Y Y K I S V H D L - D E V T S Y A L R D A F Q N L D G - - - D Q K P E S W G V L Q D T D R D R A E L R V E L Q S L S K E N E T W I R D I E A S A F V V Y L D V Q P Q N A S E R A - - -
PeS-cAT C V V L R R G H Y F K V A L K N G - T E P V S Y A T I K S Q F D E L K N - - - V D Q E G F W T S L N D N R D S W A T I R E T L I S M S D T N A E C I R V I E Q A L F V I C L D N G A P T T S S D R I - - -
FF-cAT  V A V L R R G H L F K I P L K Q A - G T I V S Y D M L K A T Y Q A V D L - - - D L E D R S W A G M L T D N R D S A G L N R Q R L L S L D S R N A V Y L E T I E A S I F V M C L D D N S P I T R A D R V - - -
Tv6-931-cAT I V V M K R G H Y V K V M P R S S N G I V A Y S K L V S I F Q A I S Q H - - - R P D G T N W T S L L T A N R D E W A W A R D E I R T G S E D G R N F V K T I E E S L F I I Y L E D V A P E T A N D R A - - -

cAT    - N H V A G Q M L H G G S K F N S G N R W F D K T L Q F I V A E D G S C G M V Y E I A A A E G P P I V A L V D H V M E - - - - - Y T K P E L V R S P M V P L P M P K K L R F N I T P E I K N D I E K A K
M-cPT-1 L S L Y G K A L L H G - - - - - N C Y N R W F D K S F T L I S F K N G Q L G L N A E A W A D A P I I G H L W E F L G T D S F H L G Y T E T G H C L G K P N P A L A P P T R L Q W D I P K Q C Q A V E S S Y
L-cPT-1 M D S Y A K S L L H G - - - - - R C Y D R W F D K S F T F V V F K N G K M G L N A E I S W A D A Q I V A H L W E Y V M S I D S L Q L G Y A E D G H C K G D I N P N I P Y P T R L Q W D I P G E C Q E V I E T S L
cPT-II  L I H L S H T M L H G - - - - - D G T N R W F D K S F N L I V A E D G T A A V H F E I A W G D G V A V L R F F E N V F R D S T Q T P A I - - - T P Q S P A A T N S S A S V E T L S F N L S G A L K A G I T A A K
cOT     - E I I A A I L I G - - - - - D P T V R W G D K S Y N L I S F S N G V F G C N C D I A P F D A M I M V N I S Y Y V D E K I - - - - - F O N E G R W K G S E K V R D I P L P E E L I F I V D E K V L N D N Q A K
PhS-cAT - - - - - T A F L F D - - - - - E N S N R W L D K T V S F I I C A N G V S A T W Y Q I A M I D G S T P N G L T Q A L S E A T V D H I E S - L E T Y - - - S - L D L V S E D F K Y I P F F A T P S I N A H I T Q L R
SS-cAT  - - - - - P H L F H A - - - - - N G F N R W S D K T I Q F S I C D N G I S A T I G E I T M L G D V T F L R L N D F V T K A I M T F D P Q - R L T N I P S H P A L M I T V S E S Y T F K T T P S I E E H T R I R
PeS-cAT - - - - - R K A Y L G - - - - - D G F N R W N D K G L Q F I I F E N G S S G Y Q V E I T M I D G L T V H R M N E W I H E S I R S H I P N - Q D H T N G T - S N G I H P T L E E Y T L T T P S I D S H I L A V R
FF-cAT  - - - - - R S G Y I L D - - - - - D S F N R W H D K T C Q I I V T A N G R S A T I F E I S M I D L M T V S L S O R L Q N A I T L D P G - N A T A S D D T I A V D S A S L K E I P L V T T D I D S R I E T L R
Tv6-931-cAT - - - - - D I F L G D - - - - - D N S N R W L D K T L S F V V C S N G V S A I W G E I T M V D G T F F G L I K A L S S P S T K S F Q T - S S G S G S D - N T T V D G D F L H I P L T T L L L S H I S P L Q

cAT    Q N L S I M I Q D L D I M M L T F H H F G K D F P K S E K L S P D A F I Q V A L Q L A Y Y R I Y S C A C A T Y - E S A S L R M F H L G R T D T I R S A S I D S L A F V K G M G - - D S T V P E Q Q V E L L R K
M-cPT-1 Q V A K A L A D D V E L Y C F O F L P F G K G L I K K C R T S P D A F V Q I A L Q L A H F R D R K F C L Y T - E A S M T R M F R E G R T E T V R S C T S E S T A F V Q A M M - - E G S H T K A D L R D L F Q K
L-cPT-1 N T A N L L A N D V D F H S P F F V A F G K G I I K K C R T S P D T F V Q L A L Q L A H Y K D M K F C L Y T - E A S M T R L F R E G R T E T V R S C T S E S T A F V Q A M M - - D P A Q T V E Q R L K L F K
cPT-II  E K F D T T V K T L S I D S I Q F O R G G K E F L K K Q L S P D A V A Q L A F O M A F L R Q M Q V A T Y - E S C T A A F K H G R T E T I R P A S I F T K R C S E A F V R D P S K H S V G L Q H M M A E
cOT     A Q Y L R E A S D L Q I A A Y A F T S F G K K L T K N K M L H P D T F I Q L A L O L A Y Y R L H P G C C Y - E T A M T R H Y H G R T E T M R S C T V E A V R W C Q S M Q - - D P S V N L R E R Q Q K M L Q
PhS-cAT K Q H L S S I S G F S Y R V H V H T T Y G A T Y I R S H K L P P K T V F Q L I C I I A A R R K F Y N P G S S W - D M A I Q R Q K R G R F D S I N V Q T A E V A A F C A I A E - - D D T V S S A E R R R L F M D
SS-cAT  A Q L K N D T S R V F R A F E I P T L N R E L F R L H K C P P N S G F Q L A I Q L A V R R Y F O R N P A A F - E P I S L N H F H K R I D V N H I M R P S I S S F C A A A A - - D T K I P T D E L R G L F F D
PeS-cAT N S H L Q T T S S R E Y S Y L T L S H F G K D F L A R H G C P I K S V F D V T I Q L A S N L Y F H N P A S W - E P M S M A H F H K G R T E I F Q G V L P S V A Q F C T T A - - - D T A V P A Q R R E M L I R
FF-cAT  Q D Y L A I T S T K Q Y V P H L I N S F G K A L L S H S A P I K A T V D L T I Q L A S R L Y F P L P A S W - E T V S T A H F H L G R P E I V Q V V L K S M V M D F C D A A L - - - D D A V P H S E S R A R L L Q
Tv6-931-cAT A Q H Q R A H D G Y T L A N L T H T Y Y A A A Y L R Q Q K L P P K T V I Q L V I C A V R R L F Y Y N P S G A V D V I S O R P F R G R T E M I Y V M T P P I Q A F C A A A E - - D P S V S G E V K R R L F L E

cAT    A V Q A H R A Y T D R A I R E A F D R H L L G L K L Q A I E D - L V S M P D I F M D T S Y A I A M - - H F N L S T S Q V - - - - - P A K T D C V M F F G P V V P D G Y G I C Y N P M E - A H
M-cPT-1 A A K K H Q N M Y R L A M T A G I D R H L F C L Y V V S K Y L - G V S S P F L A - - - - - E V L S E P W R L S T S Q I P Q S Q I R M F D P E Q H P N H L G A G G G F G P V A D D G Y G V S Y I A G E N T
L-cPT-1 A S E K H Q H M Y R L A M T A S G I D R H L F C L Y V V S K Y L - A V E S P F L K - - - - - E V L S E P W R L S T S Q I P Q S Q I R M F D P E Q H P N H L G A G G G F G P V A D D G Y G V S Y I L V G E N L
cPT-II  C S K Y H G Q L T K E A A M A G F D R H L Y A L R Y L A T A R - G L N L P E L Y L D P A Y Q M N - - H N I L S T S T L - - - - - N S P A V L S G G F A P V V P D G F G I A Y A V H D - D W
cOT     A F A K H N K M K D C S A K G F D R H L L G L L L I A K E E - G L P V P E L F T D P L F S K S G G G N F V L S T S L V - - - - - G - Y L R V G Q V V V M H V N G Y G F F Y H I R D - D R
PhS-cAT A V K S H A R F A L S T R A G M M R H V M A L K E - - V M Q P G E E L P A L Y T D P V Y L R T R - E R K V F - T S F G - - - - - D S G - A L E L G N F W A D R E A L W F S C E L G D N S A
SS-cAT  A A R I H A S N V V T T S R A H G F D R H L F - - A L E W S V R G G E T V P V L F T D P T Y S K R R - P P Q I M - T N C I - - - - - V G E - A L E G G D F D H P N G V S V T Y E T K D E F T
PeS-cAT A A N E Y T A G L Q N S G K N N Y F R L M T V L E G M P V P E - E E A L A P L R D P W L R T Y - - P R F I M - S G L T - - - - - D G G - S L D S A F A L V D P E G W I C Y S I G E N E A
FF-cAT  A A R E C N A Q I T K G T E R N Y F R L M D V I E V M A Q Q D D E K M P E L F S D P W K R G Y - - P Q L I M - Q T M V - - - - - E T K L A E D P G T F M P H S A W M M N Y V T L D D S V
Tv6-931-cAT A V K S H A R L V A L S T R A G R F W H L M A L R E - - N M P E G E L A P F A Y N D T V A R T S - - E R P V C - T S F T - - - - - E F G - L P E M G Q P Q P K S A D V W V V G V Q V F D E R V

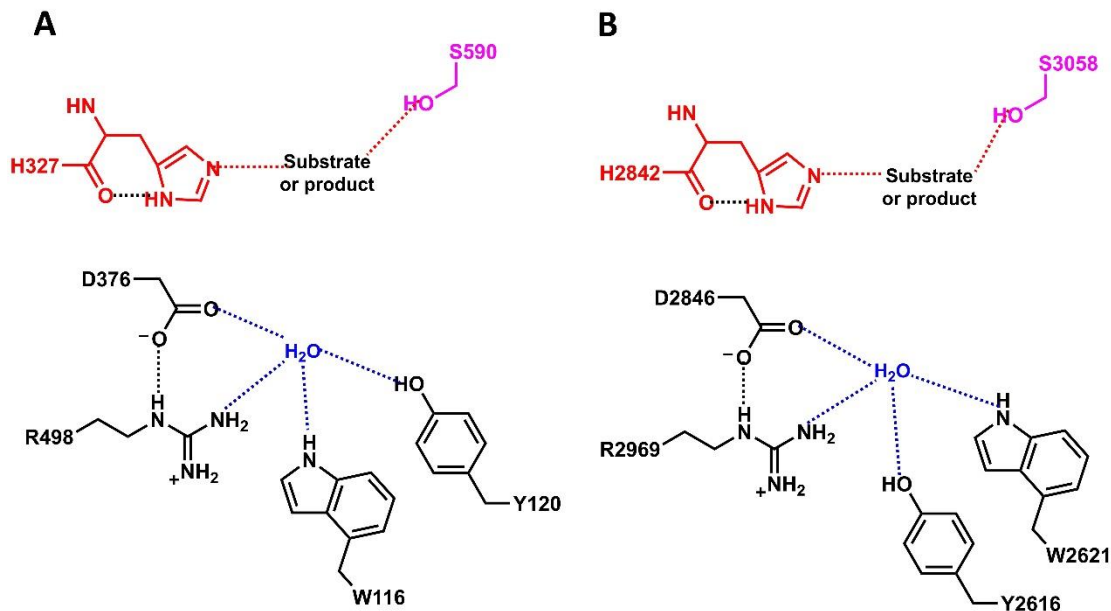
cAT    I N L S V S A Y N S C A E T N A A R M A H Y L E K A L L D M R T L L Q N H P R A K L
M-cPT-1 I F F H I S S K F S S E T N A Q R F G N H I R K A L L D I A D L F Q V P K A Y S -
L-cPT-1 I N F H I S S K F S C P E T D S H R F G R H L K E A M T D I I T L F G L S S N S K K
cPT-II  I G C N V S S Y - - - S G R N A R E F L H C V Q K C L E D I F D A L E G K A I K T -
cOT     F V V A C S A W K S C P E T D A E K L V Q L T F C A F H D M I Q L M N S T H L - - -
PhS-cAT R I F V A - - - - - N G E G R A D E V V G Y I K E A S Q V I R R I I E G L V - - - -
SS-cAT  K I S I W - - - - - G P L G Q V D K F I E V L E E S C R D I R A I I E F E P - - - -
PeS-cAT R F S I V - - - - - G P V G V S R F E E V L D E A T A L V K T L V E S S - - - -
FF-cAT  E V C F V - - - - - S P V K S A E R F G A A L D R A A E L V K N I I S A R D I - - - -
Tv6-931-cAT Q F T V I - - - - - N G E R K S E E F V E H I K A A A E I V R D I I E I N V V - - - -

```

**Figure S2.9: Sequence Alignment of the Tv6-931 cAT Domain and its Homologs**

**Note:** cAT, carnitine acyltransferase from *Mus musculus*, accession number CAA59971.1; M-cPT-1, human Muscle-type carnitine O-palmitoyltransferase 1, accession number NP\_001138609.1; L-cPT-1, human Liver-type carnitine O-palmitoyltransferase 1, accession

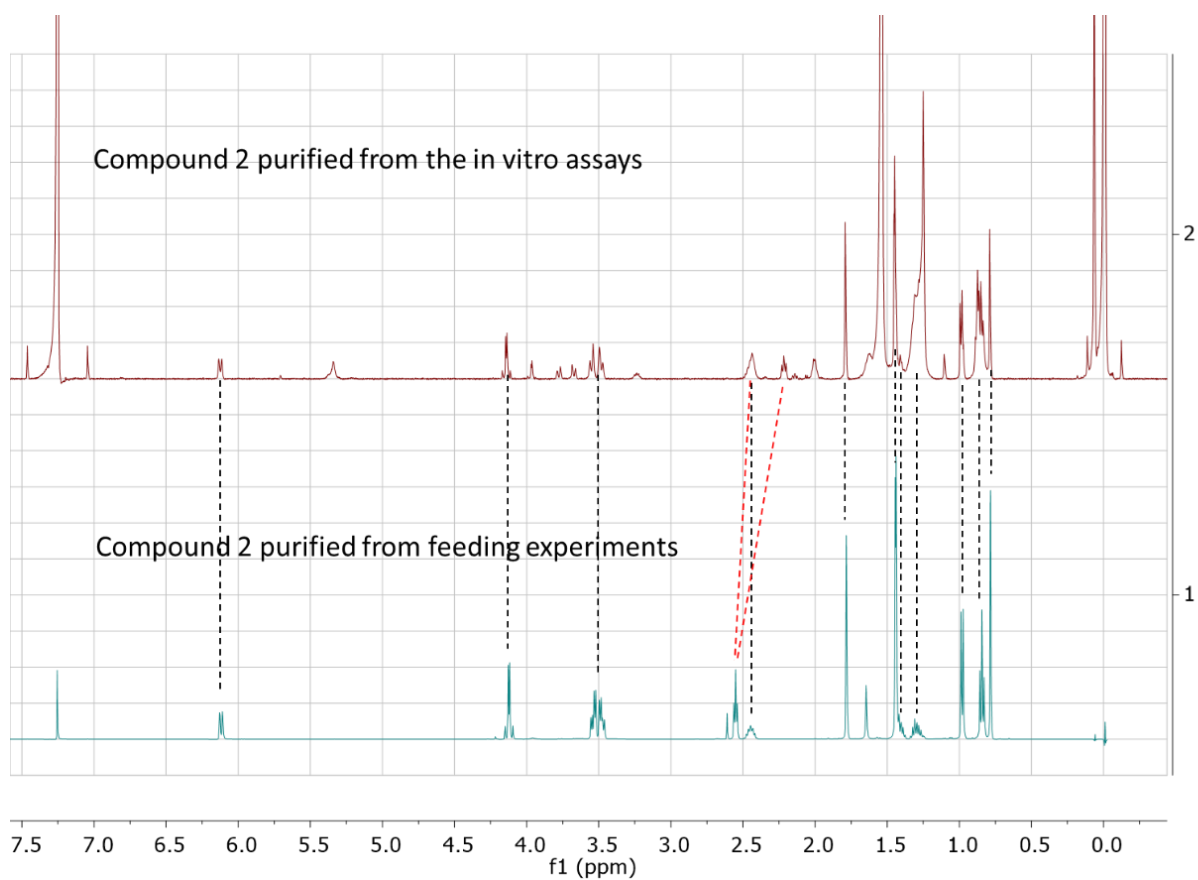
number AAC41748.1; cPT-II, carnitine palmitoyltransferase 2 from *Rattus norvegicus*, accession number EDL90417.1; cOT, human peroxisomal carnitine octanoyltransferase, accession number AAF03234.1; PhS-cAT, the carnitine acyltransferase domain of the HRPKS from *Phialocephala subalpina*, accession number CZR66375.1; SS-cAT, the carnitine acyltransferase domain of the HRPKS from *Sclerotinia sclerotiorum* 1980 UF-70, accession number APA14871.1; PeS-cAT, the carnitine acyltransferase domain of the HRPKS from *Penicillium subrubescens*, accession number OKP11696.1; FF-cAT, the carnitine acyltransferase domain of the HRPKS from *Fusarium foetens*, accession number ALQ32852.1; Tv6-931-cAT, the carnitine acyltransferase domain of Tv6-931. The conserved catalytic residue, His, is labeled in a red box. The putative conserved serine within the oxyanion hole is also labeled in a red box.



**Figure S2.10: Conserved Residues between Authentic cAT and the cAT Domain**

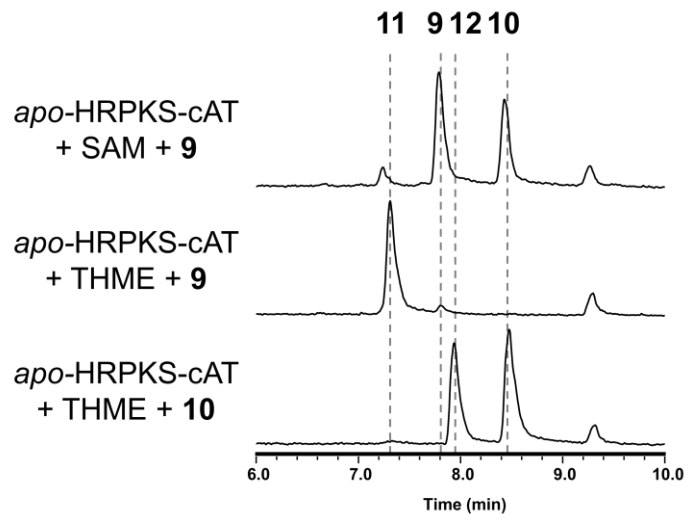
**A:** The conserved residues in the active site of the rat carnitine palmitoyltransferase 2 including the catalytic histidine, the conserved serine, and the residues that coordinate to water molecule.<sup>53</sup>

**B:** The correspondent residues are conserved in the cAT domain of Tv6-931.



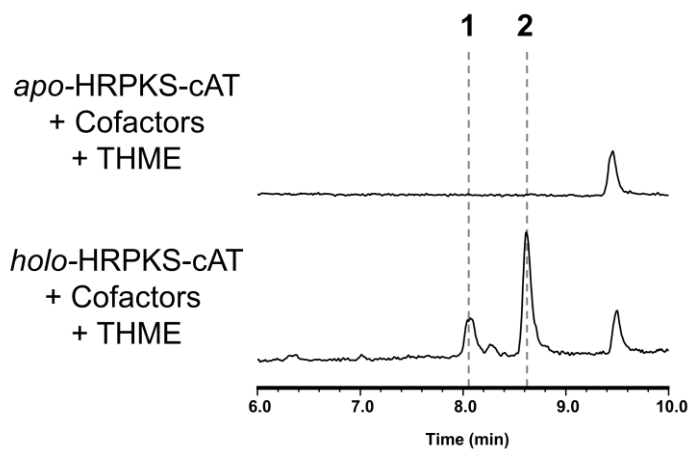
**Figure S2.11: Comparison of the  $^1\text{H}$  NMR spectra of compound 2 purified from the *in vitro* assays and the feeding experiments**

Note: the  $^1\text{H}$  NMR spectrum of compound 2 purified from the *in vitro* assays is almost identical to the spectrum of compound 2 purified from feeding experiments, except for the peaks of the proton from the two  $-\text{OH}$  group (labeled with red line in the figure). This is because of the water content of the NMR solvent ( $\text{CDCl}_3$ ) is different when collecting the NMR Data.



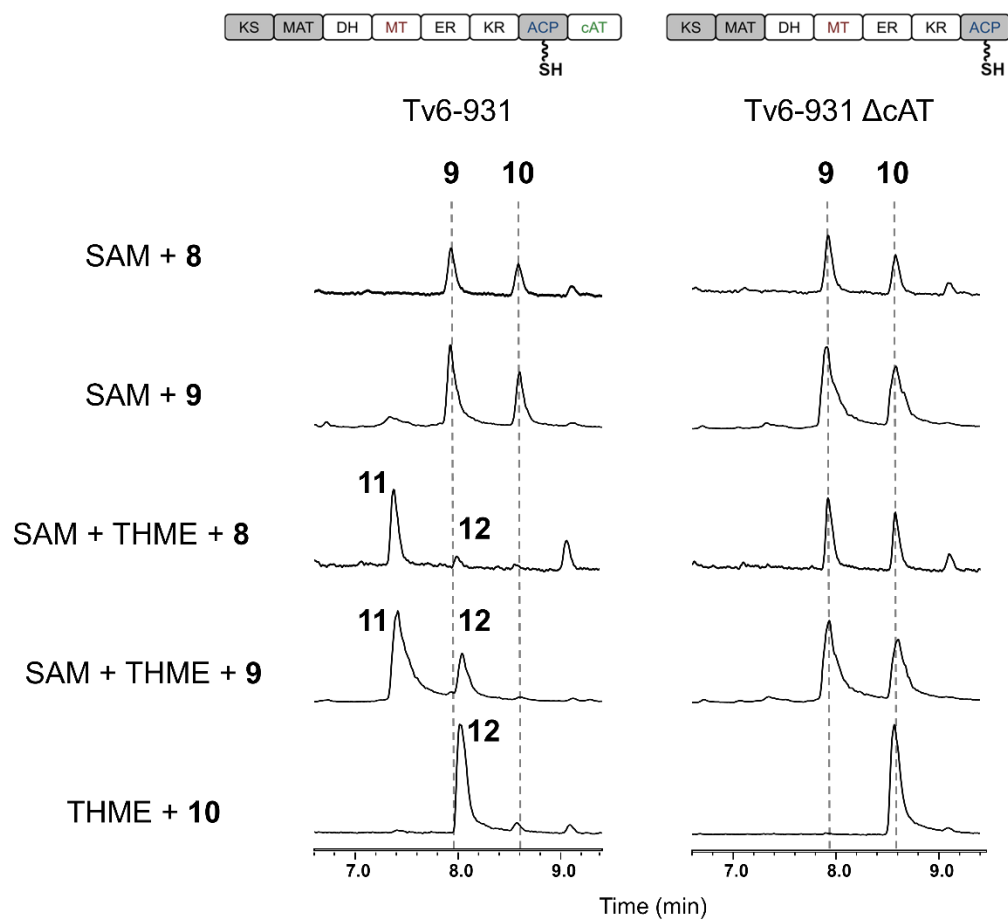
**Figure S2.12: Verification of MT and cAT Activity of *apo*-Tv6-931 Enzyme**

Using the tetraketide SNACs **9** and **10**, the MT and cAT activity of the *apo*-Tv6-931 enzyme was confirmed. SNAC **9** was readily methylated to **10** and transacylated to **11** by *apo*-Tv6-931. Likewise, SNAC **10** was also readily transacylated to **12** by *apo*-Tv6-931. (Total Ion Count)



**Figure S2.13: Abolishment of Biosynthesis in *apo*-Tv6-931 Enzyme**

Without the *holo*-form of the ACP, the biosynthesis of **1** and **2** by *apo*-Tv6-931 enzyme is completely abolished (Total Ion Count). Without any detectable product, the ACP of the *apo*-Tv6-931 enzyme is presumed to be entirely in the *apo* form.



**Figure S2.14: Methylation and Transacylation Assays with Tetraketide SNACs**

Methylation assays with tetraketide SNACs and Tv6-931 and Tv6-931  $\Delta$ cAT enzymes results in the monomethyl and dimethyl SNAC adducts. Transacylation of the SNACs to THME can occur readily in the Tv6-931 enzyme, but not the Tv6-931  $\Delta$ cAT enzyme.

## Section 2.8 Experimental Methods

### Bioinformatic Analysis:

#### *Fungal KS Phylogenetic Tree Development*

The fungal KS phylogenetic tree was generated according to the protocol detailed by Li *et al.*<sup>9</sup> The highlighted HRPKS-cAT subclade was identified using the bioinformatic tools: basic local alignment search tool (BLAST),<sup>34-35</sup> and Conserved Domain Database (CDD).<sup>36-37</sup>

#### *Fungal cAT Phylogenetic Tree Development*

A collection of 68 identified HRPKS-cAT enzymes from the fungal KS phylogenetic tree were used as the template for the phylogenetic tree. The boundaries of the ketosynthase, cAT, and other domains were identified using several bioinformatic tools such basic local alignment search tool (BLAST)<sup>34-35</sup> and Conserved Domain Database (CDD)<sup>36-37</sup>. The annotated amino acid sequence of cAT domain was then aligned using the Multiple Sequence Comparison by Log-Expectation (MUSCLE)<sup>38-39</sup> using ClustalW parameters<sup>40-41</sup>.

Phylogenetic analysis was performed using MEGA7<sup>42-43</sup> (gap opening penalty = 10, gap extension penalty = 0.1). The Gonnet matrix was used and a delay divergent cutoff was set to 30%. A Neighbor Joining tree was selected because of its ability to calculate distances to internal nodes instead of clustering and does not assume that all sequences have evolved at the same rate. For reproducibility and to prove that the tree was reliable, 1000 bootstrap replicates were applied. The scale bar indicates 0.1 changes per amino acid.

#### *Genome mining and in silico analysis for new fungal HRPKS*

The genomes of fully-sequenced fungi were readily accessed using the Joint Genome Institute (JGI) portal.<sup>44-45</sup> Using BLAST, these genomes were probed using various annotated KS domains serve as query templates. An additional search using the ACP-cAT region as a query template distinguished the novel HRPKS-cAT hybrids from traditional HRPKS enzymes. Functional domains in the translated sequences were also predicted using BLAST,<sup>34-35</sup> and CDD<sup>36-37</sup>. Protein structures were predicted using the Phyre2 server.<sup>46</sup>

## Materials and Methods

### *Microbial Strains*

The *Saccharomyces cerevisiae* BJ5464-NpgA<sup>47</sup> cells (*MAT $\alpha$  ura3-52 trp1 leu2- $\Delta$ 1 his3 $\Delta$ 200 pep4::HIS3 prb1 $\Delta$ 1.6R can1 GAL*) was used for the expression of the Tv6-931 and Tv6-931  $\Delta$ cAT enzymes and the *in vivo* biosynthesis of **1**, **2**, **5**, **6** and **7**—the Tv6-931 trihydroxymethylethane (THME), pentaerthritol (PE) and glycerol adducts, respectively. The *Saccharomyces cerevisiae* BJ5464-*apo* cells (*MAT $\alpha$  ura3-52 trp1 leu2- $\Delta$ 1 his3 $\Delta$ 200 pep4::HIS3 prb1 $\Delta$ 1.6R can1 GAL*) was used for the expression of the *apo*-Tv6-931.

The *Escherichia coli* SoluBL21<sup>TM</sup> (Genlantis) strain was used for expression of the Tv6-931 cAT domain. The *Aspergillus nidulans* A1145 strain (*pyrG89; pyroA4; nkuA::argB; riboB2*) was used for expression of the Tv6-931 enzyme and the *in vivo* biosynthesis of **1** and **2**, the Tv6-931 THME adducts.

### *HPLC-MS Analysis*

The HPLC-MS analyses were performed using a Shimadzu 2020 EVLC-MS (Phenomenex® Luna, 5 $\mu$ , 2.0  $\times$  100 mm, C-18 column) using positive and negative mode

electrospray ionization. The elution method was a linear gradient of 5-95% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes with a flow rate of 0.3 mL/min. The HPLC buffers were supplemented with 0.05% formic acid (v/v).

#### *HPLC Purification*

The HPLC purifications were performed using a Shimadzu Prominence HPLC (Phenomenex® Kinetex, 5 $\mu$ , 10.0  $\times$  250 mm, C-18 column). The elution methods were variable linear gradients in CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes with a flow rate of 3.0 mL/min.

#### *NMR Analysis*

All NMR spectra including <sup>1</sup>H, <sup>13</sup>C, COSY, HSQC, HMBC and NOESY spectra were obtained on Bruker AV500 spectrometer with a 5 mm dual cryoprobe at the UCLA Molecular Instrumentation Center. The NMR solvent, CDCl<sub>3</sub>, used for these experiments was purchased from Cambridge Isotope Laboratories, Inc.

#### *High-Resolution Mass Spectroscopy*

High-resolution mass spectra were obtained from Thermo Fisher Scientific Exactive Plus with IonSense ID-CUBE DART source at the UCLA Molecular Instrumentation Center.

#### *General Molecular Biology Experiments*

General molecular biology and cloning techniques were performed according to protocols described elsewhere.<sup>48</sup> PCR was performed using Phusion® and Q5® High-Fidelity DNA Polymerase (New England Biolabs) according to protocols recommended by the manufacturer. DNA restriction enzymes were used as recommended by the manufacturer (New England



Biolabs). PCR products were confirmed by DNA sequencing by Laragen, California, USA. TOP10 (Invitrogen), DH10B (Invitrogen) and XL1-Blue (Stratagene) *E. coli* cells were used for DNA manipulation, following standard recombinant DNA techniques. RNA extraction was performed using a RiboPure Yeast Kit (Ambion) and complementary DNA (cDNA) was prepared from the total RNA by SuperScript® II Reverse Transcriptase (Invitrogen) with random primers as described by the manufacturer.

### **Construction of Expression Plasmids**

#### *Plasmids for expression in Saccharomyces cerevisiae*

Relevant DNA primers and plasmids are listed in (see **Table S2.1** for details), respectively. The following 2 $\mu$ -based yeast expression plasmids with auxotrophic markers were used for expression: pXW55.<sup>49</sup> After codon optimization<sup>50</sup>, the coding domain sequence of the Tv6-931 gene was synthesized by Gen9. The codon-optimized gene was then inserted into the digested pXW55<sup>49</sup> plasmid (SpeI and PmlI) using *in vivo* homologous recombination to create the pTMC-1-106.<sup>51</sup>

The first-generation plasmid, pTMC-1-106, contained the Tv6-931 gene flanked by an *N*-terminal FLAG tag and *C*-terminal 6xHis Tag. Using the pTMC-1-106, various relevant PCR products were synthesized using the following primers:

pXW55-N-Term-Ura-F, pXW55-N-8xHis-R pXW55-C-Term-F, pXW55-C-Term-Ura-R, Tv6-931-N-His-HRPKS-P1-F, Tv6-931-HRPKS-P1-R, Tv6-931-HRPKS-P2-F, Tv6-931-HRPKS-P2-R, Tv6-931-HRPKS-P3-F, Tv6-931-HRPKS-cAT-Native-R, Tv6-931-HRPKS-standalone-R.

The resulting PCR products were used to assemble the second-generation plasmids, pLFH18 and pLFH19 by *in vivo* yeast homologous recombination using Frozen-EZ Yeast Transformation II Kit™ (Zymo Research). These expression plasmids include the relevant coding DNA sequence with an additional *N*-terminal 8xHis for purification.

#### *Plasmids for expression in Escherichia coli*

The Tv6-931 cAT domain was synthesized from pTMC-1-106 using the primers, Tv6-931-cAT-standalone-F and Tv6-931-cAT-standalone-R. The resulting PCR product and pET-28a(+) was digested using the restriction enzymes, NdeI and NheI. The digested DNA fragments were ligated together to produce the plasmid, pLFH17, using ExpressLink™ T4 DNA Ligase (Invitrogen). This plasmid includes the relevant coding DNA sequence and an *N*-terminal 6xHis tag for purification.

#### *Plasmids for expression in Aspergillus nidulans*

The AMA1-based fungal expression plasmid with the URA3 auxotrophic marker (pYTU) was used for expression.<sup>52</sup> The Tv6-931 HRPKS-cAT gene was synthesized from *Trichoderma virens* gDNA using the relevant DNA primers listed in **Table S2.1**:

HRPKS-cAT-gDNA-pYTU-P1-F, HRPKS-cAT-gDNA-pYTU-P1-R, HRPKS-cAT-gDNA-pYTU-P2-F, HRPKS-cAT-gDNA-pYTU-P2-R, HRPKS-cAT-gDNA-pYTU-P3-F, HRPKS-cAT-gDNA-pYTU-P3-R, HRPKS-cAT-gDNA-pYTU-P4-F, HRPKS-cAT-gDNA-pYTU-P4-R.

The empty pYTU plasmid was digested using the restriction enzymes, NotI. The linearized pYTU plasmid and the resulting PCR products were used to assemble the plasmid,

pLFH25 by *in vivo* homologous recombination using Frozen-EZ Yeast Transformation II Kit™ (Zymo Research). These expression plasmids include the native DNA sequence with introns and the native gene terminator.

### ***In vivo* Expression Experiments**

#### *Protein Expression and Purification in Saccharomyces cerevisiae*

For protein expression under the alcohol dehydrogenase (*ADH2*) promoter, the yeast expression plasmid of interest (pLFH18 or pLFH19) is first transformed into *S. cerevisiae* BJ5464-NpgA or BJ5464-*apo* cells using the Frozen-EZ Yeast Transformation II Kit™ (Zymo Research). A seed culture of the transformed *S. cerevisiae* strain was grown in 40 mL of synthetic dropout complete medium without uracil (SDC-ura). The seed culture was shaken for two days at 28°C, 250 rpm. This seed culture was then used to inoculate 4 L of YPD 1.5% media (yeast extract (10 g/L), peptone (20 g/L) and dextrose (15 g/L)) and cultured for 3 days at 28°C, 250 rpm. The cell pellet was harvested by centrifugation (5000 rpm, 10 mins), and it was then suspended in 250 mL of a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 150 mM NaCl, and 10 mM imidazole at pH 8.0.

The suspended yeast cells were lysed using sonication and the cellular debris was removed using high-speed centrifugation (17,000 rpm, 1 hr). The enzyme of interest was then purified from the supernatant using Ni-NTA agarose affinity chromatography to near homogeneity. The purified protein was concentrated, exchanged into a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub> and 50 mM NaCl at pH 8.0, aliquoted and then flash frozen in liquid nitrogen.

#### *Protein Expression and Purification in Escherichia coli*

For protein expression under the T7 promoter, the bacterial expression plasmid of interest, pLFH17, is first transformed into *E. coli* SoluBL21™ (Genlantis) cells under electroporation (2.25 kV, 400 ohms, and 25  $\mu$ F). The seed culture of the transformed *E. coli* strain was grown in 3 mL of lysogeny broth (LB) with kanamycin (35  $\mu$ g/mL) medium. A seed culture was shaken overnight at 37°C, 250 rpm. This seed culture was then used to inoculate 4 L of LB with kanamycin (35 mg/L) media and cultured to an OD<sub>600</sub> of 0.5. The culture was then incubated on ice for 10 minutes, induced with 0.1 mM isopropyl thio- $\beta$ -D-galactoside (IPTG) and shaken overnight at 16°C, 250 rpm.

The cell pellet was harvested by centrifugation (3500 rpm, 10 mins), and it was then suspended in 100 mL of a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 150 mM NaCl, and 10 mM imidazole at pH 8.0. The suspended bacterial cells were lysed using sonication and the cellular debris was removed using high-speed centrifugation (17,000 rpm, 1 hr). The enzyme of interest was then purified from the supernatant using Ni-NTA agarose affinity chromatography to near homogeneity. The purified protein was concentrated, exchanged into a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub> and 50 mM NaCl at pH 8.0, aliquoted and then flash frozen in liquid nitrogen.

For the protein expression and purification of the MatB enzyme, the protocol detailed by Ma *et al* was followed.<sup>19</sup> The purified protein was concentrated, exchanged into a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub> and 50 mM NaCl at pH 8.0, aliquoted and then flash frozen in liquid nitrogen.

#### *Biosynthesis of Tv6-931 THME (and PE) products in S.cer. BJ5464-*np*gA*

For expression of the Tv6-931 THME products, **1** and **2**, under the alcohol dehydrogenase (*ADH2*) promoter, the yeast expression plasmid of interest, pLFH19, is first

transformed into *S. cerevisiae* BJ5464-NpgA cells using the Frozen-EZ Yeast Transformation II Kit™ (Zymo Research). A seed culture of *S.cer*-pLFH19 was grown in 40 mL of synthetic dropout complete medium without uracil (SDC-ura). The seed culture was shaken for two days at 28°C, 250 rpm. This seed culture was then used to inoculate 4 L of Yeast Peptone 1.5% Dextrose (YPD 1.5%) medium and cultured for 2 days at 28°C, 250 rpm. The cell pellet was harvested by centrifugation (5000 rpm, 10 mins) and it was then suspended in 800 mL of the supernatant with 1% THME, 1% sodium acetate and 0.1% cyclopentanone (v/v). The resuspended culture was shaken for an additional 4 days at 28°C, 250 rpm.

After centrifugation to remove the yeast cells, the supernatant was extracted with 400 mL of hexanes and then extracted with 3x 800 mL of ethyl acetate. The combined organic layer was washed with 800 mL of brine (saturated NaCl) and then dried over Na<sub>2</sub>SO<sub>4</sub>. The crude extract was then evaporated to dryness *in vacuo* to produce a dark, brown oil. The crude extract was initially purified by flash column chromatography in silica using 1:1 mixture (v/v) of ethyl acetate and hexanes. The product was then further purified by semi-preparative HPLC. The elution method used was a linear gradient of 40-70% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes. The purity was confirmed by HPLC-MS and the structure was solved by NMR

#### *Biosynthesis of Tv6-931 glycerol product in Saccharomyces cerevisiae*

For expression of the Tv6-931 glycerol product **5** under the alcohol dehydrogenase (*ADH2*) promoter, the yeast expression plasmid of interest, pLFH19, is first transformed into *S. cerevisiae* BJ5464-NpgA cells using the Frozen-EZ Yeast Transformation II Kit™ (Zymo Research). A seed culture of *S.cer*-pLFH19 was grown in 40 mL of synthetic dropout complete

medium without uracil (SDC-ura). The seed culture was shaken for two days at 28°C, 250 rpm. This seed culture was then used to inoculate 4 L of Yeast Peptone 1.5% Dextrose (YPD 1.5%) medium and cultured for 2 days at 28°C, 250 rpm. The cell pellet was harvested by centrifugation (5000 rpm, 10 mins) and it was then suspended in 800 mL of the supernatant with 10% glycerol, 1% sodium acetate and 0.1% cyclopentanone (v/v). The resuspended culture was shaken for an additional 4 days at 28°C, 250 rpm.

After centrifugation to remove the yeast cells, the final supernatant was extracted with 3x 800 mL of ethyl acetate. The combined organic layer was washed with 800 mL of brine (saturated NaCl) and then dried over Na<sub>2</sub>SO<sub>4</sub>. The crude extract was then evaporated to dryness *in vacuo* to produce a dark, brown oil. The crude extract was initially purified by flash column chromatography in silica using 1:1 mixture (v/v) of ethyl acetate and hexanes (R<sub>f</sub>=0.25). The product was then further purified by semi-preparative HPLC using a C-18 reverse-phase column. The elution method used was a linear gradient of 35-55% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes. The purity was confirmed by HPLC-MS and the structure was solved by NMR.

#### *Biosynthesis of Tv6-931 THME product in Aspergillus nidulans*

For expression of the Tv6-931 THME products, **1** and **2**, under the glucoamylase (*glaA*) promoter, the fungal expression plasmid of interest, pLFH25, is first transformed into *A. nid.* A1145 cells using a PEG 4000 solution (60% PEG 4000, 50 mM CaCl<sub>2</sub>, 50 mM Tris-HCl, pH 7.5) and grown on CD agar supplemented with 1.2 M sorbitol, 0.5 µg/ml pyridoxine HCl and 0.125 µg/ml riboflavin at 37°C. Individual *A.nid*-pLFH25 transformants were then grown on CD-ST agar supplemented with 0.5 µg/ml pyridoxine HCl and 0.125 µg/ml riboflavin and

incubated at 37°C for 4 days. Afterwards, the agar was dissolved in 1 volume equivalent of acetone overnight. The slurry was then centrifuged (10,000 rpm, 10 mins) and the supernatant was evaporated *in vacuo* to remove bulk amounts of acetone. The concentrated aqueous layer was extracted 3x1 equivalent volumes of ethyl acetate. The combined organic was washed with brine (saturated NaCl) and then dried over Na<sub>2</sub>SO<sub>4</sub> to produce the crude extract of **1** and **2**.

