

CHARACTERIZING AND ENGINEERING THIOESTERASES AS BIOCATALYSTS FOR MACROCYCLIZATION

Jordan Brazeau-Henrie

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Doctorate in Philosophy degree in Chemistry

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa

Supervisor:

Dr. Christopher N. Boddy

Examining Board:

Dr. Jeffrey Keillor
Dr. Adam Damry
Dr. Roberto Chica

External Examiner:

Dr. David Sherman

ABSTRACT

Natural products (NP) present themselves as complex molecules with useful and potent biological activities. Notably, macrocyclic NP have garnered significant attention for defying the prototypical “rule of 5” yet retaining optimal pharmacological properties, and the ability to bind to targets with high affinity and specificity. While synthetic macrocyclization poses considerable challenges, nature effortlessly generates these complex molecules using thioesterase (TE) domains positioned at the C-terminus of modular biosynthetic machinery. This has spurred researchers to explore the use of TEs as *in vitro* macrocyclization biocatalysts, enabling the chemoenzymatic synthesis of numerous macrocycles. In the context of the existing literature on TEs, we proposed two models to provide a rational framework for understanding TE-mediated macrocyclization. The research presented here describes the characterization of two non-ribosomal peptide synthetase (NRPS) TE domains and their use as biocatalysts for the chemoenzymatic synthesis of two cyclic depsipeptide NP. Unlike other characterized NRPS TEs, these two demonstrated the ability to catalyze macrolactonization with minimal competing hydrolysis reactivity, making them promising biocatalysts. An uncommon tandem TE domain was also characterized, revealing its tolerance for N-terminal peptide modification, peptide sequence composition and stereochemical configuration of the nucleophile used for ring closure. These attributes make it a compelling candidate for the development of a general peptide macrocyclization biocatalyst. Finally, the first extensive engineering of a polyketide synthase (PKS) TE was undertaken via a rational protein design approach and an LCMS-based screening method. Four variants were identified with improved macrocyclization activity over wild-type (WT), highlighting the utility and success of TE protein engineering. This pioneering engineering project serves as a foundation for future TE engineering projects that will generate useful enzymes with optimized activities. Overall, the research presented further develops the field of macrocyclization biocatalysis by showcasing the utility and engineering of TE domains.

ACKNOWLEDGEMENTS

First, I would like to thank my mom, younger brother and sister for their love and support, for bearing with all my late-night studying and pretending to understand my work as I practice my presentations. I would also like to thank my loving partner for being there for me from the beginning of my PhD studies. You have been my biggest supporter, despite being an engineer and making fun of me for being a scientist.

I would like to particularly acknowledge Chris for being the best mentor a PhD student could ever ask for. You have helped me develop as a scientist, allowed me to grow as a protein engineer in your lab and encouraged me to push further. Your scientific curiosity is inspiring, I have and will always admire this quality of yours. Thank you, I will never forget all that you have done for me.

A thank you to all the Boddy lab members past and present who helped me in my research. To Jessica, I thank you for all your support throughout my time at uOttawa, for all our dinner dates and for being the best fruity friend. To Graham, I thank you for mentoring me as an undergraduate student, without you I would not be the scientist I am today. To Mike and André, for being such great collaborators. To Richenda, for being the best lab friend, for all our coffee dates and your endless support. I truly would not have made it through graduate school without you, you will forever be my werk wife. To Taylor, for being the most wonderful undergrad I had the pleasure in mentoring. To Allison, for being my fellow fruity lab mate and for all our hot girl walks. Without you, I would have likely gone mad. To Kaitlyn, former Pezacki lab member, for being the best cat mom and an endless supporter. Thank you for your beautiful quilt, I will cherish it and our wifey-ship forever. And to everyone else who contributed to my growth as a PhD student.

Btw love the shirt Jeff! You now have my eternal approval for whatever you wear when I'm not around.

“The best dissertation is a finished dissertation.”

- Ancient Grad Student Proverb

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iii
TABLE OF CONTENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
LIST OF SYMBOLS, NOMENCLATURE OR ABBREVIATIONS.....	x
CHAPTER 1 : Thioesterases as biocatalysts to access macrocycles	1
1.1 Preface.....	1
1.2 Literature review	2
1.2.1 Macrocycles are bioactive but synthetically challenging targets	2
1.2.2 Nature's macrocycle assembly lines: PKS and NRPS machinery	4
1.2.3 TE domains classification, structure and mechanism.....	9
1.2.4 Phylogenetic and structural determinants for TE off-loading	12
1.2.5 A rational framework for TE off-loading selectivity.....	16
1.2.6 Enzymes are useful biocatalytic tools for the synthesis of pharmaceuticals	18
1.2.7 TE biocatalysts: a solution to the challenges faced in macrocycle synthesis	20
1.2.8 Protein engineering: a brief overview	24
1.2.9 Towards engineering TEs, current work and limitations	25
1.3 Summary and conclusions	27
1.4 Supplementary information	29
1.5 References	34
CHAPTER 2 : Chemoenzymatic synthesis of depsipeptides using NRPS TEs.....	46
2.1 Preface.....	46
2.2 Introduction.....	47
2.3 Results and discussion	51
Synthesis of Sgd and Rzm substrates and macrocyclic authentic standards	51
Sgd- and Rzm-TE are active macrolactonization specific catalysts.....	55
Sgd-TE can effect ring-opening of authentic 2.3	59
NRPS TEs as macrolactonization biocatalysts in depsipeptide synthesis	61
2.4 Conclusions.....	62
2.5 Materials and methods.....	63
Synthetic protocols.....	63
Cloning, expression and purification	75
Biochemical assays	77
2.6 Supplementary information	80
Compound purity tables	80
Selected NMR, mass spectra and UV traces.....	81
Protein purities, yields, sequences and Rzm-TE predicted structure	98
Sgd and Rzm biosynthetic gene clusters	100
Standard curve for Sgd-TE kinetics	102
Sgd-TE product ratios and Rzm-TE relative conversion.....	102
Sgd-TE assay chromatograms and mass spectra	103
Rzm-TE assay chromatograms and mass spectra	115

Example UV trace for enzymatic assays	119
2.7 References	120
CHAPTER 3 : Characterization of an NRPS tandem TE domain in depsipeptide macrolactonization	126
3.1 Preface	126
3.2 Introduction	127
3.3 Results and discussion	133
Gac-TE1-TE2 can macrolactonize both <i>N</i> -Cbz (3.1) and <i>N</i> -acetyl (3.4) substrates.....	133
Gac-TE1 is responsible for the macrolactonization activity	135
Synthesis of Thr _x (3.7 , 3.10 and 3.13) peptide substrates.....	137
Gac-TE1-TE2 is substrate tolerant for macrolactonization.....	139
3.4 Conclusions.....	142
3.5 Materials and methods.....	143
Synthetic protocol for Thr _x peptide-SNAC.....	143
Cloning, expression and purification	146
Biochemical assays	149
3.6 Supplementary information	151
Selected NMR, mass spectra and UV traces.....	151
Protein purities, yields and sequences	154
Gac-TE1-TE2 predicted structure.....	156
<i>N</i> -Cbz peptide (3.1) assay LCMS chromatograms and mass spectra.....	157
<i>N</i> -acetyl peptide (3.4) assay LCMS chromatograms and mass spectra.....	162
Thr _x peptide (3.7 , 3.10 and 3.13) assay LCMS chromatograms and mass spectra	168
Relative macrocycle conversion for <i>N</i> -Cbz (3.1) and <i>N</i> -acetyl (3.4) substrates	174
3.7 References	175
CHAPTER 4 : Engineering the DEBS-TE as a better macrocyclization biocatalyst	180
4.1 Preface	180
Chapter contributions.....	180
4.2 Introduction	181
4.3 Results and discussion	186
DEBS-TE mutant library construction and screen.....	186
R32L, Q176M and N180D are better at macrocyclization than WT	194
WT DEBS-TE trapped acyl crystal structure	197
Proposed mutation effects to rationalize changes in macrocyclization activity	199
Combination of single mutations do not exhibit synergistic effects	203
Improved macrocyclization mutants are kinetically slower	206
4.4 Conclusions.....	208
4.5 Materials and methods.....	210
Mutant library construction and storage	210
Mutant library screen and data analysis	212
Combinatorial mutant generation and cloning	214
Expression and purification of mutants.....	216
Biochemical assays of WT DEBS-TE and mutants	218
4.6 Supplementary information	220
Constructs used and protein purities	220
Mutant library LCMS screen data and statistics	223
Michaelis-Menten kinetics and example standard curve.....	233

Example LCMS chromatograms for library screen and biochemical assays.....	245
Mutant library screen AUC raw data from MZmine2.....	259
Biochemical assays raw AUC and product ratios data.....	295
4.7 References	297
CHAPTER 5 : Conclusions and future directions	302
5.1 Preface	302
5.2 Perspective on ICS and CS models.....	303
5.3 Perspective on NRPS TEs	305
5.4 Outlook on engineering TEs.....	311
5.5 Conclusions.....	313
5.6 References	314
APPENDIX I: Permissions and manuscripts	318

LIST OF TABLES

Table 3.1. Primers used in this study.	148
Table 4.1. Saturation mutagenesis primers used for library construction.	211
Table 4.2. Primers for combinatorial mutant generation.....	216
Table 5.1. Summary of key indicators for the ICS and CS models with examples from the literature.....	304

LIST OF FIGURES

Figure 1.1. Final macrolactonization step in example total syntheses of natural products.	3
Figure 1.2. Example of PKS and NRPS macrolactone natural products.	4
Figure 1.3. Example of polyketide biosynthesis.	6
Figure 1.4. Example of non-ribosomal peptide biosynthesis and specialized domains.	8
Figure 1.5. Representation of a type I TE structure (valinomycin).	10
Figure 1.6. Canonical TE simplified catalytic mechanism.	11
Figure 1.7. Protein sequence diversity dendrogram of macrocyclizing TE domains.	13
Figure 1.8. Dendrogram zoom-in of bacterial type I PKS clade.	14
Figure 1.9. Models for TE-mediated macrocyclization.	17
Figure 1.10. Biocatalytic versus original asymmetric hydrogenation process in sitagliptin synthesis.	19
Figure 1.11. Biocatalytic versus Ru-catalyzed asymmetric hydrogenation in ipatasertib synthesis.	20
Figure 1.12. Macrocyclization of peptides catalyzed by the Tyc-TE.	22
Figure 1.13. DEBS-TE acting on non-native substrates.	24
Figure 2.1. TE-mediated release of Sgd E and Rzm A.	48
Figure 2.2. LCMS chromatograms for Sgd-TE assays.	56
Figure 2.3. LCMS chromatograms for Rzm-TE assays.	57
Figure 2.4. Michaelis-Menten kinetics of Sgd-TE.	58
Figure 2.5. Ring opening assays with Sgd-TE and macrocycle 2.3	60
Figure 2.6. TE reactions with SNAC activated substrates.	60
Figure 3.1. Biosynthesis of Gacamide A.	129
Figure 3.2. Summary of tandem TE domain activity.	131
Figure 3.3. TE catalyzed reactions and LCMS chromatograms of assays with substrates 3.1 and 3.4	134
Figure 3.4. TE1-TE2 catalyzed reactions with peptide library and assay LCMS chromatograms.	141
Figure 4.1. Structure of three example bioactive macrocyclic polyketides.	181
Figure 4.2. Biosynthesis of erythromycin A.	182
Figure 4.3. DEBS-TE <i>in vitro</i> reactions with SNAC activated substrates.	184
Figure 4.4. Chosen sites for mutagenesis and WT DEBS-TE reaction with SNAC substrate 4.1	187
Figure 4.5. Mutant library generation and screen workflow.	189
Figure 4.6. Representative scatter plots of library screen.	190
Figure 4.7. Statistical analysis of library screen.	192
Figure 4.8. Enzymatic and kinetic assay results for DEBS-TE single mutants.	195
Figure 4.9. Comparison of substrate 4.1 , 4.5 structures and WT DEBS-TE second order rate constants.	197
Figure 4.10. WT DEBS-TE trapped acyl crystal structure compared to <i>apo</i> DEBS-TE.	198
Figure 4.11. Interaction network of the DEBS-TE active site in complex with 4.1	200
Figure 4.12. DEBS-TE active site in complex with substrate 4.1	202
Figure 4.13. Enzymatic and kinetic assay results for DEBS-TE combinatorial mutants.	204
Figure 4.14. Relationship between k_{cat}/K_M and normalized M/H ratio.	207
Figure 5.1. General catalytic schemes for Rzm, SnbDE and Sgd TEs.	305
Figure 5.2. Overview of proposed peptide library screen.	307
Figure 5.3. Biaryl ether coupling in vancomycin (A) and seongsanamide (B) biosynthesis.	308
Figure 5.4. Macrocyclization activity of Gac-tandem domain on various substrates.	310

LIST OF SYMBOLS, NOMENCLATURE OR ABBREVIATIONS

6xHis:	Hexa-His
ACP:	Acyl carrier protein
alle:	Alloisoleucine
AMP:	Adenosine monophosphate
Arf:	Arthrofactin
AT:	Acyltransferase
AUC:	Area under the curve
C/E:	Dual condensation-epimerization
C:	Condensation
Cbz:	Benzyl chloroformate
CDA:	Calcium dependent antibiotic
Ces:	Cereulide
CP:	Carrier protein
C _s :	Starter condensation
Cy:	Cyclization
DAP:	2,3-diaminopropionic acid
DBU:	Diazabicycloundecene
DCM:	Dichloromethane
DEBS:	6-deoxyerythronolide B
DH:	Dehydratase
DIPEA or DIEA:	<i>N,N</i> -diisopropylethylamine
DMAP:	4-dimethylaminopyridine

DMF: Dimethylformamide

DMSO: Dimethylsulfoxide

dNTP: Deoxynucleotide triphosphate

DTNB: Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid))

E: Epimerization

EDC: 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

ee: Enantiomeric excess

EIC: Extracted ion chromatogram

Epo: Epothilone C

ER: Enoylreductase

ESI: Electrospray ionization

FA: Formic acid

FAAH: Fatty acid amide hydrolase

FAS: Fatty acid synthases

Fmoc: Fluorenylmethoxycarbonyl

Gac A: Gacamide A

gDNA: Genomic DNA

Grs: Gramicidin

HBTU: Hexafluorophosphate benzotriazole tetramethyl uronium

HOAt: 1-hydroxy-7-azabenzotriazole

HOBt: Hydroxybenzotriazole

HPLC: High-pressure liquid chromatography

HRMS: High-resolution mass spectrometry

hVISA: Vancomycin-heterointermediate *Staphylococcus aureus*

ID: Identification

i-Pr: Isopropyl alcohol

IPTG: Isopropyl- β -*D*-thiogalactopyranoside

KS or KRED: Ketoreductase

KS: Ketosynthase

LB: Lysogeny broth

LCMS: Liquid chromatography-mass spectrometry

LybB: Lysobactin

M/H: Macrocyclic to hydrolysis ratio

MDS: Molecular dynamic simulations

mLPD: Macrocyclic lipodepsipeptides

MRSA: Methicillin-resistant *Staphylococcus aureus*

MS: Mass spectrometer

MT: *N*-methyl transferases

MW: Microwave

MWCO: Molecular weight cut-off

Ni-NTA: Nickel nitrilotriacetic acid agarose

NMP: *N*-methylpyrrolidone

NP: Natural products

NR-PKS: Non-ribosomal-polyketide synthase hybrid

NRPS: Non-ribosomal synthetases

NysK: Nystatin (aka pimaricin)-forming thioesterase

ONPG: *o*-nitrophenyl- β -galactoside

P/S: Product to substrate ratio

PCP: Peptidyl carrier protein

PhSH: Thiophenol

PIK: Pikromycin

Pima: Pimaricin/nystatin

PKS: Polyketide synthases

PNK: Polynucleotide kinase

PPT: Phosphopantetheine arm

PRE: Product release enzyme

QM/MM: Quantum mechanics/molecular mechanics

RiPP: Ribosomally synthesized and post-translationally modified peptide

RP-HPLC: Reverse-phase high-pressure liquid chromatography

Rzm: Rhizomide

S→C: Catalytic Ser to Cys mutation

ScnS4: Natamycin-forming thioesterase

sfGFP: Superfolder green fluorescent protein (GFP)

Sgd: Seongsanamide

SNAC: *N*-acetylcysteamine

SnbDE: Streptogramin-forming TE

SPPS: Solid phase peptide synthesis

Srf: Surfactin

T: Thiolation

TE: Thioesterase

TFA: Trifluoroacetic acid

Txo: Teixobactin

Tyc: Tyrocidine

v_0 : Initial velocities

Vlm: Valinomycin

WT: Wild-type

β -HDA: *S*- β -hydroxydecanoic acid

Δ : Inactivation (catalytic Ser to Ala mutation)

CHAPTER 1: Thioesterases as biocatalysts to access macrocycles

1.1 Preface

The literature review presented in this chapter will summarize macrocyclic natural products, their biosynthesis, and how they are accessed by chemists using synthetic and biocatalytic methodologies. Special focus is directed towards thioesterase (TE) domains, detailing their structure, mechanism and utility as macrocyclization biocatalysts. The discussion extends to the optimization of these powerful biocatalysts through protein engineering. A manuscript based on this chapter is currently being drafted for submission as a review article by myself and Dr. Christopher N. Boddy.

1.2 Literature review

1.2.1 Macrocycles are bioactive but synthetically challenging targets

Natural products (NP) have played a significant role in drug development with many possessing potent and useful biological activities.^{1,2} In fact, 33% of all drugs approved between 1981 and 2014 are NP or NP-derived.³ They have such useful activities due to their complex and diverse structures. In particular, macrocyclic NP have significant potential as drug leads owing to their biologics-like selectivity, potency and their small molecule-like pharmacokinetics. Typically, macrocyclic NP contain ring structures ranging from 12-40 atoms,⁴ but can contain more as seen with the stambomycins.^{5,6} The macrocyclic nature of these systems contributes to their advantageous pharmacokinetic and drug-like properties. The constrained flexibility of macrocycles allows for the adoption of conformations where hydrophobic and hydrophilic groups are either buried or exposed for optimal solubility and membrane penetration.⁷⁻⁹ Moreover, this constrained conformation maximizes enthalpic contributions of binding, leading to compounds that are highly active and selective against previously “undruggable” targets.¹⁰⁻¹² While there is much interest in generating these complex molecules,^{7,13-17} macrocyclization is inherently challenging, even with the use of good internal nucleophiles such as amines (macrolactamization). This difficulty arises from the high entropic cost of end-to-end contact and is often outcompeted by formation of smaller rings and oligomers.^{18,19} The challenge is further intensified for macrolactonization, as these reactions require activation of an internal alcohol group to be sufficiently nucleophilic for ring closure.

The challenging nature of macrocyclization has prompted the development of many different and robust methodologies for macrocycle synthesis.^{19,20} Despite these efforts, the late-stage macrocyclization step in the total synthesis of an NP or NP-derived linear precursor still remains difficult, highlighting limitations in even the most robust of methods. For example, in the 1981 total synthesis of erythromycin, the final macrocyclization step required screening of different linear precursor molecules with varied

protecting groups before it was successfully cyclized by the Corey-Nicolaou macrolactonization.²¹ Currently, the Yamaguchi mixed anhydride macrolactonization and the Yonemitsu variant are the most widely used and consistent macrolactonization methods.^{19,20,22,23} Unfortunately, these reactions can still fall short in producing the desired macrocycle (**Figure 1.1**). This is exemplified in depsipeptide macrolactonization, where stereogenic carbons alpha to the activated ester are prone to epimerization, producing a mixture of diastereomeric products.²⁴ Due to the synthetic challenges in constructing macrolactones, researchers construct macrocyclic depsipeptides, which are easily accessed through solid phase peptide synthesis (SPPS) via amide bond forming macrolactamization chemistries.²⁵ Thus, there is clearly a pressing need for new approaches to access complex, structurally diverse macrolactones.

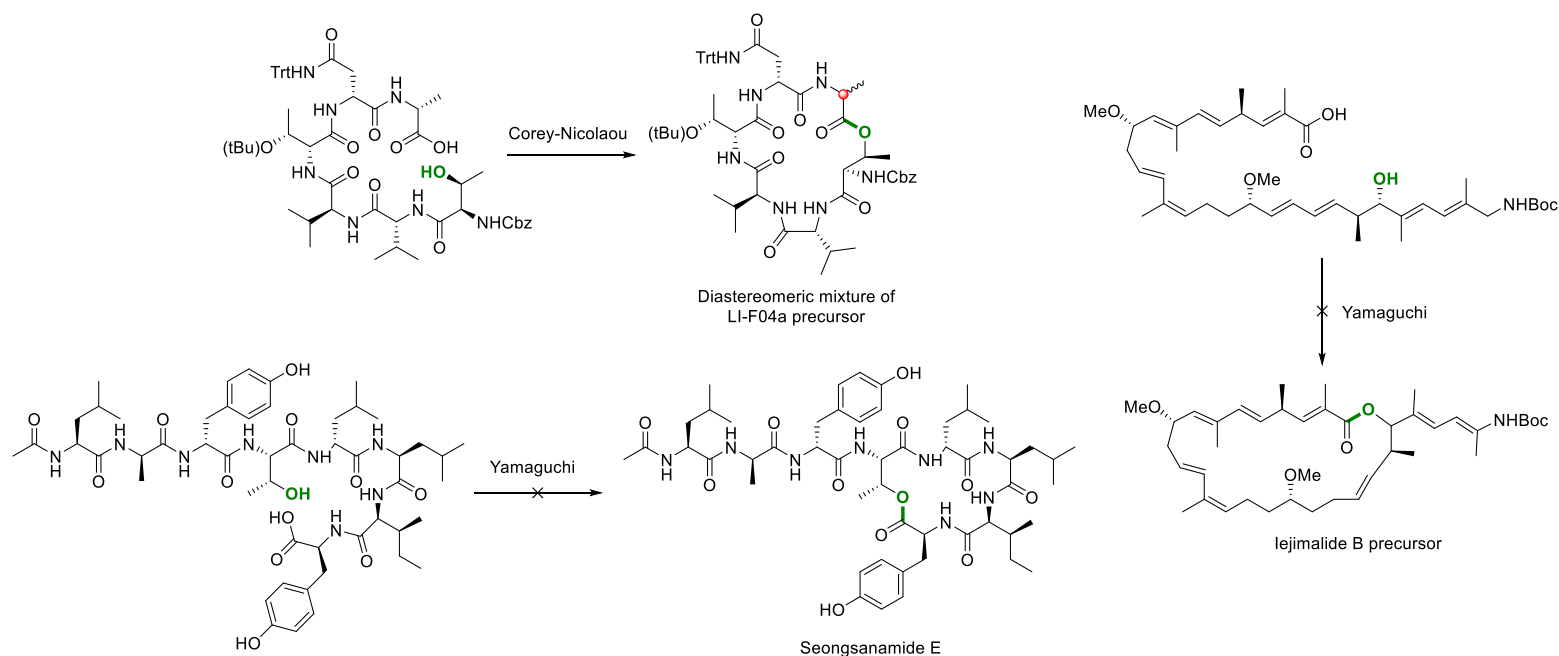


Figure 1.1. Final macrolactonization step in example total syntheses of natural products. Examples of problematic and failed macrolactonization in the total syntheses of lejimalide B,²⁶ LI-F04a²⁷ and seongsanamide E.²⁸ Note the LI-F04a macrolactonization did not fail but produced a diastereomeric mixture of products due to epimerization at the α -position (shown as red dot). The macrolactone linkage formed is shown in bold and green.

1.2.2 Nature's macrocycle assembly lines: PKS and NRPS machinery

Nature is a rich source of macrocyclic molecules, and many organisms biosynthesize them. Although they are not required for the organism's survival under normal conditions, they improve its survival fitness by providing for example, interspecies defence or the ability to compete with other organisms when resources are limited.^{29,30} Two of the most well studied biosynthetic machineries that generate macrocyclic NP are the polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS). Both are exceptional sources of macrolactones of varying ring sizes and biological activities (**Figure 1.2**).

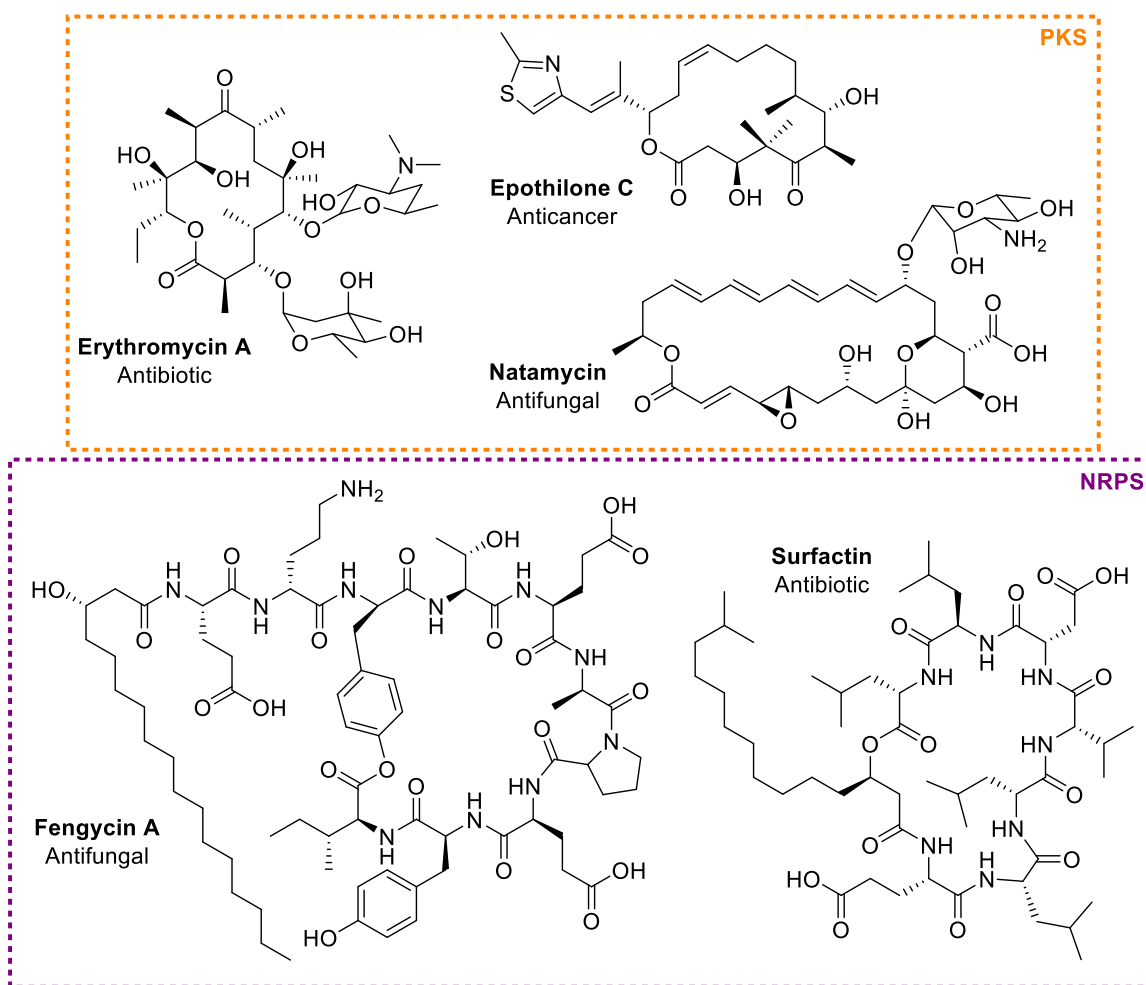


Figure 1.2. Example of PKS and NRPS macrolactone natural products.

PKS are large multienzyme complexes that build a linear polyketide precursor by adding two carbons each extension cycle. These pathways are analogous to fatty acid synthases (FAS),^{31,32} however, they can vary the oxidation state of the final molecule to yield ketones, alcohols, olefins or fully saturated chains. PKS are categorized into multiple types; type I modular (bacterial),³¹ type I iterative (fungal),³² type II,³³ and type III.³⁴ However, this chapter will focus exclusively on type I modular PKS, as they generate majority of the macrocyclic polyketides. Type I modular PKS are commonly found in bacteria and consist of multiple catalytic domains organized in modules that work cooperatively and non-iteratively for polyketide extension (**Figure 1.3A**). Polyketide biosynthesis initiates with the loading of an acyl-CoA starter unit onto the acyl carrier protein (ACP) of the first module, typically referred to as the loading module. Subsequently, the loaded acyl-chain is transferred to the active site Cys of the downstream ketosynthase (KS) domain. An acyltransferase (AT) domain then selects the appropriate malonyl-CoA extender unit, which is loaded onto the phosphopantetheine arm (PPT) prosthetic group of the downstream ACP. Once acyl and malonyl units are loaded onto the ACP and the upstream KS domain respectively, a decarboxylative Claisen-like condensation reaction catalysed by the KS domain extends the linear precursor by two carbons (i.e. 1 ketide unit). This extended product can then be transferred to the next downstream KS domain for further extension. Modules also frequently include reducing domains for sequential tailoring at the β -position, such as a ketoreductase (KR) to generate a β -hydroxyl, a dehydratase (DH) to eliminate the alcohol and create an olefin, and finally an enoylreductase (ER) to reduce the olefin into a saturated system (**Figure 1.3B**).³⁵ This tailoring occurs after decarboxylative condensation and prior to transfer to the downstream KS. Once the final extension and reductive tailoring cycle is complete, the product is transferred to a dedicated terminal release domain, forming an acyl-enzyme intermediate that is then released from the biosynthetic machinery. Typically, a thioesterase (TE) domain is responsible for this function³⁶ and will be discussed in detail in this review.

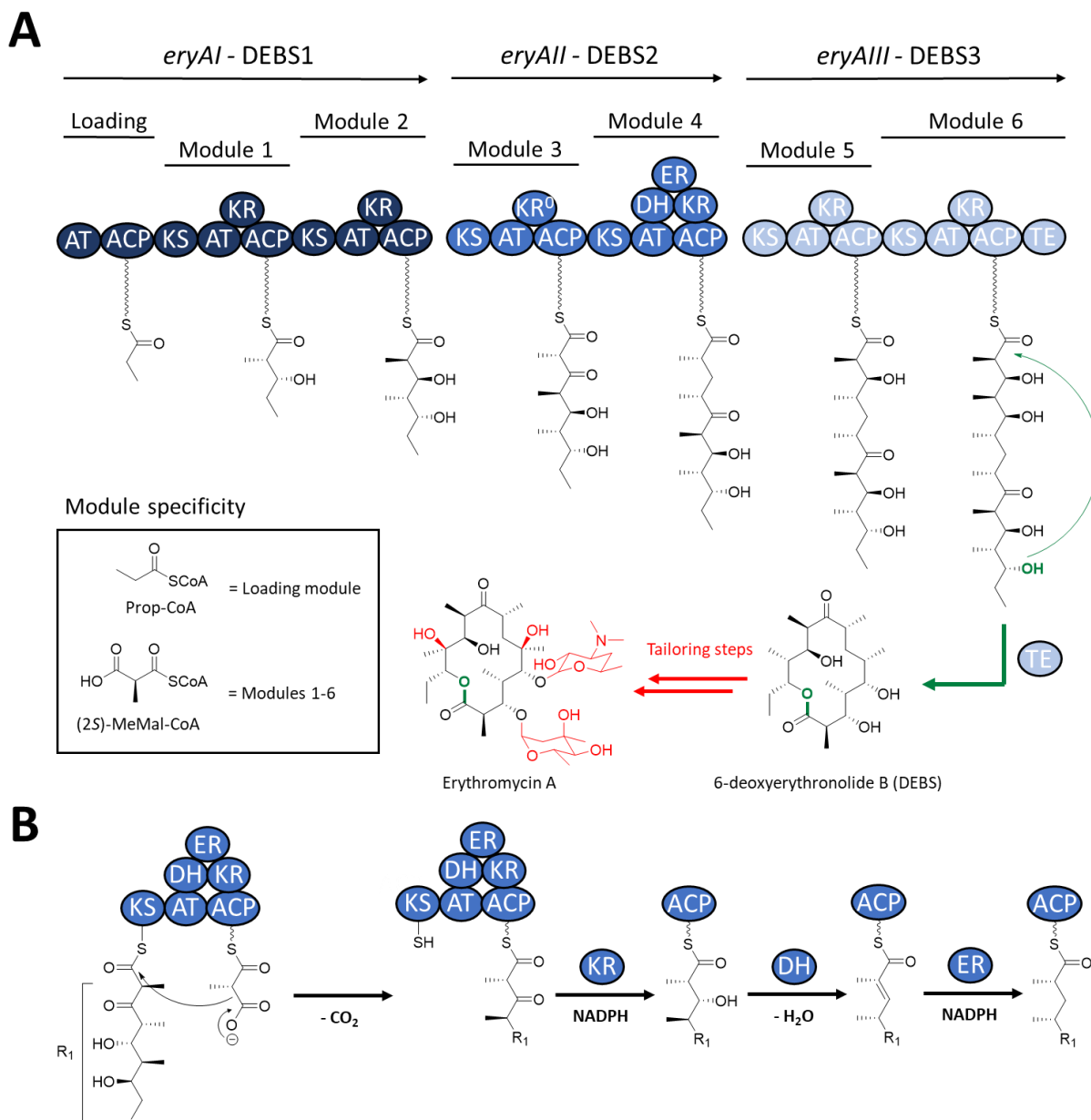


Figure 1.3. Example of polyketide biosynthesis. (A) Shown is the 6-deoxyerythronolide B (DEBS) PKS pathway with ACP thioester intermediates and the product formed after all extension and tailoring cycles. (B) The decarboxylative Claisen-like condensation of polyketide intermediate in DEBS module 4 for extension followed by sequential reductive tailoring at the β -position.

NRPS are conceptually analogous to PKS as they contain a modular architecture; however, they generate peptide-based products. In this process, the NRPS extend the growing peptide natural product by one amino acid for each module without using the ribosomal machinery. This allows them to incorporate non-proteogenic and D-configured amino acids, leading to peptide products with immense structural diversity. Building of the peptide core starts with the selection of an amino acid that is activated as an AMP-ester by an adenylation (A) domain and loaded onto the PPT prosthetic group of the peptidyl carrier protein or thiolation (PCP or T) domain. At times, the first module can contain a starter condensation domain (C_s) for condensation of an acyl-CoA ester onto the N-terminus of the starting amino acid unit. This is typically seen in the biosynthesis of macrocyclic lipopeptides where lipid chains such as *S*- β -hydroxydecanoic acid (β -HDA) are incorporated.^{37,38} An A domain then loads an amino acid onto the downstream T domain and a condensation (C) domain catalyzes condensation of the amino acid onto the upstream T domain's peptidyl chain, thereby extending the peptide (**Figure 1.4A**). While the A domain ensures amino acid loading selectivity, the C domain exhibits stereoselectivity for amide bond formation between the two amino acids. For instance, certain C domains specifically catalyze amide bonds between two L- or D-configured amino acids (^LC_L, **Figure 1.4B** or ^DC_D) or between an L- and D-configured amino acid (^DC_L).³⁹ Extra tailoring domains can be found within modules to modify the nascent peptide. Such domains can include *N*-methyl transferases (MT) for methylation of the backbone amides, X domains that act as a P450 anchor for amino acid oxidation,⁴⁰ cyclization (Cy) domains that form thiazoline and oxazoline 5-membered ring systems from Cys, Thr and Ser sidechains, and epimerization (E) domains to switch the amino acid configuration at the α -carbon (**Figure 1.4C**). Alternatively, a specialized C domain with dual condensation and epimerization (C/E) activity can epimerize the amino acid before catalyzing the condensation reaction.⁴¹ The extension and tailoring cycles continue until the peptide chain is released by a dedicated release domain, typically catalyzed by a TE domain.³⁶

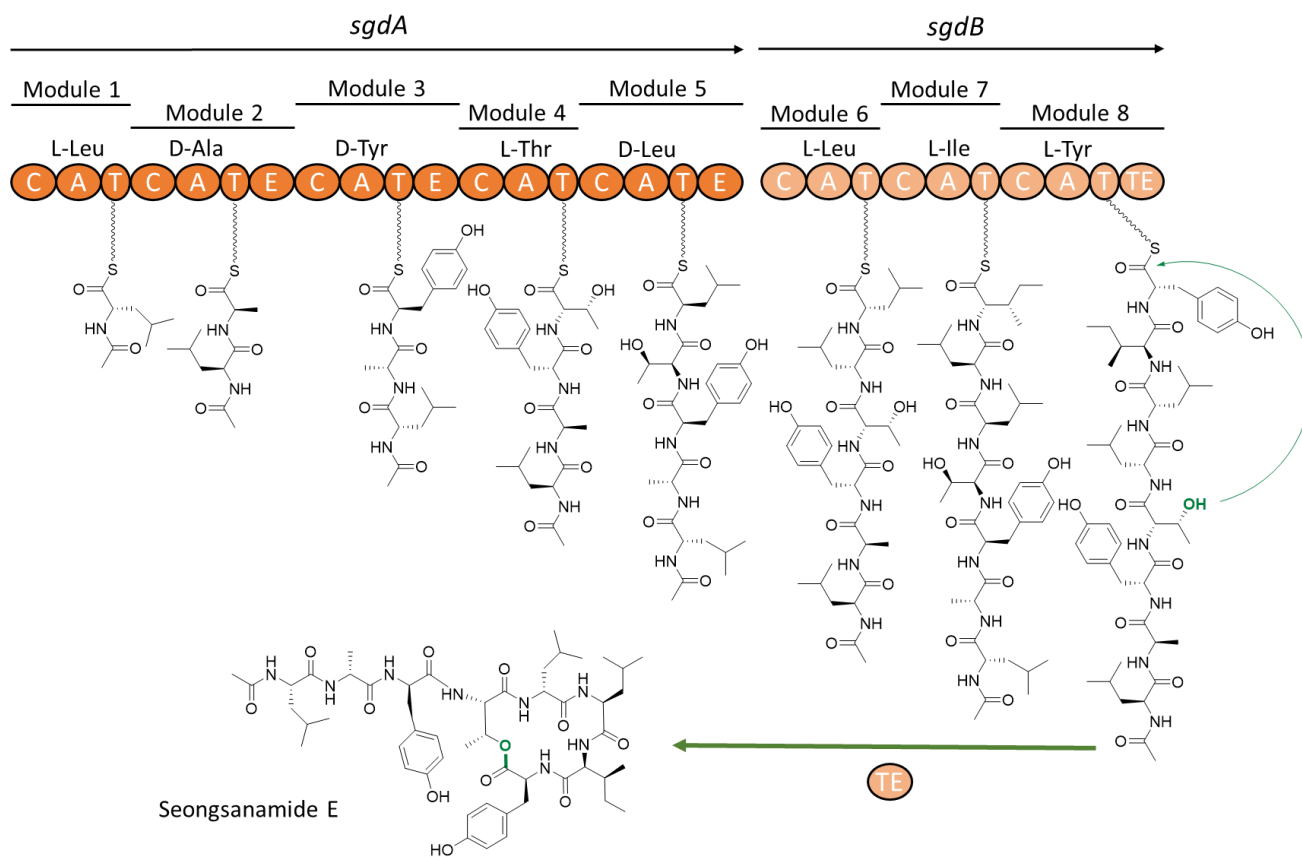
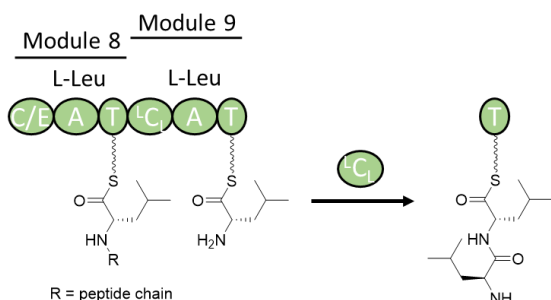
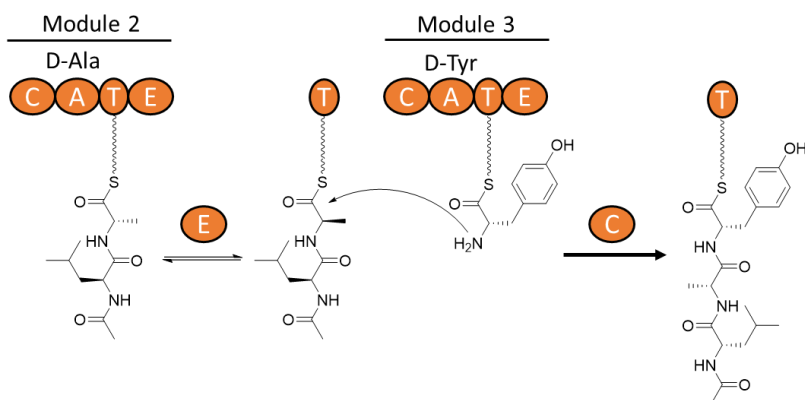
A**B****C**

Figure 1.4. Example of non-ribosomal peptide biosynthesis and specialized domains. (A) Shown is the seongsanamide (Sgd) NRPS pathway with T domain thioester intermediates and the product formed after all extension and tailoring cycles. (B) Example of $^L\text{C}_L$ catalyzed condensation between two L-configured amino acids in the gacamide A (Gac A) NRPS. (C) Epimerization in Sgd module 2 prior to condensation for peptide extension. Other tailoring domains are not shown but are described in Walsh et al. review.⁴¹

PKS and NRPS are two different strategies nature employs to build complex and biologically active macrocyclic compounds. While these megaenzymes are responsible for the construction and modifications of the linear precursors, the terminal TE domain is typically responsible for the macrocyclization reaction often required to generate a bioactive metabolite. These TE domains are nature's macrocyclization biocatalyst and are the focus for the research presented in this thesis.

1.2.3 TE domains classification, structure and mechanism

At the terminus of most PKS and NRPS pathways lies a TE (type I) domain responsible for the macrocyclic release of their cognate NP.⁴² These domains adopt an alternating $\beta/\alpha/\beta$ topology characteristic of the α/β -hydrolase fold (**Figure 1.5A**). They contain a highly twisted β -sheet forming a half barrel core with α -helices positioned around the sheet. Additional loop insertions fold into mobile subdomains called lids, which play a crucial role in shaping the active site and in regulating substrate accessibility (**Figure 1.5B**).^{36,43} Like other proteins in the α/β -hydrolase superfamily, they house a Ser-His-Asp catalytic triad required for enzyme activity.⁴⁴ In certain cases, the catalytic Ser is replaced with Cys, as observed in the polymyxin B and obafluorin forming TEs.^{45–48} While most TEs are found at the C-terminus of the biosynthetic machinery, separate standalone TE domains (type II) can also be found and are thought to prevent pathway stalling by hydrolyzing misprimed carrier protein (CP) domains.^{49–51} These type II TEs lie outside the scope of this review, but readers can find an in-depth discussion in Kotowska and Pawlik's review.⁵⁰

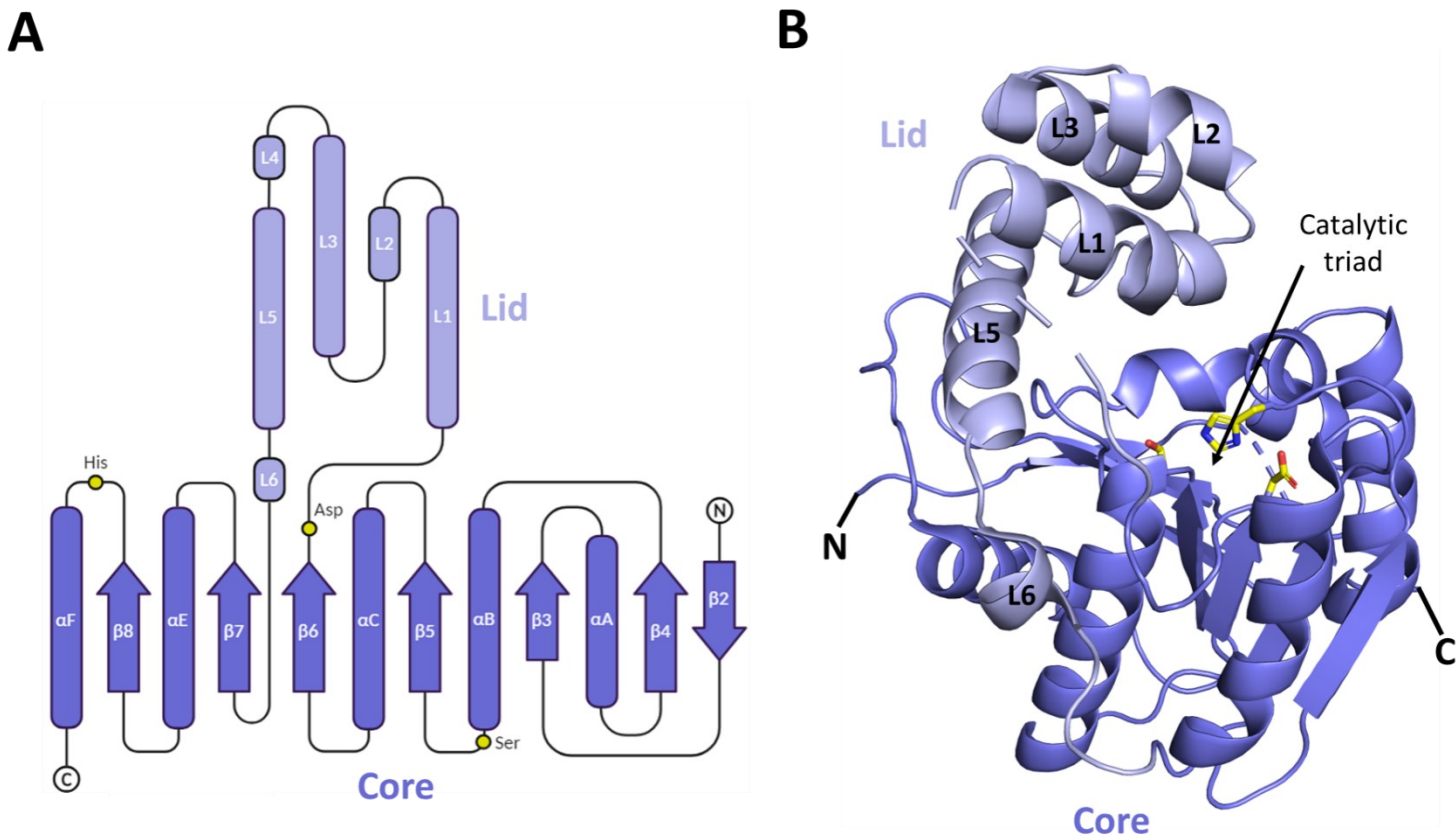


Figure 1.5. Representation of a type I TE structure (valinomycin). (A) Topology cartoon highlighting the catalytic residues (yellow dots), the lid domain and the $\beta/\alpha/\beta$ repeating core motif. Topology cartoon was adapted from Huguenin-Dezot et al.⁵² (B) Crystal structure of the WT valinomycin (VIm) TE domain (PDB: 6ECB, 1.7 Å resolution) highlighting the highly twisted β -sheet core, the catalytic triad (yellow sticks) and the mobile lid domain.

TE catalysis occurs through a two-step mechanism commonly seen in serine hydrolases,^{36,53,54} beginning with loading (step 1, **Figure 1.6**). In this step, the upstream ACP or T domain docks onto the TE, enabling the substrate tethered to the CP's PPT (substrate-PPT) to enter the TE active site.^{42,55–58} The catalytic Ser then attacks the electrophilic carbonyl of the substrate-PPT thioester, forming a charged tetrahedral intermediate stabilized by an oxyanion hole through H-bonding interactions with two backbone NH groups.⁵⁹ Concomitant to the nucleophilic attack, the catalytic Ser is deprotonated by His, acting as a general base, while Asp serves as an H-bonding partner to ensure the correct positioning of

the appropriate His tautomer. Departure of the PPT thiolate generates the acyl/peptidyl-TE intermediate, which undergoes the second step of the mechanism, release or off-loading (step 2, **Figure 1.6**). During this step, a nucleophile attacks the acyl/peptidyl-TE intermediate, releasing it from the TE domain. This step requires the catalytic His to be deprotonated to participate in nucleophile activation.^{59–62} Release of the acyl/peptidyl-TE intermediate is facilitated by either an intramolecular or intermolecular nucleophile. A charged tetrahedral intermediate forms following nucleophilic attack, collapsing to release the product and priming the TE for another catalytic cycle.^{42,63} The TE domain can either select for an intramolecular O- or N-nucleophile to release a macrocyclic product (macrocyclization) or water as an intermolecular nucleophile to release a linear free acid (hydrolysis). Furthermore, TEs can select for an endogenous O-, N-nucleophile from an upstream substrate-PPT to generate higher order oligomers (transesterification), which can then be released by hydrolysis or macrocyclization.³⁶ Some TEs can select for S- or C-nucleophiles for product release;⁶⁴ however, these will not be covered in this review.

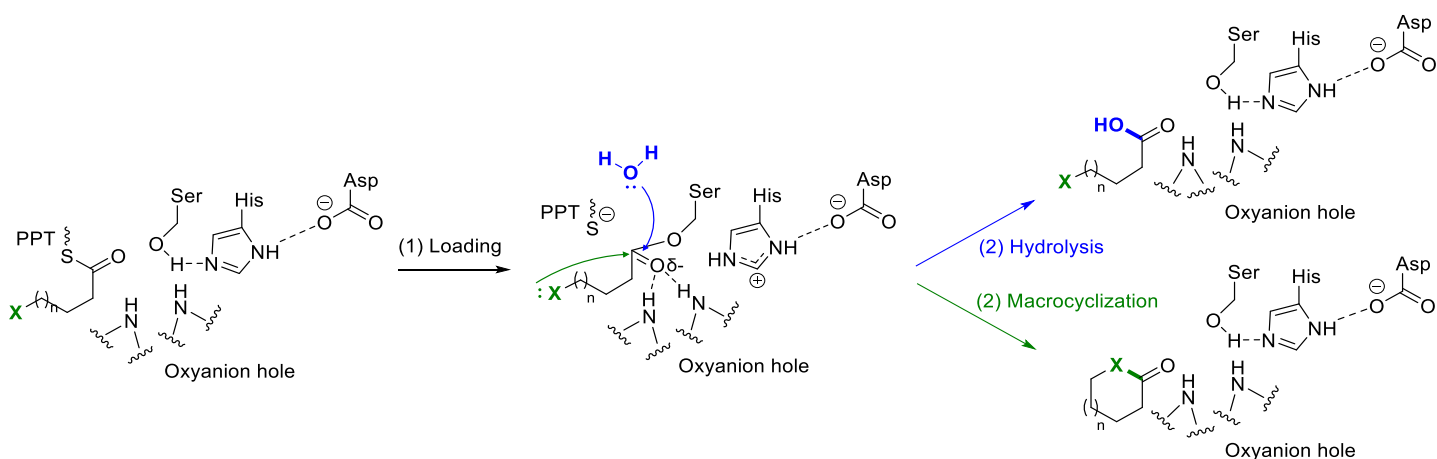


Figure 1.6. Canonical TE simplified catalytic mechanism. The catalytic triad of the TE is shown as Ser-His-Asp with the PPT shown as $\sim S$ and the oxyanion hole shown as $-NH-$. Dashed lines indicate H-bonds. Loading (1) of the substrate occurs via nucleophilic attack by the catalytic Ser. A conserved His-Asp dyad abstracts a proton from Ser for nucleophile activation. After loading, the TE can select for an intermolecular nucleophile like water (shown in blue) or an intramolecular nucleophile such as an alcohol or primary amine (depicted as X, shown in green) for hydrolytic or macrocyclic release (2) respectively.

1.2.4 Phylogenetic and structural determinants for TE off-loading

Many of the catalytic domains involved in PKS and NRPS biosynthesis such as KS, C, A and KR (to name a few) contain conserved sequence motifs that can enable predictions of substrate and or stereoselectivity using hidden Markov models.^{39,65–71} However, TE domains lack sequence motifs for predicting off-loading selectivity. Notably, phylogenetic analysis of TE sequences shows no clustering based on off-loading chemistry^{42,72} but rather cluster according to the biosynthetic machinery type and the producing organism's phylum. For example, TEs from bacterial type I PKS cluster separately from those found in NRPS. In addition, fungal PKS and *trans*AT-PKS TEs form their own sub-clusters separate from the bacterial type I PKS TEs (**Figure 1.7**), all of which is consistent with previous phylogenetic analysis.⁴² This suggests the absence of a single ancestral macrocyclization TE. As such, multiple different mechanisms controlling macrocyclization have likely evolved. These observations are in stark contrast to the more conserved hydrolyzing TEs found in FAS,^{73–75} which cluster based on substrate specificity such as alkyl chain length.⁷⁶

While phylogenetic analyses of TE sequences do not generally predict off-loading chemistry selectivity, analysis of highly similar TEs can provide some insight into the product formed. For instance, a very focused phylogenetic analysis of product release enzymes (PRE) in fungal PKS and non-ribosomal-polyketide (NR-PKS) hybrid pathways reveal well defined sub-clusters among TEs based on *f*-type or *s*-type cyclization chemistry.^{77,78} This clustering is evident in **Figure 1.7**, where Rdc (*f*-type) and Dhc (*s*-type) form their own small clusters. Similarly in bacterial type I PKS TEs, homologous sequences that cluster with strong branch support and short branch lengths often generate macrocycles of identical size as seen for the related 6-deoxyerythronolide B (DEBS) and pikromycin (PIK) TEs, which both produce 14-member macrolactones (**Figure 1.8**).^{79,80} However, this is not always the case, as seen with the nystatin (NysK) and

natamycin (SncS4) forming TEs. This further highlights the complex phylogenetic relationship of TEs and is consistent with previous studies.^{42,77,81}

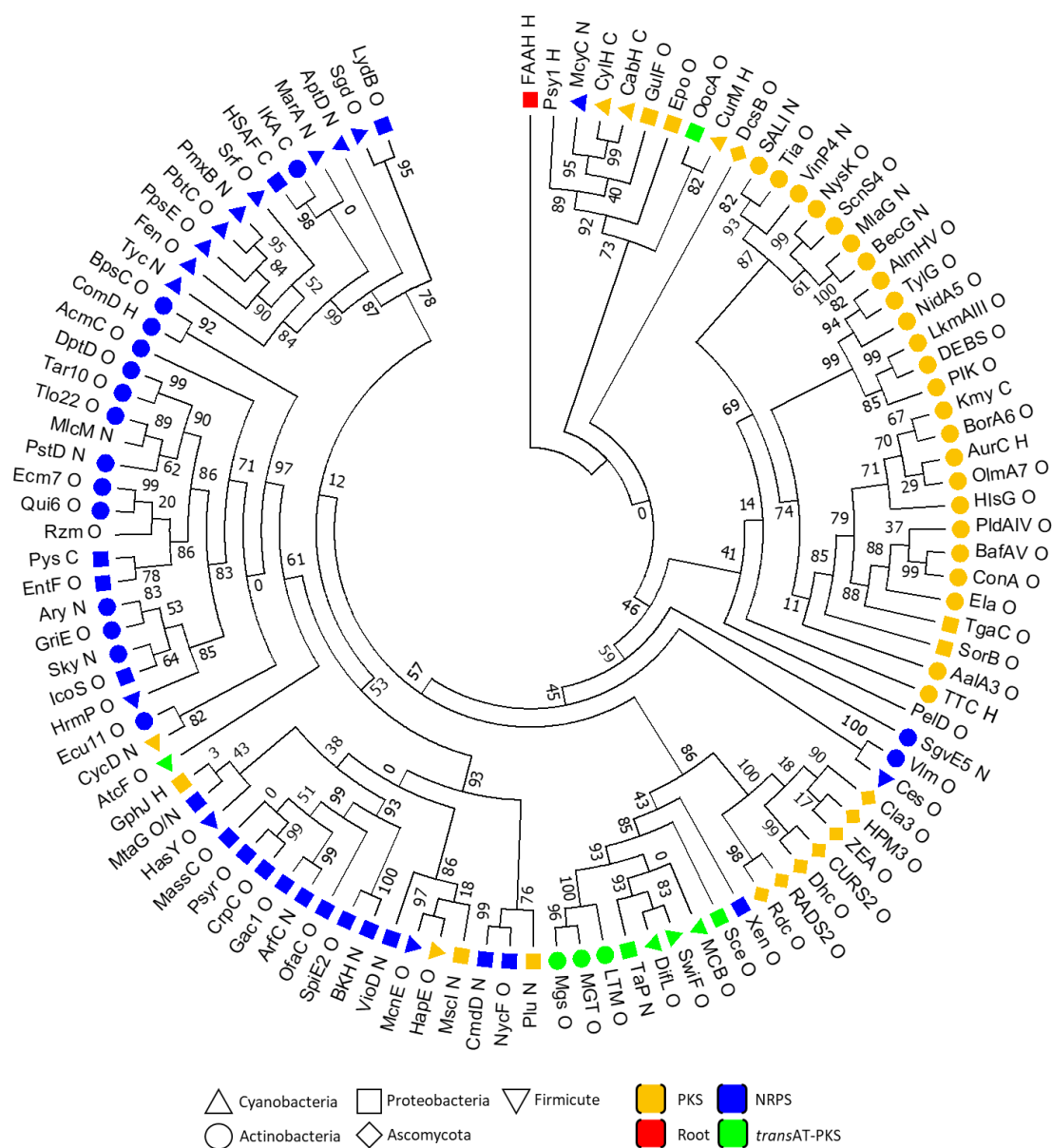


Figure 1.7. Protein sequence diversity dendrogram of macrocyclizing TE domains. 106 TE sequences were phylogenetically analyzed to construct the above dendrogram. Phylogenetic root used is the human fatty acid amide hydrolase (FAAH, red square). Branch strengths, TE abbreviations as well as the internal nucleophile are shown (O for alcohol, N for amine and C for carbon) while H denotes hydrolysis by water. The shapes display the producing organisms' phylum and the colour the type of pathway.

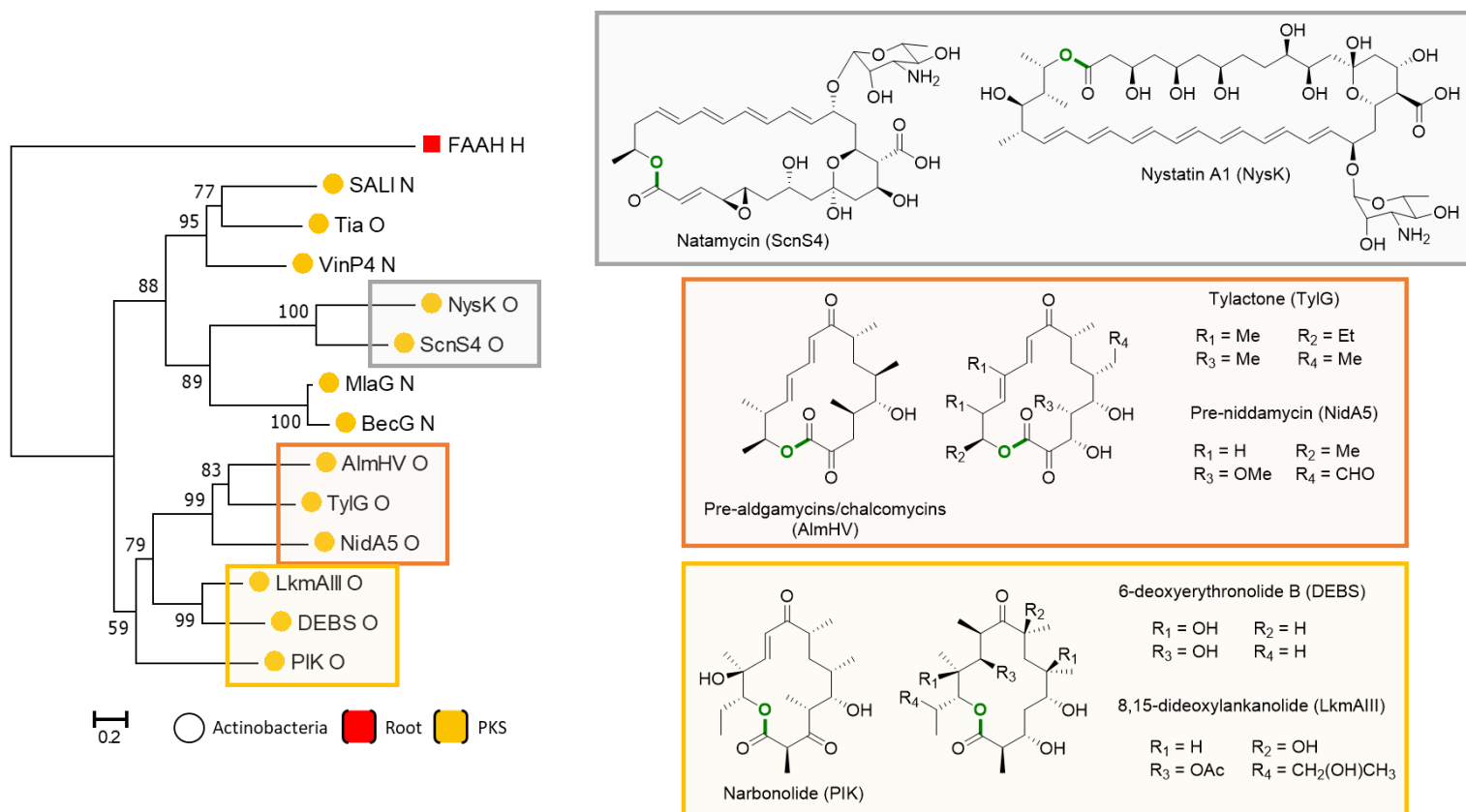


Figure 1.8. Dendrogram zoom-in of bacterial type I PKS clade. Shown is a small dendrogram comprising 13 TE sequences, representing a zoom-in of the dendrogram shown in **Figure 1.7**. Shown boxed are the macrocycle structures generated by the TEs with the bonds formed shown in green. Phylogenetic root used is the human fatty acid amide hydrolase (FAAH, red square). Branch strengths, TE abbreviations as well as the internal nucleophile are shown (O for alcohol or N for amine) while H denotes hydrolysis by water.

Phylogeny struggles to predict TE substrate selectivity and off-loading chemistry. While highly similar TE sequences likely mediate similar off-loading chemistries, effective methods to predict this are lacking and relies on experimental characterization.^{42,72} Some studies have explored active site structural features to predict off-loading chemistry, such as the active site volumes of hydrolyzing and macrocyclizing TEs.⁸² However, there was no observable correlation with active site volume and off-loading mechanism.⁴² As the lid has been implicated in controlling catalysis in serine hydrolases via significant conformational changes,^{52,83–86} it has also been investigated. Many TE crystal structures show

the lid in a distinct “open” conformation, suggesting a possible “closed” conformation to regulate active site access.^{52,87–89} Compelling evidence supporting this comes from trapped peptidyl-TE intermediate crystal structures of the bifunctional valinomycin (VIm) TE,⁵² which catalyzes transesterification and macrocyclization. During peptidyl growth, the lid remains in a dynamic open conformation, allowing active site access to upstream peptidyl monomers for transesterification. It is only in the trimerized peptidyl-TE crystal structure where the lid is observed in a closed conformation, blocking outside access to the active site and ensuring the correct oligomer is generated. In addition, the lid was shown to “curl” the peptidyl intermediate into a product-like conformation for macrocyclization.⁵² While the lid’s function in VIm and other bifunctional TEs seems clear,^{90–93} its role in monofunctional TEs like DEBS has yet to be elucidated.

At present, TE off-loading selectivity cannot be predicted based on structure, sequence, or from the context of its associated pathway.⁹⁴ There have been several attempts to experimentally determine the specific binding interactions that drive off-loading selectivity in TEs, such as those from the PIK and DEBS biosynthetic pathways. Crystal structures of both TEs in complex with phosphonate inhibitors suggests that hydrophobic packing interactions drives substrate binding and thus off-loading.^{95–97} This is consistent with a previous mutagenesis study on the DEBS-TE active site,⁹⁸ molecular dynamic simulations (MDS) and quantum mechanic/molecular mechanic (QM/MM) calculations of the PIK, DEBS and pimaricin (Pima) TEs, which suggest that hydrophobic interactions sequester the substrate away from bulk solvent, preventing hydrolysis, as the acyl-TE intermediate adopts a macrocyclic reactive conformation for macrocyclization.^{62,99,100} Unfortunately, the discrete substrate-TE interactions that drive off-loading remain elusive, presenting a valuable opportunity for future research to enhance our understanding in TE off-loading selectivity.

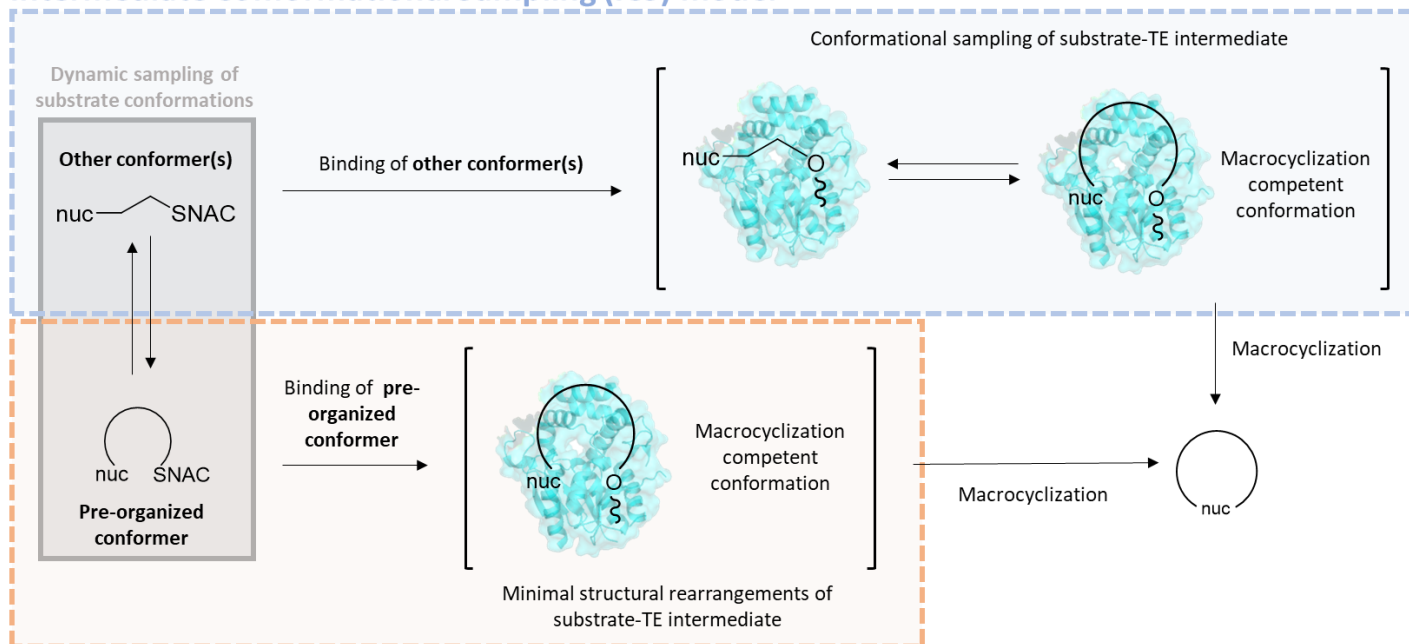
1.2.5 A rational framework for TE off-loading selectivity

The need for end-to-end contact in macrocyclization contributes to its high barrier, thus the conformation of the linear acyl-TE intermediate likely dictates the off-loading chemistry performed by the TE. MDS and QM/MM studies have illustrated this in the PIK-, DEBS- and Pima-TE, where non-productive substrate conformations favour hydrolysis.^{62,99,100} With this biochemistry in mind and the phylogenetic analysis suggesting repeated evolution of macrocyclizing TEs and thus multiple potential mechanisms controlling macrocyclization, we propose two archetypal models to rationalize TE off-loading selectivity (**Figure 1.9**). In the first model, an acyl-TE intermediate is formed after binding of a conformationally disordered substrate. This acylation step induces structural changes in the TE, akin to the induced fit mechanism,^{101,102} that restricts the substrate's conformational landscape until a product-like conformation is adopted for macrocyclization (**Figure 1.9**, blue rectangle). Vlm-TE exemplifies this model, where the lid undergoes a conformational shift in the trimerized peptidyl-TE crystal structure. The lid seals the active site and forms hydrophobic interactions with the peptidyl intermediate that limit the adoption of non-productive conformations, reduces water's access to the peptidyl-TE intermediate, and enables the substrate to “curl” into a macrocyclization competent conformation for catalysis.⁵² In this model hydrolysis occurs when water accesses the active site prior to the substrate obtaining a macrocyclization competent conformation. This model, termed intermediate conformational sampling (ICS), can thus rationalize the significant formation of hydrolysis product often reported for PKS TEs *in vitro*.^{95,97,100,103,104}

In the second model, the TE binds a pre-organized conformation of the substrate, requiring minimal structural rearrangement for macrocyclization (**Figure 1.9**, orange rectangle). This model is akin to the established conformational selection (CS) mechanism,¹⁰⁵ though with the substrate rather than the enzyme undergoing the conformational changes. Pre-organization has been observed for the tyrocidine (Tyc) and gramicidin (Grs) substrates, where a macrocyclization competent conformation is adopted in

solution prior to substrate loading.^{91,93,106–109} For instance, the Tyc-TE substrate pre-organization is facilitated by an L-Pro, which initiates a turn in the peptide substrate, and a discrete H-bond between the C-terminal L-Orn sidechain and the backbone carbonyl of an N-terminal aromatic residue (D-Phe in Tyc),⁹³ thus rationalizing the reported spontaneity in Tyc macrocyclization.¹⁰⁸ Because minimal structural rearrangement is needed for catalysis, this limits the opportunity for bulk solvent to hydrolyze the acyl-TE intermediate. Thus, these TEs are water insensitive, which is consistent with the low levels of hydrolysis seen for Tyc, Grs and other NRPS TEs *in vitro*.^{28,89,91,93,110–114}

Intermediate Conformational Sampling (ICS) model



Conformational Sampling (CS) model

Figure 1.9. Models for TE-mediated macrocyclization. *N*-acetylcysteamine (SNAC) thioester TE substrates dynamically sample conformations in solution (grey rectangle). Binding of other conformer(s) leads to a dynamic acyl-TE intermediate that sample conformations. This sampling eventually leads to the macrocyclization competent conformer for macrocyclization to occur in the ICS model (blue rectangle). Alternatively, binding of a pre-organized conformer leads to rapid macrocyclization in the CS model common to NRPS TEs (orange rectangle) with minimal competing hydrolysis reactivity. Note the “nuc” abbreviation found in all substrates denotes the internal nucleophile for macrocyclization and the term “other conformer(s)” refers to any non-cyclic conformation.

These two models differ in the timing of the substrate adopting a macrocyclization competent conformation. In the ICS model, this occurs as the acyl-TE intermediate undergoes conformational sampling. In the CS model, the substrate conformation is in equilibrium prior to binding and the enzyme selectively binds the macrocyclization competent conformation. We propose that TEs that suffer from significant hydrolytic activity, which are mostly PKS in origin such as in DEBS and PIK, likely follow this ICS model. The CS model is significantly less likely to suffer from hydrolysis, which is consistent with the low levels of observed hydrolysis in the Tyc-, Grs-TE, and other NRPS TEs. Thus, we propose that many NRPS TEs utilise this CS mechanism. Exceptions to these patterns exist, as demonstrated by the Vlm-TE⁵² that originates from an NRPS pathway yet aligns with our ICS model. Potentially the need to switch between transesterification chemistry for generation of the trimer followed by macrocyclization chemistry plays a role in the atypical reactivity of this TE.

1.2.6 Enzymes are useful biocatalytic tools for the synthesis of pharmaceuticals

Enzymes are becoming increasingly important in both industrial and academic organic syntheses. This is mainly due to their exquisite chemo-, regio-, stereo-selectivity and high reaction rates under mild conditions (ex: room temperature, neutral pH, atmospheric pressure).^{115,116} To showcase this, the following section of the review will cover two examples where enzymes have been used in the synthesis of important pharmaceuticals.

Sitagliptin is an important pharmaceutical for the treatment of type 2 diabetes and is sold by Merck as Januvia.¹¹⁷ Currently, over 240 million individuals worldwide are affected by this disease, and this number is expected to rise in the coming years. The first synthetic route to sitagliptin involved an asymmetric hydrogenation of an enamine precursor using a Rh-based chiral catalyst under high pressure.¹¹⁸ However, this step suffered from inadequate stereoselectivity and Rh contamination, necessitating additional purification steps that negatively impacted the overall yield. Following a few

iterations of enzyme engineering, a variant of an *R*-selective transaminase of a well characterized enzyme from *Arthrobacter* sp. was generated and able to convert 200 g/L of a diketone precursor into sitagliptin in 92% yield with over 99.95% enantiomeric excess (*ee*, **Figure 1.10**).¹¹⁹ Merck and Codexis jointly received the 2010 EPA Presidential Green Chemistry Challenge Award for development of this much greener and more efficient route to sitagliptin.

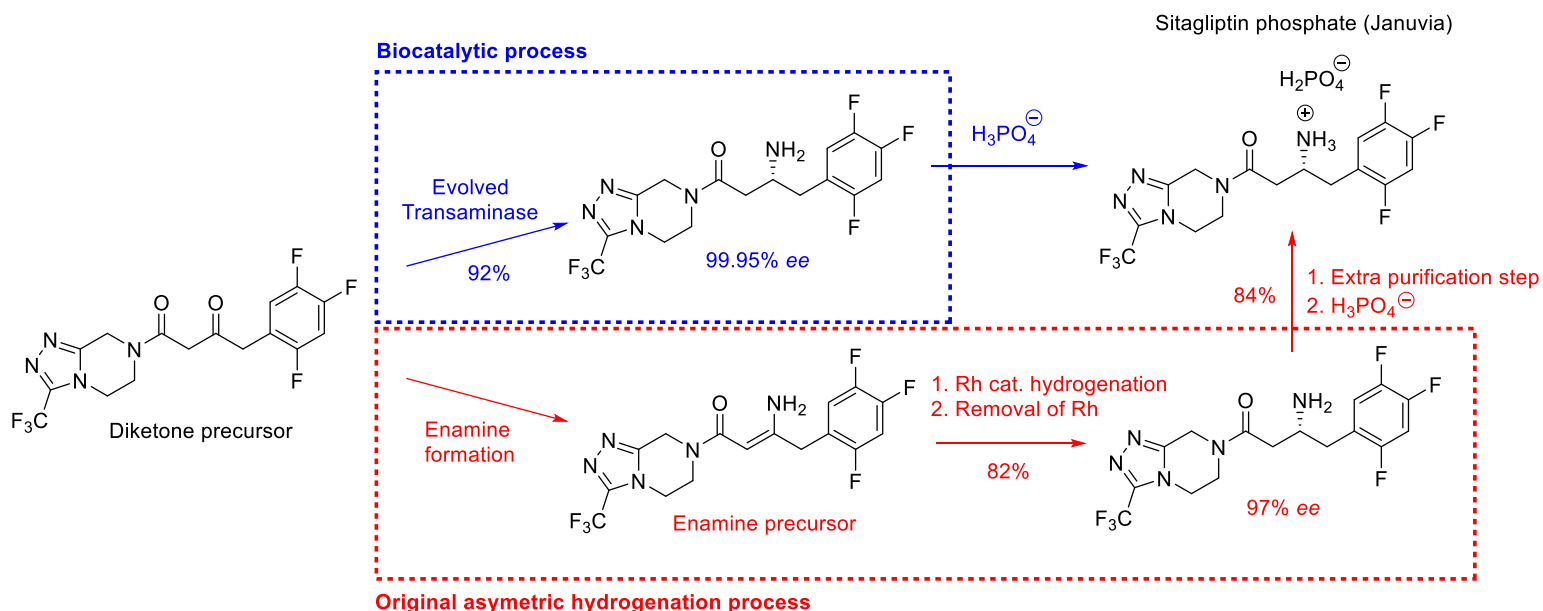


Figure 1.10. Biocatalytic versus original asymmetric hydrogenation process in sitagliptin synthesis. Shown in blue is the biocatalytic process created by Merck and Codexis compared to the original asymmetric hydrogenation process shown in red. Percentages below arrows indicate the yield. Figure was adapted from Savile et al.¹¹⁹

In 2017, Genentech and Roche collaborated in developing a ketoreductase (KRED) biocatalytic step for the synthesis of the Akt inhibitor ipatasertib, currently in phase 3 clinical trials for treatment of prostate cancer.¹²⁰ The process involved a highly diastereoselective reduction of a ketone to access a key chiral alcohol intermediate. Utilizing a commercially available and engineered KRED enzyme from Codexis, a one enzyme biocatalytic system using isopropyl alcohol as a terminal reductant for NADPH cofactor recycling proved instrumental in achieving this transformation (**Figure 1.11**).¹²¹ This biocatalytic approach

replaced a Ru-catalyzed asymmetric hydrogenation step, resulting in an improved yield, *ee* and eliminating the need for a challenging purification step to remove residual Ru.

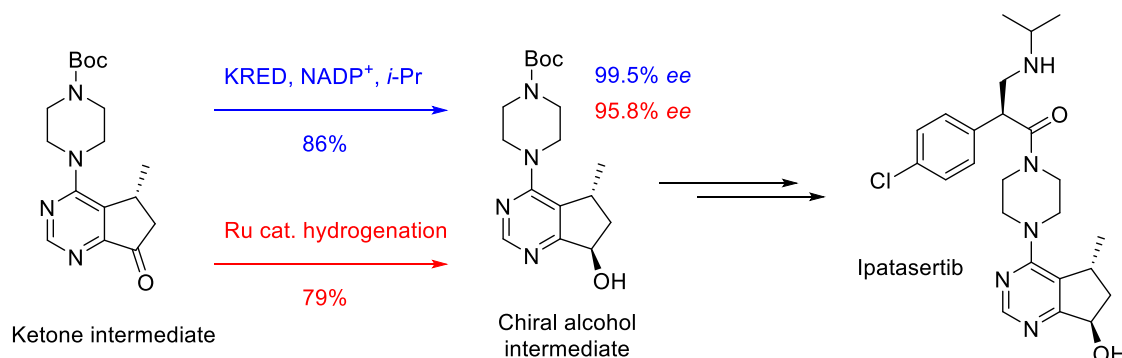


Figure 1.11. Biocatalytic versus Ru-catalyzed asymmetric hydrogenation in ipatasertib synthesis. Shown in blue is the use of a Codexis engineered KRED enzyme for production of a chiral alcohol intermediate. Shown in red is the alternative Ru-catalyzed asymmetric hydrogenation to reach the same chiral intermediate. Percentages below the arrows indicate the yield.

These examples clearly demonstrate that enzymes are powerful biocatalytic tools, that enable highly selective transformations under inert reaction conditions. Protein engineering techniques, such as directed evolution, have played a pivotal role by optimizing enzymes for better activity against non-natural substrates at high concentrations, for higher stability in reaction conditions and eliminating substrate or product inhibition.^{122–127} Enzymes across various classes have been fine-tuned for industrial applications¹²⁸ and their continued effectiveness is expected to persist for many years.

1.2.7 TE biocatalysts: a solution to the challenges faced in macrocycle synthesis

Generating macrocycles remains a synthetic challenge, even with the use of robust synthetic macrocyclization strategies.^{18,19} As shown in the previous section, enzymes can complement traditional synthetic methods to construct complex molecules in a chemo-, regio- and stereoselective way. PKS and NRPS biosynthetic machineries effectively generate macrocycles from linear intermediates that are densely functionalized with multiple nucleophiles, providing excellent chemo- and regioselectivity while

overcoming the challenges encountered in synthetic macrocyclization. Thus, PKS and NRPS provide an exceptional source for macrocyclization biocatalysts. In particular, the terminal TE domains in these pathways are typically responsible for catalyzing macrocyclization concomitant to product release from the modular biosynthetic machinery.⁴² While TEs are typically part of a megaenzyme complex, they can be excised and used as a stand-alone enzyme *in vitro* to convert linear thioester substrates into macrocycles. The first instance of a macrocyclizing TE used in this way was in the 1999 study of the DEBS-TE.¹²⁹ DEBS-TE remains one of the best studied TEs to date,^{72,97,103,129–136} displaying exceptional substrate tolerance for formation of the acyl-TE intermediate (i.e. loading).^{42,98,129,137} Subsequent biochemical studies on various TEs also demonstrated similar tolerance, suggesting substrate loading is promiscuous across TEs.^{72,91,98,111,112,129,137–140} While there are clear exceptions to this (ex: *f*-type forming Rdc-TE),^{141–143} increased substrate loading tolerance makes them promising macrocyclization biocatalysts. Herein we present two examples of the application of TEs to the synthesis of macrocycles, showcasing their utility as macrocyclization biocatalysts.

Non-ribosomal peptides typically adopt macrocyclic structures for increased binding affinity to their targets, decreased protease cleavage and increased membrane permeability.^{7–9,144} The use of non-ribosomal machinery also allows these peptides to incorporate non-proteogenic amino acids, expanding their structural diversity and leading to the generation of diverse macrocycles. A powerful use of a TE domain as a macrocyclization biocatalyst was reported by Trauger et al. in their study of the Tyc-TE. The authors were able to demonstrate the stand-alone Tyc-TE's ability to macrocyclize tyrocidine, the related non-ribosomal peptide gramicidin, and certain analogues.¹¹¹ A follow-up study further demonstrated its power as a biocatalyst, where it was shown to be immensely flexible at macrocyclizing a comprehensive library of tyrocidine analogues.¹¹⁰ This macrocyclization flexibility was due the peptide's ability to pre-organize into a cyclic conformation, so long as an aromatic residue, L-Pro and L-Orn were present at the

first, second and penultimate position of the peptide respectively (**Figure 1.12**).¹¹¹ Other analogues containing ester linkages, polyketide-like insertions and PEGA-linked peptides were all macrocyclized by the Tyc-TE,^{110,145} showcasing its power as a general peptide macrocyclization biocatalyst.

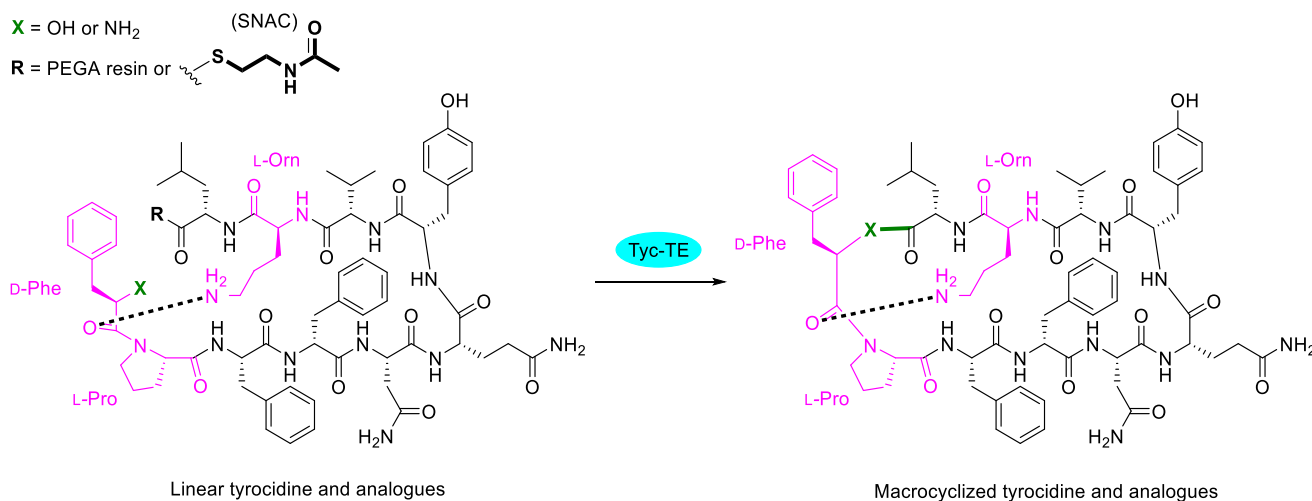


Figure 1.12. Macrocyclization of peptides catalyzed by the Tyc-TE. Simple scheme showcasing the broad substrate tolerance for Tyc-TE macrocyclization. Critical residues required for substrate pre-organization and macrocyclization are shown in pink and labelled. Note D-Phe can be replaced by any aromatic residue. All other residues not labelled can be modified to generate macrocyclic analogues. The dashed black line represents an H-bond. Figure was adapted from Trauger et al.⁹³

Studies on type I PKS TEs have shown they are substrate tolerant for loading.^{72,98,129,137,138} In particular, the DEBS-TE has shown the ability to load truncated analogues and form 6-, 8-, 12-, 14- and 16-membered macrolactones *in vivo* through domain swapping.^{130,134–136,146} Although intriguing, *in vivo* macrocyclization will not be discussed further as the focus of this thesis is on utilizing TEs as stand-alone *in vitro* biocatalysts. When taken out of its native context, the DEBS-TE still retains high loading tolerance towards synthetic substrates;^{42,98,129,137} however, it is highly regio- and stereoselective for off-loading *in vitro*. For instance, in the first example of the DEBS-TE macrocyclizing a synthetic non-native substrate, only the 14-membered macrolactone with an *R*-configured alcohol is formed (**Figure 1.13A**). Furthermore,

the TE is stereoselective even for elements distal to the nucleophile, exclusively producing the seco acid for a substrate bearing a distal *S*-configured alcohol (C11) even when the nucleophilic C13 alcohol is *R*-configured (**Figure 1.13B**).¹³³ As the *syn* and *anti* substrates undergo comparable chemical macrocyclization, their different TE-catalyzed chemistry is unlikely to be mediated by substrate pre-organization and is consistent with our proposed ICS model for macrocyclization. Additionally, when a related synthetic substrate was used to probe DEBS-TE off-loading selectivity, it was revealed that a significantly greater diversity of chemistry than expected could be accessed.⁷² The TE effectively catalyzed transesterification of the *R*-configured substrate, yielding linear and cyclic dimers, in addition to the expected hydrolysis and macrocyclization reactions (**Figure 1.13C**). This suggests PKS TEs can catalyze other reactions to access unique compounds; however, this has yet to be seen in other TEs. A critical concern affecting PKS TEs is the competing hydrolysis reactivity observed *in vitro*,^{95,97,100,103,104} restricting their use as macrocyclization biocatalysts.

TEs have proven themselves to be useful macrocyclization biocatalysts, playing a pivotal role in the total syntheses of various macrocyclic NP. The current repertoire of macrocyclization biocatalysts continues to expand with the characterization of new TEs, allowing chemists to access new macrocycles. While progress has been made in understanding TE off-loading selectivity, it still remains elusive.^{42,52,62,87,95–98,100} In addition, the competing hydrolysis reactivity observed in PKS TEs presents an issue that limits their use as macrocyclization biocatalysts. Protein engineering could prove useful in optimizing TEs to be selective for macrocyclization while increasing activity. This remains unaddressed in the literature and is discussed further in this review.

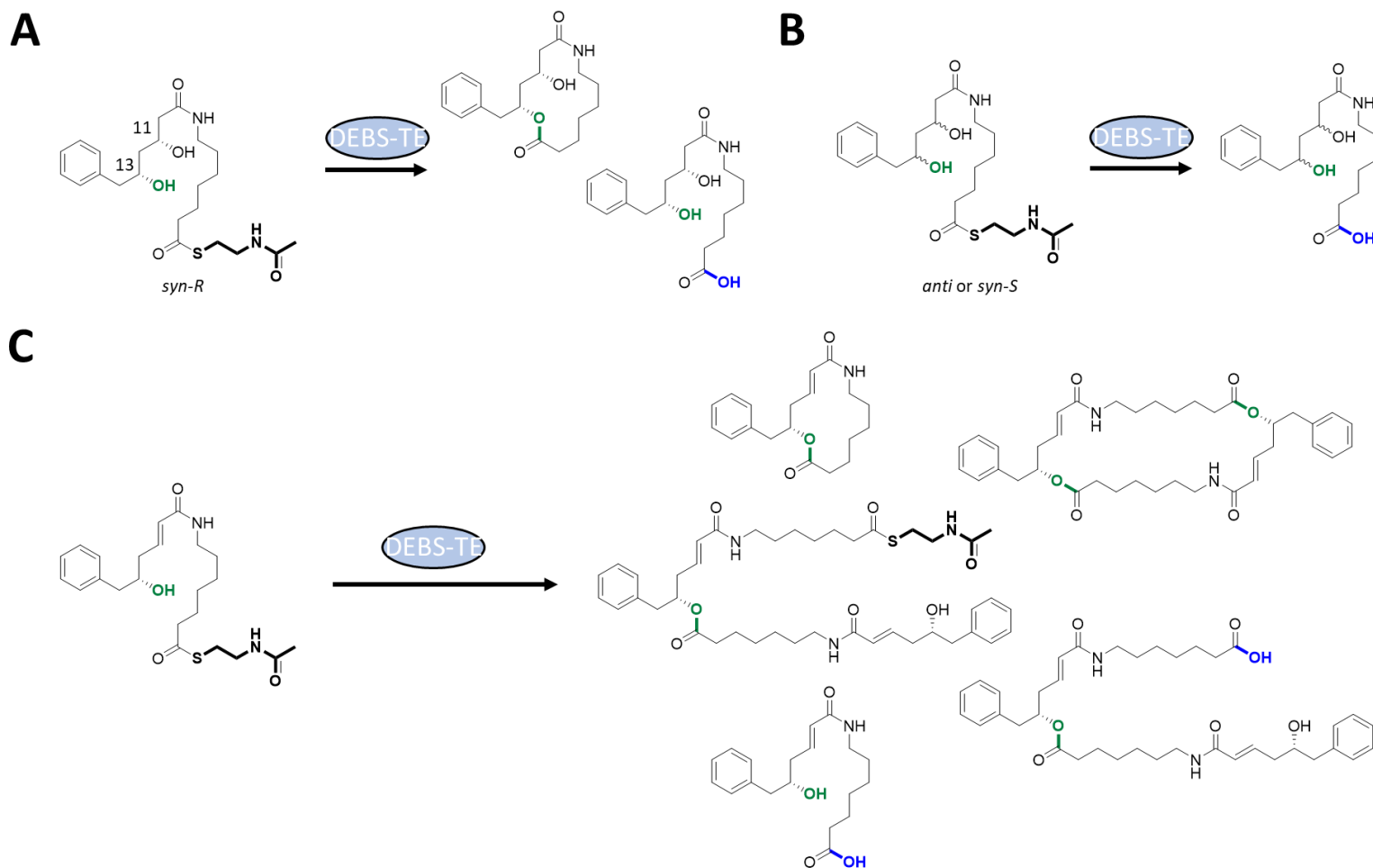


Figure 1.13. DEBS-TE acting on non-native substrates. (A) DEBS-TE producing both seco acid and 14-membered macrolactone with a *syn-R* diol substrate (B) compared to seco acid only with diastereomers of the *syn-R* diol substrate. (C) Diverse products formed by the DEBS-TE on an olefin substrate.

1.2.8 Protein engineering: a brief overview

Protein engineering has important implications for the development of biocatalysts to access complex and important pharmaceuticals. Protein engineering is typically performed either through rational protein design or by mimicking the natural process of evolution to “find” a protein with the desired function (i.e. directed evolution). Both methods have been thoroughly reviewed in the literature,^{147–150} and will briefly be addressed in this section. In rational protein design, researchers try to

predict sequences that will generate protein structures with the desired function.¹⁴⁸ This can be done *in silico* to yield non-natural sequences with unique structures (*de novo* protein design), by mutating a known protein sequence (protein redesign) or via both.^{148,151} Through this method, proteins from nature can be optimized or new proteins designed from “scratch” to meet a user defined function; however, it often necessitates an understanding of the protein subject’s structure-function relationship. Alternatively, directed evolution mimics the process of natural evolution¹⁵² through iterative cycling of two main steps: (1) diversification of a gene pool, typically performed by random mutagenesis and or gene shuffling,^{153–160} (2) followed by screening or selecting for mutants with the desired phenotype. This provides the distinct advantage of “finding” combinatorial mutations in a protein sequence that provide the desired phenotype, a feat not yet achievable through rational protein design. While both methods are different, they aren’t mutually exclusive and have been concurrently employed to produce enhanced biocatalysts in both industrial and academic settings. Consequently, protein engineering holds the potential to optimize TEs for their use as macrocyclization biocatalysts.

1.2.9 Towards engineering TEs, current work and limitations

To reach the same level of efficiency we expect in a biocatalyst, TEs would need to be engineered for optimal performance. However, this remains largely unaddressed in the literature. The most extensive efforts towards “engineering” TEs have involved mutagenesis studies on the structural determinants for loading or off-loading in DEBS,⁹⁸ PIK,^{100,138} enterobactin,¹⁶¹ surfactin,¹¹² zearalenone,¹⁴³ mycrocystin synthetase C,¹⁶² polymyxin B,⁴⁵ oocydin¹⁶³ and pimaricin TEs.¹⁶⁴ A noteworthy study by Koch et *al.* marked the first successful engineering of a TE. In this study, the stereoselective PIK-TE gained the ability to macrocyclize an epimerized non-native pentaketide substrate with improved kinetics over wild-type (WT) by substituting its catalytic Ser for a Cys (S→C).¹⁰⁰ While this is certainly promising, the S→C substitution

has yet to demonstrate the same effect in other PKS TEs and is consistent with the notion that TE-substrate interactions that drive off-loading are unique for each pair.

The limited understanding of discrete TE-substrate interactions that drive off-loading complicates TE rational protein design. Because of this, sites for mutagenesis are often chosen without clear guidance within the active site, potentially constraining the effectiveness of these studies in identifying a TE variant with the desired optimized function. More comprehensive mutagenesis and structural studies on TEs would be a welcome addition as these would better aid the selection of sites for mutagenesis in rational design and increase the prospects of finding optimized TE variants. On the other hand, directed evolution could provide an avenue to engineer TEs while bypassing the need for a detailed understanding of the discrete TE-substrate interactions that drive off-loading. Yet no directed evolution studies have been performed, presumably due to the challenge associated in developing a suitable selection method.^{150,152} This technique would thus rely on a screening method, such as LCMS based screens commonly seen in the pharmaceutical industry. This still poses a technical challenge, as LCMS based methods require a chromatographic separation step, increasing the time it takes to screen even the smallest of TE libraries. Despite this, LCMS based screens provide the advantage of utilizing automation for both library processing and data analysis. As the interest in using TEs as macrocyclization biocatalysts grows, we anticipate the emergence of TE engineering studies from both industrial and academic settings. These studies will play a crucial role in refining our current TE macrocyclization biocatalyst toolbox.

1.3 Summary and conclusions

The challenges associated with macrocyclization has plagued chemists for decades, driving the development of various synthetic methodologies to surmount them. It is only recently that researchers have explored the use of enzymes as macrocyclization biocatalysts, particularly focusing on the TEs from NRPS and PKS biosynthetic machineries. However, the mechanism by which TEs select for macrocyclization over other reactivities remains elusive. Two models were proposed to rationalize TE-mediated macrocyclization: the ICS and CS models. Common to PKS TEs is the water sensitive ICS model, where the TE constrains the conformational landscape of the acyl-TE intermediate until a reactive conformation conducive to macrocyclization is adopted. Conversely, the water inconsequential CS model, prevalent in NRPS TEs, involves substrate pre-organization to drive TE-mediated macrocyclization. Additional studies on TEs would be valuable in supporting and improving the proposed ICS and CS models. While TEs demonstrate remarkable substrate tolerance for loading, they are typically more stringent in macrocyclization. The NRPS Tyc-TE stands out as a highly tolerant macrocyclization biocatalyst, yet its substrate scope is restricted due to the stringent positional requirements for substrate pre-organization. PKS TEs like DEBS tend to be regio- and stereoselective and are often plagued with competing hydrolysis reactivity. Biochemical characterization of new macrocyclizing TEs can provide access to new macrocycles while engineering of pre-existing TEs can optimize our current repertoire. Both methods, while different, offer the ability to improve and expand the chemist's toolbox for macrocyclization.

In this thesis, the characterization and engineering of TEs is investigated for their use as macrocyclization biocatalysts. **Chapter 2** focuses on the *in vitro* biochemical characterization of two NRPS TEs, showcasing their utility as biocatalysts in the chemoenzymatic synthesis of two macrocyclic depsipeptides. **Chapter 3** focuses on the biochemical characterization of an unusual tandem NRPS TE domain to probe its function and substrate tolerance as a potential macrocyclization biocatalyst. **Chapter**

4 explores the engineering of a PKS TE for improved macrocyclization activity, employing a rational protein design approach and an LCMS based screen. These contributions expand the repertoire of TE macrocyclization biocatalysts and demonstrates their utility in macrocycle synthesis. Furthermore, this work marks a milestone in engineering a TE domain for improved activity and contributing to a deeper understanding in TE-mediated macrocyclization. Overall, the findings provide support for the proposed ICS and CS models, offering a rational framework for understanding TE-mediated macrocyclization.

1.4 Supplementary information

Construction of TE dendrograms. A total of 106 TE protein sequences were taken from the NCBI or PDB database, aligned by MUSCLE and trimmed using the MEGA7 software package.¹⁶⁵ Alignment was then phylogenetically analyzed by PhyML (v3.0)¹⁶⁶ using the BLOSUM 62 substitution matrix and assessed for branch support using the approximate likelihood ratio test (aLRT, SH-like).^{166,167} A second dendrogram was built identically as described above, except with 13 type I PKS TE protein sequences instead. Both trees used the human fatty acid amide hydrolase (FAAH) as phylogenetic root. See **Table S1** for a full list of the TE protein sequences used.

Table S1. Full list of TE protein sequences used for phylogenetic analyses. Protein sequences for macrocyclization TEs were downloaded from the NCBI or PDB databases for phylogenetic analysis. Sequences were selected if (1) the structure of the product is known and (2) the product is a macrocycle. A few hydrolyzing and tetramate forming TEs were selected to showcase how phylogeny can't predict off-loading selectivity. Shown in the table are the abbreviated names for each TE, the produced natural product, the accession or PDB code, the sequence range, the producing organism and its phylum. The nucleophile used for release is also shown, denoted either as H for hydrolysis by water, N, O or C for amine, alcohol or carbon nucleophiles respectively. TEs are coloured according to the pathway; blue for NRPS, yellow for PKS and green for *trans*AT-PKS. Red was used to denote the phylogenetic root. Note, the Sgd-TE sequence was taken from the genome of the producing organism and reverse-translated into a peptide sequence for the analysis (shown in bold).

Name	Natural Product	Nucleophile	Accession or PDB code	Sequence range	Organism	Phylum
FAAH	N/A	H	NP_001432.2	Full	<i>Homo sapiens</i>	Chordata
EntF	Enterobactin	O	CAD6014470	1066-1289	<i>Escherichia coli</i>	Proteobacteria
Sky	Skylamycin A	N	AEA30274.1	3657-3909	<i>Streptomyces</i> sp. Acta 2897	Actinobacteria
Vlm	Valinomycin	O	6ECC	Full	<i>Streptomyces tsusomaensis</i>	Actinobacteria
McyC	Mycrocystin	N	ACA83983.1	Full	<i>Planktothrix rubescens</i> No21-	Cyanobacteria
PpsE	Plipastatin	O	ASB65587.1	312-534	<i>Bacillus velezensis</i>	Firmicutes
Ces	Cereulide	O	ABD14712.1	2431-2681	<i>Bacillus cereus</i>	Firmicutes
Fen	Fengycin	O	2CGB	Full	<i>Bacillus subtilis</i>	Firmicutes

Srf	Surfactin	O	1JMK	Full	<i>Bacillus subtilis</i>	Firmicutes
Tyc	Tyrocidine	N	AAC45930.1	6254-6475	<i>Brevibacillus brevis</i>	Firmicutes
BKH	Burkholdac A	N	ABC38000.1	517-738	<i>Burkholderia thailandensis</i> E264	Proteobacteria
AcmC	Actinomycin D	O	ADG27359.1	3994-4244	<i>Streptomyces anulatus</i>	Actinobacteria
AptD	Anabaenopeptin	N	ACZ55946.1	1137-1391	<i>Anabaena</i> sp. 90	Cyanobacteria
ArfC	Arthrofactin A	N	BAC67536.1	5657-5909	<i>Pseudomonas</i> sp. MIS38	Proteobacteria
Ary	Arylomycin	N	EWS95124.1	4062-4279	<i>Streptomyces filamentosus</i> NRRL 11379	Actinobacteria
BpsC	Balhimycin	O	CAC48362.1	1622-1849	<i>Amycolatopsis balhimycina</i> DSM 5908	Actinobacteria
CrpC	Corpeptin	O	AHY86403.1	1892-2139	<i>Pseudomonas corrugata</i>	Proteobacteria
ComD	Isocomplestatin	H	AAK81827.1	1916-2164	<i>Streptomyces lavendula</i>	Actinobacteria
McnE	Cyanopeptolin	O	AAZ03554.1	1148-1412	<i>Microcystis</i> sp. NIVA-CYA 172/5	Cyanobacteria
CycD	Cyclomarin D	N	ABW00331.1	7525-7774	<i>Salinispora arenicola</i> CNS-205	Actinobacteria
DptD	Daptomycin	O	AAX31559.1	2130-2377	<i>Streptomyces filamentosus</i> NRRL 11379	Actinobacteria
Ecm7	Echinomycin	O	BAE98156.1	2905-2997	<i>Streptomyces lasalocidi</i>	Actinobacteria
PstD	Friulimicin	N	CAM56771.1	2116-2367	<i>Actinoplanes friuliensis</i>	Actinobacteria
GriE	Streptobactin	O	CBA63680.1	2218-2463	<i>Streptomyces</i> sp. ATCC 700974	Actinobacteria
HrmP	Hormaomycin	O	AEH41794.1	3213-3464	<i>Streptomyces griseoflavus</i>	Actinobacteria
LydB	Lysobactin	O	AEH59100.1	7971-8206	<i>Lysobacter</i> sp. ATCC 53042	Proteobacteria
MasC	Massetolide A	O	ABH06369.2	3513-3757	<i>Pseudomonas fluorescens</i> SS101	Proteobacteria
OfaC	Orfamide B	O	AFH75330.1	4628-4881	<i>Pseudomonas</i> sp. CMR12a	Proteobacteria
PbtC	Paenibacterin	O	AGM16414.1	2939-3157	<i>Paenibacillus</i> sp. OSY-SE	Firmicutes
PmxB	Polymyxin	N	ACA97577.1	871-1091	<i>Paenibacillus polymyxa</i>	Firmicutes
Qui6	Quinomycin	O	AET98905.1	2915-3145f	<i>Streptomyces griseovariabilis</i> subsp. <i>bandungensis</i>	Actinobacteria

Psyr	Syringomycin	O	AAV37655.1	13279-13530	<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	Proteobacteria
Tar10	Taromycin	O	AHH53508.1	2134-2378	<i>Saccharomonospora</i> sp. CNQ490	Actinobacteria
Tlo22	Telomycin	O	ALV86868.1	2145-2399	<i>Streptomyces canus</i>	Actinobacteria
MlcM	Malacidin	N	ARU08075.1	2028-2280	Unknown bacteria	
MarA	Marthiapeptide A	N	RAT94090.1	5696-5805	<i>Brevibacillus</i> sp. Leaf182	Firmicutes
Ecu11	Ecumicin	O	AIW58892.1	14997-15251	<i>Nonomuraea</i> sp. MJM5123	Actinobacteria
HasY	Hassallidin E	O	CZT62794.1	2563-2825	<i>Planktothrix sarta</i> PCC 8927	Cyanobacteria
VioD	Vioprolide A	N	AWI62629.1	1524-1775	<i>Cystobacter</i> sp.	Proteobacteria
Xen	Xenematide	O	WP_013184596.1	4706-4807	<i>Xenorhabdus nematophila</i>	Proteobacteria
IcoS	Icosalide	O	AYA44686.1	4750-4996	<i>Burkholderia gladioli</i>	Proteobacteria
Gac1	Gacamide A	O	ABA73956.1	4615-4867	<i>Pseudomonas fluorescens</i> PfO-1	Proteobacteria
Pys	Pyreudiones	C	WP_064118616.1	1043-1274	<i>Pseudomonas fluorescens</i>	Proteobacteria
Rzm	Rhizomide	O	WP_013428324	7387-end	<i>Paraburkholderia rhizoxinica</i> str. HKI 454	Pseudomonadota
Sgd	Seongsanamide E	O	CP018197.1	2629387 - 2630191	<i>Bacillus safensis</i> KCTC 12796BP	Firmicutes
SpiE2	Spiruchostatin A	O	AFR69337.1	517-739	<i>Pseudomonas</i> sp. Q71576	Proteobacteria
NycF	Nannocystin A	O	ALD82526.1	2638-2895	<i>Nannocystis</i> sp. MB1016	Proteobacteria
CmdD	Chondramide A	N	CAJ46692.1	3087-3344	<i>Chondromyces crocatus</i>	Proteobacteria
SgvE5	Griseoviridin	N	AGN74896.1	1690-1948	<i>Streptomyces griseoviridis</i>	Actinobacteria
IKA	Ikarugamycin	C	AJD77023.1	2910-3117	<i>Streptomyces</i> sp. ZJ306	Actinobacteria
MtaG	Myxothiazol	O or N	AAF19815.1	1482-1730	<i>Stigmatella aurantiaca</i> DW4/3-1	Proteobacteria
HSAF	Dihydromaltophilin	C	ABL86391.1	2907-3116	<i>Lysobacter enzymogenes</i>	Proteobacteria
Cla3	Cladosporin	O	A0A125R003.1	1805-1950	<i>Cladosporium cladosporioides</i>	Ascomycota
RADS2	Radicicol	O	KAF2180792.1	1786-2049	<i>Zopfia rhizophila</i> CBS207.26	Ascomycota
CURS2	Dehydrocurvularin	O	L7XAV2.1	1811-2058	<i>Aspergillus terreus</i>	Ascomycota
Dhc	Dehydrocurvularin	O	A0A0N7D745.1	1812-2057	<i>Alternaria cinerariae</i>	Ascomycota
HPM3	Hypothemycin	O	B3FWS8.1	1761-1951	<i>Hypomyces subiculosus</i>	Ascomycota

ZEA	Zeralenone	O	A0A098D6U0.1	1770-1932	<i>Fusarium graminearum</i> PH-1	Ascomycota
Rdc	Radicicol	O	C5H882.1	1854-2123	<i>Floropilus chiversii</i>	Ascomycota
DcsB	Decastrictine C1	O	7D78	Full	<i>Beaveria bassiana</i> ARSEF2860	Ascomycota
CurM	Curacin	H	3QIT_A	Full (chain A)	<i>Moorea producens</i> 19L	Cyanobacteria
Epo	Epothilone	O	AAF62885.1	2197-2423	<i>Sorangium cellulosum</i>	Proteobacteria
Plu	Luminmycin A	N	CAE14173.1	3888-4149	<i>Photorhabdus laumondii</i> subsp. TTO1	Proteobacteria
TgaC	Thuggacin A	O	ADH04641.1	6445-6656	<i>Sorangium cellulosum</i>	Proteobacteria
HapE	Hapalosin	O	WP_035122279	1545-1810	<i>Fischerella</i> sp. PCC 9431	Cyanobacteria
SAL1	Salinilactam	N	ABP55220.1	3501-3709	<i>Salinispora tropica</i> CNB-440	Actinobacteria
BorA6	Borrelidin	O	CAE45672.1	1524-1730	<i>Streptomyces parvulus</i>	Actinobacteria
GphJ	Gephyronic acid	H	AHA38203.1	1560-1808	<i>Archangium violaceum</i> Cb vi76	Proteobacteria
GulF	Gulmirecin A	O	AIT55263.1	2169-2403	<i>Pyxidicoccus fallax</i>	Proteobacteria
HlsG	Halstoctacosanolide A	O	BAF02927.1	1889-2098	<i>Streptomyces halstedii</i>	Actinobacteria
LkmAIII	Lankamycin	O	BAC76491.1	3084-3290	<i>Streptomyces rochei</i>	Actinobacteria
MlaG	ML-449	N	ACO94498.1	1774-1988	<i>Streptomyces</i> sp. MP39-85	Actinobacteria
ScnS4	Natamycin	O	ADX66459.1	1802-2015	<i>Streptomyces chattanoogensis</i>	Actinobacteria
NysK	Nystatin A1	O	AAF71768.1	1850-2058	<i>Streptomyces noursei</i> ATCC 11455	Actinobacteria
PldAIV	Pladienolide B	O	BAH02271.1	1592-1804	<i>Streptomyces platensis</i>	Actinobacteria
SorB	Soraphem A	O	AAA79984.2	8598-8807	<i>Sorangium cellulosum</i>	Proteobacteria
TTC	Tautomycetin	H	ABW96540.1	5548-5766	<i>Streptomyces spiroverticillatus</i>	Actinobacteria
TylG	Tylactone	O	AAB66508.1	1607-1819	<i>Streptomyces fradiae</i>	Actinobacteria
CylH	Cylindrocyclophane	C	AFV96142.1	1990-2205	<i>Cylindrospermum licheniforme</i> UTEX B 2014	Cyanobacteria
Kmy	Kendomycin	C	CAQ52624.1	3204-3419	<i>Streptomyces violaceoruber</i>	Actinobacteria
CabH	Carbamidocyclophane	C	AMB48441.1	1843-2055	<i>Nostoc</i> sp. CAVN2	Cyanobacteria

AlmHV	Aldgamycin	O	ANC94962.1	1142-1354	<i>Streptomyces</i> sp. A1(2016)	Actinobacteria
AalA3	Actinoallolide A	O	BBA66513.1	4074-4283	<i>Actinoallomurus fulvus</i>	Actinobacteria
VinP4	Vicenistatin	N	BAD08360.1	3579-3804	<i>Streptomyces halstedii</i>	Actinobacteria
BecG	BE-14106	N	ACO94470.1	1768-1982	<i>Streptomyces</i> sp. DSM 21069	Actinobacteria
Msci	Microsclerodermin M	N	AHB82059.1	2634-2899	<i>Sorangium cellulosum</i>	Proteobacteria
BafAV	Bafilomycin A	O	ADC79620.1	1883-2093	<i>Streptomyces lohii</i>	Actinobacteria
ConA	Concanamycin A	O	AAZ94391.1	1834-2045	<i>Streptomyces neyagawaensis</i>	Actinobacteria
NidA5	Niddamycin	O	AAC46028.1	1641-1856	<i>Streptomyces caeestis</i>	Actinobacteria
AurC	Aureothin	H	CAE02606.1	1986-2197	<i>Streptomyces thioluteus</i>	Actinobacteria
OlmA7	Oligomycin	O	BAB69196.1	3408-3608	<i>Streptomyces avermitilis</i>	Actinobacteria
Tia	Tiacumicin B	O	ADU86004.1	3242-3457	<i>Dactylosporangium aurantiacum</i> subsp. <i>Hamdenensis</i>	Actinobacteria
Ela	Elaiohylin	O	ATL87338.1	1872-2085	<i>Streptomyces malaysiensis</i>	Actinobacteria
PIK	Pikromycin	O	2HFK		<i>Streptomyces venezuelae</i>	Actinobacteria
DEBS	6-deoxyerythronolide B	O	1MO2		<i>Saccharopolyspora erythraea</i>	Actinobacteria
Psy1	Psymberin	H	ADA82585.1	12379-12644	Unknown bacteria	
OocA	Oocydin A	O	AFX60300.1		<i>Serratia marcescens</i>	Proteobacteria
Tap	Myxovirescins A1	N	ABF88102.1	2756-3035	<i>Myxococcus xanthus</i> DK 1622	Proteobacteria
DifL	Difficidin	O	CAG23983.1	1755-1959	<i>Bacillus velezensis</i>	Firmicutes
Mgs	Migrastatin	O	ACY01392.1	1646-1923	<i>Streptomyces platensis</i> subsp. <i>rosaceus</i>	Actinobacteria
Sce	Etnangien	O	CAN93352.1	5663-5939	<i>Sorangium cellulosum</i> So ce56	Proteobacteria
AtcF	Anthracycline	O	ALD83688.1	3167-3440	<i>Streptomyces</i> sp. CNH365	Actinobacteria
SwiF	Swinholide A	O	ATX68127.1	7702-7872	<i>Nostoc</i> sp. UHCC 0450	Cyanobacteria
LTM	Lactimidomycin	O	ACY01402.1	1796-2089	<i>Streptomyces amphibiosporus</i>	Actinobacteria
MGT	Migrostatin	O	ACY01392.1	1646-1923	<i>Streptomyces platensis</i> subsp. <i>rosaceus</i>	Actinobacteria
MCB	Macrobrevin	O	RAT98517.1	4474-4740	<i>Brevibacillus</i> sp. Leaf182	Firmicutes
PeID	Peluroside A	O	DAC82022.1	4412-4691	Unknown bacteria	

1.5 References

- (1) Beutler, J. A. Natural Products as a Foundation for Drug Discovery. *Curr. Protoc. Pharmacol.* **2019**, 86 (e46), 1–28. <https://doi.org/10.1002/0471141755.ph0911s46>.
- (2) Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Orhan, I. E.; Banach, M.; Rollinger, J. M.; Barreca, D.; Weckwerth, W.; Bauer, R.; Bayer, E. A.; Majeed, M.; Bishayee, A.; Bochkov, V.; Bonn, G. K.; Braidy, N.; Bucar, F.; Cifuentes, A.; D’Onofrio, G.; Bodkin, M.; Diederich, M.; Dinkova-Kostova, A. T.; Efferth, T.; El Bairi, K.; Arkells, N.; Fan, T. P.; Fiebich, B. L.; Freissmuth, M.; Georgiev, M. I.; Gibbons, S.; Godfrey, K. M.; Gruber, C. W.; Heer, J.; Huber, L. A.; Ibanez, E.; Kijjoo, A.; Kiss, A. K.; Lu, A.; Macias, F. A.; Miller, M. J. S.; Mocan, A.; Müller, R.; Nicoletti, F.; Perry, G.; Pittalà, V.; Rastrelli, L.; Ristow, M.; Russo, G. L.; Silva, A. S.; Schuster, D.; Sheridan, H.; Skalicka-Woźniak, K.; Skaltsounis, L.; Sobarzo-Sánchez, E.; Bredt, D. S.; Stuppner, H.; Sureda, A.; Tzvetkov, N. T.; Vacca, R. A.; Aggarwal, B. B.; Battino, M.; Giampieri, F.; Wink, M.; Wolfender, J. L.; Xiao, J.; Yeung, A. W. K.; Lizard, G.; Popp, M. A.; Heinrich, M.; Berindan-Neagoe, I.; Stadler, M.; Daglia, M.; Verpoorte, R.; Supuran, C. T. Natural Products in Drug Discovery: Advances and Opportunities. *Nat. Rev. Drug Discov.* **2021**, 20 (3), 200–216. <https://doi.org/10.1038/s41573-020-00114-z>.
- (3) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, 79 (3), 629–661. <https://doi.org/10.1021/acs.jnatprod.5b01055>.
- (4) Jimenez, D. G.; Poongavanam, V.; Kihlberg, J. Macrocycles in Drug Discovery - Learning from the Past for the Future. *J. Med. Chem.* **2023**, 66 (8), 5377–5396. <https://doi.org/10.1021/acs.jmedchem.3c00134>.
- (5) Laureti, L.; Song, L.; Huang, S.; Corre, C.; Leblond, P.; Challis, G. L.; Aigle, B. Identification of a Bioactive 51-Membered Macrolide Complex by Activation of a Silent Polyketide Synthase in *Streptomyces Ambofaciens*. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 108 (15), 6258–6263. <https://doi.org/10.1073/pnas.1019077108>.
- (6) Su, L.; Hôtel, L.; Paris, C.; Chepkirui, C.; Brachmann, A. O.; Piel, J.; Jacob, C.; Aigle, B.; Weissman, K. J. Engineering the Stambomycin Modular Polyketide Synthase Yields 37-Membered Mini-Stambomycins. *Nat. Commun.* **2022**, 13 (1), 1–16. <https://doi.org/10.1038/s41467-022-27955-z>.
- (7) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The Exploration of Macrocycles for Drug Discovery - An Underexploited Structural Class. *Nat. Rev. Drug Discov.* **2008**, 7 (7), 608–624. <https://doi.org/10.1038/nrd2590>.
- (8) Phan, J.; Shi, Z. D.; Burke, T. R.; Waugh, D. S. Crystal Structures of a High-Affinity Macrocyclic Peptide Mimetic in Complex with the Grb2 SH2 Domain. *J. Mol. Biol.* **2005**. <https://doi.org/10.1016/j.jmb.2005.08.037>.
- (9) Appavoo, S. D.; Huh, S.; Diaz, D. B.; Yudin, A. K. Conformational Control of Macrocycles by Remote Structural Modification. *Chem. Rev.* **2019**. <https://doi.org/10.1021/acs.chemrev.8b00742>.
- (10) Doak, B. C.; Zheng, J.; Dobritsch, D.; Kihlberg, J. How Beyond Rule of 5 Drugs and Clinical Candidates Bind to Their Targets. *J. Med. Chem.* **2016**, 59 (6), 2312–2327. <https://doi.org/10.1021/acs.jmedchem.5b01286>.
- (11) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macrocycles. *Nat. Chem. Biol.* **2014**, 10 (9), 723–731. <https://doi.org/10.1038/nchembio.1584>.
- (12) Dougherty, P. G.; Qian, Z.; Pei, D. Macrocycles as Protein-Protein Interaction Inhibitors. *Biochem. J.* **2017**, 474 (7), 1109–1125. <https://doi.org/10.1042/BCJ20160619.Macrocycles>.
- (13) Giordanetto, F.; Kihlberg, J. Macrocyclic Drugs and Clinical Candidates: What Can Medicinal

- Chemists Learn from Their Properties? *J. Med. Chem.* **2014**, *57* (2), 278–295.
<https://doi.org/10.1021/jm400887j>.
- (14) Mallinson, J.; Collins, I. Macrocycles in New Drug Discovery. *Future Med. Chem.* **2012**, *4* (11), 1409–1438. <https://doi.org/10.4155/fmc.12.93>.
 - (15) Cummings, M. D.; Sekharan, S. Structure-Based Macrocycle Design in Small-Molecule Drug Discovery and Simple Metrics to Identify Opportunities for Macrocyclization of Small-Molecule Ligands. *J. Med. Chem.* **2019**, *62* (15), 6843–6853.
<https://doi.org/10.1021/acs.jmedchem.8b01985>.
 - (16) Garcia Jimenez, D.; Poongavanam, V.; Kihlberg, J. Macrocycles in Drug Discovery—Learning from the Past for the Future. *J. Med. Chem.* **2023**, *66* (8), 5377–5396.
<https://doi.org/10.1021/acs.jmedchem.3c00134>.
 - (17) Marsault, E.; Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* **2011**, *54* (7), 1961–2004.
<https://doi.org/10.1021/jm1012374>.
 - (18) Mortison, J. D.; Sherman, D. H. Frontiers and Opportunities in Chemoenzymatic Synthesis. *J. Org. Chem.* **2010**, *75* (21), 7041–7051. <https://doi.org/10.1021/jo101124n>.
 - (19) Parenty, A.; Moreau, X.; Niel, G.; Campagne, J. M. Update 1 of: Macrolactonizations in the Total Synthesis of Natural Products. *Chem. Rev.* **2013**, *113* (1), PR1–PR40.
<https://doi.org/10.1021/cr300129n>.
 - (20) Fürstner, A. Lessons from Natural Product Total Synthesis: Macrocyclization and Postcyclization Strategies. *Acc. Chem. Res.* **2021**, *54* (4), 861–874.
<https://doi.org/10.1021/acs.accounts.0c00759>.
 - (21) Woodward, R. B.; Au-Yeung, B. W.; Balaram, P.; Browne, L. J.; Ward, D. E.; Card, P. J.; Chen, C. H.; Chenevert, R. B.; Fliri, A.; Frobels, K.; Gais, H.-J.; Garratt, D. G.; Hayakawa, K.; Heggie, W.; Hesson, D. P.; Hoppe, D.; Hoppe, I.; Hyatt, J. A.; Ikeda, D.; Jacobi, P. A.; Kim, K. S.; Kobuke, Y.; Kojima, K.; Krowicki, K.; Lee, V. J.; Leutert, T.; Malchenko, S.; Martens, J.; Matthews, R. S.; Ong, B. S.; Press, J. B.; Rajan Babu, T. V.; Rousseau, G.; Sauter, H. M.; Suzuki, M.; Tatsua, K.; Tolbert, L. M.; Truesdale, E. A.; Uchida, I.; Ueda, Y.; Uyehara, T.; Vasella, A. T.; Vladuchick, W. C.; Wade, P. A.; Williams, R. M.; Wong, H. N.-C. Asymmetric Total Synthesis of Erythromycin. 2. Synthesis of an Erythronolide A Lactone System. *J. Am. Chem. Soc.* **1981**, *103* (7), 3213–3215.
 - (22) Inanaga, J.; Hirata, K.; Saeki, H.; Katsuki, T.; Yamaguchi, M. A Rapid Esterification by Means of Mixed Anhydride and Its Application to Large-Ring Lactonization. *Bull. Chem. Soc. Jpn.* **1979**, *52* (7), 1989–1993. <https://doi.org/10.1246/bcsj.52.1989>.
 - (23) Hikota, M.; Sakurai, Y.; Horita, K.; Yonemitsu, O. Synthesis of Erythronolide A via a Very Efficient Macrolactonization under Usual Acylation Conditions with the Yamaguchi Reagent. *Tetrahedron Lett.* **1990**, *31* (44), 6367–6370. [https://doi.org/10.1016/S0040-4039\(00\)97066-7](https://doi.org/10.1016/S0040-4039(00)97066-7).
 - (24) Tsakos, M.; Schaffert, E. S.; Clement, L. L.; Villadsen, N. L.; Poulsen, T. B. Ester Coupling Reactions – an Enduring Challenge in the Chemical Synthesis of Bioactive Natural Products. *Nat. Prod. Rep.* **2015**, *32*, 605–632. <https://doi.org/10.1039/c4np00106k>.
 - (25) Yudin, A. K. Macrocycles: Lessons from the Distant Past, Recent Developments, and Future Directions. *Chem. Sci.* **2015**, *6* (1), 30–49. <https://doi.org/10.1039/c4sc03089c>.
 - (26) Fürstner, A.; Aïssa, C.; Chevrier, C.; Teplý, F.; Nevado, C.; Tremblay, M. Studies on lejimalide B: Preparation of the Seco Acid and Identification of the Molecule's "Achilles Heel." *Angew. Chemie - Int. Ed.* **2006**, *45* (35), 5832–5837. <https://doi.org/10.1002/anie.200601859>.
 - (27) Cochrane, J. R.; Yoon, D. H.; McErlean, C. S. P.; Jolliffe, K. A. A Macrolactonization Approach to the Total Synthesis of the Antimicrobial Cyclic Dipeptide LI-F04a and Diastereoisomeric

- Analogues. *Beilstein J. Org. Chem.* **2012**, *8*, 1344–1351. <https://doi.org/10.3762/bjoc.8.154>.
- (28) Brazeau-Henrie, J. T.; Paquette, A. R.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N. Total and Chemoenzymatic Synthesis of Seongsanamide E. *Org. Lett.* **2022**, *24* (35), 6369–6373. <https://doi.org/10.1021/acs.orglett.2c02271>.
 - (29) Demain, A. L.; Fang, A. The Natural Functions of Secondary Metabolites. *Adv. Biochem. Eng. Biotechnol.* **2000**, *69*, 1–39.
 - (30) Williams, D. H.; Stone, M. J.; Hauck, P. R.; Rahman, S. K. Why Are Secondary Metabolites (Natural Products) Biosynthesized. *J. Nat. Prod.* **1989**, *52* (6), 1189–1208. <https://doi.org/10.1021/np50066a001>.
 - (31) Smith, S.; Tsai, S.-C. The Type I Fatty Acid and Polyketide Synthases: A Tale of Two Megasyntases. *Nat. Prod. Rep.* **2007**, *24* (5), 1041–1072. <https://doi.org/10.1039/b603600g>.
 - (32) Herbst, D. A.; Townsend, C. A.; Maier, T. The Architectures of Iterative Type I PKS and FAS. *Nat. Prod. Rep.* **2018**, *35* (10), 1046–1069. <https://doi.org/10.1039/c8np00039e>.
 - (33) Wang, J.; Zhang, R.; Chen, X.; Sun, X.; Yan, Y.; Shen, X.; Yuan, Q. Biosynthesis of Aromatic Polyketides in Microorganisms Using Type II Polyketide Synthases. *Microb. Cell Fact.* **2020**, *19* (1), 1–11. <https://doi.org/10.1186/s12934-020-01367-4>.
 - (34) Shimizu, Y.; Ogata, H.; Goto, S. Type III Polyketide Synthases: Functional Classification and Phylogenomics. *ChemBioChem* **2017**, *18* (1), 50–65. <https://doi.org/10.1002/cbic.201600522>.
 - (35) Pfeifer, B. A.; Khosla, C. Biosynthesis of Polyketides in Heterologous Hosts. *Microbiol. Mol. Biol. Rev.* **2001**, *65* (1), 106–118. <https://doi.org/10.1128/mmbr.65.1.106-118.2001>.
 - (36) Adrover-Castellano, M. L.; Schmidt, J. J.; Sherman, D. H. Biosynthetic Cyclization Catalysts for the Assembly of Peptide and Polyketide Natural Products. *ChemCatChem* **2021**, *13* (9), 2095–2116. <https://doi.org/10.1002/cctc.202001886>.
 - (37) Kraas, F. I.; Giessen, T. W.; Marahiel, M. A. Exploring the Mechanism of Lipid Transfer during Biosynthesis of the Acidic Lipopeptide Antibiotic CDA. *FEBS Lett.* **2012**, *586* (3), 283–288. <https://doi.org/10.1016/j.febslet.2012.01.003>.
 - (38) Zhong, L.; Diao, X.; Zhang, N.; Li, F.; Zhou, H.; Chen, H.; Bai, X.; Ren, X.; Zhang, Y.; Wu, D.; Bian, X. Engineering and Elucidation of the Lipoinitiation Process in Nonribosomal Peptide Biosynthesis. *Nat. Commun.* **2021**, *12* (296), 1–15. <https://doi.org/10.1038/s41467-020-20548-8>.
 - (39) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic Analysis of Condensation Domains in NRPS Sheds Light on Their Functional Evolution. *BMC Evol. Biol.* **2007**, *7*, 1–15. <https://doi.org/10.1186/1471-2148-7-78>.
 - (40) Haslinger, K.; Peschke, M.; Brieke, C.; Maximowitsch, E.; Cryle, M. J. X-Domain of Peptide Synthetases Recruits Oxygenases Crucial for Glycopeptide Biosynthesis. *Nature* **2015**, *521* (7550), 105–109. <https://doi.org/10.1038/nature14141>.
 - (41) Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L.; Marshall, C. G.; Miller, D. A.; Patel, H. M. Tailoring Enzymes That Modify Nonribosomal Peptides during and after Chain Elongation on NRPS Assembly Lines. *Curr. Opin. Chem. Biol.* **2001**, *5* (5), 525–534. [https://doi.org/10.1016/S1367-5931\(00\)00235-0](https://doi.org/10.1016/S1367-5931(00)00235-0).
 - (42) Horsman, M. E.; Hari, T. P. A. A.; Boddy, C. N. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016**, *33* (2), 183–202. <https://doi.org/10.1039/c4np00148f>.
 - (43) Little, R. F.; Hertweck, C. Chain Release Mechanisms in Polyketide and Non-Ribosomal Peptide Biosynthesis. *Nat. Prod. Rep.* **2022**, *39* (1), 163–205. <https://doi.org/10.1039/d1np00035g>.
 - (44) Holmquist, M. Alpha Beta-Hydrolase Fold Enzymes Structures, Functions and Mechanisms. *Curr. Protein Pept. Sci.* **2000**, *1* (2), 209–235. <https://doi.org/10.2174/1389203003381405>.

- (45) Galea, C. A.; Roberts, K. D.; Zhu, Y.; Thompson, P. E.; Li, J.; Velkov, T. Functional Characterization of the Unique Terminal Thioesterase Domain from Polymyxin Synthetase. *Biochemistry* **2017**, *56* (4), 657–668. <https://doi.org/10.1021/acs.biochem.6b01139>.
- (46) Kotowska, M.; Pawlik, K.; Smulczyk-Krawczynszyn, A.; Bartosz-Bechowski, H.; Kuczek, K. Type II Thioesterase ScoT, Associated with Streptomyces Coelicolor A3(2) Modular Polyketide Synthase Cpk, Hydrolyzes Acyl Residues and Has a Preference for Propionate. *Appl. Environ. Microbiol.* **2009**, *75* (4), 887–896. <https://doi.org/10.1128/AEM.01371-08>.
- (47) Schaffer, J. E.; Reck, M. R.; Prasad, N. K.; Wencewicz, T. A. β -Lactone Formation during Product Release from a Nonribosomal Peptide Synthetase. *Nat. Chem. Biol.* **2017**, *13* (7), 737–744. <https://doi.org/10.1038/nchembio.2374>.
- (48) Kreidler, D. F.; Gemmell, E. M.; Schaffer, J. E.; Wencewicz, T. A.; Gulick, A. M. The Structural Basis of N-Acyl- α -Amino- β -Lactone Formation Catalyzed by a Nonribosomal Peptide Synthetase. *Nat. Commun.* **2019**, *10* (1), 1–13. <https://doi.org/10.1038/s41467-019-11383-7>.
- (49) Dillon, S. C.; Bateman, A. The Hotdog Fold: Wrapping up a Superfamily of Thioesterases and Dehydratases. *BMC Bioinformatics* **2004**, *5*, 1–14. <https://doi.org/10.1186/1471-2105-5-109>.
- (50) Kotowska, M.; Pawlik, K. Roles of Type II Thioesterases and Their Application for Secondary Metabolite Yield Improvement. *Appl. Microbiol. Biotechnol.* **2014**, *98* (18), 7735–7746.
- (51) Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A. Regeneration of Misprimed Nonribosomal Peptide Synthetases by Type II Thioesterases. *PNAS* **2002**, *99* (22), 14083–14088. <https://doi.org/10.1073/pnas.212382199>.
- (52) Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dornan, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W. Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature* **2019**, *565* (7737), 112–117. <https://doi.org/10.1038/s41586-018-0781-z>.
- (53) Rawlings, N. D.; Woessner, J. F.; Barrett, A. J. *Handbook of Proteolytic Enzymes*, 3rd ed.; Academic Press, 2012.
- (54) Long, J. Z.; Cravatt, B. F. The Metabolic Serine Hydrolases and Their Functions in Mammalian Physiology and Disease. *Chem. Rev.* **2011**, *111* (10), 6022–6063. <https://doi.org/10.1021/cr200075y>.
- (55) Tran, L.; Broadhurst, R. W.; Tosin, M.; Cavalli, A.; Weissman, K. J. Insights into Protein-Protein and Enzyme-Substrate Interactions in Modular Polyketide Synthases. *Chem. Biol.* **2010**, *17* (7), 705–716. <https://doi.org/10.1016/j.chembiol.2010.05.017>.
- (56) Liu, Y.; Zheng, T.; Bruner, S. D. Structural Basis for Phosphopantetheinyl Carrier Domain Interactions in the Terminal Module of Nonribosomal Peptide Synthetases. *Chem. Biol.* **2011**, *18* (11), 1482–1488. <https://doi.org/10.1016/j.chembiol.2011.09.018>.
- (57) Crosby, J.; Crump, M. P. The Structural Role of the Carrier Protein - Active Controller or Passive Carrier. *Nat. Prod. Rep.* **2012**, *29* (10), 1111–1137. <https://doi.org/10.1039/c2np20062g>.
- (58) Zhou, Z.; Lai, J. R.; Walsh, C. T. T. Interdomain Communication between the Thiolation and Thioesterase Domains of EntF Explored by Combinatorial Mutagenesis and Selection. *Chem. Biol.* **2006**, *13* (8), 869–879. <https://doi.org/10.1016/j.chembiol.2006.06.011>.
- (59) Korman, T. P.; Crawford, J. M.; Labonte, J. W.; Newman, A. G.; Wong, J.; Townsend, C. A.; Tsai, S. C. Structure and Function of an Iterative Polyketide Synthase Thioesterase Domain Catalyzing Claisen Cyclization in Aflatoxin Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (14), 6246–6251. <https://doi.org/10.1073/pnas.0913531107>.
- (60) Gaudelli, N. M.; Townsend, C. A. Epimerization and Substrate Gating by a TE Domain in β -Lactam Antibiotic Biosynthesis. *Nat. Chem. Biol.* **2014**, *10* (4), 251–258.

- <https://doi.org/10.1038/nchembio.1456>.
- (61) Liu, L.; Tao, W.; Bai, L.; Kim, E. S.; Zhao, Y. L.; Shi, T. Why Does Tautomycetin Thioesterase Prefer Hydrolysis to Macrocyclization? Theoretical Study on Its Catalytic Mechanism. *Catal. Sci. Technol.* **2019**, *9* (22), 6391–6403. <https://doi.org/10.1039/c9cy01355e>.
 - (62) Chen, X. P.; Shi, T.; Wang, X. L.; Wang, J.; Chen, Q.; Bai, L.; Zhao, Y. L. Theoretical Studies on the Mechanism of Thioesterase-Catalyzed Macrocyclization in Erythromycin Biosynthesis. *ACS Catal.* **2016**, *6* (7), 4369–4378. <https://doi.org/10.1021/acscatal.6b01154>.
 - (63) Rauwerdink, A.; Kazlauskas, R. J. How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: The Serine-Histidine-Aspartate Catalytic Triad of α/β -Hydrolase Fold Enzymes. *ACS Catal.* **2015**, *5* (10), 6153–6176. <https://doi.org/10.1021/acscatal.5b01539>.
 - (64) Keatinge-Clay, A. T. The Uncommon Enzymology of Cis-Acyltransferase Assembly Lines. *Chem. Rev.* **2017**, *117* (8), 5334–5366. <https://doi.org/10.1021/acs.chemrev.6b00683>.
 - (65) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. NRSPredictor2 - A Web Server for Predicting NRPS Adenylation Domain Specificity. *Nucleic Acids Res.* **2011**, *39* (SUPPL. 2), 362–367. <https://doi.org/10.1093/nar/gkr323>.
 - (66) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The Specificity-Confering Code of Adenylation Domains in Nonribosomal Peptide Synthetases. *Chem. Biol.* **1999**, *6* (8), 493–505. [https://doi.org/10.1016/S1074-5521\(99\)80082-9](https://doi.org/10.1016/S1074-5521(99)80082-9).
 - (67) Challis, G. L.; Ravel, J.; Townsend, C. A. Predictive, Structure-Based Model of Amino Acid Recognition by Nonribosomal Peptide Synthetase Adenylation Domains. *Chem. Biol.* **2000**, *7* (3), 211–224. [https://doi.org/10.1016/S1074-5521\(00\)00091-0](https://doi.org/10.1016/S1074-5521(00)00091-0).
 - (68) Yadav, G.; Gokhale, R. S.; Mohanty, D. Computational Approach for Prediction of Domain Organization and Substrate Specificity of Modular Polyketide Synthases. *J. Mol. Biol.* **2003**, *328* (2), 335–363. [https://doi.org/10.1016/S0022-2836\(03\)00232-8](https://doi.org/10.1016/S0022-2836(03)00232-8).
 - (69) Caffrey, P. Conserved Amino Acid Residues Correlating with Ketoreductase Stereospecificity in Modular Polyketide Synthases. *ChemBioChem* **2003**, *4* (7), 654–657. <https://doi.org/10.1002/cbic.200300581>.
 - (70) Ansari, M. Z.; Sharma, J.; Gokhale, R. S.; Mohanty, D. In Silico Analysis of Methyltransferase Domains Involved in Biosynthesis of Secondary Metabolites. *BMC Bioinformatics* **2008**, *9*, 1–21. <https://doi.org/10.1186/1471-2105-9-454>.
 - (71) Yadav, G.; Gokhale, R. S.; Mohanty, D. Towards Prediction of Metabolic Products of Polyketide Synthases: An In Silico Analysis. *PLoS Comput. Biol.* **2009**, *5* (4), e1000351. <https://doi.org/10.1371/journal.pcbi.1000351>.
 - (72) Hari, T. P. A.; Labana, P.; Boileau, M.; Boddy, C. N. An Evolutionary Model Encompassing Substrate Specificity and Reactivity of Type I Polyketide Synthase Thioesterases. *ChemBioChem* **2014**, *15* (18), 2656–2661. <https://doi.org/10.1002/cbic.201402475>.
 - (73) Beld, J.; Lee, D. J.; Burkart, M. D. Fatty Acid Biosynthesis Revisited: Structure Elucidation and Metabolic Engineering. *Mol. Biosyst.* **2015**, *11* (1), 38–59. <https://doi.org/10.1039/c4mb00443d>.
 - (74) Kohli, G. S.; John, U.; Van Dolah, F. M.; Murray, S. A. Evolutionary Distinctiveness of Fatty Acid and Polyketide Synthesis in Eukaryotes. *ISME J.* **2016**, *10* (8), 1877–1890. <https://doi.org/10.1038/ismej.2015.263>.
 - (75) Chakravarty, B.; Gu, Z.; Chirala, S. S.; Wakil, S. J.; Quiocho, F. A. Human Fatty Acid Synthase: Structure and Substrate Selectivity of the Thioesterase Domain. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (44), 15567–15572. <https://doi.org/10.1073/pnas.0406901101>.
 - (76) Jing, F.; Cantu, D. C.; Tvaruzkova, J.; Chipman, J. P.; Nikolau, B. J.; Yandeau-Nelson, M. D.; Reilly, P. J. Phylogenetic and Experimental Characterization of an Acyl-ACP Thioesterase Family Reveals

- Significant Diversity in Enzymatic Specificity and Activity. *BMC Biochem.* **2011**, *12* (1), 44. <https://doi.org/10.1186/1471-2091-12-44>.
- (77) Liu, L.; Zhang, Z.; Shao, C. L.; Wang, C. Y. Analysis of the Sequences, Structures, and Functions of Product-Releasing Enzyme Domains in Fungal Polyketide Synthases. *Front. Microbiol.* **2017**, *8* (SEP), 1–13. <https://doi.org/10.3389/fmicb.2017.01685>.
 - (78) Xu, Y.; Molnar, I.; Wang, C.; Wang, X.; Liu, Q.; Xu, Y. M.; Leslie Gunatilaka, A. A.; Wei, X.; Zhang, L.; Yue, Q. Intrinsic and Extrinsic Programming of Product Chain Length and Release Mode in Fungal Collaborating Iterative Polyketide Synthases. *J. Am. Chem. Soc.* **2020**, *142* (40), 17093–17104. <https://doi.org/10.1021/jacs.0c07050>.
 - (79) Xue, Y.; Zhao, L.; Liu, H. W.; Sherman, D. H. A Gene Cluster for Macrolide Antibiotic Biosynthesis in *Streptomyces venezuelae*: Architecture of Metabolic Diversity. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (21), 12111–12116. <https://doi.org/10.1073/pnas.95.21.12111>.
 - (80) Reeves, N. R.; English, R. S.; Lampel, J. S.; Post, D. A.; Boom, T. J. V. Transcriptional Organization of the Erythromycin Biosynthetic Gene Cluster of *Saccharopolyspora erythraea*. *J. Bacteriol.* **1999**, *181* (22), 7098–7106. <https://doi.org/10.1128/jb.181.22.7098-7106.1999>.
 - (81) Lackner, G.; Bohnert, M.; Wick, J.; Hoffmeister, D. Assembly of Melleolide Antibiotics Involves a Polyketide Synthase with Cross-Coupling Activity. *Chem. Biol.* **2013**, *20* (9), 1101–1106. <https://doi.org/10.1016/j.chembiol.2013.07.009>.
 - (82) Tripathi, A.; Choi, S. S.; Sherman, D. H.; Kim, E. S. Thioesterase Domain Swapping of a Linear Polyketide Tautomycin with a Macrocyclic Polyketide Pikromycin in *Streptomyces* Sp. CK4412. *J. Ind. Microbiol. Biotechnol.* **2016**, *43* (8), 1189–1193. <https://doi.org/10.1007/s10295-016-1790-2>.
 - (83) Khan, F. I.; Lan, D.; Durrani, R.; Huan, W.; Zhao, Z.; Wang, Y. The Lid Domain in Lipases: Structural and Functional Determinant of Enzymatic Properties. *Front. Bioeng. Biotechnol.* **2017**, *5* (MAR), 1–13. <https://doi.org/10.3389/fbioe.2017.00016>.
 - (84) Roussel, A.; Canaan, S.; Egloff, M. P.; Rivière, M.; Dupuis, L.; Verger, R.; Cambillau, C. Crystal Structure of Human Gastric Lipase and Model of Lysosomal Acid Lipase, Two Lipolytic Enzymes of Medical Interest. *J. Biol. Chem.* **1999**, *274* (24), 16995–17002. <https://doi.org/10.1074/jbc.274.24.16995>.
 - (85) Martinez-Gonzalez, A. I.; Díaz-Sánchez, Á. G.; De La Rosa, L. A.; Vargas-Requena, C. L.; Bustos-Jaimes, I.; Alvarez-Parrilla, E. Polyphenolic Compounds and Digestive Enzymes: In Vitro Non-Covalent Interactions. *Molecules* **2017**, *22* (4). <https://doi.org/10.3390/molecules22040669>.
 - (86) Bauer, T. L.; Buchholz, P. C. F.; Pleiss, J. The Modular Structure of α/β -Hydrolases. *FEBS J.* **2020**, *287* (5), 1035–1053. <https://doi.org/10.1111/febs.15071>.
 - (87) Bruner, S. D.; Weber, T.; Kohli, R. M.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T.; Stubbs, M. T. Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE. *Structure* **2002**, *10* (3), 301–310. [https://doi.org/10.1016/S0969-2126\(02\)00716-5](https://doi.org/10.1016/S0969-2126(02)00716-5).
 - (88) Samel, S. A.; Wagner, B.; Marahiel, M. A.; Essen, L. O. The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-Ribosomal Lipopeptide. *J. Mol. Biol.* **2006**, *359* (4), 876–889. <https://doi.org/10.1016/j.jmb.2006.03.062>.
 - (89) Yeh, E.; Lin, H.; L. Clugston, S.; M. Kohli, R.; T. Walsh, C. Enhanced Macrocyclizing Activity of the Thioesterase from Tyrocidine Synthetase in Presence of Nonionic Detergent. *Chem. Biol.* **2004**, *11* (11), 1573–1582. <https://doi.org/10.1016/j.chembiol.2004.09.003>.
 - (90) Zhou, Y.; Prediger, P.; Dias, L. C.; Murphy, A. C.; Leadlay, P. F. Macrodilide Formation by the Thioesterase of a Modular Polyketide. *Angew. Chemie - Int. Ed.* **2015**, *127* (17), 5321–5324. <https://doi.org/10.1002/ange.201500401>.

- (91) Hoyer, K. M.; Mahlert, C.; Marahiel, M. A. The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chem. Biol.* **2007**, *14* (1), 13–22. <https://doi.org/10.1016/j.chembiol.2006.10.011>.
- (92) Heberlig, G. W.; Boddy, C. N. Thioesterase from Cereulide Biosynthesis Is Responsible for Oligomerization and Macrocyclization of a Linear Tetradepsipeptide. *J. Nat. Prod.* **2020**, *83* (6), 1990–1997. <https://doi.org/10.1021/acs.jnatprod.0c00333>.
- (93) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry* **2001**, *40* (24), 7092–7098. <https://doi.org/10.1021/bi010035r>.
- (94) Conway, K. R.; Boddy, C. N. ClusterMine360: A Database of Microbial PKS/NRPS Biosynthesis. *Nucleic Acids Res.* **2013**, *41* (D1), 402–407. <https://doi.org/10.1093/nar/gks993>.
- (95) Akey, D. L.; Kittendorf, J. D.; Giraldes, J. W.; Fecik, R. A.; Sherman, D. H.; Smith, J. L. Structural Basis for Macrolactonization by the Pikromycin Thioesterase. *Nat. Chem. Biol.* **2006**, *2* (10), 537–542. <https://doi.org/10.1038/nchembio824>.
- (96) Giraldes, J. W.; Akey, D. L.; Kittendorf, J. D.; Sherman, D. H.; Smith, J. L.; Fecik, R. A. Structural and Mechanistic Insights into Polyketide Macrolactonization from Polyketide-Based Affinity Labels. *Nat. Chem. Biol.* **2006**, *2* (10), 531–536. <https://doi.org/10.1038/nchembio822>.
- (97) Argyropoulos, P.; Bergeret, F.; Pardin, C.; Reimer, J. M.; Pinto, A.; Boddy, C. N.; Schmeing, T. M. Towards a Characterization of the Structural Determinants of Specificity in the Macrocyclizing Thioesterase for Deoxyerythronolide B Biosynthesis. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (3), 486–497. <https://doi.org/10.1016/j.bbagen.2015.11.007>.
- (98) Wang, M.; Boddy, C. N. Examining the Role of Hydrogen Bonding Interactions in the Substrate Specificity for the Loading Step of Polyketide Synthase Thioesterase Domains. *Biochemistry* **2008**, *47* (45), 11793–11803. <https://doi.org/10.1021/bi800963y>.
- (99) Fan, S.; Wang, R.; Li, C.; Bai, L.; Zhao, Y. L.; Shi, T. Insight into Structural Characteristics of Protein-Substrate Interaction in Pimaricin Thioesterase. *Int. J. Mol. Sci.* **2019**, *20* (4). <https://doi.org/10.3390/ijms20040877>.
- (100) Koch, A. A.; Hansen, D. A.; Shende, V. V.; Furan, L. R.; Houk, K. N.; Jiménez-Osés, G.; Sherman, D. H. A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst. *J. Am. Chem. Soc.* **2017**, *139* (38), 13456–13465. <https://doi.org/10.1021/jacs.7b06436>.
- (101) Sullivan, S. M.; Holyoak, T. Enzymes with Lid-Gated Active Sites Must Operate by an Induced Fit Mechanism Instead of Conformational Selection. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (37), 13829–13834. <https://doi.org/10.1073/pnas.0805364105>.
- (102) Weikl, T. R.; Von Deuster, C. Selected-Fit versus Induced-Fit Protein Binding: Kinetic Differences and Mutational Analysis. *Proteins Struct. Funct. Bioinforma.* **2009**, *75* (1), 104–110. <https://doi.org/10.1002/prot.22223>.
- (103) He, W.; Wu, J.; Khosla, C.; Cane, D. E. Macrolactonization to 10-Deoxymethynolide Catalyzed by the Recombinant Thioesterase of the Picromycin/Methymycin Polyketide Synthase. *Bioorganic Med. Chem. Lett.* **2006**, *16* (2), 391–394. <https://doi.org/10.1016/j.bmcl.2005.09.077>.
- (104) Aldrich, C. C.; Venkatraman, L.; Sherman, D. H.; Fecik, R. A. Chemoenzymatic Synthesis of the Polyketide Macrolactone 10-Deoxymethynolide. *J. Am. Chem. Soc.* **2005**, *127* (25), 8910–8911. <https://doi.org/10.1021/ja0504340>.
- (105) Boehr, D. D.; Nussinov, R.; Wright, P. E. The Role of Dynamic Conformational Ensembles in Biomolecular Recognition. *Nat. Chem. Biol.* **2009**, *5* (11), 789–796. <https://doi.org/10.1038/nchembio.232>.