### ***In vitro* Enzymatic Experiments**

#### *Enzymatic Polyketide Synthesis Assays*

Enzymatic polyketide synthesis reactions consisted of 15 μM of the Tv6-931 enzyme incubated with 4 mM NADPH, 2 mM SAM, 2 mM of malonyl-coenzyme A and 5 mM of the substrate of interest (**Figure S2.3**) in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

#### *Enzymatic Pyrone Synthesis Assays*

Enzymatic pyrone synthesis reactions consisted of 15 μM of the Tv6-931 enzymes incubated with 4 mM NADPH and 2 mM of malonyl-coenzyme in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by HPLC-MS.

#### *Enzymatic <sup>13</sup>C-labeled Polyketide Synthesis Assays*

Enzymatic  $^{13}\text{C}$ -labeled polyketide synthesis reactions consisted of 15  $\mu\text{M}$  of the Tv6-931 enzyme, 20  $\mu\text{M}$  of MatB enzyme, 4 mM NADPH, 2 mM SAM, 40 mM 2- $^{13}\text{C}$  malonate, 10 mM coenzyme A, 10 mM ATP and 7 mM  $\text{MgCl}_2$  and 5 mM of the substrate of interest (**Figure S2.3**) in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by HPLC-MS.

#### *Enzymatic Methylation Assay with N-acetyl cysteamine (SNACs)*

Enzymatic methylation reactions consisted of 15  $\mu\text{M}$  of the Tv6-931, Tv6-931  $\Delta\text{cAT}$  or *apo*-Tv6-931 enzyme incubated with 5 mM SAM and 2 mM of the SNAC of interest in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by HPLC-MS.

For the enzymatic synthesis of SNAC compound **10**, the reaction was instead quenched and extracted by 3x1 equivalent volumes of ethyl acetate. The combined organic layer then dried over  $\text{Na}_2\text{SO}_4$ . The crude extract was then evaporated to dryness *in vacuo* to produce a yellow oil. The product was then purified by semi-preparative HPLC using a C-18 reverse-phase column. The elution method used was a linear gradient of 40-70% (v/v)  $\text{CH}_3\text{CN}/\text{H}_2\text{O}$  in 15 minutes, followed by 95%  $\text{CH}_3\text{CN}/\text{H}_2\text{O}$  for 3 minutes. The purity was confirmed by HPLC-MS and the structure was solved by NMR.

#### *Enzymatic Tv6-931 cAT Domain Titration Assays*

Enzymatic cAT titration reactions consisted of 15  $\mu\text{M}$  of the Tv6-931  $\Delta\text{cAT}$  enzyme incubated with various concentration of the Tv6-931 cAT enzyme (10-60  $\mu\text{M}$ ), 4 mM NADPH, 2



mM SAM, 2 mM of malonyl-coenzyme A and 5 mM of THME or Tris-Cl in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by HPLC-MS.

#### *Enzymatic Acylation Assays (SNACs)*

Enzymatic acylation reactions consisted of 15 μM of the Tv6-931, Tv6-931 ΔcAT or *apo*-Tv6-931 enzyme incubated with 5 mM THME or PE and 2 mM of the SNAC of interest in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by HPLC-MS.

#### *Enzymatic Post-Release Methylation Assays*

Enzymatic post-release methylation reactions consisted of 15 μM of the Tv6-931, Tv6-931 ΔcAT or *apo*- Tv6-931 enzyme incubated with 5 mM SAM and 2 mM of **1** in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by HPLC-MS.

For the enzymatic methylation [and isolation] of **1** to **2**, the reaction was instead quenched and extracted by 3x1 equivalent volumes of ethyl acetate. The combined organic layer then dried over Na<sub>2</sub>SO<sub>4</sub>. The crude extract was then evaporated to dryness *in vacuo* to produce a yellow oil. The product was then purified by semi-preparative HPLC using a C-18 reverse-phase column. The elution method used was a linear gradient of 40-70% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15

minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes. The purity was confirmed by HPLC-MS and the structure was confirmed by <sup>1</sup>H-NMR.

#### *Enzymatic Post-release Acyltransferase Assays*

Enzymatic post-release acylation reactions consisted of 15 μM of the Tv6-931, Tv6-931 ΔcAT or *apo*-Tv6-931 enzyme incubated with 2 mM PE and 2 mM of **1** or **2** in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer. The reaction mixture was incubated overnight at 30°C. The reaction was quenched with three volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by HPLC-MS.

#### *Synthesis of SNAC compounds*

SNAC compounds, **8** and **9**, were chemically synthesized according to the protocol described by Cacho *et al.*<sup>4</sup> SNAC compound **10** was enzymatically synthesized from **9** using a semi-preparative (10 mL) enzymatic methylation assay. The SNAC compounds were purified by semi-preparative HPLC using a C-18 reverse-phase column. The elution method used was a linear gradient of 40-70% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes.

## Section 2.9: References

1. Lim, F. Y.; Keller, N. P., Spatial and temporal control of fungal natural product synthesis. *Nat. Prod. Rep.* **2014**, *31* (10), 1277-1286.
2. Cox, R. J., Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org. Biomol. Chem.* **2007**, *5* (13), 2010-2026.
3. Chooi, Y. H.; Tang, Y., Navigating the Fungal Polyketide Chemical Space: From Genes to Molecules. *J Org Chem* **2012**, *77* (22), 9933-9953.
4. Cacho, R. A.; Thuss, J.; Xu, W.; Sanichar, R.; Gao, Z.; Nguyen, A.; Vederas, J. C.; Tang, Y., Understanding Programming of Fungal Iterative Polyketide Synthases: The Biochemical Basis for Regioselectivity by the Methyltransferase Domain in the Lovastatin Megasyntase. *J. Am. Chem. Soc.* **2015**, *137* (50), 15688-91.
5. Roberts, D. M.; Bartel, C.; Scott, A.; Ivison, D.; Simpson, T. J.; Cox, R. J., Substrate selectivity of an isolated enoyl reductase catalytic domain from an iterative highly reducing fungal polyketide synthase reveals key components of programming. *Chem. Sci.* **2017**, *8* (2), 1116-1126.
6. Liddle, E.; Scott, A.; Han, C.; Ivison, D.; Simpson, T. J.; Willis, C. L.; Cox, R. J., In vitro kinetic study of the squalestatin tetraketide synthase dehydratase reveals the stereochemical course of a fungal highly reducing polyketide synthase. *Chem. Commun.* **2017**, *53* (10), 1727-1730.
7. Zabala, A. O.; Chooi, Y. H.; Choi, M. S.; Lin, H. C.; Tang, Y., Fungal Polyketide Synthase Product Chain-Length Control by Partnering Thiohydrolase. *Acs Chem. Biol.* **2014**, *9* (7), 1576-1586.

8. Xu, W.; Chooi, Y. H.; Choi, J. W.; Li, S.; Vederas, J. C.; Da Silva, N. A.; Tang, Y., LovG: the thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. *Angew. Chem. Int. Ed.* **2013**, *52* (25), 6472-5.
9. Li, Y. F.; Tsai, K. J. S.; Harvey, C. J. B.; Li, J. J.; Ary, B. E.; Berlew, E. E.; Boehman, B. L.; Findley, D. M.; Friant, A. G.; Gardner, C. A.; Gould, M. P.; Ha, J. H.; Lilley, B. K.; McKinstry, E. L.; Nawal, S.; Parry, R. C.; Rothchild, K. W.; Silbert, S. D.; Tentilucci, M. D.; Thurston, A. M.; Wai, R. B.; Yoon, Y. J.; Aiyar, R. S.; Medema, M. H.; Hillenmeyer, M. E.; Charkoudian, L. K., Comprehensive curation and analysis of fungal biosynthetic gene clusters of published natural products. *Fungal Genet. Biol.* **2016**, *89*, 18-28.
10. Colucci, W. J.; Gandour, R. D., Carnitine Acetyltransferase - a Review of Its Biology, Enzymology, and Bioorganic Chemistry. *Bioorg. Chem.* **1988**, *16* (3), 307-334.
11. Tsai, S. C.; Lu, H. X.; 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.
12. 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. *P. Natl. Acad. Sci. U.S.A.* **2010**, *107* (14), 6246-6251.
13. Jogl, G.; Tong, L., Crystal structure of carnitine acetyltransferase and implications for the catalytic mechanism and fatty acid transport. *Cell* **2003**, *112* (1), 113-122.
14. Hsiao, Y. S.; Jogl, G.; Tong, L., Crystal structures of murine carnitine acetyltransferase in ternary complexes with its substrates. *J. Biol. Chem.* **2006**, *281* (38), 28480-28487.

15. Ramsay, R. R.; Gandour, R. D.; van der Leij, F. R., Molecular enzymology of carnitine transfer and transport. *Bba-Protein Struct. M.* **2001**, *1546* (1), 21-43.
16. Ramsay, R. R.; Naismith, J. H., A snapshot of carnitine acetyltransferase. *Trends Biochem. Sci.* **2003**, *28* (7), 343-346.
17. Ruswandi, S.; Kitani, K.; Akimitsu, K.; Tsuge, T.; Shiraishi, T.; Yamamoto, M., Structural analysis of cosmid clone pcAFT-2 carrying AFT10-1 encoding an acyl-CoA dehydrogenase involved in AF-toxin production in the strawberry pathotype of *Alternaria alternata*. *J. Gen. Plant Pathol.* **2005**, *71* (2), 107-116.
18. Kudo, F.; Matsuura, Y.; Hayashi, T.; Fukushima, M.; Eguchi, T., Genome mining of the sordarin biosynthetic gene cluster from *Sordaria araneosa* Cain ATCC 36386: characterization of cycloaraneosene synthase and GDP-6-deoxyaltrose transferase. *J. Antibiot.* **2016**, *69* (7), 541-548.
19. Ma, S. M.; Li, J. W.; Choi, J. W.; Zhou, H.; Lee, K. K.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y., Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* **2009**, *326* (5952), 589-92.
20. An, J. H.; Kim, Y. S., A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in *Rhizobium trifolii*--cloning, sequencing, and expression of the enzymes in *Escherichia coli*. *Eur J Biochem* **1998**, *257* (2), 395-402.
21. Molnar, I.; Schupp, T.; Ono, M.; Zirkle, R.; Milnamow, M.; Nowak-Thompson, B.; Engel, N.; Toupet, C.; Stratmann, A.; Cyr, D. D.; Grolach, J.; Mayo, J. M.; Hu, A.; Goff, S.; Schmid, J.; Ligon, J. M., The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. *Chem. Biol.* **2000**, *7* (2), 97-109.

22. Young, J.; Stevens, D. C.; Carmichael, R.; Tan, J.; Rachid, S.; Boddy, C. N.; Muller, R.; Taylor, R. E., Elucidation of gephyronic acid biosynthetic pathway revealed unexpected SAM-dependent methylations. *J. Nat. Prod.* **2013**, *76* (12), 2269-76.
23. Miller, D. A.; Luo, L.; Hillson, N.; Keating, T. A.; Walsh, C. T., Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem. Biol.* **2002**, *9* (3), 333-44.
24. Sudek, S.; Lopanik, N. B.; Waggoner, L. E.; Hildebrand, M.; Anderson, C.; Liu, H.; Patel, A.; Sherman, D. H.; Haygood, M. G., Identification of the putative bryostatin polyketide synthase gene cluster from "Candidatus *Endobugula sertula*", the uncultivated microbial symbiont of the marine bryozoan *Bugula neritina*. *J. Nat. Prod.* **2007**, *70* (1), 67-74.
25. Pfeifer, B. A.; Wang, C. C. C.; Walsh, C. T.; Khosla, C., Biosynthesis of yersiniabactin, a complex polyketide-nonribosomal peptide, using *Escherichia coli* as a heterologous host. *Appl. Environ. Microb.* **2003**, *69* (11), 6698-6702.
26. Piel, J., A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *P. Natl. Acad. Sci. U.S.A.* **2002**, *99* (22), 14002-14007.
27. Zhao, C. H.; Coughlin, J. M.; Ju, J. H.; Zhu, D. Q.; Wendt-Pienkowski, E.; Zhou, X. F.; Wang, Z. J.; Shen, B.; Deng, Z. X., Oxazolomycin Biosynthesis in *Streptomyces albus* JA3453 Featuring an "Acyltransferase-less" Type I Polyketide Synthase That Incorporates Two Distinct Extender Units. *J. Biol. Chem.* **2010**, *285* (26), 20097-20108.
28. Poust, S.; Phelan, R. M.; Deng, K.; Katz, L.; Petzold, C. J.; Keasling, J. D., Divergent Mechanistic Routes for the Formation of gem-Dimethyl Groups in the Biosynthesis of Complex Polyketides. *Angew. Chem. Int. Ed.* **2015**, *54* (8), 2370-2373.

29. Wagner, D. T.; Stevens, D. C.; Mehaffey, M. R.; Manion, H. R.; Taylor, R. E.; Brodbelt, J. S.; Keatinge-Clay, A. T., alpha-Methylation follows condensation in the gephyronic acid modular polyketide synthase. *Chem. Commun.* **2016**, 52 (57), 8822-5.
30. Liu, J.; Zhu, X. J.; Zhang, W. J., Identifying the Minimal Enzymes Required for Biosynthesis of Epoxyketone Proteasome Inhibitors. *Chembiochem* **2015**, 16 (18), 2585-2589.
31. Zettler, J.; Zubeil, F.; Kulik, A.; Grond, S.; Kaysser, L., Epoxomicin and Eponemycin Biosynthesis Involves gem-Dimethylation and an Acyl-CoA Dehydrogenase-Like Enzyme. *Chembiochem* **2016**, 17 (9), 792-798.
32. Poch, G. K.; Gloer, J. B., Helicascolide-a and Helicascolide-B - New Lactones from the Marine Fungus *Helicascus-Kanaloanus*. *J. Nat. Prod.* **1989**, 52 (2), 257-260.
33. Martin, J.; Crespo, G.; Gonzalez-Menendez, V.; Perez-Moreno, G.; Sanchez-Carrasco, P.; Perez-Victoria, I.; Ruiz-Perez, L. M.; Gonzalez-Pacanowska, D.; Vicente, F.; Genilloud, O.; Bills, G. F.; Reyes, F., MDN-0104, an Antiplasmodial Betaine Lipid from *Heterospora chenopodii*. *J. Nat. Prod.* **2014**, 77 (9), 2118-2123.
34. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, 215 (3), 403-410.
35. McGinnis, S.; Madden, T. L., BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res* **2004**, 32, W20-W25.
36. Marchler-Bauer, A.; Lu, S. N.; Anderson, J. B.; Chitsaz, F.; Derbyshire, M. K.; DeWeese-Scott, C.; Fong, J. H.; Geer, L. Y.; Geer, R. C.; Gonzales, N. R.; Gwadz, M.; Hurwitz, D. I.; Jackson, J. D.; Ke, Z. X.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Mullokandov, M.; Omelchenko, M. V.; Robertson, C. L.; Song, J. S.; Thanki, N.; Yamashita, R. A.; Zhang, D.

- C.; Zhang, N. G.; Zheng, C. J.; Bryant, S. H., CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **2011**, *39*, D225-D229.
37. Marchler-Bauer, A.; Derbyshire, M. K.; Gonzales, N. R.; Lu, S. N.; Chitsaz, F.; Geer, L. Y.; Geer, R. C.; He, J.; Gwadz, M.; Hurwitz, D. I.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Song, J. S.; Thanki, N.; Wang, Z. X.; Yamashita, R. A.; Zhang, D. C.; Zheng, C. J.; Bryant, S. H., CDD: NCBI's conserved domain database. *Nucleic Acids Res* **2015**, *43* (D1), D222-D226.
38. Edgar, R. C., MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **2004**, *32* (5), 1792-1797.
39. Edgar, R. C., MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *Bmc Bioinformatics* **2004**, *5*, 1-19.
40. Cheetham, J.; Dehne, F.; Pitre, S.; Rau-Chaplin, A.; Taillon, P. J., Parallel CLUSTAL W for PC clusters. *Lect Notes Comput Sc* **2003**, *2668*, 300-309.
41. Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G., Clustal W and clustal X version 2.0. *Bioinformatics* **2007**, *23* (21), 2947-2948.
42. Nei, M., Phylogenetic analysis in molecular evolutionary genetics. *Annu Rev Genet* **1996**, *30*, 371-403.
43. Tamura, K.; Dudley, J.; Nei, M.; Kumar, S., MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* **2007**, *24* (8), 1596-1599.
44. Grigoriev, I. V.; Nordberg, H.; Shabalov, I.; Aerts, A.; Cantor, M.; Goodstein, D.; Kuo, A.; Minovitsky, S.; Nikitin, R.; Ohm, R. A.; Otilar, R.; Poliakov, A.; Ratnere, I.; Riley, R.; Smirnova, T.; Rokhsar, D.; Dubchak, I., The Genome Portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Res* **2012**, *40* (D1), D26-D32.



45. Nordberg, H.; Cantor, M.; Dusheyko, S.; Hua, S.; Poliakov, A.; Shabalov, I.; Smirnova, T.; Grigoriev, I. V.; Dubchak, I., The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res* **2014**, *42* (D1), D26-D31.
46. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E., The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **2015**, *10* (6), 845-858.
47. Lee, K. K. M.; Da Silva, N. A.; Kealey, J. T., Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. *Anal Biochem* **2009**, *394* (1), 75-80.
48. Sambrook, J.; Russell, D. W.; Sambrook, J., *The condensed protocols from Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., 2006; p v, 800 p.
49. Xu, W.; Cai, X. L.; Jung, M. E.; Tang, Y., Analysis of Intact and Dissected Fungal Polyketide Synthase-Nonribosomal Peptide Synthetase in Vitro and in *Saccharomyces cerevisiae*. *Journal of the American Chemical Society* **2010**, *132* (39), 13604-13607.
50. Puigbo, P.; Guzman, E.; Romeu, A.; Garcia-Vallve, S., OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res* **2007**, *35*, W126-W131.
51. Gibson, D. G.; Benders, G. A.; Axelrod, K. C.; Zaveri, J.; Algire, M. A.; Moodie, M.; Montague, M. G.; Venter, J. C.; Smith, H. O.; Hutchison, C. A., One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *P Natl Acad Sci USA* **2008**, *105* (51), 20404-20409.
52. Yu, X.; Liu, F.; Zou, Y.; Tang, M. C.; Hang, L.; Houk, K. N.; Tang, Y., Biosynthesis of Strained Piperazine Alkaloids - Uncovering the Concise Pathway of Herquiline A. *J. Am. Chem. Soc.* **2016**.

53. Rufer, A. C.; Thoma, R.; Benz, J.; Stihle, M.; Gsell, B.; De Roo, E.; Banner, D. W.; Mueller, F.; Chomienne, O.; Hennig, M., The crystal structure of carnitine palmitoyltransferase 2 and implications for diabetes treatment. *Structure* **2006**, *14* (4), 713-23.

# Chapter 3

## **Mechanistic Investigation of the Cofactor-Independent, Reversible Transacylation of Polyketide Oxyesters to Pantetheine Thioesters Catalyzed by a Carnitine Acyltransferase-Type Domain**

- 3.1 Proposed Thioester Formation Catalyzed by a cAT-Type Domain in Polyketide Biosynthesis**
- 3.2 The Kinetic Competition Model of HRPKS and Potential Application to Tv6-931 Biosynthesis**
- 3.3 Kinetic Competition Between MT (Methylation) and cAT Domain (Esterification)**
- 3.4 Evaluation of cAT-catalyzed Oxyester-Thioester Transacylation (Polyketide “Recapture”)**
- 3.5 Chemoenzymatic Functionalization of Polyketides Using Tv6-931 Enzymes**
- 3.6 Coupling Tv6-931 HRPKS-cAT with a Downstream HRPKS for Polyketide Functionalization**
- 3.7 Discussion**
- 3.8 Supplementary Schemes**
- 3.9 Supplementary Tables**
- 3.10 Supplementary Figures**
- 3.11 Experimental Methods**
- 3.12 References**

## Section 3.1 Proposed Thioester Formation Catalyzed by a cAT-Type Domain in Polyketide Biosynthesis

As ubiquitous, biological acyl-donors, coenzyme A (CoA) thioesters comprise several of the most important metabolites and building blocks in Nature (**Figure 3.1A**). For example, acetyl-CoA serves as the pivotal entry substrate in the tricarboxylic acid (TCA) cycle in aerobic respiration.<sup>1</sup> In addition, carboxylation of acetyl-CoA generates malonyl-CoA for the biosynthesis of fatty acids and polyketides. Acetyl-CoA and other acyl-CoA molecules also play critical roles in the remodeling of chromatin, modifications of proteins, and regulation of hormones and neurotransmitters.<sup>1-3</sup> Two classic enzymatic syntheses of biological thioesters (acyl-CoA molecules) involves the dehydrogenation of  $\alpha$ -keto-carboxylic acids (NAD<sup>+</sup> dependent), and the ligation of CoA to carboxylic acids (ATP dependent).<sup>4-9</sup> A third enzymatic method to generate acyl-CoA thioesters includes the reversible transacylation between acyl-carnitine and CoA by carnitine acyltransferase (cAT) enzymes (**Figure 3.1B**).<sup>10</sup> Differing from the two classic methods which have several known examples, the cofactor-independent, reversible transacylation is mostly specific to acyl-carnitine and to a lesser extent, acetylcholine.<sup>11-12</sup> Investigation of other homologous examples of reversible, cAT-type transacylation provides opportunities to develop methodologies for the cofactor-independent, enzymatic synthesis of biological thioesters.

In primary metabolism, carnitine acyltransferase (cAT) enzymes play a critical role in the shuttling of fatty-acids across the mitochondrial membrane using reversible transacylation between CoA and carnitine (**Figure 3.1C, Figure S3.1**).<sup>10</sup> From an energetic perspective, the equilibrium constant of the transacylation between CoA (thiol) and carnitine (zwitterionic

alcohol) is surprisingly low with a near equal distribution, aiding in the reaction viability for fatty acid trafficking.<sup>11, 13-14</sup> These enzymes are widespread in aerobic organisms including all mammals, ensuring homeostatic energy production and metabolism of fatty acids.<sup>15</sup> Consequently, mutations or deficiencies in these native cAT enzymes are associated with several diseases and myopathies.<sup>14-16</sup>

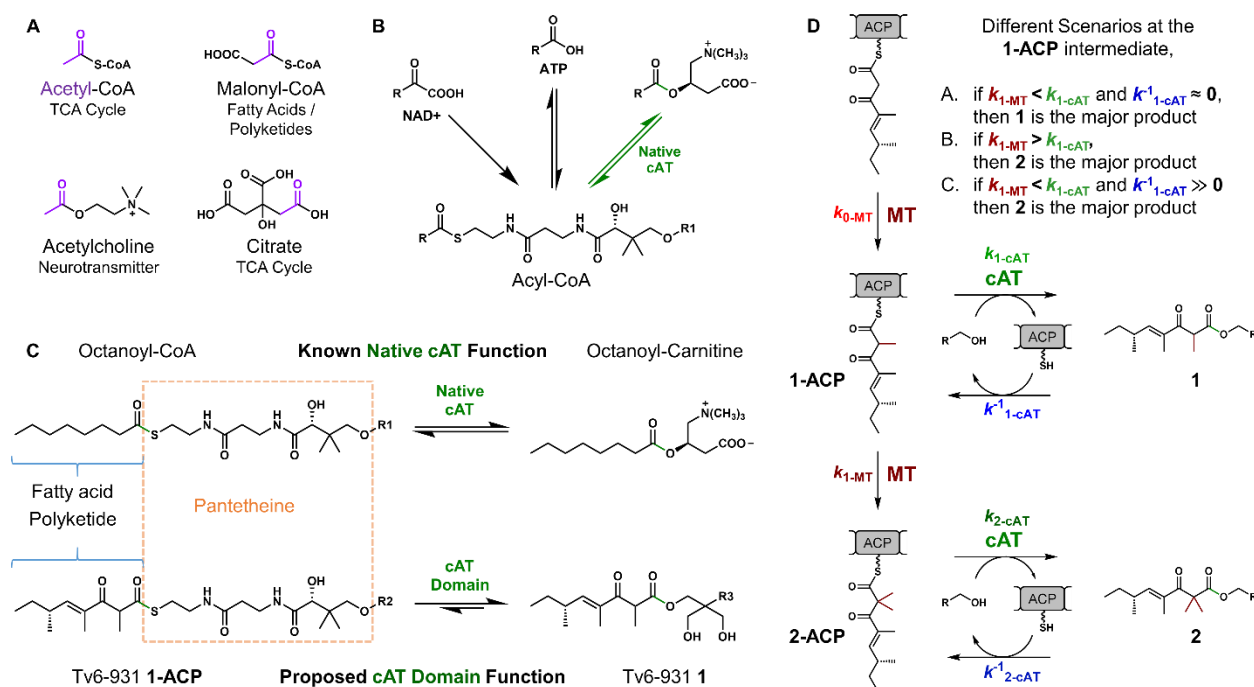
Recently, a family of fungal highly-reducing polyketide synthase (HRPKS) enzymes was discovered to harbor a fused cAT-type domain instead of the canonical thioesterase (TE) partnering enzyme (**Figure S3.2**).<sup>17</sup> In a search of 581 fungal genomes, approximately 100 HRPKS-cAT genes were identified and formed an uncharacterized branch of HRPKS enzymes.<sup>17</sup> HRPKS enzymes catalyze the biosynthesis of diverse and bioactive drugs such as the cholesterol-lowering lovastatin in an enigmatic, iterative fashion, and this appended cAT domain adds another level of complexity.<sup>18</sup>

Due to the structural similarities of the phosphopantetheine derivatives, CoA and the acyl carrier protein (ACP), the cAT domain was postulated to function as a reversible acyltransferase capable of releasing the polyketide adduct from the ACP (**Figure 3.1C**, **Figure S3.1**). HRPKS-cAT type enzymes have been speculated to catalyze the biosynthesis of other interesting polyketides such as the antifungal hypoxysordarin and phytotoxic AF-toxins.<sup>19-21</sup> Understanding the mechanism and enzymology of HRPKS-cATs can provide insight to the programming of the poorly understood HRPKS enzymes in general—importing knowledge from primary mechanism to secondary metabolites and vice versa.

Despite the abundance of HRPKS-cATs (present in ~18% of fungi), the only experimental biosynthesis was reported in 2017.<sup>17</sup> The Tv6-931 HRPKS-cAT was shown to

catalyze the biosynthesis of polyketide-polyol conjugates bearing an  $\alpha,\alpha$ -dimethyl (*gem*-dimethyl) group through a proposed polyketide release-recapture mechanism of **1** (**Figure 3.1D**). To the best of our knowledge, this is the only reported biosynthesis of a fungal polyketide containing the *gem*-dimethyl group. In the proposed mechanism, the reversibility of the fused cAT plays an important role in the formation of the rare *gem*-dimethyl moiety by recapturing the  $\alpha$ -methyl intermediate **1** for an additional methylation. (**Figure 3.1C-3.1D**). The proposed polyketide “recapture” mechanism (oxyester-thioester transacylation) distinguishes it from other known polyketide thioesterase-type (TE) partnering domains and enzymes (**Figure S3.1-S3.2**).

In the canonical TE examples, polyketide products are released from the ACP using a catalytic triad and proceeds through a covalent serine-oxyester intermediate.<sup>22-23</sup> Consequently, the TE-catalyzed release is essentially irreversible. In contrast, the non-covalent nature of cAT-type catalysis results in high reversibility (**Figure 3.1C**, **Figure S3.1**). The cAT-type domain is hypothesized to also catalyze oxyester-thioester transacylation to “recapture” the polyketide-oxyester **1** intermediate for further catalysis. In the proposed biosynthesis catalyzed by the Tv6-931 enzyme, the HRPKS components catalyze the growing polyketide chain (tetraketide) to the  $\alpha$ -methyl-enzyme bound intermediate **1-ACP** (**Figure 3.1D**). In the penultimate step, the **1-ACP** intermediate can be offloaded by a cAT-catalyzed esterification onto a polyol nucleophile to furnish  $\alpha$ -methyl oxyester **1**. Alternatively, the **1-ACP** can be methylated by the methyltransferase (MT) domain to furnish **2-ACP**, which can later be converted to the final *gem*-dimethyl oxyester **2**. Based on kinetic studies of traditional fungal HRPKS, the current paradigm suggests that the kinetic competition and substrate selectivity differences between the catalytic domains will directly dictate the distribution of polyketides **1** and **2** (**Figure S3.3**).<sup>24-25</sup>



**Figure 3.1:** Enzymatic synthesis of biological thioesters and their proposed relationship with polyketide biosynthesis **A)** Illustration of acetyl-CoA, acetyl-CoA derivatives, and their biological role. **B)** Three enzymatic methods to generate the acyl-CoA molecules. The native cAT enzyme method differs due to its reversibility and cofactor-independence. **C)** Comparison of the functions of the native cAT enzyme and the proposed cAT domain where both catalyze the transacylation (esterification) of an acyl-(phospho)pantetheine molecule to an alcohol acyl-acceptor. Only the zwitterionic, acyl-carnitine allows for near equal distribution. **D)** Proposed biosynthesis of  $\alpha,\alpha$ -dimethyl oxyester **2** catalyzed by Tv6-931. The kinetic data suggests that the polyketide recapture (oxyester-thioester transacylation  $k^{-1}_{1-cAT}$ ) in scenario C is necessary for the biosynthesis of **2** as the major product.

## Section 3.2 The Kinetic Competition Model of HRPKS and Potential

### Application to Tv6-931 Biosynthesis

The current paradigm of HRPKS programming suggests that at a particular step of the biosynthesis, the chemical outcome reflects the relative difference in kinetic rates between catalytic domains (**Figure S3.3**). For example, in the biosynthesis catalyzed by lovastatin nonaketide synthase LovB, methylation is dictated by the kinetic competition between the MT and ketoreductase (KR) domain.<sup>24</sup> Moreover, in the biosynthesis catalyzed by squalastatin

tetraketide synthase (SQTKS), the enoylreductase (ER) domain completely discriminates against the matured tetraketide intermediate compared to the diketide and triketide substrates.<sup>25</sup>

If this model was applied to the Tv6-931 biosynthesis and in the absence of cAT-catalyzed “recapture,” two possible scenarios would be envisioned (**Figure 3.1D**). In Scenario A, if the rate of methylation ( $k_{1\text{-MT}}$ ) was slower than the release of the **1-ACP** intermediate ( $k_{1\text{-cAT}}$ ), then the  $\alpha$ -methyl adduct **1** would be the major product. Conversely in Scenario B, if methylation of **1-ACP** ( $k_{1\text{-MT}}$ ) occurs faster than release ( $k_{1\text{-cAT}}$ ), then **2** would be the major product. In these reasonably straightforward scenarios, the product distribution reflects the relative kinetic rates of the MT and cAT domain for the **1-ACP** substrate.

However, both Scenario A and B have issues. Experimentally, the *gem*-dimethyl adduct **2** has been observed as the major product both *in vivo* and *in vitro*, which is inconsistent with Scenario A.<sup>17</sup> Yet, the *gem*-dimethyl containing polyketides are extremely rare. The bacterial MT domains responsible for gephyronic acid biosynthesis demonstrate exceedingly slow *gem*-dimethylation kinetic rates, casting doubt on Scenario B.<sup>26</sup>

In the previous cases with LovB and SQTKS, the kinetic competition model previously applied to catalytic domains that catalyzed irreversible reactions (**Figure S3.3**). In contrast, the proposed biosynthesis of **2** involves the cAT domain, which catalyzes reversible transacylation. Consequently, the kinetic competition of canonical HRPKSs may not sufficiently account for the biosynthesis of the Tv6-931 HRPKS-cAT enzyme.

Accordingly, Scenario C has been envisioned: the rate of methylation ( $k_{1\text{-MT}}$ ) is slower than the release of the **1-ACP** intermediate ( $k_{1\text{-cAT}}$ ) leading to the polyketide intermediate **1** to accumulate. However, an appreciable and kinetically viable transacylation or “recapture” of **1**



onto the phosphopantetheine tether of the ACP regenerates **1-ACP** (**Figure 3.1C**, **Figure 3.2**). Although the second methylation of **1-ACP** to **2-ACP** may be kinetically slow, **1** will gradually, but irreversibly convert to **2**. To the best of our knowledge, Scenario C would represent the first known example of a cAT-type mechanism on a neutral ester substrate.

Nevertheless, the proposed Scenario C has an energetic dilemma. In the reversible transacylation of acyl-CoA molecules with carnitine, the acyl-carnitine product retains surprisingly high transacylation potential probably due to solvation of the zwitterionic carnitine (**Figure 3.1C**).<sup>11, 13</sup> This plays favorably in its reversibility with an equilibrium constant near 1. In contrast, the neutral acyl-polyol **1** should exhibit significantly lower transacylation potential as a typical oxyester.<sup>27</sup> Consequently, the conversion and subsequent activation of the polyketide **1** to **1-ACP** should be thermodynamically uphill (**Figure 3.1C**). It remains to be seen if the recapture reaction ( $k_{1-cAT}^{-1*}$ ) is kinetically viable and necessary for the biosynthesis of **2**. Quantifying the kinetic parameters will resolve the different scenarios of the Tv6-931 HRPKS-cAT catalysis and will potentially highlight the key differences with canonical HRPKS enzymes.

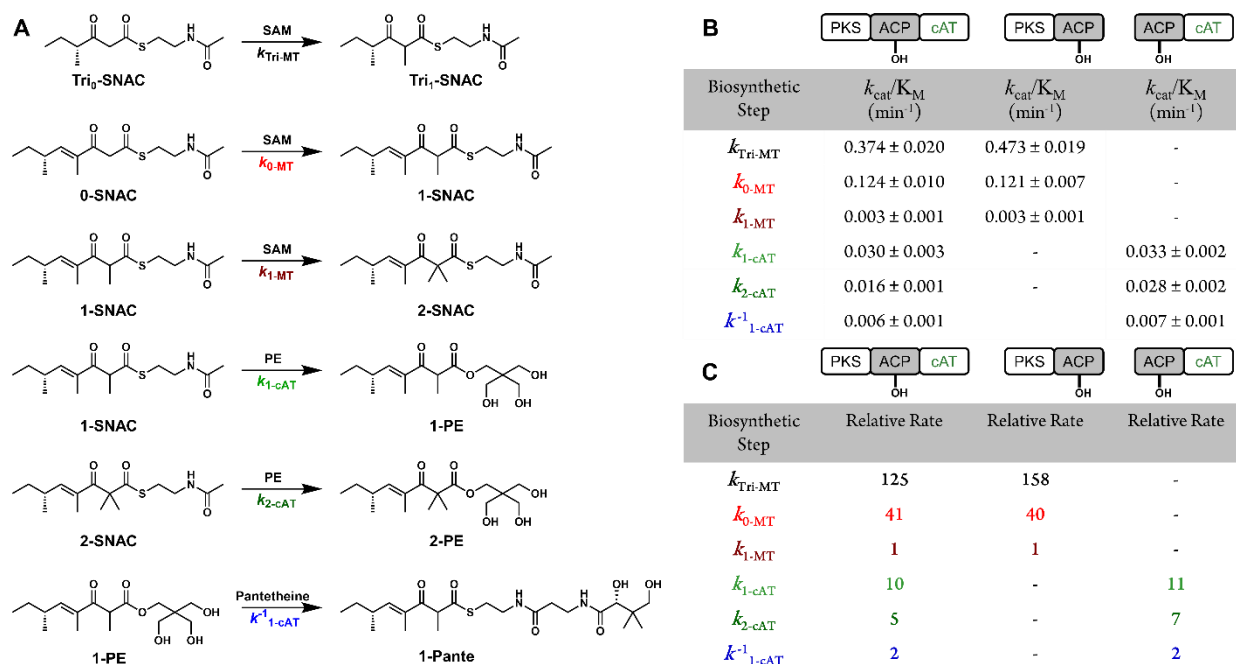
### **Section 3.3 Kinetic Competition Between MT (Methylation) and cAT Domain (Esterification)**

Accordingly, a panel of *N*-acetylcysteamine (SNAC) thioester chemical probes were synthesized as substrates of enzyme kinetic experiments of the MT and cAT domains (**Scheme S3.1-S3.6**, **Table S3.1**). These chemical probes were designed to mimic the putative ACP-bound thioester intermediates in the biosynthesis of Tv6-931 polyketides (**Figure 3.1C-3.1D**, **Figure 3.3A**). The synthesis of both enantiomers of **2-SNAC** also helped determine the stereochemistry of the distal,  $\epsilon$ -methyl group of **2** as the *R* configuration (**Figure S3.4**).

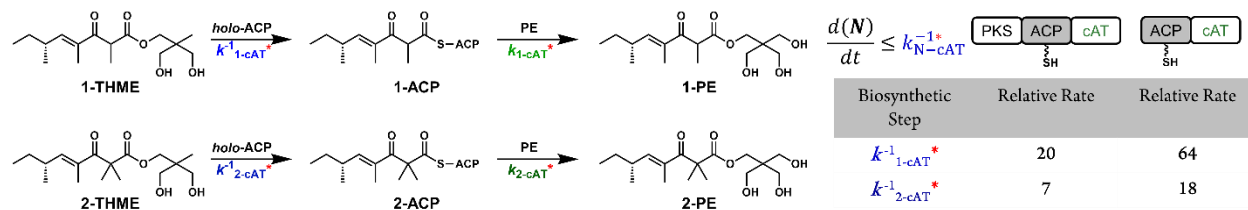
Briefly, the majority of the  $\beta$ -ketoacyl-SNAC compounds were prepared using titanium-catalyzed Aldol chemistry to furnish the  $\beta$ -hydroxy-carbonyl precursors. These  $\beta$ -hydroxy-carbonyl precursor compounds were elaborated as required then oxidized to provide the desired  $\beta$ -ketoacyl-SNAC compounds (**Scheme S3.1-S3.6**). The methyl groups in the  $\alpha$ -position of the  $\beta$ -ketoacyl-SNAC compounds were introduced during the Aldol condensation or immediately following the oxidation of the  $\beta$ -hydroxy precursors (see Section 3.11 Experimental Methods). The chiral aldehydes used for the Aldol chemistry were prepared by stereoselective  $\alpha$ -methylation on the requisite carbon backbone. The selectivity of these methylation reactions was conferred through the use of the appropriate Evans' chiral auxiliary ((*R/S*)-4-Benzyl-2-oxazolidinone), which were reductively cleaved to afford the desired enantiomerically pure  $\alpha$ -methyl alcohols. These were then oxidized under specific conditions to afford the desired aldehydes (see Section 3.11 Experimental Methods).

For these kinetic experiments, the Tv6-931 enzyme variants were transformed in *Saccharomyces cerevisiae* (BJ5464-NpgA) using yeast 2 $\mu$  plasmid vectors.<sup>17</sup> The various Tv6-931 gene constructs were expressed under the alcohol dehydrogenase 2 (ADH2) promoter in high abundance (**Figure S3.5**). For kinetic methylation ( $k_{X-MT}$ ) and esterification ( $k_{X-cAT}$ ) experiments, the *apo*-form of the Tv6-931 HRPKS-cAT enzymes were used to prevent complications with the phosphopantetheine tether of the ACP. The *holo*-form of the enzyme also exhibited higher, non-specific thioesterase activity which would interfere with the results (**Figure S3.10-S3.11, Table 3.1**). Although measuring the kinetic rates using SNACs underestimates the absolute kinetic efficiency of the enzyme, the *relative* rates and trends are informative to compare the efficiencies of the different catalytic domains.<sup>24-25, 28</sup>

For kinetic methylation ( $k_{X-MT}$ ) and esterification ( $k_{X-eAT}$ ) experiments, kinetic enzyme turnover and time course experiments were conducted under pseudo-first order kinetics with the methylating agent, *S*-adenosylmethionine (SAM), and the polyol nucleophile, pentaerythritol (PE), in five equivalents compared to the SNAC substrates (**Figure 3.2, Figure S3.6-S3.7**). While increasing the concentration of SAM [and polyol nucleophile] still linearly increased the kinetic rates, five equivalents (2.5 mM) was chosen due to cost considerations of SAM.