- (106) Brady, S. F.; Varga, S. L.; Freidinger, R. M.; Schwenk, D. A.; Mendlowski, M.; Holly, F. W.; Veber, D. F. Practical Synthesis of Cyclic Peptides, with an Example of Dependence of Cyclization Yield upon Linear Sequence. *J. Org. Chem.* **1979**, *44* (18), 3101–3105. <https://doi.org/10.1021/jo01332a003>.
- (107) Kopple, K. D. Synthesis of Cyclic Peptides. *J. Pharm. Sci.* **1972**, *61* (9), 1345–1356. <https://doi.org/10.1002/jps.2600610902>.
- (108) Bu, X.; Wu, X.; Xie, G.; Guo, Z. Synthesis of Tyrocidine A and Its Analogues by Spontaneous Cyclization in Aqueous Solution. *Org. Lett.* **2002**, *4* (17), 2893–2895. <https://doi.org/10.1021/ol0263191>.
- (109) Gibbs, C. A.; Kondejewski, H. L.; Gronwald, W.; Nip, M. A.; Hodges, S. R.; Sykes, D. B.; Wishart, S. D. Unusual β -Sheet Periodicity in Small Cyclic Peptides. *Nat. Struct. Biol.* **1998**, *5* (4), 284–288.
- (110) Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. Generality of Peptide Cyclization Catalyzed by Isolated Thioesterase Domains of Nonribosomal Peptide Synthetases. *Biochemistry* **2001**, *40* (24), 7099–7108. <https://doi.org/10.1021/bi010036j>.
- (111) Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. Peptide Cyclization Catalysed by the Thioesterase Domain of Tyrocidine Synthetase. *Nature* **2000**, *407* (6801), 215–218. <https://doi.org/10.1038/35025116>.
- (112) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. Characterization of the Surfactin Synthetase C-Terminal Thioesterase Domain as a Cyclic Depsipeptide Synthase. *Biochemistry* **2002**, *41* (45), 13350–13359. <https://doi.org/10.1021/bi026592a>.
- (113) Hou, J.; Robbel, L.; Marahiel, M. A. Identification and Characterization of the Lysobactin Biosynthetic Gene Cluster Reveals Mechanistic Insights into an Unusual Termination Module Architecture. *Chem. Biol.* **2011**, *18* (5), 655–664. <https://doi.org/10.1016/j.chembiol.2011.02.012>.
- (114) Lepetit, C. A.; Paquette, A. R.; Brazeau-Henrie, J. T.; Boddy, C. N. Total and Chemoenzymatic Synthesis of the Lipodepsipeptide Rhizomide A. *Bioorg. Med. Chem. Lett.* **2023**, *96* (Article 129506), 1–5. <https://doi.org/10.1016/j.bmcl.2023.129506>.
- (115) Wu, S.; Snajdrova, R.; Moore, J. C.; Baldenius, K.; Bornscheuer, U. T. Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chemie - Int. Ed.* **2021**, *60* (1), 88–119. <https://doi.org/10.1002/anie.202006648>.
- (116) Bell, E. L.; Finnigan, W.; France, S. P.; Green, A. P.; Hayes, M. A.; Hepworth, L. J.; Lovelock, S. L.; Niikura, H.; Osuna, S.; Romero, E.; Ryan, K. S.; Turner, N. J.; Flitsch, S. L. Biocatalysis. *Nat. Rev. Methods Prim.* **2021**, *1* (1), 1–21. <https://doi.org/10.1038/s43586-021-00044-z>.
- (117) Aschner, P.; Kipnes, M. S.; Lunceford, J. K.; Sanchez, M.; Mickel, C.; Williams-Herman, D. E. Effect of the Dipeptidyl Peptidase-4 Inhibitor Sitagliptin as Monotherapy on Glycemic Control in Patients with Type 2 Diabetes. *Diabetes Care* **2006**, *29* (12), 2632–2637. <https://doi.org/10.2337/dc06-0703>.
- (118) Hansen, K. B.; Hsiao, Y.; Xu, F.; Rivera, N.; Clausen, A.; Kubryk, M.; Krska, S.; Rosner, T.; Simmons, B.; Balsells, J.; Ikemoto, N.; Sun, Y.; Spindler, F.; Malan, C.; Grabowski, E. J. J.; Armstrong, J. D. Highly Efficient Asymmetric Synthesis of Sitagliptin. *J. Am. Chem. Soc.* **2009**, *131* (25), 8798–8804. <https://doi.org/10.1021/ja902462q>.
- (119) Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J. Biocatalytic Asymmetric Synthesis of Sitagliptin Manufacture. *Science (80-.)*. **2010**, *329* (July), 305–310. <https://doi.org/10.1126/science.1188934>.
- (120) Sweeney, C.; Bracarda, S.; Sternberg, C. N.; Chi, K. N.; Olmos, D.; Sandhu, S.; Massard, C.; Matsubara, N.; Alekseev, B.; Parnis, F.; Atduev, V.; Buchschacher, G. L.; Gafanov, R.; Corrales, L.; Borre, M.; Stroyakovskiy, D.; Alves, G. V.; Bournakis, E.; Puente, J.; Harle-Yge, M. L.; Gallo, J.;

- Chen, G.; Hanover, J.; Wongchenko, M. J.; Garcia, J.; de Bono, J. S. Ipatasertib plus Abiraterone and Prednisolone in Metastatic Castration-Resistant Prostate Cancer (IPATential150): A Multicentre, Randomised, Double-Blind, Phase 3 Trial. *Lancet* **2021**, *398* (10295), 131–142. [https://doi.org/10.1016/S0140-6736\(21\)00580-8](https://doi.org/10.1016/S0140-6736(21)00580-8).
- (121) Han, C.; Savage, S.; Al-Sayah, M.; Yajima, H.; Remarchuk, T.; Reents, R.; Wirz, B.; Iding, H.; Bachmann, S.; Fantasia, S. M.; Scalone, M.; Hell, A.; Hidber, P.; Gosselin, F. Asymmetric Synthesis of Akt Kinase Inhibitor Ipatasertib. *Org. Lett.* **2017**, *19* (18), 4806–4809. <https://doi.org/10.1021/acs.orglett.7b02228>.
- (122) Molinaro, C.; Phillips, E. M.; Xiang, B.; Milczek, E.; Shevlin, M.; Balsells, J.; Ceglia, S.; Chen, J.; Chen, L.; Chen, Q.; Fei, Z.; Hoerrner, S.; Qi, J.; De Lera Ruiz, M.; Tan, L.; Wan, B.; Yin, J. Synthesis of a CGRP Receptor Antagonist via an Asymmetric Synthesis of 3-Fluoro-4-Aminopiperidine. *J. Org. Chem.* **2019**, *84* (12), 8006–8018. <https://doi.org/10.1021/acs.joc.9b00569>.
- (123) Yasuda, N.; Cleator, E.; Kosjek, B.; Yin, J.; Xiang, B.; Chen, F.; Kuo, S. C.; Belyk, K.; Mullens, P. R.; Goodyear, A.; Edwards, J. S.; Bishop, B.; Ceglia, S.; Belardi, J.; Tan, L.; Song, Z. J.; Dimichele, L.; Reamer, R.; Cabirol, F. L.; Tang, W. L.; Liu, G. Practical Asymmetric Synthesis of a Calcitonin Gene-Related Peptide (CGRP) Receptor Antagonist Ubrogapant. *Org. Process Res. Dev.* **2017**, *21* (11), 1851–1858. <https://doi.org/10.1021/acs.oprd.7b00293>.
- (124) Schober, M.; MacDermaid, C.; Ollis, A. A.; Chang, S.; Khan, D.; Hosford, J.; Latham, J.; Ihnken, L. A. F.; Brown, M. J. B.; Fuerst, D.; Sangane, M. J.; Roiban, G. D. Chiral Synthesis of LSD1 Inhibitor GSK2879552 Enabled by Directed Evolution of an Imine Reductase. *Nat. Catal.* **2019**, *2* (10), 909–915. <https://doi.org/10.1038/s41929-019-0341-4>.
- (125) Hu, X.; Robin, S.; O'Connell, S.; Walsh, G.; Wall, J. G. Engineering of a Fungal β -Galactosidase to Remove Product Inhibition by Galactose. *Appl. Microbiol. Biotechnol.* **2010**, *87* (5), 1773–1782. <https://doi.org/10.1007/s00253-010-2662-8>.
- (126) Shang, Y. P.; Chen, Q.; Li, A. T.; Quan, S.; Xu, J. H.; Yu, H. L. Attenuated Substrate Inhibition of a Haloketone Reductase via Structure-Guided Loop Engineering. *J. Biotechnol.* **2020**, *308* (September 2019), 141–147. <https://doi.org/10.1016/j.jbiotec.2019.12.011>.
- (127) Ziegler, J.; Brandt, W.; Geißler, R.; Facchini, P. J. Removal of Substrate Inhibition and Increase in Maximal Velocity in the Short Chain Dehydrogenase/Reductase Salutaridine Reductase Involved in Morphine Biosynthesis. *J. Biol. Chem.* **2009**, *284* (39), 26758–26767. <https://doi.org/10.1074/jbc.M109.030957>.
- (128) France, S. P.; Lewis, R. D.; Martinez, C. A. The Evolving Nature of Biocatalysis in Pharmaceutical Research and Development. *JACS Au* **2023**, *3* (3), 715–735. <https://doi.org/10.1021/jacsau.2c00712>.
- (129) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. Mechanism and Specificity of the Terminal Thioesterase Domain from the Erythromycin Polyketide Synthase. *Chem. Biol.* **1999**, *6* (2), 117–125. [https://doi.org/10.1016/S1074-5521\(99\)80008-8](https://doi.org/10.1016/S1074-5521(99)80008-8).
- (130) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. Manipulation of Macrolide Ring Size by Directed Mutagenesis of a Modular Polyketide Synthase Department of Chemical Engineering. *J. Am. Chem. Soc.* **1995**, *117* (35), 9105–9106.
- (131) Tsai, S. C.; Lu, H.; Cane, D. E.; Khosla, C.; Stroud, R. M. Insights into Channel Architecture and Substrate Specificity from Crystal Structures of Two Macrocyclic-Forming Thioesterases of Modular Polyketide Synthases. *Biochemistry* **2002**, *41* (42), 12598–12606. <https://doi.org/10.1021/bi0260177>.
- (132) Tsai, S. C.; Miercke, L. J. W.; Krucinski, J.; Gokhale, R.; Chen, J. C. H.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. Crystal Structure of the Macrocyclic-Forming Thioesterase Domain of the

- Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (26), 14808–14813. <https://doi.org/10.1073/pnas.011399198>.
- (133) Pinto, A.; Wang, M.; Horsman, M.; Boddy, C. N. 6-Deoxyerythronolide B Synthase Thioesterase-Catalyzed Macrocyclization Is Highly Stereoselective. *Org. Lett.* **2012**, *14* (9), 2278–2281. <https://doi.org/10.1021/ol300707j>.
- (134) Kao, C. M.; Mcpherson, M.; Mcdaniel, R. N.; Fu, H.; Cane, D. E.; Khosla, C. Gain of Function Mutagenesis of the Erythromycin Polyketide Synthase . 2 . Engineered Biosynthesis of Departments of Chemical Engineering , Chemistry and Thesis of Polyketides , Natural Product Molecules with Remarkable Organization of These Multifunction. *J. Am. Chem. Soc.* **1997**, *119* (46), 11339–11340.
- (135) Jacobsen, J. R.; Hutchinson, C. R.; Cane, D. E.; Khosla, C. Precursor-Directed Biosynthesis of Erythromycin Analogs by an Engineered Polyketide Synthase. *Sci. Reports* **1997**, *277* (5324), 367–370.
- (136) Kao, C. M.; Lun, G.; Katz, L.; Kane, D. E.; Khosla, C. Engineered Biosynthesis of Structurally Diverse Tetraketides by a Trimodular Polyketide Synthase. *J. Am. Chem. Soc.* **1996**, *118* (38), 9184–9185.
- (137) Wang, M.; Opare, P.; Boddy, C. N. Polyketide Synthase Thioesterases Catalyze Rapid Hydrolysis of Peptidyl Thioesters. *Bioorg. Med. Chem. Lett.* **2009**, *19* (5), 1413–1415. <https://doi.org/10.1016/j.bmcl.2009.01.040>.
- (138) Lu, H.; Tsai, S. C.; Khosla, C.; Cane, D. E. Expression, Site-Directed Mutagenesis, and Steady State Kinetic Analysis of the Terminal Thioesterase Domain of the Methymycin/Picromycin Polyketide Synthase. *Biochemistry* **2002**, *41* (42), 12590–12597. <https://doi.org/10.1021/bi026006d>.
- (139) Mukhtar, T. A.; Koteva, K. P.; Wright, G. D. Chimeric Streptogramin-Tyrocycline Antibiotics That Overcome Streptogramin Resistance. *Chem. Biol.* **2005**, *12* (2), 229–235. <https://doi.org/10.1016/j.chembiol.2004.12.009>.
- (140) Kohli, R. M.; Walsh, C. T.; Burkart, M. D. Biomimetic Synthesis and Optimization of Cyclic Peptide Antibiotics. *Nature* **2002**, *418*, 658–661.
- (141) Heberlig, G. W.; Wirz, M.; Wang, M.; Boddy, C. N. Resorcylic Acid Lactone Biosynthesis Relies on a Stereotolerant Macrocyclizing Thioesterase. *Org. Lett.* **2014**, *16* (22), 5858–5861. <https://doi.org/10.1021/ol502747t>.
- (142) Heberlig, G. W.; Brown, J. T. C.; Simard, R. D.; Wirz, M.; Zhang, W.; Wang, M.; Susser, L. I.; Horsman, M. E.; Boddy, C. N. Chemoenzymatic Macrocyclic Synthesis Using Resorcylic Acid Lactone Thioesterase Domains. *Org. Biomol. Chem.* **2018**, *16* (32), 5771–5779. <https://doi.org/10.1039/c8ob01512k>.
- (143) Wang, M.; Zhou, H.; Wirz, M.; Tang, Y.; Boddy, C. N. A Thioesterase from an Iterative Fungal Polyketide Synthase Shows Macrocyclization and Cross Coupling Activity and May Play a Role in Controlling Iterative Cycling through Product Offloading. *Biochemistry* **2009**, *48* (27), 6288–6290. <https://doi.org/10.1021/bi9009049>.
- (144) Huo, L.; Rachid, S.; Stadler, M.; Wenzel, S. C.; Muller, R.; Ito, K.; Passioura, T.; Suga, H.; Langer, M.; Reck, G.; Reed, J.; Rétey, J.; Patton, G. C. G.; Paul, M.; Cooper, L. L. E.; Chatterjee, C.; Van Der Donk, W. a.; Raboisson, P.; Sambasivarao, S. V.; Sieber, S. a; Marahiel, M. a; Walsh, C. T.; Nolan, E. M.; Zhang, M.-Q.; Wilkinson, B. Learning from Nature’s Drug Factories : Nonribosomal Synthesis of Macrocyclic Peptides MINIREVIEW Learning from Nature ’ s Drug Factories : Nonribosomal Synthesis of Macrocyclic Peptides. *Biochemistry* **2008**, *18* (24), 5655–5656. <https://doi.org/10.1128/JB.185.24.7036>.
- (145) Kohli, R. M.; Burke, M. D.; Tao, J.; Walsh, C. T. Chemoenzymatic Route to Macrocyclic Hybrid Peptide/Polyketide-like Molecules. *J. Am. Chem. Soc.* **2003**, *125* (24), 7160–7161.

- <https://doi.org/10.1021/ja0352202>.
- (146) Khosla, C.; Gokhale, R. S.; Jacobsen, J. R.; Cane, D. E. Tolerance and Specificity of Polyketide Synthases. *Annu. Rev. Biochem.* **1999**, *68*, 219–253.
 - (147) Li, C.; Zhang, R.; Wang, J.; Wilson, L. M.; Yan, Y. Protein Engineering for Improving and Diversifying Natural Product Biosynthesis. *Trends Biotechnol.* **2020**, *38* (7), 729–744.
 - (148) IV, K. Rational and Semirational Protein Design. *Methods Mol. Biol.* **2018**, *1685*, 15–23. <https://doi.org/10.1007/978-1-4939-7366-8>.
 - (149) Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nat. Rev. Mol. Cell Biol.* **2019**, No. 20, 681–697. <https://doi.org/10.1038/s41580-019-0163-x>.
 - (150) Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S. T.; Zhao, H. Directed Evolution: Methodologies and Applications. *Chem. Rev.* **2021**, *121* (20), 12384–12444. <https://doi.org/10.1021/acs.chemrev.1c00260>.
 - (151) Brannigan, J. A.; Wilkinson, A. J. Protein Engineering 20 Years On. *Nat. Rev. Mol. Cell Biol.* **2002**, *3* (12), 964–970. <https://doi.org/10.1038/nrm975>.
 - (152) Packer, M. S.; Liu, D. R. Methods for the Directed Evolution of Proteins. *Nat. Rev. Genet.* **2015**, *16* (7), 379–394. <https://doi.org/10.1038/nrg3927>.
 - (153) McCullum, E. O.; Williams, B. A. R.; Zhang, J.; Chaput, J. C. Random Mutagenesis by Error-Prone PCR. *Methods Mol. Biol.* **2010**, *634* (3), 103–109. https://doi.org/10.1007/978-1-60761-652-8_7.
 - (154) Vanhercke, T.; Ampe, C.; Tirry, L.; Denolf, P. Reducing Mutational Bias in Random Protein Libraries. *Anal. Biochem.* **2005**, *339*, 9–14. <https://doi.org/10.1016/j.ab.2004.11.032>.
 - (155) Huiman, Z.; Arnold, H. F. Optimization of DNA Shuffling for High Fidelity Recombination. *Acta Polym. Sin.* **1997**, *25* (6), 1307–1308. <https://doi.org/10.1177/j.issn1000-3304.2015.14315>.
 - (156) Harayama, S. Artificial Evolution by DNA Shuffling. *Trends Biotechnol.* **1998**, *16* (2), 76–82. [https://doi.org/10.1016/S0167-7799\(97\)01158-X](https://doi.org/10.1016/S0167-7799(97)01158-X).
 - (157) Petrie, K. L.; Joyce, G. F. Deep Sequencing Analysis of Mutations Resulting from the Incorporation of DNTP Analogs. *Nucleic Acids Res.* **2010**, *38* (22), 8095–8104. <https://doi.org/10.1093/nar/gkq685>.
 - (158) Cravens, A.; Jamil, O. K.; Kong, D.; Sockolosky, J. T.; Smolke, C. D. Polymerase-Guided Base Editing Enables in Vivo Mutagenesis and Rapid Protein Engineering. *Nat. Commun.* **2021**, *12* (1), 1–12. <https://doi.org/10.1038/s41467-021-21876-z>.
 - (159) Muteeb, G.; Sen, R. Random Mutagenesis Using a Mutator Strain. In *Methods in Molecular Biology*; 2010; Vol. 634, pp 411–419. https://doi.org/10.1007/978-1-60761-652-8_29.
 - (160) Shibai, A.; Takahashi, Y.; Ishizawa, Y.; Motooka, D.; Nakamura, S.; Ying, B. W.; Tsuru, S. Mutation Accumulation under UV Radiation in Escherichia Coli. *Sci. Rep.* **2017**, *7* (1), 1–12. <https://doi.org/10.1038/s41598-017-15008-1>.
 - (161) Shaw-reid, C. A.; Kelleher, N. L.; Losey, H. C.; Gehring, A. M.; Berg, C.; Walsh, C. T. Assembly Line Enzymology by Multimodular Nonribosomal Peptide Synthetases: The Thioesterase Domain of E.Coli EntF Catalyzes Both Elongation and Cyclolactonization. *Cell Chem. Bol* **1999**, *6* (6), 385–400.
 - (162) Qiao, L.; Fang, J.; Zhu, P.; Huang, H.; Dang, C.; Pang, J.; Gao, W.; Qiu, X.; Huang, L.; Li, Y. A Novel Chemoenzymatic Approach to Produce Cilengitide Using the Thioesterase Domain from Microcystis Aeruginosa Microcystin Synthetase C. *Protein J.* **2019**, *38* (6), 658–666. <https://doi.org/10.1007/s10930-019-09864-1>.
 - (163) Fraley, A. E.; Dieterich, C. L.; Mabesoone, M. F. J.; Minas, H. A.; Meoded, R. A.; Hemmerling, F.; Piel, J. Structure of a Promiscuous Thioesterase Domain Responsible for Branching Acylation in Polyketide Biosynthesis. *Angew. Chemie - Int. Ed.* **2022**. <https://doi.org/10.1002/anie.202206385>.
 - (164) Zhou, Y.; Tao, W.; Qi, Z.; Wei, J.; Shi, T.; Kang, Q.; Zheng, J.; Zhao, Y.; Bai, L. Structural and

- Mechanistic Insights into Chain Release of the Polyene PKS Thioesterase Domain. *ACS Catal.* **2022**, *12* (1), 762–776. <https://doi.org/10.1021/acscatal.1c04991>.
- (165) Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33* (7), 1870–1874. <https://doi.org/10.1093/MOLBEV/MSW054>.
- (166) Guindon, S.; Dufayard, J. F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O.; Uindon, P. G.; Ranc, J. E. A. N. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59* (3), 307–321. <https://doi.org/10.1093/sysbio/syq010>.
- (167) Anisimova, M.; Gascuel, O. Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative. *Syst. Biol.* **2006**, *55* (4), 539–552. <https://doi.org/10.1080/10635150600755453>.

CHAPTER 2 : Chemoenzymatic synthesis of depsipeptides using NRPS TEs

2.1 Preface

The work presented in this chapter is a combination of two related published studies, “Total and Chemoenzymatic Synthesis of Seongsanamide E” *Org. Lett.* 2022, **24** (35), 6369-6373 and “Total and Chemoenzymatic Synthesis of the Lipodepsipeptide Rhizomide A” *Bioorganic Med. Chem. Lett.* 2023, **96**, #129506. Both are reprinted with permission, see **Appendix I**.

Author contributions

I expressed and purified each protein (Sgd- and Rzm-TE), collected and analysed LCMS data and performed kinetic characterization. Allison Q. O’Rourke helped with LCMS data collection with the Sgd-TE assays. Dr. André R. Paquette and Corinne A. Lepetit synthesized both Sgd and Rzm peptides, respectively, and performed all compound characterization. Dr. Michael G. Darnowski helped collect NMR data on Sgd E authentic standard. Dr. André R. Paquette, Corinne A. Lepetit, Dr. Christopher N. Boddy and I wrote the papers with input from all authors.

2.2 Introduction

Non-ribosomal peptides represent a rich and important source of macrocyclic natural products with potent and useful biological activities. For example, tyrocidine (Tyc) and gramicidin (Grs) are both used to treat superficial skin and throat infections caused by pathogenic bacteria belonging to the staphylococci and enterococci genera.¹ Over the past three decades, macrocyclic lipodepsipeptides (mLPD), *N*-acylated peptides known to contain one or more ester linkages in addition to the amide linkages,^{2,3} have gained traction. Of the many known mLPD, two caught our interest and initiated the research presented in this chapter. We envision that characterizing the TE domains responsible for their macrolactonization would provide insight into the TE biochemistry and depsipeptide biosynthesis as well as generating a concise total synthesis of these natural products while highlighting the power of these enzymes as biocatalysts.

The seongsanamides (Sgds) are a family of mLPD with 6 members, A-F, isolated from the marine sponge associated bacterium *Bacillus safensis* KCTC 12796BP.⁴ Sgds A-E are *N*-acylated octapeptides with an L-Leu-D-Ala-D-Tyr-L-Thr-D-Leu-L-Leu-L-Ile-L-Tyr sequence and a macrolactone linkage between the terminal carboxylate of L-Tyr₈ and the sidechain hydroxyl of L-Thr₄. Sgds A-D show a dose-dependent response for inhibition of β -hexosaminidase while Sgd A displays antiallergenic activities in a mouse model.⁵ Sequencing of the *B. safensis* genome identified an NRPS gene cluster^{4,6} that encoded two NRPS genes possessing adenylation (A) domain selectivity⁷⁻¹⁰ consistent with these amino acids and epimerization (E) domains appropriately placed to generate the three D-configured residues (**Figure S2.20**). Based on this putative gene cluster, it is proposed that Sgd E is generated first with the condensation of an acyl-CoA (or other CoA derivatives) onto the first residue L-Leu bound to a thiolation (T) domain by a starter condensation (C_s) domain.⁴ After generation of a linear lipooctapeptide, it is released concomitant with macrolactonisation⁴ by the C-terminal TE domain forming Sgd E (**Figure 2.1A**

and **S2.20**).¹¹ Ultimately Sgd E is proposed to be further cyclized by a P450 (SgdC),¹² forming a biaryl ether linkage between D-Tyr₃ and L-Tyr₈ via oxidative phenol-coupling, generating Sgd B (**Figure S2.20**).

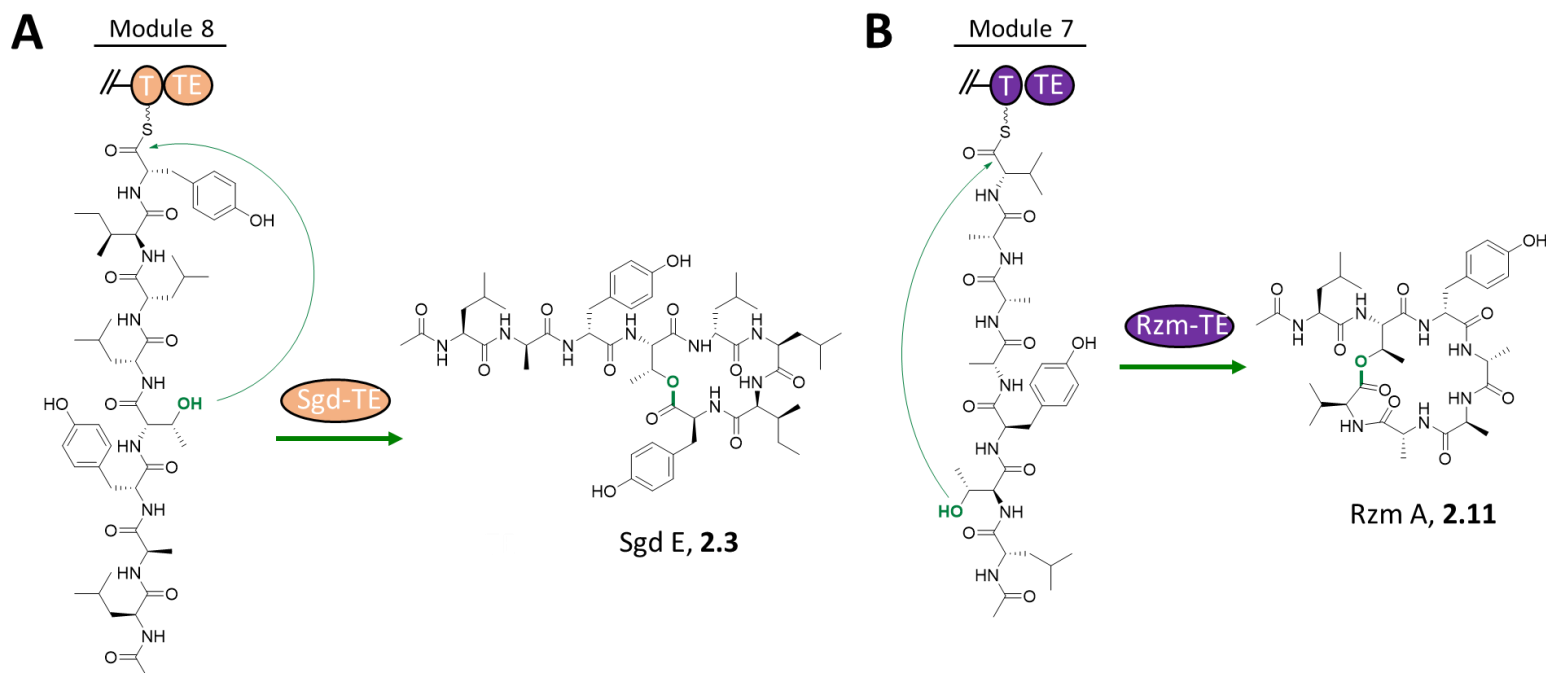


Figure 2.1. TE-mediated release of Sgd E and Rzm A. Release of Sgd E (A) and Rzm A (B) from their respective NRPS complexes via TE-mediated macrolactonization.

The rhizomides (Rzms) are a family of mLPD encoded in the genome of the β -proteobacterium *Paraburkholderia rhizoxinica* str. HKI 454.^{13,14} The strain-specific megaplasmid pBRH01 contains the NRPS gene cluster responsible for the biosynthesis of the Rzms;^{13–15} however, under typical cultivation conditions this gene cluster is silent.¹³ Installing the constitutive *apr* promoter along with the apramycin resistance gene directly upstream of the gene cluster in the native host provided a 70-fold increase in Rzm A production, enabling its isolation and structure determination.¹³ The Rzms show weak anticancer, antifungal, and antibacterial activity, suggesting they may have a highly strain specific growth inhibitory activity or play a role in intra- or inter-species communication similar to other cyclic depsipeptides.¹⁶ The Rzms are biosynthesized by a single NRPS complex (RzmA) that contains 7 modules. It begins with a C₅

domain responsible for the condensation of acetyl-CoA to a T domain bound L-Leu, yielding the peptidyl-enzyme intermediate *N*-acetyl-L-Leu-T.^{17–19} Chain extension then proceeds with the incorporation of L-Thr, D-Tyr, D-Ala, L-Ala, D-Ala and L-Val giving the linear Rzm A precursor (**Figure S2.21**). Based on the structures of Rzms A-C, the A domains of modules 5 and 6 are likely promiscuous and can activate the non-native substrates L-Ser, to yield the minor products Rzms B and C (**Figure S2.21**).¹³ The linear peptide chain is then passed to a terminal TE domain,¹¹ which is believed to be responsible for the cyclization and release of Rzms A-C through macrolactonization of the L-Thr₂ side-chain alcohol with the backbone carboxylate of L-Val₇ (**Figure 2.1B** and **S2.21**).

With the diverse and potent biological activities displayed by depsipeptides and mLPD, synthesis of these molecules is of great interest but also challenging.^{20–22} While synthesis of the linear peptidyl chain is typically straight forward, macrolactonization has been a longstanding synthetic challenge. Effective macrolactonization reactions, such as the Corey-Nicolaou and Yamaguchi esterifications routinely struggle to generate cyclic depsipeptides^{5,23,24} or lead to epimerization of the C-terminal α -carbon, generating diastereomeric products.^{25–27} To address this problem, alternative sites of ring closure via macrolactamization can be used; however, this requires more elaborate branched precursor synthesis.^{28–30} The TEs responsible for macrolactonizing natural products during their biosynthesis circumvent these synthetic challenges. Harnessing TEs as biocatalysts for the chemoenzymatic synthesis of natural products and new to nature compounds has been successfully done,^{30–33} highlighting the power of these enzymes as macrocyclization biocatalysts.

To further expand the current repertoire of TE macrocyclization biocatalysts and complete the total chemoenzymatic syntheses of Sgd E and Rzm A, the TE domains for both Sgd (Sgd-TE) and Rzm (Rzm-TE) are characterized as standalone enzymes *in vitro*. Access to an appropriate substrate, authentic standard of the natural products and recombinantly purified TE domains are required.³⁴ Using a

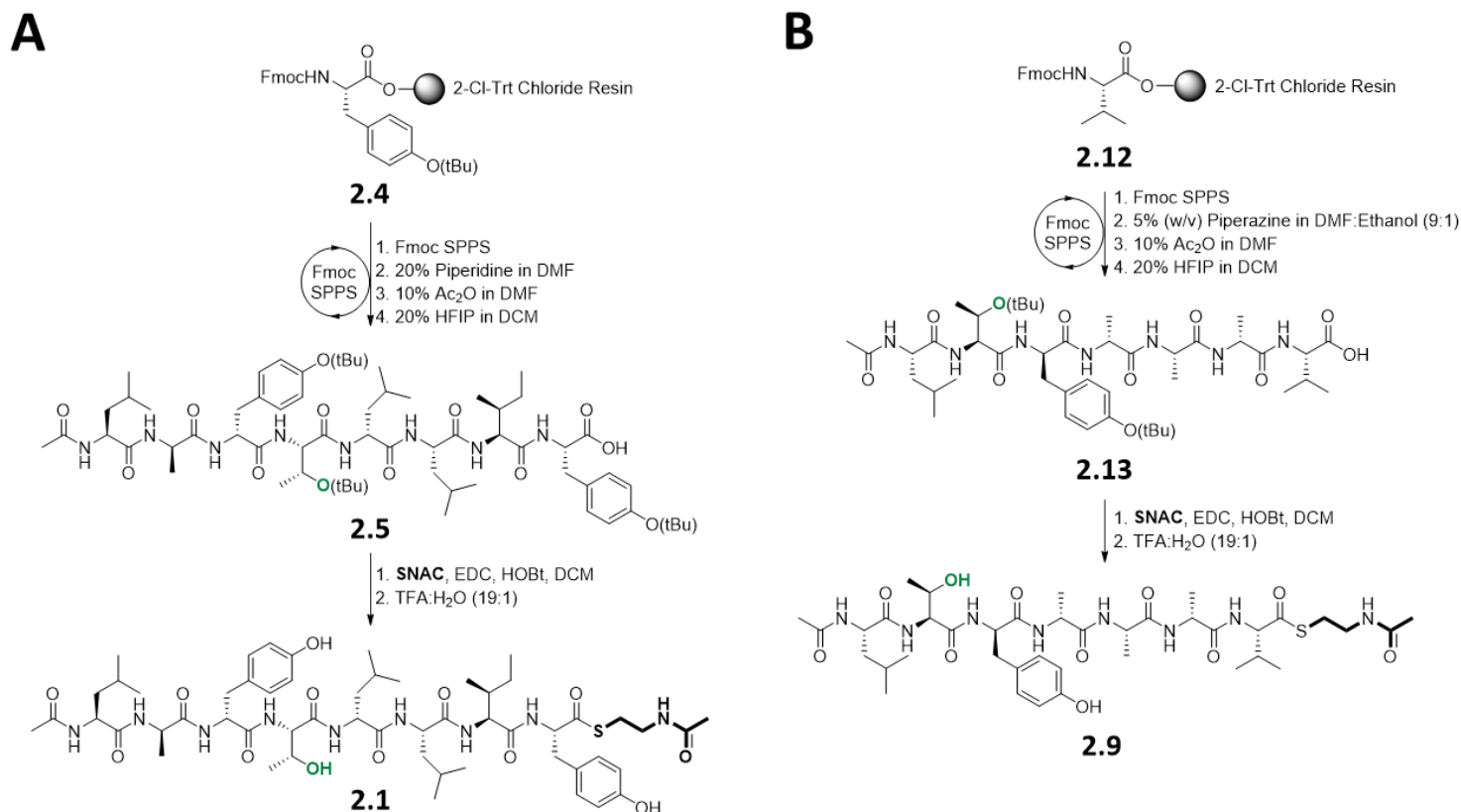
combination of solid phase peptide synthesis (SPPS) and solution phase chemistry, we were able to synthesize the linear *N*-acetylcysteamine (SNAC) activated substrates **2.1** (Sgd) and **2.9** (Rzm), as well as authentic natural product standards Sgd E **2.3** and Rzm A **2.11**. Both Sgd and Rzm TEs were recombinantly expressed and purified. *In vitro* assays with the SNAC substrates were performed and analyzed by LCMS for unambiguous detection of products. We show that both TEs are active as macrocyclization catalysts for their respective substrates with little competing hydrolysis reactivity, demonstrating the macrolactonizing power of these enzymes and experimentally confirming the gene cluster identified by Kim et al.⁴ as responsible for the biosynthesis of Sgds. We also show Sgd-TE is able to catalyze ring opening of Sgd E **2.3** forming seco acid **2.2**, a rare activity only seen for the epothilone (Epo) and surfactin (Srf) TEs.^{35,36}

2.3 Results and discussion

Synthesis of Sgd and Rzm substrates and macrocyclic authentic standards

C-terminal thioester peptides like **2.1** and **2.9** are synthetically challenging.³⁷ While linear peptides with C-terminal carboxylates or carboxamides can be routinely synthesized with automated Fmoc solid phase peptide synthesis (SPPS), C-terminal thioesters typically require specialized strategies^{38–40} or synthetic modifications.^{41,42} In our hands, the most effective approach for accessing the peptidyl-SNAC thioesters **2.1** and **2.9** (**Scheme 2.1A** and **2.1B**), is to install the thioester using solution phase chemistry on a protected linear free acid peptide.^{31,32} Thus, we generated the on resin linear protected peptides from hyper-sensitive 2-Cl-Trt Chloride resin⁴³ loaded with either C-terminal Fmoc-L-Tyr(OtBu)-OH **2.4**, for Sgd, and C-terminal Fmoc-L-Val-OH **2.12**, for Rzm. Both were subjected to SPPS under HBTU/DIEA conditions, subsequent N-terminal Fmoc-deprotection, and acetylation to generate full length protected Sgd and Rzm peptides. Cleavage from the resin under mild acidic conditions³⁰ yielded free acids **2.5**, for Sgd, and **2.13**, for Rzm, which were activated as the SNAC thioesters. The sidechain tBu protecting groups of **2.5** were removed with TFA and the substrate for Sgd, **2.1**, was purified to >95 % purity using RP-HPLC (65% yield from **2.5**, **Figure S2.2**). Similarly, **2.13** deprotection with TFA and RP-HPLC purification afforded the Rzm substrate, **2.9**, to >95 % purity (5% yield from **2.13**, **Figure S2.10**).

Scheme 2.1. SPPS of linear SNAC substrates Sgd 2.1 (A) and Rzm 2.9 (B).



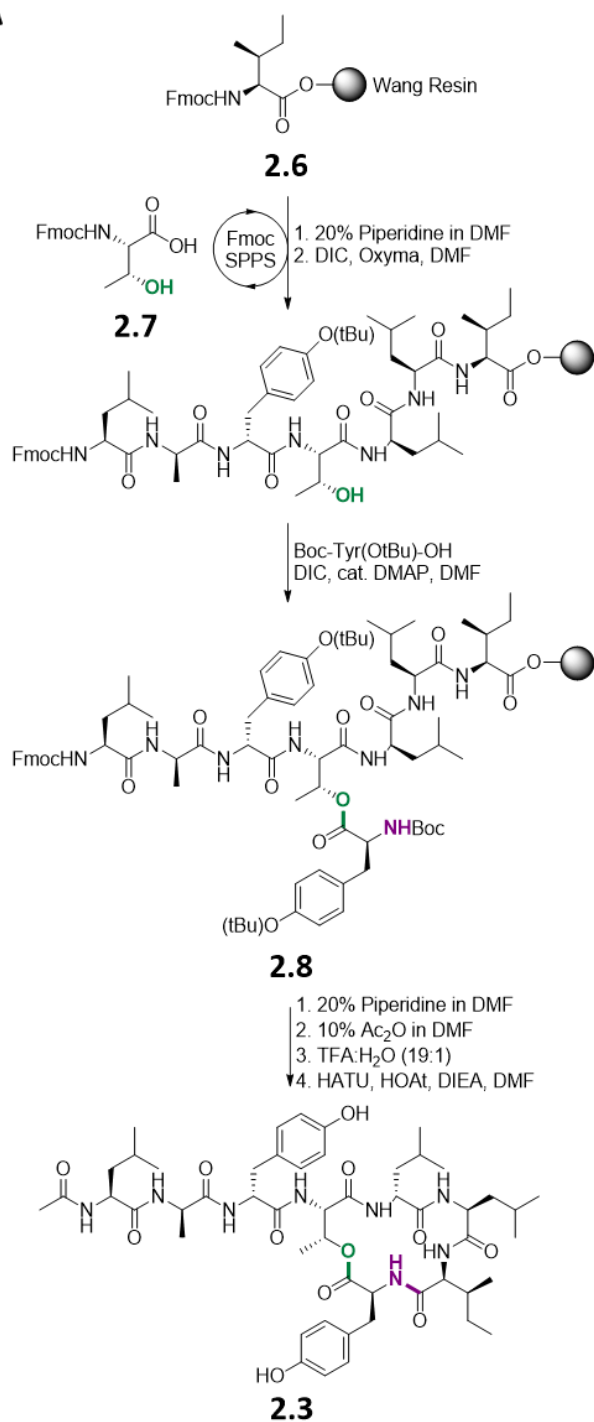
Synthesis of authentic Sgd E **2.3** was not possible via a macrolactonization strategy. All attempts to generate macrocycle from linear free acid with unmasked L-Thr β -hydroxyl group failed.⁴⁴ We thus turned our attention to the common alternative approach of macrolactamization^{45–47} via the amide bond of L-Ile₇ and L-Tyr₈ (**Scheme 2.2A**). Wang resin was loaded with Fmoc-L-Ile-OH **2.6** and underwent SPPS with DIC/Oxyma Pure conditions to extend the linear peptide sequence, incorporating Fmoc-L-Thr-OH **2.7** with an unprotected β -alcohol. The on-resin peptide ester **2.8** was formed through standard esterification conditions with Boc-L-Tyr(OtBu)-OH, followed by N-terminal deprotection and acetylation to complete the linear sequence. The peptide was fully deprotected and cleaved from the resin using TFA to afford the linear peptide ester. Macrolactamization under dilute HATU/HOAt coupling conditions successfully generated **2.3**.^{5,21} Upon LCMS analysis, two peaks with the same mass were observed eluting close

together. The isomers were separated by RP-HPLC and ¹H-NMR confirmed the first eluting isomer was the natural product Sgd E, **2.3** (**Figure S2.4** and **S2.5**). Coupling conditions known to reduce epimerization at the Ile α-carbon were screened, though no difference in product distribution was observed.²¹ Regardless authentic **2.3** could be isolated in >95% purity by RP-HPLC (32% yield from seco acid, **Figure S2.4**) and was indistinguishable from the isolated material (**Figure S2.6**).

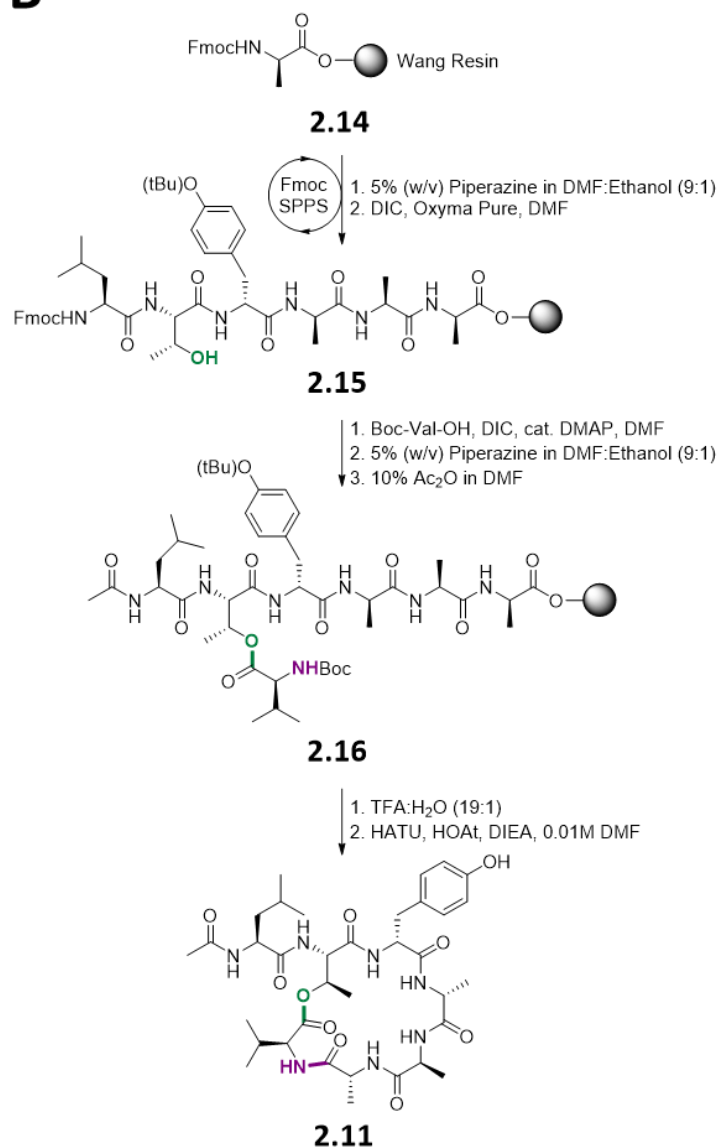
In a similar fashion, authentic Rzm A **2.11** was synthesized via macrolactamization between the L-Val₇ and D-Ala₆ amide bond (**Scheme 2.2B**). Branched peptide **2.15** was synthesized using SPPS on Wang resin, incorporating unprotected sidechain Fmoc-L-Thr-OH. On-resin esterification of L-Thr with Boc-L-Val-OH under standard DIC/DMAP conditions afforded the ester in quantitative conversion based on LCMS analysis of microcleavages. Subsequent N-terminus Fmoc deprotection and acetylation afforded peptide **2.16**, which was then cleaved off resin and fully deprotected. The crude peptide ester underwent macrolactamization under high dilution with HATU/HOAt coupling conditions to afford **2.11** following RP-HPLC purification to >95% purity (30% yield from **2.16**). Comparison of the NMR spectra recorded for our synthetic material with the NMR spectra reported for the isolated material confirmed that our product was indistinguishable from Rzm A (**Figure S2.16** and **Table S2.3**), providing a total synthesis of Rzm A and unambiguous confirmation of its chemical structure.

Scheme 2.2. SPPS of macrocyclic authentic standards Sgd E 2.3 (A) and Rzm A 2.11 (B).

A



B



Sgd- and Rzm-TE are active macrolactonization specific catalysts

To biochemically characterize the Sgd- and Rzm-TE, both were expressed and purified as stand-alone proteins. For both, codon optimized synthetic gene constructs encoding their respective C-terminal TE domains were cloned into a pET-28a(+) expression vector under the control of an inducible T7 promoter and fused to an N-terminal hexa-His (6xHis) tag. TEs were overexpressed in *Escherichia coli* BL21(DE3) and purified by metal affinity chromatography to high purity (> 95% for both, **Figure S2.17**). To assay for macrocyclization activity, Sgd-TE (5 μ M) was incubated with **2.1** (1 mM) whilst Rzm-TE (25 μ M) was incubated with **2.9** (1 mM) in 50 mM of potassium phosphate buffer (pH 7.4) at room temperature. Aliquots were taken at varying time points and analyzed by LCMS to evaluate product formation (**Figure 2.6**). Both Sgd- and Rzm-TE were shown to be catalytically active, as formation of macrocycles Sgd E **2.3** and Rzm A **2.9** were readily detected. Compared to the substrate only and boiled enzyme controls, no formation of **2.3** and **2.9** were detected within 24 hours of incubation (**Figure S2.23** and **S2.35**). **2.3** was readily detected at the 30-minute time point (**Figure 2.2A** and **S2.24**), with significant production observed after 20 hours of incubation with Sgd-TE (**Figure 2.2C** and **S2.29**). Furthermore, the macrocyclizing tolerant DEBS-TE^{48–54} was unable to macrocyclize **2.1** (**Figure S2.29**). Rzm-TE could convert 4.7% of **2.9** to **2.11** after 3 hours of incubation (**Figure 2.3A** and **S2.36**) and by 24 hours, 15% to **2.11** (**Figure 2.3B** and **S2.37**). These data are consistent with a TE-dependent rate enhancement in macrocyclization, albeit a modest enhancement for the Rzm-TE compared to other NRPS TEs such as for Tyc and Sgd.^{30,55} Because the SNAC analogue of the Sgd E precursor **2.1** was only converted into the natural product Sgd E **2.3** by Sgd-TE, it experimentally confirms that the gene cluster identified through bioinformatic analysis by Kim et al.⁴ is responsible for the biosynthesis of the Sgd family of natural products. In addition, the successful use of these TEs completes two total chemoenzymatic syntheses for Sgd E and Rzm A.

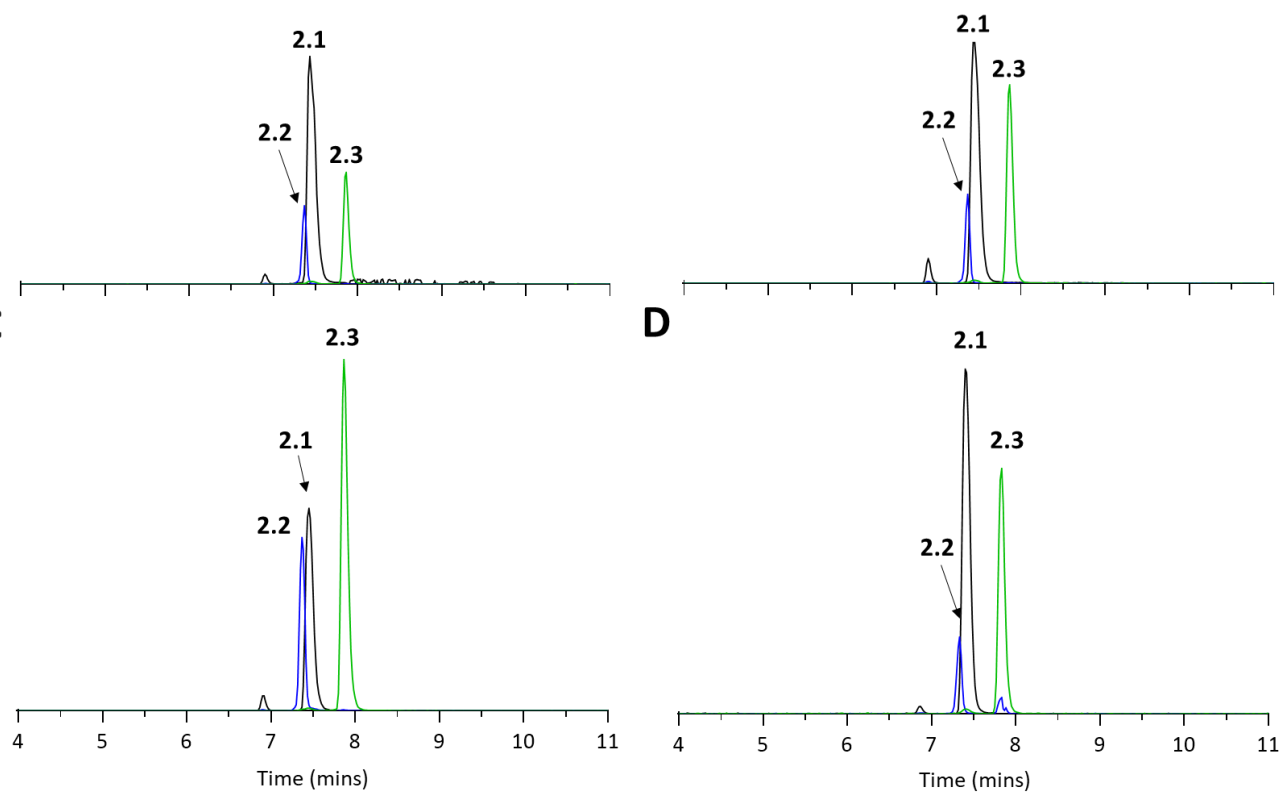
A**B****C****D**

Figure 2.2. LCMS chromatograms for Sgd-TE assays. Sgd-TE (5 μ M) assays with 1 mM of substrate **2.1** after 0.5 (A), 3 (B), and 20 (C) hours of incubation. (D) Sgd-TE 3-hour assay co-injected with authentic **2.3**. All assays were performed in 50 mM of potassium phosphate buffer (pH 7.4) at room temperature. Traces are an ion extractions for $[M + H]^+$ ions with $m/z = 994$ (**2.3**), 1012 (**2.2**) and 1113 (**2.1**).

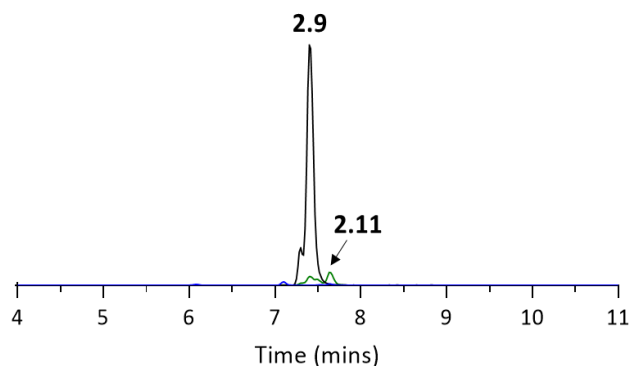
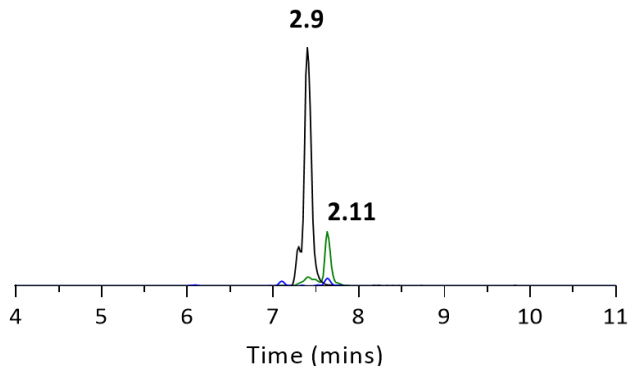
A**B**

Figure 2.3. LCMS chromatograms for Rzm-TE assays. Rzm-TE assays with 1mM of substrate **2.9** after 3 (A) and 24 (B) hours of incubation. All assays were performed in 50 mM of potassium phosphate buffer (pH 7.4) at room temperature. Traces are ion extractions for [M + H]⁺ ions with $m/z = 994$ (**2.11**) and 852 (**2.9**).

Kinetic characterization of the Sgd-TE was also performed. Sgd-TE (5 μM) was incubated for 1 hour at 25 °C with varying concentrations of substrate **2.1** (0-750 μM) in 50 mM of potassium phosphate buffer (pH 7.4). Ellman's reagent (DTNB), which tracks the release of free thiol during substrate loading onto the catalytic Ser, was included in each incubation to enable spectrophotometric and continuous quantification of Sgd-TE loading kinetics at 412 nm.^{34,53,54,56} Initial velocities were all linear over the time course of the incubations and less than 10% of **2.1** was consumed. Fitting these data to the Michaelis-Menten model showed Sgd-TE to be able to turn over **2.1** in the presence of DTNB with a $k_{cat} = (4 \pm 0.5) \times 10^{-3} \text{ s}^{-1}$ and a $K_M = 326 \pm 100 \mu\text{M}$ ($k_{cat}/K_M = 13 \pm 4 \text{ M}^{-1}\text{s}^{-1}$, **Figure 2.4**). These parameters are similar, albeit slightly lower, to those obtained from other previously characterized macrolactonizing NRPS TEs, where second order rate constants (k_{cat}/K_M) are typically 17-83 $\text{M}^{-1}\text{s}^{-1}$.^{30,32,36,57-60} In addition, in comparison to the macrolactamizing NRPS TEs Tyc and Grs, the specificity constant is $\sim 70 \times 10^3$ -fold and ~ 4.7 -fold lower respectively.⁶¹⁻⁶³ Numerous factors can affect TE rates such as the need for a PCP-TE fusion to stabilize TE

fold or the TE needs to be in their native multienzyme complexes for activity as seen in the enterobactin TE.⁶⁴

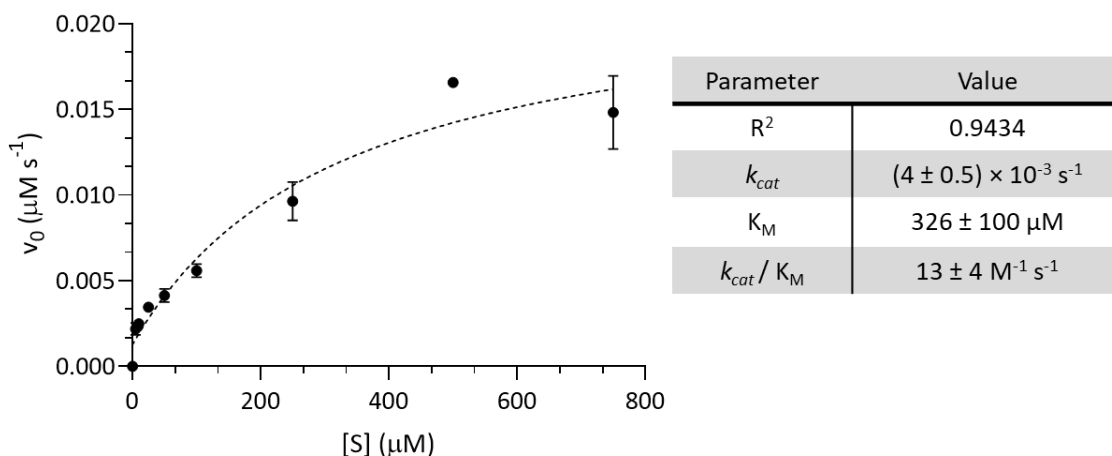


Figure 2.4. Michaelis-Menten kinetics of Sgd-TE. Sgd-TE (5 μM) was incubated with varying concentrations of substrate **2.1** (0-750 μM) and initial velocities measured for 1 hour at 25 °C in a microplate reader at 412 nm in triplicate (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted a modified Michaelis-Menten model that considers the rate of Sgd-TE background adduct formation with DTNB (dashed curve, see **Equation [2.1]**). Parameters of the fit are displayed in the table.

Kinetic characterization of the Rzm-TE with substrate **2.9** was also attempted using similar conditions for the Sgd-TE kinetics. Unfortunately, addition of DTNB into reaction mixtures led to rapid and irreversible RzmA-TE inactivation. As DTNB reacts with free thiols and RzmA-TE has four solvent accessible Cys residues (**Figure S2.18** and **S2.19A**), we hypothesize that the DTNB forms adducts with one or more of these residues and inactivates the TE. As such, kinetic characterization could not be performed; however, qualitative assessment of the extent of product formation in the product distribution assays point towards RzmA-TE being kinetically slow compared to other characterized macrolactonizing NRPS TEs (typical $k_{cat}/K_M = 17\text{-}83 \text{ M}^{-1}\text{s}^{-1}$).^{30,36,57} This is not uncommon, as the calcium dependent antibiotic (CDA) and the cereulide (Ces) forming TEs are both substantially slower, with $k_{cat}/K_M \sim 1.6 \text{ M}^{-1}\text{s}^{-1}$ and $\sim 0.3 \text{ M}^{-1}\text{s}^{-1}$, respectively.^{32,65}

Sgd-TE can effect ring-opening of authentic **2.3**

Further examination of Sgd-TE assays revealed the formation of the seco acid **2.2** could be detected above background after 1.5 hours of incubation (**Figure S2.25**), but not in the control incubations (**Figure 2.5A** and **S2.30**). After 20 hours of incubation, significantly more seco acid **2.2** was detected (**Figure 2.2C** and **S2.27**). Since both **2.3** and **2.2** are generated from the same peptidyl-TE intermediate, the relative product ratios for macrocyclization and seco acid should not change over the course of the reaction. The observation that seco acid increases with prolonged incubation (**Table S2.4**) thus suggested additional reactivity was occurring. As the Epo- and Srf-TE have been shown to effect ring opening of epothilone C (Epo C)³⁵ and surfactin (Srf),³⁶ we hypothesized that Sgd-TE could be affecting ring opening of **2.3**, ultimately leading to additional seco acid **2.2**. To test this 0.5 mM of **2.3** was incubated with Sgd-TE (5 μ M) in 50 mM of potassium phosphate buffer (pH 7.4) at room temperature. LCMS analysis showed formation of **2.2**, confirming that Sgd-TE can hydrolyze **2.3** (**Figure 2.5B** and **S2.31-34**). Compared to Srf and Sgd E, the Epo C ester bond is sterically very accessible and is known to be highly sensitive to hydrolysis, a problem that limits its serum half life in clinical settings.^{66,67} The Sgd E ester is far more hindered and thus more stable, which suggests discrete binding of **2.3** by the Sgd-TE consistent with enzyme-substrate interactions specifically stabilizing a late transition state for macrocyclization, thus binding the product in a conformation that enables attack of the active site Ser on the lactone carbonyl. This is supported by the low Michaelis constant ($K_M = 326 \pm 100 \mu\text{M}$), which is atypical of macrocyclizing PKS TEs and more similar to macrocyclizing NRPS TEs^{55,62,68,69} which, at least in the case of the Tyc-TE, bind its substrate in a product like conformation.^{55,70,71} This is consistent with our proposed CS model for TE-mediated macrocyclization. Interestingly, product inhibition of macrocyclization might be anticipated from this model; however, our kinetic characterization fits a Michaelis-Menten model without recourse to including a product inhibition step.

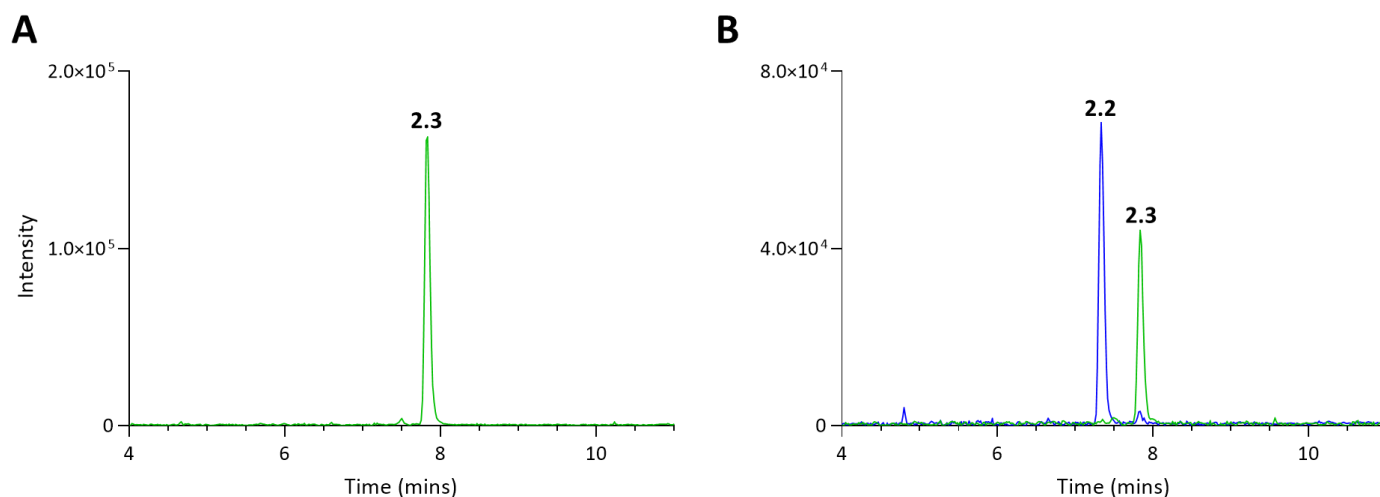


Figure 2.5. Ring opening assays with Sgd-TE and macrocycle 2.3. (A) Sgd-TE (5 μ M) assay with 0.5 mM of macrocycle 2.3 after 0 hours of incubation. Thermally denatured Sgd-TE and macrocycle only controls all showed similar results for all time points. (B) After 20 hours of incubation. All assays were performed in 50 mM of potassium phosphate buffer (pH 7.4). Traces are ion extractions for $[M + H]^+$ ions with $m/z = 994$ (2.3) and 1012 (2.2).

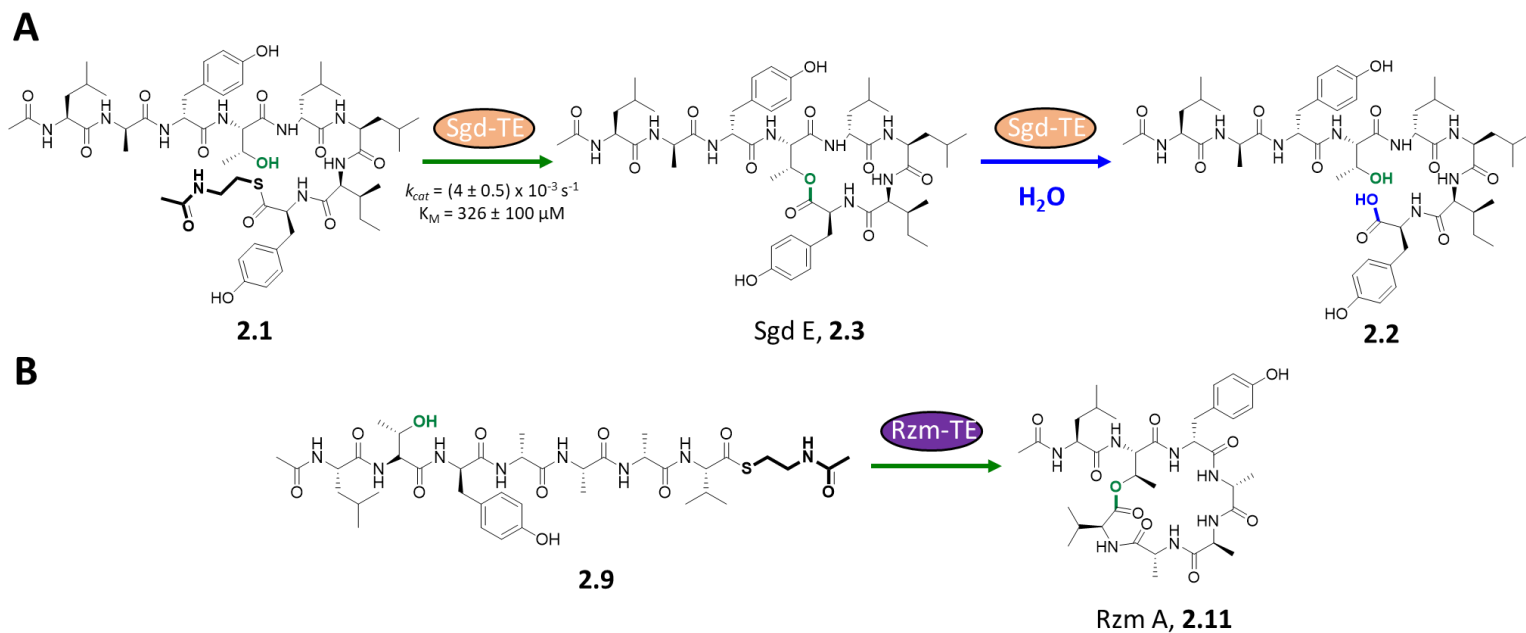


Figure 2.6. TE reactions with SNAC activated substrates. (A) Macrolactonization of SNAC substrate 2.1 catalyzed by the Sgd-TE, producing Sgd E 2.3. Kinetic parameters k_{cat} and K_M are also reported. Sgd-TE also catalyzes ring opening of 2.3 forming the seco acid 2.2. (B) Macrolactonization of SNAC substrate 2.9 catalyzed by the Rzm-TE, producing Rzm A 2.11.

NRPS TEs as macrolactonization biocatalysts in depsipeptide synthesis

Macrocyclization is a demanding type of reactivity, due to the large entropic barrier associated with end-to-end contact.^{72,73} In addition, this is made more difficult with alcohol nucleophiles, as these require activation to become sufficiently nucleophilic for macrolactonization, particularly in aqueous environments where they must outcompete water. Because of this, macrolactonization to generate the correct natural products was not possible under non-enzymatic conditions, a problem common in depsipeptide total syntheses.^{5,74–76} Ultimately we relied on macrolactamization to access authentic **2.3** and **2.11**. However, the macrocyclization to produce **2.3** suffered from unwanted isomerization, presumably due to epimerization of the activated ester.²¹ Sgd- and Rzm-TE on the other hand readily generated the depsipeptide products through macrolactonization and did not suffer from epimerization issues. In addition, the correct macrocycle is formed as the main product, which demonstrates the selectivity of these TEs in macrocyclization and thus proving themselves as powerful biocatalysts for depsipeptide synthesis.

Aside from the recently characterized Sgd and Rzm TEs, seven other macrolactonizing TEs have been used in chemoenzymatic synthesis of depsipeptides. These TEs are from the teixobactin (Txo),⁷⁷ streptogramin (SnbDE),⁵⁷ Srf,³⁶ lysobactin (LybB),⁷⁸ CDA,⁶⁵ valinomycin,³¹ and Ces³² biosynthetic pathways. Like both Sgd- and Rzm-TE, most TEs can accept SNAC substrates, though the LybB-TE required the more reactive thiophenol thioester.⁷⁸ Interestingly some are also able to process ester substrates such as the Txo⁷⁷ and Srf TEs.³⁶ These nine TEs generate 4-,⁷⁷ 5- (Sgd-TE),³⁰ 6-,^{57,79} 8-,³⁶ 9-,⁷⁸ 10-,⁶⁵ and 12-residue^{31,32} depsipeptide products. While both the SnbDE- and Rzm-TE generate 6-residue macrolactones, only the Rzm-TE catalyzes macrocyclization without substantial hydrolysis of the substrate. Despite Sgd-TE affecting ring opening of the macrocycle **2.3**, it is the only TE to date that is shown to effectively catalyze

ring closure of a 5-residue depsipeptide macrolactone. This highlights the benefit of characterizing multiple TE homologs for macrolactonization as their inherent reactivity's differ.

2.4 Conclusions

To conclude, the work presented in this chapter provided two total syntheses for the Sgd E **2.3** and Rzm A **2.11** natural products and chemoenzymatic total syntheses using the TE domains from their cognate gene clusters. First, the total synthesis of both **2.3** and **2.11** were successfully completed through a macrolactamization strategy from branched SPPS derived peptides. Consistent with the challenges of macrocyclization, the macrolactamization generating **2.3** suffered from unwanted isomerization, presumably due to epimerization of the activated ester.²¹

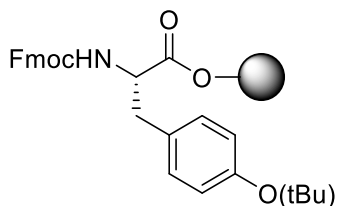
Synthesis of thioester substrates **2.1** and **2.9** was also successful with SPPS using hyper-sensitive 2-Cl-Trt Chloride resin to generate the linear protected peptides that underwent thioesterification off resin. While all attempts to chemically macrolactonize Sgd failed, the recombinantly expressed and purified Sgd-TE could form **2.3** through macrolactonization. This result confirmed the previously identified gene cluster as responsible for the biosynthesis of the Sgds family of NP. In addition, Rzm-TE was also successful in generating **2.11** through macrolactonization. This successfully completes two chemoenzymatic syntheses without any isomerization issues common in cyclic depsipeptide syntheses, highlighting the power of using TEs as macrocyclization biocatalysts. Finally, we also show that Sgd-TE can effect ring opening of **2.3** to form the seco acid **2.2**, a rare activity only seen for the Epo and Srf TEs,^{35,36} and suggests discrete binding of **2.3** with interactions specifically stabilizing a late transition state for macrocyclization. This result is well in agreement with our proposed CS model.

2.5 Materials and methods

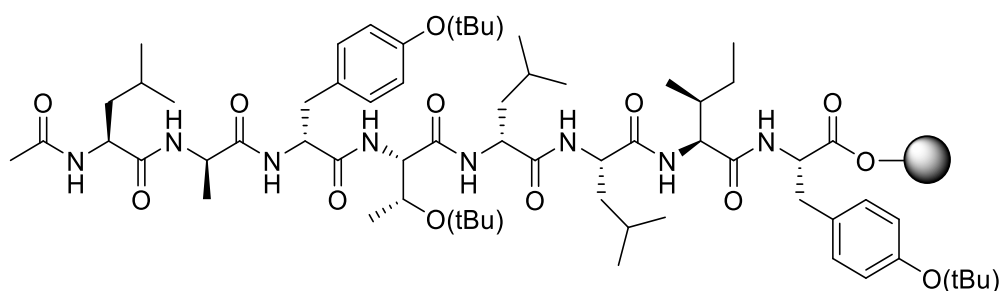
Synthetic protocols

All solvents were purchased from Fisher Scientific. All reagents were purchased at the highest available purity and used without further purification. EDC, HCl, HATU, DIEA and Boc-L-Tyr(OtBu)-OH were purchased from Oakwood Chemical (Estill, SC, USA), all Fmoc protected amino acids and HOBt were purchased from GL Biochem (Shanghai, PRC), and all other chemicals were purchased from Sigma-Aldrich (Oakville, ON, Canada or Milwaukee, WI, USA). All reactions were carried out using dry solvents under a nitrogen atmosphere unless otherwise noted. Automated peptide synthesis was performed at a 0.1 mmol scale on a CEM Liberty Blue peptide synthesizer using microwave (MW) irradiation for all steps involving heating above ambient. NMR spectroscopy was performed with either a Bruker Avance II operating at 400 MHz for ^1H and 100 MHz for ^{13}C spectra or a Bruker Avance III with a cryoprobe operating at 600 MHz for ^1H and 150 MHz for ^{13}C spectra. LCMS analysis was carried out using a Shimadzu LC20A HPLC coupled with a Shimadzu LCMS2020 single quad mass spectrometer with an ESI source. Reverse-phase purification was carried out using an Agilent 1260 infinity and a Phenomenex Kinetex 5 μm XB-O18 100Å, 150 x 10.0 mm Semi-Preparative column. High-resolution mass spectroscopy (HRMS) was conducted on a Micromass Q-TOF I (John L. Holmes Mass Spectroscopy Facility, Ottawa, ON, Canada).

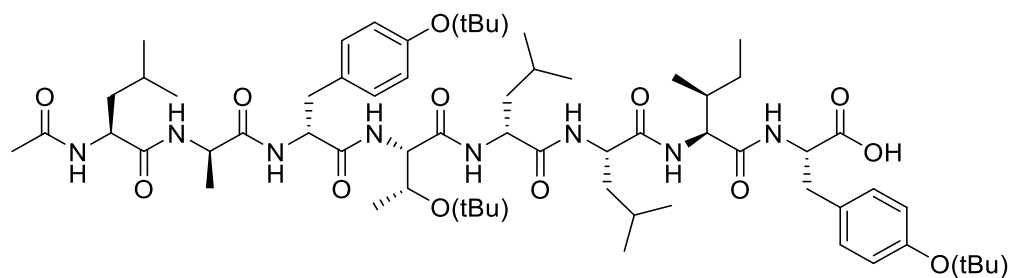
Sgd linear thioester 2.1



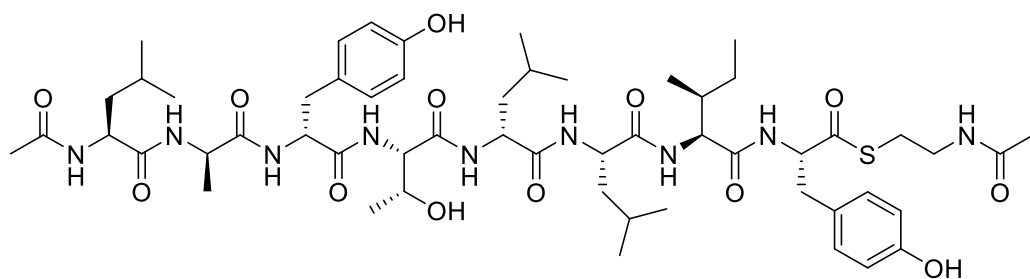
Fmoc-L-Tyr(tBu)-2-TrtCl resin (2.4). 2-Chlorotrityl chloride resin (300 mg, 200-400 mesh, 0.6 mmol/g) in a fritted peptide vessel was added DCM (5 mL), allowing the resin to swell at room temperature for 30 minutes. Fmoc-L-Tyr(tBu)-OH (230 mg, 0.5 mmol, 1 eq.) and DIEA (350 μ L, 2 mmol, 4 eq.) were added to the solution and the vessel was agitated uncapped for 20 seconds and the vessel was capped and agitated for 1.5 hours. The solution was flushed and washed with DCM (10 mL x 4). The resin was added 10 mL capping solution, DCM:MeOH:DIEA (17 mL: 2 mL: 1 mL), and agitated for 30 minutes. The solution was flushed, washed with DCM (10 mL x 4) and transferred to a round bottom flask to be dried under reduced pressure. The loading rate was determined with 2% DBU solution and spectrometric analysis (Sigma-Aldrich, Fmoc Loading Test) to afford compound **2.4** (350 mg, 0.45 mmol/g).



Protected linear Sgd E. Automated peptide synthesis was performed on a CEM Liberty Blue peptide synthesizer at a 0.1 mmol scale. Dried Fmoc-L-Tyr(tBu)-2-TrtCl resin (220 mg) was placed in the reaction vessel followed by a 5-minute swell in DMF. Couplings were performed with Fmoc-amino acid (0.2 M in DMF), HBTU (0.2 M in DMF) and DIEA (1 M in NMP) at 50 °C for 10 minutes and Fmoc deprotection were performed with 20% piperidine solution in DMF at 95 °C for 2 minutes. N-terminal acetylation was done with 10% acetic anhydride in DMF at 50 °C for 1 hour. Once complete, the resin was transferred to a fritted vessel and flushed with DCM (5 mL x 3) and dried under reduced pressure to afford crude linear peptide loaded resin. The substrate was confirmed by LCMS and used without further characterization.



Protected linear Sgd E acid (2.5). To protected linear Sgd E peptide loaded resin (100 mg) was added 5 mL cleavage solution (20% HFIP in DCM) and agitated for 30 minutes. The solution was flushed into a round bottom flask and this was repeated twice. The organic solvents were removed under reduced pressure to afford a white solid and compound **2.5** (59 mg, 74% yield) was confirmed by LCMS and ^1H NMR. ^1H NMR (400 MHz, CDCl_3) δ 8.92 (d, J = 7.6 Hz, 1H), 8.68 (d, J = 7.8 Hz, 1H), 8.59 (d, J = 9.6 Hz, 1H), 8.19 (d, J = 5.4 Hz, 1H), 7.94 (d, J = 6.2 Hz, 1H), 7.88 (d, J = 8.9 Hz, 1H), 7.05 (dd, J = 9.0, 2.6 Hz, 3H), 6.98 (d, J = 8.3 Hz, 2H), 6.86 (s, 1H), 6.58 (d, J = 8.4 Hz, 2H), 6.50 (d, J = 10.1 Hz, 1H), 5.72 (t, J = 11.7 Hz, 1H), 5.06 (t, J = 9.6 Hz, 2H), 4.95 (t, J = 6.6 Hz, 1H), 4.87 (q, J = 6.1 Hz, 1H), 4.54 (q, J = 7.6 Hz, 1H), 4.25 (t, J = 10.9 Hz, 1H), 4.12 (t, J = 4.9 Hz, 1H), 3.62 (q, J = 5.1 Hz, 1H), 3.18 – 2.91 (m, 3H), 2.52 (t, J = 12.4 Hz, 1H), 1.73 (d, J = 7.9 Hz, 1H), 1.56 (t, J = 10.3 Hz, 5H), 1.38 (td, J = 12.2, 4.5 Hz, 2H), 1.32 – 1.25 (m, 2H), 1.23 (s, 9H), 1.20 (s, 9H), 1.15 (s, 9H), 0.84 (ddt, J = 12.2, 6.3, 3.9 Hz, 27H), 0.64 (t, J = 7.3 Hz, 3H), 0.29 (d, J = 6.3 Hz, 3H).

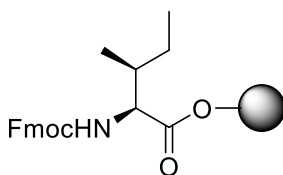


Sgd E *N*-acetylcysteamine thioester (2.1). Compound **2.5** (42 mg, 0.0365 mmol, 1 eq.) and HOBt (5.2 mg, 0.0383 mmol, 1.05 eq.) were dissolved in DCM (0.1 M) and stirred at 0 °C for 10 minutes. EDC-HCl (6 mg,

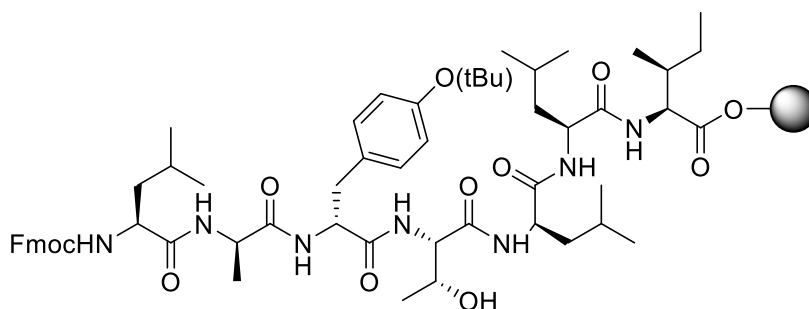
0.0383 mmol, 1.05 eq.) was added to the stirring solution and maintained at 0 °C for 30 minutes. The solution was added *N*-acetylcysteamine (5.2 mg, 0.0438 mmol, 1.2 eq.) and was stirred at room temperature overnight. The solution was diluted with DCM (5 mL) and quenched with NaHCO₃ (aq.) (5 mL). The aqueous phase was extracted with DCM (3x 5 mL) and the organic phase was separated, washed with brine (5 mL), dried over MgSO₄ and removed under reduced pressure. The thioester protected peptide was purified by column chromatography (2.5-5% MeOH in DCM, 12 mg, 26%). R_f (5% MeOH in DCM) = 0.55. The substrate was confirmed by LCMS and used without further characterization.

The purified protected thioester peptide was dissolved in 1 mL TFA:H₂O (19:1) and the solution was stirred at room temperature for 2 hours or until the reaction was judged to be complete by TLC. The solution was precipitated out of cold ether and left in the -20 °C freezer overnight. The solution was centrifuged and the supernatant was decanted and the residue was further dried by N₂ stream. The residue was purified by RP-HPLC to afford compound **2.1** (4.5 mg, 42%). ¹H NMR (600 MHz, MeOD) δ 7.08 (d, *J* = 8.5 Hz, 2H), 7.02 (d, *J* = 8.5 Hz, 2H), 6.73 – 6.67 (m, 4H), 4.70 (dd, *J* = 9.2, 5.4 Hz, 1H), 4.47 (ddd, *J* = 17.1, 9.3, 5.8 Hz, 2H), 4.39 (dd, *J* = 9.7, 4.9 Hz, 1H), 4.34 (dd, *J* = 8.5, 6.6 Hz, 1H), 4.25 (t, *J* = 7.2 Hz, 3H), 4.19 (dd, *J* = 6.5, 4.3 Hz, 1H), 4.01 (d, *J* = 4.3 Hz, 1H), 3.10 (dd, *J* = 14.4, 5.5 Hz, 1H), 3.07 – 3.01 (m, 1H), 3.02 – 2.98 (m, 4H), 2.85 (dd, *J* = 14.2, 9.2 Hz, 2H), 2.04 (s, 4H), 1.91 (s, 3H), 1.74 – 1.47 (m, 13H), 1.33 – 1.22 (m, 9H), 1.17 – 1.09 (m, 2H), 1.01 – 0.83 (m, 27H). HRMS (ESI+) exact mass calculated for C₅₅H₈₅N₉O₁₃SNa 1134.5895, found 1134.5885.

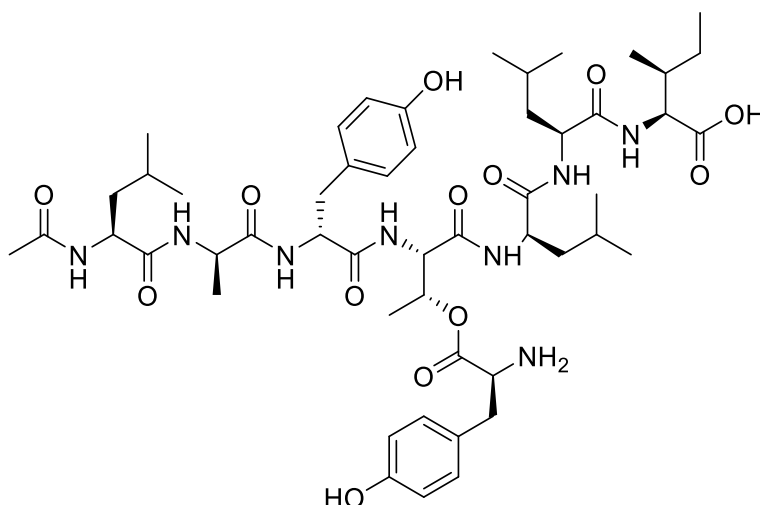
Sgd E macrocycle **2.3**



Fmoc-L-Ile-Wang resin (2.6). Wang resin (200 mg, 100-200 mesh, 1.2 mmol/g) in a fritted peptide vessel was added DMF (5 mL), allowing the resin to swell at room temperature for 30 minutes. Fmoc-L-Ile-OH (848 mg, 2.4 mmol, 10 eq.) and DIC (188 μ L, 1.2 mmol, 5 eq.) were dissolved in DCM (5 mL) in a round bottom flask and stirred at 0 °C for 20 minutes. The solvents were removed under reduced pressure and DMF (0.5 mL) and DMAP (3 mg, 0.024 mmol, 0.1 eq.) were added to the amino acid anhydride mixture which was then transferred to the swelled resin. The solution was agitated for 1 hour at room temperature then flushed, washed with DCM (10 mL x 2), DMF (10 mL x 2) and DCM (10 mL x 2) and transferred to a round bottom flask to be dried under reduced pressure. The loading rate was determined with 2% DBU solution and spectrometric analysis (Sigma-Aldrich, Fmoc Loading Test) to afford compound **2.6** (250 mg, 0.49 mmol/g).



Linear Sgd E precursor. Automated peptide synthesis was performed on a CEM Liberty Blue peptide synthesizer at a 0.1 mmol scale. Dried Fmoc-L-Ile-Wang resin (200 mg) was placed in the reaction vessel followed by a 5-minute swell in DMF. Couplings were performed with Fmoc-amino acids (0.2 M in DMF), DIC (0.2 M in DMF) and Oxyma Pure (1 M in DMF) at 95 °C for 1.5 minutes and Fmoc deprotection were performed with 20% piperidine in DMF solution at 95 °C for 2 minutes. The synthesis was paused directly following last Fmoc-L-Leu-OH and the resin was transferred to a fritted vessel and flushed with DCM (5 mL x 1).

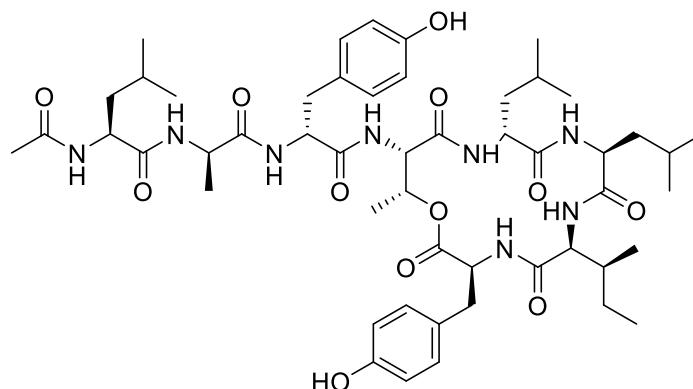


Esterification (2.8). Boc-L-Tyr(tBu)-OH (mg, 4.08 mmol, 10 eq.) was dissolved in DCM (5 mL) and cooled to 0 °C. DIC (mg, mmol, 10 eq.) was added and the solution was stirred at room temperature for 30 minutes. Half the DCM was removed under N₂ stream and DMF (2.5 mL) was added and the solution was added to the resin in the fritted vessel followed by HOBt (mg, mmol, 0.2 eq.) and the vessel was agitated at room temperature for 3 hours. The solution was flushed and washed with DCM (5 mL x 3). The substrate was confirmed by LCMS and used without further characterization.

Linear Sgd E (2.7). The resin was transferred to the automated peptide synthesis reaction vessel and synthesis was resumed. N-terminal Fmoc deprotection was performed with 20% piperidine in DMF at 95 °C for 2 minutes and acetylation was done with 10% acetic anhydride at 95 °C for 1 hour. The resin was removed from the instrument and placed in a fritted peptide vessel. The substrate was confirmed by LCMS and used without further characterization.

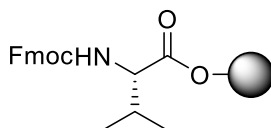
Peptide deprotection and resin cleavage. The peptide bound resin was added 1 mL TFA:H₂O (19:1) and the solution was stirred at room temperature for 2 hours. The solution was precipitated out of cold ether and left in the -20 °C freezer overnight. The solution was centrifuged, and the supernatant was decanted

and the residue was further dried by N₂ stream. Compound **2.8** (63 mg, 65%) was confirmed by LCMS and used without further purification.

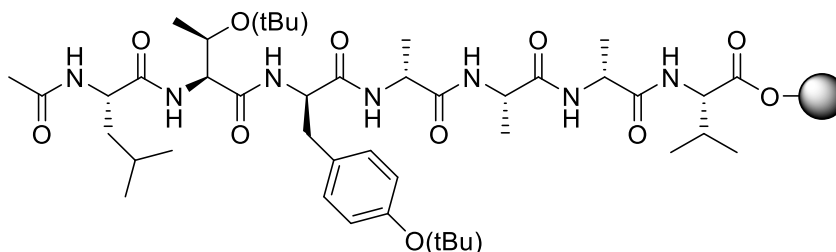


Sgd E (2.3). Linear Sgd E ester **2.8** (10.5 mg, 0.010 mmol, 1 eq.) and DIEA (2 μ L, 0.011 mmol, 1.1 eq.) were dissolved in DMF (1 mL). HATU (4 mg, 0.0105 mmol, 1.05 eq.) and HOAt (7 μ L, 1.6 M in DMF) were dissolved in DMF (100 μ L) and added to the stirring solution dropwise via syringe pump over 4 hours at room temperature. The solution was quenched with NH₄Cl and extracted with EtOAc (3x 10 mL). The combined organic fractions were dried over MgSO₄ and removed under reduced pressure. The residue was purified by RP-HPLC to afford compound **2.3** (3.1 mg, 32%). ¹H NMR (600 MHz, DMSO) δ 9.20 (d, *J* = 3.8 Hz, 2H), 8.39 (d, *J* = 7.1 Hz, 1H), 8.14 (d, *J* = 8.9 Hz, 1H), 8.09 (d, *J* = 7.6 Hz, 1H), 8.06 – 8.00 (m, 2H), 7.95 (d, *J* = 5.3 Hz, 1H), 7.73 (d, *J* = 9.3 Hz, 1H), 7.25 (d, *J* = 7.2 Hz, 1H), 7.03 (d, *J* = 8.3 Hz, 3H), 6.91 (d, *J* = 8.2 Hz, 2H), 6.64 (dd, *J* = 13.8, 8.3 Hz, 4H), 5.02 (s, 1H), 4.48 – 4.37 (m, 2H), 4.26 (p, *J* = 7.6 Hz, 3H), 4.19 (dd, *J* = 8.7, 2.7 Hz, 1H), 4.01 (d, *J* = 15.6 Hz, 1H), 3.88 (t, *J* = 9.5 Hz, 1H), 2.96 (dd, *J* = 13.8, 4.9 Hz, 1H), 2.89 – 2.83 (m, 1H), 2.79 (d, *J* = 7.6 Hz, 2H), 1.75 (s, 4H), 1.64 – 1.51 (m, 2H), 1.50 – 1.43 (m, 1H), 1.41 (t, *J* = 7.1 Hz, 4H), 1.37 (s, 0H), 1.23 (s, 6H), 1.14 (d, *J* = 7.0 Hz, 4H), 1.02 (dd, *J* = 14.3, 7.2 Hz, 1H), 0.89 (d, *J* = 6.5 Hz, 3H), 0.87 (d, *J* = 6.6 Hz, 2H), 0.86 – 0.83 (m, 3H), 0.83 (d, *J* = 3.8 Hz, 7H), 0.80 (d, *J* = 6.8 Hz, 6H). HRMS (ESI⁺) exact mass calculated for C₅₁H₇₆N₈O₁₂Na 1015.5301, found 1015.5301.

Rzm linear thioester **2.9**



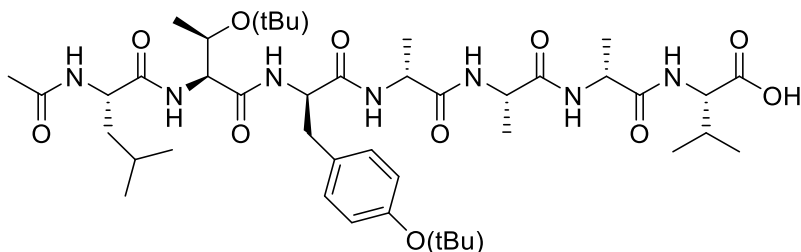
Fmoc-L-Val-2-Cl-Trt resin (2.12). 2-Chlorotrityl chloride resin (500 mg, 200-400 mesh, 1.1 mmol/g) in a fritted peptide vessel was added DCM (5 mL), allowing the resin to swell at room temperature for 30 minutes. Fmoc-L-Val-OH (224 mg, 0.66 mmol, 1.2 eq.) and DIPEA (380 μ L, 2.2 mmol, 4 eq.) were added to the solution and the vessel was agitated uncapped for 20 seconds, then the vessel was capped and agitated for 1.3 hours. The solution was flushed and washed with DCM (10 mL $\times$ 4). The resin was added 10 mL capping solution, DCM:MeOH:DIPEA (17:2:1), and agitated for 20 minutes. The solution was flushed, washed with DCM $\times$ 2, DMF $\times$ 2, DCM $\times$ 2 and transferred to a round bottom flask to be dried under reduced pressure. The loading rate was determined with 2% DBU solution and spectrometric analysis (Sigma-Aldrich, Fmoc Loading test) to afford compound **2.12** (580 mg, 0.33 mmol/g).



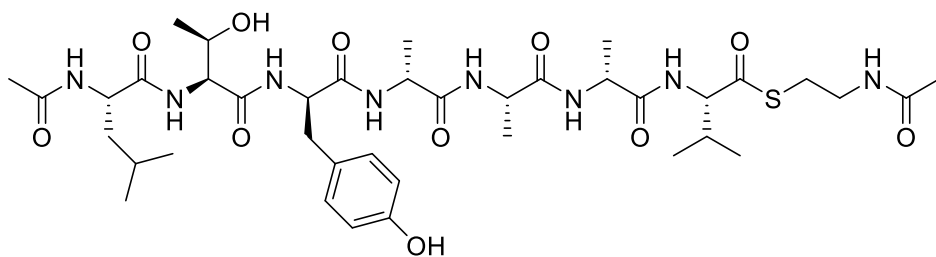
Protected linear Rzm A. Automated peptide synthesis was performed on a CEM Liberty Blue peptide synthesizer at a 0.1 mmol scale. Dried Fmoc-L-Val-2-Cl-Trt resin (303 mg) was placed in the reaction vessel followed by a 5-minute swell in DMF. Couplings were performed with Fmoc-amino acid (0.2 M in DMF), HBTU (0.2 M in DMF) and DIPEA (1 M in DMF) as a single coupling and Fmoc deprotection were performed with 5% piperazine solution in NMP:EtOH (90:10). N-terminal acetylation was done with 10% acetic anhydride in DMF at 50 $^{\circ}$ C for 1 hour. Once complete, the resin was transferred to a fritted vessel and

flushed with DCM (5 mL $\times$ 3) and dried under reduced pressure to afford crude linear peptide loaded resin.

The substrate was confirmed by LCMS and used without further characterization.

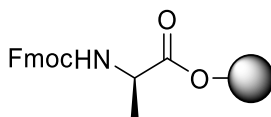


Protected linear Rzm A acid (2.13). To protected linear Rzm A peptide loaded resin (75 mg) was added 5 mL of cleavage solution (20% HFIP in DCM) and agitated for 30 minutes. The solution was flushed into a round bottom flask and this was repeated twice. The organic solvents were removed under reduced pressure and purified by column chromatography (5-10% MeOH in DCM) to afford **2.5** (47 mg, 62%) as a white solid. ^1H NMR (600 MHz, MeOD) δ 7.28 (d, J = 8.2 Hz, 2H), 6.91 (d, J = 8.5 Hz, 2H), 4.68 – 4.56 (m, 1H), 4.52 (dd, J = 11.1, 3.7 Hz, 1H), 4.47 (d, J = 7.6 Hz, 1H), 4.38 (d, J = 7.5 Hz, 1H), 4.33 (d, J = 7.1 Hz, 1H), 4.28 (d, J = 7.2 Hz, 1H), 4.17 (d, J = 5.3 Hz, 1H), 3.93 (qt, J = 6.4, 3.0 Hz, 1H), 3.19 (dd, J = 14.3, 3.7 Hz, 2H), 2.82 (dd, J = 14.3, 11.1 Hz, 1H), 2.28 – 2.18 (m, 1H), 2.01 (s, 3H), 1.74 – 1.64 (m, 1H), 1.61 (dd, J = 8.0, 6.7 Hz, 3H), 1.41 – 1.38 (m, 9H), 1.32 (s, 9H), 1.07 (s, 9H). ^{13}C NMR (151 MHz, MeOD) δ 173.95, 173.75, 173.06, 172.99 – 172.85, 172.21, 154.02, 132.35, 129.65, 123.93, 105.62, 78.13, 74.51, 66.94, 60.91 – 59.77, 59.48 – 58.36, 56.18 – 54.44, 52.23 – 51.29, 39.92, 36.79 – 36.13, 30.51, 27.86, 27.20, 24.55, 22.26, 21.15, 20.37, 18.96, 17.13, 16.73, 15.53



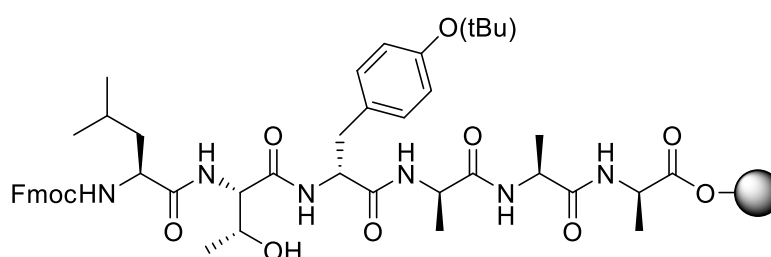
Linear Rzm A *N*-acetylcysteamine thioester (2.9). Compound **2.13** (20 mg, 0.0231 mmol, 1 eq.) and HOBT (3.29 mg, 0.0243 mmol, 1.05 eq.) were dissolved in DCM (0.1 M) and stirred at 0 °C for 10 mins. EDC·HCl (4.67 mg, 0.0243 mmol, 1.05 eq.) was added to the stirring solution and maintained at 0 °C for 30 mins. The solution was added *N*-acetylcysteamine (3.30 mg, 0.0277 mmol, 1.2 eq.) and was stirred at room temperature overnight. The solution was diluted with DCM (5 mL) and quenched with NaHCO₃ (aq.) (5 mL). The aqueous phase was extracted with DCM (3 × 5 mL) and the organic phase was separated, washed with brine (5 mL), dried over MgSO₄ and removed under reduced pressure to afford crude protected thioester peptide. The crude protected thioester peptide was dissolved in 0.5 mL TFA:H₂O (19:1) and the solution was stirred at room temperature for 2 hours or until the reaction was judged to be complete by TLC. The residue was further dried by N₂ stream before it was purified by RP-HPLC to afford compound **2.9** (1 mg, 5% from 2.13). ¹H NMR (600 MHz, DMSO) δ 8.23 (d, *J* = 8.6 Hz, 1H), 8.11 – 8.06 (m, 3H), 8.03 (d, *J* = 6.9 Hz, 2H), 7.98 (d, *J* = 8.1 Hz, 1H), 7.87 (d, *J* = 7.3 Hz, 2H), 7.02 (d, *J* = 8.0 Hz, 2H), 4.38 – 4.33 (m, 3H α), 4.30 – 4.21 (m, 3H α), 4.03 (d, *J* = 7.3 Hz, 1H α), 3.90 (s, 1H β), 3.16 – 3.10 (m, S-CH₂), 2.98 (d, *J* = 13.3 Hz, 1H β), 2.88 (t, *J* = 7.0 Hz, S-C-CH₂), 2.68 – 2.61 (m, 1H β), 2.17 – 2.11 (m, 1H), 1.84 (s, 4H), 1.77 (s, 3H), 1.62 – 1.57 (m, 1H), 1.44 (t, *J* = 7.4 Hz, 3H), 1.23 (t, *J* = 6.0 Hz, 7H), 1.20 (d, *J* = 7.1 Hz, 3H), 0.90 (d, *J* = 6.6 Hz, 4H), 0.85 (q, *J* = 6.4 Hz, 14H). HRMS (ESI+) exact mass calculated for C₃₉H₆₂N₈O₁₁S 873.4131, found 873.4156.

Rzm A macrocycle 2.11

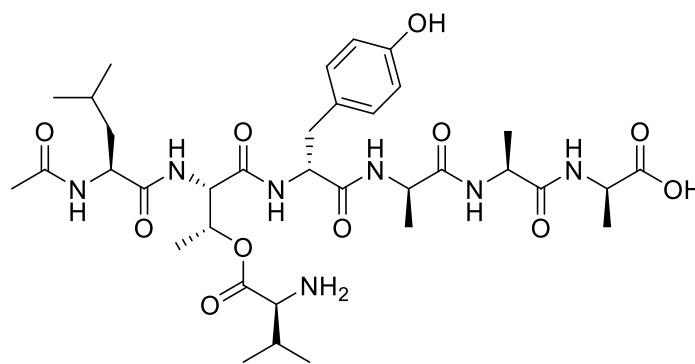


Fmoc-D-Ala-Wang resin (2.14). Wang resin (500 mg, 100-200 mesh, 0.83 mmol/g) in a fritted peptide vessel was added DMF (5 mL), allowing the resin to swell at room temperature for 30 minutes. Fmoc-D-Ala-OH (1.29 g, 4.15 mmol, 10 eq.) and DIC (325 μ L, 2.075 mmol, 5 eq.) were dissolved in DCM (12 mL) in

a round bottom flask and stirred at 0 °C for 20 minutes. The solvents were removed under reduced pressure and DMF (1 mL) and DMAP (5 mg, 0.0415 mmol, 0.1 eq.) were added to the amino acid anhydride mixture which was then transferred to the swelled resin. The solution was agitated for 1 hour at room temperature then flushed, washed with DCM (10 mL × 2), DMF (10 mL × 2) and transferred to a round bottom flask to be dried under reduced pressure. The loading rate was determined with 2% DBU solution and spectrometric analysis (Sigma-Aldrich, Fmoc-Loading Test) to afford compound **2.14** (625 mg, 0.42 mmol/g).



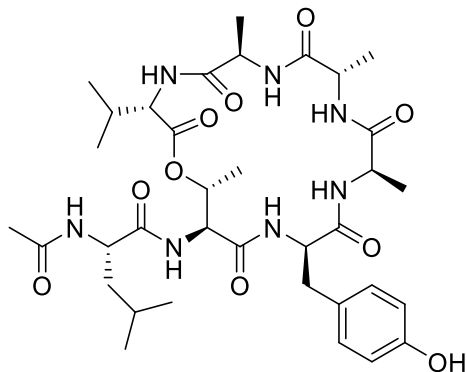
Linear Rzm A precursor (2.15). Automated peptide synthesis was performed on a CEM Liberty Blue peptide synthesizer at a 0.1 mmol scale. Dried Fmoc-D-Ala-Wang resin (238 mg) was placed in the reaction vessel followed by a 5-minute swell in DMF. Couplings were performed with Fmoc-amino acids (0.2 M in DMF), DIC (0.2 M in DMF) and Oxyma Pure (1 M in DMF) at 95 °C for 1.5 minutes and Fmoc deprotection were performed with 5% piperazine solution in NMP:EtOH (90:10) solution at 95 °C for 2 minutes. The synthesis was paused directly following the last Fmoc-L-Leu-OH, the resin was transferred to a fritted vessel and flushed with DCM (5 × 1 mL).



Esterification. Boc-L-Val-OH (108 mg, 0.5 mol, 10 eq.) was dissolved in DCM (6 mL) and cooled to 0 °C. To this solution, DIC (50 μ L, 0.5 mmol, 10 eq.) was added and the solution was stirred at room temperature for 30 minutes. Half the DCM was removed under N₂ stream and DMF (3 mL) was added and the solution was added to the resin in the fritted vessel followed by DMAP (1.2 mg, 0.01 mmol, 0.2 eq.) and the vessel was agitated at room temperature for 24 hours. The solution was flushed and washed with DCM (6 mL $\times$ 3). The substrate was confirmed by LCMS and used without further characterization.

Linear Rzm A (2.16). The resin was transferred to the automated peptide synthesis reaction vessel and synthesis was resumed. N-terminal Fmoc deprotection was performed with 5% piperazine solution in NMP:EtOH (9:1) at 95 °C for 2 minutes and acetylation was done with 10% acetic anhydride at 95 °C for 1 hour. The resin was removed from the instrument and placed in a fritted peptide vessel. The substrate was confirmed by LCMS and used without further characterization.

Peptide deprotection and resin cleavage. The peptide bound resin was added 1 mL TFA:H₂O (19:1) and the solution was stirred at room temperature for 2 hours. The solution was precipitated out of cold ether and left in the -20 °C freezer overnight. The solution was centrifuged, the supernatant decanted and the residue was further dried by N₂ stream. The substrate was confirmed by LCMS and used without further characterization.



Rzm A (2.11). Branched Rzm A ester **2.16** (20.8 mg, 0.0223 mmol, 1 eq.) and DIEA (4.3 μ L, 0.0246 mmol, 1.1 eq.) were dissolved in DMF (2.23 mL). HATU (8.9 mg, 0.023 mmol, 1.05 eq.) and HOAt (38 μ L, 0.6 M in DMF) were dissolved in DMF (230 μ L) and added to the stirring solution dropwise via syringe pump over 22 hours at room temperature. The solution was quenched with NH_4Cl and extracted with EtOAc (3 $\times$ 20 mL). The combined organic fractions were dried over MgSO_4 and removed under reduced pressure. The residue was purified by RP-HPLC to afford compound **2.11** (4.9 mg, 30%). ^1H NMR (600 MHz, DMSO) δ 8.66 (d, J = 7.2 Hz, 1H), 8.38 (d, J = 8.1 Hz, 1H), 8.21 (d, J = 5.1 Hz, 1H), 8.09 (d, J = 7.7 Hz, 1H), 7.80 (d, J = 8.9 Hz, 1H), 7.40 (d, J = 9.1 Hz, 1H), 7.33 (d, J = 7.8 Hz, 1H), 7.08 (d, J = 6.8 Hz, 2H), 6.67 (d, J = 6.6 Hz, 2H), 5.24 (qd, J = 6.5, 2.4 Hz, 1H), 4.53 (d, J = 2.2 Hz, 1H), 4.50 (dd, J = 9.0, 2.4 Hz, 1H), 4.43 (td, J = 8.1, 4.8 Hz, 1H), 4.33 (p, J = 7.2 Hz, 1H), 4.27 – 4.22 (m, 1H), 4.22 – 4.19 (m, 1H), 4.07 (dd, J = 7.0, 5.2 Hz, 1H), 3.03 (dd, J = 14.3, 4.7 Hz, 1H), 2.87 (dd, J = 14.3, 8.9 Hz, 1H), 1.91 (s, 3H), 1.90 – 1.86 (m, 1H), 1.71 – 1.62 (m, 1H), 1.58 (ddd, J = 13.7, 10.5, 4.7 Hz, 1H), 1.46 (ddd, J = 13.9, 9.3, 4.7 Hz, 1H), 1.33 (d, J = 6.9 Hz, 3H), 1.27 (dd, J = 11.6, 7.2 Hz, 6H), 1.06 (d, J = 6.5 Hz, 3H), 0.94 (d, J = 6.6 Hz, 3H), 0.88 (d, J = 6.5 Hz, 3H), 0.83 (d, J = 6.7 Hz, 3H), 0.77 (d, J = 6.7 Hz, 3H). ^{13}C NMR (151 MHz, DMSO) δ 173.22, 173.08, 172.33, 172.15, 171.38, 170.53, 169.94, 168.62, 158.86, 158.62, 156.43, 130.52, 127.57, 115.51, 58.00, 55.66, 55.04, 51.40, 49.97, 49.54, 49.13, 36.54, 36.52, 30.78, 24.51, 23.62, 22.85, 21.70, 19.40, 19.36, 18.40, 18.05, 17.61, 16.75, 15.95. HRMS (ESI+) exact mass calculated for $\text{C}_{35}\text{H}_{53}\text{N}_7\text{O}_{10}$ 754.3750, found 754.3752.

Cloning, expression and purification

Cloning and expression. Two synthetic gene constructs were designed: one using the seongsanamide (Sgd) TE coding sequence from *Bacillus safensis* genome⁴ (CP018197.1 range 2,619,962 - 2,620,768) and codon optimized for *E.coli* (GeneArt, ThermoFisher). A second was designed using the rhizomide thioesterase (Rzm-TE) peptide sequence (WP_013428324, range 7387-end), reverse translated into its coding DNA sequence using EMBOSS Backtranseq webserver and codon optimized for *E. coli* (GeneArt,

ThermoFisher). Constructs were cloned into NdeI and EcoRI linearized pET-28a(+) expression vector, generating pJBH19 and pJBH55 for Sgd- and Rzm-TE respectively. Resulting constructs were confirmed by Sanger sequencing (Génome Québec). Expression vectors were transformed into *E. coli* BL21(DE3) chemically competent cells for protein expression. A flask containing 400 mL for Sgd-TE and 100 mL for Rzm-TE of LB media supplemented with 50 µg/mL of kanamycin were inoculated with 0.5 % v/v of dense transformed overnight cultures. Cultures were grown at 37 °C with 200 rpm shaking to an OD₆₀₀ of 0.6-0.8. Protein expression was then induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cultures were then incubated at 16 °C with 200 rpm shaking for 18 hours. DEBS-TE was expressed from the pRS33 construct as previously described.^{53,54}

at -80 °C. Protein concentration was determined by Bradford assay (#5000006, Bio-Rad). DEBS-TE was purified as previously described.⁵³

For Rzm-TE, cells were harvested by centrifugation at 4,000 g and resuspended in 2.5 mL of lysis buffer (50 mM potassium phosphate, 1 mg/mL lysozyme, 1 µg/mL of pepstatin A and 1 µg/mL leupeptin, pH 7.4). Cells were lysed by sonication on ice and cell debris removed by centrifugation at 12,000 g for 1 hour. Resulting lysate was incubated with 250 µL of Ni-NTA resin slurry (#30210, QIAGEN) on a rocking platform shaker for 1.5 hours. Lysate was then loaded into a gravity flow column and resin allowed to settle by gravity. Purification buffer (50 mM potassium phosphate, pH 7.4) was used at the indicated imidazole concentrations for resin washing and protein elution. Resin was first washed once with 0 mM imidazole and once more with 20 mM imidazole. Protein was eluted twice with 100 mM imidazole and then twice more with 250 mM imidazole. Protein collected during elution was dialyzed in a Slide-A-Lyzer 10 kDa MWCO dialysis cassette (#66810, ThermoFisher) as per the manufacturer's recommended protocol in 500× sample volume of 50 mM potassium phosphate buffer (pH 7.4). Dialyzed sample was collected and used the same day.

Biochemical assays

Michaelis-Menten kinetics of Sgd-TE. Continuous enzymatic assays using DTNB (#D8130, Sigma-Aldrich) to track release of SNAC free thiol upon substrate loading was performed in a half-area 96-well plate (50 mM potassium phosphate buffer pH 7.4, 7.5% v/v DMSO, 4% v/v DTNB, 5 µM Sgd-TE) and varying concentrations of substrate **2.1** (0 to 750 µM) in triplicate. First, Sgd-TE was dialysed with dialysis buffer (5 spins, 14,000 g, 15 minutes, 4 °C) to remove glycerol using Amicon Ultra-0.5 10 kDa MWCO tubes (#UFC5010, Millipore). A 2X substrate master mix (containing potassium phosphate buffer, DMSO and substrate) and Sgd-TE master mix (containing potassium phosphate buffer, DTNB and Sgd-TE) were prepared separately and arrayed into a half-area 96-well plate. Sgd-TE master mix was mixed into each

substrate master mix well and absorbance measurements were taken every 15 seconds for 1 hour at 412 nm and 25 °C in a microplate reader. An SNAC thiolate standard curve (0 to 100 µM) was generated to convert absorbance readings into concentrations (µM). Initial velocities were plotted against substrate concentration and the data fitted to a modified Michaelis-Menten model that considers the rate of Sgd-TE background adduct formation by DTNB (v_{bck} , see **Equation [2.1]**) using GraphPad Prism 8.

$$v_0 = \left(\frac{V_{\max} \cdot [S]}{K_M + [S]} \right) + v_{bck} \quad \text{Equation [2.1]}$$

Product distribution and macrocycle ring opening assays for Sgd-TE. Sgd-TE was dialysed with dialysis buffer (5 spins, 14,000 g, 15 minutes, 4 °C) to remove glycerol using Amicon Ultra-0.5 10 kDa MWCO tubes (#UFC5010, Millipore). Assays were carried out in eppendorf tubes in a 60 µL final reaction volume (50 mM potassium phosphate buffer pH 7.4, 7.5% v/v DMSO, 1 mM substrate **2.1**, 5 µM Sgd-TE). A substrate and negative control were prepared similarly as described above with the addition of water and heat denatured Sgd-TE (5 µM) in place of functional Sgd-TE respectively. A set of assays was prepared using recombinantly purified DEBS-TE, substrate only and heat denatured TE controls. Macrocycle ring opening assays were prepared similarly as described above except with 0.5 mM macrocycle authentic standard **2.3** in place of substrate. Reactions were left to incubate on the benchtop at room temperature for turnover.

Aliquots at time points 0, 0.5, 1.5, 3 and 20 hours were taken and quenched in 3 volumes of LCMS Optima grade acetonitrile (MeCN) supplemented with 0.05% v/v LCMS Optima grade formic acid (FA). DEBS-TE assays were quenched after 20 hours only. Quenched aliquots were left on the benchtop for 15 minutes, then centrifuged at max speed for 10 minutes at room temperature before analysis by LCMS (Shimadzu Prominence UFLC system with SPD-20A Prominence UV/VIS Detector, LCMS-2020 single quadrupole MS). Chromatographic separations were done using an Agilent Eclipse Plus C18 (3.5-micron, 150×4.6 mm) reverse phase column in ESI+ mode with a flow rate of 1 mL/min. The mobile phase was

composed of LCMS Optima grade H₂O (A) and MeCN (B) both supplemented with 0.05% v/v FA. Samples were monitored at 210 and 254 nm with an elution gradient of 10% B to 95% B over 15 minutes.

Product distribution assays for Rzm-TE. Assays were carried out in triplicate in eppendorf tubes in a 60 µL final reaction volume (50 mM potassium phosphate buffer pH 7.4, 10% v/v DMSO, 1 mM substrate **2.9**, 25 µM Rzm-TE). A substrate and negative control were prepared similarly as described above with the addition of water and heat denatured Rzm-TE (25 µM) in place of functional Rzm-TE respectively. Reactions were left to incubate on the benchtop at room temperature for turnover. Aliquots at time points 0, 3 and 24 hours were taken and quenched in 3 volumes of LCMS Optima grade acetonitrile (MeCN) supplemented with 0.05% v/v LCMS Optima grade formic acid (FA). Quenched aliquots were left on the benchtop for 15 minutes, then centrifuged at max speed for 10 minutes at room temperature before analysis by LCMS on a Shimadzu LC20A HPLC coupled with a Shimadzu LCMS2020 single quad mass spectrometer and an ESI source. Chromatographic separations were done using an Agilent Eclipse XDB-C18 (5-micron, 150 × 4.6 mm) reverse phase column in ESI+ mode with a flow rate of 1 mL/min. The mobile phase was composed of LCMS Optima grade H₂O (A) and MeCN (B) both supplemented with 0.05% v/v FA. Samples were monitored at 210 and 254 nm with an elution gradient of 10% B to 75% B over 10 minutes before increasing to 95% B over 2 minutes.

2.6 Supplementary information

Compound purity tables

Table S2.1. Purity and expected/observed m/z of Sgd peptides used.

Compounds	Molecular weight	Purity	Expected m/z	Observed m/z
SNAC-Sgd E 2.1	1112.40	99 %	1134.5895	1134.5885
Sgd E hydrolysis 2.2 *	1011.23	N/A	N/A	N/A
Sgd E 2.3	993.21	99 %	1015.5301	1015.5301

* **2.2** was not synthesized, detection only on low-res Shimadzu LCMS-2020 single quadrupole.

Table S2.2. Purity and expected/observed m/z of Rzm peptides used.

Compounds	Molecular weight	Purity	Expected m/z	Observed m/z
SNAC-Rzm A 2.9	851.03	93 %	873.4131	873.4156
Rzm A hydrolysis 2.10 *	749.86	N/A	N/A	N/A
Rzm A 2.11	731.85	93 %	754.3750	754.3752

* **2.10** was not synthesized, detection only on low-res Shimadzu LCMS-2020 single quadrupole.

Selected NMR, mass spectra and UV traces

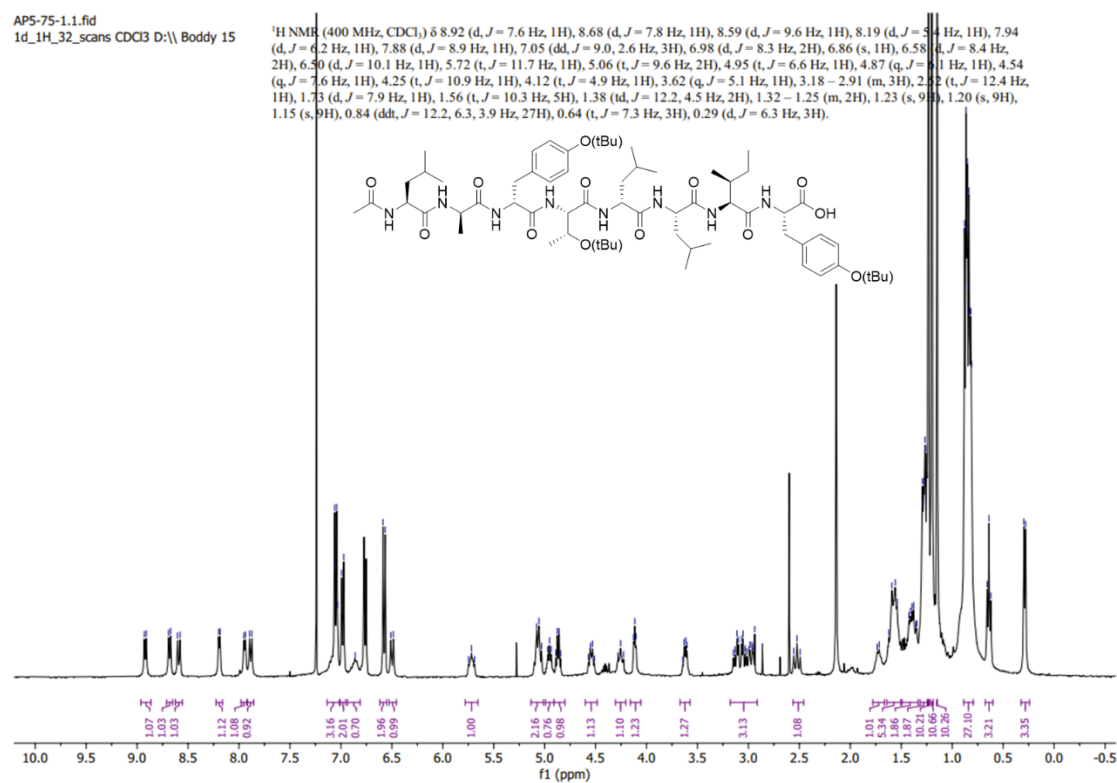


Figure S2.1. ¹H NMR spectrum (400 MHz) of compound 2.5 in CDCl₃.

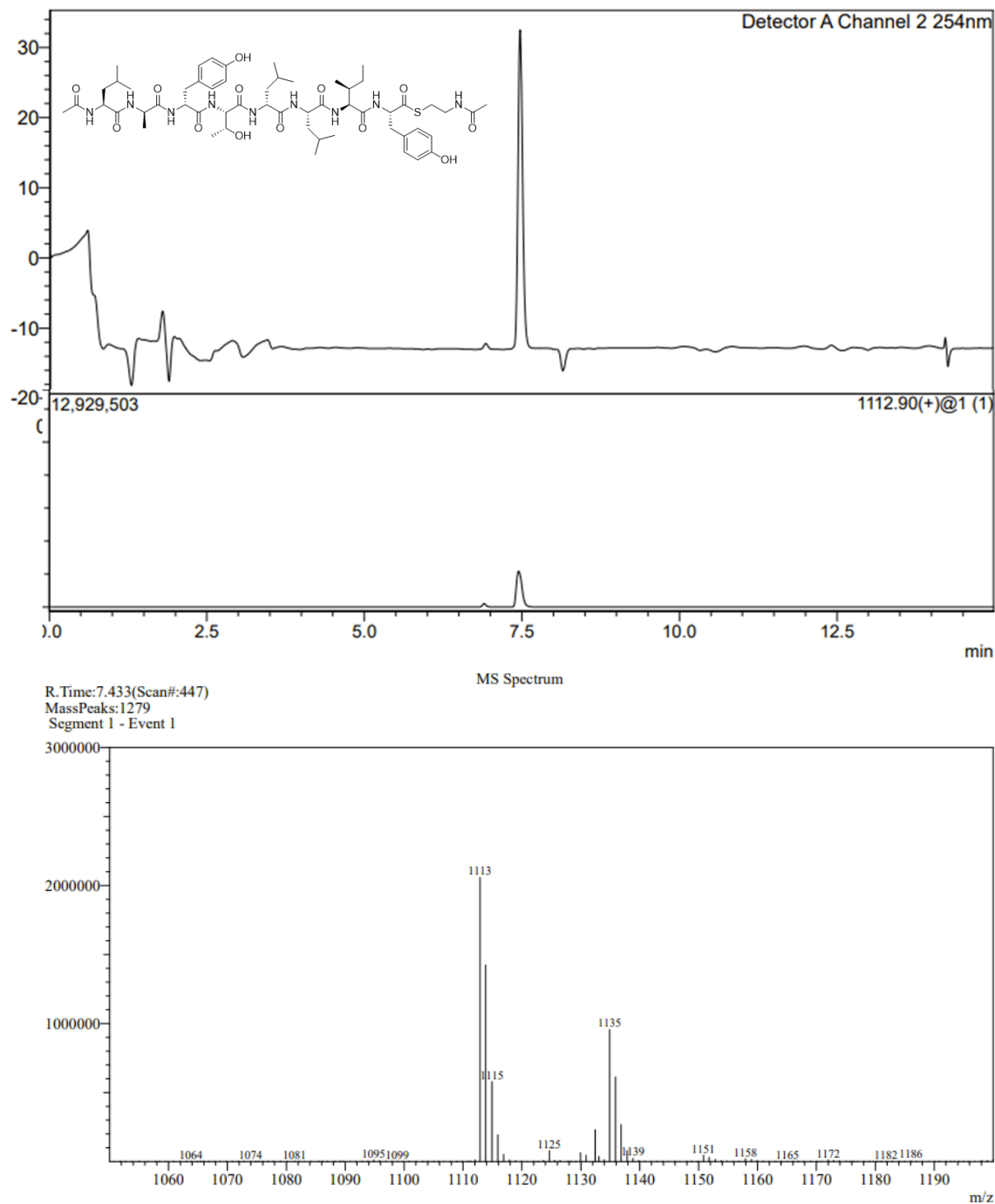


Figure S2.2. LCMS, UV traces and MS spectra of compound 2.1. UV trace was acquired at 254 nm and LCMS trace is an extracted ion chromatogram for $[M + H]^+$ ion with an $m/z = 1113$. Ion with $m/z = 1135$ corresponds to $[M + Na]^+$.

AP5-26-3.1003.fid
1H

¹H NMR (600 MHz, MeOD) δ 7.08 (d, J = 8.5 Hz, 2H), 7.02 (d, J = 8.5 Hz, 2H), 6.73 – 6.67 (m, 4H), 4.70 (dd, J = 9.2, 5.4 Hz, 1H), 4.47 (ddd, J = 17.1, 9.3, 5.8 Hz, 2H), 4.39 (dd, J = 9.7, 4.9 Hz, 1H), 4.34 (dd, J = 8.5, 6.6 Hz, 1H), 4.25 (t, J = 7.2 Hz, 3H), 4.19 (dd, J = 6.5, 4.3 Hz, 1H), 4.01 (d, J = 4.3 Hz, 1H), 3.10 (dd, J = 14.4, 5.5 Hz, 1H), 3.07 – 3.01 (m, 1H), 3.02 – 2.98 (m, 4H), 2.85 (dd, J = 14.2, 9.2 Hz, 2H), 2.04 (s, 4H), 1.91 (s, 3H), 1.74 – 1.47 (m, 13H), 1.33 – 1.22 (m, 9H), 1.17 – 1.09 (m, 2H), 1.01 – 0.83 (m, 27H).

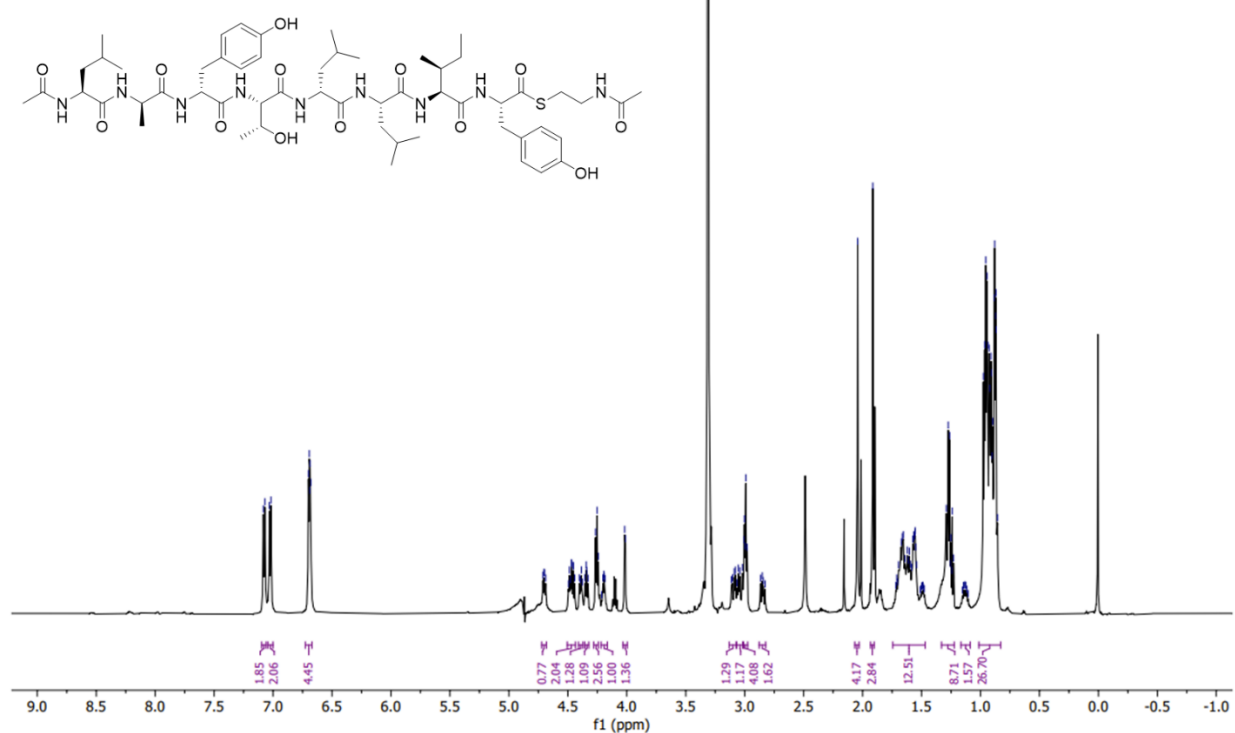


Figure S2.3. ¹H NMR spectrum (600 MHz) of compound 2.1 in MeOD-d₄.

AP5-136-1.10.fid
1d_1H DMSO (C:\data\guest\nmr) nmr 4

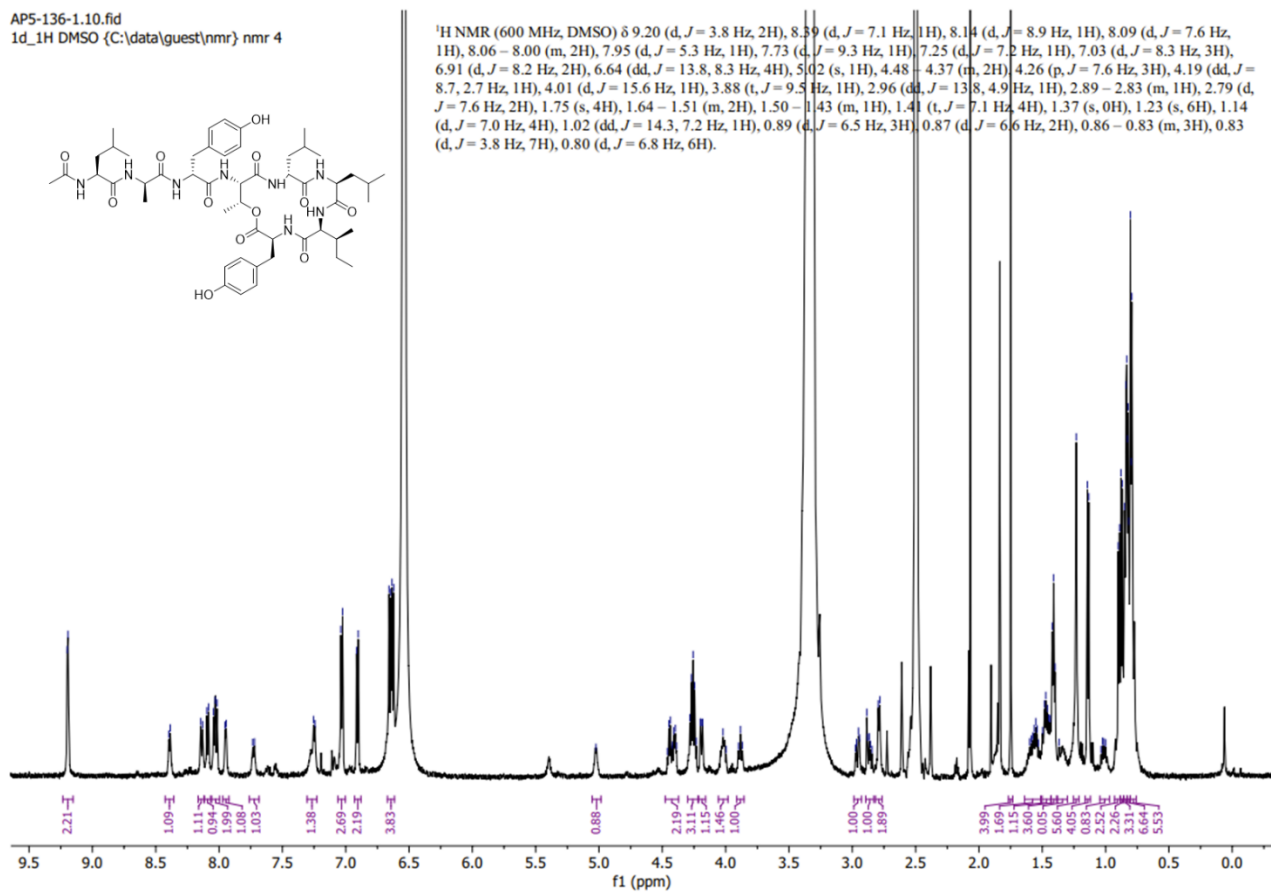
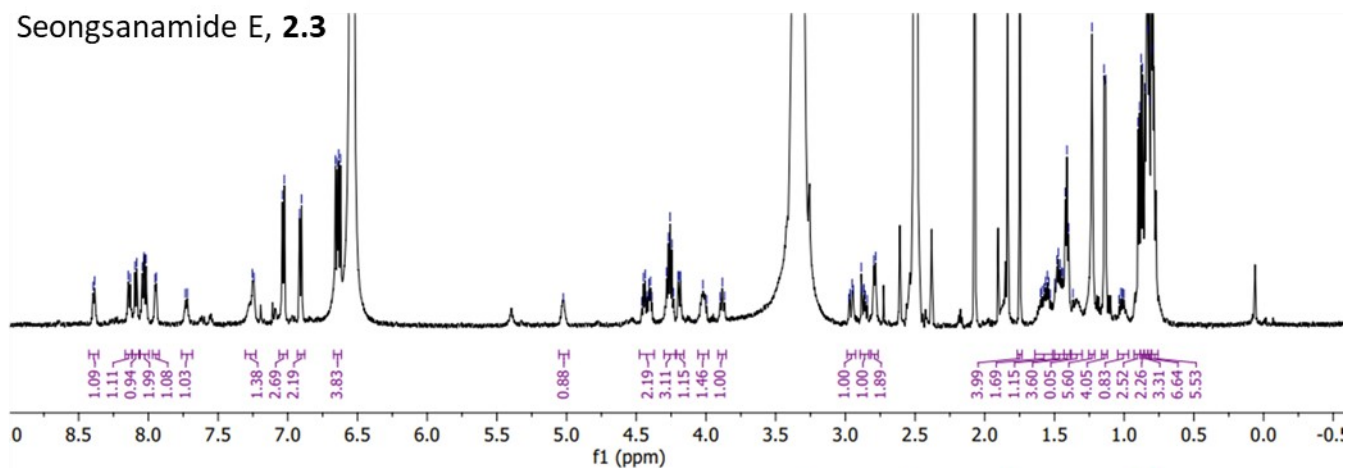


Figure S2.5. ¹H NMR spectrum (600 MHz) of compound 2.3 in DMSO-d₆.

Seongsanamide E, **2.3**



Isolated Sgd E; *Org. Lett.* 2018, 20, 23, 7539-7543

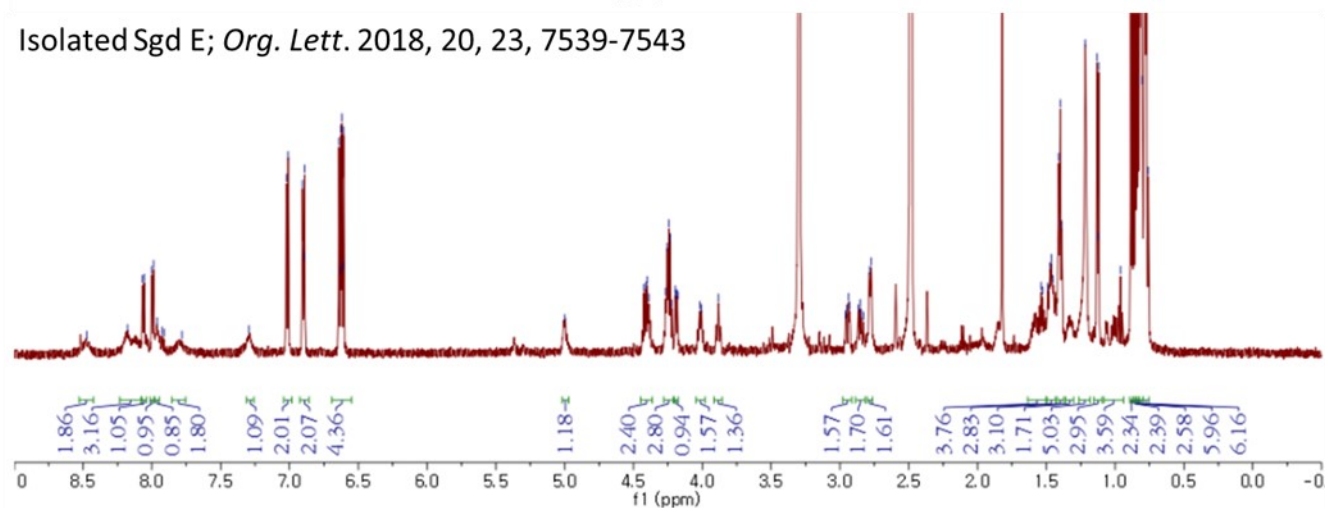
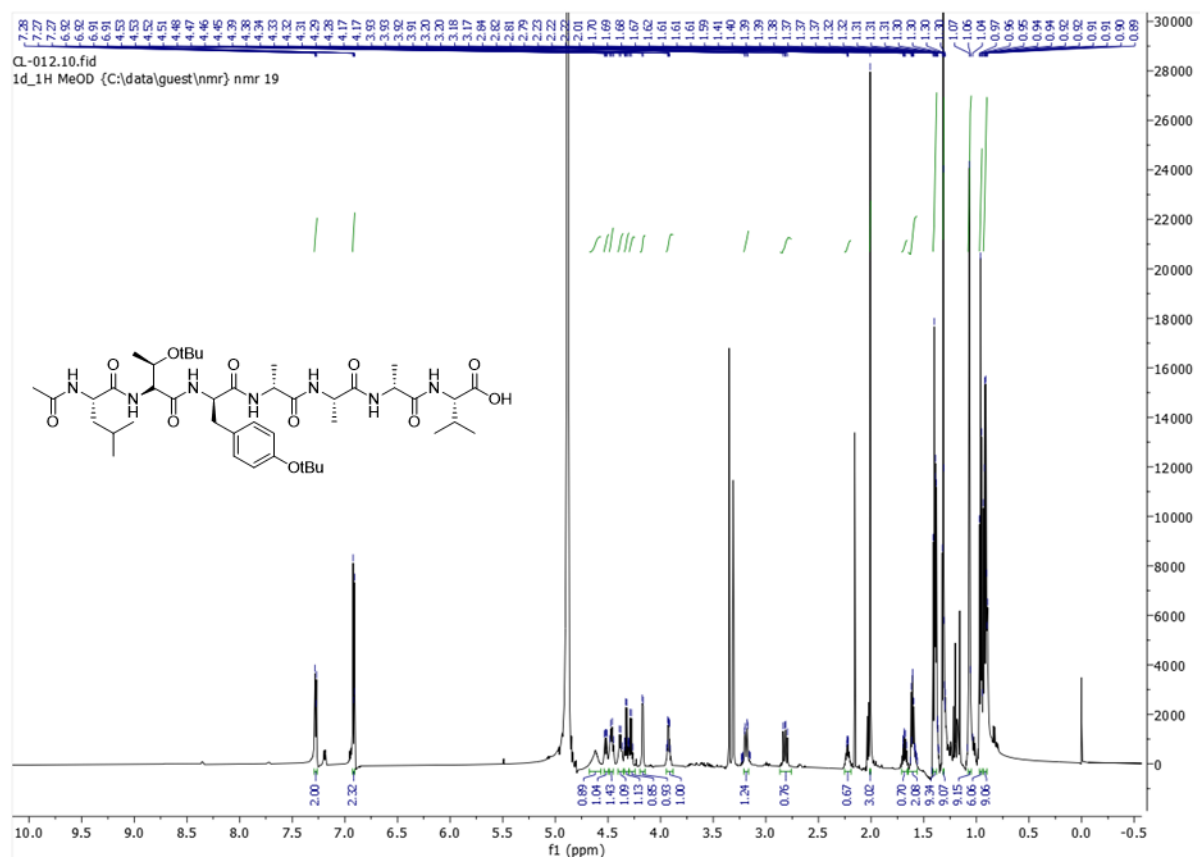


Figure S2.6. ¹H-NMR comparison of Sgd E (**2.3**) with isolated material.



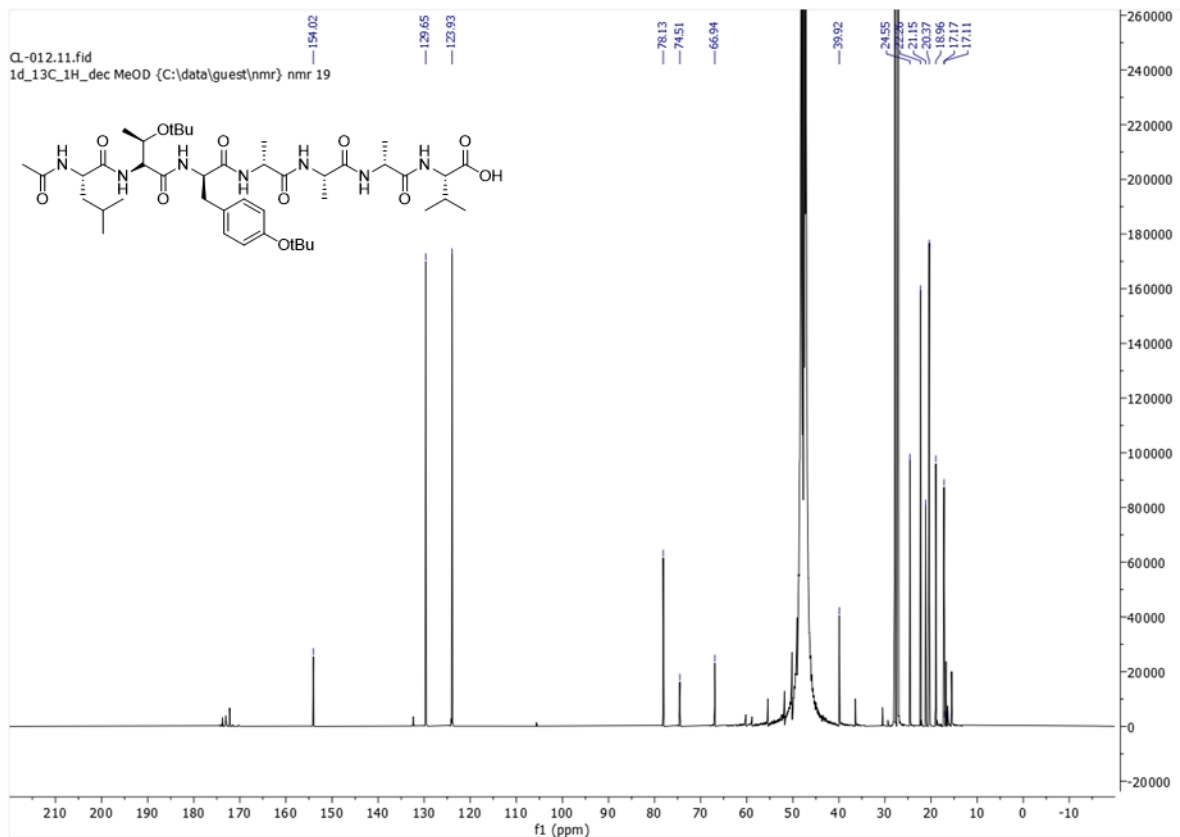


Figure S2.8. ^{13}C NMR spectrum (150 MHz) of compound 2.13 in MeOD.

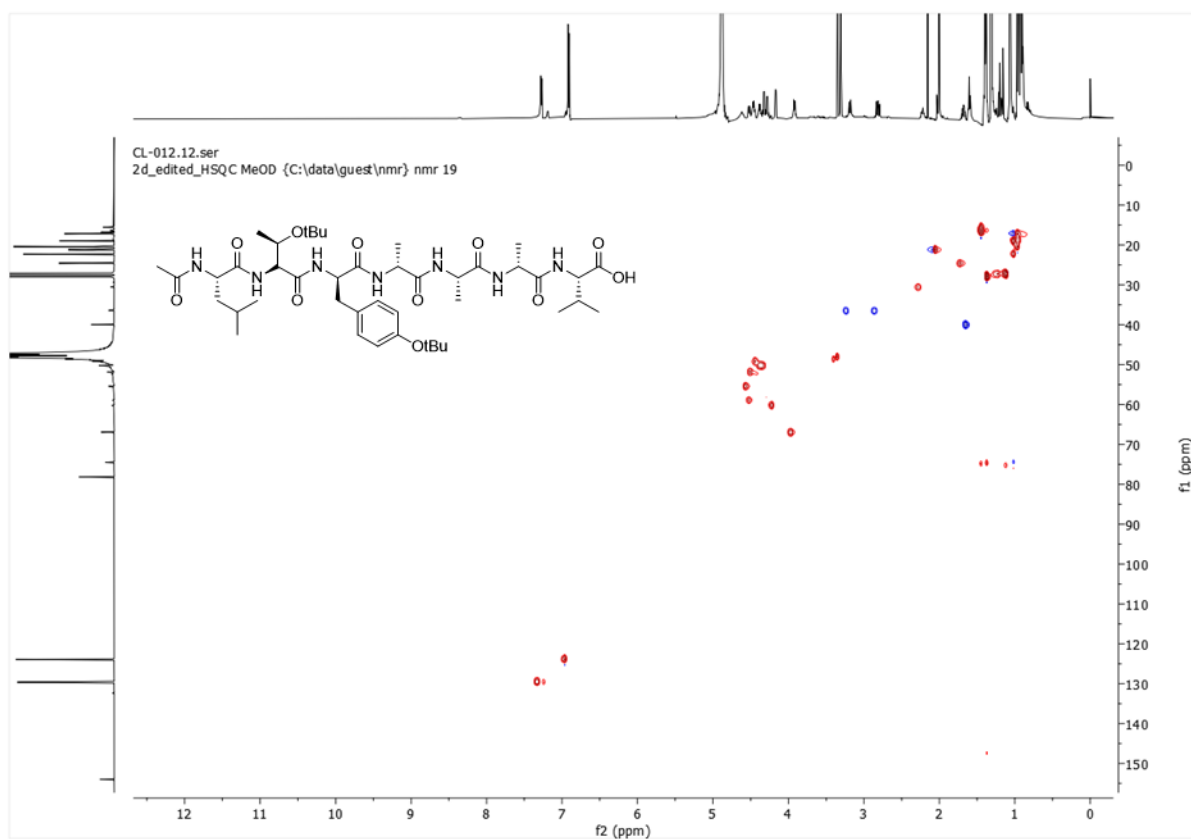


Figure S2.9. HSQC NMR spectrum of compound 2.13 in MeOD.

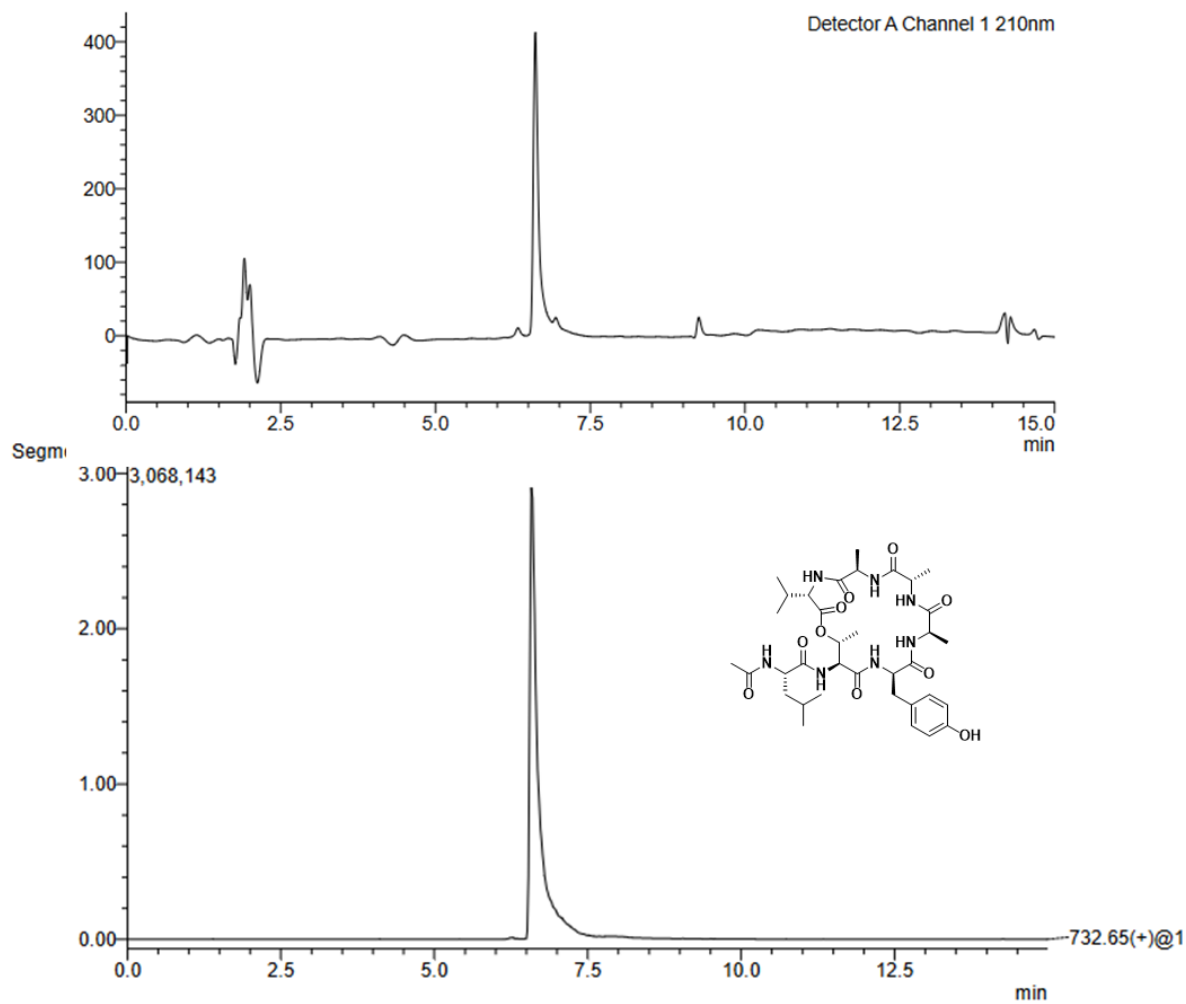


Figure S2.12. LCMS and UV traces of compound 2.11. UV trace was acquired at 210 nm and LCMS trace is an extracted ion chromatogram for $[M + H]^+$ ion with an $m/z = 733$.