**Figure 3.2:** Kinetic experiments using SNAC probes with MT and cAT Domains **A**) Enzymatic reactions using SNAC thioester chemical probes and **1-PE** used to measure the various kinetic rates related to the Tv6-931 biosynthesis. **B**) Kinetic rates ( $k_{\text{cat}}/K_M$ , min<sup>-1</sup>) extrapolated from the saturation curve kinetic reactions. **C**) Relative kinetic rates normalized to the  $k_{1-MT}$ .



**Figure 3.3:** Kinetic transesterification experiments to estimate polyketide recapture **A**) Alternative measurements of the polyketide “recapture” rate or transacylation of oxyesters **1** and **2** to form polyketide-ACP adducts, **1-ACP** and **2-ACP**. The formation of the polyketide-ACP adducts were estimated by using an overall transesterification.

a) *Enzymatic Kinetic Assays (MT vs cAT): whole protein*

Under these conditions, the various parameters at the penultimate steps of the biosynthesis were measured (**Figure 3.1C**). In the kinetic methylation experiments, the conversions of **Tri<sub>0</sub>-SNAC** to **Tri<sub>1</sub>-SNAC** and also **0-SNAC** to **1-SNAC** proceeded readily ( $k_{\text{Tri-MT}}$  and  $k_{0\text{-MT}}$ ). However, the methylation of **1-SNAC** to **2-SNAC** was shown to be significantly slower (**Figure 3.2**), requiring extended time intervals for minimal conversion (**Figure S3.6**). These results were consistent with MT domains of bacterial PKSs that catalyzed dimethylation very slowly.<sup>26</sup> When comparing the results of the release (esterification) of the SNAC compounds, the esterification rates ( $k_{1\text{-cAT}}$  and  $k_{2\text{-cAT}}$ ) by the cAT domain are comparatively much slower than the methylation rates of **Tri<sub>0</sub>-SNAC** to **Tri<sub>1</sub>-SNAC** and the conversion of **0-SNAC** to **1-SNAC** ( $k_{\text{Tri-MT}}$  and  $k_{0\text{-MT}}$ ). Nevertheless, esterification rates are still significantly faster than the conversion of **1-SNAC** to **2-SNAC** ( $k_{1\text{-MT}}$ ). The kinetic data indicates that the rate-determining step and bottleneck of the biosynthesis is the second methylation at the  $\alpha$ -position.

b) *Enzymatic Kinetic Assays (MT vs cAT): individual components*

To delineate the role of the MT and cAT domain from each other, the HRPKS-cAT were dissected to their individual HRPKS and cAT components. Due to enzyme solubility considerations, the dissected cAT component was purified as the ACP-cAT di-domain. The methylation rates at the tetraketide steps were fairly consistent whether using the whole Tv6-931 protein or the  $\Delta\text{cAT}$  variant (**Figure 3.2B-3.2C**). Using the dissected domain, the esterification rates (release) of the **1-SNAC** ( $k_{1\text{-cAT}}$ ) and **2-SNAC** ( $k_{2\text{-cAT}}$ ) exhibited marginal to moderate improvements, respectively, compared to the whole protein. Nevertheless, both esterification rates pale in comparison to the methylation rates of **Tri<sub>0</sub>-SNAC** and **0-SNAC** ( $k_{\text{Tri-MT}}$  and  $k_{0\text{-MT}}$ ),

but significantly exceeds the methylation of **1-SNAC** ( $k_{1\text{-MT}}$ ). Consequently, the relative trends of the kinetic rates were consistent and clear: the rate-limiting step of the biosynthesis is the second methylation event at the  $\alpha$ -carbon (**Figure 3.2, Figure S3.6-S3.7**). The data acquired from using the whole protein and the dissected components both show that  $k_{1\text{-MT}} < k_{1\text{-cAT}}$ , which directly conflicts with Scenario B (**Figure 3.1D**).

### **Section 3.4 Evaluation of cAT-catalyzed Oxyester-Thioester Transacylation (Polyketide “Recapture”)**

#### a) *Measuring Direct Acylation of Pantetheine (holo-ACP Surrogate)*

After excluding the likelihood of Scenario B, the difference between Scenarios A and C is the viability of the “recapture” reaction, where the polyketide-oxyester is converted to the higher-energy polyketide-ACP thioesters, **1-ACP** and **2-ACP** (**Figure 3.1D**). Considering the recapture reaction is ACP-dependent, measuring the direct acylation of the ACP may be more difficult compared to methylation and esterification. Nevertheless, pantetheine and SNAC are both *holo*-ACP mimics, and we envisioned that the cAT domain could use these ACP mimics to “recapture” **1** via oxyester-thioester transacylation. Accordingly, incubation of **1-PE**, *apo*-HRPKS-cAT or *apo*-ACP-cAT Tv6-931 enzyme, with pantetheine and SNAC produce the corresponding thioesters **1-Pante** and **1-SNAC**, in low to moderate quantities, respectively (**Figure S3.12, Table 3.1B**).

The enzymatic production of **1-Pante** using the *apo*-ACP-cAT provides direct evidence for the conversion of the polyketide-oxyester to the polyketide-pantetheine thioester as proposed in the “recapture” mechanism (**Figure 3.1C-3.1D**). Subsequently, kinetic studies were performed with both the *apo*-HRPKS-cAT and the *apo*-ACP-cAT. For both enzyme variants, **1-PE** was

incubated with pantetheine according to the same pseudo-first order kinetics as the methylation and release assays (**Figure 3.2, Figure S3.8**). Despite the equilibrium limitation of directly converting an oxyester to a thioester (**Figure 3.1C**), the rate of the recapture ( $k_{1-\text{cAT}}^{-1}$ ) to form the **1-Pante** adduct was comparable to the second methylation ( $k_{1-\text{MT}}$ ). Together, the kinetic data supports Scenario C that the rate-limiting second methylation ( $k_{1-\text{MT}}$ ) necessitates that a kinetic competent enzymatic recapture ( $k_{1-\text{cAT}}^{-1}$ ) to regenerate the **1-ACP** intermediate. The polyketide “recapture” ensures that the *gem*-dimethyl adduct **2** is the major biosynthetic product—consistent with *in vivo* and *in vitro* experiments (**Figure 3.1D, Figure 3.2C**).

#### b) *Measuring Enzymatic Transesterification Rate*

Although using pantetheine as a chemical surrogate for the ACP provides a convenient and fair estimation of  $k_{1-\text{cAT}}^{-1}$  value to evaluate the biosynthesis (**Figure 3.2A**), we suspect that it significantly underestimates the cAT domain’s ability to recapture a polyketide to the bona fide substrate (*holo*-ACP). To conservatively estimate the rate of polyketide-ACP formation, the recapture mechanism (transacylation to ACP) was coupled to a subsequent esterification for an overall transesterification (**Figure 3.3**). In the transesterification, the compounded rate denoted as  $\frac{d(N)}{dt}$  will be correlated to the rate of formation of the polyketide-ACP adduct,  $k_{\text{N-cAT}}^{-1*}$ , since the the rate of polyketide-ACP formation cannot proceed slower than overall transesterification. If the formation of the polyketide-ACP adduct is rate limiting in the transesterification, then  $\frac{d(N)}{dt}$  conservatively approximates the  $k_{\text{N-cAT}}^{-1*}$  rate. Since the transesterification is a compound rate and requires a thioester linkage to the *holo*-ACP, this “recapture” rates will be denoted as  $k_{\text{N-cAT}}^{-1*}$  to highlight the differences.

Using the Tv6-931 *holo*-HRPKS-cAT enzyme, the trishydroxymethylethane (THME) adducts, **1-THME** and **2-THME**, were incubated with PE and monitored for the formation of **1-PE** and **2-PE**, respectively (**Figure 3.3**, **Figure S3.9**). The formation of **1-PE** and **2-PE**,  $k_{1\text{-cAT}}^{-1*}$  and  $k_{2\text{-cAT}}^{-1*}$ , is significantly improved compared to the direct acylation of pantetheine and proceeds at a rate at least comparable to  $k_{1\text{-cAT}}$  and  $k_{2\text{-cAT}}$ . Notably, the trend of  $k_{1\text{-cAT}}^{-1*}$  and  $k_{2\text{-cAT}}^{-1*}$  are similar to  $k_{1\text{-cAT}}$  and  $k_{2\text{-cAT}}$  in that the reactions with the  $\alpha$ -methyl adducts proceeded more quickly than their *gem*-dimethyl counterparts. This data further supports that the “recapture” rate is kinetically relevant in the biosynthesis. Although the rates of polyketide “recapture” ( $k_{\text{N-cAT}}^{-1*}$ ) using *holo*-enzyme cannot be directly compared to the methylation and esterification with SNAC and *apo*-enzyme, the significant levels of transacylation still provides compelling evidence that polyketide “recapture” contributes to the biosynthesis since the rate of *gem*-dimethyl formation ( $k_{1\text{-MT}}$ ) is slow (**Figure 3.1D** Scenario C, **Figure 3.3C**).

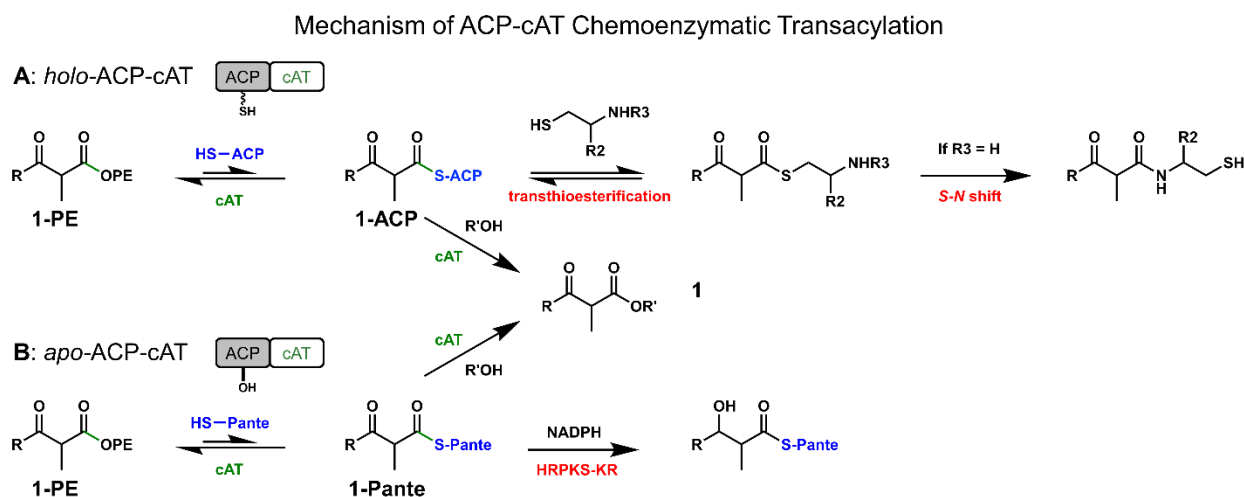
## Section 3.5 Chemoenzymatic Functionalization of Polyketides Using Tv6-931

### Enzymes

Similar to the esterification reactions ( $k_{1\text{-cAT}}$  and  $k_{2\text{-cAT}}$ ), the transesterification results, ( $k_{1\text{-cAT}}^{-1*}$  and  $k_{2\text{-cAT}}^{-1*}$ ) proceeds faster with the  $\alpha$ -methyl adduct compared to its *gem*-dimethyl counterparts (**Figure 3.3**). Interestingly, unlike the methylation and esterification assays where whole and dissected proteins gave similar responses, the dissected protein exhibited significantly faster polyketide transesterification compared to the whole enzyme (**Figure 3.3**). We speculate that the differences can be attributed to factors such as a complicated, compound kinetic rate and using the *holo*-enzyme form instead of the *apo* form. Although the origin of the rate enhancement is currently not well-determined, it provides an opportunity to readily access and



trap the high-energy polyketide-ACP intermediates (**1-ACP** and **2-ACP**) using chemoenzymatic methods (**Scheme 3.1A**). Combined with the direct acylation of pantetheine (**Scheme 3.1B**), the Tv6-931 enzyme provides two methods to access higher-energy thioester intermediates (*holo*-ACP and pantetheine).



**Scheme 3.1:** ACP-cAT-catalyzed transacylation of **1-PE**. **A)** Using the *holo*-ACP-cAT enzyme, the **1-ACP** intermediate can be chemically trapped by other thiols to furnish the corresponding transthioesterification adduct or enzymatically functionalized to produce the corresponding oxyester **1**. **B)** Using the *apo*-ACP-cAT enzyme, pantetheine can be used a substrate (ACP-surrogate) for direct transaction to produce **1-Pante** which can be enzymatically functionalized to **1** or further functionalized by other HRPKS enzymes (KR domain used as an example).

#### *Chemoenzymatic Trapping of the Acyl-ACP Intermediates (holo-ACP-cAT)*

Although the transacylation and transesterification reactions (**Figure 3.1D**, **Figure 3.3**) have been shown to be cAT and *holo*-ACP dependent, the **1-ACP** and **2-ACP** intermediates have not yet been chemically trapped.<sup>17</sup> Capturing the high-energy polyketide-ACP adduct for derivatization provides additional evidence for the presence of the key reactive enzyme-bound intermediate **1-ACP** and opens opportunities for the Tv6-931 enzyme and potentially other HRPKS-cAT enzymes as biocatalysts for polyketide functionalization.

A natural choice for trapping the thioester **1-ACP** is with a second thiol for thiol-thioester exchange (**Scheme 3.1A**, **Scheme S3.7**). Thioester exchange usually proceeds in a facile manner, kinetically outpaces spontaneous hydrolysis, and serves as an important reaction in biological processes.<sup>29</sup> When **1-PE** was incubated with the *holo*-ACP-cAT and SNAC, **1-SNAC** was produced in low yields and readily detected by LCMS (**Table 3.1A** Entry 2a, **Figure S3.10**). Accordingly, other free thiols such as *N*-acetylcysteine and pantetheine also produced the corresponding thioester in low yields (**Table 3.1A** Entry 2, **Figure S3.10**). As expected, the total conversion of the reversible reaction is greatly limited by the equilibrium distribution between oxyesters and thioesters (**Figure 3.1C**).

Consequently, we aimed to overcome these energetic limitations by coupling the energetically unfavorable thioesterification with an exergonic reaction to shift the overall reaction equilibrium distribution. One commonly employed technique to drive a transthioesterification reaction is to couple the reaction with a chemical *S-N* shift to produce the corresponding amide—the energetic basis for native chemical ligation (NCL) (**Scheme 3.1A**).<sup>30</sup> Belecki *et al* used an analogous strategy and cysteamine in high concentration to intercept and detect several PKS enzyme-bound intermediates in the biosynthesis of the calicheamicin.<sup>31</sup>

Compared to the calicheamicin example, a key distinction of this mechanism is the ability to load a specific polyketide group onto the ACP using the cAT and consequently, to intercept a distinct biosynthetic intermediate (**1-ACP**). Incubation of **1-PE** with *holo*-ACP-cAT and various cysteamine-derivatives capable of *S-N* shift resulted in significantly improved transacylation yields (**Table 3.1A** Entry 3, **Figure S3.10**). Methylation of the sulfur atom of the cysteamine nucleophiles completely abolishes transacylation suggesting that the mechanism proceeds

through the transthioesterification and *S-N* shift cascade instead of direct amidation (**Table 3.1A** Entry 4, **Scheme 3.1A**).

Although the amide functionalization was reasonable, the polyketide thioester adducts were only produced in low yields. Instead, the major product was the shunt hydrolyzed product **1-COOH** in most cases (**Table 3.1A**, **Scheme S3.7**). In the absence of a kinetically suitable nucleophile, the cAT also catalyzes hydrolysis of the ACP-bound thioester (**Table 3.1A**, Entry 1 and 4). Conversely, if **1-ACP** undergoes cAT-catalyzed esterification, hydrolysis is barely observed. (**Table 3.1A** Entry 5). As anticipated from the kinetic transesterification assays ( $k_{1-cAT}^{-1*}$ ), the enzymatic transesterification proceeds readily and in high yield.

**A**

**B**

Entry	Nucleophile	Transacylation (%)	Hydrolysis (%)
1	H <sub>2</sub> O	-	47
2a	HS-CH <sub>2</sub> -NHAc	5	34
2b	Pantetheine	4	33
2c	HS-CH <sub>2</sub> -NHAc   COOH	6	37
3a	HS-CH <sub>2</sub> -NH <sub>2</sub>	22	19
3b	HS-CH <sub>2</sub> -NH <sub>2</sub>   COOH	25	33
3c	HS-CH <sub>2</sub> -NH <sub>2</sub>   COOH   NH <sub>2</sub>	27	28
3d	HS-CH <sub>2</sub> -NH <sub>2</sub>   COOMe   NH <sub>2</sub>	50	5
4a	S-CH <sub>2</sub> -NH <sub>2</sub>   COOH	N.D.	62
4b	S-CH <sub>2</sub> -NH <sub>2</sub>   COOH   NH <sub>2</sub>	N.D.	59
5a	HO-CH <sub>2</sub> -NH <sub>2</sub>   OH   OH	80	2
5b	HO-CH <sub>2</sub> -NH <sub>2</sub>   OH   OH   OH	93	2

Entry	Nucleophile	Transacylation (%)	Hydrolysis (%)
1	H <sub>2</sub> O	-	N.D.
2a	HS-CH <sub>2</sub> -NHAc	5	N.D.
2b	Pantetheine	20	2
2c	HS-CH <sub>2</sub> -NHAc   COOH	N.D.	N.D.
3-5	See Table 2A	N.D.	N.D.

**Table 3.1:** Transacylation of **1-PE** by various nucleophiles using **A**) *holo*-ACP-cAT (forming **1-ACP in situ**) and **B**) *apo*-ACP-cAT. Entry 1: Water. Entry 2: Thiols not capable of *S-N* shift. Entry 3: Thiols capable of *S-N* shift. Entry 4: *S*-methyl nucleophiles. Entry 5: Polyol nucleophile substrates accepted by the cAT domain. Note: entry 1a (SNAC) and 1b (Pantetheine) are poor and improved chemical mimics of *holo*-ACP, respectively. N.D. = Not Detected.

Similar to **1-PE**, chemoenzymatic transacylation reactions were also carried out with the *gem*-dimethyl **2-PE** with the *holo*-ACP-cAT enzyme (**Scheme S3.8A**). Using the **2-PE** acyl donor, the chemical transthioesterification-based transacylation reactions proceeded with less efficiency (**Table S3.5, Figure S3.11**). Significant Michael-addition side products were now formed possibly accelerated by the *gem*-dimethyl group. Analogous to the **1-PE** example, the enzymatic esterification of **2-PE** also proceeded in a facile manner compared to their chemical transthioesterification counterpart (**Figure S3.11**).

a) *Enzymatic Synthesis of Polyketide-Pantetheine Derivative (apo-ACP-cAT)*

In parallel to the enzymatic reactions with the *holo*-ACP-cAT, negative control reactions were conducted with the *apo*-ACP-cAT and without any enzyme. Based on the mechanism highlighted in **Scheme 3.1A**, the **1-ACP** intermediate would not be formed, subsequent transthioesterification would not take place, and transacylation should be abolished. Accordingly, the negative controls without enzyme exhibited no transacylation. For the *apo*-ACP-cAT, transacylation was also abolished in all cases except for SNAC and pantetheine (**Table 3.1B, Figure S3.12**). Unlike the other nucleophiles, SNAC and pantetheine both mimic the *holo*-ACP as evident in the “recapture” kinetic assays. Interestingly, when the *apo*-ACP-cAT enzyme is supplemented with exogenous pantetheine to serve as a *holo*-ACP surrogate, the enzymatic transesterification and chemical transthioesterification activity are restored (**Scheme S3.7, Figure S3.13-S3.14**). Therefore, the transacylation of **1-PE** with pantetheine and SNAC catalyzed by *apo*-ACP-cAT should proceed directly on the substrate in a single enzymatic step unlike the two-step process using the *holo*-ACP-cAT (**Scheme 3.1**).

We hypothesize that the improvement of catalysis for pantetheine with the *apo*-ACP-cAT compared to the *holo*-ACP-cAT can be due to the preferred binding of the phosphopantetheine tether of the *holo*-ACP to the cAT active site compared to exogenous pantetheine. Consequently, the transacylation of **1-PE** would react faster with the *holo*-ACP (native substrate), and the formation of **1-ACP** would greatly outcompete the formation of **1-Pante**. For both *apo*-HRPKS-cAT and the *apo*-ACP-cAT enzymes, the formation of **1-ACP** greatly outpaces the formation of **1-Pante**, further supporting the substrate preference hypothesis (**Figure 3.3, Figure S3.8**). This result intuitively explains why the *apo*-ACP-cAT catalyzes the formation of **1-Pante** in higher yields than the *holo* form. A similar phenomenon was observed with a *trans*-acting  $\psi$ -ACP-MT didomain in the biosynthesis of tricholignan A, and the *apo* state of the  $\psi$ -ACP was critical to the catalysis.<sup>32</sup> Conversion of the *apo*- $\psi$ -ACP-MT to the *holo* state nearly abolished the methylation activity, probably due to competition and interference with the inline PKS-ACP.

Since the “recapture” kinetic results and transacylation data were promising, we explored the *apo*-ACP-cAT enzymes as viable catalysts to produce polyketide-pantetheine derivatives. After minor optimization, we observe 28% yield (one-pot) using 0.5 mM **1-PE**, 10 mM pantetheine, and 30  $\mu$ M *apo*-ACP-cAT (6 mol%) (**Scheme S3.9**). Despite the equilibrium limitation of converting an oxyester to thioester (**Figure 3.1C**), the reaction produces modest, but reasonable yields. In addition, pantetheine can be readily produced from the reduction of the inexpensive and commercially-available pantethine (disulfide) in high yields. Moreover, both the starting acyl-donor (**1-PE**) and enzyme catalyst (*apo*-ACP-cAT) can be harvested from simple yeast fermentation at >10 mg/L.<sup>33-34</sup>

Although the conversion percentage is modest, the mild conditions allow for recovery of the majority of the starting material (**Scheme S3.9**), leading to an overall, higher-adjusted yield

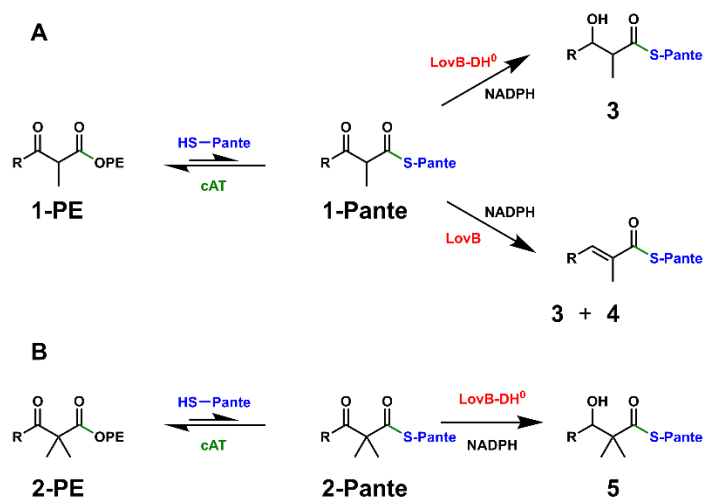
(28% yield in 45% conversion, >60% adjusted yield). In addition, this methodology represents one of the few chemoenzymatic strategies to produce polyketide pantetheine derivatives. While there are many chemical syntheses towards polyketide-pantetheine derivatives, they typically require several synthetic steps.<sup>28, 35</sup> One synthetic steps may involve the hydrolysis of **1** to **1-COOH**, which is prone to decarboxylation. While there are also other efficient, chemoenzymatic strategies towards polyketide-CoA derivatives, they also requires multiple, synthetic steps [and enzymes].<sup>35</sup> In contrast, this methodology proceeds with one enzyme in one-step from the commercially-available pantetheine or pantethine (disulfide) without requiring other cofactors or protecting group manipulations. Accordingly, the ability to chemoenzymatically and quickly synthesize these polyketide-pantetheine conjugates even in low yields provides a complementary and accessible method to probe PKS enzymology.

Using the **2-PE** and **0-PE**, the catalysis and chemistry was also executed to produce the analogous **2-Pante** and **0-Pante**, albeit with some diminished yield and catalytic efficiency (**Figure S3.12**, **Figure S3.15**). Nevertheless, this illustrates that the *apo*-ACP-cAT exhibits some mild substrate flexibility. Additional enzyme engineering can address several issues to improve the Tv6-931 *apo*-ACP-cAT enzyme as oxyster transacylation biocatalysts by broadening the substrate scope, improving catalytic efficiency and tailoring the binding pocket to preferentially bind pantetheine instead of the *holo*-ACP.

### **Section 3.6 Coupling Tv6-931 HRPKS-cAT with a Downstream HRPKS for Polyketide Functionalization**

In addition to NCL-type coupling and cAT-catalyzed esterification, we tested coupling the transacylation of **1-Pante**, with another exergonic enzymatic reaction such as ketoreduction

for further polyketide diversification. By varying the upstream acyl-donor and the downstream enzyme, different combinations of polyketide-pantetheine probes may be produced. Starting with a polyketide-PE oxyester conjugate, the corresponding polyketide-pantetheine conjugate can be generated for further functionalization and to probe their suitability as substrates for different HRPKS catalytic domains.



**Scheme 3.2:** Coupling the ACP-cAT enzyme with another functionalization enzyme. Using **1-PE** (A) or **2-PE** (B), pantetheine, *apo*-ACP-cAT, NADPH, and LovB enzyme variants, a small variety polyketide-pantetheine derivatives can be enzymatically synthesized.

Since the LovB enzyme is a classical HRPKS example, it was a logical choice for this proof-of-concept model. In addition, the LovB enzyme also utilizes both the ketoreductase (KR) and dehydratase (DH) domains on an  $\alpha$ -methyl tetraketide intermediate. To first examine only the KR domain of the LovB HRPKS, we used a mutant variant of LovB with the dehydratase (DH) deactivated. This LovB-DH<sup>0</sup> mutant variant allows for inspection of the KR domain without possible interference from the DH and other downstream domains. When LovB-DH<sup>0</sup> was incubated with **1-Pante** generated *in situ* using the Tv6-931 *apo*-ACP-cAT, a significant amount of **1-Pante** was reduced to its  $\beta$ -hydroxyl pantetheine analog **3** (**Figure S3.16**). Notably,

the high reduction efficiency and results were comparable to the control reactions that used **1-Pante** directly. (**Figure S3.17**).

Furthermore, if the **1-Pante** (*in situ*) was incubated with the native LovB enzyme, the functional DH domain will also catalyze the dehydration of the resulting  $\beta$ -hydroxyl adduct **3** to the diene **4** in low levels (**Figure S3.16**). The extent of dehydration is most likely limited by the inherent substrate preference of the partnering enzyme (LovB). In addition, when the *apo*-ACP-cAT was incubated with **2-PE** instead, the analogous **2-Pante** adduct was also sufficiently produced *in situ* (**Figure S3.18**). This adduct can also be reduced by LovB-DH<sup>0</sup> or native LovB to generate the  $\beta$ -hydroxyl adduct **5**. Due to the *gem*-dimethyl moiety, compound **5** lacks an acidic proton at the  $\alpha$ -position and cannot be further dehydrated.

This proof of concept model shows that starting with the acyl-pantetheine enzymatic synthon like **1-PE** (acyl-donor), different pantetheine thioesters can be easily access: **1-Pante**, **3**, and **4**. Likewise, starting with **2-PE**, then **2-Pante** and **5** can be also sufficiently produced. Despite the kinetically slow and thermodynamically uphill oxyester-thioester transacylation, this platform shows significant potential to synthesize acyl-pantetheine compounds as chemical probes and functionalized polyketides (**Scheme 3.1-3.2**). Since this enzymatic system can offer similar insight to their acyl-SNAC analogs, using the cAT domain offers an attractive and complementary method to thioester chemical probes especially given the longer half-lives of oxyesters. Coupling the formation of acyl-pantetheine (*in situ*) with another HRPKS enzyme may enable more modularity and programming control for future engineering efforts.



## Section 3.7 Discussion

Although there has been significant research in the cofactor-independent, oxyester-thioester transacylation mechanism of cAT-type enzymes, the scope of the research has been mainly limited to the trafficking of fatty acids and primary metabolism.<sup>10-11, 14</sup> In this report, we apply our understanding of cAT-catalyzed transacylation towards a polyketide “recapture” in secondary metabolism. Using an assortment of chemical SNAC probes (**Table S3.1**) and a variety of Tv6-931 enzymes, we demonstrated that the second methylation ( $k_{1-MT}$ ) is the rate-limiting step in the biosynthesis of **2** (**Figure 3.1D**, **Figure 3.2**). According to the traditional HRPKS kinetic competition model, the main product should be **1** since  $k_{1-cAT}$  is approximately ten-folds faster than  $k_{1-MT}$  (**Figure 3.2C**). Despite the thermodynamic considerations (**Figure 3.1C**), we showed that the polyketide “recapture” mechanism (oxyester-thioester transacylation) catalyzed by the cAT-like domain is kinetically viable and necessary (**Figure 3.2-3.3**). This extra variable adds complexity to the biosynthesis to the Tv6-931 enzyme and shows that the enzymatic programming of HRPKS-cAT enzymes may need additional considerations.

From a mechanistic point of view, cAT domains and enzymes present an interesting method of synthesizing higher-energy thioester bonds. Traditionally, these acyl-pantetheine derivatives require the coupling of adenosyl triphosphate (ATP) or other activating agents to convert a carboxylic acid to the corresponding thioester.<sup>36</sup> It is intriguing to postulate why Nature would devise this elaborate mechanism to install a *gem*-dimethyl group and a greater purpose with HRPKS-cAT enzymes. Since it appears that fatty acid synthases (FAS) and cAT enzymes share vital roles in fatty acid metabolism, it is plausible that these HRPKS enzymes began utilizing enzymatic tools afforded to FAS, their distant relatives. In the Tv6-931 example, this

partnership of enzymes circumvents complications with decarboxylation and poor methylation rates.

Using the *holo*-ACP-cAT, enzymatic transesterification of polyketides **1** and **2** with other polyols proceed with high efficiency, and chemo-enzymatic amidation with cysteamine-derivatives proceed with moderate efficiency via *S-N* shift. In addition, using the *apo*-ACP-cAT, a variety of polyketide-pantetheine compounds could be enzymatically synthesized. Using one enzyme motif (two forms), three different chemical bonds (oxyesters, thioesters, amides) can be formed in an enzymatic or chemoenzymatic fashion.

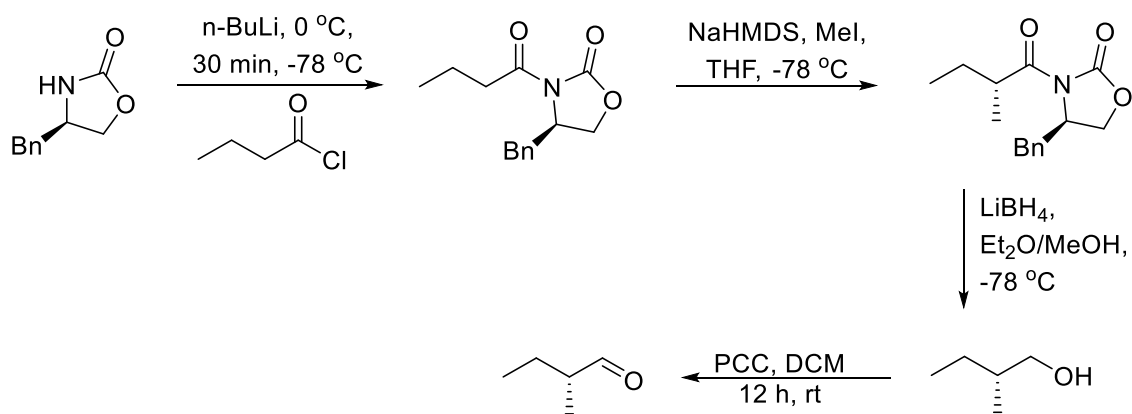
The facile enzymatic syntheses of acyl-pantetheine derivatives from the Tv6-931 enzyme provides unique access to privileged chemical probes.<sup>36</sup> The resulting pantetheine thioesters can then be subsequently modified by other downstream HRPKS enzymes. Mixing and matching different HRPKS enzymes through the acyl-pantetheine intermediates bypass the ACP-recognition issues that hampers other combinatorial biosynthetic approaches. This enzymatic, proof-of-concept provides a modular synthesis of an assortment of acyl-pantetheine compounds using HRPKS enzymes and expands the molecular toolbox available to biosynthetic chemist.

As an additional benefit, the half-lives of non-activated esters such as **1** and **2** are usually several orders of magnitude longer than thioesters such as acyl-SNAC and acyl-pantetheine compounds. We observed that the **1-SNAC** and **2-SNAC** are significantly more prone to degradation by hydrolysis, and thiol-Michael addition from residual SNAC. On the other hand, after more than a year of repeated use, we do not notice appreciable levels of degradation of **1** and **2**. Since the synthesis of PE derivatives have abundant precedence, new opportunities for

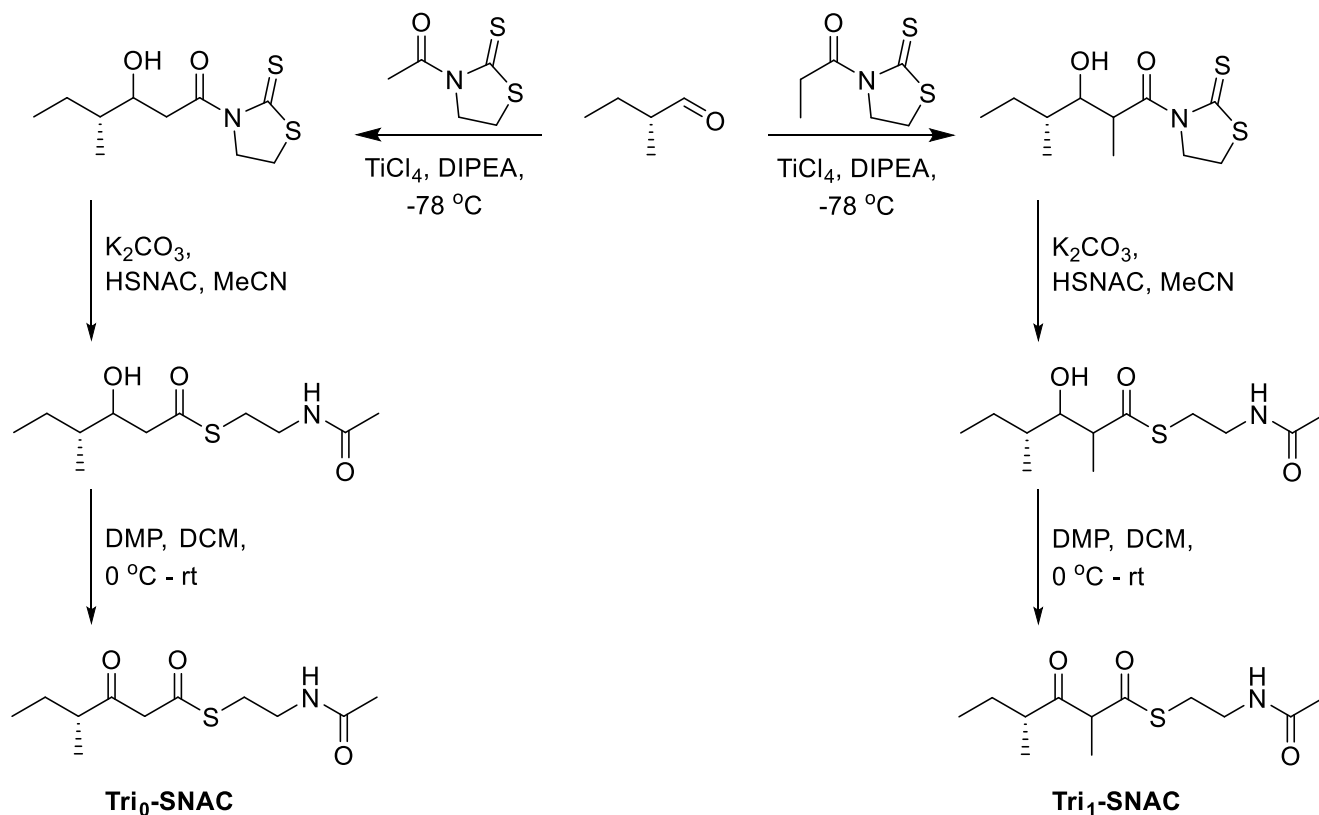
chemically stable polyketide probes and enzymatic synthons can be realized from this complementary platform after optimization and engineering efforts.<sup>37-42</sup>

Finally, this platform demonstrates that from exploiting only one HRPKS-cAT system, a variety of unnatural polyketides-conjugates have already been demonstrated. With approximately 100 other HRPKS-cAT enzymes currently unexplored, the substrate scope of the starting acyl-donors are bound to increase. The knowledge gained from this mechanistic study will provide insight into other members of the HRPKS-cAT family such as the putative biosynthetic enzymes for hypoxysordarin and AF-toxin. By mixing and matching different HRPKS enzymes, a versatile method to probe HRPKS functionalities can be envisioned. Preliminary data suggests that the substrate scope may be modest but tolerated (**Figure S3.15**). Coupling the polyketide substrates and products of upstream HRPKS-cAT enzymes with a library of other downstream HRPKS enzymes, exciting new possibilities, chemical diversity and molecular probes of polyketides are waiting to be (re)captured.