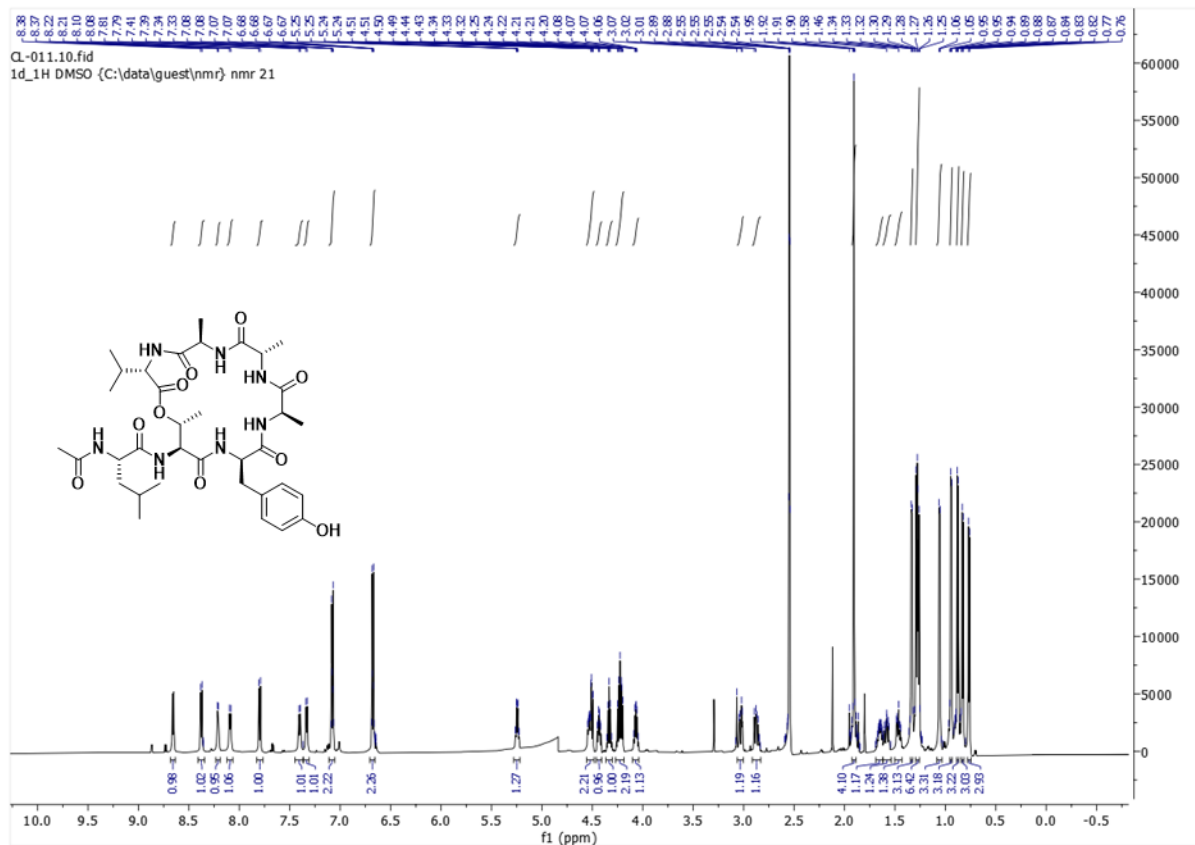


Figure S2.13. ^1H NMR spectrum (600 MHz) of compound 2.11 in DMSO-d_6 .

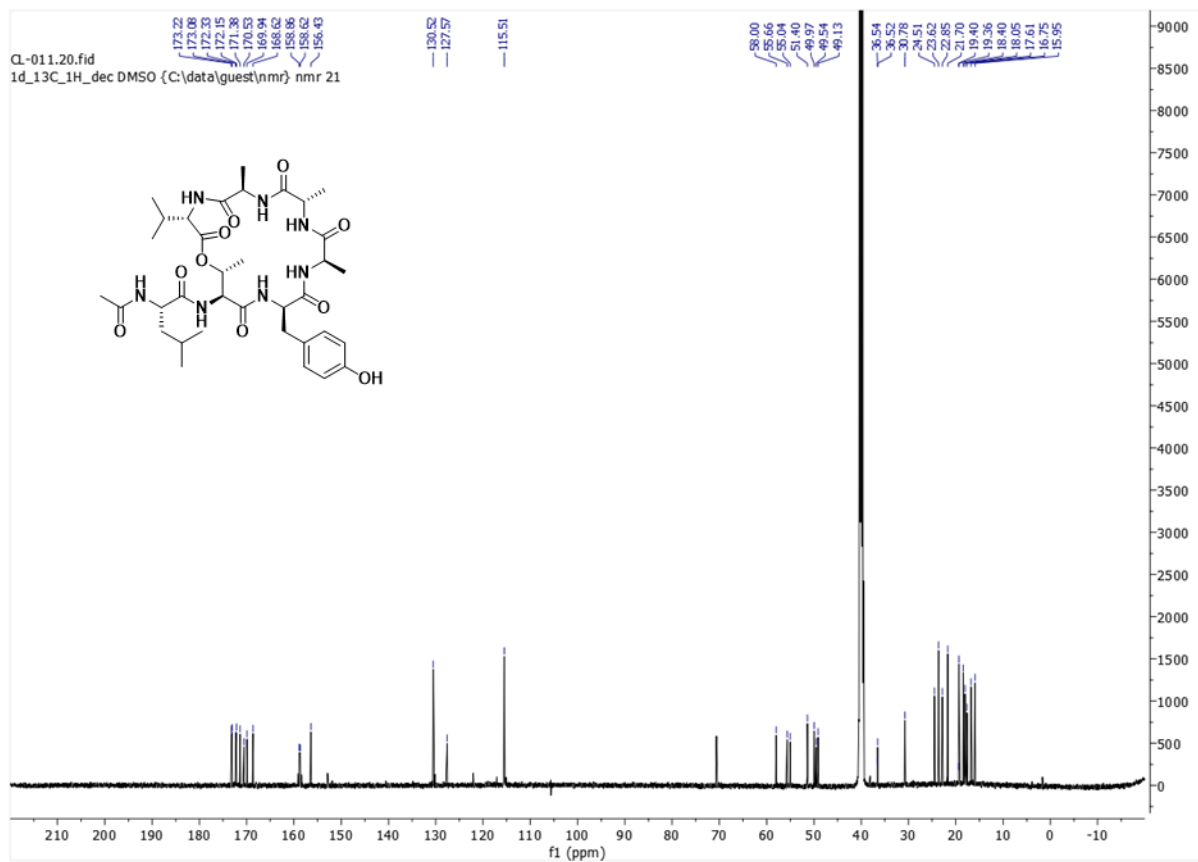


Figure S2.14. ^{13}C NMR spectrum (150 MHz) of compound 2.11 in DMSO-d_6 .

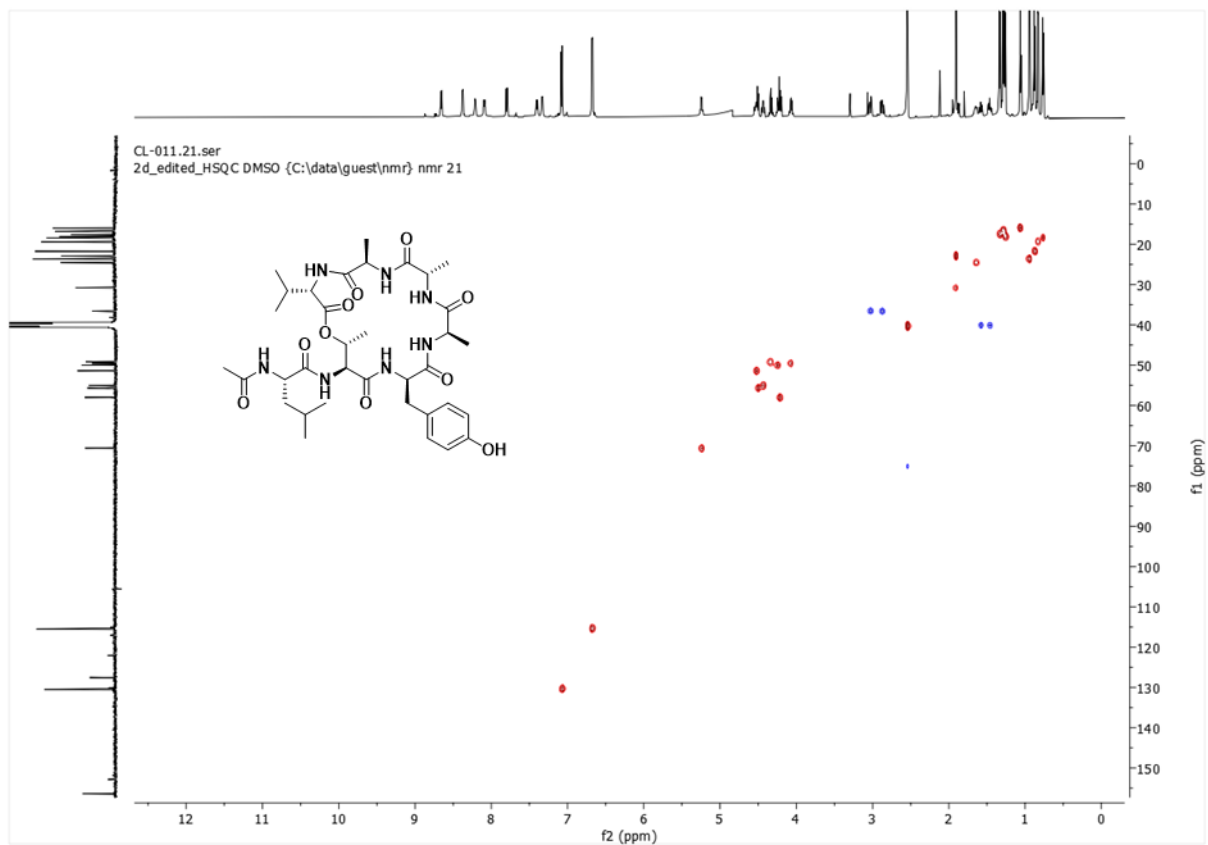


Figure S2.15. HSQC NMR spectrum of compound 2.11 in DMSO- d_6 .

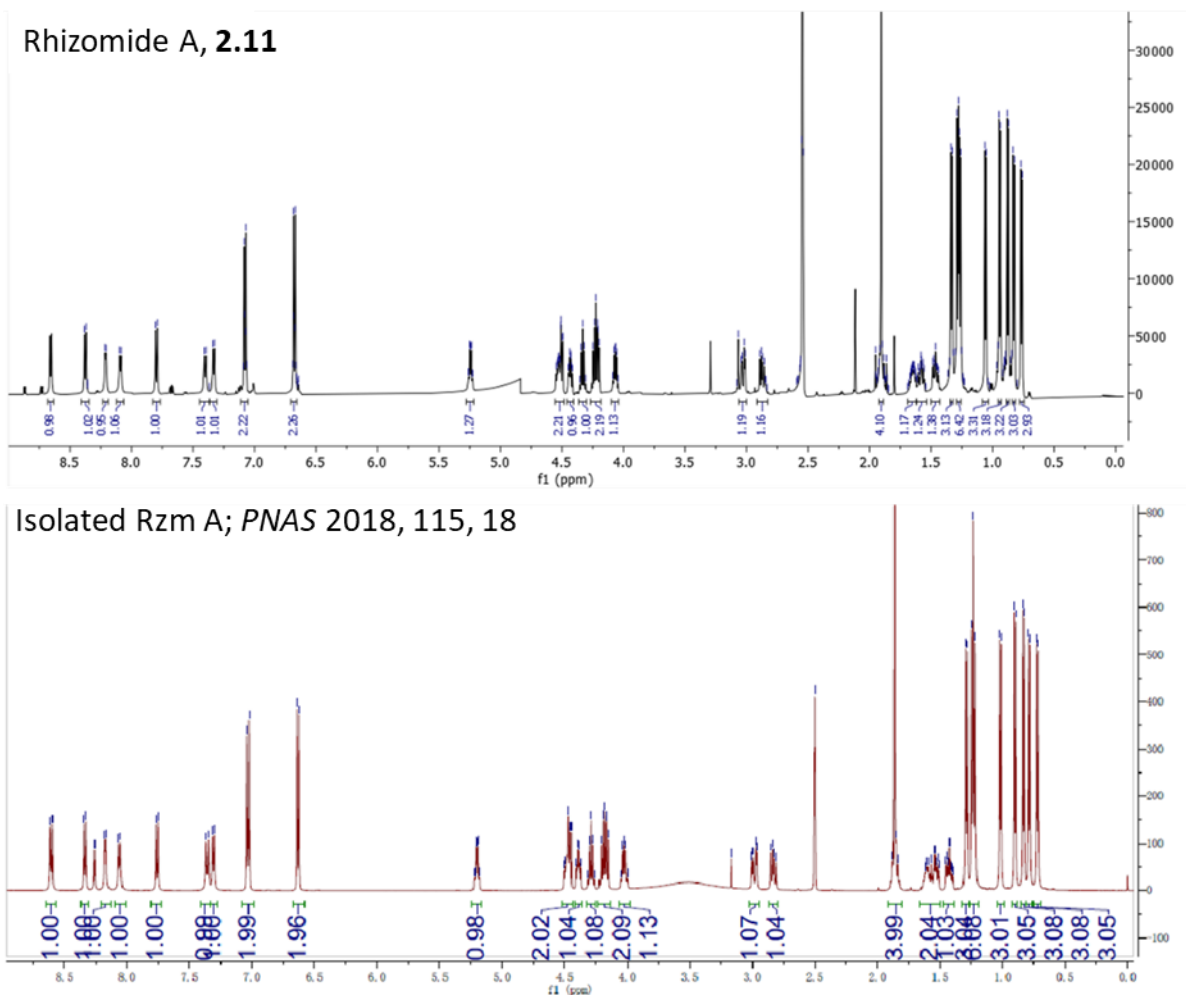


Figure S2.16. ^1H NMR comparison of Rzm A (**2.11**) with isolated material.

Table S2.3. ^1H (600 MHz) and ^{13}C NMR (150 MHz) data of Rzm A (2.11) in DMSO- d_6 .

	No.	δ_c (th)	δ_c (exp)	$\Delta\delta_c$	δ_H (th)	δ_H (exp)	$\Delta\delta_H$	Mult. (J in Hz)
HAc	1	170.0 C	170.53 C	0.53				
	2	22.3 CH ₃	22.57 CH ₃	0.27	1.86	1.90	0.04	s
Leu	1	172.7 C	173.22 C	0.48				
	2	50.9 CH	49.9 CH	1.00	4.48	4.53	0.05	d (2.2)
	3a	39.5 CH ₂	40.13 CH ₂	0.63	1.53	1.57	0.04	m
	3b				1.42	1.46	0.04	m
	4	24.0 CH	24.53 CH	0.53	1.61	1.63	0.02	m
	5	23.1 CH ₃	23.6 CH ₃	0.50	0.89	0.94	0.05	d (6.6)
	6	21.2 CH ₃	21.7 CH ₃	0.50	0.83	0.88	0.05	d (6.5)
	NH				8.33	8.38	0.05	d (8.1)
Thr	1	168.1 C	168.6 C	0.50				
	2	55.2 CH	55.69 CH	0.49	4.44	4.50	0.06	dd (2.4, 9.0)
	3	70.1 CH	70.63 CH	0.53	5.20	5.24	0.04	qd (2.4, 6.5)
	4	15.4 CH ₃	15.9 CH ₃	0.50	1.00	1.06	0.06	d (6.5)
	NH				7.76	7.80	0.04	d (8.9)
Tyr	1	170.8 C	171.38 C	0.58				
	2	54.6 CH	55.04 CH	0.44	4.38	4.43	0.05	td (4.8, 8.1)
	3a	36.0 CH ₂	36.54 CH ₂	0.54	2.98	3.02	0.04	dd (4.7, 14.3)
	3b				2.83	2.87	0.04	dd (8.9, 14.3)
	4	127.0 C	127.57 C	0.57				
	5/9	130.0 CH	130.52 CH	0.52	7.02	7.06	0.04	d (6.8)
	6/8	115.0 CH	115.56 CH	0.56	6.63	6.67	0.04	d (6.6)
	7	155.9 C	156.43 C	0.53				
	NH				7.30	7.33	0.03	d (7.8)
	OH				8.26			
Ala 1	1	172.5 C	173.0 C	0.5				
	2	49.0 CH	49.54 CH	0.54	4.02	4.07	0.05	dd (5.2, 7.0)
	3	17.1 CH ₃	17.61 CH ₃	0.51	1.29	1.33	0.04	d (6.9)
	NH				8.17	8.21	0.08	d (7.7)
Ala 2	1	171.8 C	172.3 C	0.5				
	2	48.6 CH	49.13 CH	0.53	4.29	4.33	0.04	p (7.2)
	3	16.2 CH ₃	16.75 CH ₃	0.55	1.22	1.27	0.05	dd (7.2, 11.6)
	NH				8.60	8.66	0.06	d (7.2)
Ala 3	1	171.6 C	172.15 C	0.55				
	2	49.4 CH	49.97 CH	0.57	4.17	4.24	0.07	m
	3	17.5 CH ₃	17.61 CH ₃	0.11	1.23	1.33	0.10	d (6.9)
	NH				7.35	7.33	0.02	d (7.8)
Val	1	169.4 C	169.9 C	0.5				
	2	57.5 CH	58 CH	0.5	4.19	4.21	0.02	m
	3	30.3 CH	30.78 CH	0.48	1.85	1.87	0.02	m
	4	18.8 CH ₃	19.36 CH ₃	0.56	0.78	0.83	0.05	d (7.6)
	5	17.9 CH ₃	18.4 CH ₃	0.5	0.72	0.77	0.05	d (7.6)
	NH				8.05	8.09 7.80	0.04	d (5.1) d (8.9)

Protein purities, yields, sequences and Rzm-TE predicted structure

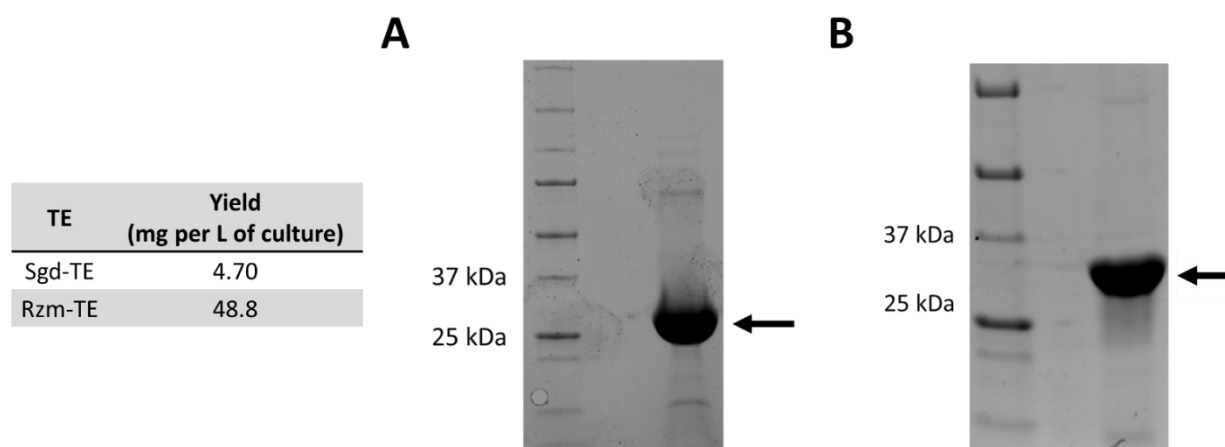


Figure S2.17. SDS-PAGE analysis of purified recombinant Sgd-TE (A) and Rzm-TE (B). Expected molecular weights for Sgd- and Rzm-TE are approximately 33 and 32 kDa respectively. In both panels, left lanes correspond to Precision Plus Protein Unstained Protein Standards (#1610363, Bio-Rad), right lanes correspond to the purified samples (loaded ~33 µg for each). Protein yields are shown in the table to the left.

>Sgd-TE
 MGSS**HHHHHH**SSGLVPRGSHMKKKDKTEKNLELI IKLRES DGAPFFFIHPASGNTLCYME LAKLLDNACS
 LYGINFPYVEDRSMENIRLEDLSRLYLKKIKEIQPKGPYMLGGW**S**IGGLFAYEIANQLIEQGEEVSLLSL
 ID**T**AFPRAEYILSKEESI IQVLNSISAHLILDPNKEKEIAEVT FKNLINS PKIRQMLPPGLEVEDIKTLS
 NIFHANTKVS AEYKIEKQYRGLIYFEAQEGTRLGNKWSSYLEGDLKNIQVPGK**H**AEMMDSPSVNIIANKL
 KEEILSATYI-

>Rzm-TE
 MGSS**HHHHHH**SSGLVPRGSHMLARHLVKRDGIQENSFSVLLPLKPTGSRPALF**C**IHPGFGLSWSYMG LSS
 YLHPDQPLYGLQARGLDGTSPLASTLDEMVS DYLEHMRRVQPKGPY**C**LLGW**S**FGGNVAHSM AVRLEHQGE
 KVAL LALL**D**STPLVGSNAFAEKENQLDQAMVRDVFAYHYSNELVSAMDEQLLENTEEIAKNNL**C**IVRDYS
 PSQYGGNALLFRATVAEAG**C**ETLVSPDAWNPYVFGGIEVHDVH**C**SHGEMLKPEPTAIIGSTLAC**C**KLDKWE
 SQQAQKVNT EELK-

Figure S2.18. Sgd- and Rzm-TE translated protein sequences. Shown in bold is the N-terminal 6xHis-tag used for purification. Predicted catalytic triad for both TEs are highlighted in yellow. For the Rzm-TE sequence, all Cys residues are highlighted in magenta (6 total), with solvent accessible Cys residues underlined.

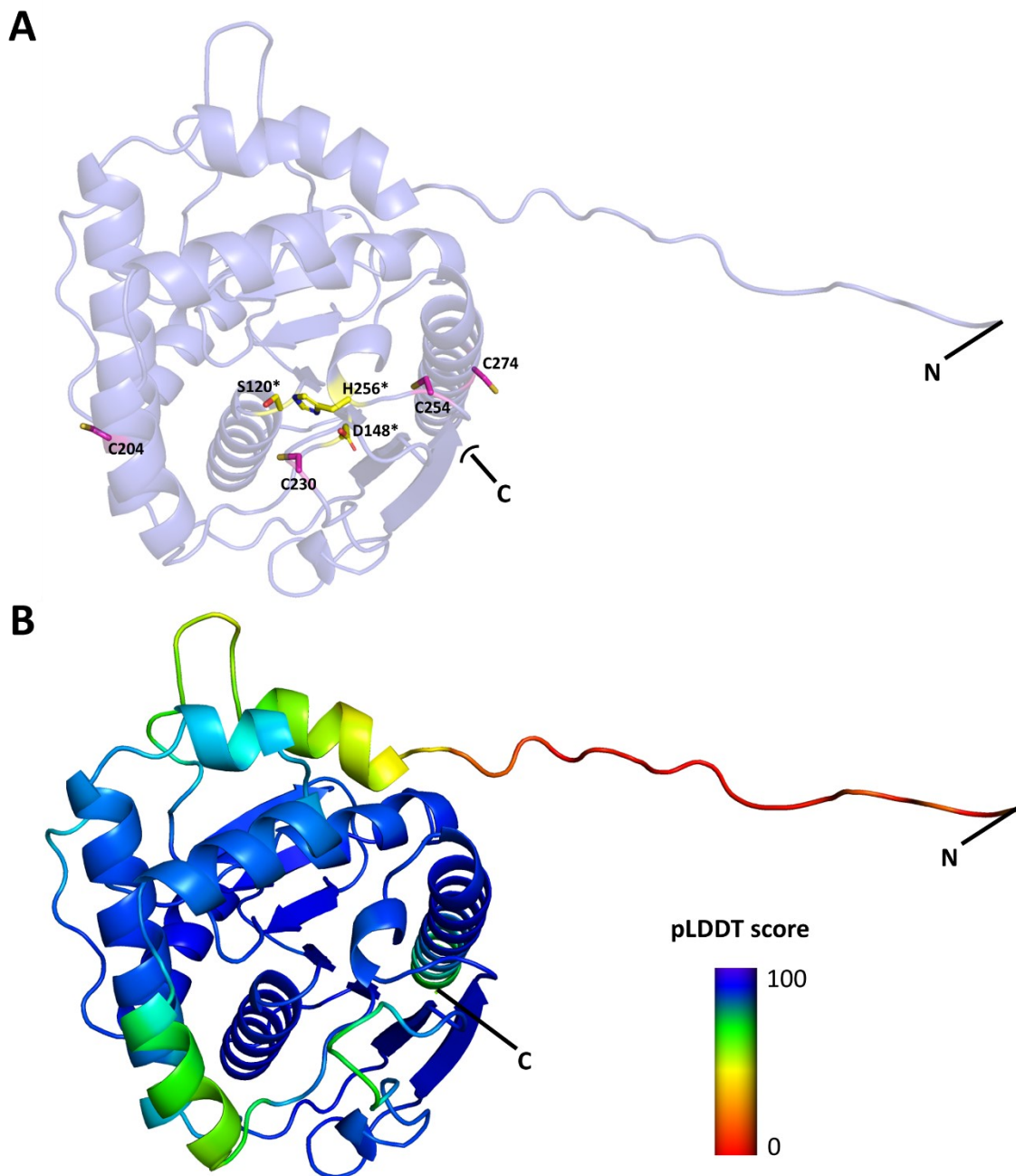


Figure S2.19. Rzm-TE AlphaFold2 predicted structure. Rzm-TE is predicted to adopt an α/β -hydrolase fold commonly adopted by TE domains. Prediction was generated from the AlphaFold2 (v2.3.2) Google Collab webserver using standard settings.⁸⁰ (A) Structure is shown as a purple cartoon with the predicted catalytic triad shown as yellow sticks. Only the four solvent accessible Cys residues are shown as magenta sticks. (B) Predicted structure with residues coloured according to AlphaFold2's per-residue confidence metric pLDDT.

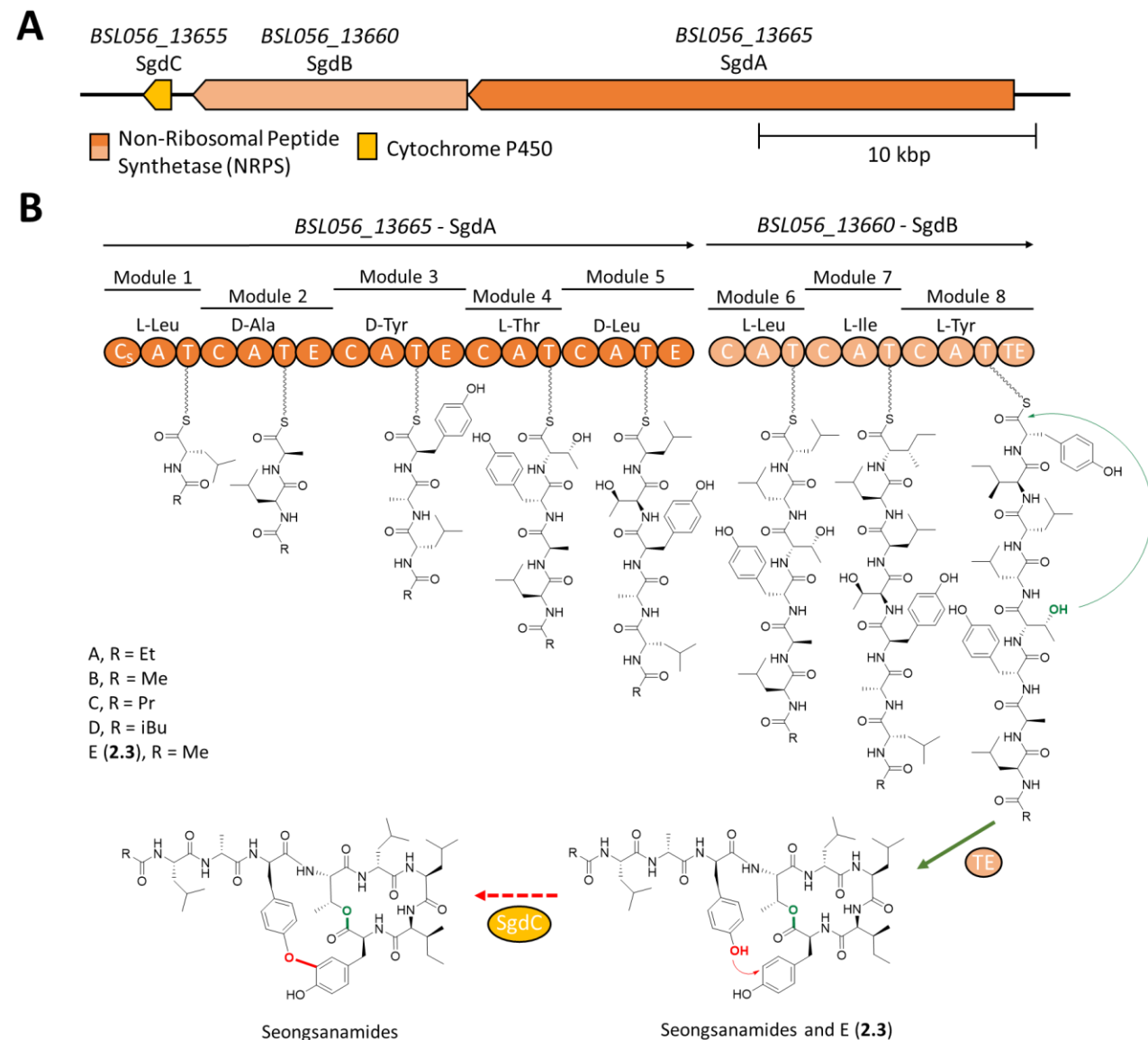


Figure S2.20. Biosynthesis of seongsanamides A-E. (A) The BL056_13665 (*sgdA*), BSL06_13660 (*sgdB*) and BSL056_13655 (*sgdC*) genes involved in Sgd A-E biosynthesis. The ends of the gene cluster contain ~2 kb regions with a low GC content (21%) within a 200 bp sliding window. No resistance, transport and tailoring genes have been identified beyond the cluster. (B) Biosynthesis of seongsanamide (Sgd) analogues A-E and proposed formation of Sgd A-D biaryl-ether analogues. Sgd E (**2.3**) has only been observed as a monocyclic macrocycle. Abbreviations; Cs, starter condensation domain; C, condensation domain; A, adenylation domain; T, thiolation or peptidyl carrier domain; E, epimerization domain; TE, thioesterase domain.

Standard curve for Sgd-TE kinetics

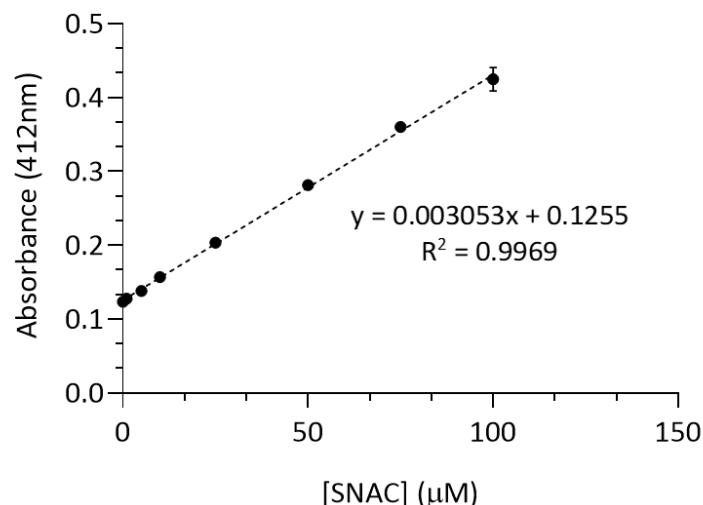


Figure S2.22. SNAC standard curve. Single absorbance measurements at 412 nm were taken in a half-area 96-well plate (mean $\pm$ s.d., $n = 3$ technical replicates). Mixtures containing varying concentrations of *N*-acetylcysteamine thiolate (SNAC, 0 to 100 μ M) were prepared in triplicate in 50 mM of phosphate buffer (pH 7.4) and 7.5% v/v DMSO with 5 μ M Sgd-TE. 4% v/v DTNB reacts with SNAC thiolate forming a yellow coloured TNB anion that can be quantified by spectrophotometry.^{34,53,54,56} Absorbance readings were plotted against SNAC concentration to yield the above standard curve.

Sgd-TE product ratios and Rzm-TE relative conversion

Table S2.4. Calculated product ratios of Sgd-TE assays. Product ratios were calculated from areas under the curve (AUCs) by dividing the respective AUCs as shown below. AUC values were obtained from extracted ion chromatograms for $[M + H]^+$ ions with $m/z = 994$ (**2.3**), 1012 (**2.2**) and 1113 (**2.1**).

Time (hours)	Substrate 2.1 assays		Ring opening assays 2.3
	2.2 / 2.1	2.3 / 2.2	2.2 / 2.3
0	0.16193	N/A	N/A
0.5	0.18765	1.76690	0.02805
1.5	0.19552	3.07404	0.09162
3	0.24973	3.43211	0.16143
20	0.60648	2.45733	1.53637

Table S2.5. Relative product conversion of Rzm-TE and substrate 2.11 only assays. Product conversion (%) was calculated from areas under the curve (AUCs) by dividing the AUC of **2.11** by the total area. AUC values were obtained from extracted ion chromatograms for $[M + H]^+$ ions with $m/z = 733$ (**2.11**) and 852 (**2.9**).

Time (hours)	2.11 only control				Rzm-TE + 2.11 assays			
	2.11	2.9	Total	Conversion (%)	2.11	2.9	Total	Conversion (%)
0	260214	28851767	29111981	0.89	187972	9671989	9859961	1.91
3	168659	19460353	19629012	0.86	457589	9230897	9688486	4.72
24	142885	33989916	34132801	0.42	1813820	10073585	11887405	15.26

Sgd-TE assay chromatograms and mass spectra

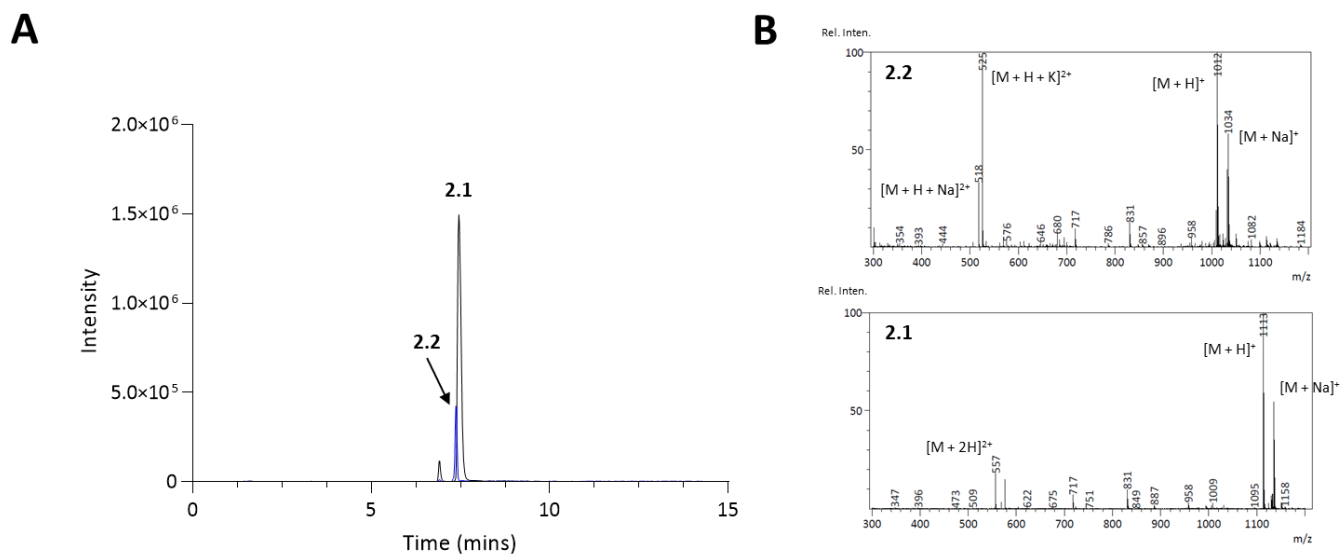


Figure S2.23. Sgd-TE 0-hour incubation with substrate 2.1. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 0 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 1012$ (**2.2**) and 1113 (**2.1**). (B) Mass spectra for substrate **2.1** and hydrolysis product **2.2**, mass adducts are identified in each spectra. Thermally denatured and substrate only controls all showed similar results for all time points.

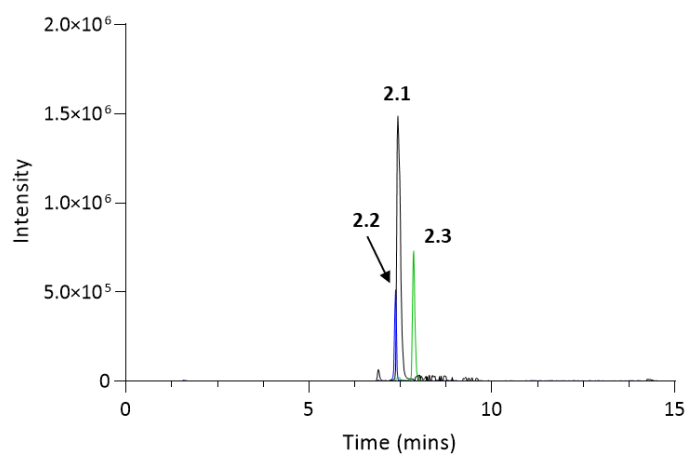
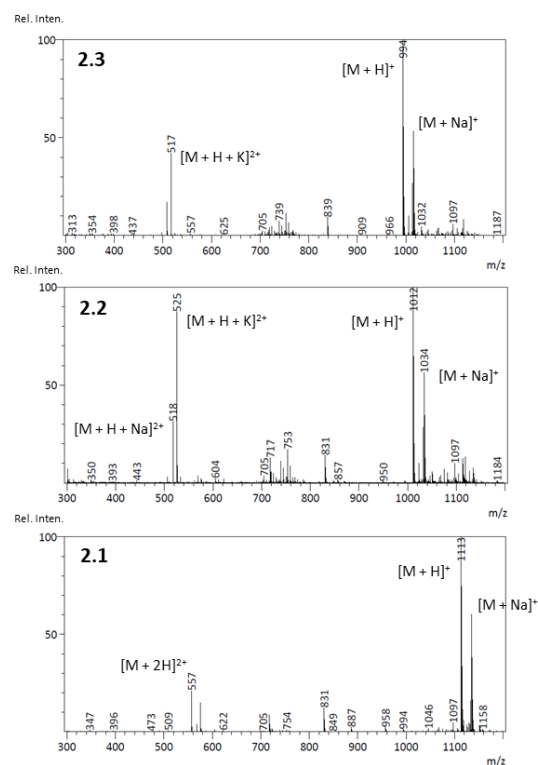
A**B**

Figure S2.24. Sgd-TE 0.5-hour incubation with substrate 2.1. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 0.5 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with m/z = 994 (**2.3**), 1012 (**2.2**) and 1113 (**2.1**). (B) Mass spectra for substrate **2.1**, hydrolysis product **2.2** and macrocyclic product **2.3**, mass adducts are identified in each spectra.

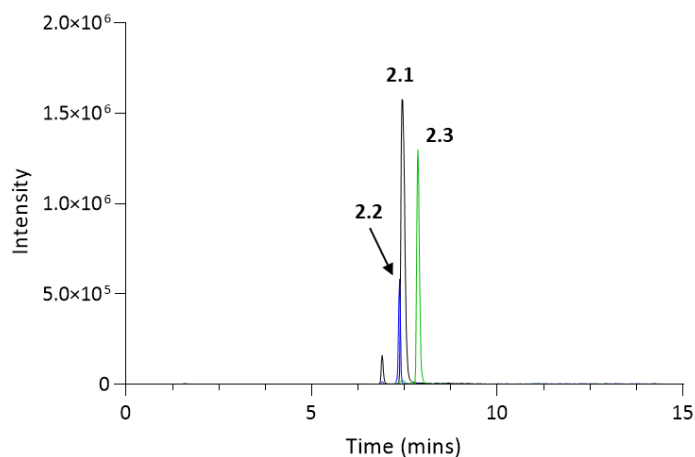
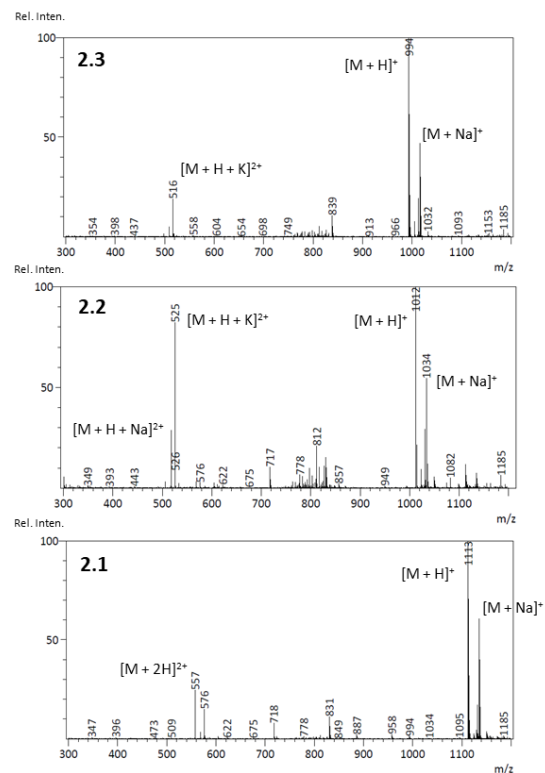
A**B**

Figure S2.25. Sgd-TE 1.5-hour incubation with substrate 2.1. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 1.5 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**), 1012 (**2.2**) and 1113 (**2.1**). (B) Mass spectra for substrate **2.1**, hydrolysis product **2.2** and macrocyclic product **2.3**, mass adducts are identified in each spectra.

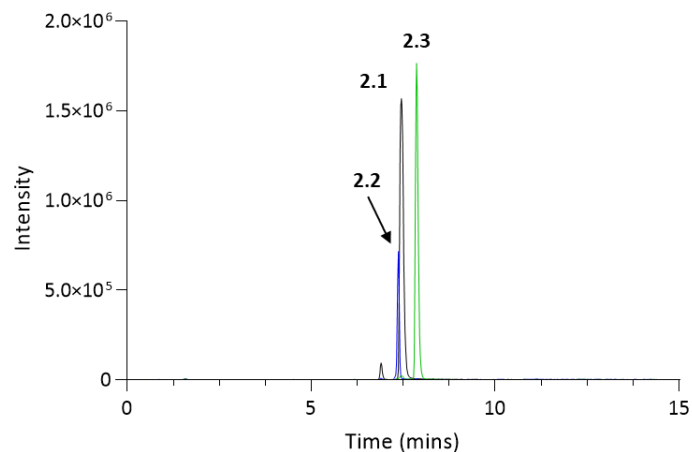
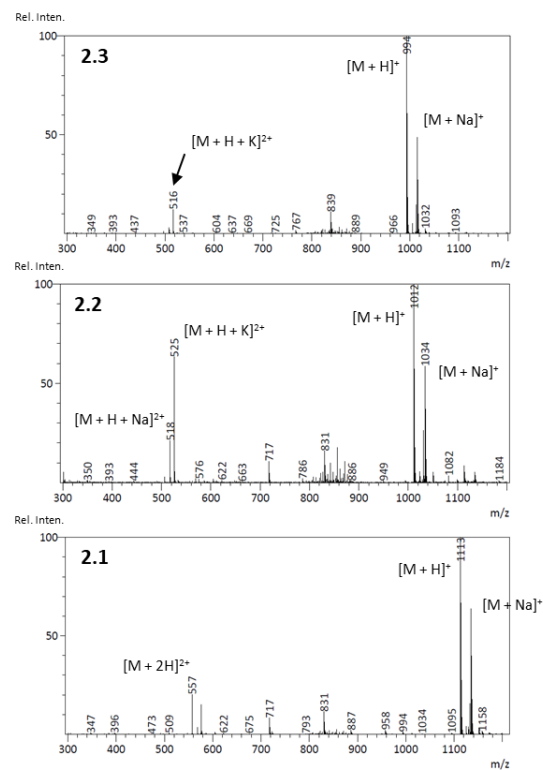
A**B**

Figure S2.26. Sgd-TE 3-hour incubation with substrate 2.1. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 3 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**), 1012 (**2.2**) and 1113 (**2.1**). (B) Mass spectra for substrate **2.1**, hydrolysis product **2.2** and macrocyclic product **2.3**, mass adducts are identified in each spectra.

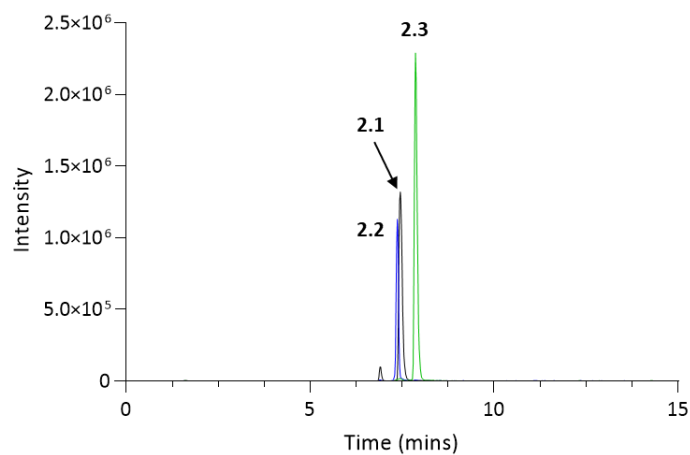
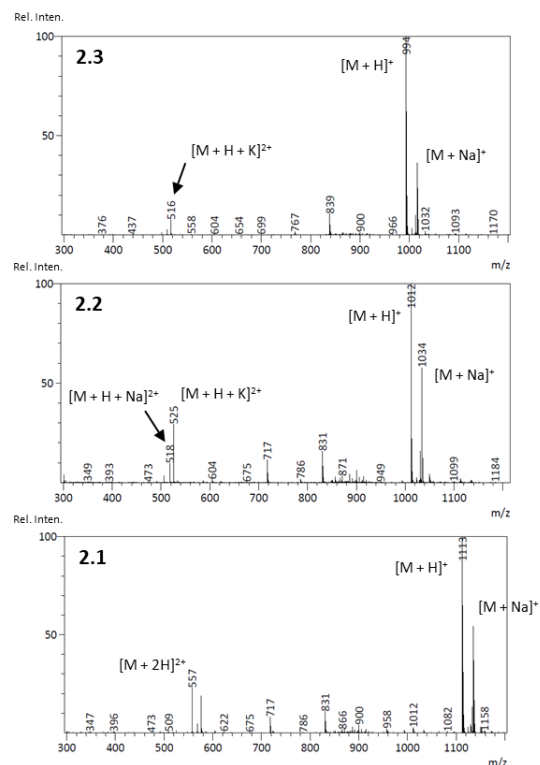
A**B**

Figure S2.27. Sgd-TE 20-hour incubation with substrate 2.1. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 20 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with m/z = 994 (**2.3**), 1012 (**2.2**) and 1113 (**2.1**). (B) Mass spectra for substrate **2.1**, hydrolysis product **2.2** and macrocyclic product **2.3**, mass adducts are identified in each spectra.

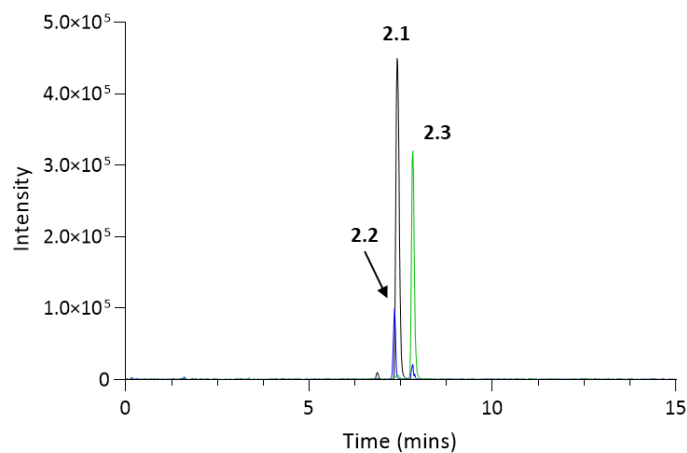
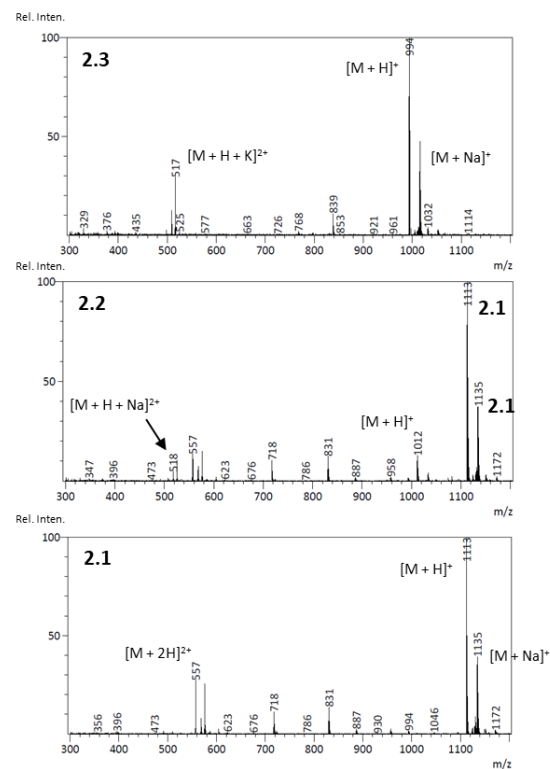
A**B**

Figure S2.28. Sgd-TE 3-hour assay co-injection with macrocycle 2.3. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 3 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4), co-injected with macrocycle **2.3**. Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**), 1012 (**2.2**) and 1113 (**2.1**). (B) Mass spectra for substrate **2.1**, hydrolysis product **2.2** and macrocyclic product **2.3**, mass adducts are identified in each spectra.

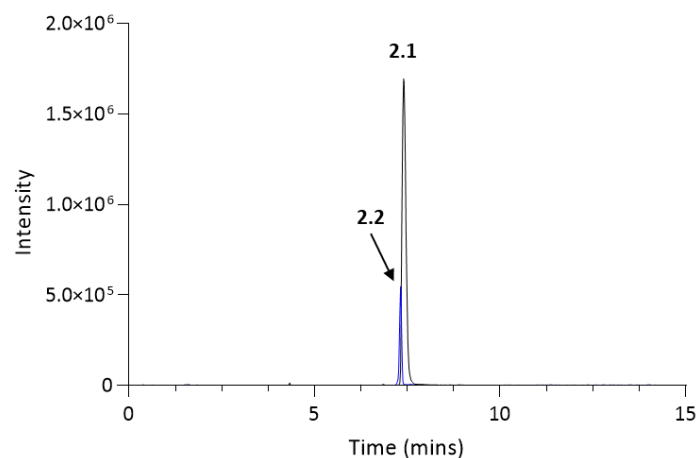
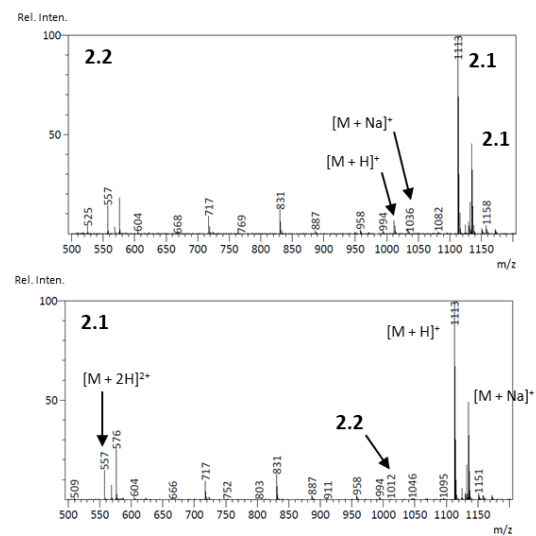
A**B**

Figure S2.29. DEBS-TE 20-hour incubation with substrate 2.1. (A) Extracted ion chromatogram for DEBS-TE (5 μ M) assay after 20 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 1012$ (**2.2**) and 1113 (**2.1**). (B) Mass spectra for substrate **2.1** and hydrolysis product **2.2**, mass adducts are identified in each spectra. Thermally denatured and substrate only controls all showed similar results.

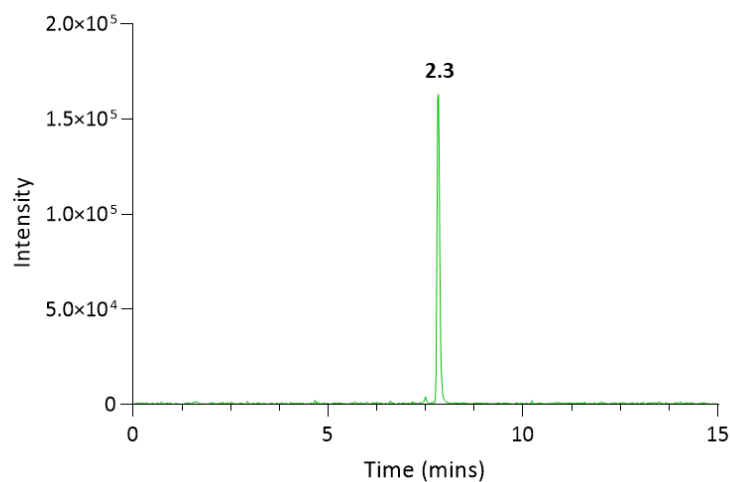
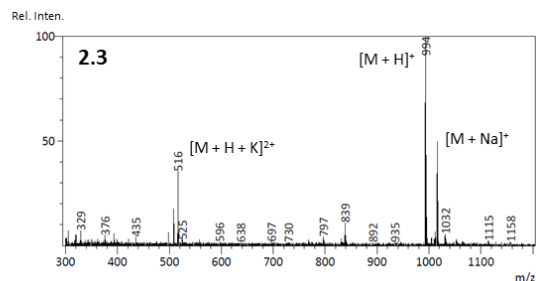
A**B**

Figure S2.30. Sgd-TE 0-hour incubation with macrocycle 2.3. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 0 hours of incubation with 0.5 mM of **2.3** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**). (B) Mass spectra for macrocycle authentic standard **2.3**, mass adducts are identified. Thermally denatured and macrocycle only controls all showed similar results for all time points.

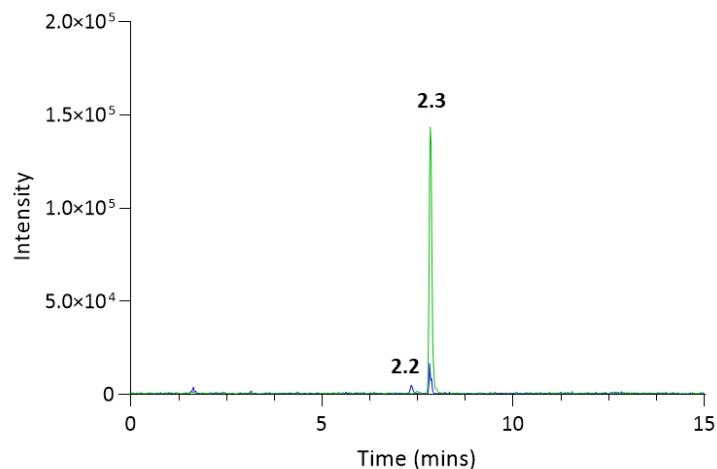
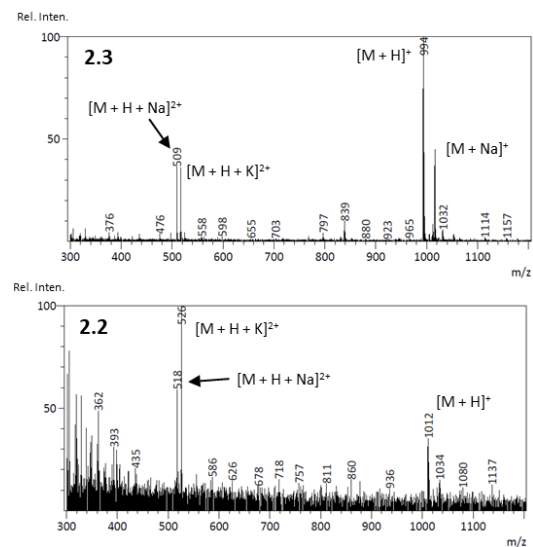
A**B**

Figure S2.31. Sgd-TE 0.5-hour incubation with macrocycle 2.3. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 0.5 hours of incubation with 0.5 mM of **2.3** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**) and 1012 (**2.2**). (B) Mass spectra for macrocycle authentic standard **2.3** and hydrolysis product **2.2**, mass adducts are identified in each spectra.

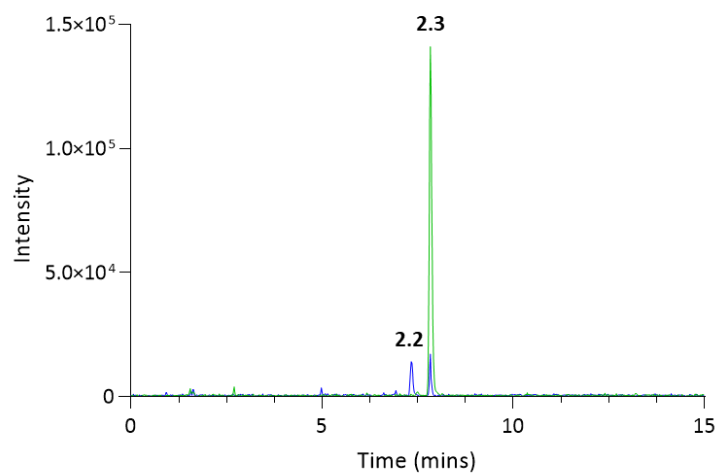
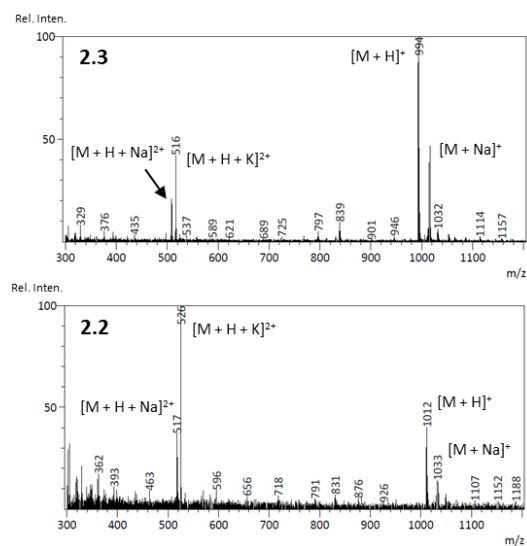
A**B**

Figure S2.32. Sgd-TE 1.5-hour incubation with macrocycle 2.3. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 1.5 hours of incubation with 0.5 mM of **2.3** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**) and 1012 (**2.2**). (B) Mass spectra for macrocycle authentic standard **2.3** and hydrolysis product **2.2**, mass adducts are identified in each spectra.

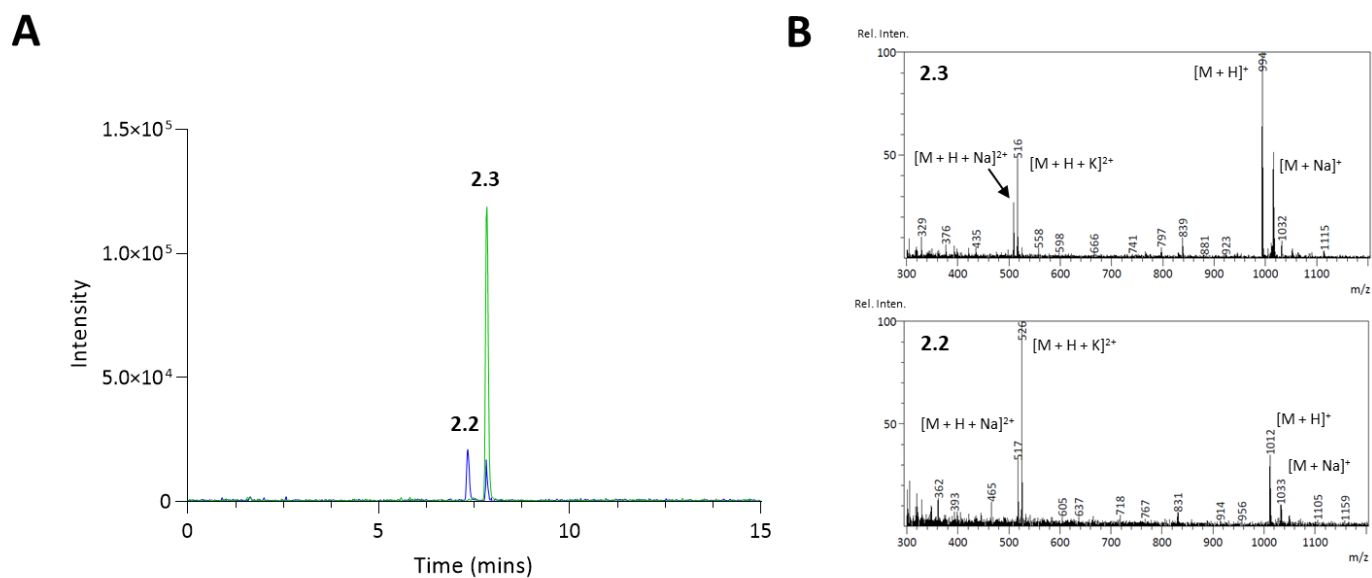


Figure S2.33. Sgd-TE 3-hour incubation with macrocycle 2.3. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 3 hours of incubation with 0.5 mM of **2.3** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**) and 1012 (**2.2**). (B) Mass spectra for macrocycle authentic standard **2.3** and hydrolysis product **2.2**, mass adducts are identified in each spectra.

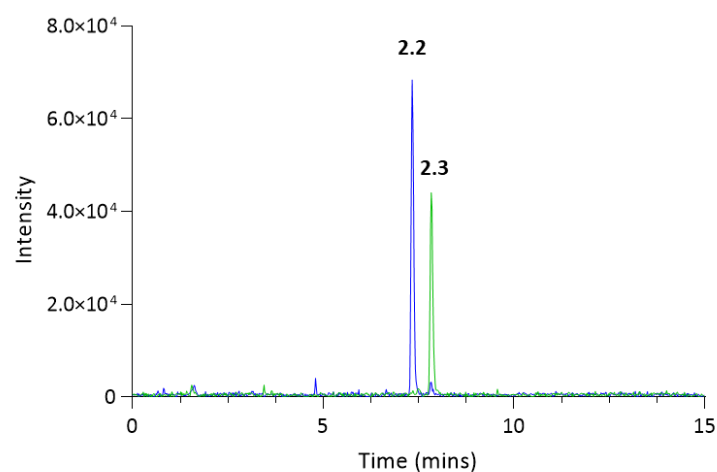
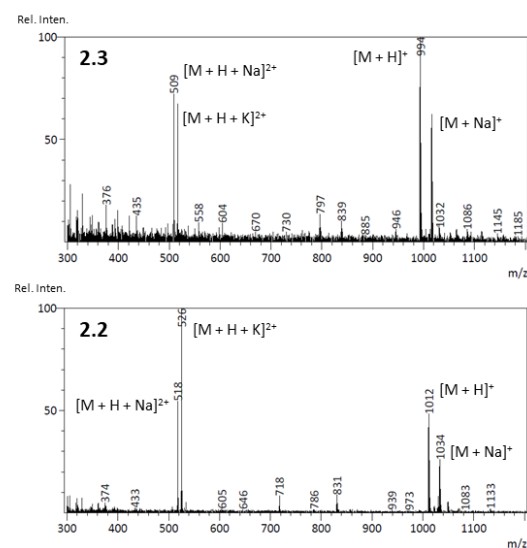
A**B**

Figure S2.34. Sgd-TE 20-hour incubation with macrocycle 2.3. (A) Extracted ion chromatogram for Sgd-TE (5 μ M) assay after 20 hours of incubation with 0.5 mM of **2.3** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 994$ (**2.3**) and 1012 (**2.2**). (B) Mass spectra for macrocycle authentic standard **2.3** and hydrolysis product **2.2**, mass adducts are identified in each spectra.

Rzm-TE assay chromatograms and mass spectra

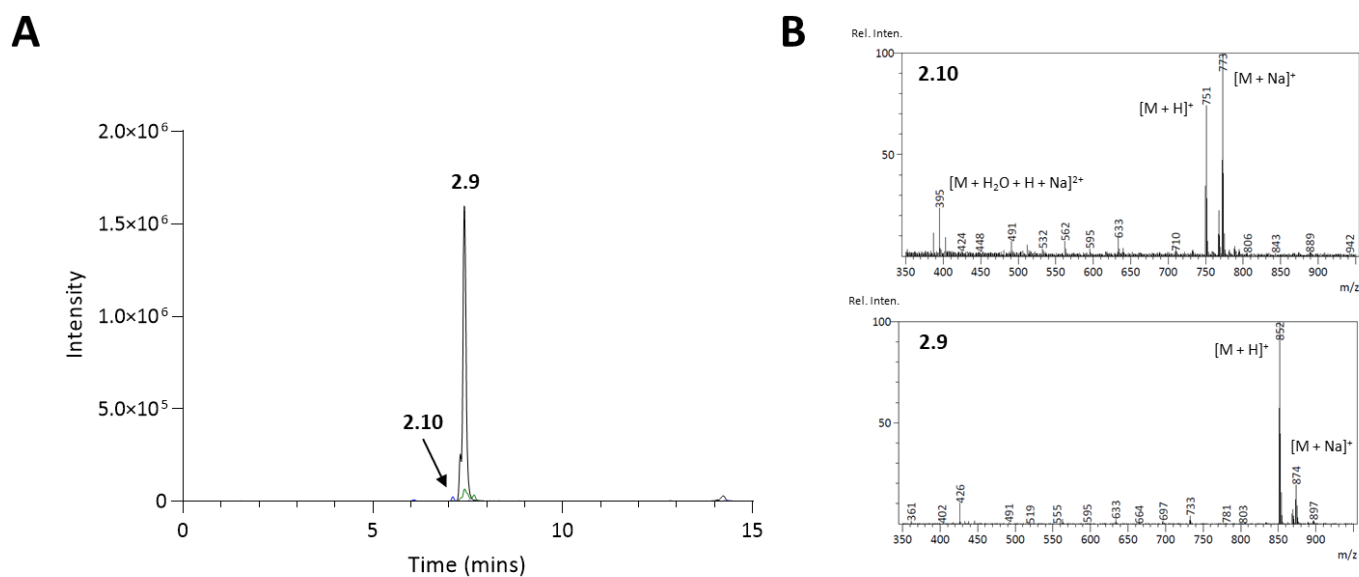


Figure S2.35. Rzm-TE 0-hour incubation with substrate 2.9. (A) Extracted ion chromatogram for Rzm-TE (25 μ M) assay after 0 hours of incubation with 1 mM of **2.9** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 750$ (**2.10**), 852 (**2.9**). (B) Mass spectra for substrate **2.9** and hydrolysis product **2.10**, mass adducts are identified in each spectra. Thermally denatured and substrate only controls all showed similar results for all time points.

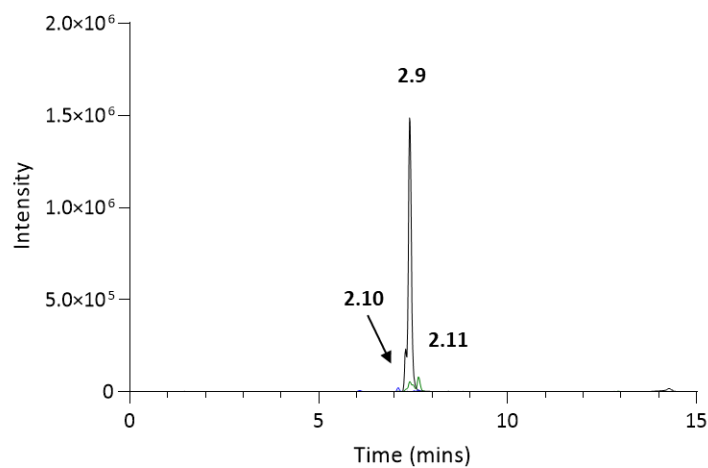
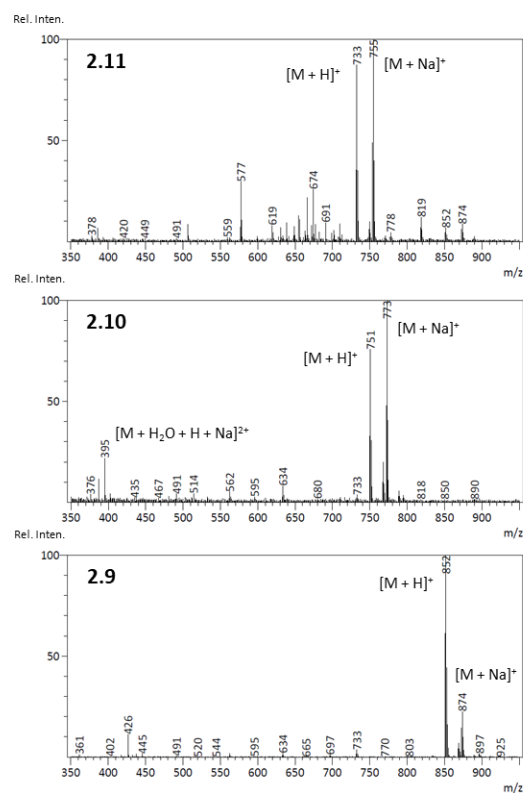
A**B**

Figure S2.36. Rzm-TE 3-hour incubation with substrate 2.9. (A) Extracted ion chromatogram for Rzm-TE (25 μ M) assay after 3 hours of incubation with 1 mM of **2.9** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 733$ (**2.11**), 750 (**2.10**) and 852 (**2.9**). (B) Mass spectra for substrate **2.9**, hydrolysis product **2.10** and macrocyclic product **2.11**, mass adducts are identified in each spectra.

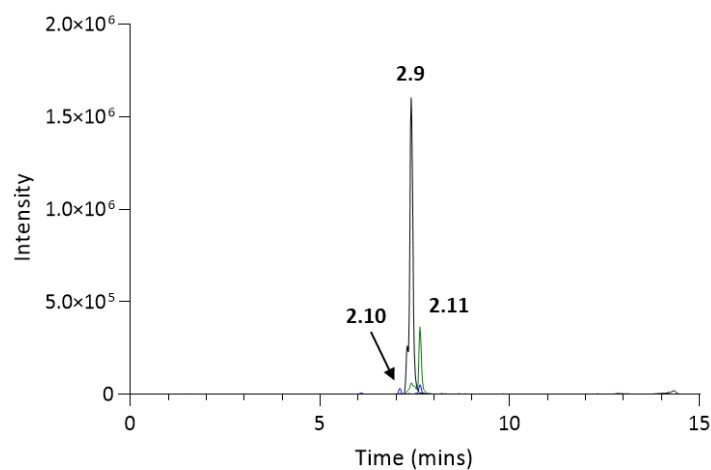
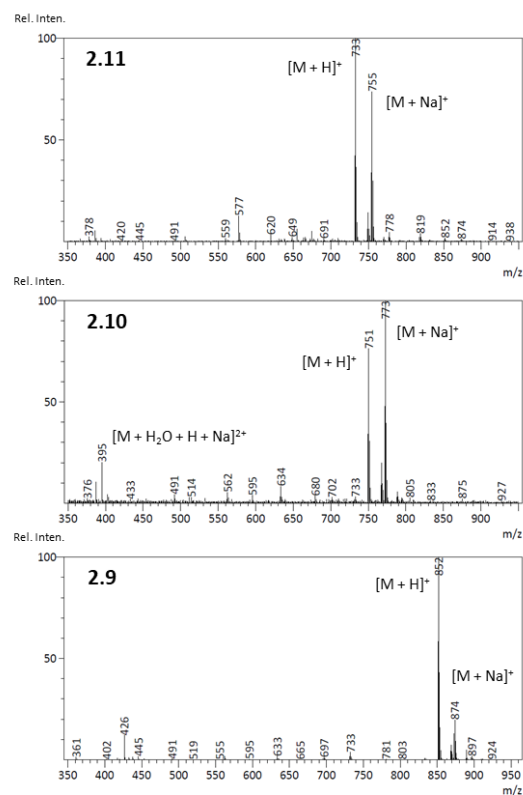
A**B**

Figure S2.37. Rzm-TE 24-hour incubation with substrate 2.9. (A) Extracted ion chromatogram for Rzm-TE (25 μ M) assay after 24 hours of incubation with 1 mM of **2.9** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 733$ (**2.11**), 750 (**2.10**) and 852 (**2.9**). (B) Mass spectra for substrate **2.9**, hydrolysis product **2.10** and macrocyclic product **2.11**, mass adducts are identified in each spectra.

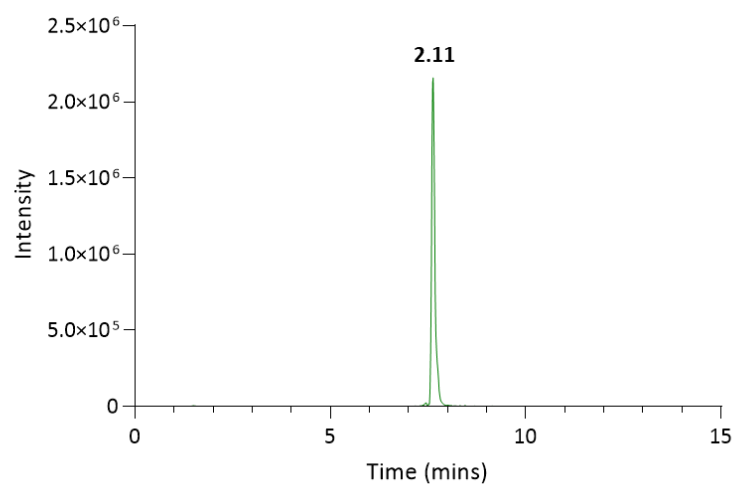
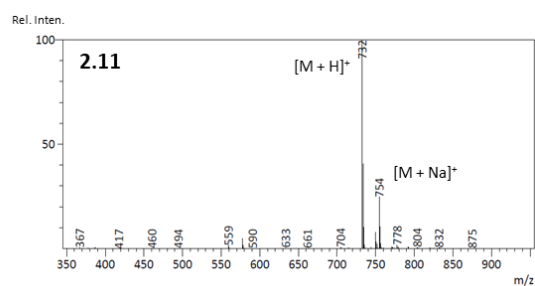
A**B**

Figure S2.38. Macrocycle 2.11 only control. (A) Extracted ion chromatogram for macrocycle **2.11** (1 mM) only control assay after 0 hours of incubation in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 733$ (**2.11**). (B) Mass spectra for macrocycle authentic standard **2.11**, mass adducts are identified.

Example UV trace for enzymatic assays

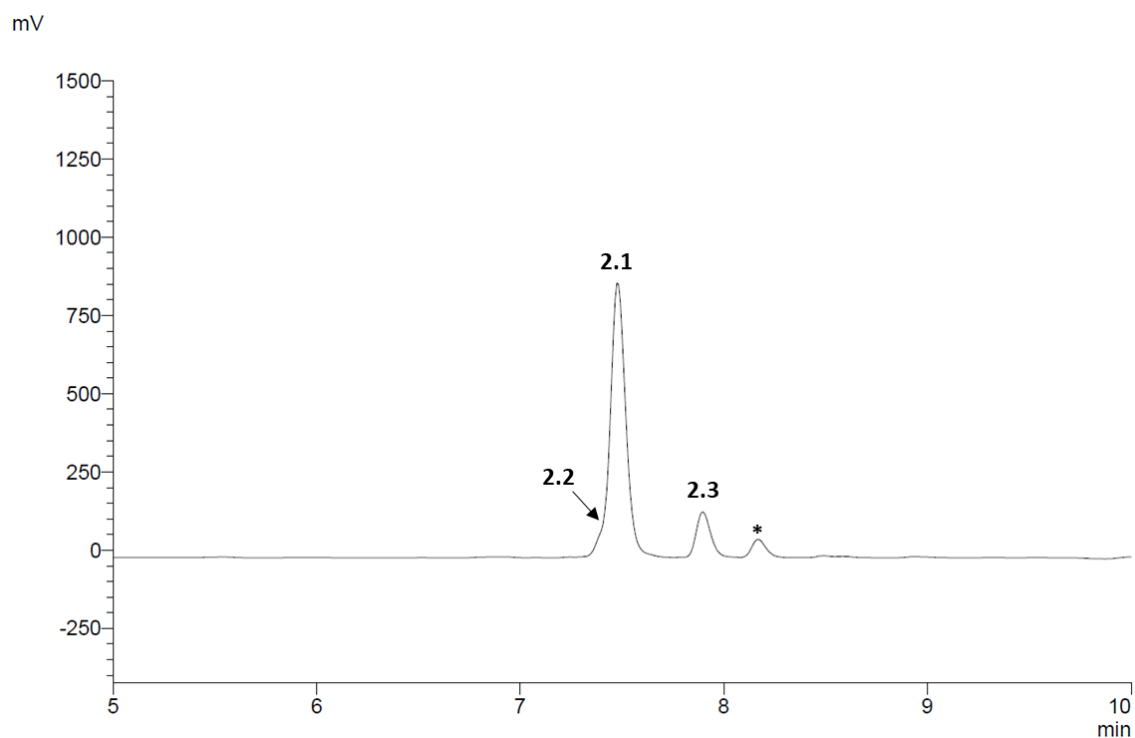


Figure S2.42. UV trace of Sgd-TE 3-hour incubation with substrate 2.1. UV trace was acquired at 210 nm for the Sgd-TE (5 μ M) assay after 3 hours of incubation with 1 mM of **2.1** in potassium phosphate buffer (50 mM, pH 7.4). Peaks corresponding to the substrate **2.1**, hydrolysis **2.2** and macrocycle **2.3** products are identified. Both **2.2** and **2.1** peaks overlap leading to a small shoulder peak corresponding to **2.2**, as shown. Peak labeled with an asterisk is an injector peak.

2.7 References

- (1) Wenzel, M.; Rautenbach, M.; Vosloo, J. A.; Siersma, T.; Aisenbrey, C. H. M.; Zaitseva, E.; Laubscher, W. E.; van Rensburg, W.; Behrends, J. C.; Bechinger, B.; Hamoen, L. W. The Multifaceted Antibacterial Mechanisms of the Pioneering Peptide Antibiotics Tyrocidine and Gramicidin S. *MBio* **2018**, *9* (5). <https://doi.org/10.1128/mBio.00802-18>.
- (2) Evidente, A. Bioactive Lipodepsipeptides Produced by Bacteria and Fungi. *Int. J. Mol. Sci.* **2022**, *23* (20), 12342. <https://doi.org/10.3390/ijms232012342>.
- (3) Bionda, N.; Pitteloud, J. P.; Cudic, P. Cyclic Lipodepsipeptides: A New Class of Antibacterial Agents in the Battle against Resistant Bacteria. *Future Med. Chem.* **2013**, *5* (11), 1311–1330. <https://doi.org/10.4155/fmc.13.86>.
- (4) Kim, G. J.; Li, X.; Kim, S. H.; Yang, I.; Hahn, D.; Chin, J.; Nam, S. J.; Nam, J. W.; Nam, D. H.; Oh, D. C.; Chang, H. W.; Choi, H. Seongsanamides A-D: Antiallergic Bicyclic Peptides from *Bacillus Safensis* KCTC 12796BP. *Org. Lett.* **2018**, *20* (23), 7539–7543. <https://doi.org/10.1021/acs.orglett.8b03293>.
- (5) Shabani, S.; Hutton, C. A. Total Synthesis of Seongsanamide B. *Org. Lett.* **2020**, *22* (11), 4557–4561. <https://doi.org/10.1021/acs.orglett.0c01642>.
- (6) Hanh, N. P. K.; Kim, S. H.; Kim, G. J.; Choi, H.; Nam, D. H. The Complete Genome Sequence of a Marine Sponge-Associated Bacteria, *Bacillus Safensis* KCTC 12796BP, Which Produces the Anti-Allergic Compounds. *Korean J. Microbiol.* **2018**, *54* (4), 448–452. <https://doi.org/10.7845/kjm.2018.8069>.
- (7) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The Specificity-Confering Code of Adenylation Domains in Nonribosomal Peptide Synthetases. *Chem. Biol.* **1999**, *6* (8), 493–505. [https://doi.org/10.1016/S1074-5521\(99\)80082-9](https://doi.org/10.1016/S1074-5521(99)80082-9).
- (8) Kries, H.; Niquille, D. L.; Hilvert, D. A Subdomain Swap Strategy for Reengineering Nonribosomal Peptides. *Chem. Biol.* **2015**, *22* (5), 640–648. <https://doi.org/10.1016/j.chembiol.2015.04.015>.
- (9) Crüsemann, M.; Kohlhaas, C.; Piel, J. Evolution-Guided Engineering of Nonribosomal Peptide Synthetase Adenylation Domains. *Chem. Sci.* **2013**, *4* (3), 1041–1045. <https://doi.org/10.1039/c2sc21722h>.
- (10) Challis, G. L.; Ravel, J.; Townsend, C. A. Predictive, Structure-Based Model of Amino Acid Recognition by Nonribosomal Peptide Synthetase Adenylation Domains. *Chem. Biol.* **2000**, *7* (3), 211–224. [https://doi.org/10.1016/S1074-5521\(00\)00091-0](https://doi.org/10.1016/S1074-5521(00)00091-0).
- (11) Horsman, M. E.; Hari, T. P. A. A.; Boddy, C. N. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016**, *33* (2), 183–202. <https://doi.org/10.1039/c4np00148f>.
- (12) Podust, L. M.; Sherman, D. H. Diversity of P450 Enzymes in the Biosynthesis of Natural Products. *Nat. Prod. Rep.* **2012**, *29* (10), 1251–1266.
- (13) Wang, X.; Zhou, H.; Chen, H.; Jing, X.; Zheng, W.; Li, R.; Sun, T.; Liu, J.; Fu, J. Discovery of Recombinases Enables Genome Mining of Cryptic Biosynthetic Gene Clusters in Burkholderiales Species. *PNAS* **2018**, *115* (18), E4255–E4263. <https://doi.org/10.1073/pnas.1720941115>.
- (14) Lackner, G.; Moebius, N.; Partida-Martinez, L. P.; Boland, S.; Hertweck, C. Evolution of an Endofungal Lifestyle: Deductions from the Burkholderia Rhizoxinica Genome. *BMC Genomics* **2011**, *12* (210), 1–13. <https://doi.org/10.1186/1471-2164-12-210>.
- (15) Lackner, G.; Moebius, N.; Partida-Martinez, L.; Hertweck, C. Complete Genome Sequence of Burkholderia Rhizoxinica, an Endosymbiont of Rhizopus Microsporus. *J. Bacteriol.* **2011**, *193* (3), 783–784. <https://doi.org/10.1128/JB.01318-10>.

- (16) Alonzo, D. A.; Schmeing, T. M. Biosynthesis of Depsipeptides, or Depsi: The Peptides with Varied Generations. *Protein Sci.* **2020**, *29* (12), 2316–2347. <https://doi.org/10.1002/pro.3979>.
- (17) Chooi, Y. H.; Tang, Y. Adding the Lipo to Lipopeptides: Do More with Less. *Chem. Biol.* **2010**, *17* (8), 791–793. <https://doi.org/10.1016/j.chembiol.2010.08.001>.
- (18) Zhong, L.; Diao, X.; Zhang, N.; Li, F.; Zhou, H.; Chen, H.; Bai, X.; Ren, X.; Zhang, Y.; Wu, D.; Bian, X. Engineering and Elucidation of the Lipoinitiation Process in Nonribosomal Peptide Biosynthesis. *Nat. Commun.* **2021**, *12* (296), 1–15. <https://doi.org/10.1038/s41467-020-20548-8>.
- (19) Imker, H. J.; Krahn, D.; Clerc, J.; Kaiser, M.; Walsh, C. T. N-Acylation during Glidobactin Biosynthesis by the Tridomain Nonribosomal Peptide Synthetase Module GlbF. *Chem. Biol.* **2010**, *17* (10), 1077–1083. <https://doi.org/10.1016/j.chembiol.2010.08.007>.
- (20) Shabani, S.; Hutton, C. A. Depsipeptide Synthesis Using a Late-Stage Ag(I)-Promoted Macrolactonisation of Peptide Thioamides. *Chem. Commun.* **2021**, *57* (16), 2081–2084. <https://doi.org/10.1039/d0cc07747j>.
- (21) Kaur, H.; Harris, P. W. R.; Little, P. J.; Brimble, M. A. Total Synthesis of the Cyclic Depsipeptide YM-280193, a Platelet Aggregation Inhibitor. *Org. Lett.* **2015**, *17* (3), 492–495. <https://doi.org/10.1021/ol503507g>.
- (22) Hawkins, P. M. E.; Tran, W.; Nagalingam, G.; Cheung, C. Y.; Giltrap, A. M.; Cook, G. M.; Britton, W. J.; Payne, R. J. Total Synthesis and Antimycobacterial Activity of Ohmyungsamycin A, Deoxyecumicin, and Ecumicin. *Chem. - A Eur. J.* **2020**, *26* (66), 15200–15205. <https://doi.org/10.1002/chem.202002408>.
- (23) Han, F.; Guo, Y.; Ye, T. Total Synthesis of Antiallergic Bicyclic Peptide Seongsanamide A. *Org. Chem. Front.* **2020**, *7* (13), 1658–1662. <https://doi.org/10.1039/d0qo00531b>.
- (24) Yudin, A. K. Macrocycles: Lessons from the Distant Past, Recent Developments, and Future Directions. *Chem. Sci.* **2015**, *6* (1), 30–49. <https://doi.org/10.1039/c4sc03089c>.
- (25) Ghosh, A. K.; Xu, C. X. A Convergent Synthesis of the Proposed Structure of Antitumor Depsipeptide Stereocalpin A. *Org. Lett.* **2009**, *11* (9), 1963–1966. <https://doi.org/10.1021/ol900412u>.
- (26) Cochrane, J. R.; McErlean, C. S. P.; Jolliffe, K. A. Total Synthesis and Assignment of the Side Chain Stereochemistry of LI-F04a: An Antimicrobial Cyclic Depsipeptide. *Org. Lett.* **2010**, *12* (15), 3394–3397. <https://doi.org/10.1021/ol101254m>.
- (27) Kopple, K. D. Synthesis of Cyclic Peptides. *J. Pharm. Sci.* **1972**, *61* (9), 1345–1356. <https://doi.org/10.1002/jps.2600610902>.
- (28) Kuranaga, T.; Enomoto, A.; Tan, H.; Fujita, K.; Wakimoto, T. Total Synthesis of Theonellaeptolide Id. *Org. Lett.* **2017**, *19* (6), 1366–1369. <https://doi.org/10.1021/acs.orglett.7b00249>.
- (29) Bando, Y.; Hou, Y.; Seyfarth, L.; Probst, J.; Götze, S.; Bogacz, M.; Hellmich, U. A.; Stallforth, P.; Mittag, M.; Arndt, H. D. Total Synthesis and Structure Correction of the Cyclic Lipodepsipeptide Orfamide A. *Chem. - A Eur. J.* **2022**, *28* (20), 2–6. <https://doi.org/10.1002/chem.202104417>.
- (30) Brazeau-Henrie, J. T.; Paquette, A. R.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N. Total and Chemoenzymatic Synthesis of Seongsanamide E. *Org. Lett.* **2022**, *24* (35), 6369–6373. <https://doi.org/10.1021/acs.orglett.2c02271>.
- (31) Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dornan, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W. Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature* **2019**, *565* (7737), 112–117. <https://doi.org/10.1038/s41586-018-0781-z>.
- (32) Heberlig, G. W.; Boddy, C. N. Thioesterase from Cereulide Biosynthesis Is Responsible for Oligomerization and Macrocyclization of a Linear Tetradepsipeptide. *J. Nat. Prod.* **2020**, *83* (6),

- 1990–1997. <https://doi.org/10.1021/acs.jnatprod.0c00333>.
- (33) Heberlig, G. W.; Brown, J. T. C.; Simard, R. D.; Wirz, M.; Zhang, W.; Wang, M.; Susser, L. I.; Horsman, M. E.; Boddy, C. N. Chemoenzymatic Macrocyclic Synthesis Using Resorcylic Acid Lactone Thioesterase Domains. *Org. Biomol. Chem.* **2018**, *16* (32), 5771–5779. <https://doi.org/10.1039/c8ob01512k>.
 - (34) Brazeau-Henrie, J. T.; Paquette, A. R.; Boddy, C. N. In Vitro Biochemical Characterization of Excised Macrocyclizing Thioesterase Domains from Non-Ribosomal Peptide Synthetases. In *Non-Ribosomal Peptide Biosynthesis and Engineering*; 2023; Vol. 2670, pp 101–125.
 - (35) Boddy, C. N.; Schneider, T. L.; Hotta, K.; Walsh, C. T.; Khosla, C. Epothilone C Macrolactonization and Hydrolysis Are Catalyzed by the Isolated Thioesterase Domain of Epothilone Polyketide Synthase. *J. Am. Chem. Soc.* **2003**, *125* (12), 3428–3429. <https://doi.org/10.1021/ja0298646>.
 - (36) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. Characterization of the Surfactin Synthetase C-Terminal Thioesterase Domain as a Cyclic Depsipeptide Synthase. *Biochemistry* **2002**, *41* (45), 13350–13359. <https://doi.org/10.1021/bi026592a>.
 - (37) Dawson, P. E.; Kent, S. B. H. Synthesis of Native Proteins by Chemical Ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
 - (38) Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. Solid Phase Synthesis of Peptide C-Terminal Thioesters by Fmoc/t-Bu Chemistry. *J. Am. Chem. Soc.* **1999**, *121* (49), 11369–11374. <https://doi.org/10.1021/ja992668n>.
 - (39) Zheng, J. S.; Chang, H. N.; Wang, F. L.; Liu, L. Fmoc Synthesis of Peptide Thioesters without Post-Chain-Assembly Manipulation. *J. Am. Chem. Soc.* **2011**, *133* (29), 11080–11083. <https://doi.org/10.1021/ja204088a>.
 - (40) Blanco-Canosa, J. B.; Dawson, P. E. An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation. *Angew. Chemie - Int. Ed.* **2008**, *47* (36), 6851–6855. <https://doi.org/10.1002/anie.200705471>.
 - (41) Zheng, J. S.; Tang, S.; Qi, Y. K.; Wang, Z. P.; Liu, L. Chemical Synthesis of Proteins Using Peptide Hydrazides as Thioester Surrogates. *Nat. Protoc.* **2013**, *8* (12), 2483–2495. <https://doi.org/10.1038/nprot.2013.152>.
 - (42) Elashal, H. E.; Sim, Y. E.; Raj, M. Serine Promoted Synthesis of Peptide Thioester-Precursor on Solid Support for Native Chemical Ligation. *Chem. Sci.* **2016**, *8* (1), 117–123. <https://doi.org/10.1039/C6SC02162J>.
 - (43) Von Eggelkraut-Gottanka, R.; Klose, A.; Beck-Sickinger, A. G.; Beyermann, M. Peptide Athioester Formation Using Standard Fmoc-Chemistry. *Tetrahedron Lett.* **2003**, *44* (17), 3551–3554. [https://doi.org/10.1016/S0040-4039\(03\)00582-3](https://doi.org/10.1016/S0040-4039(03)00582-3).
 - (44) Chandra, J.; Manne, S. R.; Mondal, S.; Mandal, B. (E)-Ethyl-2-Cyano-2-(((2,4,6-Trichlorobenzoyl)Oxy)Imino)Acetate: A Modified Yamaguchi Reagent for Enantioselective Esterification, Thioesterification, Amidation, and Peptide Synthesis. *ACS Omega* **2018**, *3* (6), 6120–6133. <https://doi.org/10.1021/acsomega.8b00732>.
 - (45) Gimenez, D.; Phelan, A.; Murphy, C. D.; Cobb, S. L. Fengycin A Analogues with Enhanced Chemical Stability and Antifungal Properties. *Org. Lett.* **2021**, *23* (12), 4672–4676. <https://doi.org/10.1021/acs.orglett.1c01387>.
 - (46) Hawkins, P. M. E.; Giltrap, A. M.; Nagalingam, G.; Britton, W. J.; Payne, R. J. Total Synthesis of Ecumicin. *Org. Lett.* **2018**, *20* (4), 1019–1022. <https://doi.org/10.1021/acs.orglett.7b03967>.
 - (47) Moreira, R.; Wolfe, J.; Taylor, S. D. A High-Yielding Solid-Phase Total Synthesis of Daptomycin Using a Fmoc SPPS Stable Kynurenine Synthron. *Org. Biomol. Chem.* **2021**, *19* (14), 3144–3153. <https://doi.org/10.1039/d0ob02504f>.

- (48) Argyropoulos, P.; Bergeret, F.; Pardin, C.; Reimer, J. M.; Pinto, A.; Boddy, C. N.; Schmeing, T. M. Towards a Characterization of the Structural Determinants of Specificity in the Macrocyclizing Thioesterase for Deoxyerythronolide B Biosynthesis. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (3), 486–497. <https://doi.org/10.1016/j.bbagen.2015.11.007>.
- (49) Lu, H.; Tsai, S. C.; Khosla, C.; Cane, D. E. Expression, Site-Directed Mutagenesis, and Steady State Kinetic Analysis of the Terminal Thioesterase Domain of the Methymycin/Picromycin Polyketide Synthase. *Biochemistry* **2002**, *41* (42), 12590–12597. <https://doi.org/10.1021/bi026006d>.
- (50) Tsai, S. C.; Miercke, L. J. W.; Krucinski, J.; Gokhale, R.; Chen, J. C. H.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. Crystal Structure of the Macrocyclic-Forming Thioesterase Domain of the Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (26), 14808–14813. <https://doi.org/10.1073/pnas.011399198>.
- (51) Pinto, A.; Wang, M.; Horsman, M.; Boddy, C. N. 6-Deoxyerythronolide B Synthase Thioesterase-Catalyzed Macrocyclization Is Highly Stereoselective. *Org. Lett.* **2012**, *14* (9), 2278–2281. <https://doi.org/10.1021/ol300707j>.
- (52) He, W.; Wu, J.; Khosla, C.; Cane, D. E. Macrolactonization to 10-Deoxymethynolide Catalyzed by the Recombinant Thioesterase of the Picromycin/Methymycin Polyketide Synthase. *Bioorganic Med. Chem. Lett.* **2006**, *16* (2), 391–394. <https://doi.org/10.1016/j.bmcl.2005.09.077>.
- (53) Wang, M.; Boddy, C. N. Examining the Role of Hydrogen Bonding Interactions in the Substrate Specificity for the Loading Step of Polyketide Synthase Thioesterase Domains. *Biochemistry* **2008**, *47* (45), 11793–11803. <https://doi.org/10.1021/bi800963y>.
- (54) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. Mechanism and Specificity of the Terminal Thioesterase Domain from the Erythromycin Polyketide Synthase. *Chem. Biol.* **1999**, *6* (2), 117–125. [https://doi.org/10.1016/S1074-5521\(99\)80008-8](https://doi.org/10.1016/S1074-5521(99)80008-8).
- (55) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry* **2001**, *40* (24), 7092–7098. <https://doi.org/10.1021/bi010035r>.
- (56) Hari, T. P. A.; Labana, P.; Boileau, M.; Boddy, C. N. An Evolutionary Model Encompassing Substrate Specificity and Reactivity of Type I Polyketide Synthase Thioesterases. *ChemBioChem* **2014**, *15* (18), 2656–2661. <https://doi.org/10.1002/cbic.201402475>.
- (57) Mahlert, C.; Sieber, S. A.; Grünwald, J.; Marahiel, M. A. Chemoenzymatic Approach to Enantiopure Streptogramin B Variants: Characterization of Stereoselective Pristinamycin I Cyclase from *Streptomyces Pristinaespiralis*. *J. Am. Chem. Soc.* **2005**, *127* (26), 9571–9580. <https://doi.org/10.1021/ja051254t>.
- (58) Samel, S. A.; Wagner, B.; Marahiel, M. A.; Essen, L. O. The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-Ribosomal Lipopeptide. *J. Mol. Biol.* **2006**, *359* (4), 876–889. <https://doi.org/10.1016/j.jmb.2006.03.062>.
- (59) Mulyani, S.; Egel, E.; Kittel, C.; Turkanovic, S.; Wohlleben, W.; Süssmuth, R. D.; Van Pée, K. H. The Thioesterase Bhp Is Involved in the Formation of β -Hydroxytyrosine during Balhimycin Biosynthesis in *Amycolatopsis Balhimycina*. *ChemBioChem* **2010**, *11* (2), 266–271. <https://doi.org/10.1002/cbic.200900600>.
- (60) Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A. Regeneration of Misprimed Nonribosomal Peptide Synthetases by Type II Thioesterases. *PNAS* **2002**, *99* (22), 14083–14088. <https://doi.org/10.1073/pnas.212382199>.
- (61) Lin, H.; Thayer, D. A.; Wong, C. H.; Walsh, C. T. Macrolactamization of Glycosylated Peptide Thioesters by the Thioesterase Domain of Tyrocidine Synthetase. *Chem. Biol.* **2004**, *11* (12), 1635–1642. <https://doi.org/10.1016/j.chembiol.2004.09.015>.

- (62) Hoyer, K. M.; Mahlert, C.; Marahiel, M. A. The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chem. Biol.* **2007**, *14* (1), 13–22. <https://doi.org/10.1016/j.chembiol.2006.10.011>.
- (63) Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. Generality of Peptide Cyclization Catalyzed by Isolated Thioesterase Domains of Nonribosomal Peptide Synthetases. *Biochemistry* **2001**, *40* (24), 7099–7108. <https://doi.org/10.1021/bi010036j>.
- (64) Gehring, A. M.; Mori, I.; Walsh, C. T. Reconstitution and Characterization of the Escherichia Coli Enterobactin Synthetase from EntB, EntE, and EntF. *Biochemistry* **1998**, *37* (8), 2648–2659. <https://doi.org/10.1021/bi9726584>.
- (65) Grünewald, J.; Sieber, S. A.; Marahiel, M. A. Chemo- and Regioselective Peptide Cyclization Triggered by the N-Terminal Fatty Acid Chain Length: The Recombinant Cyclase of the Calcium-Dependent Antibiotic from Streptomyces Coelicolor. *Biochemistry* **2004**, *43* (10), 2915–2925. <https://doi.org/10.1021/bi036140d>.
- (66) Chou, T. C.; O'Connor, O. A.; Tong, W. P.; Guan, Y.; Zhang, Z. G.; Stachel, S. J.; Lee, C.; Danishefsky, S. J. The Synthesis, Discovery, and Development of a Highly Promising Class of Microtubule Stabilization Agents: Curative Effects of Desoxyepothilones B and F against Human Tumor Xenografts in Nude Mice. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (14), 8113–8118. <https://doi.org/10.1073/pnas.131153098>.
- (67) Blum, W.; Aichholz, R.; Ramstein, P.; Kühnöl, J.; Brügggen, J.; O'Reilly, T.; Flörsheimer, A. In Vivo Metabolism of Epothilone B in Tumor-Bearing Nude Mice: Identification of Three New Epothilone B Metabolites by Capillary High-Pressure Liquid Chromatography/Mass Spectrometry/Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15* (1), 41–49. [https://doi.org/10.1002/1097-0231\(20010115\)15:1<41::AID-RCM190>3.0.CO;2-W](https://doi.org/10.1002/1097-0231(20010115)15:1<41::AID-RCM190>3.0.CO;2-W).
- (68) Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. Peptide Cyclization Catalysed by the Thioesterase Domain of Tyrocidine Synthetase. *Nature* **2000**, *407* (6801), 215–218. <https://doi.org/10.1038/35025116>.
- (69) Bruner, S. D.; Weber, T.; Kohli, R. M.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T.; Stubbs, M. T. Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE. *Structure* **2002**, *10* (3), 301–310. [https://doi.org/10.1016/S0969-2126\(02\)00716-5](https://doi.org/10.1016/S0969-2126(02)00716-5).
- (70) Bu, X.; Wu, X.; Xie, G.; Guo, Z. Synthesis of Tyrocidine A and Its Analogues by Spontaneous Cyclization in Aqueous Solution. *Org. Lett.* **2002**, *4* (17), 2893–2895. <https://doi.org/10.1021/ol0263191>.
- (71) Kuo, M. C.; Gibbons, W. A. Determination of Individual Side-Chain Conformations, Tertiary Conformations, and Molecular Topography of Tyrocidine A from Scalar Coupling Constants and Chemical Shifts†. *Biochemistry* **1979**, *18* (26), 5855–5867. <https://doi.org/10.1021/bi00593a016>.
- (72) Parenty, A.; Moreau, X.; Campagne, J. M. Macrolactonizations in the Total Synthesis of Natural Products. *Chem. Rev.* **2006**, *106* (3), 911–939. <https://doi.org/10.1021/cr0301402>.
- (73) Mortison, J. D.; Sherman, D. H. Frontiers and Opportunities in Chemoenzymatic Synthesis. *J. Org. Chem.* **2010**, *75* (21), 7041–7051. <https://doi.org/10.1021/jo101124n>.
- (74) Kaneda, M.; Sueyoshi, K.; Teruya, T.; Ohno, H.; Fujii, N.; Oishi, S. Total Synthesis of Odoamide, a Novel Cyclic Depsipeptide, from an Okinawan Marine Cyanobacterium. *Org. Biomol. Chem.* **2016**, *14* (38), 9093–9104. <https://doi.org/10.1039/c6ob01583b>.
- (75) Grab, H. A.; Kirsch, V. C.; Sieber, S. A.; Bach, T. Total Synthesis of the Cyclic Depsipeptide Vioprolide D via Its (Z)-Diastereoisomer. *Angew. Chemie - Int. Ed.* **2020**, *59* (30), 12357–12361. <https://doi.org/10.1002/anie.202002328>.
- (76) Wang, D.; Song, S.; Tian, Y.; Xu, Y.; Miao, Z.; Zhang, A. Total Synthesis of the Marine Cyclic

- Depsipeptide Viequeamide A. *J. Nat. Prod.* **2013**, 76 (5), 974–978. <https://doi.org/10.1021/np4001027>.
- (77) Mandalapu, D.; Ji, X.; Chen, J.; Guo, C.; Liu, W. Q.; Ding, W.; Zhou, J.; Zhang, Q. Thioesterase-Mediated Synthesis of Teixobactin Analogues: Mechanism and Substrate Specificity. *J. Org. Chem.* **2018**, 83 (13), 7271–7275. <https://doi.org/10.1021/acs.joc.7b02462>.
- (78) Hou, J.; Robbel, L.; Marahiel, M. A. Identification and Characterization of the Lysobactin Biosynthetic Gene Cluster Reveals Mechanistic Insights into an Unusual Termination Module Architecture. *Chem. Biol.* **2011**, 18 (5), 655–664. <https://doi.org/10.1016/j.chembiol.2011.02.012>.
- (79) Lepetit, C. A.; Paquette, A. R.; Brazeau-Henrie, J. T.; Boddy, C. N. Total and Chemoenzymatic Synthesis of the Lipodepsipeptide Rhizomide A. *Bioorg. Med. Chem. Lett.* **2023**, 96 (Article 129506), 1–5. <https://doi.org/10.1016/j.bmcl.2023.129506>.
- (80) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, 596 (7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.

CHAPTER 3 : Characterization of an NRPS tandem TE domain in depsipeptide macrolactonization

3.1 Preface

The work presented in this chapter characterizes the macrolactonization ability and substrate tolerance of an unusual tandem TE domain in the gacamide A biosynthetic gene cluster. Further experiments are required to finalize this study for publication; however the data from this chapter clearly characterize the key unique aspects of the Gac-tandem TE domains and their substrate scope.

Chapter contributions

I performed all cloning, mutagenesis, product distribution assays and data analysis. Dr. André R. Paquette synthesized the *N*-Cbz and *N*-acetyl peptide substrates. Corinne A. Lepetit synthesized the Thr_x peptide substrates. Dr. Mark Silby (University of Massachusetts Dartmouth) kindly provided me with the *Pseudomonas fluorescens* Pf0-1 host organism for cloning of the gacamide A tandem TE domain. Dr. Christopher N. Boddy, Dr. André R. Paquette and I devised the project.

3.2 Introduction

The worldwide emergence of antibiotic resistant bacterial pathogens has proven to be a problem in the modern world. Of the many antimicrobial compounds discovered in nature, macrocyclic lipodepsipeptides (mLPD) show great potential to tackle this problem. For example, daptomycin has potent antimicrobial activity against methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-heterointermediate *S. aureus* (hVISA) strains by forming ion permeable pores in bacterial membranes, inducing cell death.¹⁻⁴ mLPD are composed of a peptide with one or more ester linkages between two amino acids, forming a macrocyclic core, and a lipid tail bound at the N-terminus. Typically, the peptide core is biosynthesized by an NRPS, offering substantial structural versatility through variations in the selection of amino acids incorporated, their configuration (D or L), the number and the presence of non-proteogenic amino acids. Moreover, the diversity can be further augmented with the incorporation of different N-terminal lipids, which can differ in length, composition (ex: β -OH, branched lipids) and in saturation.⁵⁻⁷ Because of this, mLPD also display a wide range of useful biological activities, including antimicrobial, antiviral, antifungal, antitumor and immunomodulator.^{8,9}

Recently, a biosynthetic gene cluster responsible for production of the mLPD gacamide A (Gac A) was discovered through genome mining of the *Pseudomonas fluorescens* Pf0-1 genome. In the WT strain, this gene cluster is silent due to a nonfunctional GacS/GacA global regulator housing a point mutation in the *gacA* gene.^{10,11} Incorporation of a functional *gacA* gene in the WT strain allowed for production of Gac A, its isolation and structural determination.¹¹ Gac A does not show any important biological activity, but rather plays a role in increasing swarming and hemolytic activity of *P. fluorescens*,¹¹ much like the surface-active rhamnolipids found in pseudomonads.^{12,13} Gac A is biosynthesized by three NRPS complexes housing a total of 11 modules that assemble a peptide core with the D-Leu-D-Asp-D-Gln-D-*alle* (*allolle*)-D-Leu-D-Gln-D-Ser-L-Leu-L-Leu-D-Ser-L-Ile sequence (**Figure 3.1**). A starter condensation domain (C_s) is

positioned at module 1 for condensation of an *S*- β -hydroxydecanoic acid (β -HDA) onto the N-terminus of D-Leu₁.^{14–16} Most condensation (C) domains in the gene cluster are dual epimerization-condensation (C/E) domains that convert the upstream loaded L-amino acids into their D enantiomers before condensation for NRP extension.^{17,18} Modules 9 and 10; however, are exceptions, featuring C domains that specifically catalyze condensation between two L-amino acids (^LC_L).¹⁸ Finally, the linear peptide precursor is passed onto a tandem TE domain (TE1-TE2) for its release via macrolactonization between the D-Ser₇ and L-Ile₁₁ residues (**Figure 3.1**).

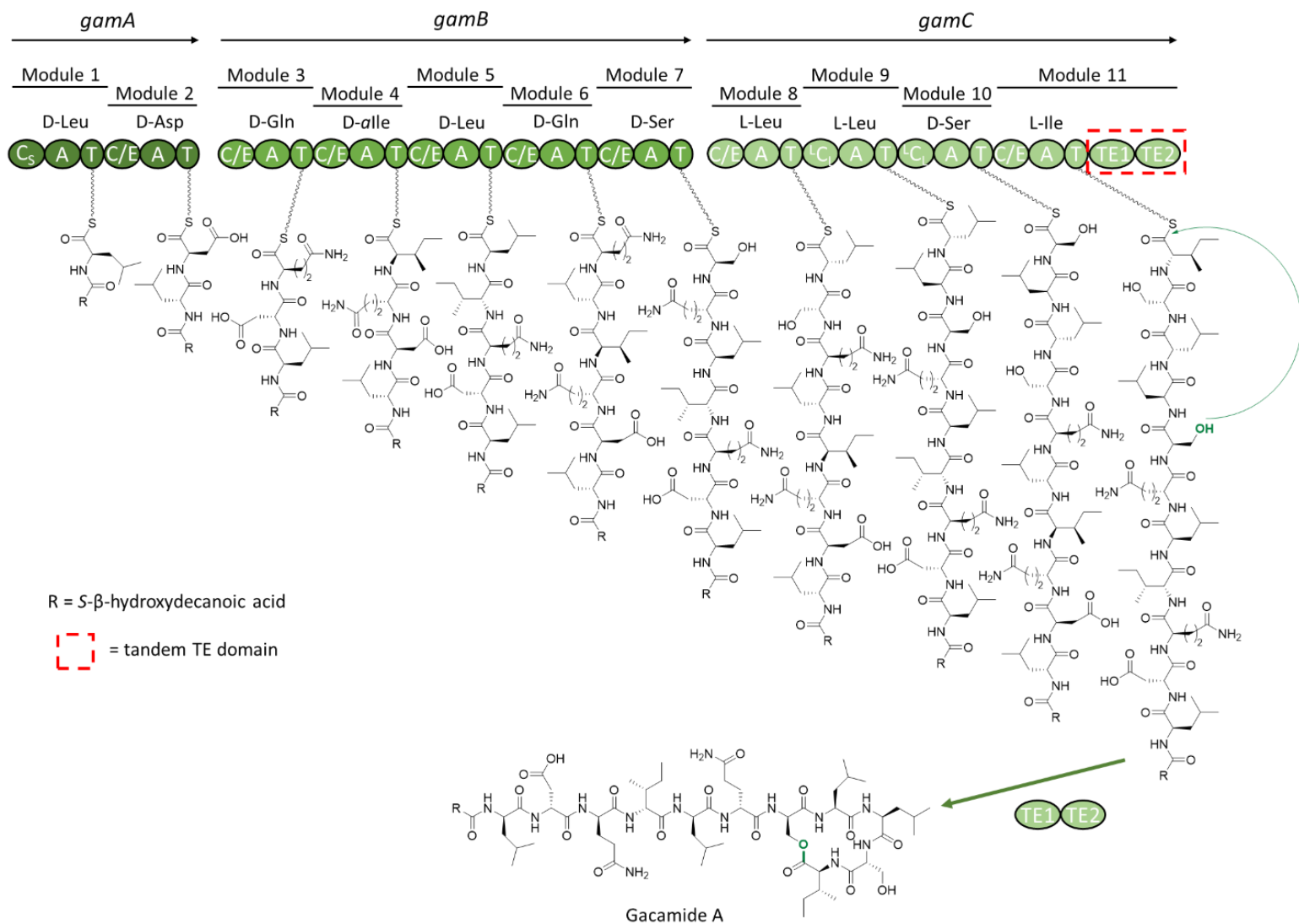


Figure 3.1. Biosynthesis of Gacamide A. The biosynthetic gene cluster contains 3 genes (*gamA*, *gamB* and *gamC*) that encode for their respective NRPS complexes. Dual C/E domains are appropriately positioned for conversion of loaded L-amino acids to their D-amino acid enantiomers. Modules 9 and 10 contain ^LC_L domains that specifically catalyze amide bond formation between two L-amino acids. Module 11 encodes a tandem TE domain (TE1-TE2) responsible for the macrocyclic release of gacamide A (Gac). Abbreviations; C_s, starter condensation domain; C, condensation domain; C/E, dual epimerization-condensation domain; A, adenylation domain; T, thiolation or peptidyl carrier domain; ^LC_L, condensation domain with L-amino acid specificity; TE, thioesterase domain.

The Gac A NRPS gene cluster presents an uncommon tandem TE domain (TE1-TE2) at its C-terminus instead of the canonical singular TE domain. Such domains have only been seen in mLPD biosynthetic machinery producing teixobactin (Txo),¹⁹ arthrofactin (Arf),²⁰ lysobactin (LybB),²¹ malleipeptin,²² syringopeptin²³ and massetolide A.²⁴ While it has been suggested that the TE1s are responsible for macrolactonization,²⁵ the available biochemical studies on the tandem systems from Txo, Arf and LybB biosynthesis cast doubt on this hypothesis (**Figure 3.2**). For instance, inactivation of LybB-TE2 in the tandem system yields no discernable impact on lysobactin production, suggesting TE1 is solely responsible for its formation.²¹ However, Arf-TE2 inactivation leads to a near-complete cessation of arthrofactin production.²⁰ This is strikingly different from Txo, where inactivating either TE does not affect product formation; however, the use of either TE1 or TE2 as stand-alone enzymes results in no detectable formation of product.¹⁹ Given the divergent findings from these studies, we set out to examine the role of the tandem TE domains in Gac A biosynthesis.

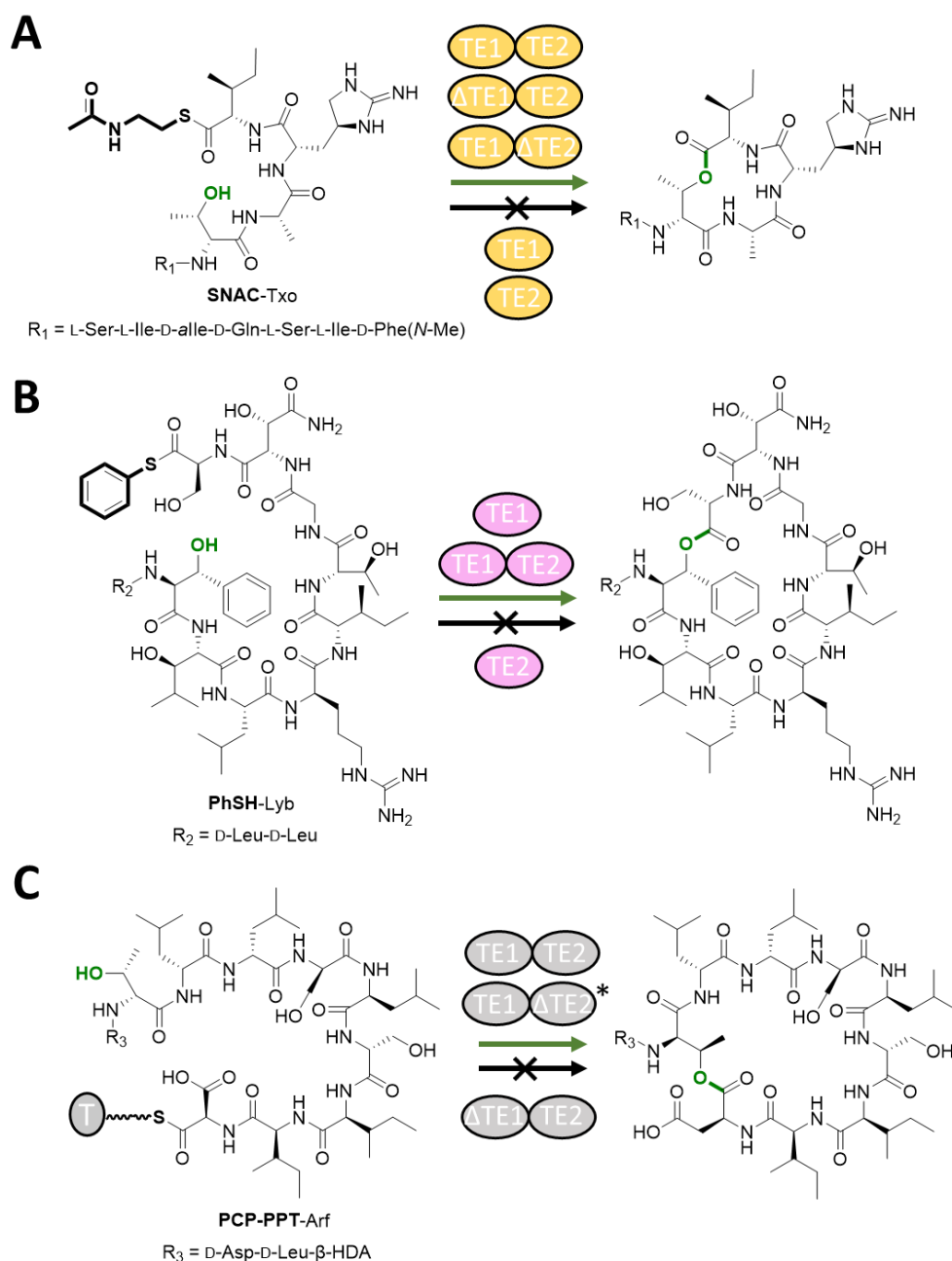


Figure 3.2. Summary of tandem TE domain activity. *In vitro* activity of the teixobactin (Txo, yellow, panel A), lysobactin (LybB, pink, panel B) and *in vivo* activity for arthrofactin (Arf, grey, panel C) tandem TE, inactivated mutant or stand-alone TE domains on thioester activated peptide substrates. ΔTE1 and ΔTE2 denotes enzyme inactivation by substituting their respective catalytic Ser for an Ala. The crossed out black arrows represent a condition that does not lead to formation of the macrocycle. Abbreviations; SNAC, *N*-acetylcysteamine; PhSH, thiophenol; *alle*, *alloisoleucine*. * construct has a 95% reduction in macrocycle production.²⁰

To study the macrocyclization capabilities of the Gac tandem TE domain (TE1-TE2), TE1-TE2, TE1 and TE1/TE2 inactivation mutants were characterized as stand-alone enzymes *in vitro*. Appropriate substrates containing only the residues required for macrocyclization were synthesized as *N*-acetylcysteamine (SNAC) activated *N*-Cbz (**3.1**) or *N*-acetyl (**3.4**) protected peptides. Enzymes were recombinantly expressed and purified for enzymatic assays with substrates **3.1** or **3.4** and analyzed by LCMS for unambiguous detection of products. We show TE1-TE2 is an active macrolactonization catalyst with both **3.1** and **3.4**, suggesting TE1-TE2 can tolerate N-terminal peptide modifications. Additionally, we establish that TE1 is responsible for the macrolactonization activity observed in the tandem domain. We also show the spontaneous macrocyclization of substrate **3.1**, suggesting substrate pre-organization drives macrocyclization as stipulated in our proposed CS model. Finally, a small *de novo* library of SNAC-peptides was synthesized to explore TE1-TE2 substrate tolerance. Incubations with the library showed TE1-TE2 macrolactonized two of the three peptides to form 5- and 7-residue non-native macrolactones. This suggests TE1-TE2 is ring size, stereo and sequence tolerant, making it a promising general peptide macrocyclization biocatalyst.

3.3 Results and discussion

Gac-TE1-TE2 can macrolactonize both *N*-Cbz (**3.1**) and *N*-acetyl (**3.4**) substrates

To biochemically characterize the Gac-TE tandem domain, the tandem domain (TE1-TE2) was expressed and purified as a stand-alone protein. The DNA coding sequence of TE1-TE2 was amplified by PCR from *Pseudomonas fluorescens* Pf0-1 genome and cloned into a pET-28a(+) expression vector under the control of an inducible T7 promoter fused to an N-terminal hexa-His (6xHis) tag. The resulting construct was overexpressed in *Escherichia coli* BL21(DE3) and purified by metal affinity chromatography to high purity (> 95%, **Figure S3.4**). Testing for macrolactonization activity was carried out by incubating TE1-TE2 (5 μ M) with either *N*-Cbz substrate **3.1** (0.9 mM) or *N*-acetyl substrate **3.4** (0.5 mM), both of which were provided to me by Dr. André Paquette, in 50 mM of potassium phosphate buffer (pH 7.4) at room temperature. Aliquots were taken at varying time points and analyzed by LCMS to evaluate product formation. In all cases, TE1-TE2 is able to convert substrates **3.1** and **3.4** into their respective macrocycles **3.3** and **3.6** (**Figure 3.3A**), which were absent in the substrate only and boiled enzyme controls within 22 hours of incubation (**Figure S3.7** and **S3.12**). After 3 hours of incubation, **3.3** was readily detected while no formation of **3.6** was observed, presumably due to the limit of detection of the instrument (**Figure S3.8** and **S3.13**). Further incubation of 22 hours leads to higher production of **3.3** and detection of **3.6** (**Figure 3.3B-C**, **S3.9** and **S3.15**). Minimal formation of *N*-Cbz hydrolysis product **3.2** was observed while *N*-acetyl hydrolysis product **3.5** was not detected, highlighting TE1-TE2's potential as a macrolactonization specific biocatalyst. The ability to process both substrates suggests TE1-TE2 can tolerate N-terminal peptide modifications, an attractive feature for accessing novel macrocyclic peptides.

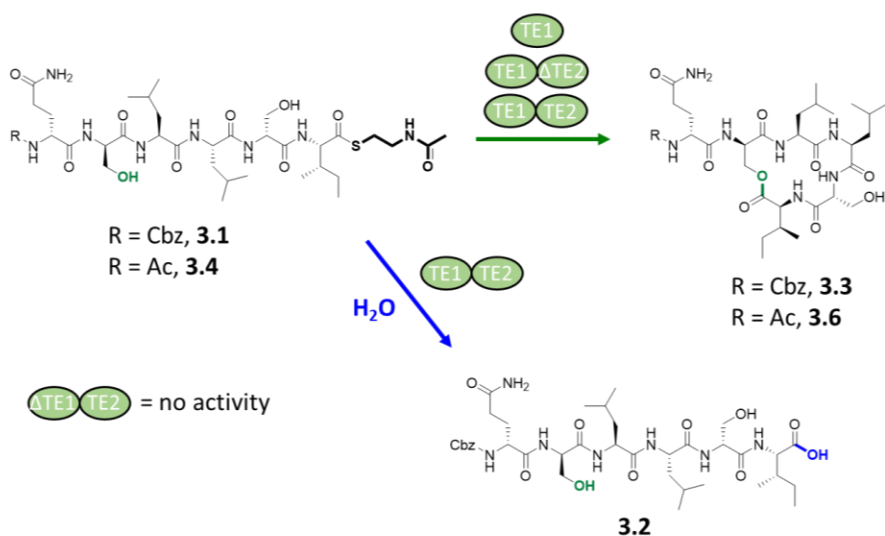
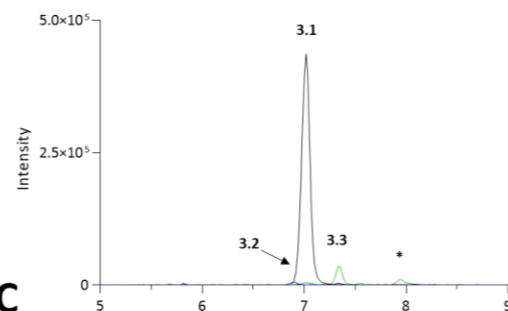
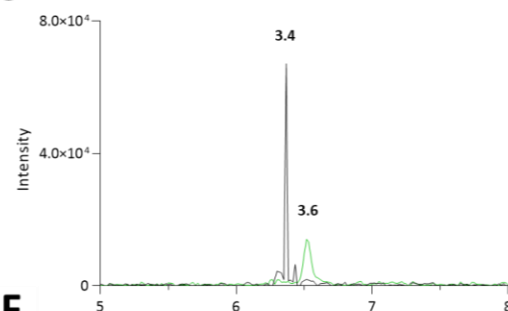
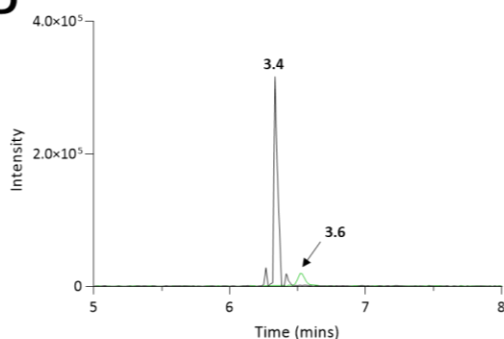
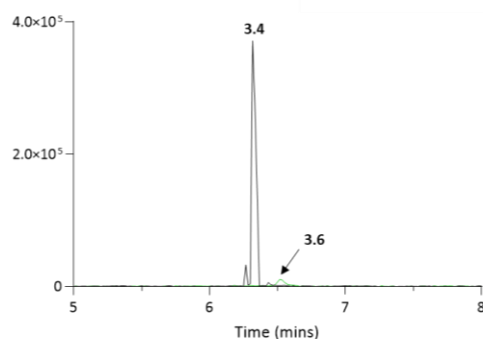
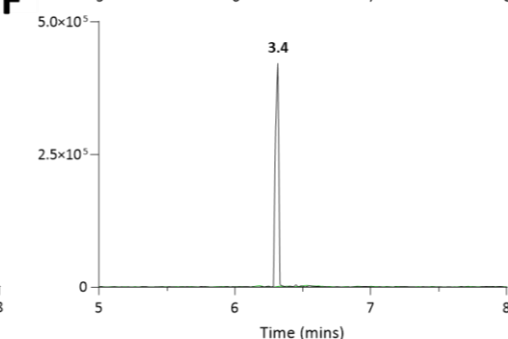
A**B****C****D****E****F**

Figure 3.3. TE catalyzed reactions and LCMS chromatograms of assays with substrates **3.1 and **3.4**.** (A) Overview of reactions catalyzed by the various Gac-TE constructs with substrates **3.1** and **3.4**. (B-F) Extracted ion chromatogram (EIC) analysis of assays containing 0.9 mM of substrate **3.1** incubated with 5 μM of TE1-TE2 (B) or 0.5 mM of substrate **3.4** incubated with TE1-TE2 (C), TE1 (D), ΔTE2 (E) and ΔTE1 (F). Assays were performed in 50 mM of potassium phosphate buffer (pH 7.4) and incubated for 22 hours. Trace in panel B is an ion extraction for $[\text{M} + \text{H}]^+$ ions with $m/z = 777$ (**3.3**), 795 (**3.2**) and 896 (**3.1**). Traces in panels C-F are ion extractions for $[\text{M} + \text{H}]^+$ ions with $m/z = 684$ (**3.6**) and 804 (**3.4**). The peak labelled by an asterisk is an injector peak.

Gac-TE1 is responsible for the macrolactonization activity

While not common, tandem TE domains have been identified in mLPD biosynthetic gene clusters of many pseudomonads.^{19–24} To date, only the Txo- and LybB-tandem TEs have been characterized *in vitro*, whereas Arf-tandem TE has been studied *in vivo*. However, the activity of TE1 and TE2 in these tandem domains differ (**Figure 3.2**).^{19–21} To investigate the activity of the individual TE domains in the Gac tandem system, the coding DNA sequence of TE1 and TE2 were independently cloned similarly as described for TE1-TE2 for expression, purification and enzymatic assays. Unfortunately, all attempts to obtain TE2 as a stand-alone enzyme were unsuccessful. This is not uncommon as similar results were seen in Txo-TE2, suggesting TE1 may be required for proper folding of TE2.¹⁹ Thus, in lieu of obtaining stand-alone TE2, inactivation mutants were generated by substituting the active site Ser of each TE domain in TE1-TE2 to Ala via site directed mutagenesis (denoted as Δ TE1 and Δ TE2).²⁶ TE1, Δ TE1 and Δ TE2 constructs were overexpressed and purified by metal affinity chromatography to high purity (> 95%, **Figure S3.4**) followed by enzymatic assays under identical conditions to TE1-TE2. After 22 hours of incubation with substrate **3.4**, production of macrocycle **3.6** is observed in the TE1 and Δ TE2 assays (**Figure 3.3D-E, S3.9 and S3.15**) but not in the Δ TE1 assays (**Figure 3.3F and S3.16**). These data suggest TE1 is responsible for the macrolactonization activity in the tandem domain.

In the LybB study, TE1 was shown to be responsible for the macrolactonization activity while TE2 is proteolytically cleaved from the tandem domain in its heterologous expression host presumably to hydrolyze misprimed carrier proteins (CP) in the NRPS complex, a function akin to type II TEs.^{21,27,28} Despite Gac-TE1 behaving similarly as LybB-TE1, Gac-TE2 is not shown to be proteolytically cleaved from the tandem domain (**Figure S3.4**). However, this could simply be due to the lack of the host specific protease, if any, for TE2 cleavage. This would need to be investigated in the host organism. In the Arf gene cluster, the fully functional tandem domain exhibits enhanced arthrofactin production *in vivo* compared to

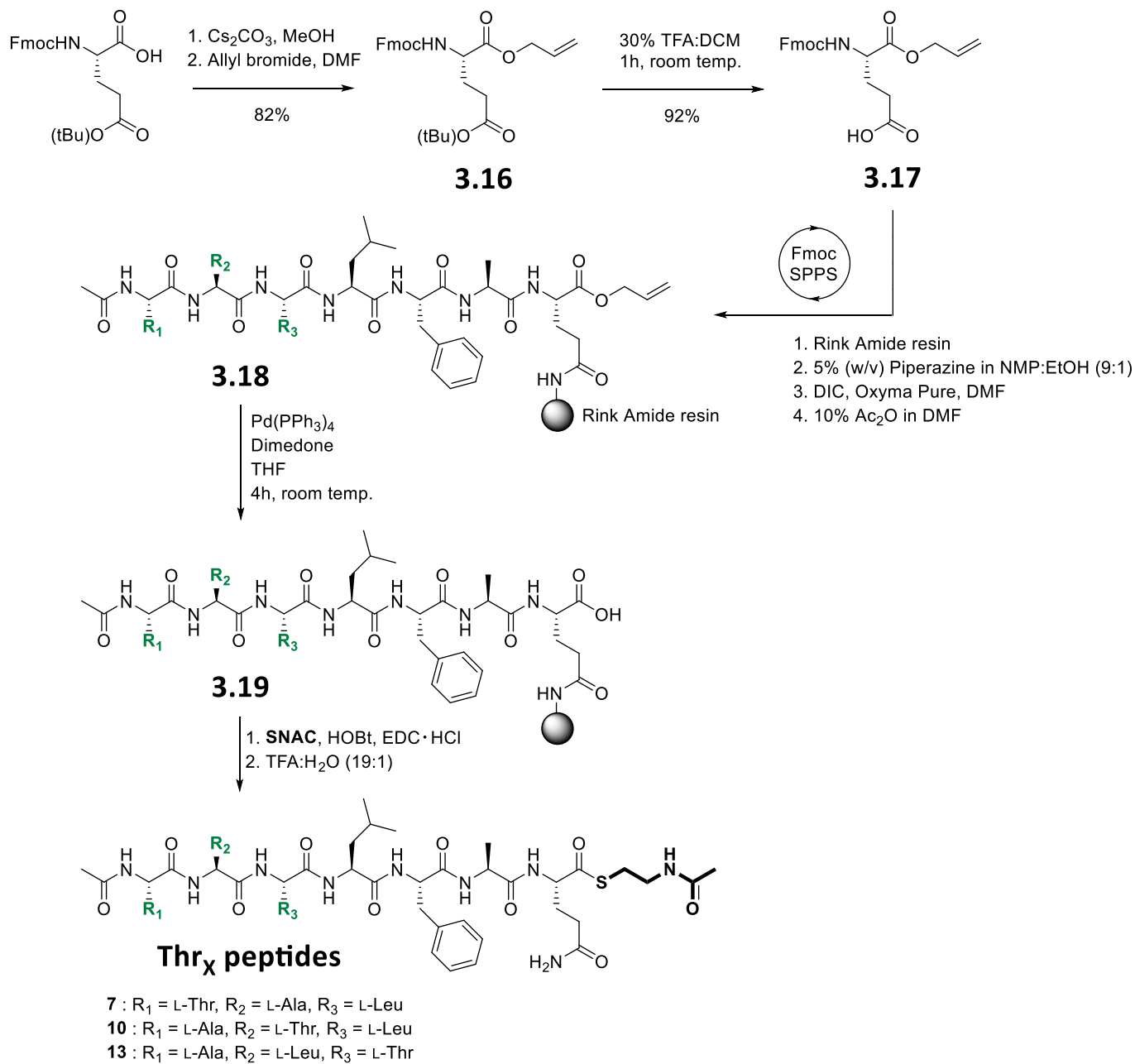
inactivated TE1 or TE2 variants.²⁰ This is consistent with the enhanced NP production observed in clusters containing tandem CP domains, such as FAS producing polyunsaturated lipids^{29,30} and the PKS producing naphthopyrone,³¹ mupirocin³² and curacin,³³ which suggests tandem domains are meant to increase production of the NP. To our surprise, TE1-TE2 only converts 2.8% of **3.4** into **3.6** whilst the stand-alone TE1 roughly converts twice as much (5.2%, **Table S3.1**). As TE2 likely requires TE1 for proper folding, it is plausible that TE2 decreases TE1-TE2 stability in the assay conditions, making it a shorter-lived catalyst compared to the stand-alone TE1, and thus converting less substrate into macrocycle. Alternatively, the short linker connecting TE1 to TE2 may facilitate discrete contacts between the two TEs, potentially constraining TE1's structural flexibility. Consequently, TE1 might struggle to adopt a catalytically active conformation within the tandem construct, thus reducing its activity. Based on these data, the precise function of TE2 remains elusive and warrants further investigation as a stand-alone enzyme.

During the 22-hour incubation of the substrate **3.1** only control, 1.3% spontaneously converted into macrocycle **3.3** (**Table S3.1** and **Figure S3.10**). This is substantially less macrocycle than was generated in the incubations with TE1, ΔTE2 and TE1-TE2 (~6-7% conversion, **Table S3.1**). After 5 days of incubation, 3.8% of **3.1** was converted into macrocycle (**Table S3.1** and **Figure S3.11**). Spontaneous macrocyclization has been observed for the ribosomally synthesized and post-translationally modified peptide (RiPP) lantibiotic, patellamide A,^{35,36} and the non-ribosomal peptide tyrocidine (Tyc)³⁴ where substrate pre-organization is believed to contribute to its spontaneity. In addition, enzyme mediated macrocyclization of Tyc is also likely driven by substrate pre-organization.^{37,38} This suggests substrate **3.1** undergoes pre-organization, effectively paying the entropic cost associated with macrocyclization³⁹ and is consistent with our proposed CS model. Because both peptides **3.1** and **3.4** have identical sequences, its likely that **3.4** also can undergo spontaneous macrocyclization through substrate pre-organization; however, this would need to be experimentally verified.

Synthesis of Thr_x (**3.7**, **3.10** and **3.13**) peptide substrates

Building upon the observed substrate tolerance for macrolactonization of substrates **3.1** and **3.4**, we aimed to further investigate TE1-TE2's substrate tolerance against a small *de novo* peptide library. This library was designed to introduce variations in amino acid composition, peptide length, internal nucleophile position and stereochemical configuration. For this, three *N*-acetylated peptides containing 7 L-configured amino acids most commonly found in macrocyclic depsipeptides were generated (**Scheme 3.1**). The internal nucleophile L-Thr was placed at either position 1 (Thr₁, **3.7**), 2 (Thr₂, **3.10**) or 3 (Thr₃, **3.13**) to generate either 7-, 6- or 5-residue macrocyclic depsipeptides. This yielded 3 peptide substrates that significantly differed from the native Gac A peptide, making these *de novo* peptides. A solid phase peptide synthesis (SPPS) strategy was employed by anchoring the unprotected C-terminal L-Glu side chain onto Rink Amide resin, allowing us to perform C-terminal thioesterification on resin.^{40,41} Thus, C-terminally protected allyl-L-Glu **3.17** was loaded onto Rink Amide resin for standard SPPS under HBTU/DIEA conditions, subsequent N-terminal Fmoc-deprotection, acetylation and C-terminal allyl deprotection to generate full length protected Thr_x peptides **3.19**. On resin thioesterification, followed by a global deprotection of L-Thr side chains and peptide cleavage from resin under TFA⁴² afforded peptides **3.7**, **3.10** and **3.13**, which were used without any purification for enzymatic assays.

Scheme 3.1. SPPS of linear SNAC Thr_x (3.7, 3.10 and 3.13) peptide substrates



Gac-TE1-TE2 is substrate tolerant for macrolactonization

To investigate the substrate tolerance of the Gac tandem domain, enzymatic assays with TE1-TE2 and crude mixtures of the Thr_x peptide substrates (10% v/v) were performed under similar conditions as the *N*-Cbz **3.1** and *N*-acetyl **3.4** assays. Singlecate data was collected by incubating enzymes at room temperature for 18 hours before analysis by LCMS. Crude mixtures of each peptide were used to eliminate the time-consuming purification step and enabling quick testing of TE1-TE2. In all incubations, hydrolysis products **3.8**, **3.11** and **3.14** were generated as substantial products (**Figure 3.4A-C**, **S3.21**, **S3.22** and **S3.23**). This is not surprising as TE loading is known to be substrate promiscuous^{43–46} and necessitates off-loading to prevent pathway stalling, often by hydrolysis. These observations are consistent with the screening hypothesis describing the evolution of natural product diversity.^{45,47} Only incubations with peptides **3.7** and **3.13** showed formation of a product with an *m/z* consistent with a 7- and 5-residue macrolactone (**Figure 3.4A** and **C**). This is intriguing as TE1-TE2 demonstrates the ability to macrolactonize using the sidechain alcohol of L-Thr instead of the naturally occurring D-Ser nucleophile, indicating a stereo tolerance for the nucleophile. This stands in stark contrast to the Txo-TE1-TE2, where changes in the stereochemical configuration of the internal nucleophile abolished macrocyclic activity.¹⁹ Additionally, non-native 7- (**3.9**) and 5-residue (**3.15**) macrocycles are formed and suggests TE1-TE2 exhibits some tolerance for the size of the generated ring. However, when L-Thr was placed at position 2 (**3.10**), formation of macrocycle product **3.12** was not observed (**Figure 3.4B**). Because these assays used crude peptide substrate mixtures, substrate **3.7** and hydrolysis product **3.10** eluted at two different retention times (labelled as **3.7'** and **3.10'**), presumably due to other compounds and or solvents present in the mixture. Assays using purified peptide substrates will need to be performed in order to confirm the results presented herein.

A total of 9 NRPS TEs have been characterized and used as macrocyclization biocatalysts in depsipeptide syntheses, generating 4-, 5-, 6-, 8-, 9-, 10- and 12-residue macrocyclic rings.^{19,21,48–54} Gac-TE1-TE2 is the only characterized domain to date shown to form a 7-residue macrolactone using a *de novo* peptide substrate that shares little similarity to the natural Gac A peptide, albeit with substantial hydrolysis product. The *in trans* acting surugamide (SurE) TE is the only other shown to macrocyclize non-native peptides; however, its functionality is restricted in the formation of macrolactams.^{55,56} Thus, TE1-TE2 shows promise as a macrolactonization biocatalyst; it does not form oligomers, as seen with the DEBS-TE,⁵⁷ is substrate tolerant for the stereochemical configuration of the alcohol nucleophile, and is ring size tolerant. With optimization via protein engineering, it can be a useful addition to our current repertoire of macrocyclization biocatalysts, as it can provide us access to 7-residue depsipeptide macrolactones.

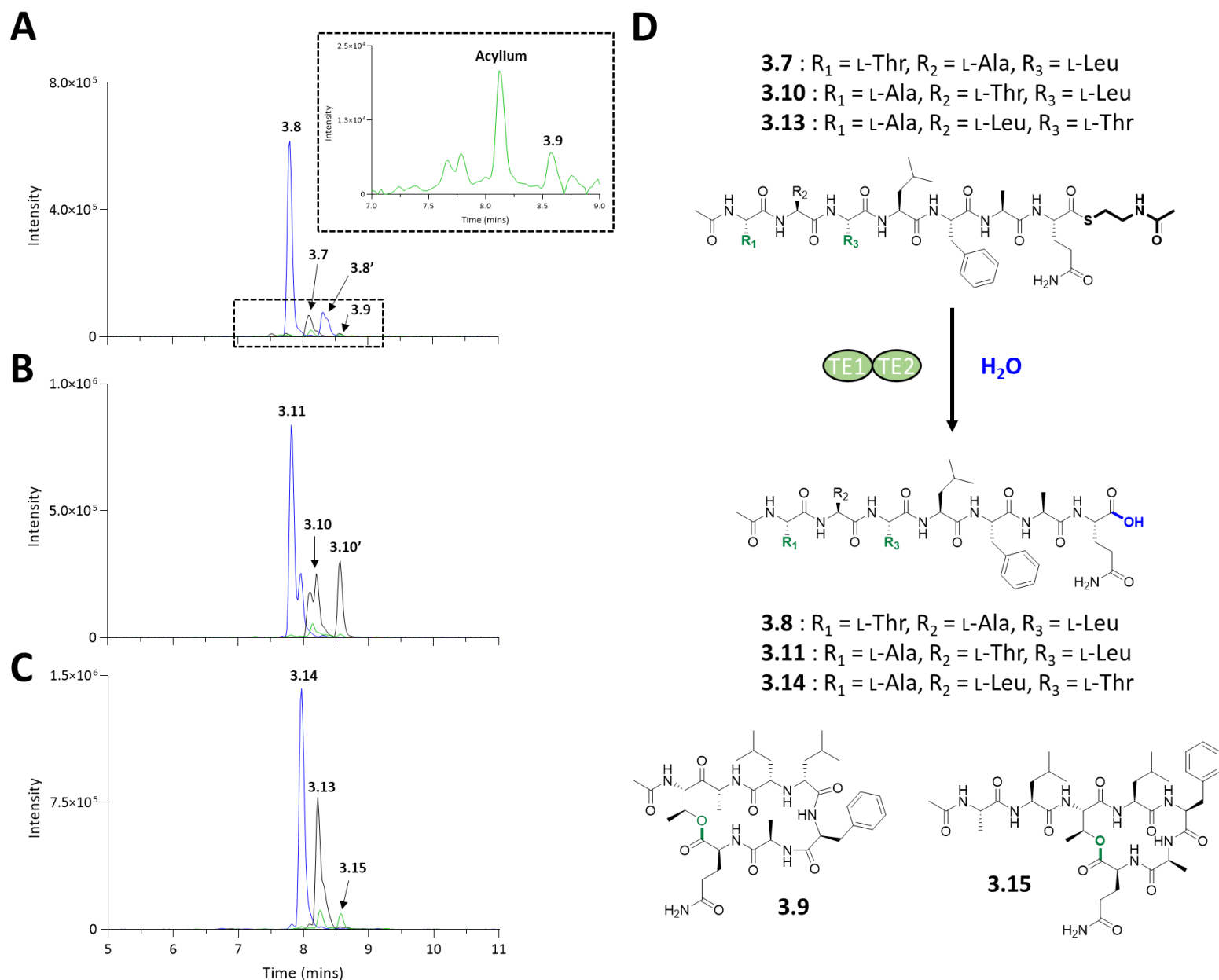


Figure 3.4. TE1-TE2 catalyzed reactions with peptide library and assay LCMS chromatograms. (A-C) Extracted ion chromatogram (EIC) analysis of TE1-TE2 (5 μM) assays with 10% v/v of crude substrates **3.7** (A), **3.10** (B) and **3.13** (C) after 22 hours of incubation in 50 mM of phosphate buffer (pH 7.4). Traces are ion extractions for $[M + H]^+$ ions with $m/z = 788$ (**3.9** and **3.15**), 806 (**3.8**, **3.11** and **3.14**) and 907 (**3.7**, **3.10** and **3.13**). Shown boxed is a zoom-in of the trace with detection of macrocyclic product **3.9** (panel A) formation compared to the acylium ion. (D) Overview of reactions catalyzed by TE1-TE2 on substrates **3.7**, **3.10** and **3.13** (peptide library).

3.4 Conclusions

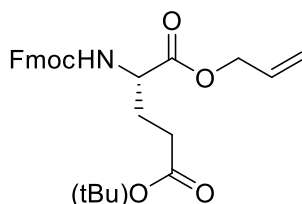
To conclude, the preliminary work presented in this chapter showcased the macrolactonization activity of the unusual Gac-tandem TE domain and its broad substrate tolerance. First, TE1-TE2 was shown to be an active macrolactonization biocatalyst, converting two N-terminally protected peptides **3.1** and **3.4** into macrocycles, suggesting TE1-TE2 is tolerant to N-terminal peptide modifications. Expression and purification of TE1 as a stand-alone enzyme and TE1-TE2 inactivation mutants revealed TE1 to be responsible for the macrocyclization activity. In addition, TE1 was shown to produce more macrocycle compared to the tandem domain. Possible sources for this difference may be that the TE1-TE2 structure is less stable or is less likely to adopt a catalytically active conformation due to an increased structural rigidity in the tandem construct. Interestingly, substrate **3.1** was shown to spontaneously form macrocycle **3.3** without enzyme catalyst. This suggests substrate pre-organization drives macrocyclization, which is consistent with our proposed CS model. Finally, TE1-TE2 was shown to macrolactonize *de novo* peptide substrates **3.7** and **3.13** into 5- and 7-residue macrolactones, respectively, using a non-native L-Thr internal nucleophile. This highlights TE1-TE2 as a promising macrocyclization biocatalyst, showcasing substrate tolerance at the N-terminus, stereo tolerance for the nucleophile and flexibility in forming rings of different sizes. While this work provides some valuable insight, further investigations into the Gac-tandem TE domains and their potential applications in biocatalysis are required.

3.5 Materials and methods

Synthetic protocol for Thr_x peptide-SNAC

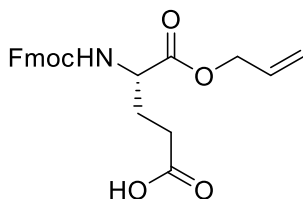
All solvents were purchased from Fisher Scientific. All reagents were purchased at the highest available purity and used without further purification. EDC, HCl, HATU, DIEA and Boc-L-Tyr(OtBu)-OH were purchased from Oakwood Chemical (Estill, SC, USA), all Fmoc protected amino acids and HOBt were purchased from GL Biochem (Shanghai, PRC), and all other chemicals were purchased from Sigma-Aldrich (Oakville, ON, Canada or Milwaukee, WI, USA). All reactions were carried out using dry solvents under a nitrogen atmosphere unless otherwise noted. Automated peptide synthesis was performed on a CEM Liberty Blue peptide synthesizer using microwave (MW) irradiation for all steps involving heating above ambient at a 0.1 mmol scale. NMR spectroscopy was performed with either a Bruker Avance II operating at 400 MHz for ¹H and 100 MHz for ¹³C spectra or a Bruker Avance III with a cryoprobe operating at 600 MHz for ¹H and 150 MHz for ¹³C spectra. LCMS analysis was carried out using a Shimadzu LC20A HPLC coupled with a Shimadzu LCMS2020 single quad mass spectrometer with an ESI source. Reverse-phase purification was carried out using an Agilent 1260 infinity and a Phenomenex Kinetex 5 μm XB-O18 100Å, 150 x 10.0 mm Semi-Preparative column. High-resolution mass spectroscopy (HRMS) was conducted on a Micromass Q-TOF I (John L. Holmes Mass Spectroscopy Facility, Ottawa, ON, Canada).