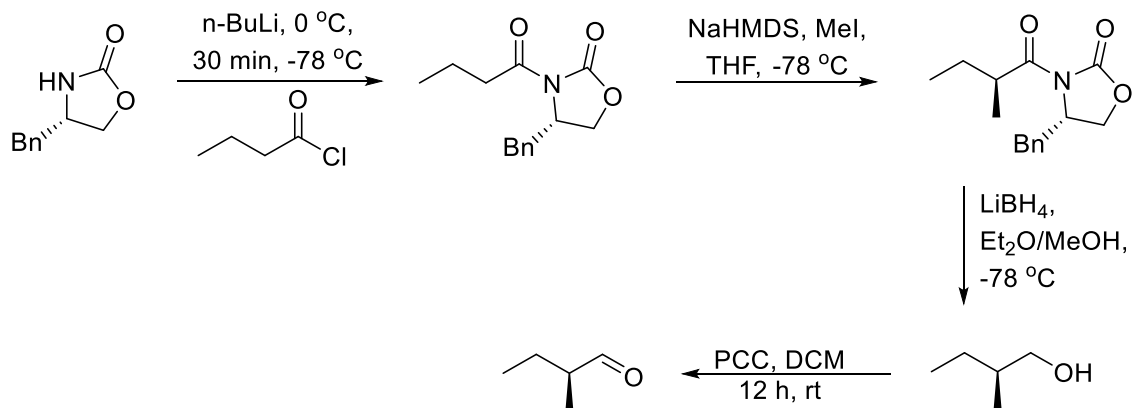
### Section 3.8 Supplementary Schemes



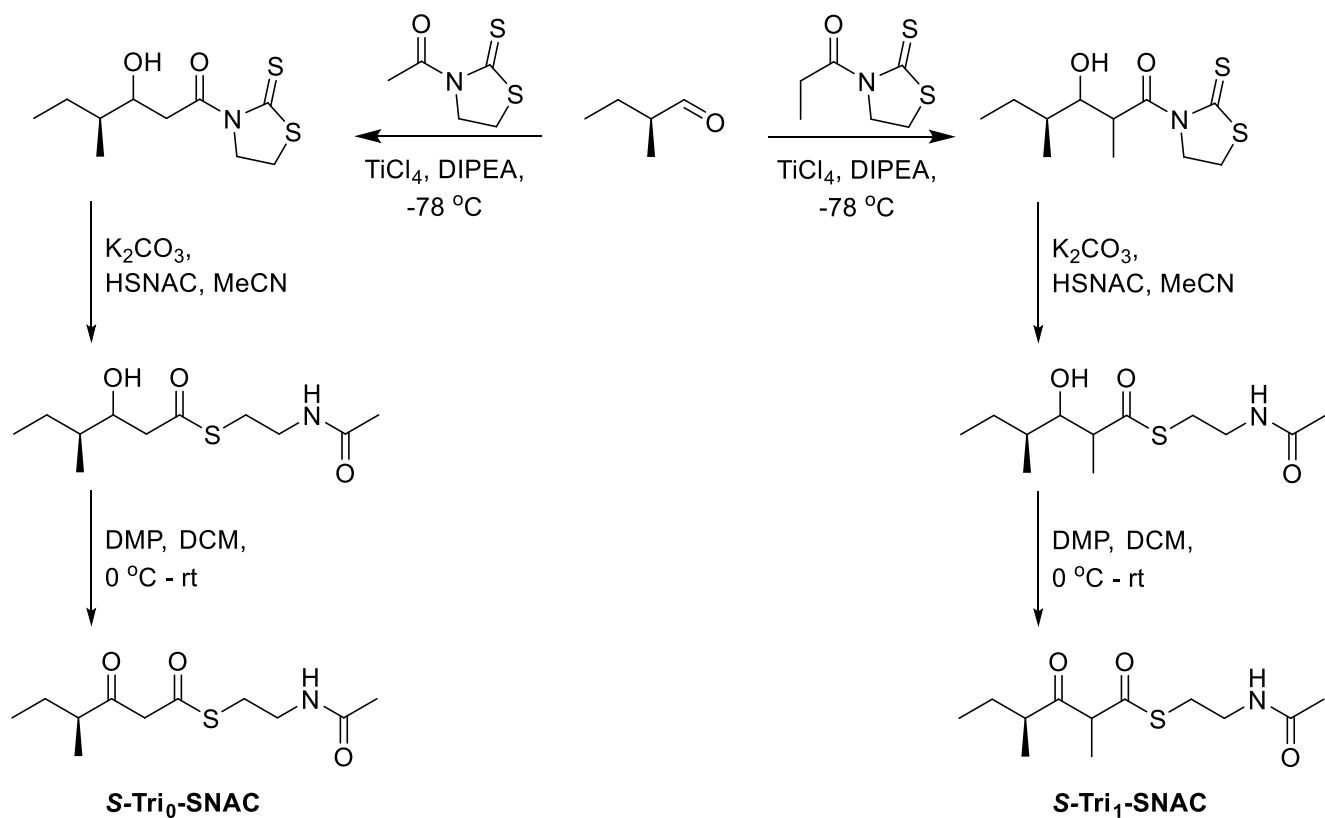
**Scheme S3.1:** Synthetic scheme for the preparation of (*R*)-2-methylbutanol



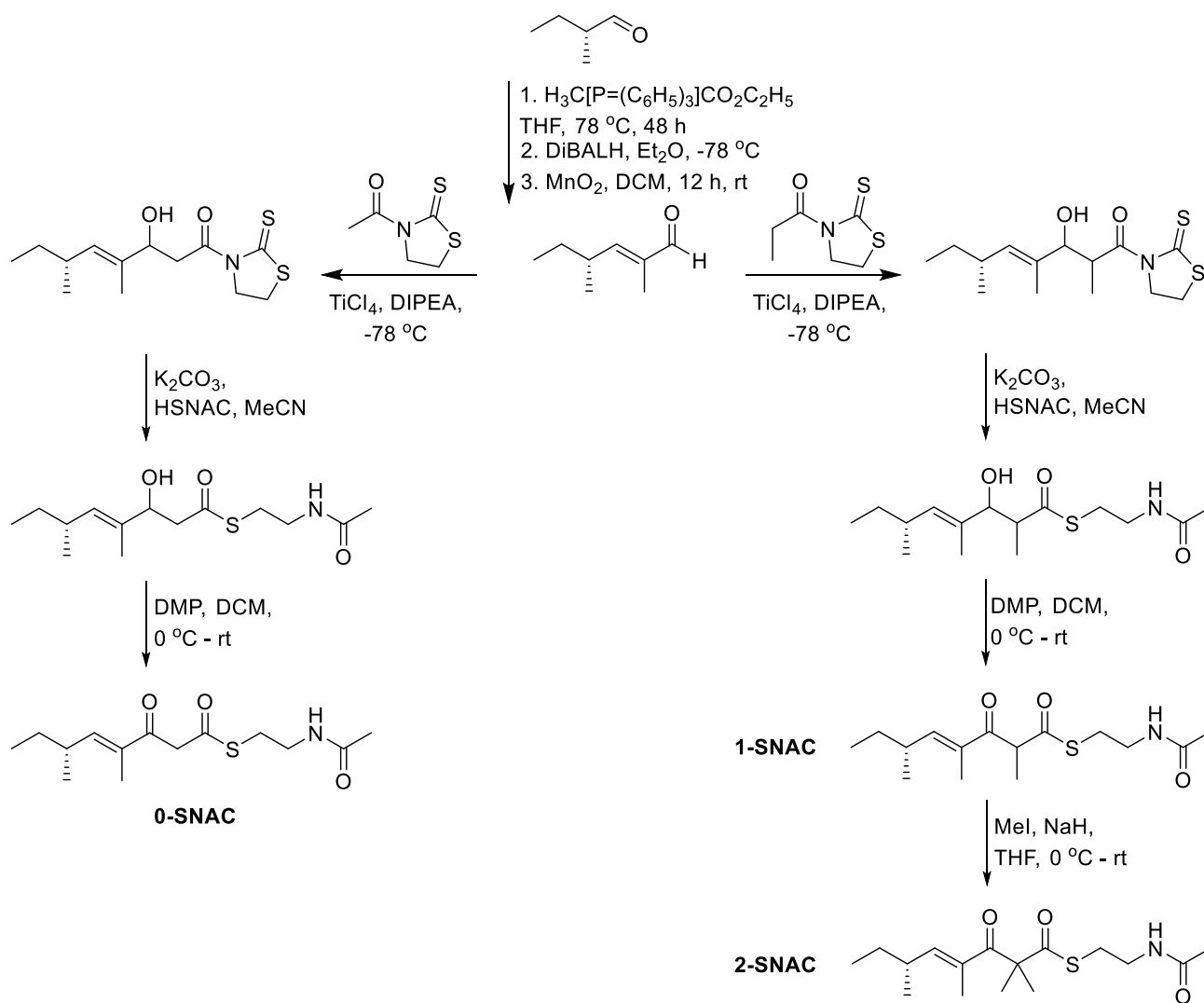
**Scheme S3.2:** Synthetic scheme for the preparation of **Tri<sub>0</sub>-SNAC** and **Tri<sub>1</sub>-SNAC**



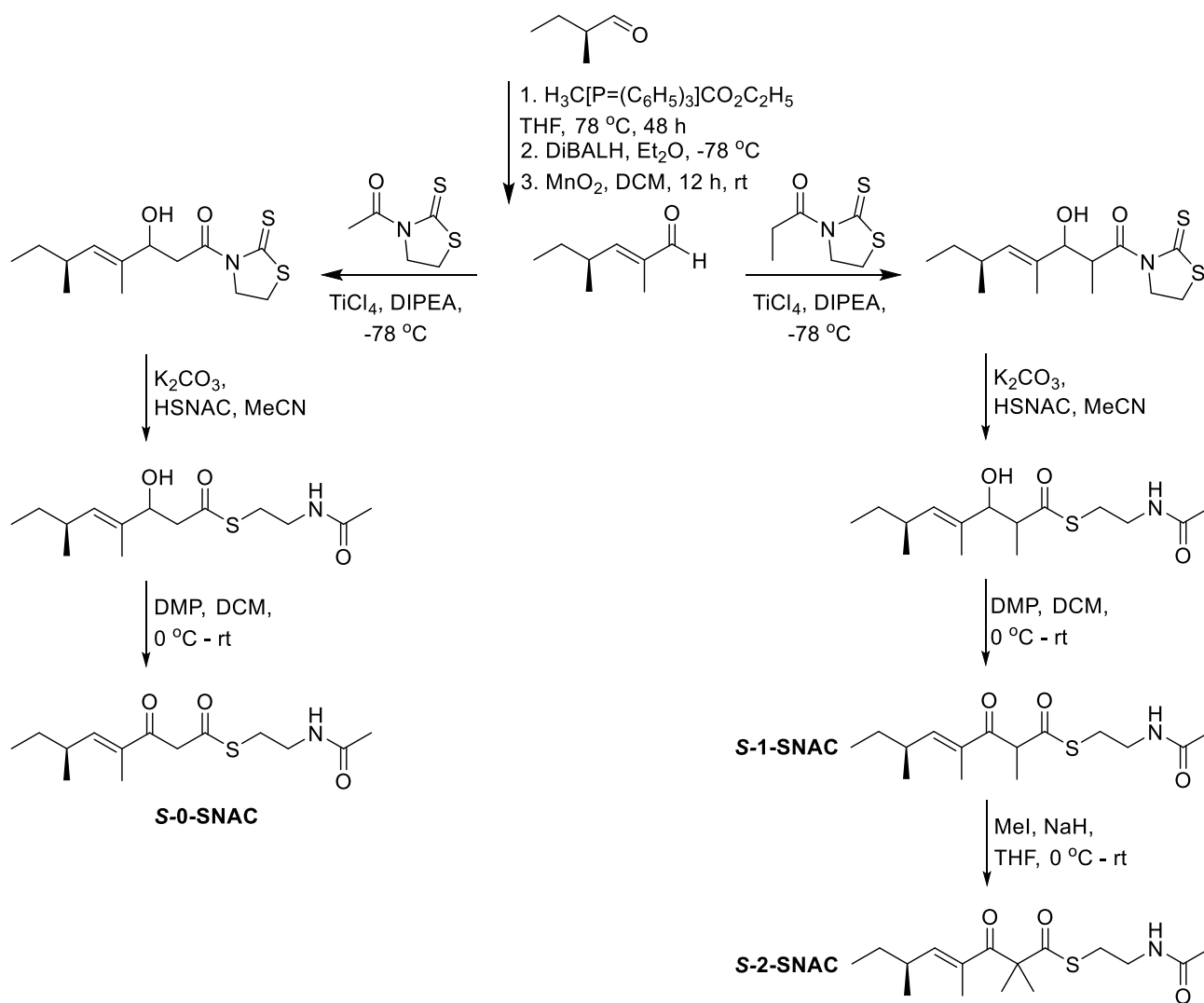
**Scheme S3.3:** Synthetic scheme for the preparation of (*R*)-2-methylbutanal



**Scheme S3.4:** Synthetic scheme for the preparation of *S*-Tri<sub>0</sub>-SNAC and *S*-Tri<sub>1</sub>-SNAC

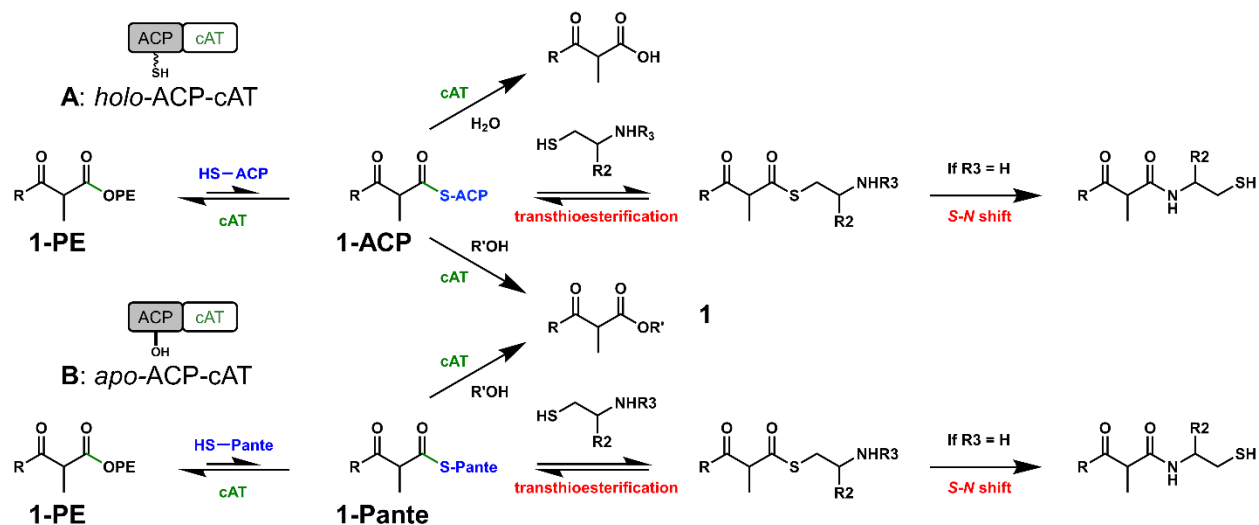


**Scheme S3.5:** Synthetic scheme for the preparation of **0-SNAC**, **1-SNAC** and **2-SNAC**



**Scheme S3.6:** Synthetic scheme for the preparation of *S*-0-SNAC, *S*-1-SNAC and *S*-2-SNAC

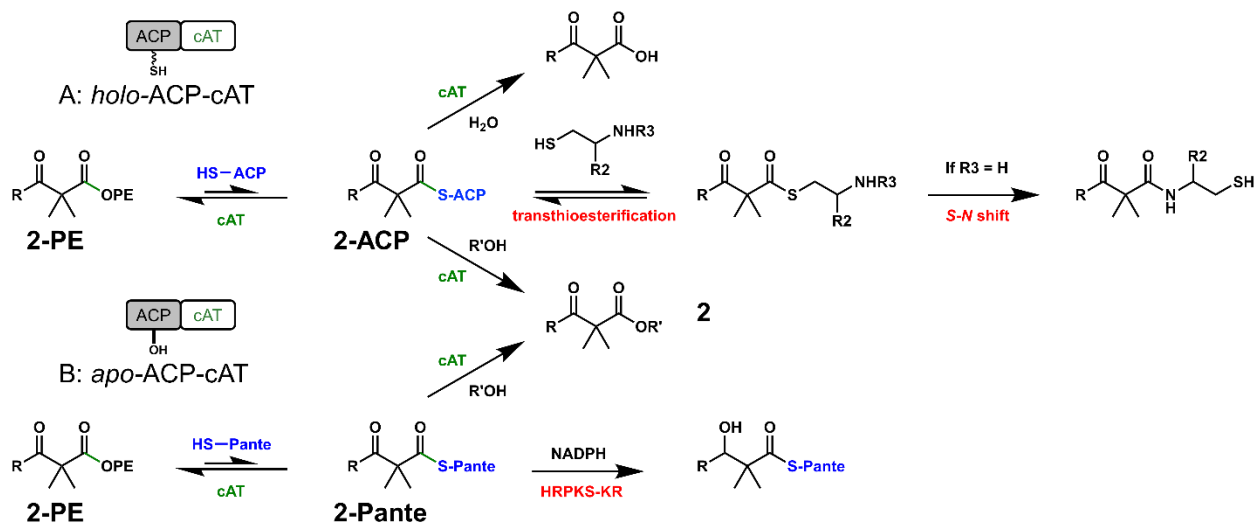
Mechanism of ACP-cAT Chemoenzymatic Transacylation



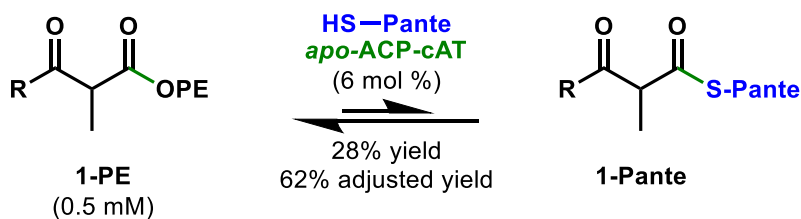
**Scheme S3.7:** Mechanism of ACP-cAT Chemoenzymatic Transacylation Using **1-PE**: A) Using the holo-ACP-cAT enzyme, the **1-ACP** intermediate can be chemically trapped by other thiols to furnish the corresponding transthioesterification adduct or enzymatically functionalized to produce the corresponding oxyster **1**. B) Using the apo-ACP-cAT enzyme, pantetheine can be used a substrate (ACP-surrogate) for direct transaction to produce **1-Pante** which can be chemically trapped by other thiols to furnish the corresponding transthioesterification adduct or enzymatically functionalized to produce the corresponding oxyster **1**.



Mechanism of ACP-cAT Chemoenzymatic Transacylation



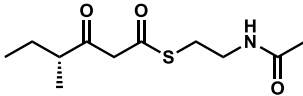
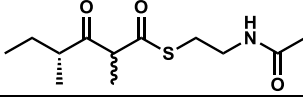
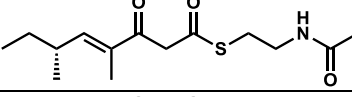
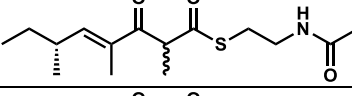
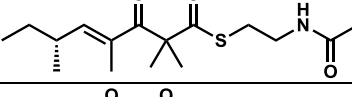
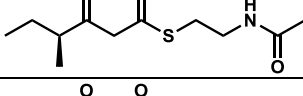
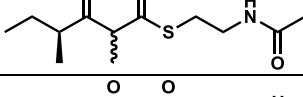
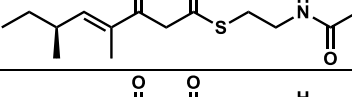
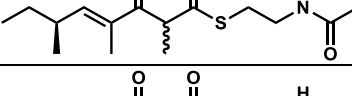
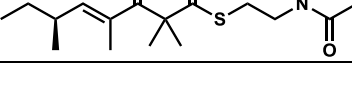
**Scheme S3.8:** Mechanism of ACP-cAT Chemoenzymatic Transacylation Using **2-PE**: A) Using the holo-ACP-cAT enzyme, the **2-ACP** intermediate can be chemically trapped by other thiols to furnish the corresponding transthioesterification adduct or enzymatically functionalized to produce the corresponding oxyster **2**. B) Using the apo-ACP-cAT enzyme, pantetheine can be used a substrate (ACP-surrogate) for direct transaction to produce **2-Pante** which can be enzymatically functionalized to **2** or further functionalized by other HRPKS enzymes.



**Scheme S3.9:** Enzymatic synthesis of **1-Pante** catalyzed by Tv6-931 apo-ACP-cAT enzyme: **1-PE** (0.5 mM), pantetheine (10 mM), apo-ACP-cAT (30  $\mu$ M), 2 mM DTT,  $\text{Na}_2\text{HPO}_4$  (pH 8.0) at 30°C overnight. The majority of the starting substrate, **1-PE**, can be recovered and reused in a subsequent reaction.

## Section 3.9 Supplementary Tables

**Table S3.1: List of SNAC Compounds**

Compound Structure	Compound Label
	<b>Tri<sub>0</sub>-SNAC</b>
	<b>Tri<sub>1</sub>-SNAC</b>
	<b>0-SNAC</b>
	<b>1-SNAC</b>
	<b>2-SNAC</b>
	<b>S-Tri<sub>0</sub>-SNAC</b>
	<b>S-Tri<sub>1</sub>-SNAC</b>
	<b>S-0-SNAC</b>
	<b>S-1-SNAC</b>
	<b>S-2-SNAC</b>

**Table S3.2: DNA Primers Used**

<b>Primer Name</b>	<b>Sequence of Primer (5' to 3')</b>
pXW55-N-Term-Ura-F	AGCAGCACGTTCCCTTATATG
pXW55-N-8xHis-R	ACTAGTGTGATGGTGTGATGGTGTGATGGCTAGCCATATGGT ATTACG
pXW55-C-Term-F	TAATTTAAATGACAAATTTGTGCG
pXW55-C-Term-Ura-R	ATGTCGAAAGCTACATATAAG
Tv6-931-N-His- HRPKS-P1-F	AATACCATATGGCTAGCCATCACCATCACCATCACCATCACACTA GTATGTCTCCACAAG
Tv6-931-HRPKS-P1-R	AGTCAACATATTGGAAAGATATG
Tv6-931-HRPKS-P2-F	CGAAGTAATCATATCTTTCCAA
Tv6-931-HRPKS-P2-R	AAATCACCGAAAATCCTTTTCA
Tv6-931-HRPKS-P3-F	GAGATTGTTGAAAAGGATTTTCG
Tv6-931-HRPKS-cAT- Native-R	TGAAGGCATCGGTCCGCACAAATTTGTCATTTAAATTAACAACA TTGATCTCAATGATA
Tv6-931-HRPKS- standalone-R	CGCACAAATTTGTCATTTAAATTAGTGATGGTGTGATGGTGTGATGCAC GGTAACGAAGGCAGC
Tv6-931-ACP_cAT- XW55-F	TGAAGGCATCGGTCCGCACAAATTTGTCATTTAAATTAACAACA TTGATCTCAATGATA

**Table S3.3: DNA Plasmids Used**

<b>Plasmid Name</b>	<b>Plasmid Backbone</b>	<b>Description of Plasmid</b>
pLFH18	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -8xHis-Tv6-931 $\Delta$ cAT gene under the <i>ADH2</i> promoter
pLFH19	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -8xHis-Tv6-931 gene under the <i>ADH2</i> promoter
pLFH29	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -8xHis-Tv6-931 ACP-cAT gene under the <i>ADH2</i> promoter
YEplovB-6His-DH <sup>0</sup>	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -FLAG-LovB-DH <sup>0</sup> -C-6xHis gene under the <i>ADH2</i> promoter
YEplovB-6His	pXW55	2 $\mu$ <i>S. cer.</i> expression plasmid with <i>N</i> -FLAG-LovB-C-6xHis gene under the <i>ADH2</i> promoter

**Table S3.4: Expression Strains Used**

<b>Strain Name</b>	<b>Organism</b>	<b>Description of Expression Strain</b>
<i>S.cer</i> -pLFH18- <i>apo</i>	<i>S. cer.</i>	<i>S. cer.</i> BJ5464- <i>apo</i> strain expressing the pLFH19 plasmid
<i>S.cer</i> -pLFH19- <i>apo</i>	<i>S. cer.</i>	<i>S. cer.</i> BJ5464- <i>apo</i> strain expressing the pLFH19 plasmid
<i>S.cer</i> -pLFH29- <i>apo</i>	<i>S. cer.</i>	<i>S. cer.</i> BJ5464- <i>apo</i> strain expressing the pLFH29 plasmid
<i>S.cer</i> -pLFH19- <i>holo</i>	<i>S. cer.</i>	<i>S. cer.</i> BJ5464-npgA strain expressing the pLFH19 plasmid
<i>S.cer</i> -pLFH29- <i>holo</i>	<i>S. cer.</i>	<i>S. cer.</i> BJ5464-npgA strain expressing the pLFH29 plasmid
<i>S.cer</i> -LovB-DH <sup>0</sup>	<i>S. cer.</i>	<i>S. cer.</i> BJ5464-npgA strain expressing the YEplovB-6His-DH <sup>0</sup> plasmid
<i>S.cer</i> -LovB	<i>S. cer.</i>	<i>S. cer.</i> BJ5464-npgA strain expressing the YEplovB-6His plasmid

**Table S3.5: Transacylation of 2-PE by various nucleophiles**

**A**

**B**

Entry	Nucleophile	Transacylation (%)	Hydrolysis (%)
1a		4	47
1b	Pantetheine	2	46
2a		Mixture of Side Products	25
2b		Mixture of Side Products	20
3a		67	9
3b		94	2

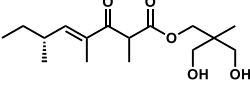
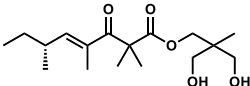
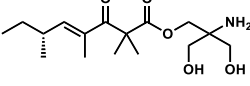
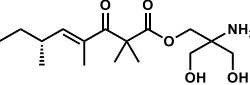
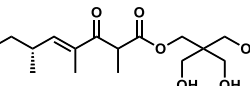
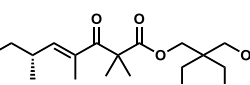
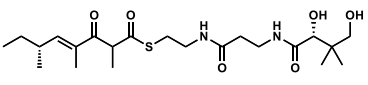
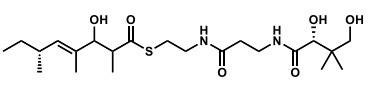
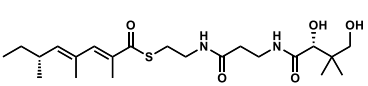
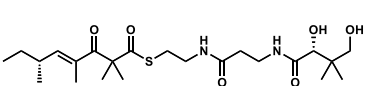
Entry	Nucleophile	Transacylation (%)	Hydrolysis (%)
1a		1	N.D.
1b	Pantetheine	7	N.D.
2-3	See Table S3.5A	N.D.	N.D.

Transacylation of **2-PE** by various nucleophiles using **A**) *holo*-ACP-cAT (forming **1-ACP in situ**) and **B**) *apo*-ACP-cAT. Entry 1: Thiols not capable of *S-N* shift. Entry 2: Thiols capable of *S-N* shift. Entry 3: Polyol nucleophile substrates accepted by the cAT domain.

Note 1: entry 1a (SNAC) and 1b (Pantetheine) are poor and improved chemical mimics of *holo*-ACP, respectively. N.D. = Not Detected.

Note 2: unlike the transacylation of **1-PE** (Table 3.1), the *gem*-dimethyl **2-PE** acyl-donor and its adducts are significantly more prone to Michael addition, forming side products (Figure S3.11).

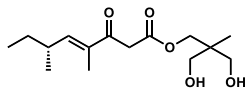
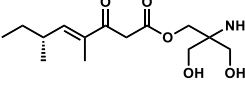
**Table S3.6: HPLC-MS Mass Data of Compounds**

Molecule	Compound Structure	Chemical Formula	Adducts Found	Expected Mass (m/z)	Mass Found (m/z)
<b>1-THME</b>		C <sub>16</sub> H <sub>28</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	301	301
			[M+Na] <sup>+</sup>	323	323
			[M+H-H <sub>2</sub> O] <sup>+</sup>	283	283
			[M-H] <sup>-</sup>	299	299
<b>2-THME</b>		C <sub>17</sub> H <sub>30</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	315	315
			[M+Na] <sup>+</sup>	337	337
			[M+H-H <sub>2</sub> O] <sup>+</sup>	297	297
<b>1-Tris</b>		C <sub>15</sub> H <sub>27</sub> NO <sub>5</sub>	[M+H] <sup>+</sup>	302	316
			[M+Na] <sup>+</sup>	324	338
			[M+H-H <sub>2</sub> O] <sup>+</sup>	284	284
<b>2-Tris</b>		C <sub>16</sub> H <sub>29</sub> NO <sub>5</sub>	[M+H] <sup>+</sup>	316	316
			[M+Na] <sup>+</sup>	338	338
			[M+H-H <sub>2</sub> O] <sup>+</sup>	298	298
<b>1-PE</b>		C <sub>16</sub> H <sub>28</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	317	317
			[M+Na] <sup>+</sup>	339	339
			[M+H-H <sub>2</sub> O] <sup>+</sup>	299	299
			[M-H] <sup>-</sup>	315	315
<b>2-PE</b>		C <sub>17</sub> H <sub>30</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	331	331
			[M+Na] <sup>+</sup>	353	353
			[M+H-H <sub>2</sub> O] <sup>+</sup>	313	313
<b>1-Pante</b>		C <sub>22</sub> H <sub>38</sub> N <sub>2</sub> O <sub>6</sub> S	[M+H] <sup>+</sup>	459	459
			[M+Na] <sup>+</sup>	481	481
			[M+H-H <sub>2</sub> O] <sup>+</sup>	441	441
			[M-H] <sup>-</sup>	457	457
<b>3</b>		C <sub>22</sub> H <sub>40</sub> N <sub>2</sub> O <sub>6</sub> S	[M+H] <sup>+</sup>	461	461
			[M+Na] <sup>+</sup>	483	483
			[M+H-H <sub>2</sub> O] <sup>+</sup>	443	443
			[M-H] <sup>-</sup>	459	459
<b>4</b>		C <sub>22</sub> H <sub>38</sub> N <sub>2</sub> O <sub>5</sub> S	[M+H] <sup>+</sup>	443	443
			[M+Na] <sup>+</sup>	465	465
			[M-H] <sup>-</sup>	441	441
			[M+Cl] <sup>-</sup>	477	477
<b>2-Pante</b>		C <sub>23</sub> H <sub>40</sub> N <sub>2</sub> O <sub>6</sub> S	[M+H] <sup>+</sup>	473	473
			[M+Na] <sup>+</sup>	495	495
			[M+H-H <sub>2</sub> O] <sup>+</sup>	455	455
			[M-H] <sup>-</sup>	471	471

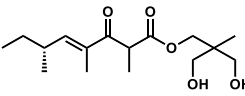
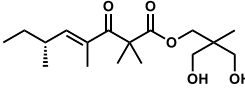
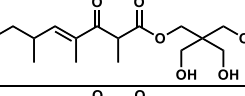
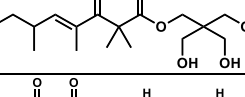
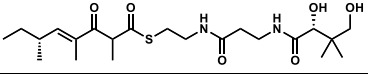
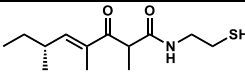
Molecule	Compound Structure	Chemical Formula	Adducts Found	Expected Mass (m/z)	Mass Found (m/z)
<b>5</b>		C <sub>23</sub> H <sub>42</sub> N <sub>2</sub> O <sub>6</sub> S	[M+H] <sup>+</sup>	475	475
			[M+Na] <sup>+</sup>	497	497
			[M+H-H <sub>2</sub> O] <sup>+</sup>	457	457
<b>1-COOH</b>		C <sub>11</sub> H <sub>18</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	199	199
			[M+H-CO <sub>2</sub> ] <sup>+</sup>	155	155
<b>1-SNAC</b>		C <sub>15</sub> H <sub>25</sub> NO <sub>3</sub> S	[M+H] <sup>+</sup>	300	300
			[M+Na] <sup>+</sup>	322	322
			[M-H] <sup>-</sup>	298	298
<b>2-COOH</b>		C <sub>12</sub> H <sub>20</sub> O <sub>3</sub>	[M+H] <sup>+</sup>	213	213
			[M+H-CO <sub>2</sub> ] <sup>+</sup>	169	169
<b>2-SNAC</b>		C <sub>16</sub> H <sub>27</sub> NO <sub>3</sub> S	[M+H] <sup>+</sup>	314	314
			[M+Na] <sup>+</sup>	336	336
<b>6</b>		C <sub>16</sub> H <sub>25</sub> NO <sub>5</sub> S	[M+H] <sup>+</sup>	344	344
			[M+Na] <sup>+</sup>	366	366
			[M-H] <sup>-</sup>	342	342
			[M+Cl] <sup>-</sup>	378	378
<b>7</b>		C <sub>13</sub> H <sub>23</sub> NO <sub>2</sub> S	[M+H] <sup>+</sup>	258	258
			[M+Na] <sup>+</sup>	280	280
			[M-H] <sup>-</sup>	256	256
			[M+Cl] <sup>-</sup>	292	292
<b>7'</b>		C <sub>15</sub> H <sub>28</sub> N <sub>2</sub> O <sub>2</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	333	333
			[M+Na] <sup>+</sup>	355	355
			[M-H] <sup>-</sup>	331	331
<b>7b</b>		C <sub>26</sub> H <sub>44</sub> N <sub>2</sub> O <sub>4</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	513	513
			[M+Na] <sup>+</sup>	535	535
			[M-H] <sup>-</sup>	511	511
<b>7c</b>		C <sub>25</sub> H <sub>45</sub> N <sub>3</sub> O <sub>6</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	534	534
			[M+Na] <sup>+</sup>	556	556
			[M-H] <sup>-</sup>	532	532
<b>8</b>		C <sub>14</sub> H <sub>23</sub> NO <sub>4</sub> S	[M+H] <sup>+</sup>	302	302
			[M+Na] <sup>+</sup>	324	324
			[M-H] <sup>-</sup>	300	300
<b>8'</b>		C <sub>17</sub> H <sub>28</sub> N <sub>2</sub> O <sub>6</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	421	42
			[M-H] <sup>-</sup>	419	419
<b>9</b>		C <sub>15</sub> H <sub>25</sub> NO <sub>4</sub> S	[M+H] <sup>+</sup>	316	316
			[M+Na] <sup>+</sup>	337	337
			[M-H] <sup>-</sup>	314	314
			[M+Cl] <sup>-</sup>	350	350

Molecule	Compound Structure	Chemical Formula	Adducts Found	Expected Mass (m/z)	Mass Found (m/z)
<b>9'</b>		C <sub>19</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	449	449
			[M-H] <sup>-</sup>	447	447
<b>10</b>		C <sub>19</sub> H <sub>32</sub> N <sub>2</sub> O <sub>6</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	449	449
			[M+Na] <sup>+</sup>	471	471
			[M-H] <sup>-</sup>	447	447
			[M+Cl] <sup>-</sup>	483	483
<b>11</b>		C <sub>16</sub> H <sub>30</sub> N <sub>2</sub> O <sub>2</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	347	347
			[M+Na] <sup>+</sup>	369	369
			[M+Cl] <sup>-</sup>	381	381
<b>11b</b>		C <sub>19</sub> H <sub>37</sub> NO <sub>6</sub> S	[M+H] <sup>+</sup>	408	408
			[M+Na] <sup>+</sup>	430	430
			[M+Cl] <sup>-</sup>	442	442
<b>11c</b>		C <sub>13</sub> H <sub>27</sub> NOS	[M+H] <sup>+</sup>	246	246
<b>12</b>		C <sub>20</sub> H <sub>34</sub> N <sub>2</sub> O <sub>6</sub> S <sub>2</sub>	[M+H] <sup>+</sup>	463	463
			[M+Na] <sup>+</sup>	485	485
			[M-H] <sup>-</sup>	461	461
			[M+Cl] <sup>-</sup>	497	497
<b>12b</b>		C <sub>21</sub> H <sub>39</sub> NO <sub>8</sub> S	[M+H] <sup>+</sup>	466	466
			[M+Na] <sup>+</sup>	488	488
<b>12c</b>		C <sub>15</sub> H <sub>29</sub> NO <sub>3</sub> S	[M+H] <sup>+</sup>	304	304
			[M+Na] <sup>+</sup>	326	326
<b>13</b>		C <sub>16</sub> H <sub>29</sub> NO <sub>5</sub>	[M+H] <sup>+</sup>	316	316
			[M+Na] <sup>+</sup>	338	338
			[M-H] <sup>-</sup>	314	314
<b>0-PE</b>		C <sub>15</sub> H <sub>26</sub> O <sub>6</sub>	[M+H] <sup>+</sup>	303	303
			[M+Na] <sup>+</sup>	325	325
			[M+H-H <sub>2</sub> O] <sup>+</sup>	285	285
			[M-H] <sup>-</sup>	301	301
<b>0-SNAC</b>		C <sub>14</sub> H <sub>23</sub> NO <sub>3</sub> S	[M+H] <sup>+</sup>	286	286
			[M+Na] <sup>+</sup>	308	308
			[M-H] <sup>-</sup>	284	284
<b>0-Pante</b>		C <sub>21</sub> H <sub>36</sub> N <sub>2</sub> O <sub>6</sub> S	[M+H] <sup>+</sup>	445	445
			[M+Na] <sup>+</sup>	467	467
			[M-H] <sup>-</sup>	443	443

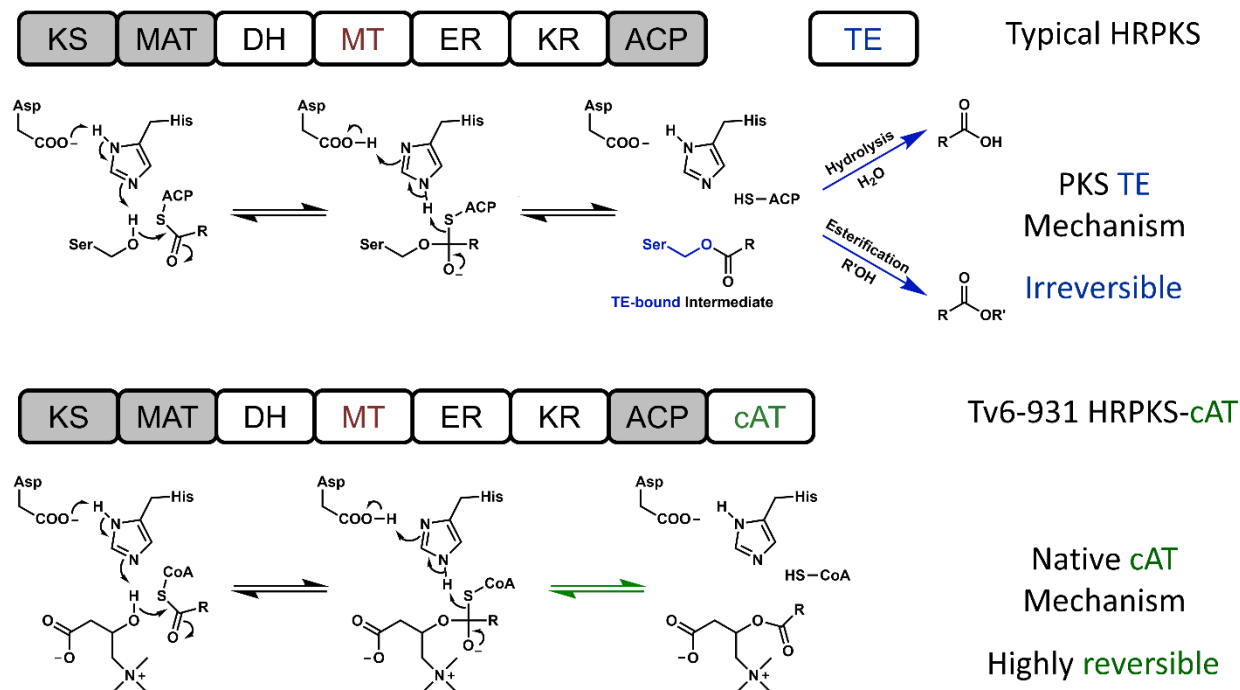


Molecule	Compound Structure	Chemical Formula	Adducts Found	Expected Mass (m/z)	Mass Found (m/z)
<b>0-THME</b>		C <sub>15</sub> H <sub>26</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	287	288
			[M+Na] <sup>+</sup>	309	309
			[M+H-H <sub>2</sub> O] <sup>+</sup>	269	269
			[M-H] <sup>-</sup>	285	285
<b>0-Tris</b>		C <sub>14</sub> H <sub>25</sub> NO <sub>5</sub>	[M+H] <sup>+</sup>	288	288
			[M+Na] <sup>+</sup>	310	310
			[M+H-H <sub>2</sub> O] <sup>+</sup>	270	270
			[M-H] <sup>-</sup>	286	286

**Table S3.7: HR-MS Mass Data of Compounds**

Molecule	Compound Structure	Chemical Formula	Adducts Found	Expected Mass (m/z)	Mass Found (m/z)
<b>1-THME</b>		C <sub>16</sub> H <sub>28</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	301.2015	301.1995
			[M+Na] <sup>+</sup>	318.2281	318.2259
			[M+H-H <sub>2</sub> O] <sup>+</sup>	283.1909	283.1888
			[M-H] <sup>-</sup>	299.1859	299.1852
<b>2-THME</b>		C <sub>17</sub> H <sub>30</sub> O <sub>5</sub>	[M+H] <sup>+</sup>	315.2172	315.2149
			[M+Na] <sup>+</sup>	332.2437	332.2415
			[M+H-H <sub>2</sub> O] <sup>+</sup>	297.2066	297.2043
			[M-H] <sup>-</sup>	313.2015	313.2010
<b>1-PE</b>		C <sub>16</sub> H <sub>28</sub> O <sub>6</sub>	[M+Na] <sup>+</sup>	339.1783	339.1768
			[M+H-H <sub>2</sub> O] <sup>+</sup>	299.1859	299.1837
<b>2-PE</b>		C <sub>17</sub> H <sub>30</sub> O <sub>6</sub>	[M+Na] <sup>+</sup>	353.1940	353.1928
			[M+H-H <sub>2</sub> O] <sup>+</sup>	313.2015	313.1985
<b>1-Pante</b>		C <sub>22</sub> H <sub>38</sub> N <sub>2</sub> O <sub>6</sub> S	[M+Na] <sup>+</sup>	481.2348	481.2336
<b>7</b>		C <sub>13</sub> H <sub>23</sub> N O <sub>2</sub> S	[M+Na] <sup>+</sup>	280.1347	280.1322

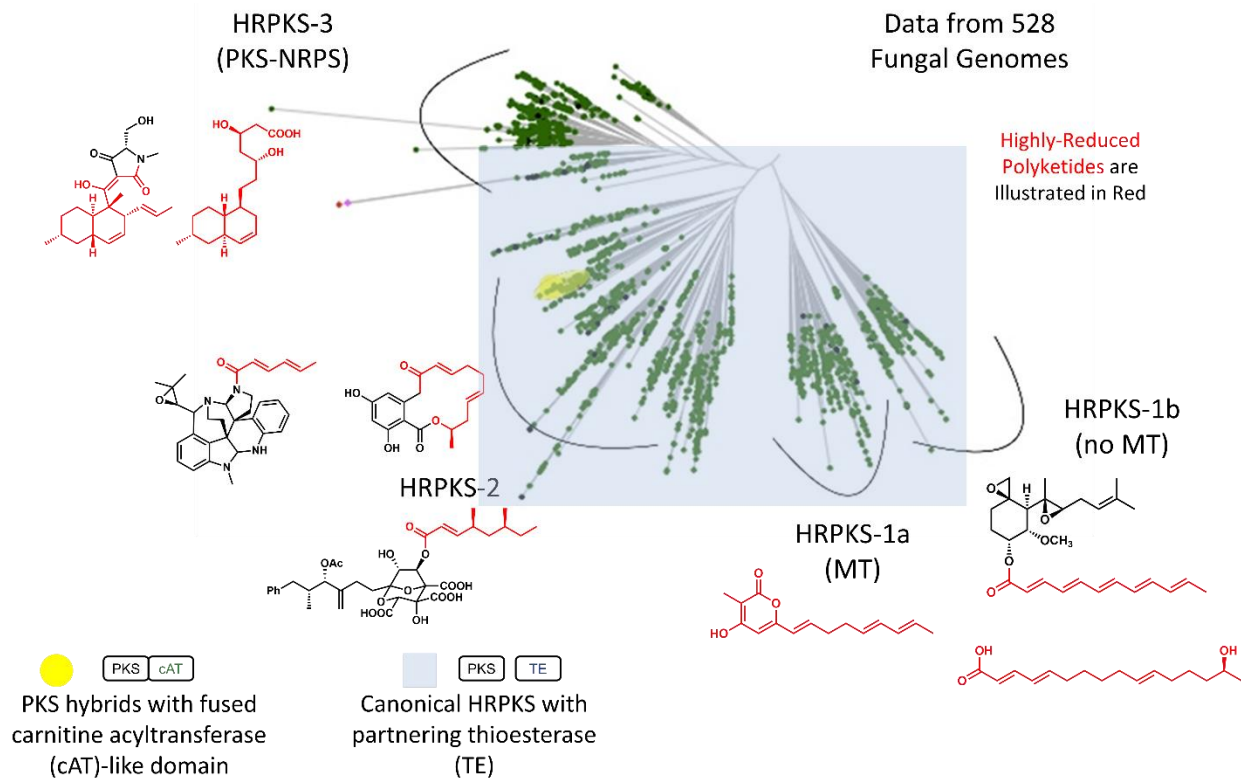
## Section 3.10 Supplementary Figures



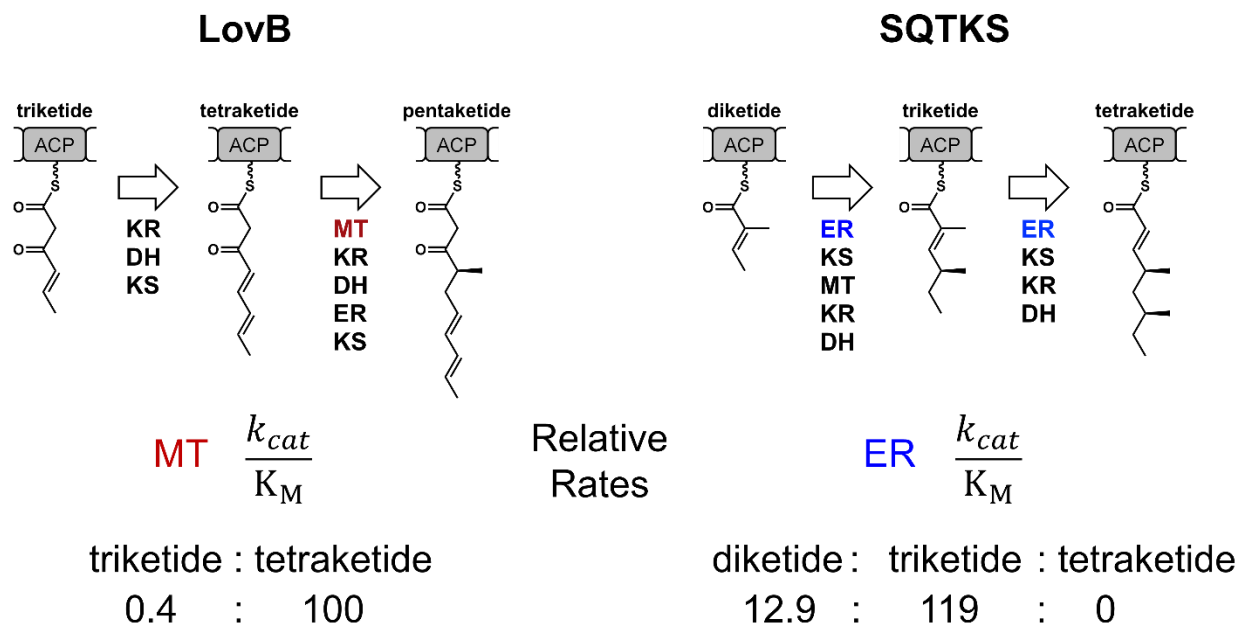
**Figure S3.1: Mechanistic Comparison of the PKS TE Domain and Native cAT Enzyme:**

Top: Mechanism of the PKS TE domain where esterification proceeds through a covalent bond with a serine of the catalytic triad. The TE-catalyzed esterification is essentially irreversible.

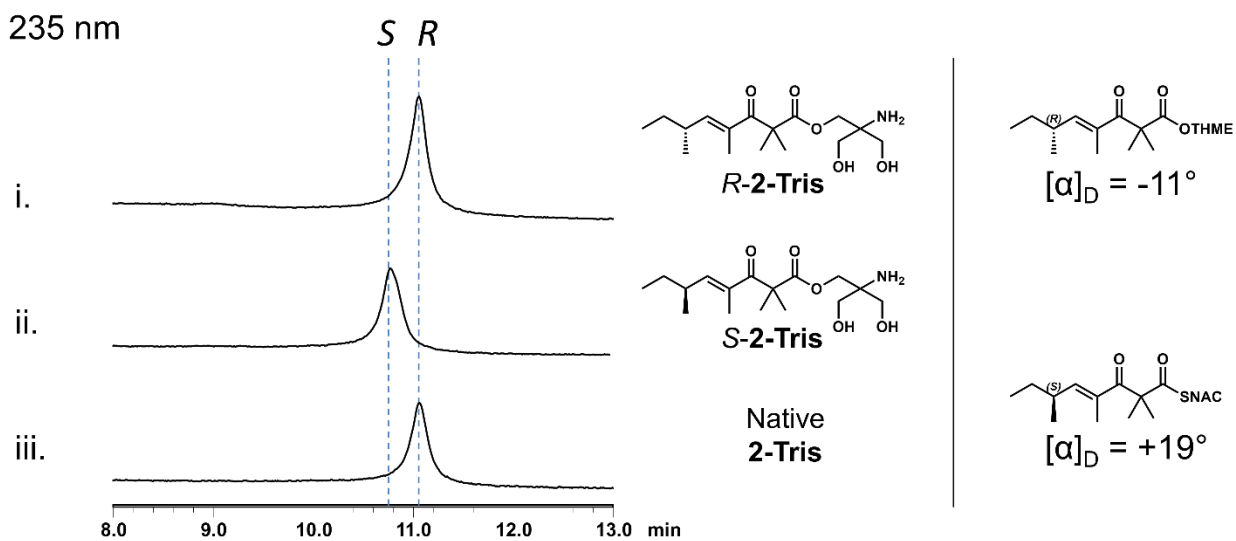
Bottom: Mechanism of the native cAT domain where esterification proceeds non-covalently. The cAT catalyzed esterification is highly reversible



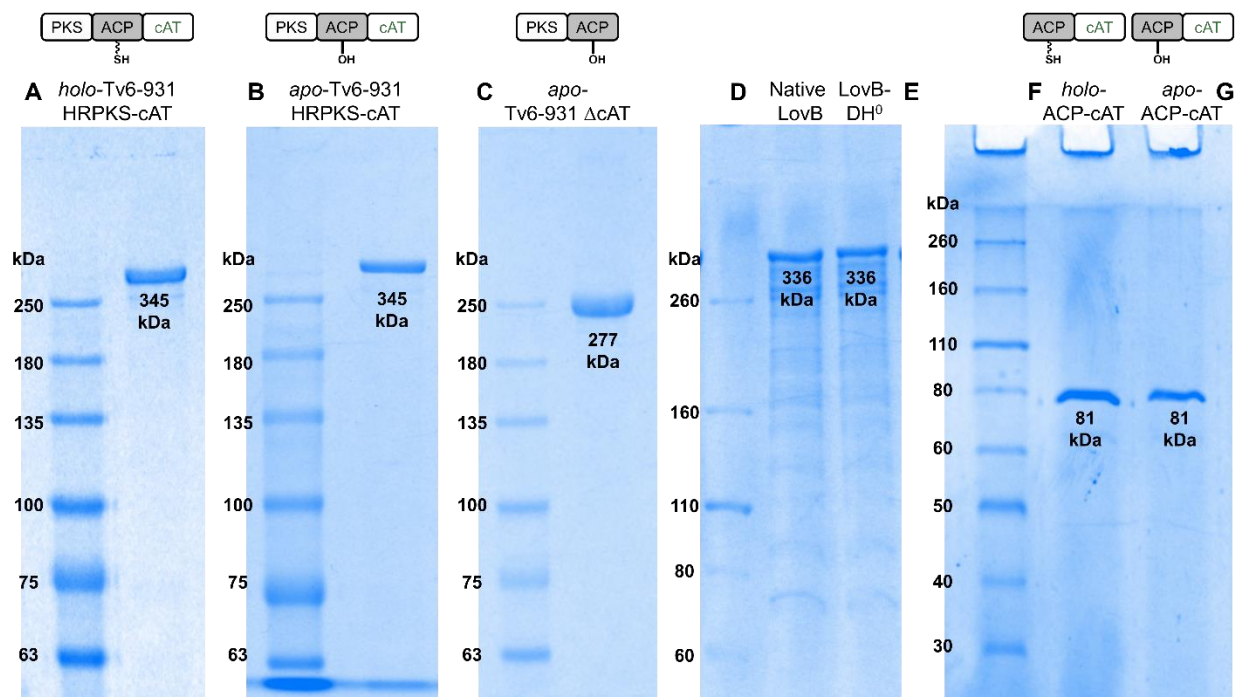
**Figure S3.2: Fungal KS Phylogenetic Tree with Highlighted HRPKS-cAT Fusion Enzymes:** Phylogenetic Tree of Fungal HRPKS enzymes discovered from DNA sequencing. The region shaded in blue depicts HRPKS enzymes that utilize partnering TE domains. The subclade highlighted in the yellow represents HRPKS-cAT fusion enzymes.



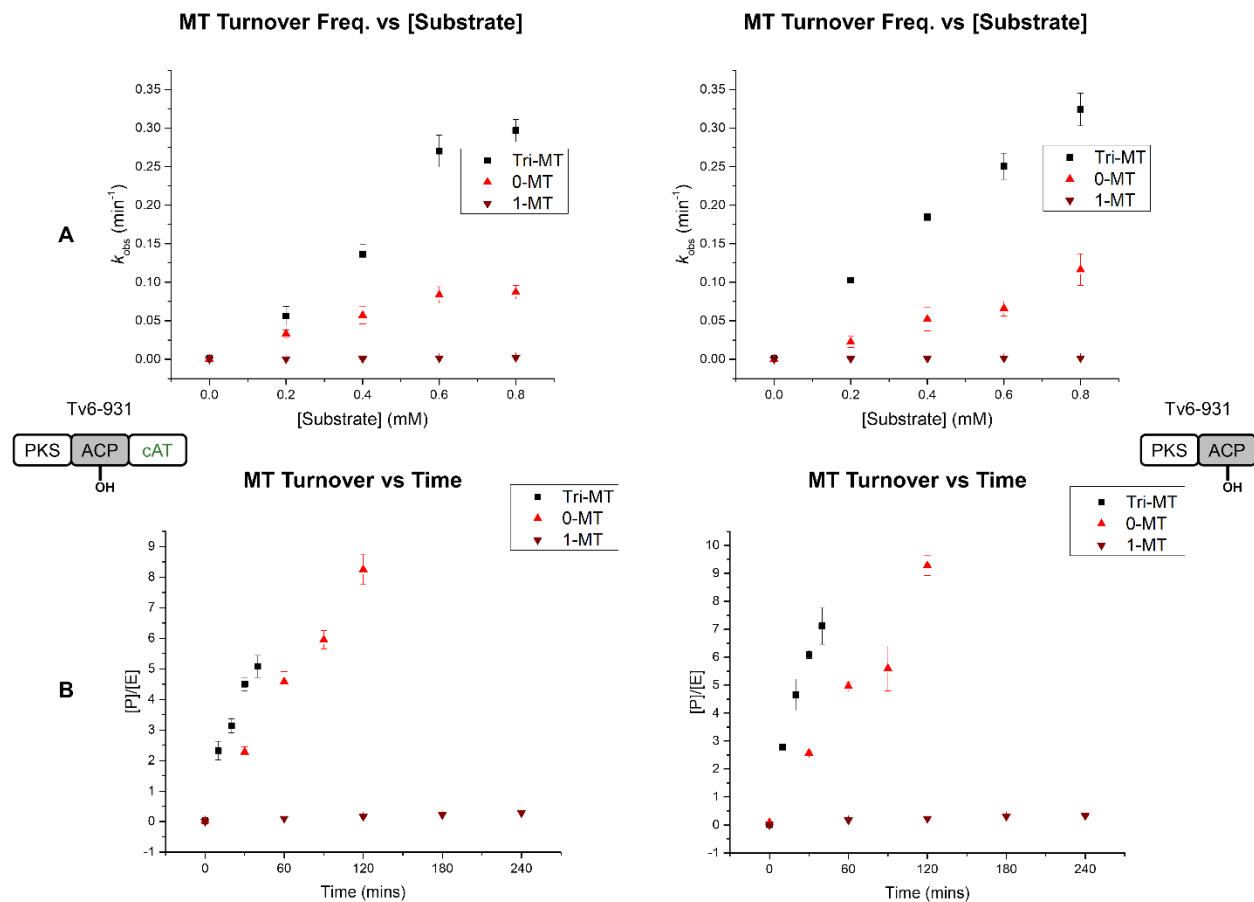
**Figure S3.3: Kinetic Competition Model of HRPKS:** Examples of the PKS Kinetic Competition Model of catalytic domains to explain key steps in the biosynthesis of dihydromonacolin L (LovB) and squalestatin tetraketide-arm (SQTKS) respectively.



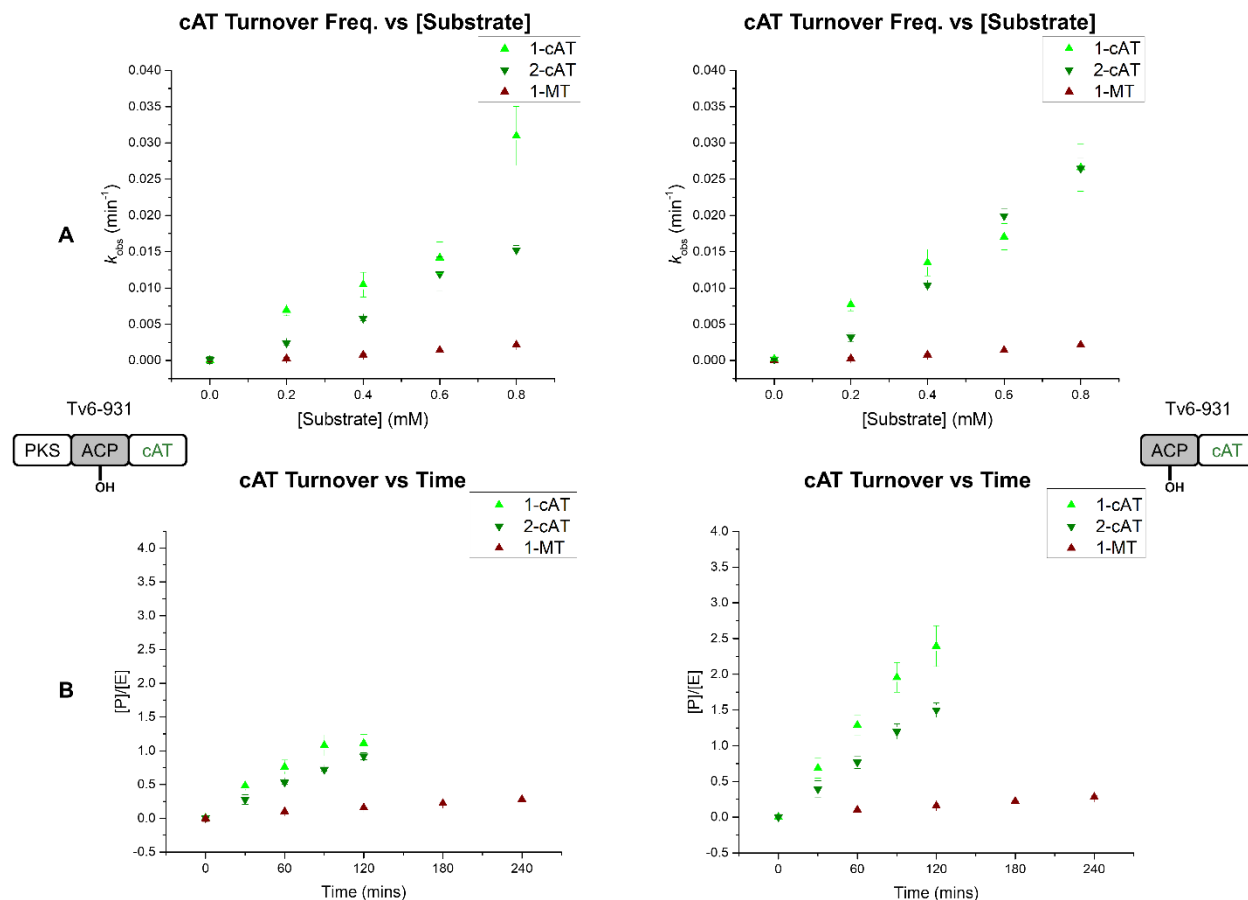
**Figure S3.4: Determination of the Stereochemistry of the Distal,  $\epsilon$ -Methyl Group:** The stereochemistry of the distal,  $\epsilon$ -methyl group of compound **2** was assigned by the separation on a chiral HPLC column, and by measuring the optical rotation.



**Figure S3.5: Heterologous Expression and Purification of Tv6-931 and LovB Enzyme Variants:** Illustration of SDS-PAGE gel images from the heterologous expression and purification of Tv6-931 and LovB enzyme variants. The samples are labeled as followed: **A)** *holo*-Tv6-931, **B)** *apo*-Tv6-931, **C)** *apo*-Tv6-931  $\Delta$ cAT, **D)** native LovB, **E)** LovB-DH<sup>0</sup>, **F)** *holo*-Tv6-931 ACP-cAT, **G)** *apo*-Tv6-931 ACP-cAT. Note 1: For samples **A-E**, a 6% acrylamide gel was used; for samples **F-G**, a 9% acrylamide gel was used. Note 2: For samples **A-C**, BLUEstain™ protein ladder was used; for samples **D-G**, Novex™ protein ladder was used.

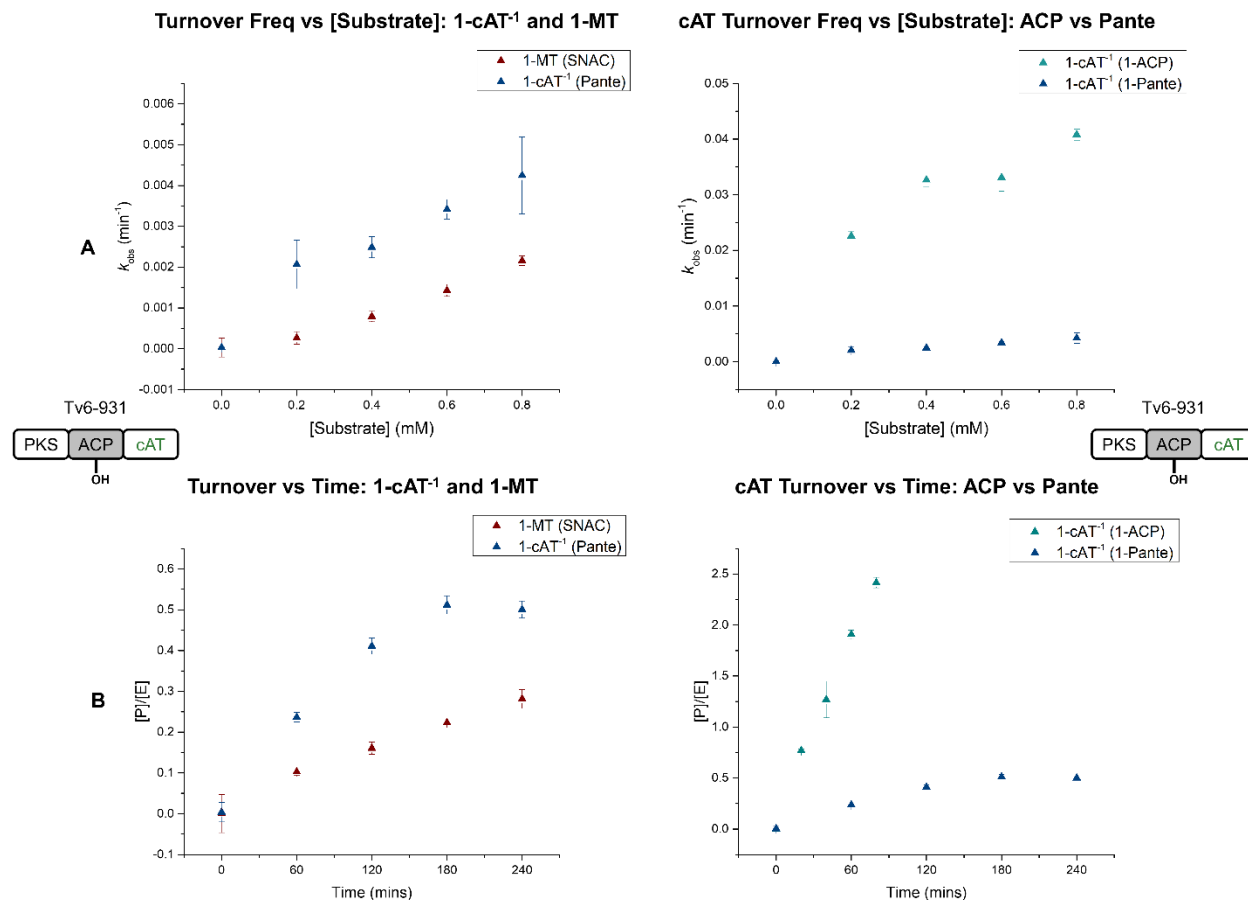


**Figure S3.6: Enzymatic Methylation Kinetic Assay:** A) Saturation curve and B) time course experiment of methylation in the biosynthesis of **1** and **2** using SNAC thioester chemical probes under pseudo-first order kinetic conditions. The saturation curve assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM SAM and increasing concentration of SNAC substrates. The time course assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM SAM, and 0.5 mM of SNAC substrates.



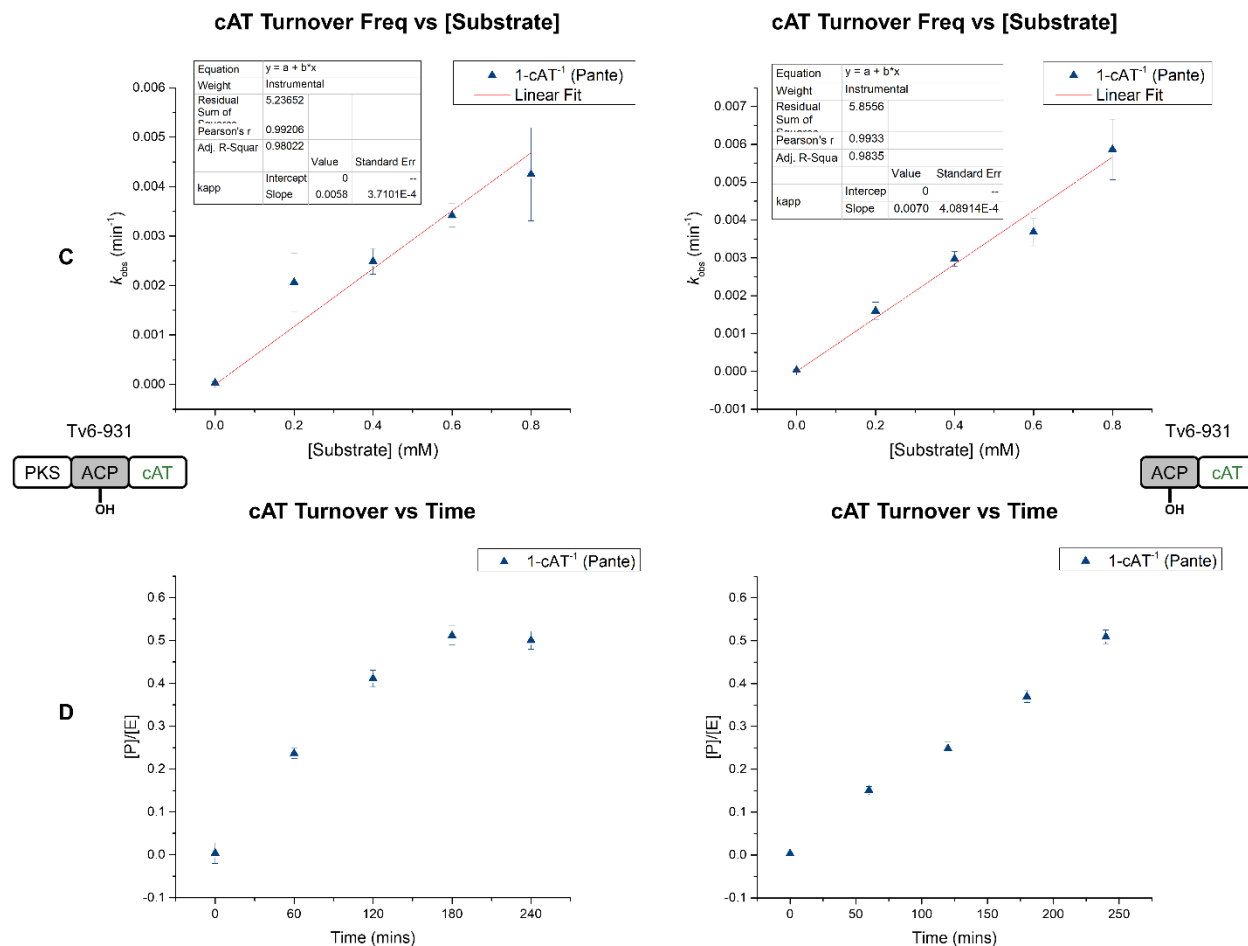
**Figure S3.7: Enzymatic Release (Esterification) Kinetic Assay: A)** Saturation curve and **B)** time course experiment of esterification (release) in the biosynthesis of **1** and **2** using SNAC thioester chemical probes under pseudo-first order kinetic conditions. The saturation curve assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM PE and increasing concentration of SNAC substrates. The time course assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM PE, and 0.5 mM of SNAC substrates. Note: methylation of 1-SNAC using *apo*-Tv6-931 HRPKS-cAT enzyme ( $k_{1\text{-MT}}$ ) was added for comparative purposes (see **Figure S3.6**)



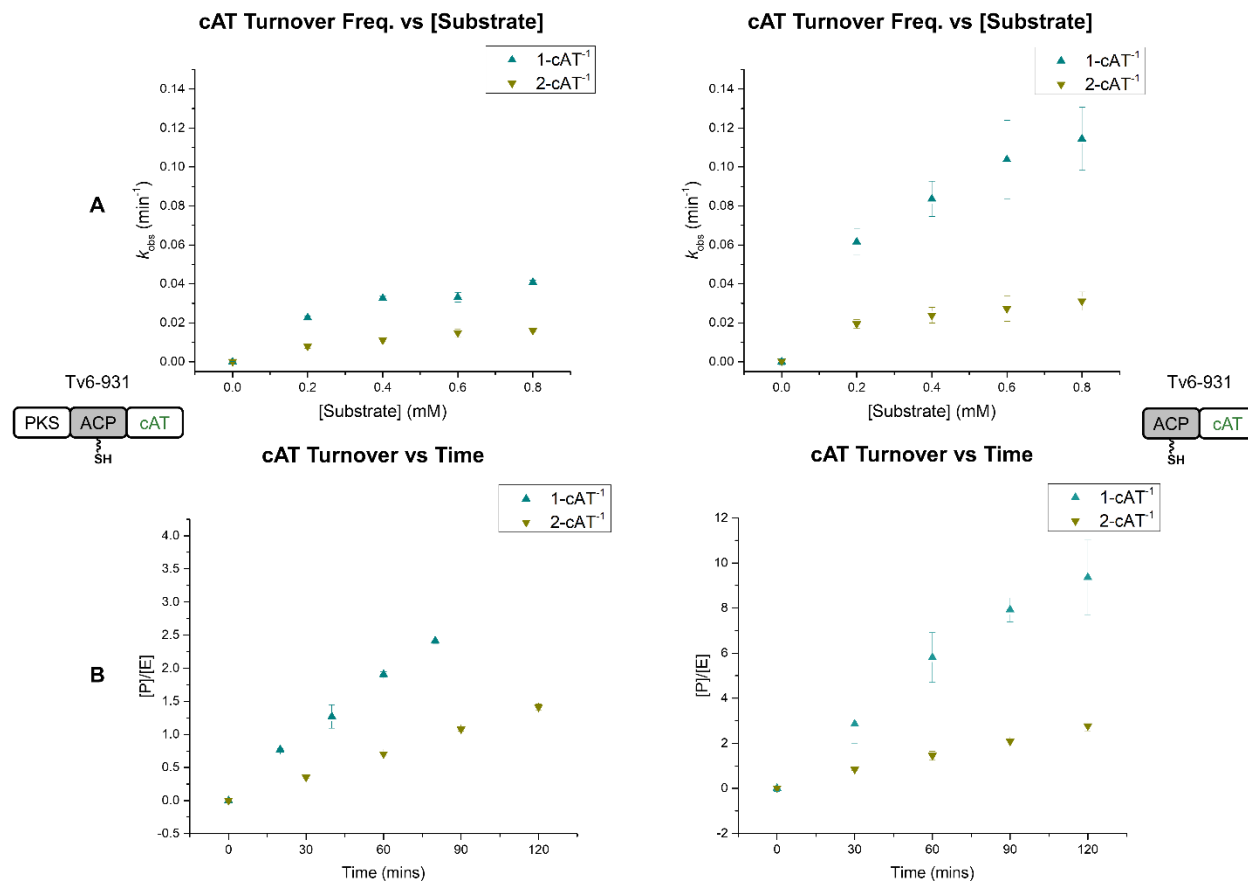


**Figure S3.8: Enzymatic Recapture (Oxyester-Thioester Transacylation) Kinetic Assay:**

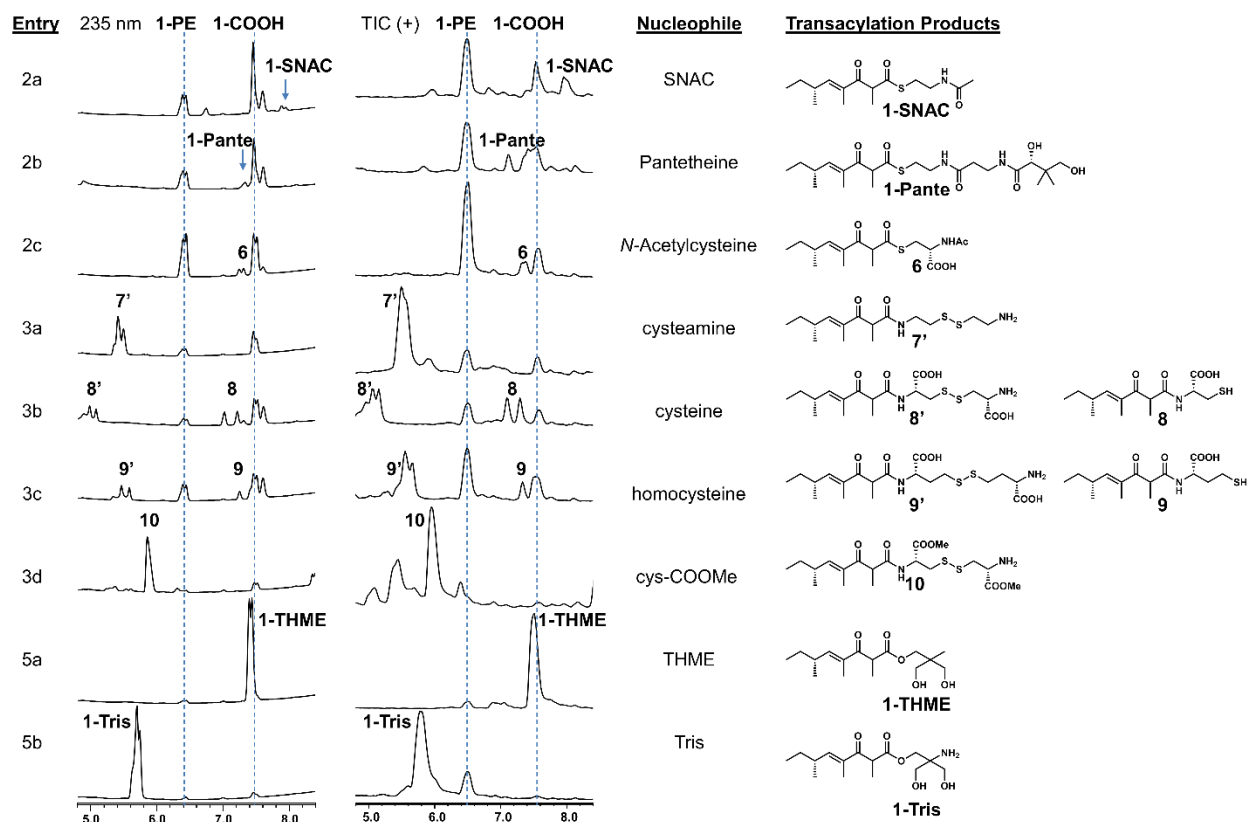
Left: **A)** Saturation curve and **B)** time course experiment of the polyketide “recapture” rate using pantetheine compared to methylation of **1-SNAC**. The polyketide “recapture” assays were performed using 15  $\mu$ M of Tv6-931 *apo*-HRPKS-cAT enzyme, 2.5 mM PE or pantetheine, and 0.5 mM of 1-PE. Right: **A)** Saturation curve and **B)** time course experiment of the polyketide “recapture” rate using *holo*-ACP forming **1-ACP** compared with pantetheine forming **1-Pante**. Note: Left: methylation of **1-SNAC** using *apo*-Tv6-931 HRPKS-cAT enzyme ( $k_{1-MT}$ ) was added for comparative purposes (see **Figure S3.5**). Right: recapture of *holo*-ACP using *apo*-Tv6-931 HRPKS-cAT enzyme ( $k_{1-cAT}^{-1}$ ) was added for comparative purposes (see **Figure S3.8**)



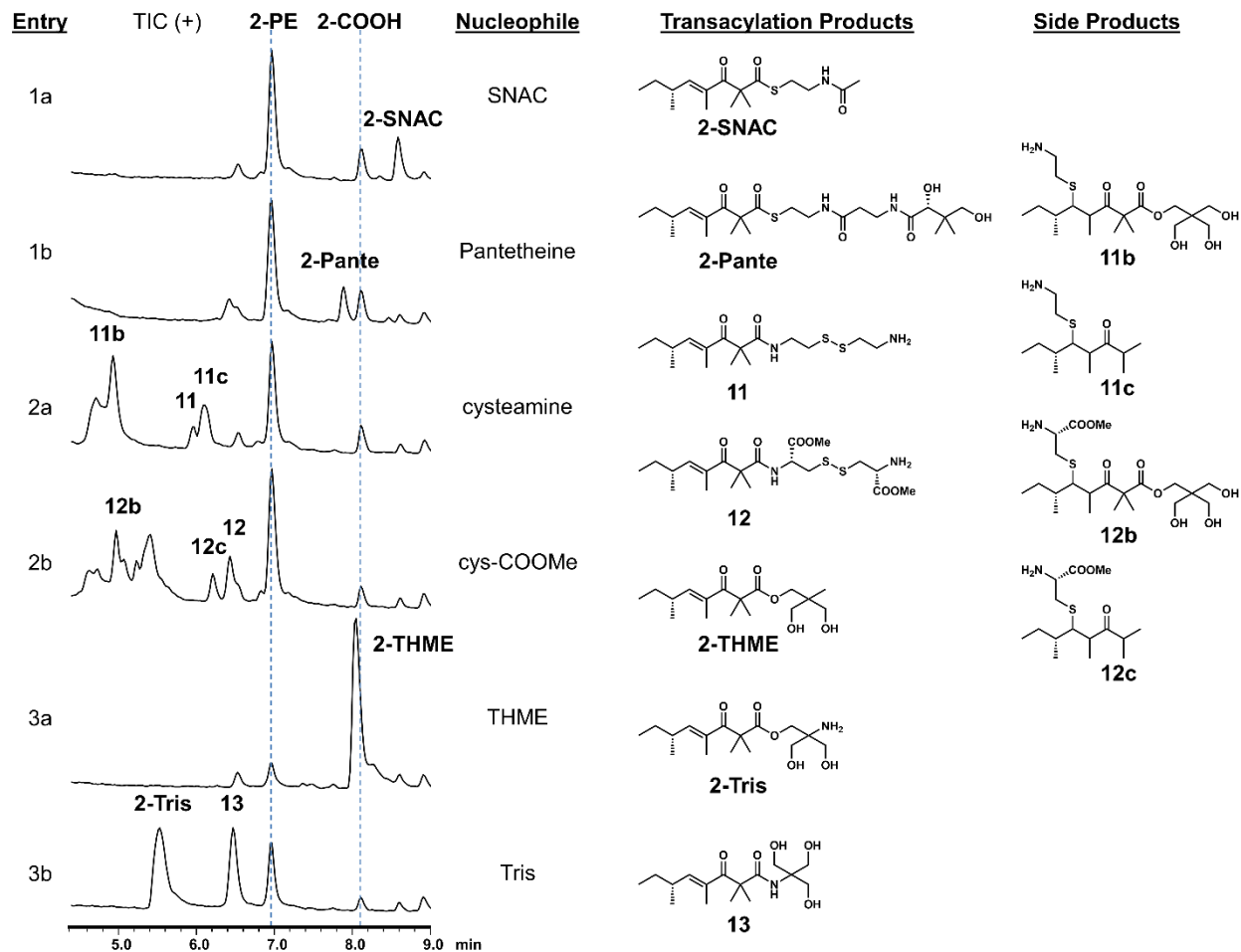
**Figure S3.8: Enzymatic Recapture (Oxyester-Thioester Transacylation) Kinetic Assay (Continued):** **C)** Saturation curve and calculation of  $k_{cat}/K_M$  and **B)** time course experiment of polyketide “recapture” using **1-PE** and pantetheine under pseudo-first order kinetic conditions. The saturation curve assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM pantetheine and increasing concentration of **1-PE**. The time course assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM pantetheine, and 0.5 mM of **1-PE**.



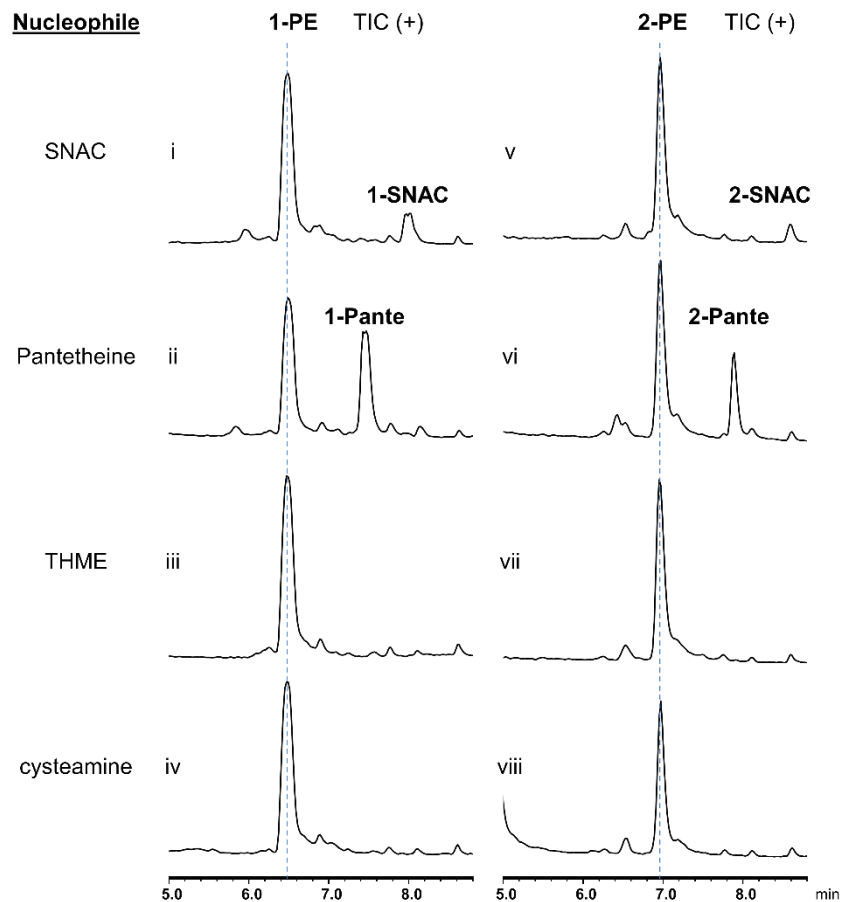
**Figure S3.9: Enzymatic Recapture (Transesterification) Kinetic Assay: A)** Saturation curve and **B)** time course experiment of the polyketide “recapture” rate or transacylation of oxyesters **1** and **2** to form polyketide-ACP adducts, **1-ACP** and **2-ACP**. The formation of the polyketide-ACP adducts were estimated using an overall transesterification under pseudo-first order kinetic conditions. The saturation curve assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM PE, and increasing concentrations of **1-THME** or **2-THME**. The time course assays were performed using 15  $\mu\text{M}$  of enzyme, 2.5 mM PE, and 0.5 mM of **1-THME** or **2-THME**.