Thr_x peptide thioesters **3.7**, **3.10** and **3.13**

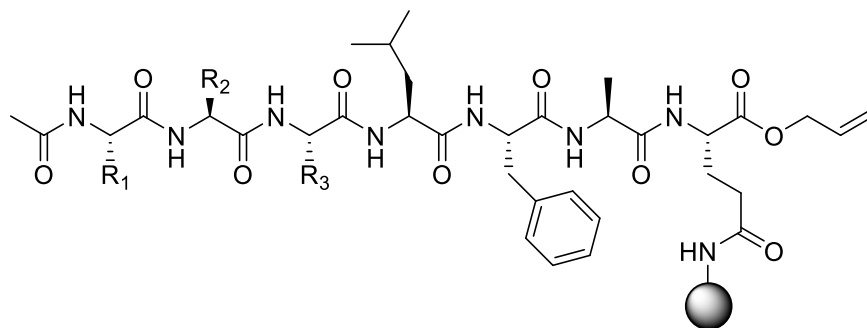


Allyl protection of L-Glu (3.16). Fmoc-L-Glu-(OtBu)-OH (1.57 g, 3.69 mmol, 1 eq.) was dissolved in MeOH (12 mL). Cs₂CO₃ (601 mg, 1.84 mmol, 0.5 eq.) was added into the solution and the reaction was stirred for

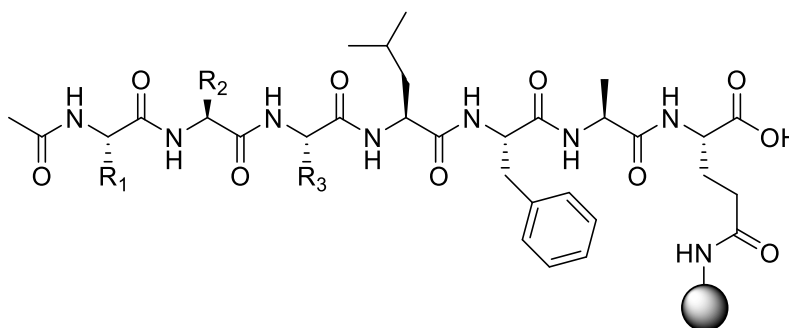
30 minutes. Subsequently, MeOH was removed in vacuo. The residue was redissolved in DMF (15.7 mL). Allyl bromide (478 μ L, 5.53 mmol, 1.5 eq.) was added dropwise and the reaction mixture was stirred overnight. EtOAc was added and the solution was extracted with a solution of 10% NaHCO₃ and brine. The organic layer was dried over MgSO₄ and concentrated. The protected L-Glu was purified by column chromatography (5-10% AcOEt in Hexane) to afford compound **3.16** as a white solid (1.4708 mg, 86%). R_f (10% MeOH in DCM) = 0.47. ¹H NMR (600 MHz, CDCl₃) δ 7.79 (dq, J = 7.6, 1.0 Hz, 2H), 7.62 (t, J = 6.5 Hz, 2H), 7.42 (tq, J = 7.5, 1.1 Hz, 2H), 7.34 (tt, J = 7.5, 1.1 Hz, 2H), 6.01 – 5.87 (m, 1H), 5.49 (d, J = 8.3 Hz, 1H), 5.41 – 5.33 (m, 1H), 5.29 (dq, J = 10.5, 1.2 Hz, 1H), 4.68 (d, J = 5.8 Hz, 2H), 4.49 – 4.34 (m, 3H), 4.25 (t, J = 7.1 Hz, 1H), 2.46 – 2.27 (m, 2H), 2.21 (dd, J = 13.7, 6.9 Hz, 1H), 2.00 (d, J = 7.5 Hz, 1H), 1.47 (s, 9H). ¹³C NMR (151 MHz, CDCl₃) δ 171.75, 156.00, 143.73, 141.33, 131.46, 127.73, 127.09, 125.11, 120.00, 119.04, 80.91, 67.10, 66.15, 53.60, 47.17, 31.49, 28.09, 27.58.



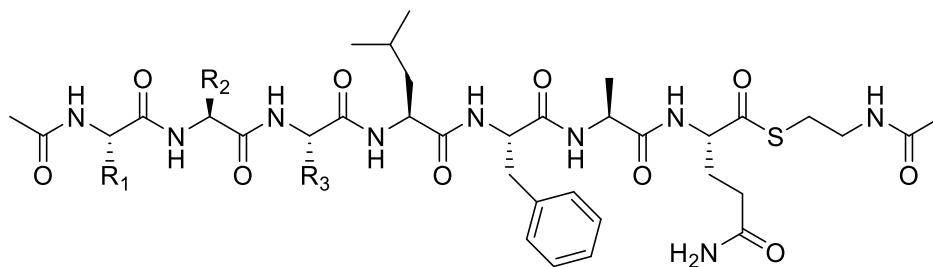
L-Glu deprotection (3.17). Compound **3.16** (1.47 mg, 3.159 mmol) was dissolved in a solution of 30% TFA:DCM (25 mL) and the reaction was stirred at room temperature for 2 hours. Usually, no starting material was detected by thin layer chromatography and the reaction was removed under reduced pressure. Diethyl ether was then added and evaporated under vacuum three times. Compound **3.17** was purified by column chromatography (5-10% MeOH in DCM) to afford a white solid (951 mg, 92 %). R_f (10% MeOH in DCM) = 0.25. ¹H NMR (300 MHz, MeOD) δ 7.83 – 7.73 (m, 2H), 7.65 (dd, J = 7.4, 3.5 Hz, 2H), 7.42 – 7.22 (m, 4H), 5.91 (ddt, J = 16.3, 10.8, 5.6 Hz, 1H), 5.39 – 5.14 (m, 2H), 4.60 (dt, J = 5.6, 1.5 Hz, 2H), 4.32 (t, J = 6.9 Hz, 2H), 4.28 – 4.15 (m, 2H), 2.37 (t, J = 7.4 Hz, 2H), 2.15 (td, J = 13.4, 7.4 Hz, 1H), 1.91 (ddd, J = 16.6, 14.1, 7.2 Hz, 1H).



Protected linear peptide (3.18). Automated peptide synthesis was performed on a CEM Liberty Blue peptide synthesizer at a 0.1 mmol scale. Rink amide resin (167 mg) was placed in the reaction vessel followed by a 5-minute swell in DMF. Couplings were performed with Fmoc-L-amino acid (0.2 M in DMF), DIC (0.2 M in DMF) and Oxyma Pure (1 M in DMF) at 95 °C for 1.5 minutes and Fmoc deprotection were performed with 5% piperazine solution in NMP:EtOH (90:10). N-terminal acetylation was done with 10% acetic anhydride in DMF at 50 °C for 1 hour. Once complete, the resin was transferred to a fritted vessel and flushed with DCM (5 mL x 3) and dried under reduced pressure to afford crude linear peptide loaded resin. The compound was confirmed by LCMS and used without further characterization.



Allyl deprotection on the linear peptide (3.19). Compound **3.18** (0.1 mmol) was swelled in 3 mL of THF for 20 minutes in the fritted vessel peptide. Pd(PPh₃)₄ (11.55 mg, 0.01 mmol, 0.1 eq.) and dimedone (42.0 mg, 0.3 mmol, 3 eq.) were added and gently stirred for 3.5 hours away from light. The on-resin peptide was then washed with DCM x 3, DMF x 3 and finally with DCM x 3. The compound was confirmed by LCMS and used without further characterization.



Thr_x peptide-SNAC (3.7, 3.10 and 3.13). Compound **3.19** (0.1 mmol), HOBt (14.2 mg, 0.105 mmol, 1.05 eq.) and EDC·HCl (20.1 mg, 0.105 mmol, 1.05 eq.) were added to a fritted peptide vessel with DCM (0.1 M) and stirred for 20 minutes. *N*-acetylcysteamine (14.3 mg, 0.12 mmol, 1.2 eq.) was dissolved in DCM (1 mL) and added to the fritted vessel containing the peptide and gently stirred for 2 hours. The compound was confirmed by LCMS and used without further characterization. The crude thioester peptide on resin was dissolved in 2 mL TFA:H₂O (19:1) and the solution was allowed to swell at room temperature for 2 hours. The fritted peptide vessel was flushed in a flask and the solvent was removed under reduced pressure with co-evaporation using ether. The substrates were confirmed by LCMS and used without purification.

Cloning, expression and purification

Cloning of Gac-TE1, -TE2 and -tandem (TE1-TE2) domain. A single *Pseudomonas fluorescens* Pf0-1 colony, strain kindly provided to us by Dr. Mark Silby (University of Massachusetts Dartmouth), was inoculated into LB media supplemented with 100 µg/mL of ampicillin and grown overnight at 37 °C with 200 rpm shaking. Genomic DNA (gDNA) was extracted from dense overnight culture using the Monarch Genomic DNA Purification Kit (#T3010, NEB) as per the manufacturer's recommended protocol. The coding DNA sequence of the tandem TE domain (TE1-TE2) and the individual TE domains (TE1 and TE2) from the gacamide A (Gac) biosynthetic gene cluster (CP000094.2 range 2,532,969 - 2,534,285) was amplified from the gDNA by PCR using the primers shown in **Table 3.1**. Each reaction contained the following: the correct

forward and reverse primer pair (both 0.5 μ M), Q5 reaction buffer (1X), equimolar mixture of dNTPs (200 μ M), gDNA (1 ng), Q5 High-Fidelity DNA Polymerase (0.5 units, #M0491 NEB). Negative controls were prepared identically except without the addition of polymerase. PCR was performed as per the manufacturer's recommended conditions, with an annealing temperature of 69 °C. Amplification of DNA by PCR was confirmed via gel electrophoresis and the bands corresponding to each TE construct were extracted using the Monarch DNA Gel Extraction Kit (#T1020, NEB). Extracted constructs were cloned into NdeI and EcoRI linearized pET-28a(+) expression vector, generating pJBH14, pJBH15 and pJBH53 for the Gac-TE1, Gac-TE2 and Gac-tandem domains respectively. All constructs were confirmed by Sanger sequencing (G  nome Qu  bec).

Generation of TE1 and TE2 inactivated (Δ) variants. Site-directed mutagenesis²⁶ was performed using the primers shown in **Table 3.1**. Every reaction was set in duplicate and contained the following: the correct forward and reverse primer pair (both 0.5 μ M), Q5 Reaction Buffer (1X), dNTPs (200 μ M), pJBH55 template (1 ng), Q5 High-Fidelity DNA Polymerase (0.5 units, #M0491 NEB). Negative controls were prepared identically except without the addition of polymerase. Full plasmid PCR was performed for every reaction as per the manufacturer's recommended conditions, with an annealing temperature of 74°C for 50 seconds. Amplification of DNA by PCR was confirmed via gel electrophoresis and every reaction purified using the E.Z.N.A. Cycle-Pure Kit (#D6492, Omega Bio-tek) as described in their product manual followed by an overnight DpnI digestion (10 units, #R0176 NEB) at 37°C. Digested products were then purified once more using the E.Z.N.A. Cycle-Pure Kit (#D6492, Omega Bio-tek) as described in their product manual. Eluted products were then phosphorylated for 1 hour at 37°C and ligated for 2 hours at room temperature, generating constructs pJBH53-1 (Δ TE1) and -2 (Δ TE2), before their transformation into *E. coli* XL1-blue chemically competent cells. Each phosphorylation reaction contained the following: the eluted product, T4 DNA Ligase Buffer (1X, #B0202 NEB), T4 Polynucleotide Kinase (2.5 units, #M0201 NEB). T4 DNA Ligase

(5 units, #M0202 NEB) was added to the same phosphorylation reactions prior to the 2-hour ligation incubation. A negative control was also prepared and treated similarly as described above except without the addition of T4 PNK and T4 DNA ligase to account for background colonies after transformation.

Table 3.1. Primers used in this study. *P. fluorescens* Pf0-1's complete genome (CP000094.2) at positions 2,532,969-2,534,285 was used for primer design for cloning of Gac-TE1, -TE2 and -tandem domains. Mutagenesis primers for TE inactivation were designed to substitute the catalytic Ser of one TE in the tandem domain with Ala. Primer pairs were designed accordingly with a T_m of ~65°C and ordered from Integrated DNA Technologies with standard desalting. All primers are shown 5' to 3' with the forward primer shown on top and the reverse shown on the bottom in each row. Shown underlined is the mutagenic codon used for TE inactivation.

Plasmid	TE construct generated	Oligos (Forward/Reverse)
pJBH14	Gac-TE1	TCGTACATATGCAGGATGACGGTCTGGTCGTGGTGCTGCTGACGATCC ATTCGGAATTCTTACAATGCCTGGCTGAGCGCCT
pJBH15	Gac-TE2	TCGTACATATGTATCAGCCGCTGCTGACGATCC ATTCGGAATTCTTATCAGGACAACACTTGACCCGC
pJBH53	Gac-tandem (TE1-TE2)	TCGTACATATGCAGGATGACGGTCTGGTCGTGGTGCTGCTGACGATCC ATTCGGAATTCTTATCAGGACAACACTTGACCCGC
pJBH53-1	ΔTE1	<u>GCATTTGGCGGGGTGCTGGCCTAC</u> CCAGCCAGCAAGGCGATAAGGCCCC
pJBH53-2	ΔTE2	<u>GCATTCGGTGGTTGGGTGGCTTACGCC</u> ATGCCCCGATCAGGTGCACCGGTCCT

Expression. Constructs were transformed into *E. coli* BL21(DE3) chemically competent cells for protein expression. A flask containing 400 mL of LB media supplemented with 100 µg/mL of ampicillin were inoculated with 0.5 % v/v of dense transformed overnight cultures. Cultures were grown at 37 °C with 200 rpm shaking to an OD₆₀₀ of 0.6-0.8. Protein expression was then induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cultures were then incubated at 16 °C with 200 rpm shaking for 18 hours.

Protein purification. Entire procedure was performed at 4 °C and all plasticware used were pre-chilled. Cells were harvested by centrifugation at 4,000 g and resuspended in 10 mL of lysis buffer (100 mM sodium phosphate, 300 mM NaCl, 10% w/v glycerol, 1 mg/mL lysozyme, 1 µg/mL of pepstatin A and 1 µg/mL leupeptin, pH 7.4). Cells were lysed by sonication on ice and cell debris removed by centrifugation at 12,000 g. Resulting lysate was incubated with 800 µL of Ni-NTA resin slurry (#30210, QIAGEN) on a rocking platform shaker for 1.5 hours. Lysate was then loaded into a gravity flow column and resin allowed to settle by gravity. Purification buffer (100 mM Tris, 300 mM NaCl, pH 7.4) was used at the indicated imidazole concentrations for resin washing and protein elution. Resin was first washed once with 0 mM imidazole and once more with 20 mM imidazole. Protein was eluted twice with 100 mM imidazole and then twice more with 250 mM imidazole. Protein collected during elution was concentrated and the buffer exchanged for dialysis buffer (100 mM Tris, 300 mM NaCl, pH 7.4) by centrifugation at 4,000 g using a Vivaspinn-20 10 kDa MWCO centrifugal concentrator tube (#VS2002, Sartorius). An equal volume of 60% w/v glycerol was added to protein concentrate, flash frozen in liquid nitrogen then stored at -80 °C. Protein concentration was determined by Bradford assay (#5000006, Bio-Rad). All attempts to purify TE2 were unsuccessful.

Biochemical assays

Product distribution assays with purified substrates. Assays were carried out in eppendorf tubes in a 70 µL final reaction volume (50 mM potassium phosphate buffer pH 7.4, 10% v/v DMSO, 0.9 mM substrate **3.1** or 0.5 mM of substrate **3.4**, 5 µM enzyme) in triplicate. A substrate and negative control were prepared similarly as described above with the addition of water and heat denatured enzyme (5 µM) in place of functional enzyme respectively. Reactions were left to incubate on the benchtop at room temperature for turnover. Aliquots at time points 0, 3 and 22 hours were taken and quenched in 3 volumes of LCMS Optima grade acetonitrile (MeCN) supplemented with 0.05% v/v LCMS Optima grade formic acid (FA).

Substrate **3.4** only control had an additional aliquot quenched after 120 hours of incubation. Quenched aliquots were left on the benchtop for 15 minutes, then centrifuged at max speed for 10 minutes at room temperature before analysis by LCMS (Shimadzu Prominence UFLC system with SPD-20A Prominence UV/VIS Detector, LCMS-2020 single quadrupole MS). Chromatographic separations were done using an Agilent Eclipse Plus C18 (3.5-micron, 150×4.6 mm) reverse phase column in ESI+ mode with a flow rate of 1 mL/min. The mobile phase was composed of LCMS Optima grade H₂O (A) and MeCN (B) both supplemented with 0.05% v/v FA. Samples were monitored at 210 and 254 nm with an elution gradient of 10% B to 99% B over 15 minutes.

Preliminary product distribution assays with crude Thr_x peptide substrates. Preliminary assays were performed on crude mixtures of Thr_x peptide (**3.7**, **3.10** and **3.13**) substrates. Assays were carried out in eppendorf tubes in a 20 µL final reaction volume (50 mM potassium phosphate buffer pH 7.4, 10% v/v DMSO, 10% v/v crude substrate mixture, 5 µM enzyme). A crude substrate only control was prepared similarly as described above. Only a single replicate was made, and reactions were left to incubate on the benchtop at room temperature for turnover. After 18 hours, reactions were quenched in 3 volumes of LCMS Optima grade acetonitrile (MeCN) supplemented with 0.05% v/v LCMS Optima grade formic acid (FA). Quenched reactions were left on the benchtop for 15 minutes, then centrifuged at max speed for 10 minutes at room temperature before analysis by LCMS (Shimadzu Prominence UFLC system with SPD-20A Prominence UV/VIS Detector, LCMS-2020 single quadrupole MS). Chromatographic separations were done using an Agilent Eclipse Plus C18 (3.5-micron, 150×4.6 mm) reverse phase column in ESI+ mode with a flow rate of 1 mL/min. The mobile phase was composed of LCMS Optima grade H₂O (A) and MeCN (B) both supplemented with 0.05% v/v FA. Samples were monitored at 210 and 254 nm with an elution gradient of 10% B to 99% B over 15 minutes.

Selected NMR, mass spectra and UV traces



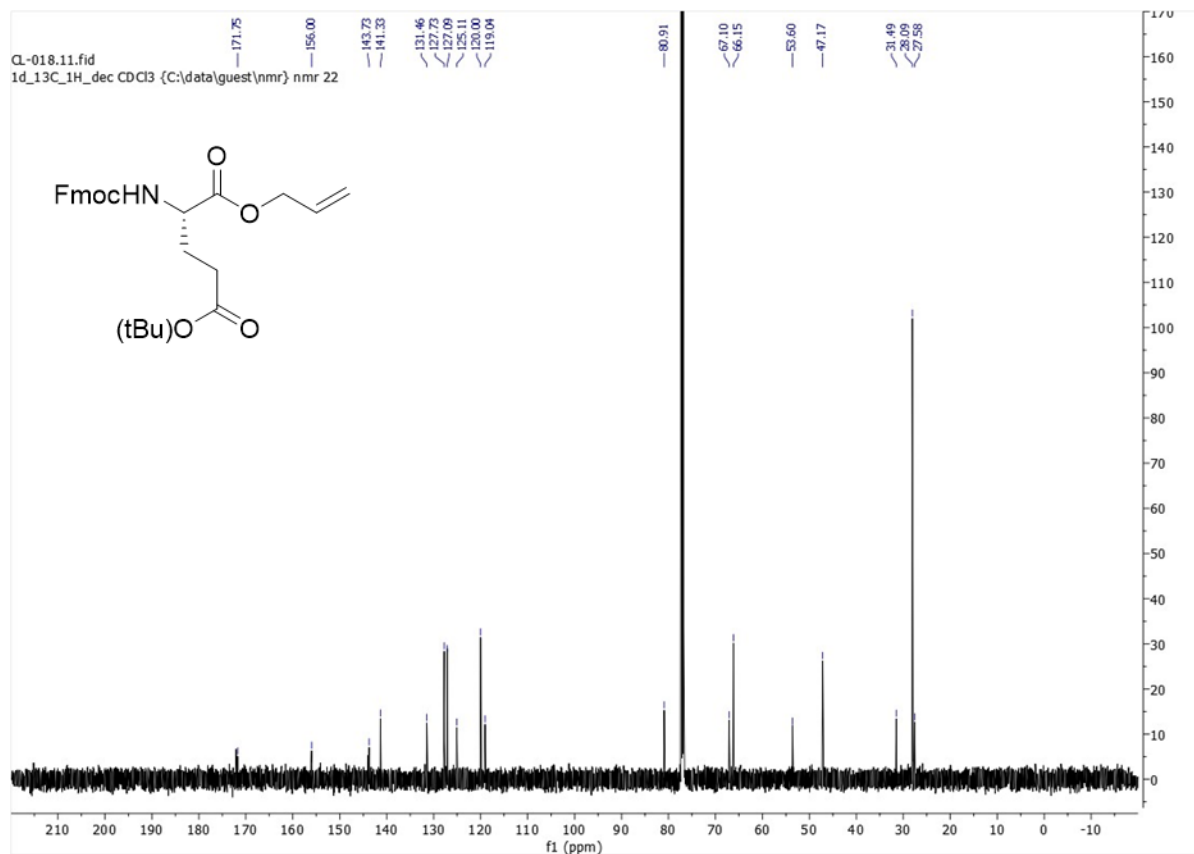


Figure S3.2. ¹³C NMR spectrum (150 MHz) of compound 3.16 in CDCl₃.

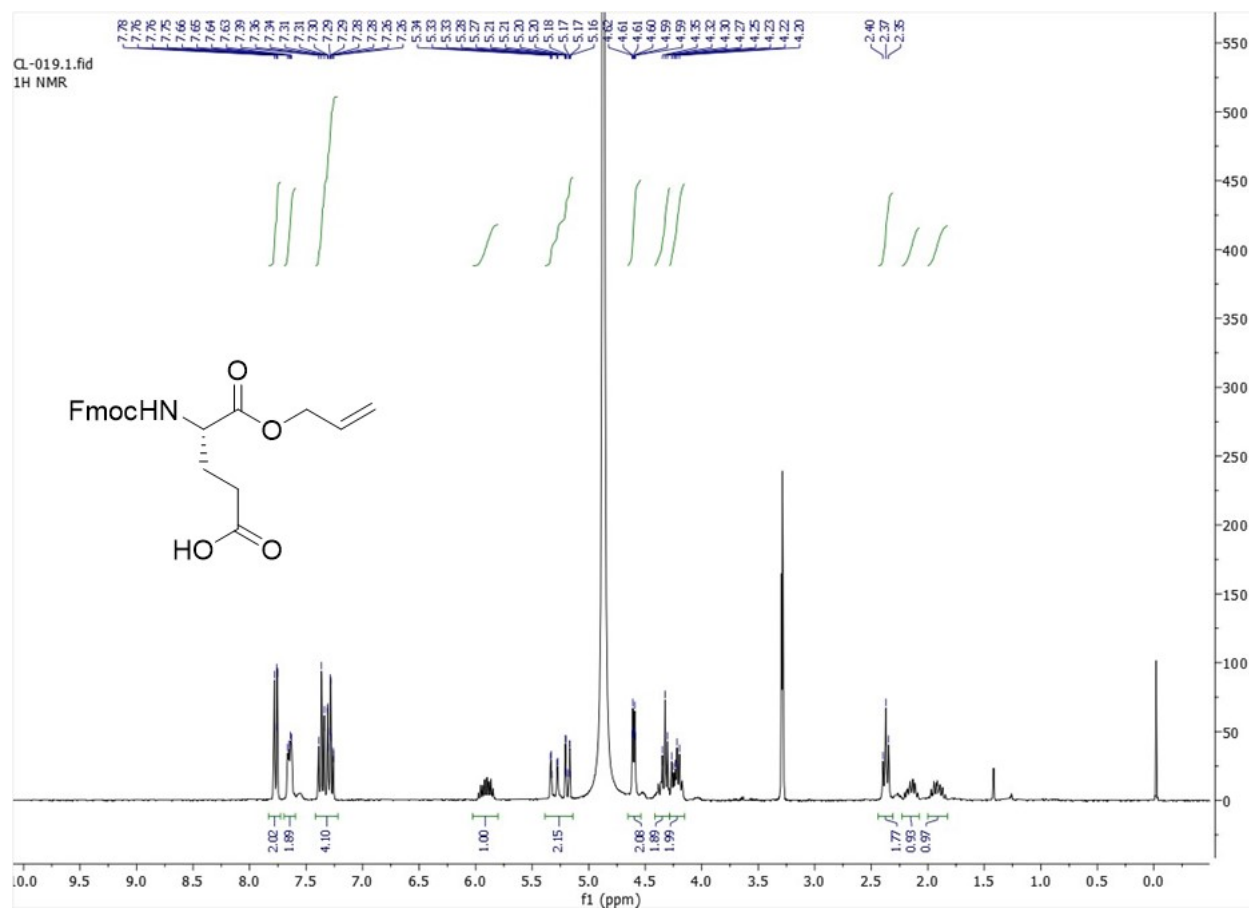


Figure S3.3. ^1H NMR spectrum (600 MHz) of compound 3.17 in MeOD.

Protein purities, yields and sequences

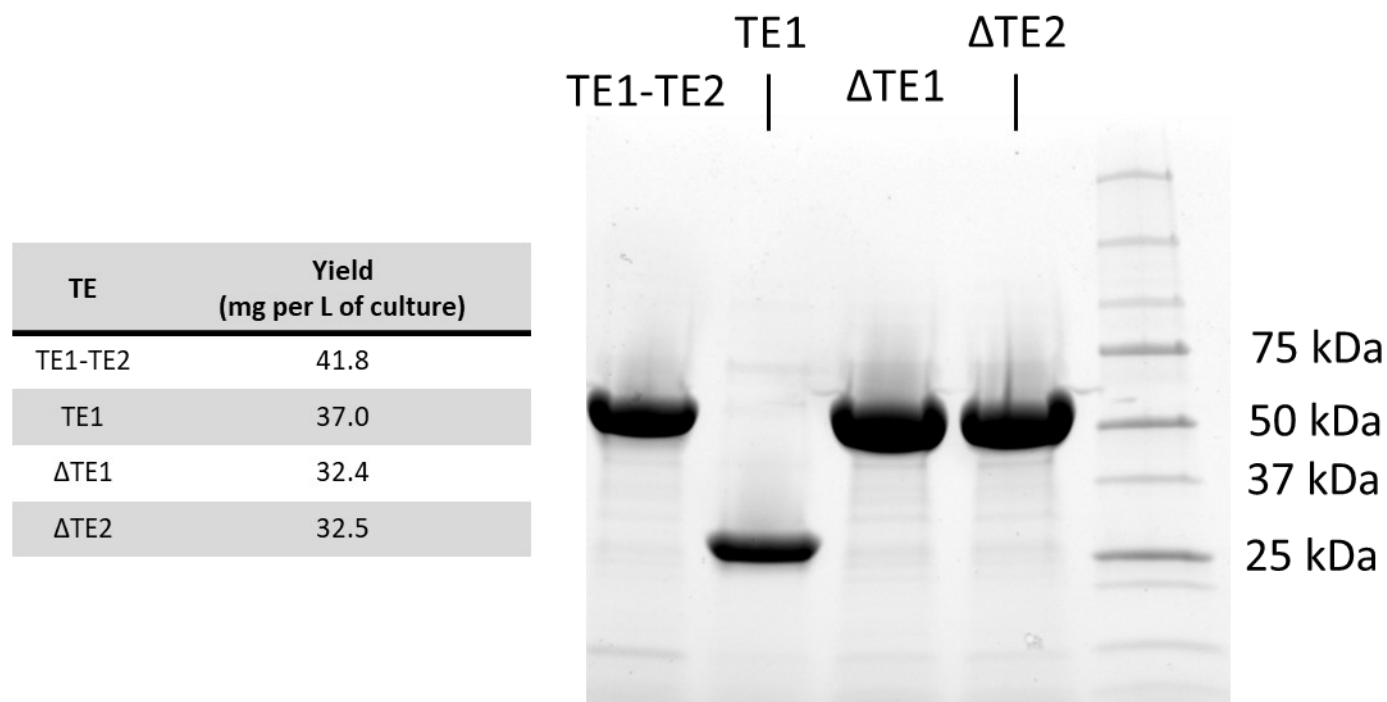


Figure S3.4. SDS-PAGE analysis of purified recombinant Gac-TE constructs. Expected molecular weights for TE1 is approximately 33 kDa while TE1-TE2, Δ TE1 and Δ TE2 are 64 kDa. The rightmost lane of gel corresponds to Precision Plus Protein Unstained Protein Standards (#1610363, Bio-Rad), all other lanes contain protein samples and are labeled on top. Protein yields are shown in the table to the left.

>Gac-TE1-TE2

MGSS**HHHHHH**SSGLVPRGSHMQDDGLVVVRAGGAGTPLFLMHEFSGMDTYFPALGLHIEGDFPIYGLTGI
 PFGQAPLRTMECLAARMVEIIRSVQPQGPYRLAGW**S**FGGVLAYEVAMQLLGMDETVAFVGLI**D**SFVPRLT
 DQGKARWSDEHAHKRQLLQHCTLYWAARGDAGQSALAHLALLQAELEQIDFQPLLQRCREEQLLHPDMAA
 AAPQDLWHYLDREVAHGHALAHYRINPLGLTLHLFRAQEHPTLSRRHPTLGWSDEIASTLLHCIDVPGD
HQSMMQAPHVHVLGQALSQALQA**STAIESRAYQF**LLTIQTGHAGHAPIFCVPGAGDSVTGFIHLTEALGP
 EWPIYGLQPRGLDGATVPHSRVESAAATFYLRAL**E**QLYPQGPVHLIGH**S**FGGWVAYAMAAQLQAAGREVAS
 LTLI**D**SESPGGNSVVGKPYTTTAALLRLIESLHLSSGKSLGIDALAFEADESTQMRLLDGMVHAGVLS
 ERASSDAMQGPVRTFATALRTVYRPAAYTGVPRLALVDDPTLDASGNQREQAAMLEGWQREARDLAVWY
 GPGN**F**FTILKAPNVFSLAAWWHDGLTVAAGQVLS-

>Gac-TE1

MGSS**HHHHHH**SSGLVPRGSHMQDDGLVVVRAGGAGTPLFLMHEFSGMDTYFPALGLHIEGDFPIYGLTGI
 PFGQAPLRTMECLAARMVEIIRSVQPQGPYRLAGW**S**FGGVLAYEVAMQLLGMDETVAFVGLI**D**SFVPRLT
 DQGKARWSDEHAHKRQLLQHCTLYWAARGDAGQSALAHLALLQAELEQIDFQPLLQRCREEQLLHPDMAA
 AAPQDLWHYLDREVAHGHALAHYRINPLGLTLHLFRAQEHPTLSRRHPTLGWSDEIASTLLHCIDVPGD
HQSMMQAPHVHVLGQALSQAL-

>Gac-ΔTE1-TE2 (ΔTE1)

MGSS**HHHHHH**SSGLVPRGSHMQDDGLVVVRAGGAGTPLFLMHEFSGMDTYFPALGLHIEGDFPIYGLTGI
 PFGQAPLRTMECLAARMVEIIRSVQPQGPYRLAGW**A**FGGVLAYEVAMQLLGMDETVAFVGLI**D**SFVPRLT
 DQGKARWSDEHAHKRQLLQHCTLYWAARGDAGQSALAHLALLQAELEQIDFQPLLQRCREEQLLHPDMAA
 AAPQDLWHYLDREVAHGHALAHYRINPLGLTLHLFRAQEHPTLSRRHPTLGWSDEIASTLLHCIDVPGD
HQSMMQAPHVHVLGQALSQALQA**STAIESRAYQF**LLTIQTGHAGHAPIFCVPGAGDSVTGFIHLTEALGP
 EWPIYGLQPRGLDGATVPHSRVESAAATFYLRAL**E**QLYPQGPVHLIGH**S**FGGWVAYAMAAQLQAAGREVAS
 LTLI**D**SESPGGNSVVGKPYTTTAALLRLIESLHLSSGKSLGIDALAFEADESTQMRLLDGMVHAGVLS
 ERASSDAMQGPVRTFATALRTVYRPAAYTGVPRLALVDDPTLDASGNQREQAAMLEGWQREARDLAVWY
 GPGN**F**FTILKAPNVFSLAAWWHDGLTVAAGQVLS-

>Gac-TE1-ΔTE2 (ΔTE2)

MGSS**HHHHHH**SSGLVPRGSHMQDDGLVVVRAGGAGTPLFLMHEFSGMDTYFPALGLHIEGDFPIYGLTGI
 PFGQAPLRTMECLAARMVEIIRSVQPQGPYRLAGW**S**FGGVLAYEVAMQLLGMDETVAFVGLI**D**SFVPRLT
 DQGKARWSDEHAHKRQLLQHCTLYWAARGDAGQSALAHLALLQAELEQIDFQPLLQRCREEQLLHPDMAA
 AAPQDLWHYLDREVAHGHALAHYRINPLGLTLHLFRAQEHPTLSRRHPTLGWSDEIASTLLHCIDVPGD
HQSMMQAPHVHVLGQALSQALQA**STAIESRAYQF**LLTIQTGHAGHAPIFCVPGAGDSVTGFIHLTEALGP
 EWPIYGLQPRGLDGATVPHSRVESAAATFYLRAL**E**QLYPQGPVHLIGH**A**FGGWVAYAMAAQLQAAGREVAS
 LTLI**D**SESPGGNSVVGKPYTTTAALLRLIESLHLSSGKSLGIDALAFEADESTQMRLLDGMVHAGVLS
 ERASSDAMQGPVRTFATALRTVYRPAAYTGVPRLALVDDPTLDASGNQREQAAMLEGWQREARDLAVWY
 GPGN**F**FTILKAPNVFSLAAWWHDGLTVAAGQVLS-

Figure S3.5 Gac-TE translated protein sequences. Shown in bold is the N-terminal 6xHis-tag used for purification.

Predicted catalytic triad for TE1 is highlighted in yellow and in green for TE2. The linker joining both TE1 to TE2 is highlighted in magenta. For the inactivated mutants, the catalytic Ser was mutated to Ala and is shown underlined.

Gac-TE1-TE2 predicted structure

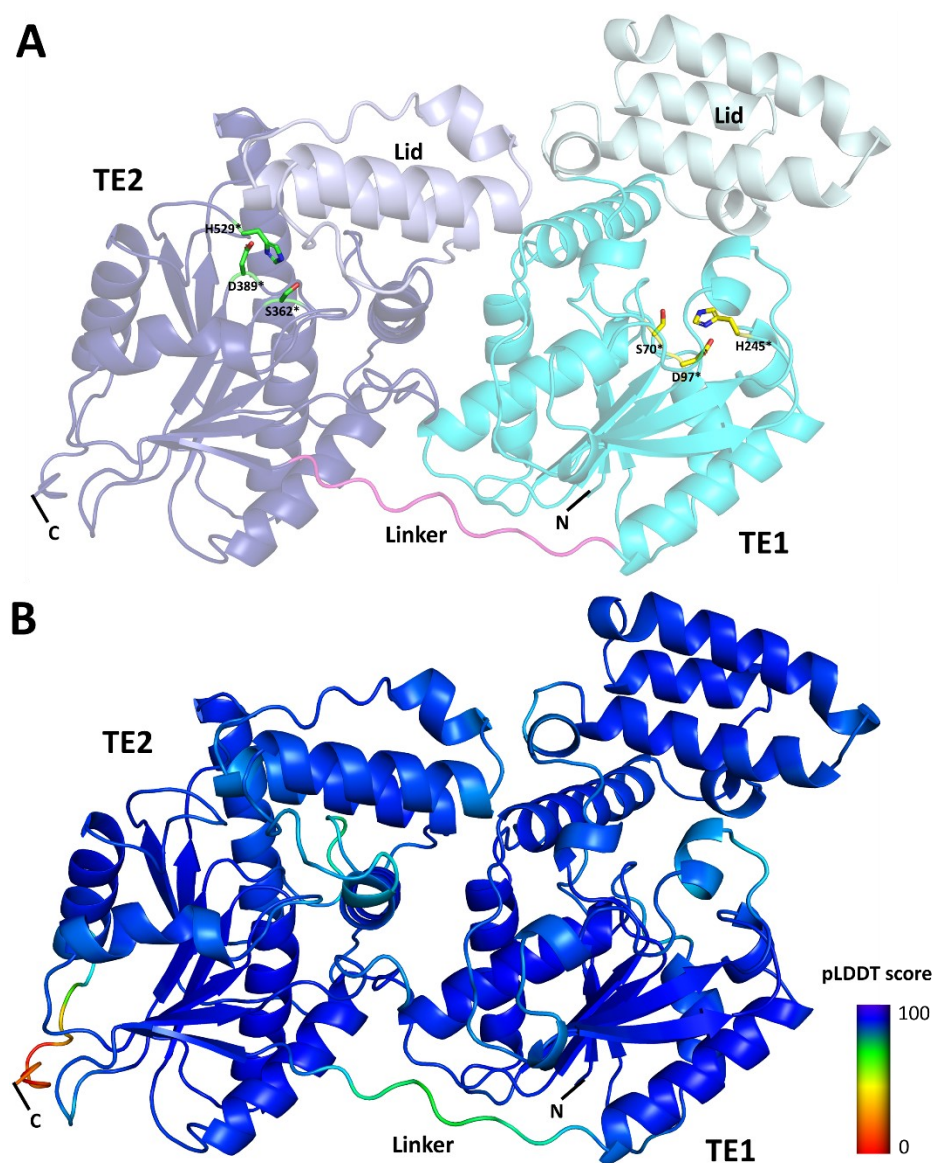


Figure S3.6. Gac-TE1-TE2 AlphaFold2 predicted structure. (A) Gac-TE1-TE2 contains two full length TE domains linked together by a 9 amino acid linker (magenta). Both TE domains are shown to adopt an α/β -hydrolase fold commonly adopted by TE domains, shown as a cartoon. Two distinct active sites are predicted in each TE housing their own catalytic triads (yellow and green sticks) and lids, suggesting they can function independently from each other. Prediction was generated from the AlphaFold2 (v2.3.2) Google Collab webserver using standard settings.⁵⁸ (B) Predicted structure with residues coloured according to AlphaFold2's per-residue confidence metric pLDDT.

***N*-Cbz peptide (**3.1**) assay LCMS chromatograms and mass spectra**

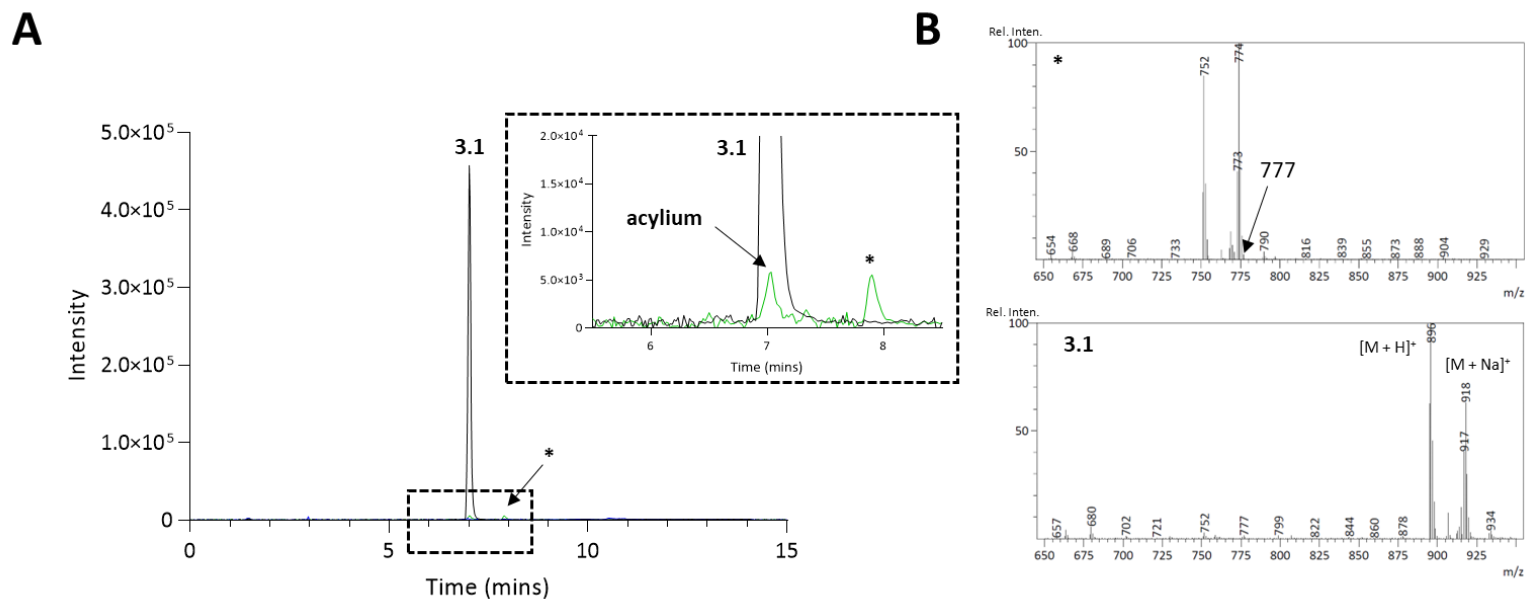


Figure S3.7. TE1-TE2 0-hour incubation with substrate **3.1.** (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 0 hours of incubation with 0.9 mM of **3.1** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 896$ (**3.1**). Shown boxed is a zoom-in of the trace showcasing the injector peak, labelled with an asterisk, relative to the acylium ion. (B) Mass spectra for substrate **3.1** and injector peak (*), mass adducts are identified the spectra for **3.1** only. Thermally denatured controls all showed similar results for all time points. All enzymes and substrate only control showed similar results after 0 hours of incubation under identical conditions.

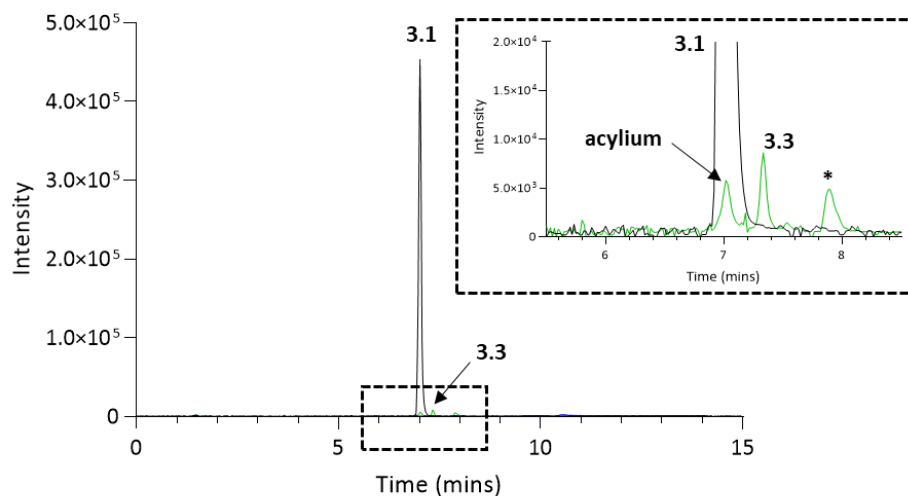
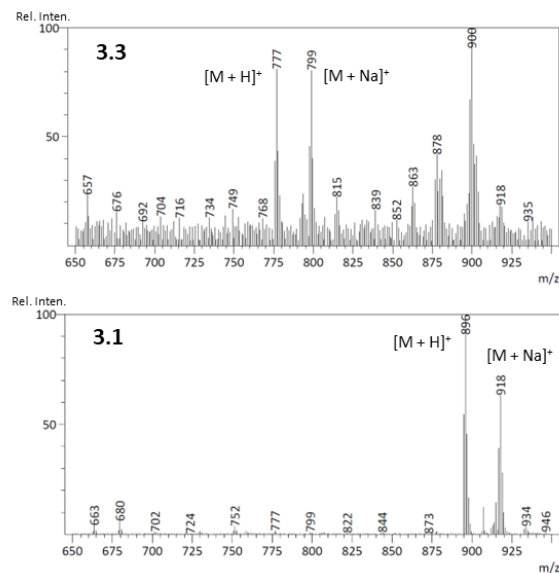
A**B**

Figure S3.8. TE1-TE2 3-hour incubation with substrate 3.1. (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 3 hours of incubation with 0.9 mM of **3.1** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 777$ (**3.3**) and 896 (**3.1**). Shown boxed is a zoom-in of the trace with detection of macrocytic product **3.3** formation compared to the acylium ion. Peak labelled with an asterisk is an injector peak seen in the 0-hour incubation (**Figure S3.7**). (B) Mass spectra for substrate **3.1** and macrocytic product **3.3**, mass adducts are identified in each spectra. Δ TE2 showed similar results after 22 hours of incubation under identical conditions.

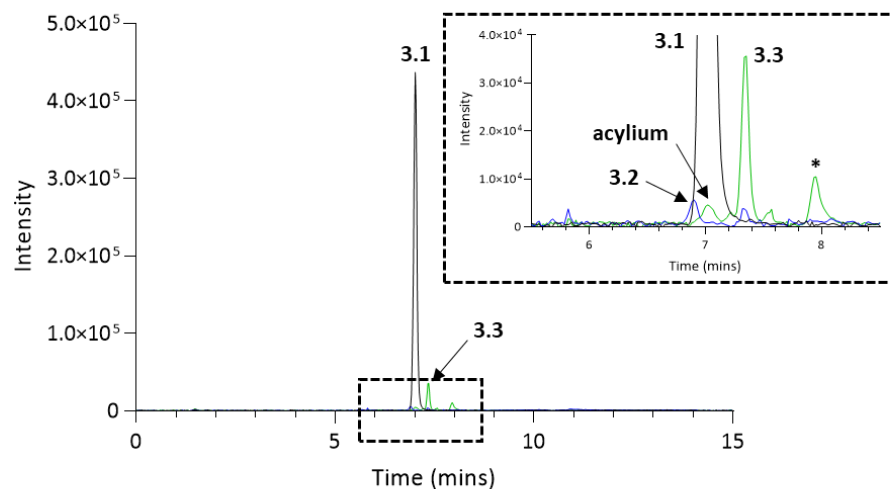
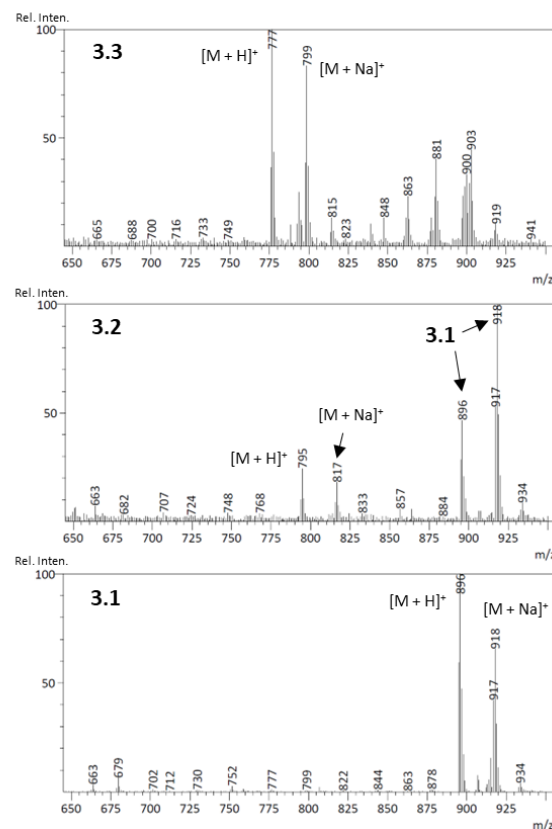
A**B**

Figure S3.9. TE1-TE2 22-hour incubation with substrate 3.1. (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 22 hours of incubation with 0.9 mM of **3.1** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 777$ (**3.3**), 795 (**3.2**) and 896 (**3.1**). Shown boxed is a zoom-in of the trace with detection of hydrolysis product **3.2** and macrocyclic product **3.3** formation compared to the acylium ion. Peak labelled with an asterisk is an injector peak seen in the 0-hour incubation (**Figure S3.7**). (B) Mass spectra for substrate **3.1**, hydrolysis product **3.2** and macrocyclic product **3.3**, mass adducts are identified in each spectra. Δ TE2 showed similar results after 22 hours of incubation under identical conditions.

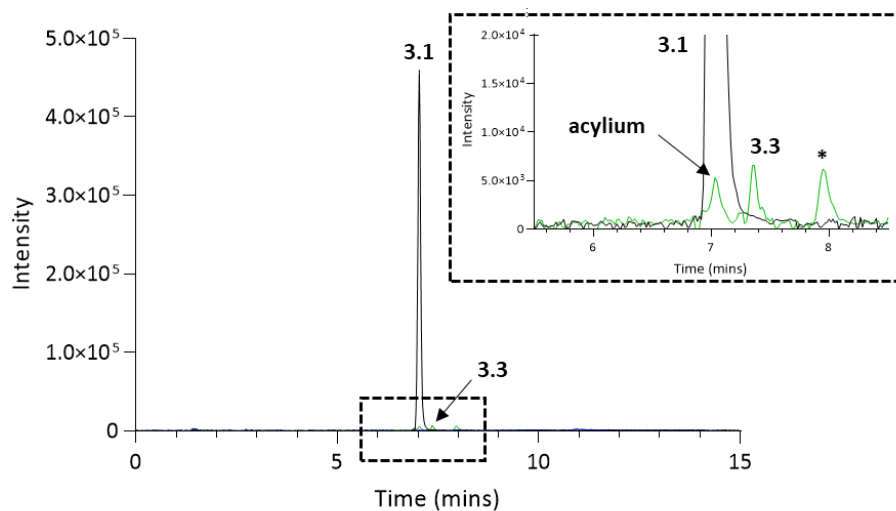
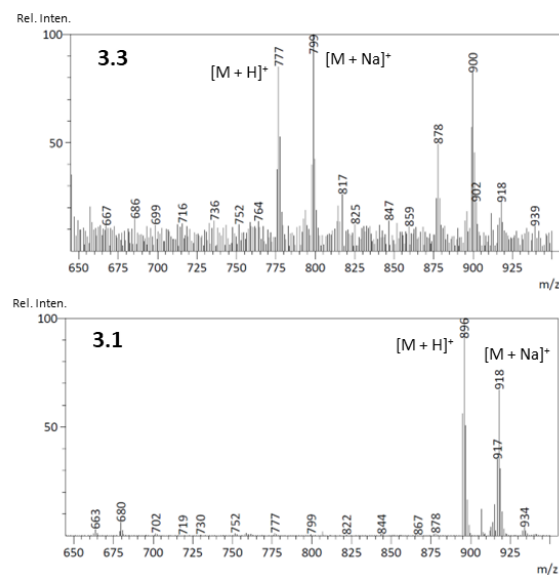
A**B**

Figure S3.10. Substrate 3.1 only control 22-hour incubation. (A) Extracted ion chromatogram for substrate **3.1** (0.9 mM) only control assay after 22 hours of incubation in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 777$ (**3.3**) and 896 (**3.1**). Shown boxed is a zoom-in of the trace with detection of macrocyclic product **3.3** formation compared to the acylium ion. Peak labelled with an asterisk is an injector peak seen in the 0-hour incubation (**Figure S3.7**). (B) Mass spectra for substrate **3.1** and macrocyclic product **3.3**, mass adducts are identified in each spectra.

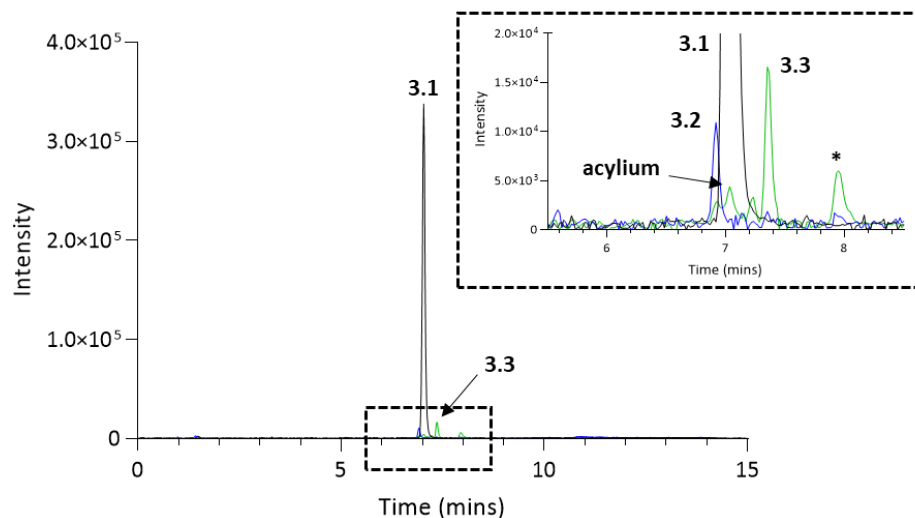
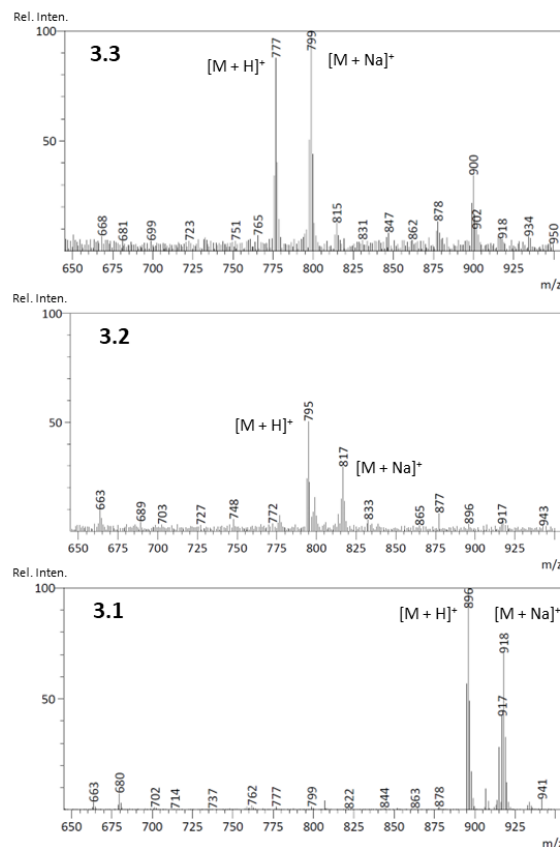
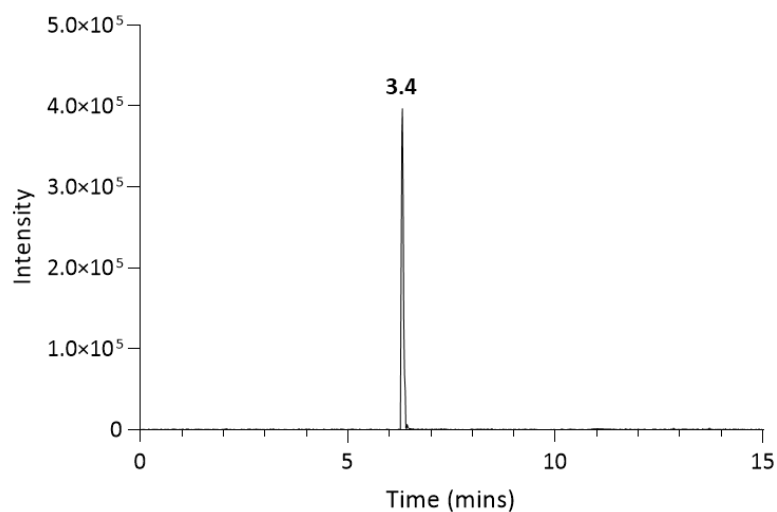
A**B**

Figure S3.11. Substrate 3.1 only control 5-day incubation. (A) Extracted ion chromatogram for substrate **3.1** (0.9 mM) only control assay after 5 days of incubation in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 777$ (**3.3**) and 896 (**3.1**). Shown boxed is a zoom-in of the trace with detection of hydrolysis product **3.2** and macrocyclic product **3.3** formation compared to the acylium ion. Peak labelled with an asterisk is an injector peak seen in the 0-hour incubation (**Figure S3.7**). (B) Mass spectra for substrate **3.1**, hydrolysis product **3.2** and macrocyclic product **3.3**, mass adducts are identified in each spectra.

N-acetyl peptide (**3.4**) assay LCMS chromatograms and mass spectra

A



B

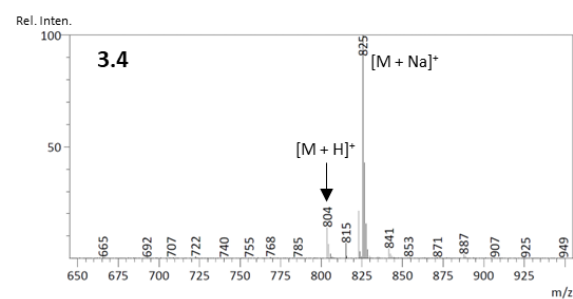


Figure S3.12 TE1-TE2 0-hour incubation with substrate **3.4.** (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 0 hours of incubation with 0.5 mM of **3.4** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 804$ (**3.4**). (B) Mass spectra for substrate **3.4**, mass adducts are identified in each spectra. Thermally denatured and substrate only controls all showed similar results for all time points. All enzymes showed similar results after 0 hours of incubation under identical conditions.

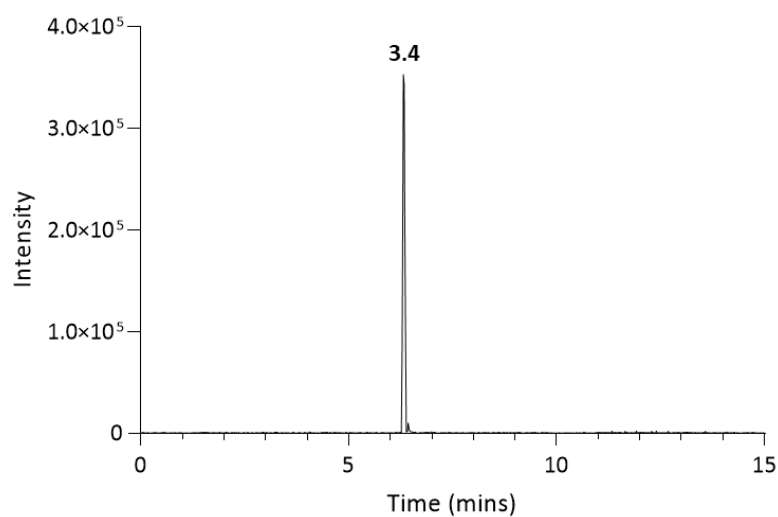
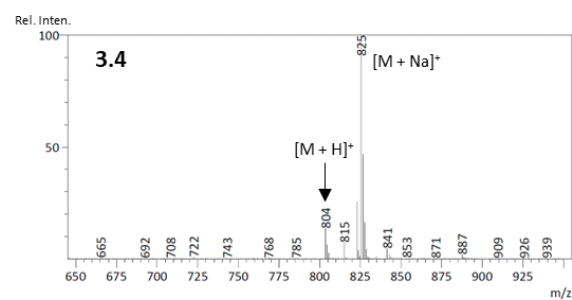
A**B**

Figure S3.13. TE1-TE2 3-hour incubation with substrate 3.4. (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 3 hours of incubation with 0.5 mM of **3.4** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 804$ (**3.4**). (B) Mass spectra for substrate **3.4**, mass adducts are identified in each spectra. All enzymes showed similar results after 3 hours of incubation under identical conditions.

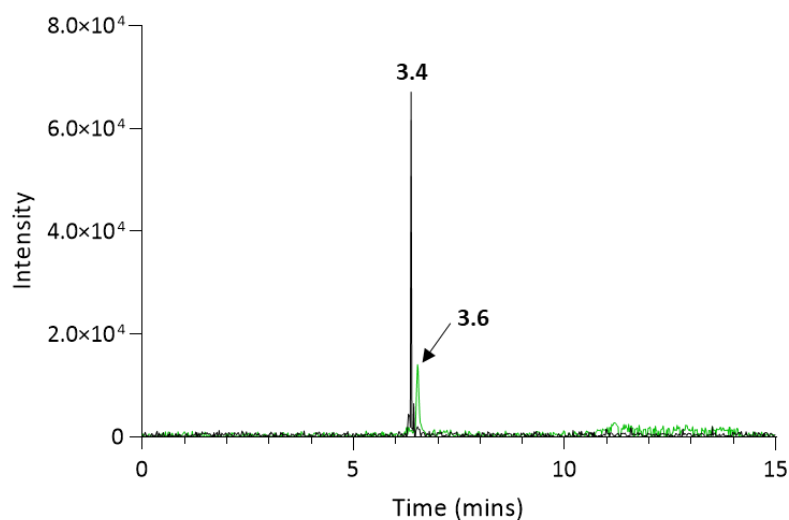
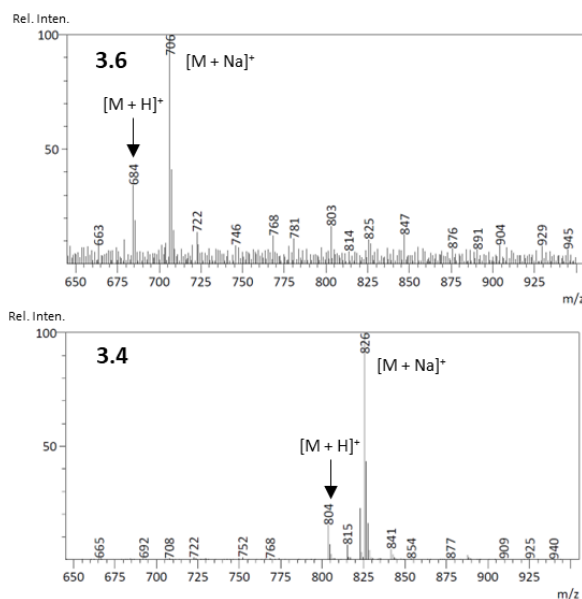
A**B**

Figure S3.14. TE1-TE2 22-hour incubation with substrate 3.4. (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 22 hours of incubation with 0.5 mM of **3.4** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 684$ (**3.6**) and 804 (**3.4**). (B) Mass spectra for substrate **3.4** and macrocyclic product **3.6**, mass adducts are identified in each spectra. Δ TE2 showed similar results after 22 hours of incubation under identical conditions. The peak denoted with an asterisk is an injector peak.

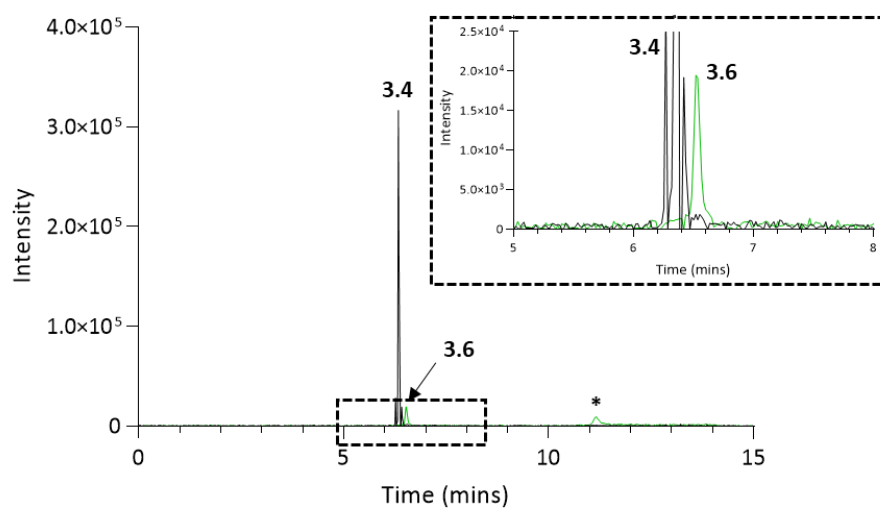
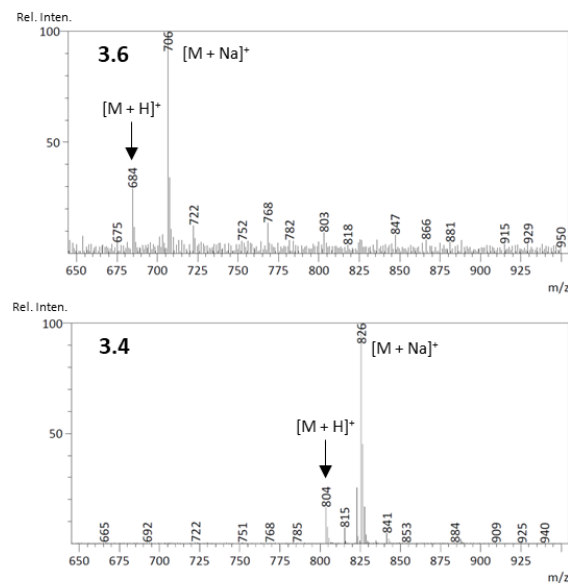
A**B**

Figure S3.15. TE1 22-hour incubation with substrate 3.4. (A) Extracted ion chromatogram for TE1 (5 μ M) assay after 22 hours of incubation with 0.5 mM of **3.4** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 684$ (**3.6**) and 804 (**3.4**). Shown boxed is a zoom-in of the trace with detection of macrocyclic product **3.6** formation compared to the substrate **3.4**. (B) Mass spectra for substrate **3.4** and macrocyclic product **3.6**, mass adducts are identified in each spectra. The peak denoted with an asterisk is an injector peak.

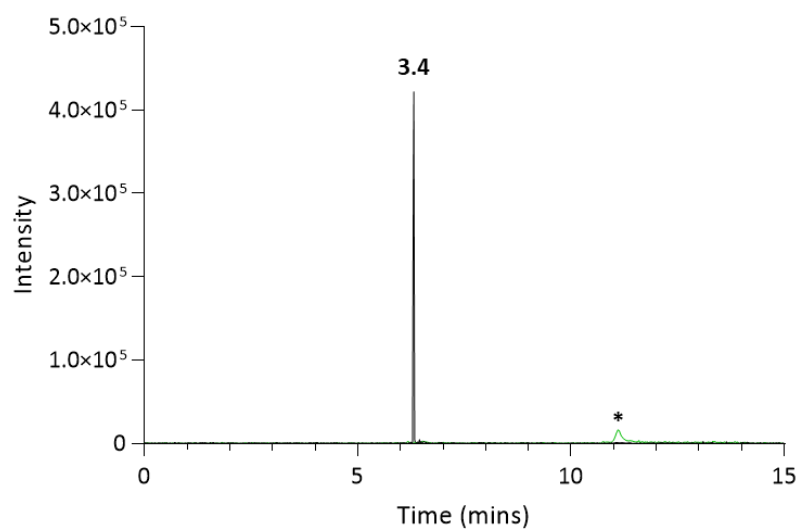
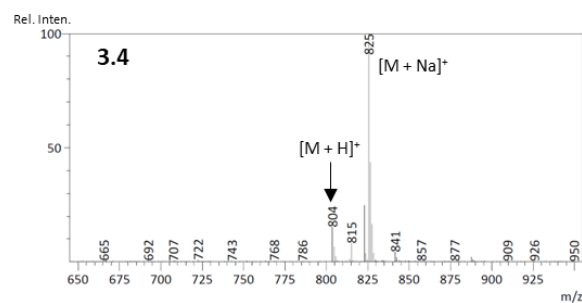
A**B**

Figure S3.16. Δ TE1 22-hour incubation with substrate 3.4. (A) Extracted ion chromatogram for Δ TE1 (5 μ M) assay after 22 hours of incubation with 0.5 mM of **3.4** in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 804$ (**3.4**). (B) Mass spectra for substrate **3.4**, mass adducts are identified in each spectra. All time points for Δ TE1 showed similar results under identical conditions. The peak denoted with an asterisk is an injector peak.

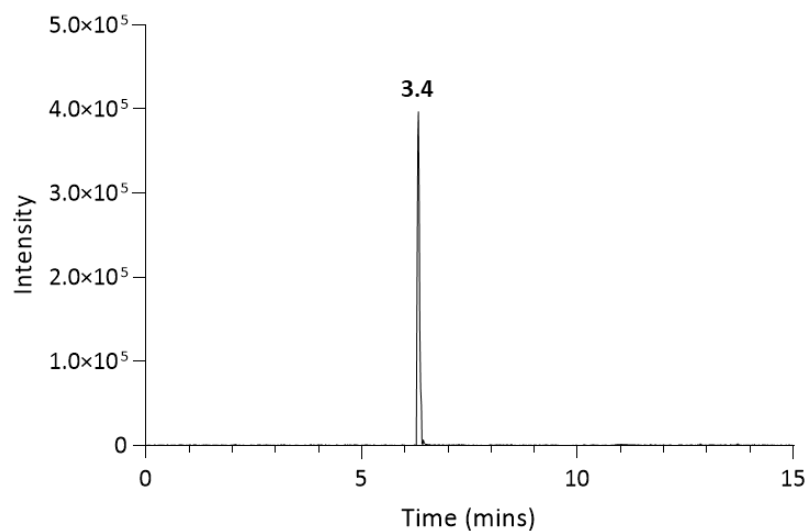
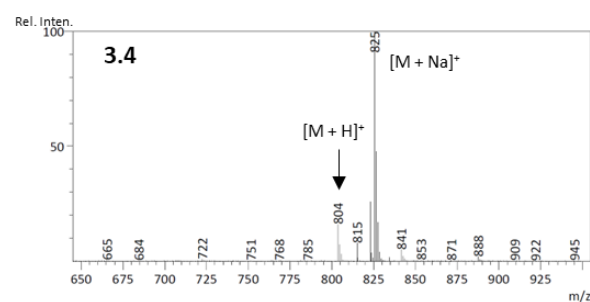
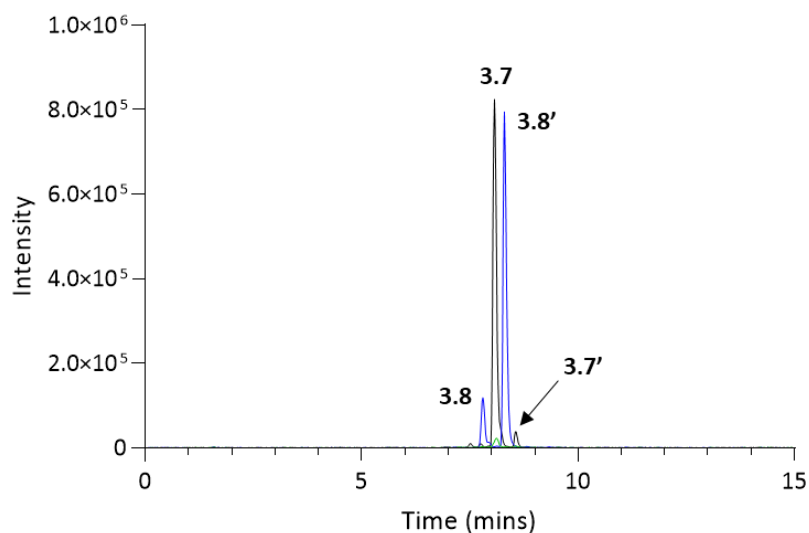
A**B**

Figure S3.17. Substrate 3.4 only control 22-hour incubation. (A) Extracted ion chromatogram for substrate **3.4** (0.5 mM) only control assay after 22 hours of incubation in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 804$ (**3.4**). (B) Mass spectra for substrate **3.4**, mass adducts are identified in the spectra. Substrate **3.4** only 5-day incubation showed similar results under identical conditions.

Thr_x peptide (3.7, 3.10 and 3.13) assay LCMS chromatograms and mass spectra

A



B

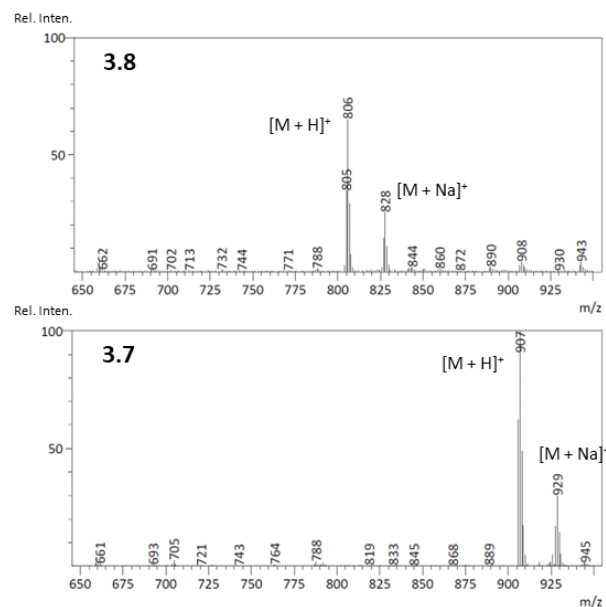


Figure S3.18. Crude substrate 3.7 only control 18-hour incubation. (A) Extracted ion chromatogram for crude **3.7** mixture (10% v/v) only control assay after 18 hours of incubation in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 806$ (**3.8** and **3.8'**) and 907 (**3.7** and **3.7'**). Peaks **3.7'** and **3.8'** correspond to the substrate and hydrolysis product at different retention times. (B) Mass spectra for substrate **3.7/3.7'** and hydrolysis product **3.8/3.8'**, mass adducts are identified in each spectra.