**Figure S3.10: Chemoenzymatic Functionalization of 1-PE Using *holo*-ACP-cAT:** Various transacylation of cysteamine derivatives and polyol with 1-PE using *holo*-ACP-cAT. Different thioester, amide and oxyester polyketide conjugates were observed.

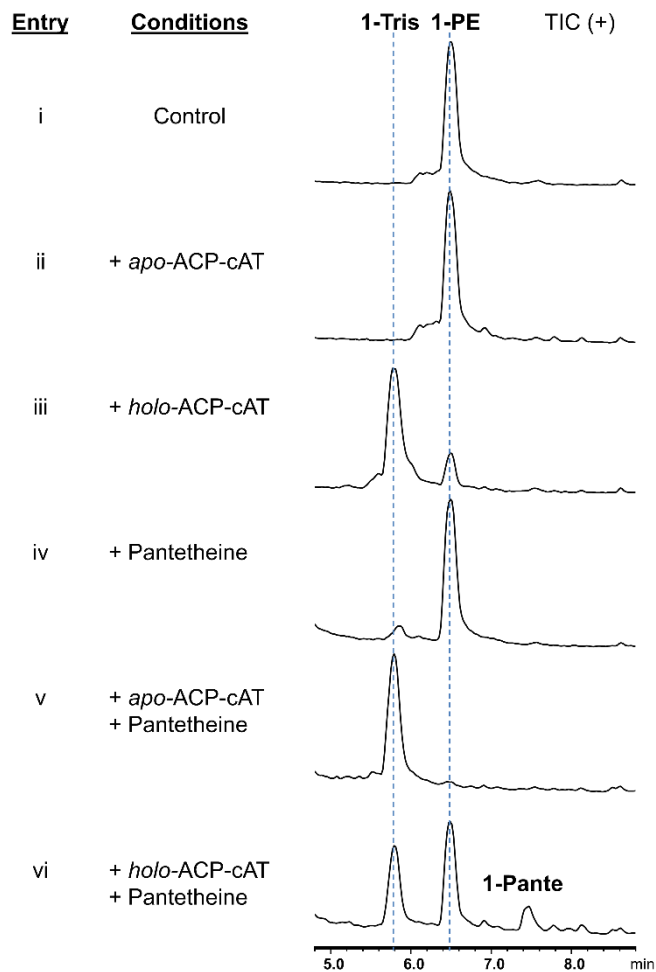


**Figure S3.11: Chemoenzymatic Functionalization of 2-PE Using *holo*-ACP-cAT:** Various transacylation of cysteamine derivatives and polyol with **2-PE** using *holo*-ACP-cAT (TIC). Different thioester, amide and oxyster polyketide conjugates were observed. Putative side products from Michael addition are also illustrated.



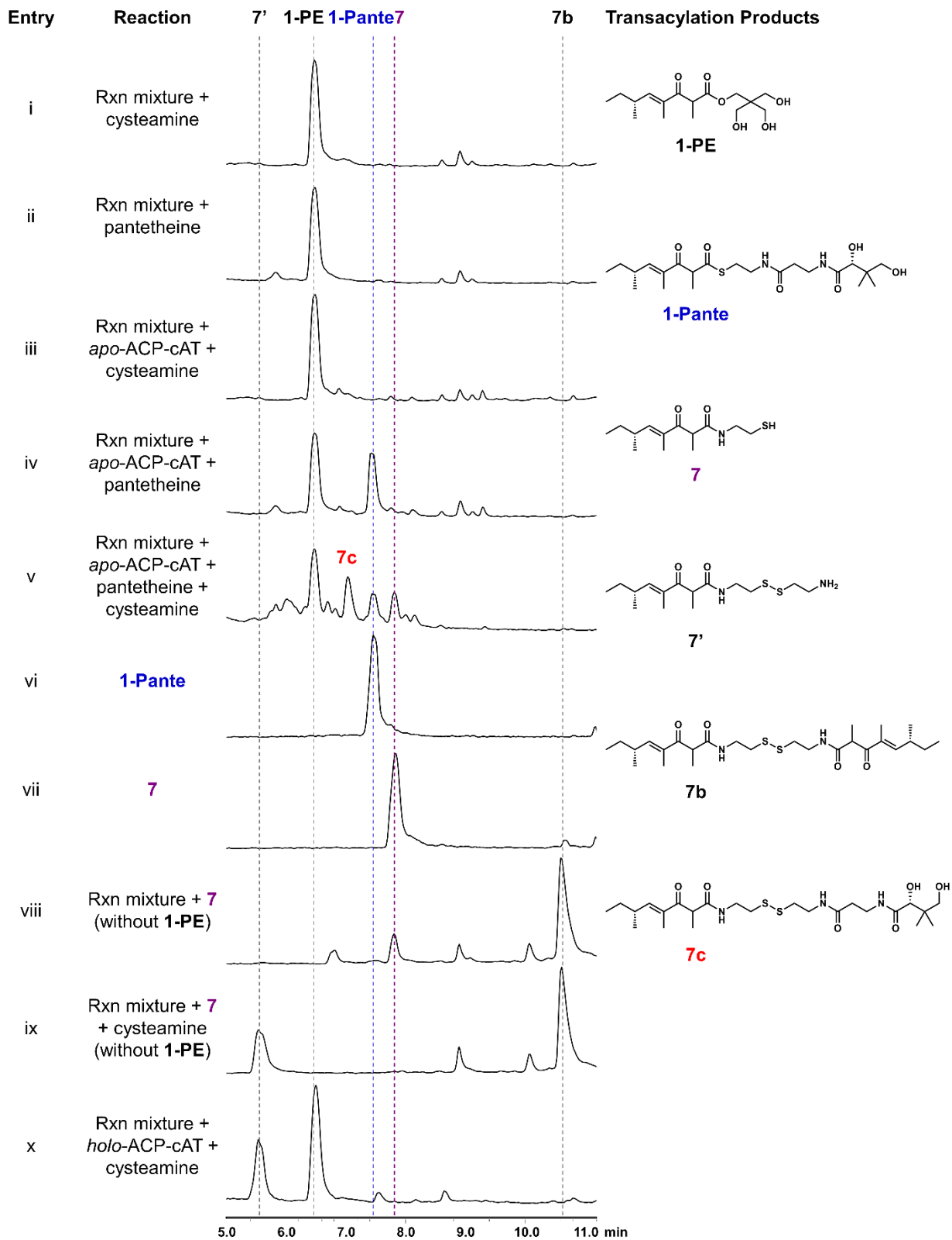
**Figure S3.12: Enzymatic Functionalization of 1-PE and 2-PE Using *apo*-ACP-cAT:**

Transacylation of SNAC and pantetheine with **1-PE** and **2-PE** using *apo*-ACP-cAT. Most of the transacylation reactions observed using *holo*-ACP-cAT were abolished (**Figure S3.9**). The only polyketide-thioester conjugates observed using *apo*-ACP-cAT were **1-SNAC**, **1-Pante**, **2-SNAC**, and **2-Pante**.



**Figure S3.13: Supplementation of Pantetheine Restores *apo*-ACP-cAT Transesterification Activity:** Supplementation of the *apo*-ACP-cAT with exogenous pantetheine restores the enzymatic transesterification activity. Control reaction: 0.25 mM **1-PE**, 10 mM Tris, 2mM DTT, 20 mM Na<sub>2</sub>HPO<sub>4</sub> (pH 8.0). Additives: 30 μM ACP-cAT (*apo* or *holo*) and 10 mM pantetheine.

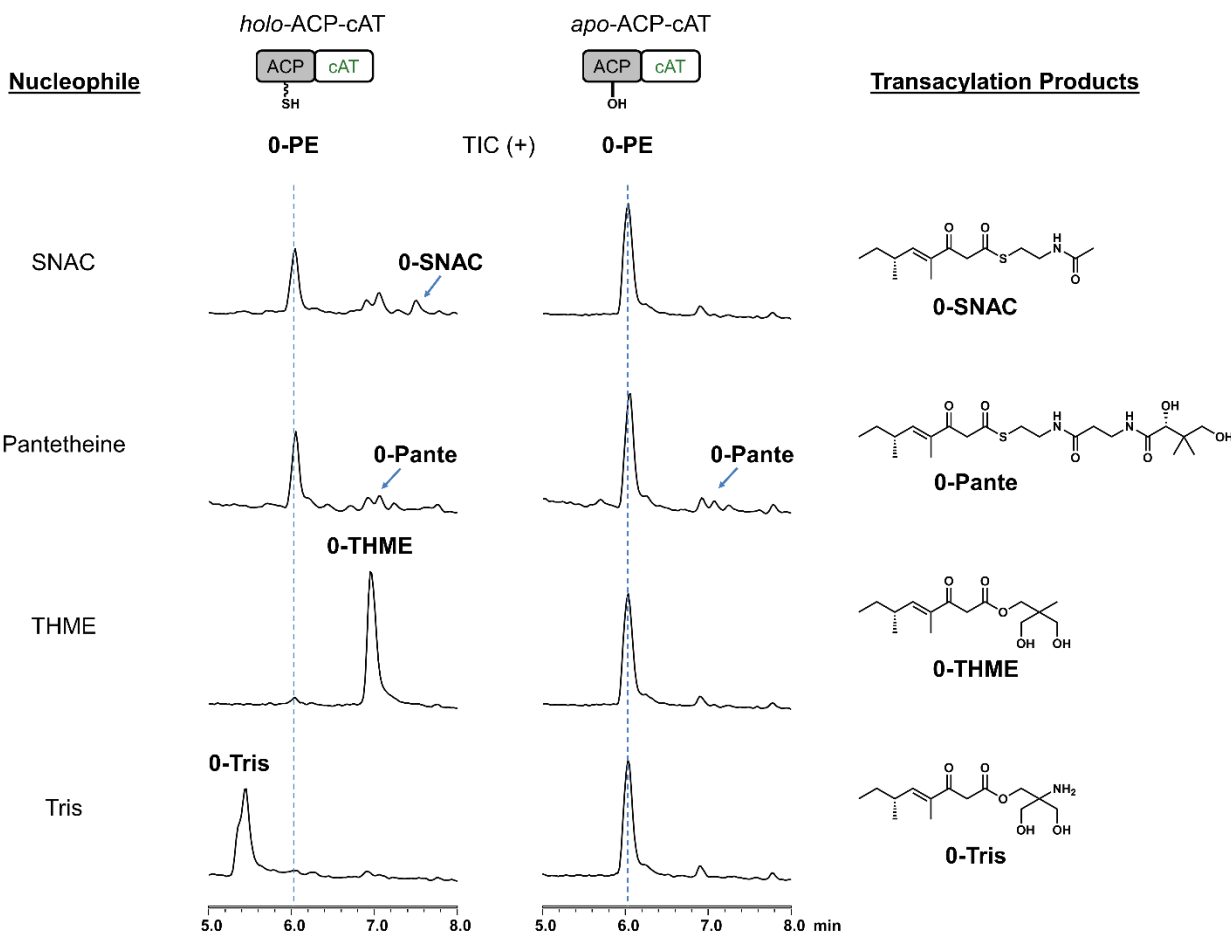
**Figure S3.14: Supplementation of Pantetheine Restores *apo*-ACP-cAT Transthioesterification Activity**



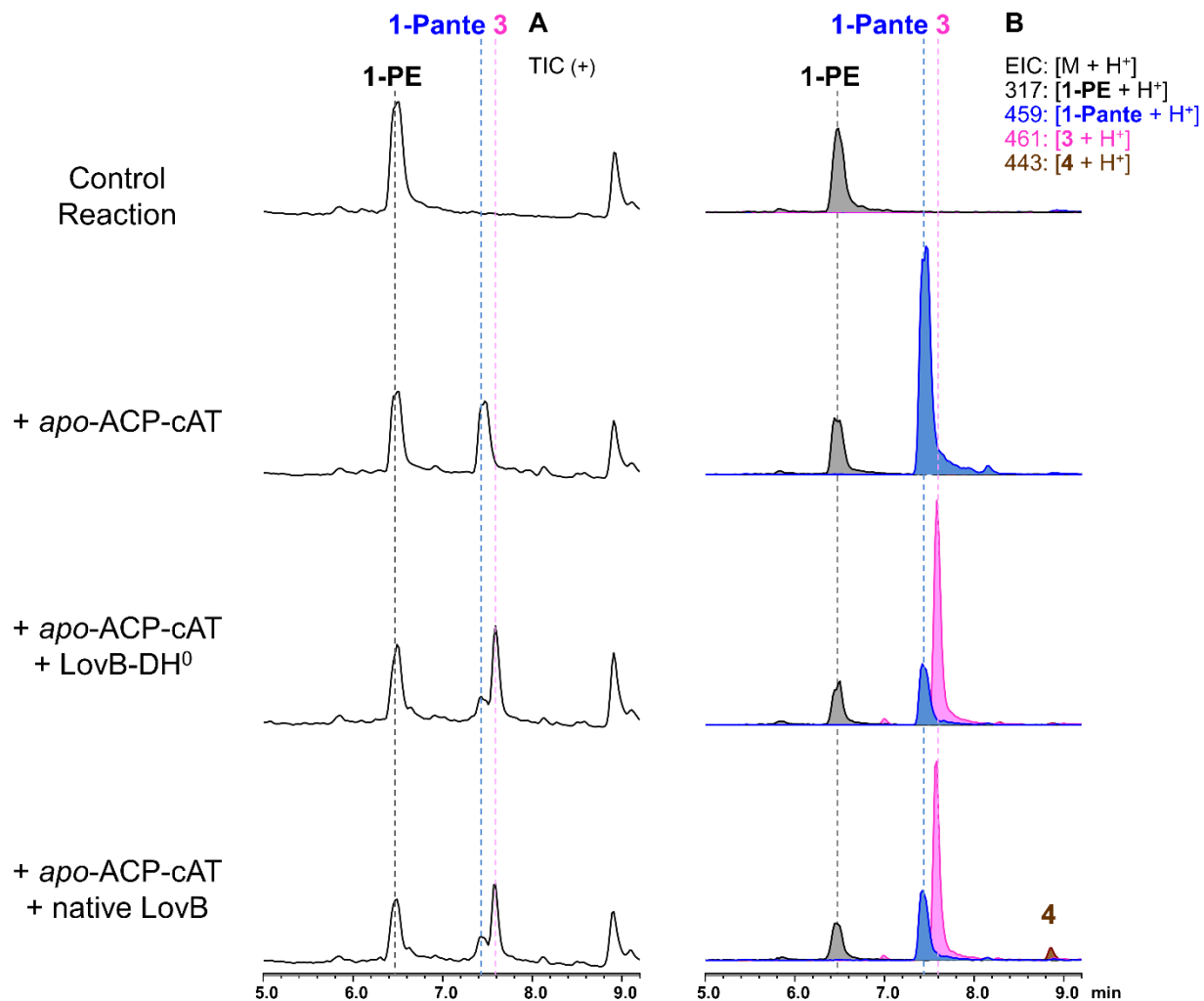


### Figure S3.14: Supplementation of Pantetheine Restores *apo*-ACP-cAT

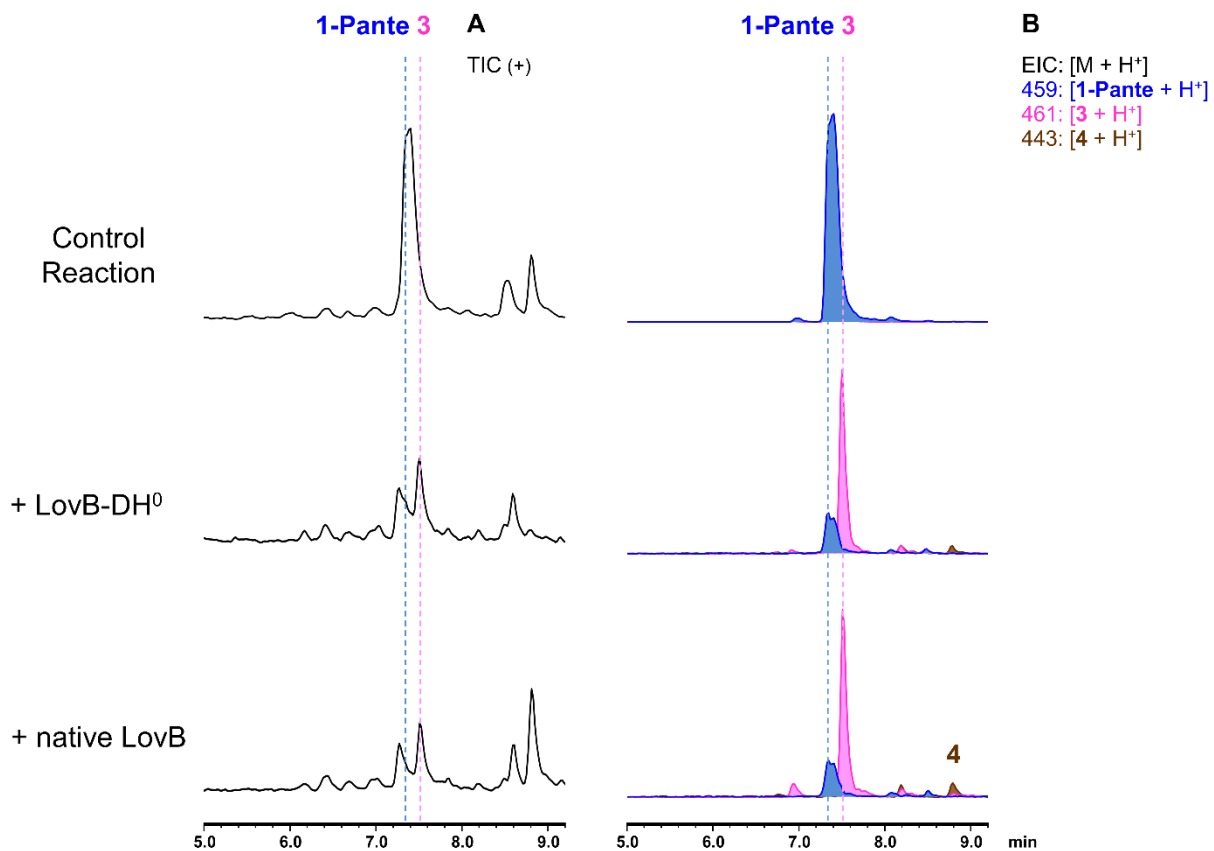
**Transthioesterification Activity:** Supplementation of the *apo*-ACP-cAT with exogenous pantetheine restores the chemical transthoesterification activity. Rxn mixture: 0.25 mM **1-PE**, 2 mM DTT, 20 mM Na<sub>2</sub>HPO<sub>4</sub> (pH 8.0). Additives: 30 μM *apo*-ACP-cAT (entries iii-v), 2 mM cysteamine (entries i, iii, v, ix), 10 mM cysteamine (entry x), 10 mM pantetheine (entries ii, iv, v), 30 μM *holo*-ACP-cAT. Other: 0.25 mM **1-Pante** (entry vi), 0.25 mM **7** (entry vii-ix). Note: Entry vii was prepared and quenched immediately upon purification of **7** due to susceptibility to oxidation to **7c**



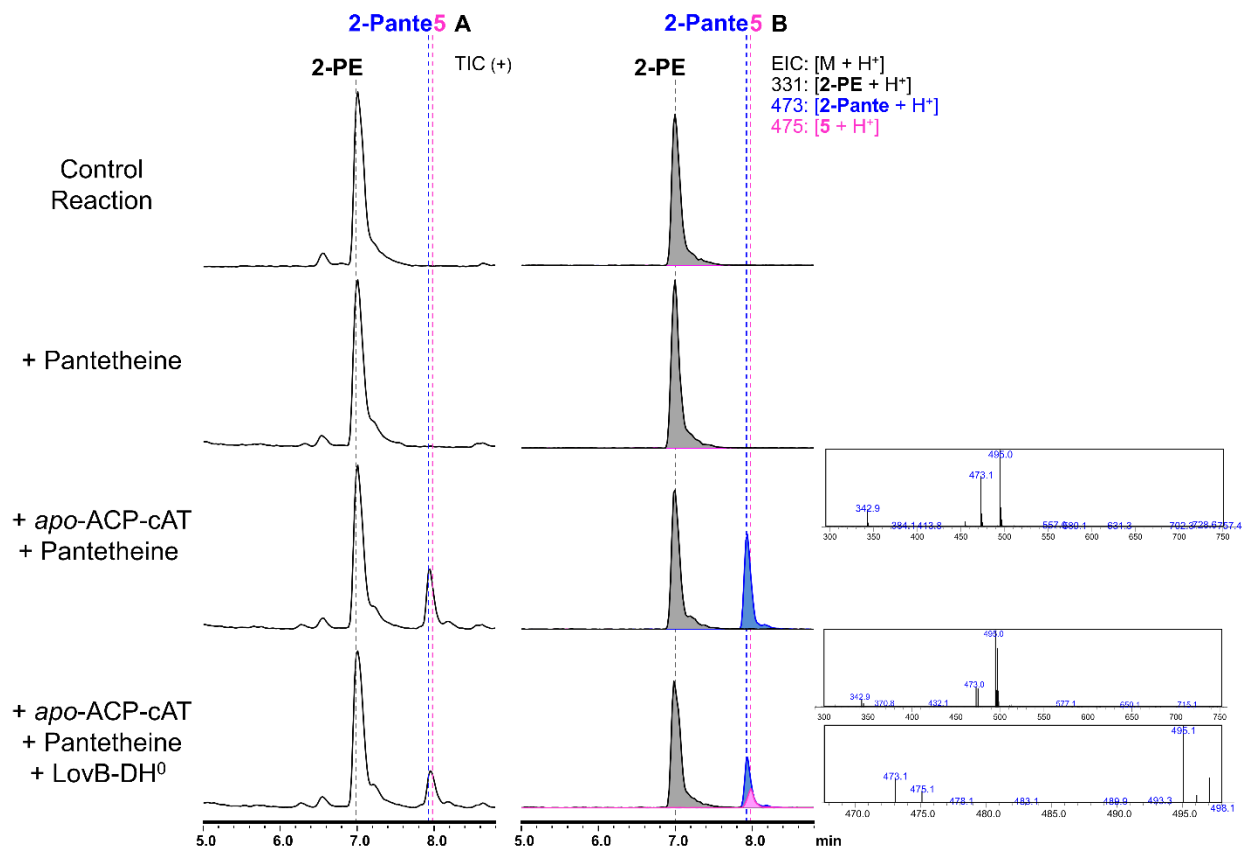
**Figure S3.15: Enzymatic Functionalization of 0-PE Using *holo*-ACP:** Various transacylation of cysteamine derivatives and polyol with **0-PE** using *holo*-ACP-cAT (TIC). Different thioester and oxyester polyketide conjugates were observed. The cAT-catalyzed esterification of **0-PE** was abolished when using the *apo* form of the enzyme.



**Figure S3.16:** Coupling ACP-cAT with LovB enzyme with **1-PE** for chemoenzymatic synthesis of **1-Pante**, **3**, and **4**: A) TIC (+), B) EIC (+) traces. Control Reaction: 0.5 mM **1-PE**, 10 mM pantetheine, 2 mM NADPH, 2 mM DTT, 30 mM Na<sub>2</sub>HPO<sub>4</sub> (pH 8.0) Additives: 30 μM *apo*-ACP-cAT, 10 μM LovB-DH<sup>0</sup> and 10 μM native LovB.



**Figure S3.17:** Conversion of **1-Pante** by LovB enzyme for chemoenzymatic synthesis of **3**, and **4**: A) TIC (+), B) EIC (+) traces. Control Reaction: 0.5 mM **1-Pante**, 2 mM NADPH, 20 mM Na<sub>2</sub>HPO<sub>4</sub> (pH 8.0) Additives: 10 μM LovB-DH<sup>0</sup> and 10 μM native LovB.



**Figure S3.18:** Coupling ACP-cAT with LovB enzyme with **2-PE** for chemoenzymatic synthesis of **2-Pante**, and **5**: A) TIC (+), B) EIC (+) traces. Control Reaction: 0.5 mM **2-PE**, 10 mM pantetheine, 2 mM NADPH, 2 mM DTT, 30 mM Na<sub>2</sub>HPO<sub>4</sub> (pH 8.0) Additives: 30 μM *apo*-ACP-cAT, and 10 μM LovB-DH<sup>0</sup>.

## Section 3.11 Experimental Methods

### Bioinformatic Analysis:

#### *Fungal KS Phylogenetic Tree Development*

The fungal KS phylogenetic tree was generated according to the protocol detailed by Li *et al.*<sup>43</sup> The highlighted HRPKS-cAT subclade was identified using the bioinformatic tools: basic local alignment search tool (BLAST),<sup>44-45</sup> and Conserved Domain Database (CDD).<sup>46-47</sup>

#### *Genome mining and in silico analysis for new fungal HRPKS*

The genomes of fully-sequenced fungi were readily accessed using the Joint Genome Institute (JGI) portal.<sup>48-49</sup> Using BLAST, these genomes were probed using various annotated KS domains serve as query templates. An additional search using the ACP-cAT region as a query template distinguished the novel HRPKS-cAT hybrids from traditional HRPKS enzymes. Functional domains in the translated sequences were also predicted using BLAST,<sup>44-45</sup> and CDD<sup>46-47</sup>. Protein structures were predicted using the Phyre2 server.<sup>50</sup>

### Materials and Methods

#### *Microbial Strains*

The *Saccharomyces cerevisiae* BJ5464-NpgA<sup>51</sup> cells (*MAT $\alpha$  ura3-52 trp1 leu2- $\Delta$ 1 his3 $\Delta$ 200 pep4::HIS3 prb1 $\Delta$ 1.6R can1 GAL*) was used for the expression of the Tv6-931 whole protein, Tv6-931 ACP-cAT, LovB-DH<sup>0</sup>, and native LovB enzyme. This strain was also used in the *in vivo* biosynthesis of Tv6-931 polyketide adducts, **1** ( $\alpha$ -methyl) and **2** ( $\alpha,\alpha$ -dimethyl polyketide).

The *Saccharomyces cerevisiae* BJ5464-*apo* cells (*MAT $\alpha$  ura3-52 trp1 leu2- $\Delta$ 1 his3 $\Delta$ 200 pep4::HIS3 prb1 $\Delta$ 1.6R can1 GAL*) was used for the expression of the *apo*-Tv6-931, *apo*-Tv6-931  $\Delta$ cAT, and *apo*-Tv6-931 ACP-cAT enzymes.

#### *HPLC-MS Data Collection*

The HPLC-MS data collection was performed using a Shimadzu 2020 EVLC-MS (Phenomenex® Luna, 5  $\mu$ m, 2.0  $\times$  100 mm, C-18 column) using positive and negative mode electrospray ionization. The elution method was a linear gradient of 5-95% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes with a flow rate of 0.3 mL/min. The HPLC buffers were supplemented with 0.1% formic acid (v/v).

#### *Analytic HPLC Data Collection*

The analytic HPLC-MS data collection was performed using a Shimadzu Prominence HPLC (Phenomenex® Luna, 5  $\mu$ m, 2.0  $\times$  100 mm, C-18 column). The elution method was a linear gradient of 5-95% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes with a flow rate of 0.3 mL/min. The HPLC buffers were supplemented with 0.05% formic acid (v/v).

#### *Analytic Chiral HPLC Data Collection*

The analytic HPLC-MS data collection was performed using a Shimadzu Prominence HPLC (Phenomenex® Lux Cellulose-1, 5  $\mu$ m, 4.6 x 250 mm column). The elution method was a linear gradient of 8-13% (v/v) isopropanol (IPA)/hexanes in 15 minutes, followed by 95% isopropanol (IPA)/hexanes for 3 minutes with a flow rate of 1.0 mL/min. The HPLC buffers were supplemented with 0.1% triethylamine (v/v).

### *HPLC Purification*

The HPLC purifications were performed using a Shimadzu Prominence HPLC (Phenomenex® Kinetex, 5  $\mu$ m, 10.0  $\times$  250 mm, C-18 column). The elution methods were variable linear gradients in CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes with a flow rate of 3.0 mL/min.

### *NMR Analysis*

All NMR spectra including <sup>1</sup>H, <sup>13</sup>C, COSY, HSQC, HMBC and NOESY spectra were obtained on Bruker AV500 spectrometer with a 5 mm dual cryoprobe at the UCLA Molecular Instrumentation Center. The NMR solvent, CDCl<sub>3</sub>, used for these experiments was purchased from Cambridge Isotope Laboratories, Inc.

### *High-Resolution Mass Spectroscopy*

High-resolution mass spectra were obtained from Thermo Fisher Scientific Exactive Plus with IonSense ID-CUBE DART source (APCI) and Waters LCT Premier (ESI) at the UCLA Molecular Instrumentation Center.

### *General Molecular Biology Experiments*

General molecular biology and cloning techniques were performed according to protocols described elsewhere.<sup>52</sup> PCR was performed using Phusion® and Q5® High-Fidelity DNA Polymerase (New England Biolabs) according to protocols recommended by the manufacturer. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs). PCR products were confirmed by DNA sequencing by Laragen, California, USA.

TOP10 (Invitrogen), DH10B (Invitrogen) and XL1-Blue (Stratagene) *E. coli* cells were used for DNA manipulation, following standard recombinant DNA techniques.

### **Synthesis of SNAC Substrates:**

Note: For clarity and simplicity, an abbreviated version of the “Synthesis of SNAC Substrates” was included. For the full version, please refer to the Supplementary Information of the corresponding manuscript.

### *General Synthetic Procedures*

Reactions involving either air or moisture sensitive reactants were conducted under an atmosphere of argon. All solvents and chemicals were reagent grade and used as supplied from Sigma-Aldrich unless otherwise stated. DMP was sourced from AK Scientific at 95% purity. Anhydrous solvents required were dried according to the procedures outlined in Perrin and Armarego.<sup>53</sup> Removal of solvent was performed under reduced pressure, below 40 °C, and in special cases highlighted at 0 °C using a Büchi rotary evaporator. All reactions and fractions from column chromatography were monitored by thin layer chromatography (TLC). Analytical TLC was done on glass plates (5 × 1.5 cm) pre-coated (0.25 mm) with silica gel (normal SiO<sub>2</sub>, Merck 60 F254). Compounds were visualized by exposure to UV light and/ by exposing the plates to KMnO<sub>4</sub> solution, followed by heating. Flash chromatography was performed on silica gel (EM Science, 60Å pore size, 230-400 mesh).

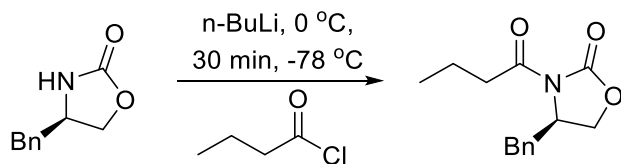
### *Spectroscopic Analyses*

Nuclear magnetic resonance (NMR) spectra were obtained on Varian Inova 500 or 700 MHz spectrometer. <sup>1</sup>H NMR chemical shifts are reported in parts per million (ppm) using the



residual proton resonance of solvents as reference:  $\text{CDCl}_3$   $\delta$  7.26, or  $\text{CD}_2\text{Cl}_2$   $\delta$  5.32.  $^{13}\text{C}$  NMR chemical shifts are reported relative to  $\text{CDCl}_3$   $\delta$  77.1, or  $\text{CD}_2\text{Cl}_2$   $\delta$  53.8. Optical rotations were measured on a Perkin Elmer 241 polarimeter with a microcell (10 cm, 1 mL) at 25 °C. Infrared spectra (IR) were recorded on a Nicolet Magna 750. Cast Film refers to the evaporation of a solution on a NaCl plate. Mass spectra were recorded on a Kratos IMS-50 (high resolution, electron impact ionization (EI)) or by using an Agilent Technologies 6220 orthogonal acceleration TOF instrument equipped with +ve and -ve ion ESI ionization source, and full-scan MS (high resolution analysis) with two-point lock mass correction operating mode. The instrument inlet was an Agilent Technologies 1200 SL HPLC system.

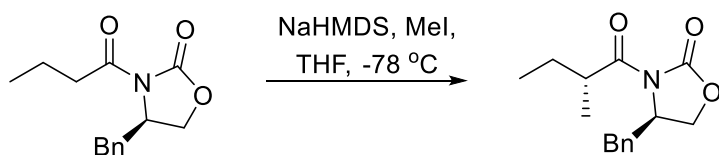
#### (*R*)-4-benzyl-3-butyryloxazolidin-2-one



This known compound was prepared following a literature protocol.<sup>24</sup> (*R*)-4-benzyl-oxazolidin-2-one (7.00 g, 39.5 mmol) was dissolved in dry THF (150 mL) and cooled to -78 °C under argon while stirring. *n*-BuLi (2.5 M, 15.8 mL, 39.5 mmol) was slowly added *via* syringe and the reaction mixture was stirred for 30 min. Butyryl chloride (5.31 mL, 51.4 mmol) was added and the reaction mixture was stirred for 2 hr at -78 °C before removing the cooling bath allowing the mixture to warm to room temperature. The reaction was quenched with saturated  $\text{NH}_4\text{Cl}$  solution (50 mL) and the THF was removed *in vacuo*. The aqueous layer was extracted with EtOAc (3  $\times$  50 mL). The combined organic extracts were washed with saturated  $\text{NaHCO}_3$  solution (2  $\times$  20 mL), brine (2  $\times$  20 mL), dried over  $\text{Na}_2\text{SO}_4$ , and concentrated *in vacuo*. The residue was purified by flash chromatography ( $\text{SiO}_2$ , 10–20% EtOAc in hexanes), yielding

the product as a colourless oil,  $R_f = 0.35$  in 20% EtOAc in hexanes, (10.1 g, 85%).  $[\alpha]_D^{25} = -48.4$  ( $c$  1.0,  $\text{CHCl}_3$ ); IR ( $\text{CHCl}_3$ , cast) 3064, 2965, 2933, 2876, 1783, 1701  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 700 MHz)  $\delta$  7.34 – 7.32 (m, 2H), 7.28 – 7.26 (m, 1H), 7.21 – 7.20 (m, 2H), 4.69 – 4.66 (m, 1H), 4.21 – 4.15 (m, 2H), 3.30 (dd,  $J = 13.3, 2.8$  Hz, 1H), 2.97 – 2.94 (m, 1H), 2.90 – 2.86 (m, 1H), 2.77 (dd,  $J = 13.3, 9.8$  Hz, 1H), 1.75 – 1.71 (m, 2H), 1.01 (t,  $J = 7.7$  Hz, 3H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 176 MHz)  $\delta$  173.4, 153.6, 135.5, 129.5, 129.1, 127.5, 66.3, 55.3, 38.1, 37.5, 17.8, 13.8. HRMS (ESI) Calcd for  $\text{C}_{14}\text{H}_{17}\text{NO}_3\text{Na}$   $[\text{M}+\text{Na}]^+$  270.1101, found 270.1103.

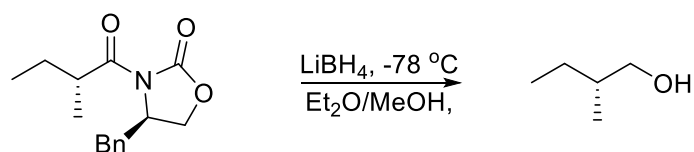
**(R)-4-benzyl-3-((R)-2-methylbutanoyl)oxazolidin-2-one**



This known compound was prepared following a literature protocol.<sup>24</sup> (R)-4-benzyl-3-butyloxazolidin-2-one (10.0 g, 40.4 mmol) was dissolved in dry THF (200 mL) and cooled to  $-78$  °C under argon while stirring. NaHMDS (1 M, 44.5 mL, 44.5 mmol) was slowly added *via* syringe and the reaction was stirred for 30 min. Methyl iodide (7.60 mL, 121 mmol) was added slowly and the reaction stirred for 3 hr. The cooling bath was then removed allowing the reaction mixture to warm to room temperature then quenched with brine (100 mL). The THF was removed *in vacuo* and the aqueous extracted with EtOAc (3 x 60 mL). The combined organic phases were dried with  $\text{Na}_2\text{SO}_4$ , concentrated *in vacuo*, and purified by column chromatography ( $\text{SiO}_2$ , 5–10% EtOAc in hexanes), yielding the product as a pale-yellow oil,  $R_f = 0.36$  in 20% EtOAc in hexanes, (7.9 g, 75%).  $[\alpha]_D^{25} = -68.03$  ( $c$  1.0,  $\text{CHCl}_3$ ); IR ( $\text{CHCl}_3$ , cast) 3029, 2969, 2934, 2877, 1781, 1698  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz)  $\delta$  7.35 – 7.32 (m, 2H), 7.29 – 7.26 (m, 1H), 7.22 – 7.21 (m, 2H), 4.70 – 4.66 (m, 1H), 4.22 – 4.15 (m, 2H), 3.66 – 3.62 (m, 1H),

3.27 (dd,  $J = 13.5, 3.0$  Hz, 1H), 2.77 (dd,  $J = 13.0, 9.5$  Hz, 1H), 1.80 – 1.75 (m, 1H), 1.50 – 1.45 (m, 1H), 1.22 (d,  $J = 7.0$  Hz, 3H), 0.93 (t,  $J = 7.5$  Hz, 3H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz)  $\delta$  177.3, 153.2, 135.5, 129.6, 129.1, 127.5, 66.2, 55.5, 39.3, 38.1, 26.6, 17.0, 11.8. HRMS (ESI) Calcd for  $\text{C}_{15}\text{H}_{19}\text{NO}_3\text{Na}$   $[\text{M}+\text{Na}]^+$  284.1257, found 284.1256.

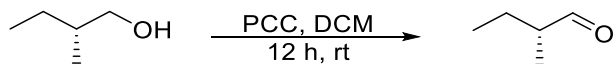
### (*R*)-2-Methylbutan-1-ol



This known compound was prepared following a modified literature protocol.<sup>24</sup> (*R*)-4-benzyl-3-((*R*)-2-methylbutanoyl)oxazolidin-2-one (7.80 g, 29.9 mmol) was dissolved in dry ether (100 mL) and cooled to  $-20$  °C under argon while stirring. Methanol (1.45 mL, 35.8 mmol) was then added followed by slow addition of  $\text{LiBH}_4$  (2 M, 17.9 g, 35.8 mmol) to the reaction mixture. The reaction mixture was then moved to an ice bath at  $0$  °C and stirred for 2 hr. A solution of  $\text{NaOH}$  (1 M, 40 mL) was slowly added to the reaction mixture while stirring, until two clear layers were observed. The layers were separated and the aqueous layer was extracted with ether (3 x 20 mL). The organic extracts were washed with brine (2 x 20 mL), dried with  $\text{Na}_2\text{SO}_4$ , and concentrated *in vacuo* using an ice bath, the solvent was collected from the trap and concentrated separately. The combined residue was purified by column chromatography ( $\text{SiO}_2$ , 30% ether in pentanes), yielding the product as a colourless oil,  $R_f = 0.46$  in 50% EtOAc in hexanes, (2.44 g, 93%). Due to the extremely volatile nature of the compound it was used immediately.  $[\alpha]_{\text{D}}^{25} = 6.1$  ( $c$  0.95,  $\text{CH}_2\text{Cl}_2$ ); IR ( $\text{CH}_2\text{Cl}_2$ , cast) 3442, 2960, 2928, 2857, 1274  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CD}_2\text{Cl}_2$ , 500 MHz)  $\delta$  3.43 (dd,  $J = 10.5, 5.8$  Hz, 1H, H4), 3.35 (dd,  $J = 10.5, 6.4$  Hz, 1H, H4'), 1.51 – 1.44 (m, 1H, H3), 1.43 – 1.36 (m, 1H, H2), 1.31 (br s, 1H, H5), 1.14 – 1.08

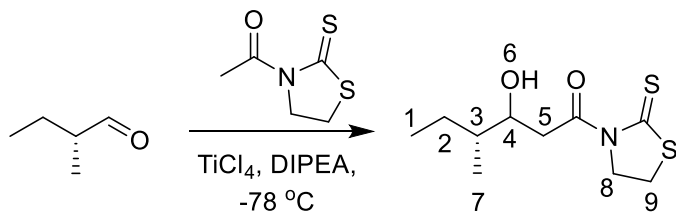
(m, 1H, H2'), 0.87 (t,  $J = 7.5$  Hz, 3H, H1), 0.86 (d,  $J = 6.5$  Hz, 3H, H3).  $^{13}\text{C}$  NMR ( $\text{CD}_2\text{Cl}_2$ , 125 MHz)  $\delta$  68.2, 37.8, 26.1, 16.2, 11.5. HRMS (EI) Calcd for  $\text{C}_5\text{H}_{12}\text{O}$   $[\text{M}]^+$  88.0888, found 88.0878.

### (*R*)-2-Methylbutanal



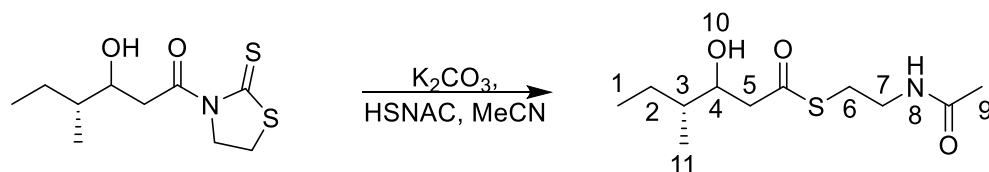
This known compound was prepared following a modified literature protocol.<sup>54</sup> To a suspension of PCC (8.94 g, 41.5 mmol) in 100 mL of dry DCM was added a solution of (*R*)-2-Methylbutan-1-ol (2.44 g, 27.7 mmol) in 10 mL of dry DCM. The reaction mixture was then stirred at room temperature for 12 hr, before 10 g of silica gel was added and allowed to stir for another 30 min. Ether (100 mL) was then added and the reaction mixture was filtered through a pad of celite. The filtrate was then concentrated *in vacuo* at 0 °C with the solvent collected and concentrated separately. The combined crude mixture was purified by passing through a plug of silica using a gradient of 20–30% ether in pentanes. The pure product; (*R*)-2-methylbutanal was isolated as a pale yellow oil,  $R_f = 0.80$  in DCM, (2.2 g, 90%). Due to the extremely volatile nature of the compound it was used immediately in the following step.  $^1\text{H}$  NMR ( $\text{CD}_2\text{Cl}_2$ , 500 MHz)  $\delta$  9.60 (d,  $J = 1.9$  Hz, 1H, H4), 2.31 – 2.19 (m, 1H, H3), 1.73 (dq,  $J = 13.9, 7.5, 6.4$  Hz, 1H, H2), 1.46 – 1.37 (m, 1H, H2'), 1.06 (d,  $J = 7.0$  Hz, 3H, H5), 0.93 (t,  $J = 6.5$  Hz, 3H, H1).  $^{13}\text{C}$  NMR ( $\text{CD}_2\text{Cl}_2$ , 125 MHz)  $\delta$  205.5, 48.1, 23.9, 13.0, 11.5.

### 3-Hydroxy-4*R*-methyl-1-(2-thioxothiazolidin-3-yl)hexan-1-one



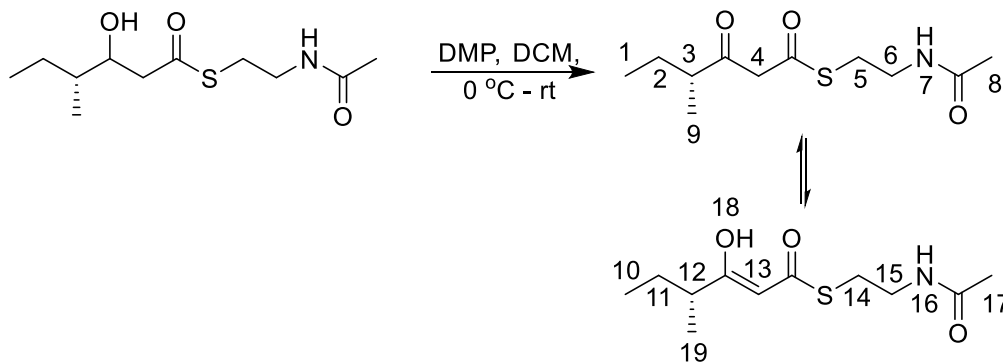
This new compound was synthesized using an established procedure.<sup>55</sup> To a stirred solution of *N*-acetyl-thiazolidine-2-thione (0.23 g, 1.4 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (25 mL) was added TiCl<sub>4</sub> (1.0 M solution in CH<sub>2</sub>Cl<sub>2</sub>, 1.4 mL, 1.4 mmol) at 0 °C under argon. The reaction mixture was stirred for 5 min to dissolve, then cooled to -78 °C. Diisopropylethylamine (0.29 mL, 1.7 mmol) was added and the reaction stirred for 1 hr. A solution of (*R*)-2-methylbutanal (0.11 g, 1.3 mmol) in CH<sub>2</sub>Cl<sub>2</sub> was added dropwise and the reaction stirred for 15 min at -78 °C, then 2 h at rt. The reaction was quenched by addition of half-saturated NH<sub>4</sub>Cl solution (25 mL). Layers were separated and the aqueous extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 25 mL). The combined organic phases were washed with brine (100 mL), dried with MgSO<sub>4</sub>, concentrated *in vacuo*, and purified by column chromatography (SiO<sub>2</sub>, 30% EtOAc in hexanes), yielding the product as a yellow oil, a 60:40 mixture of diastereomers, R<sub>f</sub> = 0.11 in 30% EtOAc in hexanes, (0.19 g, 54%). [α]<sub>D</sub><sup>25</sup> = 3.0 (*c* 1.0, CHCl<sub>3</sub>); IR (CHCl<sub>3</sub>, cast) 3499, 2961, 2932, 2875, 1697, 1368, 1281, 1157, 1053 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 4.65 – 4.55 (m, 2H, H8), 4.07 – 3.97 (m, 1H, H5), 3.50 – 3.34 (m, 2H, H7), 3.31 – 3.26 (m, 2H, H9), 2.86 (br, 1H, H6), 1.61 – 1.44 (m, 2H, H2+H3), 1.22 – 1.17 (m, 1H, H2'), 0.94 – 0.90 (m, 6H, H4+H1). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 202.0, 174.7, 71.1, 55.8, 43.4, 39.9, 28.4, 25.6, 14.1, 11.8. HRMS (ESI) Calcd for C<sub>10</sub>H<sub>17</sub>NO<sub>2</sub>S<sub>2</sub>Na [M+Na]<sup>+</sup> 270.0593, found 270.0599.

### **S-2-Acetamidoethyl 3-hydroxy-4*R*-methylhexanethioate**



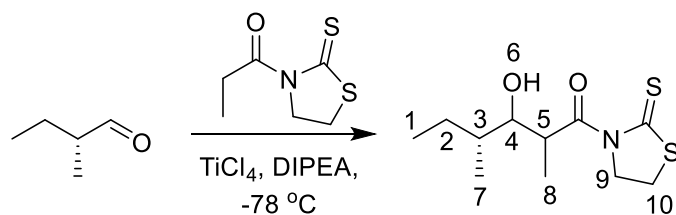
This new compound was synthesized using an established procedure.<sup>55</sup> To a stirred solution of 3-hydroxy-4*R*-methyl-1-(2-thioxothiazolidin-3-yl)hexan-1-one (0.135 g, 0.55 mmol) in 5 mL of dry acetonitrile was added K<sub>2</sub>CO<sub>3</sub> (0.226 g, 1.64 mmol) followed by *N*-acetylcysteamine (0.070 mL, 0.65 mmol). The reaction was stirred for 30 min before removing the solvent *in vacuo*. The residue was suspended in CH<sub>2</sub>Cl<sub>2</sub>, washed with brine, and dried with MgSO<sub>4</sub> before concentrating *in vacuo*. The residue was purified by column chromatography (SiO<sub>2</sub>, gradient elution from 50-100% EtOAc in hexanes), yielding the product as a colourless oil, R<sub>f</sub> = 0.18 in EtOAc, (0.089 g, 66%). [α]<sub>D</sub><sup>25</sup> = 1.70 (*c* 1.0, CHCl<sub>3</sub>); IR (CHCl<sub>3</sub>, cast) 3309, 3089, 2962, 2933, 1688, 1658, 1552 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 5.84 (br, 1H, H8), 4.03 – 3.98 (m, 0.7H, H4), 3.96 – 3.91 (m, 0.3H, H4'), 3.50 – 3.39 (m, 2H, H7), 3.10 – 3.00 (m, 2H, H6), 2.76 – 2.66 (m, 2H, H5), 2.62 (br, 0.3H, H10), 2.49 (br, 0.7H, H10'), 1.97 (s, 3H, H9), 1.56 – 1.48 (m, 1H, H2), 1.48 – 1.40 (m, 1H, H3), 1.22 – 1.12 (m, 1H, H2'), 0.94 – 0.88 (m, 6H, H1+H11). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 200.2, 200.0, 170.4, 72.4, 71.8, 48.7, 47.7, 40.1, 40.0, 39.3, 28.9, 25.6, 24.9, 23.3, 14.5, 13.8, 11.8, 11.5. HRMS (ESI) Calcd for C<sub>11</sub>H<sub>21</sub>NO<sub>3</sub>SNa [M+Na]<sup>+</sup> 270.1134, found 270.1135.

***S*-2-Acetamidoethyl 4*R*-methyl-3-oxohexanethioate (Trio-SNAC)**



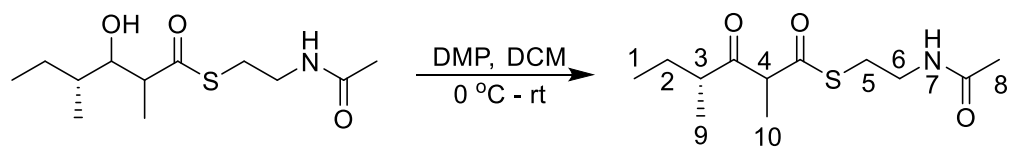
This new compound was synthesized using an established procedure.<sup>24</sup> To a stirred solution of *S*-2-acetamidoethyl 3-hydroxy-4*R*-methylhexanethioate (0.076 g, 0.31 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (5 mL) at 0 °C was added a suspension of Dess-Martin periodinane (0.047 g, 0.34 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub>. The reaction was stirred for 30 min at 0 °C then another 2.5 hr at rt before quenching with 1:1 saturated NaHCO<sub>3</sub> solution / saturated Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> solution (10 mL). The phases were separated and the aqueous extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 10 mL). Organic phases were combined, washed with brine (40 mL), and dried with Na<sub>2</sub>SO<sub>4</sub> before concentrating *in vacuo*. The residue was purified by column chromatography (SiO<sub>2</sub>, gradient elution from 70-100% EtOAc in hexanes), yielding the product as a colourless oil, a 2:1 mixture of keto-enol tautomers, R<sub>f</sub> = 0.25 in EtOAc, (0.042 g, 56%). [α]<sub>D</sub><sup>25</sup> = -8.0 (c 0.14, CH<sub>2</sub>Cl<sub>2</sub>); IR (CH<sub>2</sub>Cl<sub>2</sub>, cast) 3296, 3078, 2968, 2933, 1720, 1657, 1613, 1551 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 5.98 (br, 1H, H7+H16), 4.45 (s, 0.3H, H13) 3.73 (s, 1.4H, H4), 3.49 – 3.44 (m, 2H, H6+H15), 3.10 – 3.07 (m, 2H, H5+H14), 2.59 – 2.53 (m, 0.7H, H3), 2.14 – 2.07 (m, 0.3H, H12), 1.97 (s, 3H, H8+H17), 1.75 – 1.59 (m, 2H, H2+H11+H3+H12), 1.49 – 1.39 (m, 1H, H2'+H11'), 1.14 – 1.10 (m, 3H, H9), 0.91 – 0.88 (m, 3H, H1+H10). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 206.0, 194.5, 192.6, 181.4, 170.5, 170.3, 98.3, 55.5, 48.6, 41.1, 40.0, 39.4, 29.3, 27.9, 27.2, 25.6, 23.2, 23.1, 17.6, 15.4, 11.8, 11.5. HRMS (ESI) Calcd for C<sub>11</sub>H<sub>19</sub>NO<sub>3</sub>SNa [M+Na]<sup>+</sup> 268.0978, found 268.0977.

### 3-Hydroxy-2,4*R*-dimethyl-1-(2-thioxothiazolidin-3-yl)hexan-1-one



This new compound was synthesized using an established procedure.<sup>55</sup> To a stirred solution of *N*-propionyl-thiazolidine-2-thione (0.25 g, 1.4 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (25 mL) was added TiCl<sub>4</sub> (1.0 M solution in CH<sub>2</sub>Cl<sub>2</sub>, 1.4 mL, 1.4 mmol) at 0 °C under argon. The reaction was stirred for 5 min to dissolve, then cooled to -78 °C. Diisopropylethylamine (0.29 mL, 1.7 mmol) was added and the reaction stirred for 1 hr. A solution of (*R*)-2-methylbutanal (0.11 g, 1.3 mmol) in CH<sub>2</sub>Cl<sub>2</sub> was added dropwise and the reaction stirred for 15 min at -78 °C, then 2 hr at rt. The reaction was quenched by addition of half-saturated NH<sub>4</sub>Cl solution (25 mL). Layers were separated and the aqueous extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 25 mL). The combined organic phases were washed with brine (100 mL), dried with MgSO<sub>4</sub>, concentrated *in vacuo*, and purified by column chromatography (SiO<sub>2</sub>, 30% EtOAc in hexanes), yielding the product as a yellow oil, a 70:30 mixture of diastereomers, R<sub>f</sub> = 0.20 in 30% EtOAc in hexanes, (0.21 g, 57%). [α]<sub>D</sub><sup>25</sup> = -2.9 (*c* 1.0, CHCl<sub>3</sub>); IR (CHCl<sub>3</sub>, cast) 3490, 2963, 2933, 2875, 1697, 1366, 1279, 1155, 1060 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 4.93 – 4.80 (m, 1H, H4), 4.59 – 4.52 (m, 2H, H9), 3.80 – 3.66 (m, 1H, H5), 3.33 – 3.27 (m, 2H, H10), 2.74 (br, 0.7H, H6), 2.32 (br, 0.3H, H6'), 1.81 – 1.75 (m, 1H, H2), 1.53 – 1.38 (m, 1H, H2'+H3), 1.19 (d, *J* = 7.0 Hz, 3H, H8), 0.99 – 0.83 (m, 6H, H1+H7). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 201.8, 179.8, 75.1, 56.5, 41.1, 37.2, 28.2, 25.3, 14.8, 11.1, 9.8. HRMS (ESI) Calcd for C<sub>11</sub>H<sub>19</sub>NO<sub>2</sub>S<sub>2</sub>Na [M+Na]<sup>+</sup> 284.0749, found 284.0753.

### ***S*-2-Acetamidoethyl 2,4*R*-dimethyl-3-oxohexanethioate (Tri<sub>1</sub>-SNAC)**

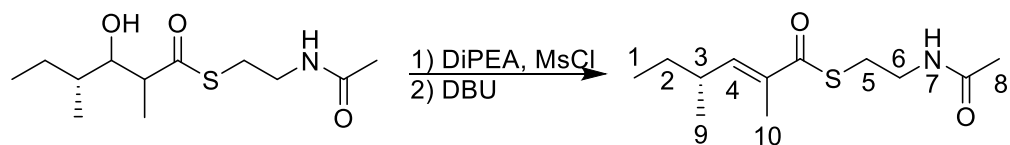


This new compound was synthesized using an established procedure.<sup>55</sup> To a stirred solution of *S*-2-acetamidoethyl 3-hydroxy-2,4*R*-dimethylhexanethioate (0.083 g, 0.34 mmol) in



dry CH<sub>2</sub>Cl<sub>2</sub> (10 mL) at 0 °C was added Dess-Martin periodinane (0.051 g, 0.37 mmol). The reaction was stirred for 30 min at 0 °C, then another 2.5 hr before quenching with 10 mL of 1:1 saturated NaHCO<sub>3</sub> solution / saturated Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> solution. The phases were separated and the aqueous extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 10 mL). Organic phases were combined, washed with brine (40 mL), and dried with Na<sub>2</sub>SO<sub>4</sub> before concentrating *in vacuo*. The residue was purified by column chromatography (SiO<sub>2</sub>, gradient elution from 70-100% EtOAc in hexanes), yielding the product as a colourless oil, R<sub>f</sub> = 0.30 in EtOAc, (0.053 g, 64%). [α]<sub>D</sub><sup>25</sup> = -5.6 (c 0.11, CH<sub>2</sub>Cl<sub>2</sub>); IR (CHCl<sub>3</sub>, cast) 3297, 3077, 2968, 2922, 1722, 1660, 1600, 1551 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 5.85 (br, 1H, H7), 3.97 – 3.91 (m, 1H, H4), 3.50 – 3.35 (m, 2H, H6), 3.12 – 3.00 (m, 2H, H5), 2.72 – 2.64 (m, 1H, H3), 1.96 (s, 3H, H8), 1.76 – 1.64 (m, 1H, H2), 1.41 – 1.35 (m, 4H, H2'+H10), 1.09 (d, *J* = 7.0 Hz, 1.5H, H9), 1.08 (d, *J* = 7.0 Hz, 1.5H, H9'), 0.87 (t, *J* = 7.0 Hz, 1.5H, H1), 0.86 (t, *J* = 7.0 Hz, 1.5H, H1'). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 208.6, 208.5, 196.8, 196.6, 170.4, 60.1, 59.9, 47.4, 47.3, 39.3, 28.9, 26.1, 25.6, 23.2, 16.4, 16.0, 13.9, 13.8, 11.6. HRMS (ESI) Calcd for C<sub>12</sub>H<sub>21</sub>NO<sub>3</sub>SNa [M+Na]<sup>+</sup> 282.1134, found 282.1133.

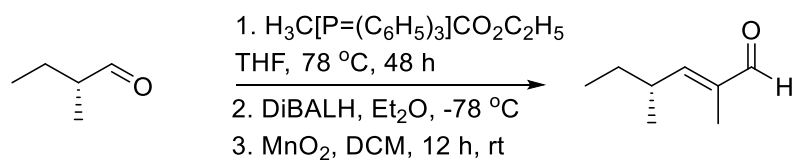
**(*R,E*)-*S*-2-Acetamidoethyl 2,4-dimethylhex-2-enethioate**



This known compound was synthesized using an established procedure.<sup>56</sup> To a stirred solution of *S*-2-acetamidoethyl 3-hydroxy-2,4*R*-dimethylhexanethioate (0.69 g, 2.6 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (20 mL) under argon at 0 °C was added diisopropylethylamine (0.91 mL, 5.2 mmol). After 5 min, methanesulfonyl chloride (0.62 mL, 7.8 mmol) was then added and the reaction stirred for 1 hr. The cooling bath was removed and 1,8-diazabicyclo[5.4.0]undec-7-ene (1.55

mL, 10.4 mmol) was added slowly. Evolution of heat was noted. The reaction was stirred for 16 h at rt before cooling to 0 °C and quenching with 10% citric acid solution (20 mL). The phases were separated and the aqueous extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 20 mL). Organic phases were combined, washed with saturated NaHCO<sub>3</sub> solution (80 mL), brine (80 mL), and dried with MgSO<sub>4</sub> before concentrating *in vacuo*. The residue was purified by column chromatography (SiO<sub>2</sub>, gradient elution from 50-100% EtOAc in hexanes), yielding the product as a yellow oil, R<sub>f</sub> = 0.33 in EtOAc, (0.47 g, 74%). [ $\alpha$ ]<sub>D</sub><sup>25</sup> = -29.7 (c 0.90, CHCl<sub>3</sub>); IR (CHCl<sub>3</sub>, cast) 3289, 3080, 2962, 2929, 1658, 1598, 1551, 1456 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz)  $\delta$  6.53 (dq, *J* = 10.0, 1.4 Hz, 1H, H4), 5.90 (br, 1H, H7), 3.47 – 3.43 (m, 2H, H6), 3.07 – 3.05 (m, 2H, H5), 2.48 – 2.42 (m, 1H, H3), 1.96 (s, 3H, H8), 1.43 (d, *J* = 1.4 Hz, 3H, H10), 1.49 – 1.41 (m, 1H, H2), 1.40 – 1.31 (m, 1H, H2'), 1.02 (d, *J* = 6.5 Hz, 3H, H9), 0.86 (t, *J* = 7.5 Hz, 3H, H1). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz)  $\delta$  194.3, 170.2, 147.6, 135.4, 39.9, 35.1, 29.6, 28.5, 23.3, 19.6, 12.7, 11.9. HRMS (ESI) Calcd for C<sub>12</sub>H<sub>21</sub>NO<sub>2</sub>SNa [M+Na]<sup>+</sup> 266.1185, found 266.1181.

### **(*R,E*)-2,4-dimethylhex-2-enal**



To a flame dried round bottom flask containing 70 mL of dry THF, (*R*)-2-methylbutanal (2.15 g, 25.0 mmol) was dissolved, followed by Ethyl 2-(triphenylphosphoranylidene)propionate (18.1 g, 49.9 mmol). The reaction mixture was then refluxed at 70 °C for 48 hr, then cooled and 70 mL of Et<sub>2</sub>O was added to precipitate out the triphenyl phosphine oxide. The precipitate was then filtered off, and the filtrate was concentrated *in vacuo* at 0 °C. The crude product was then

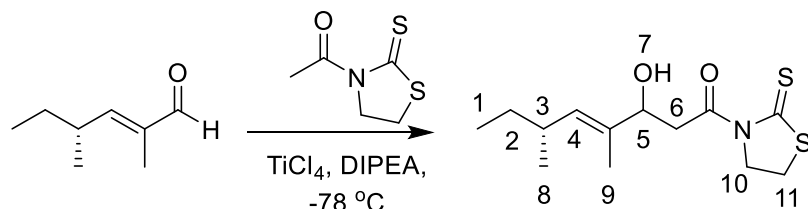
purified on silica column using an eluent system of 20% ether in pentanes to provide (*R,E*)-ethyl 2,4-dimethylhex-2-enoate as a clear oil,  $R_f = 0.32$  in 20:1 pentanes/diethyl ether, (3.47, 82%).

Due to the extremely volatile nature of the compound it was used immediately in the following step. To a flame dried round bottom flask containing 100 mL of dry DCM, (*R,E*)-ethyl 2,4-dimethylhex-2-enoate (3.45 g, 20.3 mmol) was dissolved and cooled to  $-78\text{ }^\circ\text{C}$  while stirring under an argon atmosphere. To the cooled solution, DiBALH (1 M, 41.6 ml, 41.6 mmol) was slowly added while stirring over 20 min. The reaction mixture was then stirred for 2 hr. The reaction mixture was then quenched by slow addition of 20 mL of MeOH, followed by the addition of 40 mL of saturated potassium sodium tartrate then stirred until two distinct layers were observed. The layers were then separated and the aqueous layer was extracted with Et<sub>2</sub>O (3 x 20 mL), then combined and washed with brine (3 x 20 mL), then dried over MgSO<sub>4</sub>. The solvent was removed *in vacuo* at  $0\text{ }^\circ\text{C}$  to provide (*R,E*)-2,4-dimethylhex-2-en-1-ol (2.54 g, 98%), as a single spot by TLC,  $R_f = 0.35$  in 3:1 pentanes/diethyl ether eluent.

Due to the extremely volatile nature of the compound it was used immediately in the following step. To a flame dried round bottom flask containing 100 mL of dry DCM, was added (*R,E*)-2,4-dimethylhex-2-en-1-ol (2.50 g, 19.5 mmol) from the previous step, followed by the slow addition of MnO<sub>2</sub> (5.09 g, 58.5 mmol). The reaction mixture was then sealed under argon and stirred for 12 h at room temperature before it was filtered through a pad of celite, and concentrated *in vacuo* to yield the pure product (*R,E*)-2,4-dimethylhex-2-enal (1.12 g, 63%),  $R_f = 0.6$  (in 20% ethyl acetate in hexanes). Due to the extremely volatile nature of the compound it was used immediately in the following step.  $[\alpha]_D^{25} = -18.4$  (*c* 1.0, CH<sub>2</sub>Cl<sub>2</sub>); IR (CH<sub>2</sub>Cl<sub>2</sub>, cast) 2964, 2930, 2876, 1689, 1644, 1388 cm<sup>-1</sup>; <sup>1</sup>H NMR (CD<sub>2</sub>Cl<sub>2</sub>, 500 MHz)  $\delta$  9.38 (s, 1H, H5), 6.25 (dq, *J* = 10.0, 1.5 Hz, 1H, H4), 2.68 – 2.57 (m, 1H, H3), 1.72 (d, *J* = 1.5 Hz, 3H, H7), 1.54

– 1.46 (m, 1H, H2), 1.43 – 1.34 (m, 1H, H2'), 1.05 (d,  $J = 6.5$  Hz, 3H, H6), 0.88 (t,  $J = 7.5$  Hz, 3H, H1).  $^{13}\text{C}$  NMR ( $\text{CD}_2\text{Cl}_2$ , 125 MHz)  $\delta$  195.7, 160.6, 138.5, 35.5, 29.9, 19.6, 12.0, 9.5.

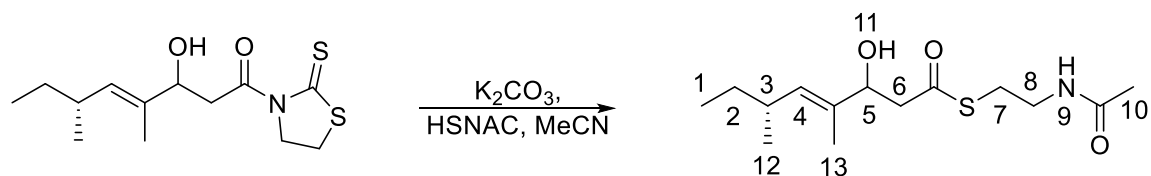
**(6*R*,*E*)-3-Hydroxy-4,6-dimethyl-1-(2-thioxothiazolidin-3-yl)oct-4-en-1-one**



This new compound was synthesized using a modified procedure.<sup>24</sup> To a stirred solution of *N*-acetyl-thiazolidine-2-thione (0.014 g, 0.09 mmol) in dry  $\text{CH}_2\text{Cl}_2$  (5 mL) was added  $\text{TiCl}_4$  (1.0 M solution in  $\text{CH}_2\text{Cl}_2$ , 0.10 mL, 0.10 mmol) at  $0\text{ }^\circ\text{C}$  under argon. The reaction mixture was stirred for 5 min to dissolve, then cooled to  $-78\text{ }^\circ\text{C}$ . Diisopropylethylamine (0.018 mL, 0.10 mmol) was added and the reaction stirred for 30 min. A solution of (*R*,*E*)-2,4-dimethylhex-2-enal (0.010 g, 0.08 mmol) in  $\text{CH}_2\text{Cl}_2$  was added dropwise and the reaction stirred for 15 min at  $-78\text{ }^\circ\text{C}$ , then 1.5 hr at rt. The reaction was quenched by addition of 10% citric acid solution (5 mL). Layers were separated and the aqueous extracted with  $\text{CH}_2\text{Cl}_2$  (3 x 4 mL). The combined organic phases were dried with  $\text{MgSO}_4$ , concentrated *in vacuo*, and purified by column chromatography ( $\text{SiO}_2$ , 30% EtOAc in hexanes), yielding the product as a yellow oil,  $R_f = 0.16$  in 30% EtOAc in hexanes, (0.013 g, 57%).  $[\alpha]_D^{25} = -12.2$  ( $c$  0.27,  $\text{CHCl}_3$ ); IR ( $\text{CHCl}_3$ , cast) 3467, 2957, 2924, 1700, 1460  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz)  $\delta$  5.27 – 5.23 (m, 1H, H4), 4.62 – 4.52 (m, 3H, H10+H5), 3.58 – 3.52 (m, 1H, H6), 3.48 – 3.44 (m, 3H, H6'), 3.34 – 3.26 (m, 2H, H11), 2.67 (br, 1H, H7), 2.32 – 2.24 (m, 1H, H3), 1.67 (s, 3H, H9), 1.39 – 1.29 (m, 1H, H2), 1.30 – 1.17 (m, 1H, H2'), 0.95 – 0.91 (m, 3H, H8), 0.86 – 0.81 (m, 3H, H1).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz)  $\delta$

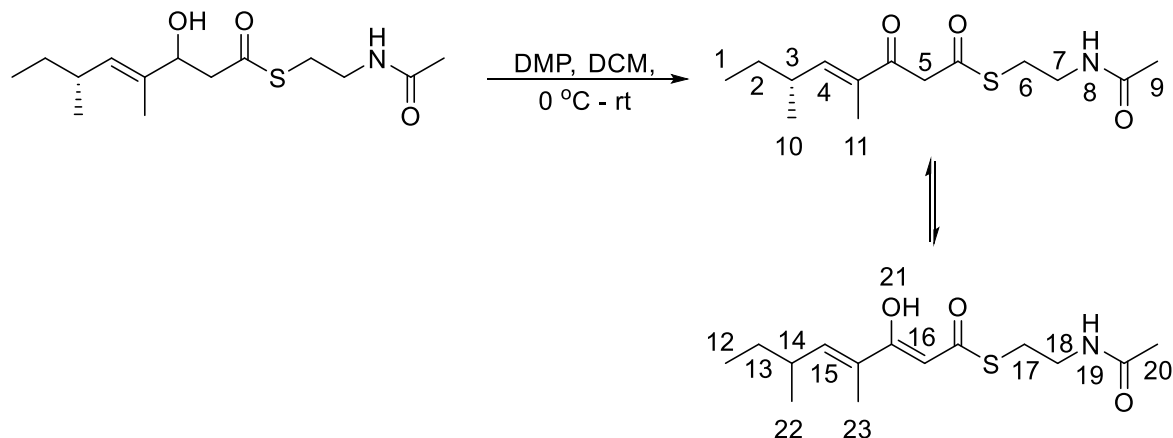
201.9, 174.0, 134.1, 133.6, 55.8, 44.6, 33.7, 30.3, 28.6, 20.6, 12.5, 12.0. HRMS (ESI) Calcd for  $C_{13}H_{21}NO_2S_2Na$   $[M+Na]^+$  310.0906, found 310.0910.

**(6*R*,*E*)-*S*-2-Acetamidoethyl 3-hydroxy-4,6-dimethyloct-4-enethioate**



This new compound was synthesized using a modified procedure.<sup>24</sup> To a stirred solution of (6*R*,*E*)-3-hydroxy-4,6-dimethyl-1-(2-thioxothiazolidin-3-yl)oct-4-en-1-one (0.0034g, 0.012 mmol) in dry acetonitrile (1 mL) was added  $K_2CO_3$  (0.049 g, 0.035 mmol) followed by *N*-acetylcysteamine (0.0015 mL, 0.014 mmol). The reaction was stirred for 30 min before removing the solvent *in vacuo*. The residue was suspended in  $CH_2Cl_2$ , washed with brine, and dried with  $MgSO_4$  before concentrating *in vacuo*. The residue was purified by column chromatography ( $SiO_2$ , gradient elution from 50-100% EtOAc in hexanes), yielding the product as a colourless oil,  $R_f = 0.30$  in EtOAc, (0.0022 g, 65%).  $[\alpha]_D^{25} = -14.4$  ( $c$  0.11,  $CHCl_3$ ); IR ( $CHCl_3$ , cast) 3295, 3091, 2959, 2927, 2872, 1690, 1658, 1553  $cm^{-1}$ ;  $^1H$  NMR ( $CDCl_3$ , 500 MHz)  $\delta$  5.80 (br, 1H, H9), 5.26 – 5.22 (m, 1H, H4), 4.50 – 4.46 (m, 1H, H5), 3.50 – 3.40 (m, 2H, H8), 3.10 – 3.00 (m, 2H, H7), 2.88 – 2.81 (m, 1H, H6), 2.79 – 2.72 (m, 1H, H9'), 2.35 (br, 1H, H11), 2.32 – 2.22 (m, 1H, H3), 1.97 (s, 3H, H10), 1.64 (d,  $J = 1.0$  Hz, 3H, H13), 1.38 – 1.29 (m, 1H, H2), 1.28 – 1.16 (m, 1H, H2'), 0.94 – 0.90 (m, 3H, H8), 0.85 – 0.79 (m, 3H, H1).  $^{13}C$  NMR ( $CDCl_3$ , 125 MHz)  $\delta$  199.0, 170.4, 133.9, 133.5, 77.3, 49.8, 39.5, 33.8, 30.2, 28.9, 23.3, 20.5, 12.2, 11.9. HRMS (ESI) Calcd for  $C_{14}H_{25}NO_3SNa$   $[M+Na]^+$  310.1447, found 310.1447.

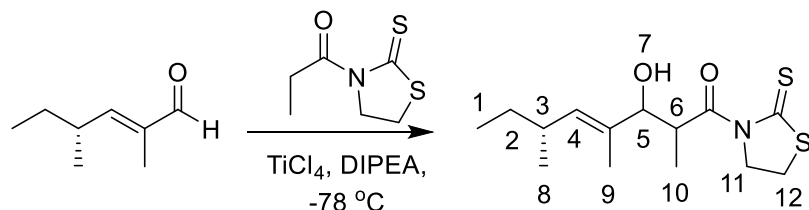
**(*R,E*)-*S*-2-Acetamidoethyl 4,6-dimethyl-3-oxooct-4-enethioate (0-SNAC)**



This new compound was synthesized using an established procedure.<sup>24</sup> To a stirred solution of (*6R,E*)-*S*-2-acetamidoethyl 3-hydroxy-4,6-dimethyloct-4-enethioate (0.0022 g, 0.0077 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (1 mL) at 0°C was added Dess-Martin periodinane (0.0036 g, 0.0085 mmol). The reaction was stirred for 16 h before quenching with 1:1 saturated NaHCO<sub>3</sub> solution / saturated Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> solution (2 mL). The phases were separated and the aqueous extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 2 mL). Organic phases were combined and dried with MgSO<sub>4</sub> before concentrating *in vacuo*. The residue was purified by column chromatography (SiO<sub>2</sub>, 50% EtOAc in hexanes), yielding the product as a colourless oil, a 2:1 mixture of keto-enol tautomers, R<sub>f</sub> = 0.41 in EtOAc, (0.0019 g, 87%). [α]<sub>D</sub><sup>25</sup> = -23.2 (*c* 0.062, CHCl<sub>3</sub>); IR (CHCl<sub>3</sub>, cast) 3303, 3076, 2962, 2928, 1660, 1582, 1556 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 12.76 (br, 0.3H, H21), 6.44 – 6.42 (m, 0.3H, H15), 6.40 – 6.37 (m, 0.7H, H15), 5.96 (br, 1H, H10+H23), 5.64 (s, 0.3H, H16), 3.99, 3.96 (ABq, *J* = 15.0 Hz, 1.4H, H5), 3.50 – 3.44 (m, 2H, H7+H18), 3.12 – 3.08 (m, 2H, H6+H17), 2.58 – 2.42 (m, 1H, H3+H14), 1.98 (s, 2.1H, H9), 1.97 (s, 0.9H, H20), 1.80 (d, *J* = 1.4 Hz, 2.1H, H11), 1.78 (d, *J* = 1.4 Hz, 0.9H, H23), 1.51 – 1.42 (m, 1H, H2+H13), 1.28 – 1.19 (m, 1H, H2'+H13'), 1.04 (d, *J* = 7.0 Hz, 2.1H, H10), 1.01 (d, *J* = 7.0 Hz, 0.9H, H22), 0.89 – 0.84 (m, 3H, H1+H12). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 194.6, 193.7, 193.3, 170.9, 170.5, 170.3, 151.7,

144.8, 135.8, 126.9, 96.9), 52.6, 40.0, 39.3, 35.5, 35.0, 29.8, 29.6, 29.2, 28.0, 23.3, 19.9, 19.5, 12.2, 12.0, 11.9, 11.5, 11.0. HRMS (ESI) Calcd for C<sub>14</sub>H<sub>23</sub>NO<sub>3</sub>SNa [M+Na]<sup>+</sup> 308.1291, found 308.1289.

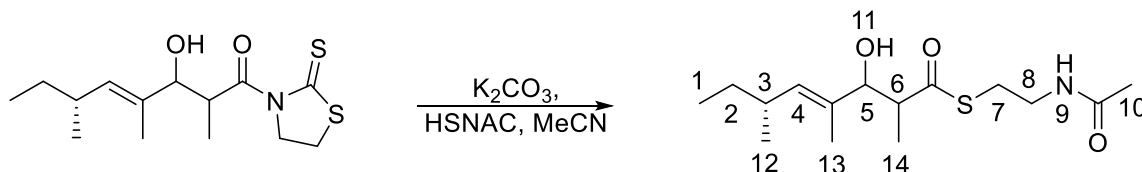
**(6*R*,*E*)-3-hydroxy-2,4,6-trimethyl-1-(2-thioxothiazolidin-3-yl)oct-4-en-1-one**



This new compound was prepared following a literature protocol.<sup>24</sup> To a stirred solution of *N*-propionyl-thiazolidine-2-thione (1.56 g, 8.90 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (100 mL) was added TiCl<sub>4</sub> (1.0 M solution in CH<sub>2</sub>Cl<sub>2</sub>, 9.79 mL, 9.79 mmol) at 0 °C under argon. The reaction was stirred for 10 min, then cooled to -78 °C and stirred for 15 min. Diisopropylethylamine (1.35 mL, 10.7 mmol) was added and the reaction stirred for 30 min. A solution of (*R*,*E*)-2,4-dimethylhex-2-enal (1.12 g, 8.90 mmol) in CH<sub>2</sub>Cl<sub>2</sub> was added dropwise and the reaction stirred for 1.5 hr at -78 °C, then 1 hr at rt. The reaction was quenched by addition of 10% citric acid solution (20 mL). Layers were separated and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 20 mL). The combined organic phases were then dried with MgSO<sub>4</sub> and concentrated *in vacuo*. The crude residue was purified by column chromatography (SiO<sub>2</sub>, 20% EtOAc in hexanes), yielding the product as a yellow oil and a 50:50 mixture of diastereomers, R<sub>f</sub> = 0.19 in 30% EtOAc in hexanes, (2.07 g, 77%). IR (CHCl<sub>3</sub>, cast) 3496, 2959, 2924, 1701, 1455 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 5.29 – 5.28 (m, 1H, H<sub>4</sub>), 4.80 – 4.77 (m, 1H, H<sub>6</sub>), 4.57 – 4.47 (m, 2H, H<sub>11</sub>), 4.38 – 4.36 (m, 1H, H<sub>5</sub>), 3.31 – 3.22 (m, 2H, H<sub>12</sub>), 2.49 – 2.42 (m, 1H, H<sub>7</sub>), 2.30 – 2.28 (m, 1H, H<sub>3</sub>), 1.60 (s, 3H, H<sub>9</sub>), 1.37 – 1.21 (m, 2H, H<sub>2</sub>), 1.16 (d, *J* = 7.0 Hz, 3H, H<sub>10</sub>), 0.93 –

0.92 (m, 3H, H8), 0.84 – 0.81 (m, 3H, H1). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 201.8, 178.9, 133.4, 132.5, 56.6, 42.4, 33.9, 30.4, 28.4, 20.8, 13.5, 12.2, 11.1. HRMS (ESI) Calcd for C<sub>14</sub>H<sub>23</sub>NO<sub>2</sub>·S<sub>2</sub>Na [M+Na]<sup>+</sup> 324.1062, found 324.1062.

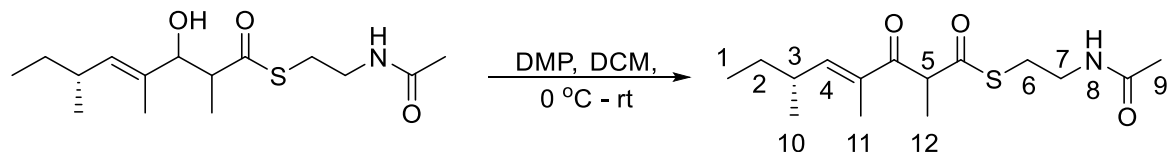
**(6*R*,*E*)-S-2-acetamidoethyl 3-hydroxy-2,4,6-trimethyloct-4-enethioate**



This new compound was prepared following a literature protocol.<sup>24</sup> To a stirred solution of (6*R*,*E*)-3-hydroxy-2,4,6-trimethyl-1-(2-thioxothiazolidin-3-yl)oct-4-en-1-one (2.00 g, 6.97 mmol) in dry acetonitrile (30 mL) at rt was added K<sub>2</sub>CO<sub>3</sub> (2.41 g, 17.4 mmol) followed by *N*-acetylcysteamine (2.49 g, 20.9 mmol). The reaction was stirred for 20 min before the solvent was removed *in vacuo*. The residue was taken up in DCM (20 mL), washed with brine (2 x 10 mL), then dried with MgSO<sub>4</sub> before concentrating *in vacuo*. The residue was purified by column chromatography (SiO<sub>2</sub>, gradient elution from 50-100% EtOAc in hexanes), yielding the product as a colourless oil, R<sub>f</sub> = 0.31 in EtOAc, (2.10 g, 87%). IR (CHCl<sub>3</sub>, cast) 3308, 3088, 2960, 2930, 1658, 1551 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 5.79 (br s, 1H, H9), 5.26 – 5.23 (m, 1H, H4), 4.31 – 4.29 (m, 1H, H5), 3.48 – 3.40 (m, 2H, H8), 3.06 – 2.96 (m, 2H, H7), 2.91 – 2.85 (m, 1H, H6), 2.30 – 2.26 (m, 1H, H3), 2.19 – 2.11 (m, 1H, H11), 1.96 (s, 3H, H10), 1.61 (s, 3H, H13), 1.39 – 1.17 (m, 5H, H2, 14), 0.93 – 0.90 (m, 3H, H12), 0.85 – 0.79 (m, 3H, H1). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 203.7, 170.4, 134.4, 132.3, 76.9, 51.8, 39.7, 34.0, 30.3, 28.7, 23.4, 20.7, 13.3, 12.2, 11.6. HRMS (ESI) Calcd for C<sub>15</sub>H<sub>27</sub>NO<sub>3</sub>SNa [M+Na]<sup>+</sup> 324.1604, found 324.1607.

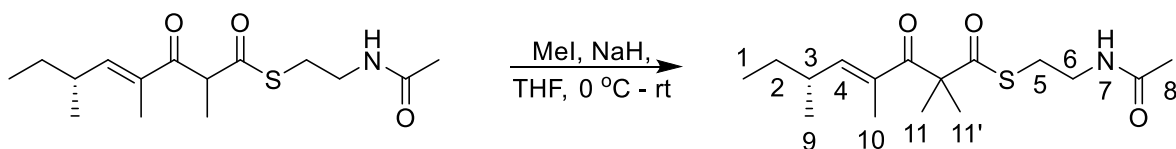
**(6*R*,*E*)-S-2-acetamidoethyl 2,4,6-trimethyl-3-oxooct-4-enethioate (1-SNAC)**





This new compound was prepared following a literature protocol.<sup>24</sup> To a stirred solution of *(6R,E)*-*S*-2-acetamidoethyl 3-hydroxy-2,4,6-trimethyloct-4-enethioate (1.60 g, 5.31 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (50 mL) at 0 °C was added Dess-Martin periodinane (2.25 g, 5.31 mmol). The reaction mixture was stirred for 30 min at 0 °C, then a further 3 hr at rt before quenching with 1:1 saturated NaHCO<sub>3</sub> solution/saturated Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> solution (20 mL). The phases were separated and the aqueous layer extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 15 mL). Organic phases were combined, washed with brine (2 x 15 mL), then dried with Na<sub>2</sub>SO<sub>4</sub> before concentrating *in vacuo*. The residue was purified by column chromatography (SiO<sub>2</sub>, gradient elution from 30% EtOAc in hexanes), yielding the product as a colourless oil, *R*<sub>f</sub> = 0.39 in EtOAc, (1.59 g, 85%). IR (CH<sub>3</sub>OH, cast) 3290, 2962, 2932, 1689, 1660, 1547 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz) δ 6.44 (d, *J* = 9.1 Hz 1H, H4), 5.99 (br, 1H, H8), 4.45 – 4.43 (m, 1H, H5), 3.46 – 3.32 (m, 2H, H7), 3.07 – 2.98 (m, 2H, H6), 2.50 – 2.49 (m, 1H, H3), 1.95 (s, 3H, H9), 1.79 (s, 3H, H11), 1.47 – 1.32 (m, 5H, H2,12), 1.02 (d, *J* = 7.0 Hz, 3H, H10), 0.87 – 0.84 (m, 3H, H1). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 197.4, 196.6, 170.5, 150.7, 135.4, 54.9, 39.5, 35.7, 29.8, 28.8, 23.2, 19.7, 15.0, 12.1, 11.9. HRMS (ESI) Calcd for C<sub>15</sub>H<sub>25</sub>NO<sub>3</sub>SNa [M+Na]<sup>+</sup> 322.1447, found 322.1445.

***(R,E)*-*S*-2-acetamidoethyl 2,2,4,6-tetramethyl-3-oxooct-4-enethioate (2-SNAC)**



This new compound was synthesized using a modified protocol.<sup>57</sup> Sodium hydride (24.0 mg, 0.60 mmol, 60% suspension in oil) was weighed into a flame dried round bottom flask and washed with anhydrous Et<sub>2</sub>O (3 x 2 mL) to remove the bulk of the oil. A solution of (6*R*,*E*)-*S*-2-acetamidoethyl 2,4,6-trimethyl-3-oxooct-4-enethioate (100 mg, 0.334 mmol) in dry THF (15 mL) was added to the vial. The contents were cooled to 0 °C and iodomethane (0.10 mL, 1.7 mmol) was added. The reaction was stirred under argon for 4 hr before quenching with 10% citric acid solution (5 mL). The mixture was extracted with EtOAc (3 x 10 mL). Organic phases were combined, dried with MgSO<sub>4</sub>, then concentrated *in vacuo*. The residue was purified by flash chromatography (SiO<sub>2</sub>, 50% EtOAc in hexanes) to provide a pale-yellow oil, R<sub>f</sub> = 0.43 in EtOAc, (91.1 mg, 87%). [ $\alpha$ ]<sub>D</sub><sup>25</sup> = -14.4 (*c* 1.0, CHCl<sub>3</sub>); IR (CHCl<sub>3</sub>, cast) 3293, 2962, 2932, 1679, 1660 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz)  $\delta$  6.09 (d, *J* = 9.8, 1.4 Hz, 1H, H4), 5.79 (br s, 1H, H7), 3.44 – 3.40 (m, 2H, H6), 3.07 – 2.96 (m, 2H, H5), 2.47 – 2.43 (m, 1H, H3), 1.96 (s, 3H, H8), 1.78 (s, 3H, H10), 1.47 – 1.39 (m, 7H, H2,11+H11'), 1.30 – 1.23 (m, 1H, H2'), 0.97 (m, 3H, H9), 0.84 (t, *J* = 7.0 Hz, 3H, H1). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz)  $\delta$  203.1, 199.6, 170.3, 149.5, 133.0, 61.2, 39.7, 35.5, 29.8, 28.7, 25.5, 24.5, 23.4, 19.7, 12.9, 12.0. HRMS (ESI) Calcd for C<sub>16</sub>H<sub>27</sub>NO<sub>3</sub>.SNa [M+Na]<sup>+</sup> 336.1604, found 336.1605.

## Construction of Expression Plasmids

### *Construction of pLFH18, pLFH19, and pLFH29 Expression Plasmid*

Relevant DNA primers and plasmids are listed in **Table S3.2-S3.3**. The following 2 $\mu$ -based yeast expression plasmids with auxotrophic markers were used for expression: pXW55.<sup>58</sup> After codon optimization<sup>59</sup>, the coding domain sequence of the Tv6-931 gene was synthesized by Gen9. The codon-optimized gene was then inserted into the digested pXW55<sup>58</sup> plasmid (SpeI and PmlI) using *in vivo* homologous recombination to create the pTMC-1-106.<sup>60</sup>

The first-generation plasmid, pTMC-1-106, contained the Tv6-931 gene flanked by an *N*-terminal FLAG tag and *C*-terminal 6xHis Tag.<sup>17</sup> Using the pTMC-1-106, various relevant PCR products were synthesized using the following primers:

pXW55-N-Term-Ura-F, pXW55-N-8xHis-R pXW55-C-Term-F, pXW55-C-Term-Ura-R, Tv6-931-N-His-HRPKS-P1-F, Tv6-931-HRPKS-P1-R, Tv6-931-HRPKS-P2-F, Tv6-931-HRPKS-P2-R, Tv6-931-HRPKS-P3-F, Tv6-931-HRPKS-cAT-Native-R, Tv6-931-HRPKS-standalone-R, Tv6-931-ACP\_cAT-XW55-F.

The resulting PCR products were used to assemble the second-generation plasmids, pLFH18, pLFH19, and pLFH29 by *in vivo* yeast homologous recombination using Frozen-EZ Yeast Transformation II Kit<sup>TM</sup> (Zymo Research). These expression plasmids include the relevant coding DNA sequence with an additional *N*-terminal 8xHis for purification.

### *Construction of YEplovB-6His-DH<sup>0</sup> Expression Plasmid*

The YEplovB-6His plasmid was cloned according to a known protocol.<sup>61</sup>

The YEplovB-6His-DH<sup>0</sup> plasmid was constructed from the YEplovB-6His plasmid according to a known protocol.<sup>24</sup> Briefly, the H985A point mutation was achieved through quick change mutagenesis incorporated within PCR primers.

### **Chiral Resolution of $\epsilon$ -methyl on Tv6-931 Adduct 2**

The chiral  $\epsilon$ -methyl on **2** was determined by synthesizing both the **R-2-Tris** and **S-2-Tris** standard and comparing the separation against the authentic **2-Tris** sample using analytic chiral HPLC column. In addition, optical rotation measurements were used to support the *R* stereochemical assignment.

#### *Enzymatic synthesis of **R-2-Tris** and **S-2-Tris** standards*

Enzymatic esterification reactions consisted of 15  $\mu$ M of *apo*-Tv6-931 enzyme incubated with 0.5 mM of **R-2-SNAC** or **S-2-SNAC** in 10 mM of Tris buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. Afterwards, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by analytic chiral HPLC.

#### *Enzymatic biosynthesis of authentic **2-Tris***

Enzymatic biosynthesis of authentic **2-Tris** reactions consisted of 15  $\mu$ M of Tv6-931 enzyme incubated with 4 mM NADPH, 2 mM SAM, 2 mM malonyl-coenzyme A in 10 mM of Tris buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. Afterwards, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes. The supernatant was separated and then analyzed by analytic chiral HPLC.

#### *Optical Rotation Measurement of **2-THME***

Optical rotation data was collected on a Jasco P-1020 system. For compound **2-THME**,  $[\alpha]_D^{25} = -11^\circ$  ( $c=0.05$ , MeOH), For compound **S-2-SNAC**,  $[\alpha]_D^{25} = +19^\circ$  ( $c=0.05$ , MeOH),  $[\alpha]_D^{25} = +29^\circ$  ( $c=0.16$ , CHCl<sub>3</sub>). For compound **R-2-SNAC**,  $[\alpha]_D^{25} = -14^\circ$  ( $c=1.0$ , CHCl<sub>3</sub>).

## ***In vivo* Expression Experiments**

### *Protein Expression and Purification in Saccharomyces cerevisiae*

For protein expression under the alcohol dehydrogenase (*ADH2*) promoter, the yeast expression plasmid of interest is first transformed into *S. cerevisiae* BJ5464-NpgA or BJ5464-*apo* cells using the Frozen-EZ Yeast Transformation II Kit™ (Zymo Research) (**Table S3.3-S3.4**). A seed culture of the transformed *S. cerevisiae* strain was grown in 45 mL of synthetic dropout complete medium without uracil (SDC-ura). The seed culture was shaken for two days at 28°C, 250 rpm. This seed culture was then used to inoculate 3 L of YPD 1.5% media (yeast extract (10 g/L), peptone (20 g/L) and dextrose (15 g/L)) and cultured for 3 days at 28°C, 250 rpm. The cell pellet was harvested by centrifugation (5000 rpm, 10 mins), and it was then suspended in 250 mL of a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 150 mM NaCl, and 10 mM imidazole at pH 8.0.

The suspended yeast cells were lysed using sonication and the cellular debris was removed using high-speed centrifugation (17,000 rpm, 1 hr) and passed through a 0.2 μM filter. The enzyme of interest was then purified from the supernatant using Ni-NTA agarose affinity chromatography to near homogeneity. The purified protein was concentrated, exchanged into a buffer containing 50 mM NaH<sub>2</sub>PO<sub>4</sub> and 50 mM NaCl at pH 8.0, aliquoted and then flash frozen in liquid nitrogen.

*Biosynthesis of Tv6-931 polyketides (1 and 2) in S.cer. BJ5464-npgA*

For expression of the Tv6-931 polyketides, **1** and **2**, under the alcohol dehydrogenase (*ADH2*) promoter, the yeast expression plasmid of interest, pLFH19, is first transformed into *S. cerevisiae* BJ5464-NpgA cells using the Frozen-EZ Yeast Transformation II Kit™ (Zymo Research). A seed culture of *S.cer*-pLFH19 was grown in 30 mL of synthetic dropout complete medium without uracil (SDC-ura). The seed culture was shaken for two days at 28°C, 250 rpm. This seed culture was then used to inoculate 2 x 1 L of Yeast Peptone 1.5% Dextrose (YPD 1.5%) medium and cultured at 28°C, 250 rpm. After 1 day, 10 g of THME or PE (1% w/v) was added to each 1 L culture. The cultures were shaken for an additional 2 days at 28°C, 250 rpm.

After centrifugation to remove the yeast cells, the supernatant was extracted with 6 x 500 mL of ethyl acetate. The combined organic layer was washed with 500 mL of brine (saturated NaCl) and then dried over Na<sub>2</sub>SO<sub>4</sub>. The crude extract was then evaporated to dryness *in vacuo* to produce a dark, brown oil. The crude extract was initially purified by flash column chromatography in silica. For **1-THME** and **2-THME**, the mobile phase was 1:1 ethyl acetate/hexanes (v/v). For **1-PE** and **2-PE**, the mobile phase was ethyl acetate. The product was then further purified by semi-preparative HPLC. The elution method used was a linear gradient of 40-70% (v/v) CH<sub>3</sub>CN/H<sub>2</sub>O in 15 minutes, followed by 95% CH<sub>3</sub>CN/H<sub>2</sub>O for 3 minutes. The purity was confirmed by HPLC-MS and the structure was determined by NMR.

## ***In vitro* Enzymatic Experiments**

### *Enzymatic Methylation Kinetic Assay with SNAC Compounds*

#### *a) Time course*

Enzymatic methylation time course reactions consisted of 15 μM of *apo*-Tv6-931 or *apo*-Tv6-931 Δ*cAT* enzyme incubated with 2.5 mM SAM and 0.5 mM of the SNAC

of interest in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated at 30°C. At different time points, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate.

*b) Saturation curve*

Enzymatic methylation saturation curve reactions consisted of 15 µM of *apo*-Tv6-931 or *apo*-Tv6-931 ΔcAT enzyme incubated with 2.5 mM SAM and 0.2-0.8 mM of the SNAC of interest in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated at 30°C. At a specific time point, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate. Using the Origin software, the  $k_{cat}/K_M$  value was estimated from the slope of the turnover frequency against the SNAC substrate concentration.

*Enzymatic Esterification Kinetic Assay with N-acetyl cysteamine (SNAC) Compounds*

*a) Time course*

Enzymatic esterification time course reactions consisted of 15 µM of *apo*-Tv6-931 or Tv6-931 *apo*-ACP-cAT enzyme incubated with 2.5 mM PE and 0.5 mM of the SNAC of interest in 10 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated at 30°C. At different time points, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The

supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate.

*b) Saturation curve*

Enzymatic esterification saturation curve reactions consisted of 15  $\mu\text{M}$  of *apo*-Tv6-931 and Tv6-931 *apo*-ACP-cAT enzyme incubated with 2.5 mM PE and 0.2-0.8 mM of the SNAC of interest in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer (pH 8.0). The reaction mixture was incubated at 30°C. At a specific time point, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate. Using the Origin software, the  $k_{\text{cat}}/K_{\text{M}}$  value was estimated from the slope of the turnover frequency against the SNAC substrate concentration.

*Enzymatic Recapture (1-Pante synthesis) Kinetic Assay with 1-PE*

*a) Time course*

Enzymatic recapture (**1-Pante** synthesis) time course reactions consisted of 15  $\mu\text{M}$  of *apo*-Tv6-931 or Tv6-931 *apo*-ACP-cAT enzyme incubated with 2.5 mM pantetheine and 0.5 mM **1-PE** in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer (pH 8.0). The reaction mixture was incubated at 30°C. At different time points, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate.

*b) Saturation curve*

Enzymatic recapture (**1-Pante** synthesis) saturation curve reactions consisted of 15



$\mu\text{M}$  of *apo*-Tv6-931 and Tv6-931 *apo*-ACP-cAT enzyme incubated with 2.5 mM pantetheine and 0.5 mM **1-PE** in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer (pH 8.0). The reaction mixture was incubated at 30°C. At a specific time point, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate. Using the Origin software, the  $k_{\text{cat}}/K_M$  value was estimated from the slope of the turnover frequency against the polyketide substrate concentration.

Note: commercial pantetheine was used in these reactions.

#### *Enzymatic Recapture (Transesterification) Kinetic Assay with 1-THME and 2-THME*

##### *a) Time course*

Enzymatic recapture (transesterification) time course reactions consisted of 15  $\mu\text{M}$  of *holo*-Tv6-931 or Tv6-931 *holo*-ACP-cAT enzyme incubated with 2.5 mM PE and 0.5 mM **1-THME** or **2-THME** in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer (pH 8.0). The reaction mixture was incubated at 30°C. At different time points, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate.

##### *b) Saturation curve*

Enzymatic recapture (transesterification) saturation curve reactions consisted of 15  $\mu\text{M}$  of *holo*-Tv6-931 or Tv6-931 *holo*-ACP-cAT enzyme incubated with 2.5 mM PE and 0.2-0.8 mM **1-THME** or **2-THME** in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer (pH 8.0). The

reaction mixture was incubated at 30°C. At a specific time point, the reaction mixture was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC. Each measurement was performed in triplicate. Using the Origin software, the  $k_{cat}/K_M$  value was estimated from the slope of the turnover frequency against the polyketide substrate concentration.

#### *Chemoenzymatic Functionalization Using holo-ACP-cAT Enzyme of 1-PE and 2-PE*

Chemoenzymatic functionalization reactions consisted of 30  $\mu$ M of the Tv6-931 *holo*-ACP-cAT enzyme incubated with 0.25 mM of **1-PE** or **2-PE**, 10 mM nucleophile (**Table 3.1A**, **Table S3.5A**), 2 mM DTT in 20 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

#### *Enzymatic Functionalization Using apo-ACP-cAT Enzyme of 1-PE and 2-PE*

Enzymatic functionalization reactions consisted of 30  $\mu$ M of the Tv6-931 *apo*-ACP-cAT enzyme incubated with 0.25 mM of **1-PE** or **2-PE**, 10 mM nucleophile (**Table 3.1B**, **Table S3.5B**) in 20 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

Note: Addition of 2 mM DTT to the reaction does not produce noticeable changes in the results.

#### *Enzymatic Transesterification Using apo-ACP-cAT with Supplementation of Pantetheine*

Enzymatic transesterification reaction consisted of 30  $\mu$ M of the Tv6-931 *apo*-ACP-cAT enzyme incubated with 0.25 mM of **1-PE**, 10 mM Tris, 10 mM pantetheine, 2 mM DTT in 20 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

Note: As an alternative to commercial pantetheine, a 100 mM pantetheine solution could be prepared by the reduction of commercial pantethine (disulfide) according to a modified literature protocol.<sup>34</sup>

#### *Chemoenzymatic Transthioesterification Using apo-ACP-cAT with Supplementation of Pantetheine*

Chemoenzymatic transthioesterification reaction consisted of 30  $\mu$ M of the Tv6-931 *apo*-ACP-cAT enzyme incubated with 0.25 mM of **1-PE**, 10 mM pantetheine, 2 mM DTT in 20 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated for 12 hours at 30°C. Afterwards, 2 mM of cysteamine was added to the reaction mixture; the reaction mixture was incubated for an additional 12 hours at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

Control reactions (**Figure S3.14** entries viii and ix) consisted of 0.25 mM of **7**, 2 mM of cysteamine (entry ix) in 20 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction was incubated for 12 hours at 30°C.

Note: As an alternative to commercial pantetheine, a 100 mM pantetheine solution could be prepared by the reduction of commercial pantethine (disulfide) according to a modified literature protocol.<sup>34</sup>

#### *Enzymatic Functionalization Using apo-ACP-cAT Enzyme of 0-PE*

The compound **0-PE** was prepared by an enzymatic esterification reaction. Enzymatic esterification reactions consisted of 15  $\mu$ M of *apo*-Tv6-931 enzyme incubated with 0.5 mM of **0-SNAC**, 5 mM PE in 10 mM of  $\text{NaH}_2\text{PO}_4$  buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. Afterwards, the reaction mixture was quenched with quenched and extracted by 6 x 1 equivalent volumes of ethyl acetate. The combined organic layer then dried over  $\text{Na}_2\text{SO}_4$ . The crude extract was then evaporated to dryness *in vacuo* to produce an off-white solid. The product was then purified by semi-preparative HPLC using a C-18 reverse-phase column. The elution method used was a linear gradient of 40-70% (v/v)  $\text{CH}_3\text{CN}/\text{H}_2\text{O}$  in 15 minutes, followed by 95%  $\text{CH}_3\text{CN}/\text{H}_2\text{O}$  for 3 minutes. The purity was confirmed by HPLC-MS and the structure was confirmed by  $^1\text{H-NMR}$ .

Enzymatic functionalization reactions consisted of 30  $\mu$ M of the Tv6-931 *holo*-ACP-cAT enzyme incubated with 0.25 mM of **0-PE**, 10 mM nucleophile (**Figure S3.15**) in 20 mM of  $\text{NaH}_2\text{PO}_4$  buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

#### *Chemoenzymatic Synthesis of 3 and 4 Using ACP-cAT and LovB Enzymes*

Chemoenzymatic synthesis of **3** and **4** reactions consisted of 30  $\mu$ M Tv6-931 *apo*-ACP-cAT enzyme, 10 mM LovB-DH<sup>0</sup> or native LovB incubated with 0.5 mM **1-PE**, 10 mM

pantetheine, 2 mM NADPH, 2 mM DTT in 30 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

Alternatively, chemoenzymatic synthesis reactions with **1-Pante** consist of 10 mM LovB-DH<sup>0</sup> or native LovB incubated with 0.5 mM **1-Pante**, 2 mM NADPH in 20 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

Note: As an alternative to commercial pantetheine, a 100 mM pantetheine solution could be prepared by the reduction of commercial pantethine (disulfide) according to a modified literature protocol.<sup>34</sup>

#### *Chemoenzymatic Synthesis of 5 Using ACP-cAT and LovB Enzymes*

Chemoenzymatic synthesis of **5** reactions consisted of 30 μM Tv6-931 *apo*-ACP-cAT enzyme, 10 mM LovB-DH<sup>0</sup> or native LovB incubated with 0.5 mM **2-PE**, 10 mM pantetheine, 2 mM NADPH, 2 mM DTT in 30 mM of NaH<sub>2</sub>PO<sub>4</sub> buffer (pH 8.0). The reaction mixture was incubated overnight at 30°C. The reaction was quenched with six volumes of acetonitrile and centrifuged at 14,000 rpm for 8 minutes at 20°C. The supernatant was separated and then analyzed by HPLC-MS.

Note: As an alternative to commercial pantetheine, a 100 mM pantetheine solution could be prepared by the reduction of commercial pantethine (disulfide) according to a modified literature protocol.<sup>34</sup>

## Section 3.12 References

1. Pietrocola, F.; Galluzzi, L.; Bravo-San Pedro, J. M.; Madeo, F.; Kroemer, G., Acetyl coenzyme A: a central metabolite and second messenger. *Cell Metab* **2015**, *21* (6), 805-21.
2. Li, L. O.; Klett, E. L.; Coleman, R. A., Acyl-CoA synthesis, lipid metabolism and lipotoxicity. *Biochim Biophys Acta* **2010**, *1801* (3), 246-51.
3. Grevengoed, T. J.; Klett, E. L.; Coleman, R. A., Acyl-CoA metabolism and partitioning. *Annu Rev Nutr* **2014**, *34*, 1-30.
4. Patel, M. S.; Nemeria, N. S.; Furey, W.; Jordan, F., The pyruvate dehydrogenase complexes: structure-based function and regulation. *J Biol Chem* **2014**, *289* (24), 16615-23.
5. McCartney, R. G.; Rice, J. E.; Sanderson, S. J.; Bunik, V.; Lindsay, H.; Lindsay, J. G., Subunit interactions in the mammalian alpha-ketoglutarate dehydrogenase complex. Evidence for direct association of the alpha-ketoglutarate dehydrogenase and dihydrolipoamide dehydrogenase components. *J Biol Chem* **1998**, *273* (37), 24158-64.
6. Chuang, D. T.; Hu, C. W.; Ku, L. S.; Markovitz, P. J.; Cox, R. P., Subunit structure of the dihydrolipoyl transacylase component of branched-chain alpha-keto acid dehydrogenase complex from bovine liver. Characterization of the inner transacylase core. *J Biol Chem* **1985**, *260* (25), 13779-86.
7. Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M.; Escalante-Semerena, J. C., The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A. *Biochemistry* **2003**, *42* (10), 2866-73.
8. Jogl, G.; Tong, L., Crystal structure of yeast acetyl-coenzyme A synthetase in complex with AMP. *Biochemistry* **2004**, *43* (6), 1425-31.

9. Hisanaga, Y.; Ago, H.; Nakagawa, N.; Hamada, K.; Ida, K.; Yamamoto, M.; Hori, T.; Arii, Y.; Sugahara, M.; Kuramitsu, S.; Yokoyama, S.; Miyano, M., Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer. *J Biol Chem* **2004**, *279* (30), 31717-26.
10. Jogl, G.; Hsiao, Y. S.; Tong, L., Structure and function of carnitine acyltransferases. *Ann N Y Acad Sci* **2004**, *1033*, 17-29.
11. Pieklik, J. R.; Guynn, R. W., Equilibrium constants of the reactions of choline acetyltransferase, carnitine acetyltransferase, and acetylcholinesterase under physiological conditions. *J Biol Chem* **1975**, *250* (12), 4445-50.
12. Hersh, L. B., Kinetic studies of the choline acetyltransferase reaction using isotope exchange at equilibrium. *J Biol Chem* **1982**, *257* (21), 12820-5.
13. RosasGarcia, V. M.; Gandour, R. D., Conformationally-dependent free energies of solvation. An explanation for the large group-transfer potential of acetylcarnitine. *J Am Chem Soc* **1997**, *119* (32), 7587-7588.
14. Ramsay, R. R.; Gandour, R. D.; van der Leij, F. R., Molecular enzymology of carnitine transfer and transport. *Biochim Biophys Acta* **2001**, *1546* (1), 21-43.
15. Gobin, S.; Thuillier, L.; Jogl, G.; Faye, A.; Tong, L.; Chi, M.; Bonnefont, J. P.; Girard, J.; Prip-Buus, C., Functional and structural basis of carnitine palmitoyltransferase 1A deficiency. *J Biol Chem* **2003**, *278* (50), 50428-34.
16. Corti, S.; Bordoni, A.; Ronchi, D.; Musumeci, O.; Aguenouz, M.; Toscano, A.; Lamperti, C.; Bresolin, N.; Comi, G. P., Clinical features and new molecular findings in Carnitine Palmitoyltransferase II (CPT II) deficiency. *J Neurol Sci* **2008**, *266* (1-2), 97-103.

17. Hang, L.; Tang, M. C.; Harvey, C. J. B.; Page, C. G.; Li, J.; Hung, Y. S.; Liu, N.; Hillenmeyer, M. E.; Tang, Y., Reversible Product Release and Recapture by a Fungal Polyketide Synthase Using a Carnitine Acyltransferase Domain. *Angew Chem Int Ed Engl* **2017**, *56* (32), 9556-9560.
18. Chooi, Y. H.; Tang, Y., Navigating the fungal polyketide chemical space: from genes to molecules. *J Org Chem* **2012**, *77* (22), 9933-53.
19. Kudo, F.; Matsuura, Y.; Hayashi, T.; Fukushima, M.; Eguchi, T., Genome mining of the sordarin biosynthetic gene cluster from *Sordaria araneosa* Cain ATCC 36386: characterization of cycloaraneosene synthase and GDP-6-deoxyaltrose transferase. *J Antibiot (Tokyo)* **2016**, *69* (7), 541-8.
20. Ruswandi, S.; Kitani, K.; Akimitsu, K.; Tsuge, T.; Shiraishi, T.; Yamamoto, M., Structural analysis of cosmid clone pcAFT-2 carrying AFT10-1 encoding an acyl-CoA dehydrogenase involved in AF-toxin production in the strawberry pathotype of *Alternaria alternata*. *Journal of General Plant Pathology* **2005**, *71* (2), 107-116.
21. Brown, D. W.; Butchko, R. A.; Baker, S. E.; Proctor, R. H., Phylogenomic and functional domain analysis of polyketide synthases in *Fusarium*. *Fungal Biol* **2012**, *116* (2), 318-31.
22. 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), 14808-13.
23. 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), 6246-51.
24. Cacho, R. A.; Thuss, J.; Xu, W.; Sanichar, R.; Gao, Z.; Nguyen, A.; Vederas, J. C.; Tang, Y., Understanding Programming of Fungal Iterative Polyketide Synthases: The Biochemical Basis for Regioselectivity by the Methyltransferase Domain in the Lovastatin Megasyntase. *J Am Chem Soc* **2015**, *137* (50), 15688-91.
25. Roberts, D. M.; Bartel, C.; Scott, A.; Ivison, D.; Simpson, T. J.; Cox, R. J., Substrate selectivity of an isolated enoyl reductase catalytic domain from an iterative highly reducing fungal polyketide synthase reveals key components of programming. *Chem Sci* **2017**, *8* (2), 1116-1126.
26. Wagner, D. T.; Stevens, D. C.; Mehaffey, M. R.; Manion, H. R.; Taylor, R. E.; Brodbelt, J. S.; Keatinge-Clay, A. T., alpha-Methylation follows condensation in the gephyronic acid modular polyketide synthase. *Chem Commun (Camb)* **2016**, *52* (57), 8822-5.
27. Jencks, W. P.; Cordes, S.; Carriuolo, J., The free energy of thiol ester hydrolysis. *J Biol Chem* **1960**, *235*, 3608-14.
28. Liddle, E.; Scott, A.; Han, L. C.; Ivison, D.; Simpson, T. J.; Willis, C. L.; Cox, R. J., In vitro kinetic study of the squalestatin tetraketide synthase dehydratase reveals the stereochemical course of a fungal highly reducing polyketide synthase. *Chem Commun (Camb)* **2017**, *53* (10), 1727-1730.
29. Bracher, P. J.; Snyder, P. W.; Bohall, B. R.; Whitesides, G. M., The relative rates of thiol-thioester exchange and hydrolysis for alkyl and aryl thioalkanoates in water. *Orig Life Evol Biosph* **2011**, *41* (5), 399-412.

30. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B., Synthesis of proteins by native chemical ligation. *Science* **1994**, *266* (5186), 776-9.
31. Belecki, K.; Townsend, C. A., Biochemical determination of enzyme-bound metabolites: preferential accumulation of a programmed octaketide on the enediyne polyketide synthase CalE8. *J Am Chem Soc* **2013**, *135* (38), 14339-48.
32. Chen, M.; Liu, Q.; Gao, S.-S.; Young, A. E.; Jacobsen, S. E.; Tang, Y., Genome mining and biosynthesis of a polyketide from a biofertilizer fungus that can facilitate reductive iron assimilation in plant. *Proceedings of the National Academy of Sciences* **2019**, *116* (12), 5499-5504.
33. Yan, X.; Akinnusi, T. O.; Larsen, A. T.; Auclair, K., Synthesis of 4'-aminopantetheine and derivatives to probe aminoglycoside N-6'-acetyltransferase. *Org Biomol Chem* **2011**, *9* (5), 1538-46.
34. Wittwer, C. T.; Wyse, B. W.; Hansen, R. G., Enzymatic hydrolysis of pantetheine. *Methods Enzymol* **1986**, *122*, 36-43.
35. Agarwal, V.; Diethelm, S.; Ray, L.; Garg, N.; Awakawa, T.; Dorrestein, P. C.; Moore, B. S., Chemoenzymatic Synthesis of Acyl Coenzyme A Substrates Enables in Situ Labeling of Small Molecules and Proteins. *Org Lett* **2015**, *17* (18), 4452-5.
36. Franke, J.; Hertweck, C., Biomimetic Thioesters as Probes for Enzymatic Assembly Lines: Synthesis, Applications, and Challenges. *Cell Chem Biol* **2016**, *23* (10), 1179-1192.
37. Roy, R.; Touaibia, M., 3.36 - Application of Multivalent Mannosylated Dendrimers in Glycobiology. In *Comprehensive Glycoscience*, Kamerling, H., Ed. Elsevier: Oxford, 2007; pp 821-870.

38. Kuchurov, I. V.; Arabadzhi, S. S.; Zharkov, M. N.; Fershtat, L. L.; Zlotin, S. G., Sustainable Synthesis of Polynitroesters in the Freon Medium and their in Vitro Evaluation as Potential Nitric Oxide Donors. *ACS Sustainable Chemistry & Engineering* **2018**, *6* (2), 2535-2540.
39. Kumara Swamy, K. C.; Kumaraswamy, S.; Senthil Kumar, K.; Muthiah, C., Cyclic chlorophosphites as scaffolds for the one-pot synthesis of  $\alpha$ -aminophosphonates under solvent-free conditions. *Tetrahedron Letters* **2005**, *46* (19), 3347-3351.
40. Pike, R. D.; Reinecke, B. A.; Dellinger, M. E.; Wiles, A. B.; Harper, J. D.; Cole, J. R.; Dendramis, K. A.; Borne, B. D.; Harris, J. L.; Pennington, W. T., Bicyclic Phosphite Esters from Pentaerythritol and Dipentaerythritol: New Bridging Ligands in Organometallic and Inorganic Chemistry. *Organometallics* **2004**, *23* (9), 1986-1990.
41. Padias, A. B.; Hall, H. K., Synthesis and polymerization of pentaerythritol monoacrylate and methacrylate and their bicyclic ortho esters. *Macromolecules* **1982**, *15* (2), 217-223.
42. Dunn, T. J.; Neumann, W. L.; Rogic, M. M.; Woulfe, S. R., Versatile methods for the synthesis of differentially functionalized pentaerythritol amine derivatives. *The Journal of Organic Chemistry* **1990**, *55* (26), 6368-6373.
43. Li, Y. F.; Tsai, K. J. S.; Harvey, C. J. B.; Li, J. J.; Ary, B. E.; Berlew, E. E.; Boehman, B. L.; Findley, D. M.; Friant, A. G.; Gardner, C. A.; Gould, M. P.; Ha, J. H.; Lilley, B. K.; McKinstry, E. L.; Nawal, S.; Parry, R. C.; Rothchild, K. W.; Silbert, S. D.; Tentilucci, M. D.; Thurston, A. M.; Wai, R. B.; Yoon, Y. J.; Aiyar, R. S.; Medema, M. H.; Hillenmeyer, M. E.; Charkoudian, L. K., Comprehensive curation and analysis of fungal biosynthetic gene clusters of published natural products. *Fungal Genet. Biol.* **2016**, *89*, 18-28.
44. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215* (3), 403-410.

45. McGinnis, S.; Madden, T. L., BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res* **2004**, *32*, W20-W25.
46. Marchler-Bauer, A.; Lu, S. N.; Anderson, J. B.; Chitsaz, F.; Derbyshire, M. K.; DeWeese-Scott, C.; Fong, J. H.; Geer, L. Y.; Geer, R. C.; Gonzales, N. R.; Gwadz, M.; Hurwitz, D. I.; Jackson, J. D.; Ke, Z. X.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Mullokandov, M.; Omelchenko, M. V.; Robertson, C. L.; Song, J. S.; Thanki, N.; Yamashita, R. A.; Zhang, D. C.; Zhang, N. G.; Zheng, C. J.; Bryant, S. H., CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **2011**, *39*, D225-D229.
47. Marchler-Bauer, A.; Derbyshire, M. K.; Gonzales, N. R.; Lu, S. N.; Chitsaz, F.; Geer, L. Y.; Geer, R. C.; He, J.; Gwadz, M.; Hurwitz, D. I.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Song, J. S.; Thanki, N.; Wang, Z. X.; Yamashita, R. A.; Zhang, D. C.; Zheng, C. J.; Bryant, S. H., CDD: NCBI's conserved domain database. *Nucleic Acids Res* **2015**, *43* (D1), D222-D226.
48. Grigoriev, I. V.; Nordberg, H.; Shabalov, I.; Aerts, A.; Cantor, M.; Goodstein, D.; Kuo, A.; Minovitsky, S.; Nikitin, R.; Ohm, R. A.; Otilar, R.; Poliakov, A.; Ratnere, I.; Riley, R.; Smirnova, T.; Rokhsar, D.; Dubchak, I., The Genome Portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Res* **2012**, *40* (D1), D26-D32.
49. Nordberg, H.; Cantor, M.; Dusheyko, S.; Hua, S.; Poliakov, A.; Shabalov, I.; Smirnova, T.; Grigoriev, I. V.; Dubchak, I., The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res* **2014**, *42* (D1), D26-D31.
50. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E., The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **2015**, *10* (6), 845-858.

51. Lee, K. K. M.; Da Silva, N. A.; Kealey, J. T., Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. *Anal Biochem* **2009**, *394* (1), 75-80.
52. Sambrook, J.; Russell, D. W.; Sambrook, J., *The condensed protocols from Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., 2006; p v, 800 p.
53. Wiemer, B., D. D. Perrin and W. L. F. Armarego: Purification of Laboratory Chemicals. 3. Aufl., Oxford. Pergamon Press, 1988, 391 S., zahlr. Tab., \$ 37,50, ISBN 0-08-034714-2. *Acta hydrochimica et hydrobiologica* **1989**, *17* (6), 632-632.
54. Corey, E. J.; Suggs, J. W., Pyridinium chlorochromate. An efficient reagent for oxidation of primary and secondary alcohols to carbonyl compounds. *Tetrahedron Letters* **1975**, *16* (31), 2647-2650.
55. Cochrane, R. V.; Sanichar, R.; Lambkin, G. R.; Reiz, B.; Xu, W.; Tang, Y.; Vederas, J. C., Production of New Cladosporin Analogues by Reconstitution of the Polyketide Synthases Responsible for the Biosynthesis of this Antimalarial Agent. *Angew Chem Int Ed Engl* **2016**, *55* (2), 664-8.
56. Cane, D. E.; Kudo, F.; Kinoshita, K.; Khosla, C., Precursor-directed biosynthesis: biochemical basis of the remarkable selectivity of the erythromycin polyketide synthase toward unsaturated triketides. *Chem Biol* **2002**, *9* (1), 131-42.
57. Nhu Lam, M.; Dudekula, D.; Durham, B.; Collingwood, N.; Brown, E. C.; Nagarajan, R., Insights into beta-ketoacyl-chain recognition for beta-ketoacyl-ACP utilizing AHL synthases. *Chem Commun (Camb)* **2018**, *54* (64), 8838-8841.

58. Xu, W.; Cai, X. L.; Jung, M. E.; Tang, Y., Analysis of Intact and Dissected Fungal Polyketide Synthase-Nonribosomal Peptide Synthetase in Vitro and in *Saccharomyces cerevisiae*. *Journal of the American Chemical Society* **2010**, *132* (39), 13604-13607.
59. Puigbo, P.; Guzman, E.; Romeu, A.; Garcia-Vallve, S., OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res* **2007**, *35*, W126-W131.
60. Gibson, D. G.; Benders, G. A.; Axelrod, K. C.; Zaveri, J.; Algire, M. A.; Moodie, M.; Montague, M. G.; Venter, J. C.; Smith, H. O.; Hutchison, C. A., One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *P Natl Acad Sci USA* **2008**, *105* (51), 20404-20409.
61. Ma, S. M.; Li, J. W.; Choi, J. W.; Zhou, H.; Lee, K. K.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y., Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* **2009**, *326* (5952), 589-92.