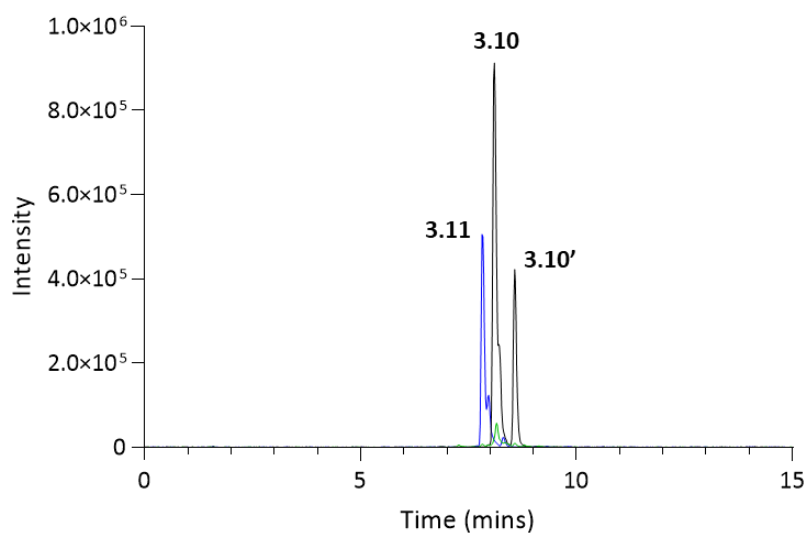
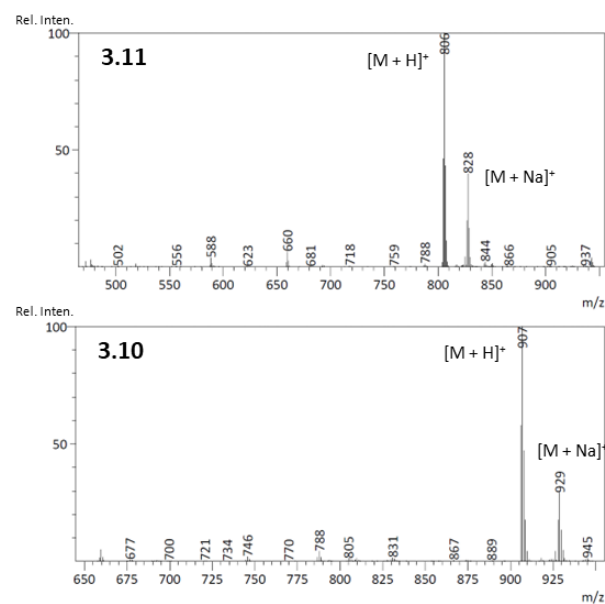
A**B**

Figure S3.19. Crude substrate 3.10 only control 18-hour incubation. (A) Extracted ion chromatogram for crude **3.10** mixture (10% v/v) only control assay after 18 hours of incubation in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 806$ (**3.11**) and 907 (**3.10** and **3.10'**). Peak **3.10'** correspond to the substrate at a different retention time. (B) Mass spectra for substrate **3.10/3.10'** and hydrolysis product **3.11**, mass adducts are identified in each spectra.

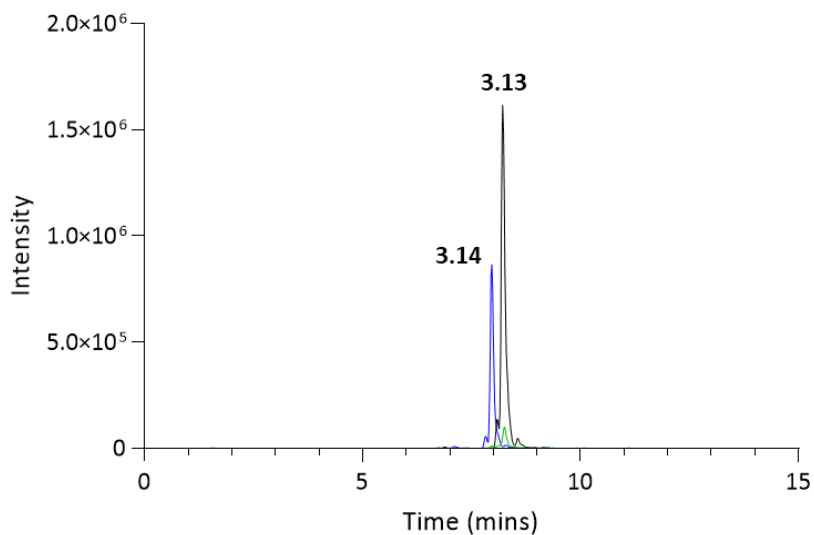
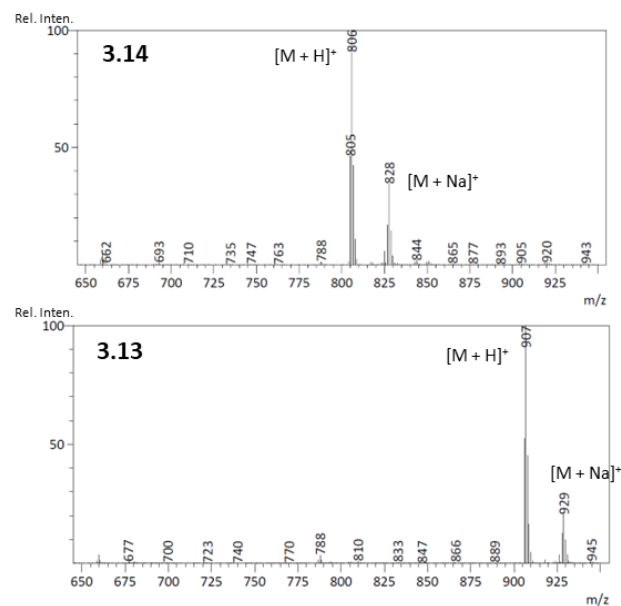
A**B**

Figure S3.20. Crude substrate 3.13 only control 18-hour incubation. (A) Extracted ion chromatogram for crude **3.13** mixture (10% v/v) only control assay after 18 hours of incubation in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M+H]^+$ ions with m/z = 806 (**3.14**) and 907 (**3.13**). (B) Mass spectra for substrate **3.13** and hydrolysis product **3.14**, mass adducts are identified in each spectra.

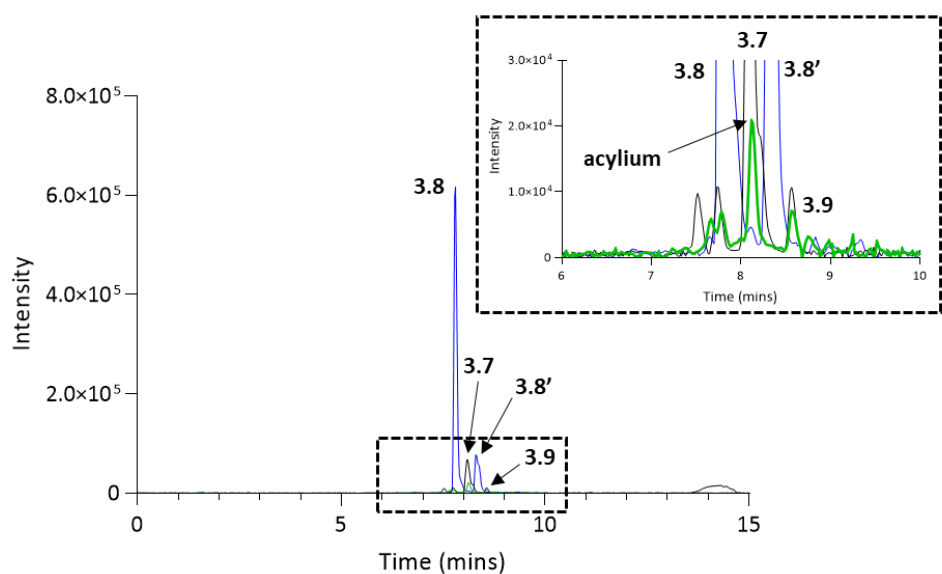
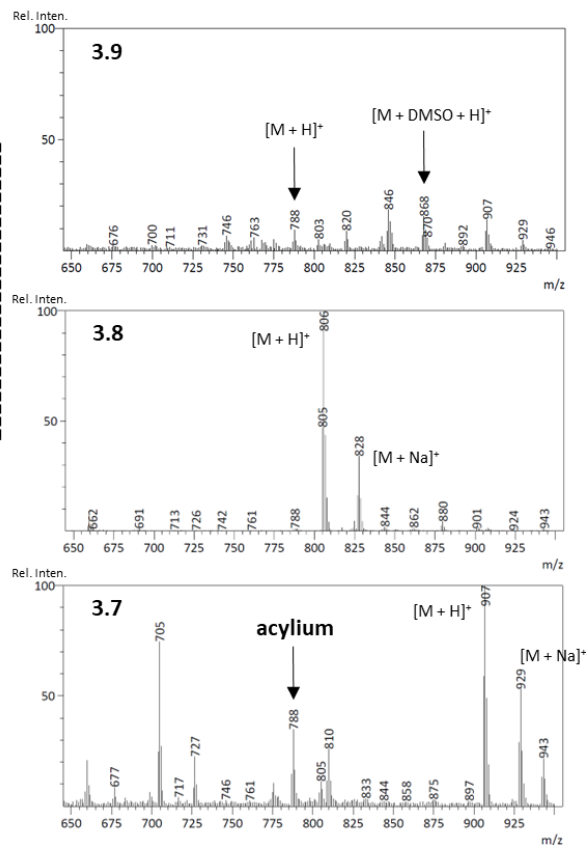
A**B**

Figure S3.21. TE1-TE2 18-hour incubation with crude substrate 3.7. (A) Extracted ion chromatogram for TE1-TE2 (5 μM) assay after 18 hours of incubation with 10% v/v of crude **3.7** mixture in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[\text{M} + \text{H}]^+$ ions with $m/z = 788$ (**3.9**), 806 (**3.8** and **3.8'**) and 907 (**3.7**). Peak **3.8'** correspond to the hydrolysis product at a different retention time. Shown boxed is a zoom-in of the trace with detection of macrocyclic product **3.9** (bolded green curve) formation compared to the acylium ion. (B) Mass spectra for substrate **3.7**, hydrolysis product **3.8/3.8'** and macrocyclic product **3.9**, mass adducts are identified in each spectra.

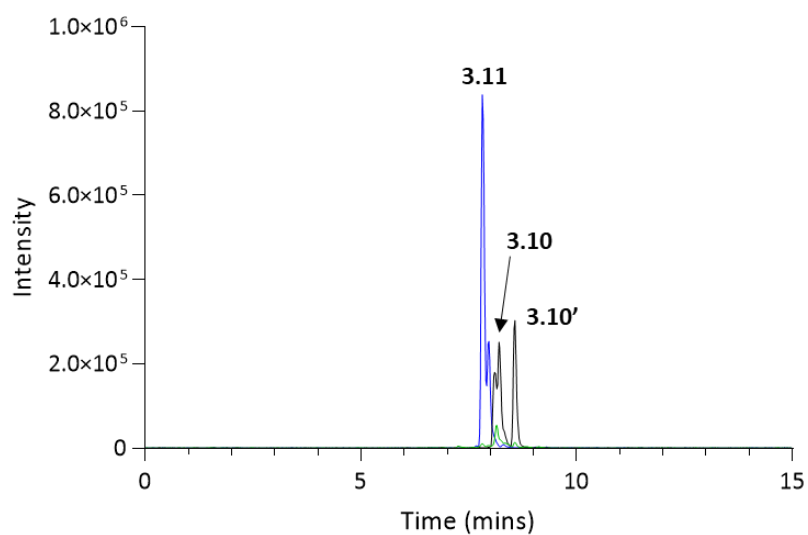
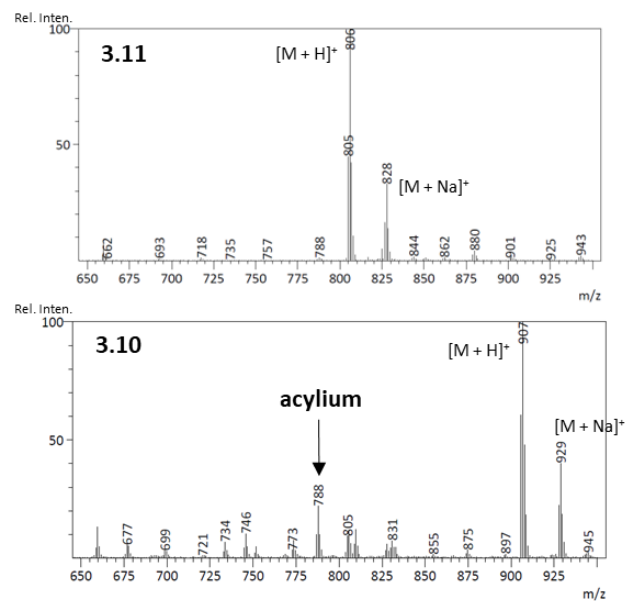
A**B**

Figure S3.22. TE1-TE2 18-hour incubation with crude substrate 3.10. (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 18 hours of incubation with 10% v/v of crude **3.10** mixture in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 806$ (**3.11**) and 907 (**3.10** and **3.10'**). Peak **3.10'** correspond to the substrate at a different retention time. (B) Mass spectra for substrate **3.10/3.10'** and hydrolysis product **3.11**, mass adducts are identified in each spectra.

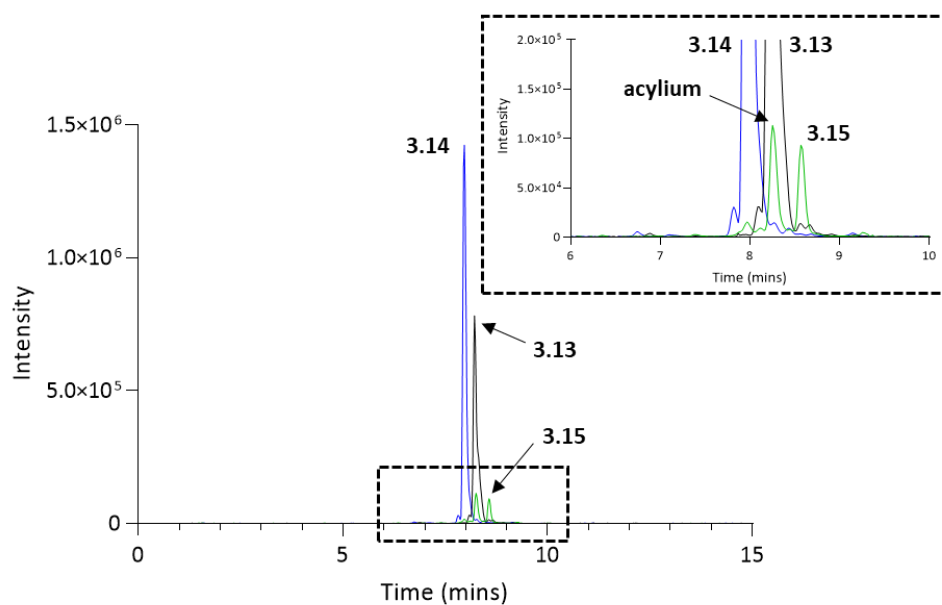
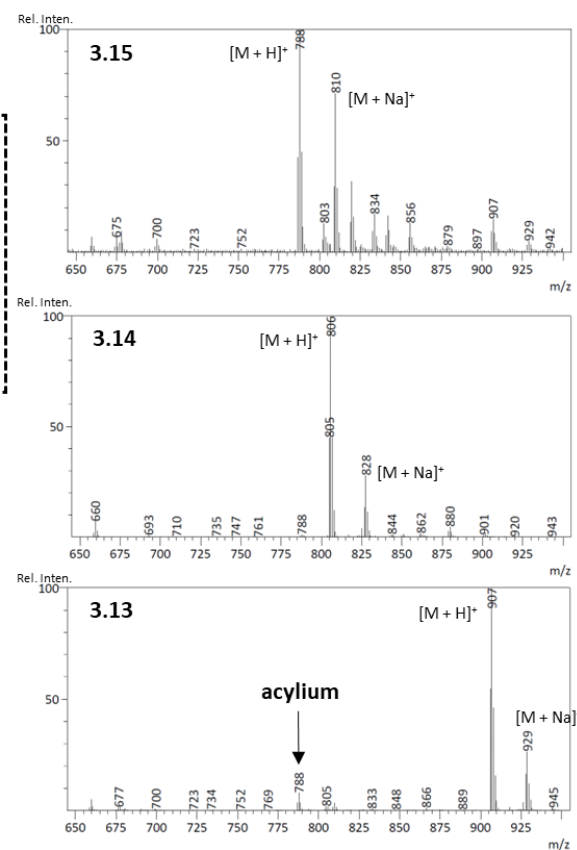
A**B**

Figure S3.23. TE1-TE2 18-hour incubation with crude substrate 3.13. (A) Extracted ion chromatogram for TE1-TE2 (5 μ M) assay after 18 hours of incubation with 10% v/v of crude **3.13** mixture in phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 788$ (**3.15**), 806 (**3.14**) and 907 (**3.13**). Shown boxed is a zoom-in of the trace with detection of macrocyclic product **3.15** formation compared to the acylium ion. (B) Mass spectra for substrate **3.13**, hydrolysis product **3.14** and macrocyclic product **3.15**, mass adducts are identified in each spectra.

Relative macrocycle conversion for *N*-Cbz (**3.1**) and *N*-acetyl (**3.4**) substrates

Table S3.1. Raw AUCs and relative macrocycle percent conversion. Enzymatic assays with either TE1-TE2, TE1, ΔTE1 or ΔTE2 were carried out in triplicate (technical replicates) in potassium phosphate buffer (50 mM, pH 7.4) with either 0.9 mM of **3.1** or 0.5 mM of **3.4**. Aliquots at time points 0, 3 and 22 hours were quenched and analyzed by LCMS. Substrate **3.1** only control was also quenched after 120 hours of incubation. AUCs were calculated from extracted ion chromatograms for [M + H]⁺ ions with *m/z* = 896 (**3.1**), 795 (**3.2**), 777 (**3.3**), 804 (**3.4**), 703 (**3.5**) and 684 (**3.6**). Macrocycle percent conversion was calculated by dividing the average macrocycle (**3.3** or **3.6**) AUC from the sum of averages of substrate (**3.1** or **3.4**), hydrolysis (**3.2** or **3.5**) and macrocycle AUCs.

	Sample	Incubation time (h)	Substrate AUC			Macrocycle AUC			Hydrolysis AUC			% conversion of macrocycle
			R1	R2	R3	R1	R2	R3	R1	R2	R3	
<i>N</i> -Cbz, 3.1	3.1 only	0	2241344	2196617	2131368	0	0	0	0	0	0	0.0
		3	2209295	2215210	2184090	0	0	0	0	0	0	0.0
		22	2378166	2282028	2352038	31242	32115	33844	35490	27856	32033	1.3
		120	1615709	1621094	1667288	69097	69846	65863	105148	98395	95558	3.8
	TE1-TE2	0	2498489	2327431	2211129	0	0	0	0	0	0	0.0
		3	2231739	2205133	2157034	37877	43529	37511	0	0	0	1.8
		22	2108993	2255497	2549881	130660	154434	174313	64163	66429	73075	6.1
	TE1	0	2315102	2311023	2280952	0	0	0	0	0	0	0.0
		3	2260571	2340213	2338671	51238	48693	50910	0	0	0	2.1
		22	2232787	2289379	2177794	171415	172583	177851	107279	115654	115780	6.9
<i>N</i> -acetyl, 3.4	3.4 only	0	1510862	1503156	1238951	0	0	0	0	0	0	0.0
		3	1271128	1205960	1304904	0	0	0	0	0	0	0.0
		22	1953501	1991374	1516318	0	0	0	0	0	0	0.0
	TE1-TE2	0	1399002	1962679	1231681	0	0	0	0	0	0	0.0
		3	1306607	1182178	1345653	0	0	0	0	0	0	0.0
		22	2316075	2170358	2120000	65085	62715	60088	0	0	0	2.8
	TE1	0	2005008	1690557	654355	0	0	0	0	0	0	0.0
		3	1362039	1509405	1395653	0	0	0	0	0	0	0.0
		22	1713411	1349008	1401406	85071	80134	80732	0	0	0	5.2
	ΔTE1	0	1463102	2351318	1038956	0	0	0	0	0	0	0.0
		3	1421534	1366844	1237274	0	0	0	0	0	0	0.0
		22	1854693	1865444	1238480	0	0	0	0	0	0	0.0
	ΔTE2	0	1183784	1900976	1492034	0	0	0	0	0	0	0.0
		3	1772191	1248436	1546905	0	0	0	0	0	0	0.0
		22	1388249	2089801	1320677	43244	48396	45177	0	0	0	2.8

3.7 References

- (1) Rose, W. E.; Leonard, S. N.; Sakoulas, G.; Kaatz, G. W.; Zervos, M. J.; Sheth, A.; Carpenter, C. F.; Rybak, M. J. Daptomycin Activity against Staphylococcus Aureus Following Vancomycin Exposure in an in Vitro Pharmacodynamic Model with Simulated Endocardial Vegetations. *Antimicrob. Agents Chemother.* **2008**, *52* (3), 831–836. <https://doi.org/10.1128/AAC.00869-07>.
- (2) Miller, W. R.; Bayer, A. S.; Arias, C. A. Mechanism of Action and Resistance to Daptomycin in Staphylococcus Aureus and Enterococci. *Cold Spring Harb. Perspect. Med.* **2016**, *6* (11), 1–16. <https://doi.org/10.1101/cshperspect.a026997>.
- (3) Sauermann, R.; Rothenburger, M.; Graninger, W.; Joukhadar, C. Daptomycin: A Review 4 Years after First Approval. *Pharmacology* **2008**, *81* (2), 79–91. <https://doi.org/10.1159/000109868>.
- (4) Johnson, A. Daptomycin in the Treatment of Skin, Soft-Tissue and Invasive Infections Due to Gram-Positive Bacteria. *Future Microbiol.* **2006**, *1* (3), 255–265. <https://doi.org/10.2217/17460913.1.3.255>.
- (5) Raaijmakers, J. M.; Bruijn, I. De; Kock, M. J. D. De. Cyclic Lipopeptide Production by Plant-Associated Pseudomonas Spp.: Diversity , Activity , Biosynthesis , and Regulation. *Mol. Plant-Microbe Interact.* **2006**, *19* (7), 699–710.
- (6) Nielsen, T. H.; Sørensen, D.; Tobiasen, C.; Andersen, J. B.; Christophersen, C.; Givskov, M.; Sørensen, J. Antibiotic and Biosurfactant Properties of Cyclic Lipopeptides Produced by Fluorescent Pseudomonas Spp. from the Sugar Beet Rhizosphere. *Appl. Environ. Microbiol.* **2002**, *68* (7), 3416–3423. <https://doi.org/10.1128/AEM.68.7.3416>.
- (7) Bionda, N.; Pitteloud, J. P.; Cudic, P. Cyclic Lipodepsipeptides: A New Class of Antibacterial Agents in the Battle against Resistant Bacteria. *Future Med. Chem.* **2013**, *5* (11), 1311–1330. <https://doi.org/10.4155/fmc.13.86>.
- (8) Schneider, T.; Müller, A.; Miess, H.; Gross, H. Cyclic Lipopeptides as Antibacterial Agents - Potent Antibiotic Activity Mediated by Intriguing Mode of Actions. *Int. J. Med. Microbiol.* **2014**, *304* (1), 37–43. <https://doi.org/10.1016/j.ijmm.2013.08.009>.
- (9) Patel, S.; Ahmed, S.; Eswari, J. S. Therapeutic Cyclic Lipopeptides Mining from Microbes : Latest Strides and Hurdles. *World J. Microbiol. Biotechnol.* **2015**, *31* (8), 1177–1193. <https://doi.org/10.1007/s11274-015-1880-8>.
- (10) Loper, J. E.; Hassan, K. A.; Mavrodi, D. V.; Li, E. W. D.; Lim, C. K.; Shaffer, B. T.; Elbourne, L. D. H.; Stockwell, V. O.; Hartney, S. L.; Breakwell, K.; Henkels, M. D.; Tetu, S. G.; Rangel, L. I.; Kidarsa, T. A.; Wilson, N. L.; Van, J. E.; Mortel, D.; Song, C.; Blumhagen, R.; Radune, D.; Hostetler, J. B.; Brinkac, L. M.; Durkin, A. S.; Kluepfel, D. A.; Wechter, W. P.; Anderson, A. J.; Kim, Y. C.; Iii, L. S. P.; Pierson, E. A.; Lindow, S. E.; Kobayashi, D. Y. Comparative Genomics of Plant-Associated Pseudomonas Spp.: Insights into Diversity and Inheritance of Traits Involved in Multitrophic Interactions. *PLoS Genet.* **2012**, *8* (7), e1002784. <https://doi.org/10.1371/journal.pgen.1002784>.
- (11) Jahanshah, G.; Yan, Q.; Gerhardt, H.; Pataj, Z.; Lämmerhofer, M.; Pianet, I.; Josten, M.; Sahl, H. G.; Silby, M. W.; Loper, J. E.; Gross, H. Discovery of the Cyclic Lipopeptide Gacamide a by Genome Mining and Repair of the Defective Gaca Regulator in Pseudomonas Fluorescens Pf0-1. *J. Nat. Prod.* **2019**, *82* (2), 301–308. <https://doi.org/10.1021/acs.jnatprod.8b00747>.
- (12) Thakur, P.; Saini, N. K.; Thakur, V. K.; Gupta, V. K.; Saini, R. V.; Saini, A. K. Rhamnolipid the Glycolipid Biosurfactant: Emerging Trends and Promising Strategies in the Field of Biotechnology and Biomedicine. *Microb. Cell Fact.* **2021**, *20* (1), 1–15. <https://doi.org/10.1186/s12934-020-01497-9>.
- (13) Soberón-Chávez, G.; Lépine, F.; Déziel, E. Production of Rhamnolipids by Pseudomonas

- Aeruginosa. *Appl. Microbiol. Biotechnol.* **2005**, *68* (6), 718–725. <https://doi.org/10.1007/s00253-005-0150-3>.
- (14) Kraas, F. I.; Giessen, T. W.; Marahiel, M. A. Exploring the Mechanism of Lipid Transfer during Biosynthesis of the Acidic Lipopeptide Antibiotic CDA. *FEBS Lett.* **2012**, *586* (3), 283–288. <https://doi.org/10.1016/j.febslet.2012.01.003>.
 - (15) Chooi, Y. H.; Tang, Y. Adding the Lipo to Lipopeptides: Do More with Less. *Chem. Biol.* **2010**, *17* (8), 791–793. <https://doi.org/10.1016/j.chembiol.2010.08.001>.
 - (16) Zhong, L.; Diao, X.; Zhang, N.; Li, F.; Zhou, H.; Chen, H.; Bai, X.; Ren, X.; Zhang, Y.; Wu, D.; Bian, X. Engineering and Elucidation of the Lipoinitiation Process in Nonribosomal Peptide Biosynthesis. *Nat. Commun.* **2021**, *12* (296), 1–15. <https://doi.org/10.1038/s41467-020-20548-8>.
 - (17) Balibar, C. J.; Vaillancourt, F. H.; Walsh, C. T. Generation of D Amino Acid Residues in Assembly of Arthrofactin by Dual Condensation/Epimerization Domains. *Chem. Biol.* **2005**, *12* (11), 1189–1200. <https://doi.org/10.1016/j.chembiol.2005.08.010>.
 - (18) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic Analysis of Condensation Domains in NRPS Sheds Light on Their Functional Evolution. *BMC Evol. Biol.* **2007**, *7*, 1–15. <https://doi.org/10.1186/1471-2148-7-78>.
 - (19) Mandalapu, D.; Ji, X.; Chen, J.; Guo, C.; Liu, W. Q.; Ding, W.; Zhou, J.; Zhang, Q. Thioesterase-Mediated Synthesis of Teixobactin Analogues: Mechanism and Substrate Specificity. *J. Org. Chem.* **2018**, *83* (13), 7271–7275. <https://doi.org/10.1021/acs.joc.7b02462>.
 - (20) Roongsawang, N.; Washio, K.; Morikawa, M. In Vivo Characterization of Tandem C-Terminal Thioesterase Domains in Arthrofactin Synthetase. *ChemBioChem* **2007**, *8* (5), 501–512. <https://doi.org/10.1002/cbic.200600465>.
 - (21) Hou, J.; Robbel, L.; Marahiel, M. A. Identification and Characterization of the Lysobactin Biosynthetic Gene Cluster Reveals Mechanistic Insights into an Unusual Termination Module Architecture. *Chem. Biol.* **2011**, *18* (5), 655–664. <https://doi.org/10.1016/j.chembiol.2011.02.012>.
 - (22) Biggins, J. B.; Kang, H. S.; Ternei, M. A.; Deshazer, D.; Brady, S. F. The Chemical Arsenal of Burkholderia Pseudomallei Is Essential for Pathogenicity. *J. Am. Chem. Soc.* **2014**, *136* (26), 9484–9490. <https://doi.org/10.1021/ja504617n>.
 - (23) Scholz-Schroeder, B. K.; Soule, J. D.; Gross, D. C. The SypA, SypB, and SypC Synthetase Genes Encode Twenty-Two Modules Involved in the Nonribosomal Peptide Synthesis of Syringopeptin by Pseudomonas Syringae Pv. Syringae B301D. *Mol. Plant-Microbe Interact.* **2003**, *16* (4), 271–280. <https://doi.org/10.1094/MPMI.2003.16.4.271>.
 - (24) De Bruijn, I.; De Kock, M. J. D.; De Waard, P.; Van Beek, T. A.; Raaijmakers, J. M. Massetolide A Biosynthesis in Pseudomonas Fluorescens. *J. Bacteriol.* **2008**, *190* (8), 2777–2789. <https://doi.org/10.1128/JB.01563-07>.
 - (25) Little, R. F.; Hertweck, C. Chain Release Mechanisms in Polyketide and Non-Ribosomal Peptide Biosynthesis. *Nat. Prod. Rep.* **2022**, *39* (1), 163–205. <https://doi.org/10.1039/d1np00035g>.
 - (26) Fisher, C. L.; Pei, G. K. Modification of a PCR-Based Site-Directed Mutagenesis Method. *Biotechniques* **1997**, *23* (4), 570–574.
 - (27) Little, R.; Trottmann, F.; Preissler, M.; Hertweck, C. An Intramodular Thioesterase Domain Catalyses Chain Release in the Biosynthesis of a Cytotoxic Virulence Factor. *RSC Chem. Biol.* **2022**, *3* (9), 1121–1128. <https://doi.org/10.1039/d2cb00121g>.
 - (28) Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A. Regeneration of Misprimed Nonribosomal Peptide Synthetases by Type II Thioesterases. *PNAS* **2002**, *99* (22), 14083–14088. <https://doi.org/10.1073/pnas.212382199>.
 - (29) Hayashi, S.; Satoh, Y.; Ujihara, T.; Takata, Y.; Dairy, T. Enhanced Production of Polyunsaturated

- Fatty Acids by Enzyme Engineering of Tandem Acyl Carrier Proteins. *Sci. Rep.* **2016**, *6* (Article 35441), 1–6. <https://doi.org/10.1038/srep35441>.
- (30) Jiang, H.; Zirkle, R.; Metz, J. G.; Braun, L.; Richter, L.; Van Lanen, S. G.; Shen, B. The Role of Tandem Acyl Carrier Protein Domains in Polyunsaturated Fatty Acid Biosynthesis. *J. Am. Chem. Soc.* **2008**, *130* (20), 6336–6337. <https://doi.org/10.1021/ja801911t>.
 - (31) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. Identification of Claisen Cyclase Domain in Fungal Polyketide Synthase WA, a Naphthopyrone Synthase of *Aspergillus nidulans*. *Chem. Biol.* **2001**, *8* (2), 189–197. [https://doi.org/10.1016/S1074-5521\(00\)90068-1](https://doi.org/10.1016/S1074-5521(00)90068-1).
 - (32) Rahman, A. S.; Hotherhall, J.; Crosby, J.; Simpson, T. J.; Thomas, C. M. Tandemly Duplicated Acyl Carrier Proteins, Which Increase Polyketide Antibiotic Production, Can Apparently Function Either in Parallel or in Series. *J. Biol. Chem.* **2005**, *280* (8), 6399–6408. <https://doi.org/10.1074/jbc.M409814200>.
 - (33) Gu, L.; Eisman, E. B.; Dutta, S.; Franzmann, T. M.; Walter, S.; Gerwick, W. H.; Skiniotis, G.; Sherman, D. H. Tandem Acyl Carrier Proteins in the Curacin Biosynthetic Pathway Promote Consecutive Multienzyme Reactions with a Synergistic Effect. *Angew. Chemie - Int. Ed.* **2011**, *50* (12), 2795–2798. <https://doi.org/10.1002/anie.201005280>.
 - (34) Bu, X.; Wu, X.; Xie, G.; Guo, Z. Synthesis of Tyrocidine A and Its Analogues by Spontaneous Cyclization in Aqueous Solution. *Org. Lett.* **2002**, *4* (17), 2893–2895. <https://doi.org/10.1021/ol0263191>.
 - (35) Koehnke, J.; Bent, A. F.; Houssen, W. E.; Mann, G.; Jaspars, M.; Naismith, J. H. The Structural Biology of Patellamide Biosynthesis. *Curr. Opin. Struct. Biol.* **2014**, *29*, 112–121. <https://doi.org/10.1016/j.sbi.2014.10.006>.
 - (36) Milne, B. F.; Long, P. F.; Starcevic, A.; Hranueli, D.; Jaspars, M. Spontaneity in the Patellamide Biosynthetic Pathway. *Org. Biomol. Chem.* **2006**, *4* (4), 631–638. <https://doi.org/10.1039/b515938e>.
 - (37) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry* **2001**, *40* (24), 7092–7098. <https://doi.org/10.1021/bi010035r>.
 - (38) Yeh, E.; Lin, H.; L. Clugston, S.; M. Kohli, R.; T. Walsh, C. Enhanced Macrocyclizing Activity of the Thioesterase from Tyrocidine Synthetase in Presence of Nonionic Detergent. *Chem. Biol.* **2004**, *11* (11), 1573–1582. <https://doi.org/10.1016/j.chembiol.2004.09.003>.
 - (39) Martí-Centelles, V.; Pandey, M. D.; Burguete, M. I.; Luis, S. V. Macrocyclization Reactions: The Importance of Conformational, Configurational, and Template-Induced Preorganization. *Chem. Rev.* **2015**, *115* (16), 8736–8834. <https://doi.org/10.1021/acs.chemrev.5b00056>.
 - (40) Ficht, S.; Payne, R. J.; Guy, R. T.; Wong, C. H. Solid-Phase Synthesis of Peptide and Glycopeptide Thioesters through Side-Chain-Anchoring Strategies. *Chem. - A Eur. J.* **2008**, *14* (12), 3620–3629. <https://doi.org/10.1002/chem.200701978>.
 - (41) Derbal, S.; Hensler, M.; Fang, W.; Nizet, V.; Ghedira, K.; Nefzi, A. On Resin Amino Acid Side Chain Attachment Strategy for the Head to Tail Synthesis of New Glutamine Containing Gramicidin-S Analogs and Their Antimicrobial Activity. *Bioorganic Med. Chem. Lett.* **2010**, *20* (19), 5701–5704. <https://doi.org/10.1016/j.bmcl.2010.08.015>.
 - (42) Arbour, C. A.; Mendoza, L. G.; Stockdill, J. L. Recent Advances in the Synthesis of C-Terminally Modified Peptides. *Org. Biomol. Chem.* **2020**, *18* (37), 7253–7272. <https://doi.org/10.1039/d0ob01417f>.
 - (43) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. Mechanism and Specificity of the Terminal Thioesterase Domain from the Erythromycin Polyketide Synthase. *Chem. Biol.* **1999**, *6* (2), 117–

125. [https://doi.org/10.1016/S1074-5521\(99\)80008-8](https://doi.org/10.1016/S1074-5521(99)80008-8).
- (44) Wang, M.; Boddy, C. N. Examining the Role of Hydrogen Bonding Interactions in the Substrate Specificity for the Loading Step of Polyketide Synthase Thioesterase Domains. *Biochemistry* **2008**, *47* (45), 11793–11803. <https://doi.org/10.1021/bi800963y>.
- (45) Horsman, M. E.; Hari, T. P. A. A.; Boddy, C. N. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016**, *33* (2), 183–202. <https://doi.org/10.1039/c4np00148f>.
- (46) Wang, M.; Opere, P.; Boddy, C. N. Polyketide Synthase Thioesterases Catalyze Rapid Hydrolysis of Peptidyl Thioesters. *Bioorg. Med. Chem. Lett.* **2009**, *19* (5), 1413–1415. <https://doi.org/10.1016/j.bmcl.2009.01.040>.
- (47) Firn, R. D.; Jones, C. G. Natural Products – a Simple Model to Explain Chemical Diversity. *Nat. Prod. Rep.* **2003**, *20*, 382–391. <https://doi.org/10.1039/b208815k>.
- (48) Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dornan, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W. Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature* **2019**, *565* (7737), 112–117. <https://doi.org/10.1038/s41586-018-0781-z>.
- (49) Brazeau-Henrie, J. T.; Paquette, A. R.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N. Total and Chemoenzymatic Synthesis of Seongsanamide E. *Org. Lett.* **2022**, *24* (35), 6369–6373. <https://doi.org/10.1021/acs.orglett.2c02271>.
- (50) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. Characterization of the Surfactin Synthetase C-Terminal Thioesterase Domain as a Cyclic Depsipeptide Synthase. *Biochemistry* **2002**, *41* (45), 13350–13359. <https://doi.org/10.1021/bi026592a>.
- (51) Lepetit, C. A.; Paquette, A. R.; Brazeau-Henrie, J. T.; Boddy, C. N. Total and Chemoenzymatic Synthesis of the Lipodepsipeptide Rhizomide A. *Bioorg. Med. Chem. Lett.* **2023**, *96* (Article 129506), 1–5. <https://doi.org/10.1016/j.bmcl.2023.129506>.
- (52) Mahlert, C.; Sieber, S. A.; Grünwald, J.; Marahiel, M. A. Chemoenzymatic Approach to Enantiopure Streptogramin B Variants: Characterization of Stereoselective Pristinamycin I Cyclase from *Streptomyces pristinaespiralis*. *J. Am. Chem. Soc.* **2005**, *127* (26), 9571–9580. <https://doi.org/10.1021/ja051254t>.
- (53) Heberlig, G. W.; Boddy, C. N. Thioesterase from Cereulide Biosynthesis Is Responsible for Oligomerization and Macrocyclization of a Linear Tetradepsipeptide. *J. Nat. Prod.* **2020**, *83* (6), 1990–1997. <https://doi.org/10.1021/acs.jnatprod.0c00333>.
- (54) Grünwald, J.; Sieber, S. A.; Marahiel, M. A. Chemo- and Regioselective Peptide Cyclization Triggered by the N-Terminal Fatty Acid Chain Length: The Recombinant Cyclase of the Calcium-Dependent Antibiotic from *Streptomyces coelicolor*. *Biochemistry* **2004**, *43* (10), 2915–2925. <https://doi.org/10.1021/bi036140d>.
- (55) Zhou, Y.; Lin, X.; Xu, C.; Shen, Y.; Wang, S. P.; Liao, H.; Li, L.; Deng, H.; Lin, H. W. Investigation of Penicillin Binding Protein (PBP)-like Peptide Cyclase and Hydrolase in Surugamide Non-Ribosomal Peptide Biosynthesis. *Cell Chem. Biol.* **2019**, *26* (5), 737–744.e4. <https://doi.org/10.1016/j.chembiol.2019.02.010>.
- (56) Kobayashi, M.; Fujita, K.; Matsuda, K.; Wakimoto, T. Streamlined Chemoenzymatic Synthesis of Cyclic Peptides by Non-Ribosomal Peptide Cyclases. *J. Am. Chem. Soc.* **2023**, *145* (6), 3270–3275. <https://doi.org/10.1021/jacs.2c11082>.
- (57) Hari, T. P. A.; Labana, P.; Boileau, M.; Boddy, C. N. An Evolutionary Model Encompassing Substrate Specificity and Reactivity of Type I Polyketide Synthase Thioesterases. *ChemBioChem* **2014**, *15* (18), 2656–2661. <https://doi.org/10.1002/cbic.201402475>.

- (58) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.

CHAPTER 4 : Engineering the DEBS-TE as a better macrocyclization biocatalyst

4.1 Preface

The work presented in this chapter is the culmination of our work on engineering the DEBS-TE via rational protein design to generate mutants with improved macrocyclization activity. A manuscript is currently being drafted, in collaboration with Dr. Schmeing's lab at McGill University (Montréal, Québec). The Schmeing lab has solved a high-resolution structure of the wild-type (WT) DEBS-TE with our substrate **4.1** covalently linked to an active site 2,3-diaminopropionic acid (DAP) residue and will solve the structure of some of our DEBS-TE mutants with substrate **4.1**.

Chapter contributions

I performed all experiments and data analysis with the exception of substrate synthesis. Dr. Graham W. Heberlig devised the synthetic route. Substrate **4.1** was kindly provided to me by Dr. Christopher N. Boddy, who synthesized **4.1** and performed all compound characterization. WT DEBS-TE trapped acyl crystal structure with **4.1** was acquired and kindly provided to me by our collaborators in Dr. Schmeing's lab. Both Dr. Christopher N. Boddy and I devised the project.

Polyketides are a class of natural products produced by many organisms such as bacteria, fungi, and plants.¹⁻³ This class of natural products presents itself as one of the largest in nature with remarkable chemical diversity and many useful biological activities. For example, ossamycin has potent antifungal and cytotoxic activity by inhibiting the mitochondrial F_1F_0 -ATPase.⁴ Erythromycin A is a potent broad-spectrum antibiotic that inhibits the 50S ribosomal subunit, effectively stalling protein synthesis in many bacterial pathogens.⁵ Finally, epothilone A is a potential anticancer agent that stabilizes microtubules by binding to tubulin and inducing mitotic arrest.⁶



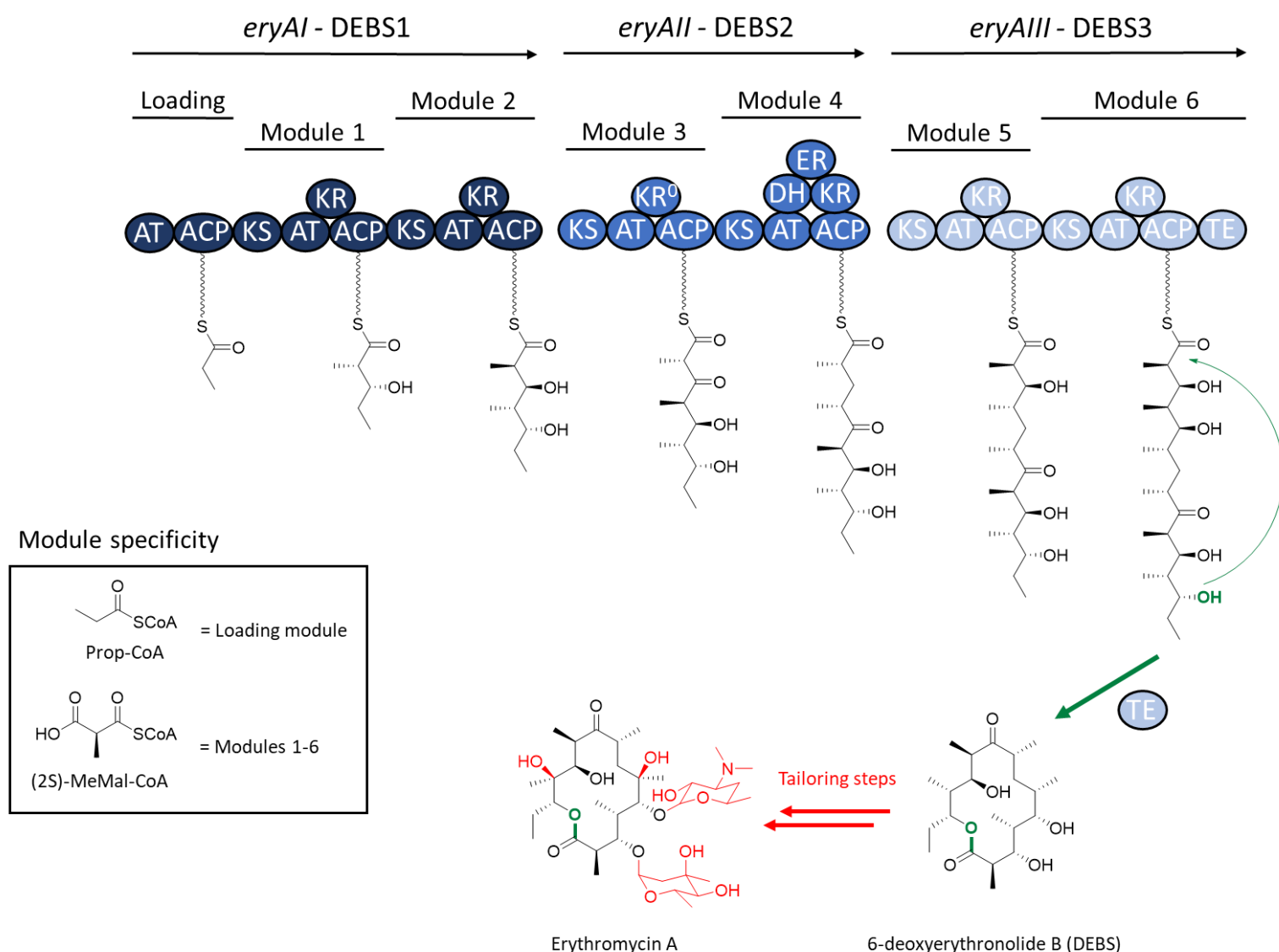


Figure 4.2. Biosynthesis of erythromycin A. The erythromycin biosynthetic gene cluster contains three PKS genes, *eryAI-eryAIII*, responsible for the biosynthesis of the macrocyclic precursor to erythromycin A, 6-deoxyerythronolide B (DEBS). A C-terminal TE domain is responsible for the macrocyclic release of DEBS for post-assembly tailoring. Specificity for ketide building blocks for every module is shown boxed. Abbreviations; prop-CoA, propionyl-CoA; MeMal-CoA, methylmalonyl-CoA; AT, acyltransferase domain; ACP, acyl carrier protein; KS, ketosynthase domain; KR, ketoreductase domain; KR⁰, inactive ketoreductase domain; DH, dehydratase domain; ER, enoylreductase domain; TE, thioesterase domain.

The most well studied polyketide to date is the macrolide antibiotic erythromycin A isolated from *Saccharopolyspora erythraea* (formerly *Streptomyces erythraeus*),^{7,8} with numerous published studies on its PKS machinery and on its cognate TE domain.^{9–19} Its biosynthesis takes place through a type I PKS biosynthetic gene cluster encoding three PKS genes that assemble the polyketide from acyl thioester intermediates loaded onto the phosphopantetheine arm (PPT) of ACP domains and then shuttled between the active sites of the catalytic domains. The PKS possesses thio-templated logic through the AT domains, incorporating propionyl-CoA as the starting unit followed by methylmalonyl-CoA for polyketide extension with KR, DH and ER domains appropriately positioned for the correct oxidation state (**Figure 4.2**).^{20,21} The final stage of polyketide biosynthesis is off-loading from the biosynthetic machinery typically through the action of a C-terminal TE domain.²² In this case, the TE catalyzes the release concomitant to macrolactonization of the linear precursor, producing 6-deoxyerythronolide B (DEBS) that is eventually converted into erythromycin A (**Figure 4.2**).

While there is substantial interest in creating macrocyclic structures synthetically, the process has proven to be challenging. This difficulty arises due to the need for end-to-end contact, resulting in high entropic costs and activation barriers, in which enables competing oligomerization reactivity. Furthermore, macrocycles are densely functionalized, creating distinctive electronic and steric environments that present unique barriers to macrocyclization.^{23–25} The first reported synthesis of erythromycin A, although successful, required testing over 17 variations of the Corey-Nicolaou macrolactonization through the systematic switching of protecting groups on the linear precursor.²⁶ The TE domain responsible for macrocyclization in the DEBS pathway presents itself as a powerful alternative to chemical macrocyclization. For example, it has considerable catalytic flexibility with the ability to form 6-, 8-, 12-, 14- and 16-membered macrolactones when artificially fused to the C-terminus of PKS modules.^{11,17–19,27} In addition, the standalone TE can effectively convert linear thioester activated

substrates into macrocycles *in vitro* (**Figure 4.3A**),^{9,10,14–16} showing promise as a biocatalyst. However, the DEBS-TE has been shown to be stereoselective for the nucleophilic alcohol as well as for distal alcohol groups (**Figure 4.3B**) and shows a preference for substrates featuring a carbonyl in the polyketide backbone.^{9,10,14–16} While it forms the desired macrocycle, it can also generate substantial hydrolysis product (**Figure 4.3A**).

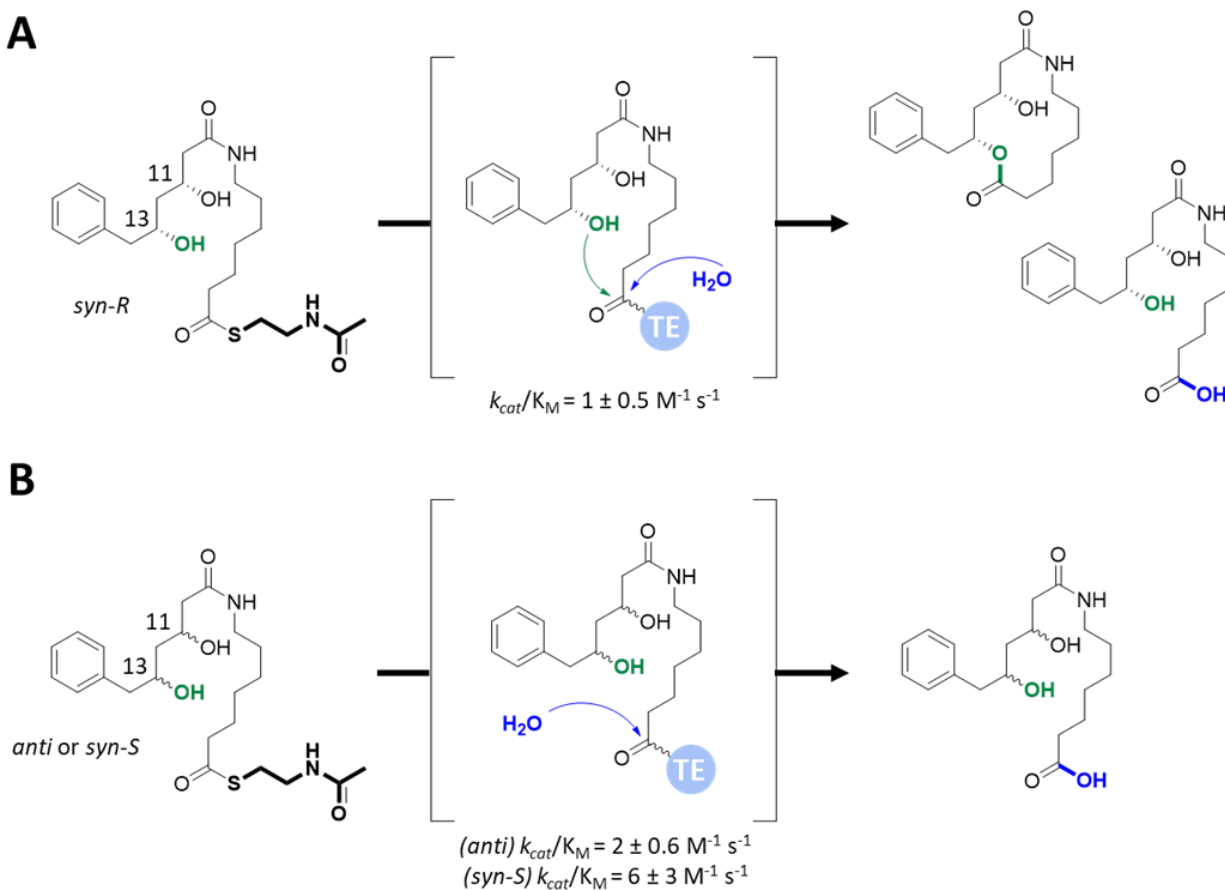


Figure 4.3. DEBS-TE *in vitro* reactions with SNAC activated substrates. (A) Macrolactonization and hydrolysis catalyzed by the DEBS-TE on an *R*-configured *syn* diol substrate. (B) DEBS-TE only catalyzes hydrolysis of an *S*-configured *syn* diol substrate and its *anti* diastereomer (*S*-configured C13) due to its stereoselectivity. Values for k_{cat}/K_M were obtained from Pinto *et al.* DEBS-TE study.¹⁵

Studies on the DEBS-TE structure in complex with diphenyl phosphonate inhibitors have been performed in the hopes of understanding the structural determinants for the off-loading chemistry. The

structural analysis suggested that enzyme-substrate binding is likely dominated by hydrophobic packing interactions, aligning with the structural analysis from a related TE domain in the pikromycin (PIK-TE) biosynthetic gene cluster and insights from a previous mutagenesis study on the DEBS-TE.^{9,28–30} While these inhibitors are structurally analogous to substrates known to undergo macrocyclization, the discrete substrate-TE interactions that dictate the off-loading chemistry remain elusive. As a result, it remains challenging to predict whether a particular substrate will undergo macrocyclization or hydrolysis prior to experimental characterization.^{15,31}

While the DEBS-TE shows promise as a biocatalyst, the substantial competing hydrolysis reactivity presents a problem. In many cases protein engineering has been used to optimize promising biocatalyst candidates,^{32,33} yet, optimization of PKS TE off-loading reactivity through protein engineering remains unaddressed. This is presumably due to the complexity in creating a high-throughput screen/selection, given the absence of a straightforward way to link formation of the macrocycle to cell viability or generation of a fluorescence signal. Herein, we sought to engineer the DEBS-TE to improve its macrocyclization activity towards a linear polyketide substrate mimic **4.1**. A rational protein design approach was taken, where 9 sites within the active site were chosen for site saturation mutagenesis. A comprehensive 1,323 clone library was built, ensuring full coverage of all possible mutants, and screened via LCMS for unambiguous detection of macrocycle **4.3** and hydrolysis **4.2** products formation. The screen identified three single mutants with improved production of **4.3**, as demonstrated by normalized **4.3** to **4.2** (M/H) ratios of product distribution assays with purified enzymes. Furthermore, combinatorial mutants were generated to assess for possible mutation synergy. Although no apparent synergy was detected, a fourth mutant with improved macrocyclization activity was identified. Finally, the WT DEBS-TE crystal structure in complex with substrate **4.1** was acquired, providing key insights on residues important for macrocyclization activity.

4.3 Results and discussion

DEBS-TE mutant library construction and screen

The objective for engineering of the wild-type (WT) DEBS-TE was the generation of enzyme variants with improved macrocycle **4.3** over hydrolysis **4.2** product (M/H) ratios when incubated with substrate **4.1** *in vitro*. A rational protein design approach was taken through site saturation mutagenesis of 9 residues in the active site as identified from a high-resolution crystal structure (PDB: 5D3K, 1.7Å resolution, **Figure 4.4A-B**).⁹ The structural determinants for off-loading chemistry and discrete enzyme-substrate interactions are not known. Consequently, we selected these residues for being solvent accessible and their proposed hydrophobic packing interactions with the substrate,^{9,28} increasing our chances of yielding mutants with changes in off-loading chemistry selectivity. All residues are positioned on the two core ($\alpha 2$ and $\alpha 10$) and lid ($\alpha 6/7$) α -helices that shape the entrance cavity, except for A77. Because there is no literature precedence for substrate pre-organization and DEBS-TE has notable hydrolysis reactivity, it likely adheres to our proposed ICS model where solvent access to the acyl-TE intermediate as it samples conformations generates substantial hydrolysis. Thus, we believe mutagenesis at these sites will not only change enzyme-substrate interactions, but also affect the conformation of the entrance cavity. This can result in mutants that potentially “gate” access of solvent into the active site, slowing acyl-TE hydrolysis rates and thus increasing the M/H ratio.

Site saturation mutagenesis³⁴ was performed at each position using the NNS degenerate codon, encoding all 20 canonical amino acids while minimizing the oversampling required in library screening. To screen all 171 mutants, a library comprising 1,323 clones was generated for ~4.6× oversampling (147 clones per site for >99% library coverage).^{35,36} Transformation into *Escherichia coli* BL21(DE3) chemically competent cells yielded clones that were arrayed into 96-well plates containing LB media with ampicillin

(100 $\mu\text{g/mL}$) and 30% w/v glycerol for long term storage as master plates (**Figure 4.5A**) for subsequent library screening and hit identification.

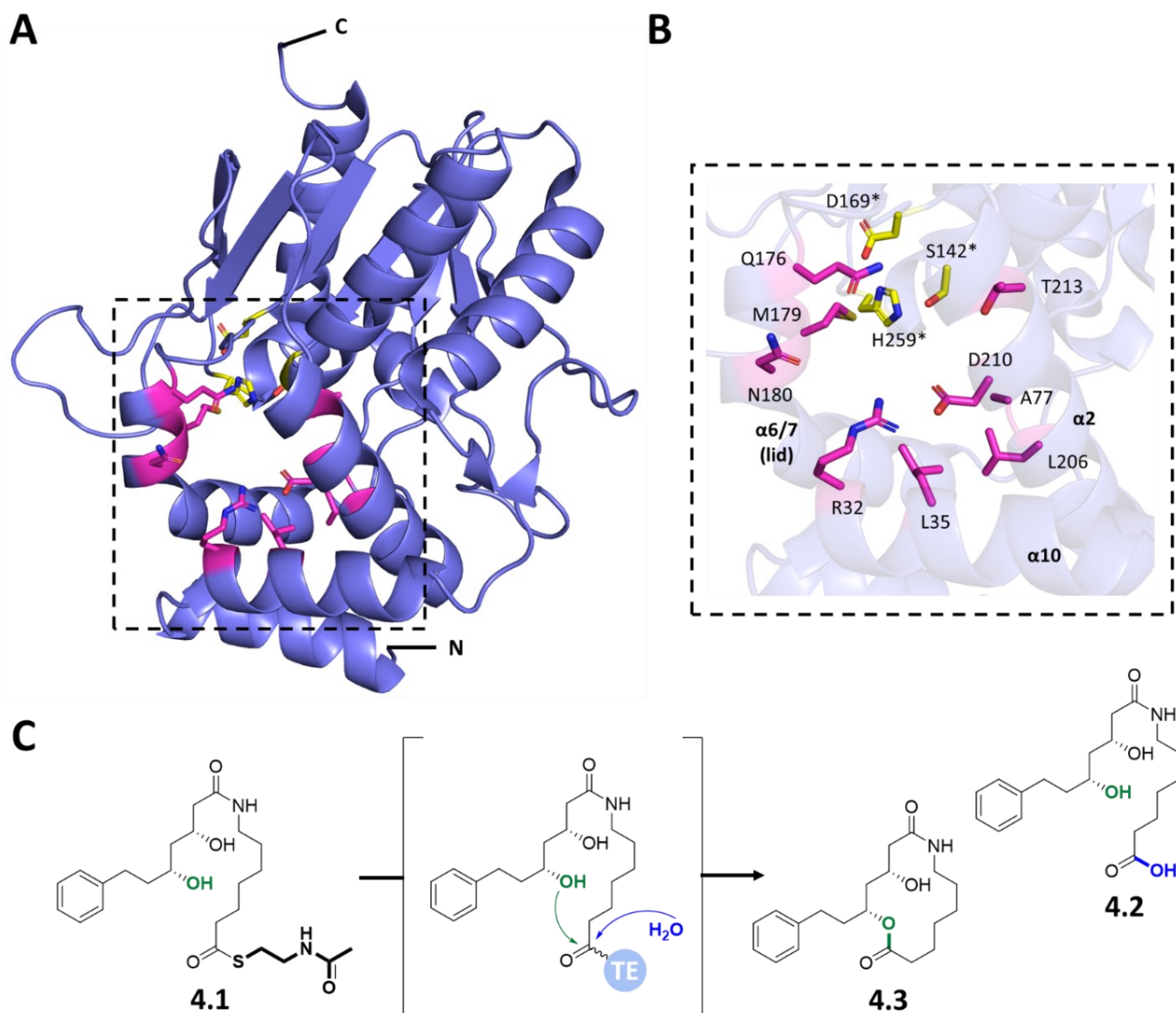


Figure 4.4. Chosen sites for mutagenesis and WT DEBS-TE reaction with SNAC substrate 4.1. (A) Crystal structure of the WT *apo* DEBS-TE (PDB: 5D3K, 1.7 Å resolution) shown as a purple cartoon with the active site boxed. (B) Zoom-in of the active site with the catalytic triad shown as yellow sticks. Sites chosen for site saturation mutagenesis are shown as magenta sticks and are solvent accessible. The structural elements comprising the entrance cavity are labelled in bold. (C) Macrolactonization and hydrolysis of SNAC substrate **4.1** catalyzed by the WT DEBS-TE, producing macrolactone **4.3** and seco acid **4.2**.

To date, no TE has been engineered via directed evolution, requiring us to develop our own selection/screen method for DEBS-TE engineering. For selection to be viable, production of **4.2** would need to be toxic *in vivo* or **4.3** would need to exhibit antibiotic activity. Unfortunately, products **4.2** and **4.3** have not shown these activities. Alternatively, production of **4.2** or **4.3** would need to be indirectly linked to a life-death scenario *in vivo* using complementation-based selection (ex: PACE, CPR).^{37–39} However, there isn't a clear and direct way to do this. While generation of a fluorescence signal would offer the ability to screen mutants in a high-throughput fashion, there is currently no practical method to selectively link fluorescence to macrocyclization. Thus, high-throughput screening and selection methods could not be used. We resorted to an LCMS based screen, which is generally a lower throughput method compared to the aforementioned.³⁷ However, this method offered several advantages such as the use of an autosampler for automated library processing and separation of products **4.2** and **4.3** through reverse-phase chromatography. The high sensitivity of the mass spectrometer (MS) also enabled us to unambiguously detect and quantify small changes in product formation, ensuring that mutants with improved macrocyclic activity were captured.

To screen the clonal library (**Figure 4.5B**), master plates were pin replicated into 96-deep well plates containing LB media and ampicillin (100 µg/mL) for bacterial growth, protein expression and lysis by freeze-thawing.⁴⁰ The resulting lysate was taken and incubated with **4.1** (1 mM) in potassium phosphate buffer (50 mM, pH 7.4) at room temperature for 24 hours before analysis by LCMS. Area under the curve (AUC) for **4.2** and **4.3** products were preprocessed by MZmine2 from extracted ion chromatograms and plotted as a scatter plot for subsequent analysis of screen data and hit identification (**Figure 4.6**).

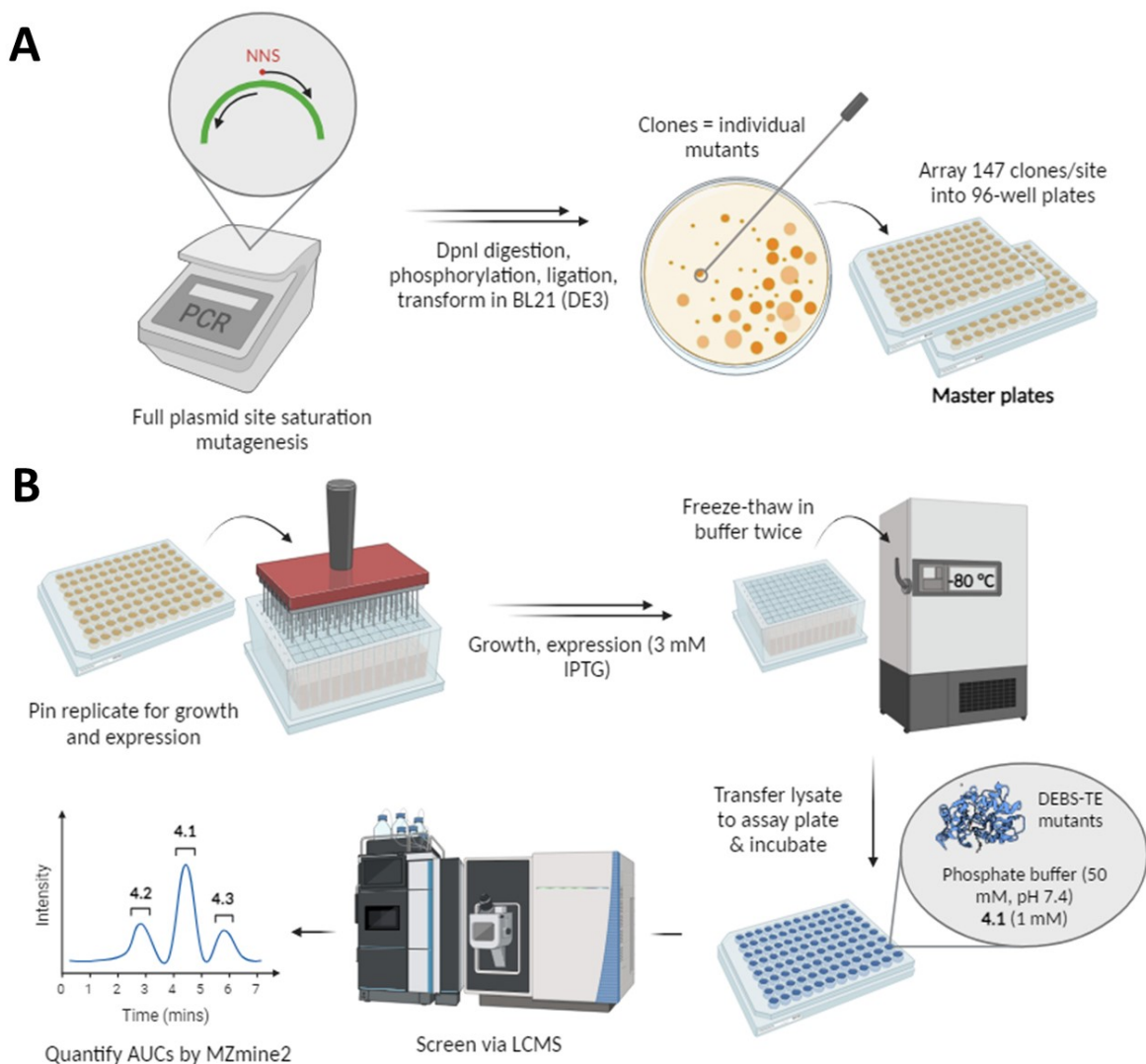
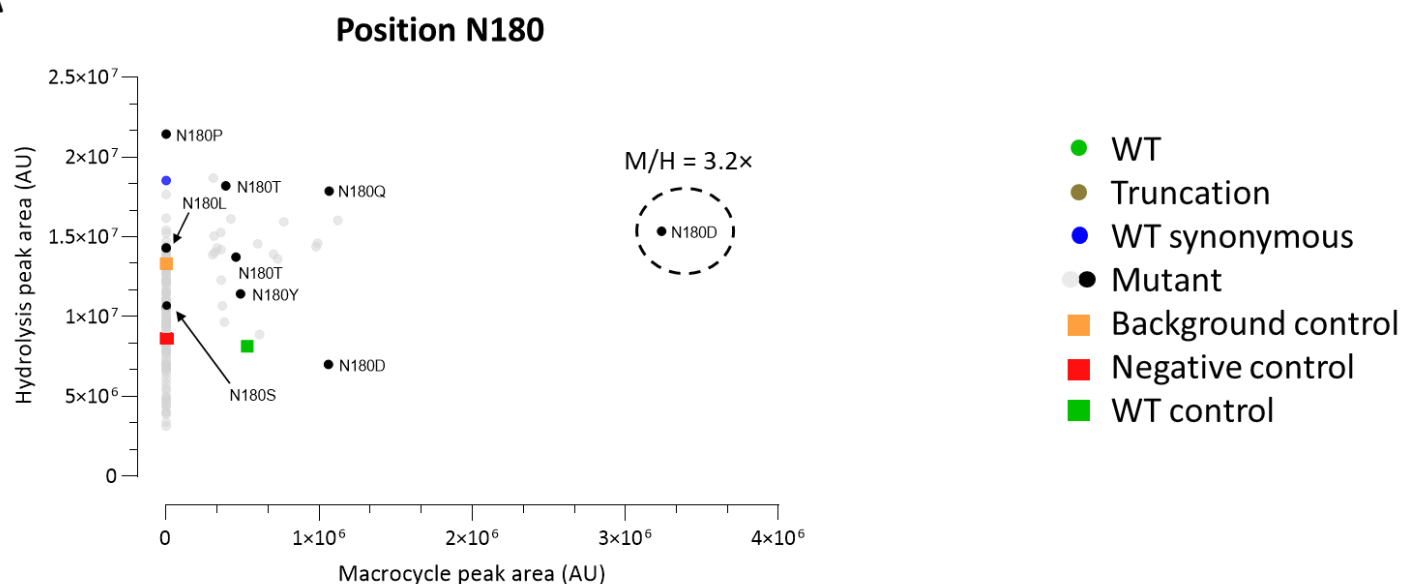


Figure 4.5. Mutant library generation and screen workflow. (A) Mutant library construction via site saturation mutagenesis (full plasmid PCR) of the WT DEBS-TE coding sequence. DpnI digestion removes parental template, while phosphorylation allows for ligation of mutated PCR products and cloning. Transformation of ligated products into *E. coli* BL21(DE3) cells yields clones, each corresponding to a possible DEBS-TE mutant. A total of 147 clones for each position are arrayed into 96-well plates for long-term storage as master plates. (B) Clonal library screen by LCMS. Clones are pin replicated from master plates into 96-deep well plates for growth, expression and lysis by two cycles of freeze thawing in buffer. Cell lysate is transferred into a 96-well plate containing reaction buffer and substrate **4.1** for incubation at room temperature. Assays are quenched, analyzed by LCMS, hydrolysis **4.2** and macrocycle **4.3** product AUC extracted by MZmine2 for analysis and mutant hit identification.

A



B

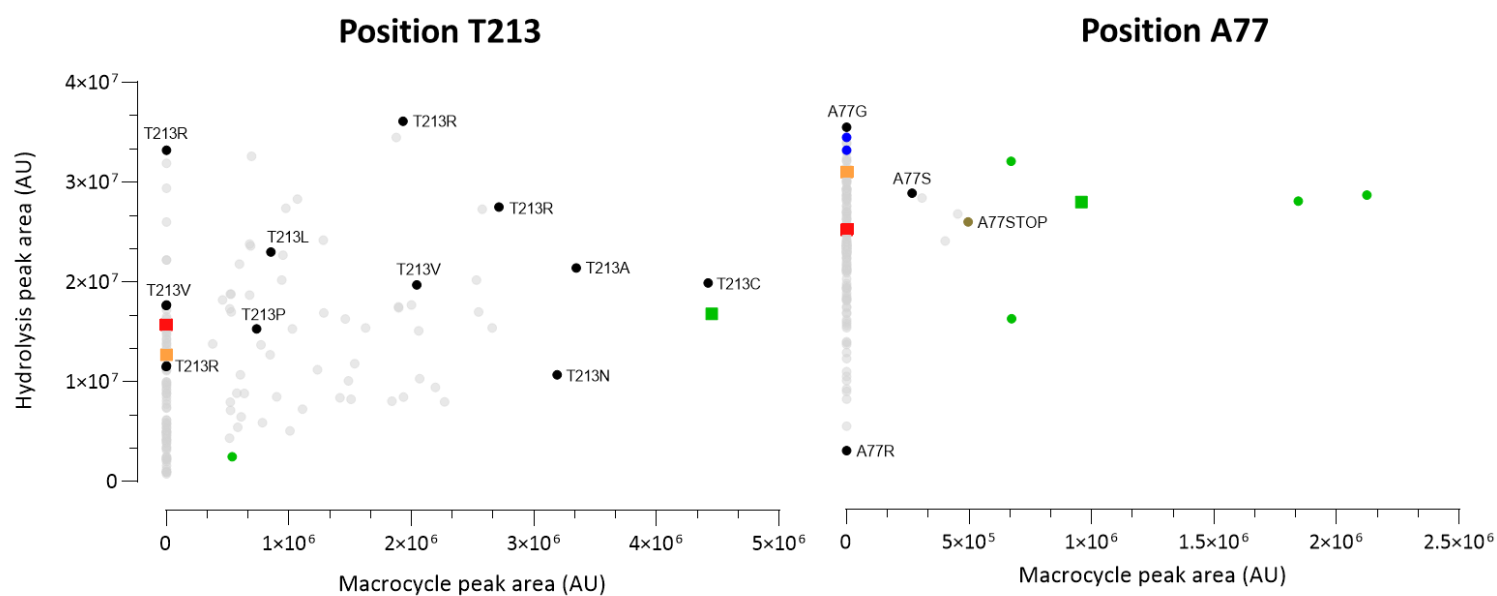


Figure 4.6. Representative scatter plots of library screen. Scatter plots of hydrolysis (4.2) and macrocycle (4.3) peak areas obtained from MZmine2 preprocessing of LCMS data. Symbol and colour legend is shown to the right in panel A. (A) Data for position N180, showcasing a mutant (N180D, circled) with a 3.2-fold increase in macrocycle to hydrolysis product ratio (M/H) over WT. Positions R32 and Q176M showed similar results. (B) Data for positions T213 and A77, showcasing positions retaining or abolishing macrocyclic activity. All other positions showed similar results to position A77.

Based on the data acquired from the screen, 10 clones per position were selected for sequencing. Our selection prioritized clones with the highest M/H ratios (top macrocyclization), although this was not always possible, and those with superior **4.2** AUC over WT (top hydrolysis). Analysis of the screen data pinpointed three variants (R32L, Q176M and N180D) that displayed enhanced macrocyclization activity (**Figure S4.4, S4.7** and **4.6A**). These variants were further confirmed by LCMS analysis of assays with recombinantly purified protein. In general, all positions showed clones with reduced macrocyclization activity alongside a considerably larger group of clones with high hydrolysis activity. This suggests mutagenesis at these positions affect the substrate's ability to adopt a macrocyclization competent conformation and or enhances solvent accessibility to the active site for increased hydrolysis. Moreover, this suggests these positions can undergo mutation without much impact on the fold, and is consistent with a previous mutagenesis study.²⁸ Further studies on these hydrolyzing variants can provide useful insight into how hydrolysis is promoted. Position T213 stands as the sole exception to this trend, as it still retains a substantial population of macrocyclization active clones (**Figure 4.6B**).

Our screen suffered from variability and yielded a negative Z' (**Figure 4.7**), suggesting overlap between positive and negative control signals.^{41,42} While there is essentially no variability in the negative control data, given there is no macrocycle formed, there is large variability in the positive control data which creates a large standard deviation (σ). The Z' calculation uses the 3σ limit for the positive and negative control data, which if normally distributed should include >99.7% of the data. Thus, a Z' between 0.5 and 1 indicates that the positive and negative control data will be easily distinguishable and thus the screen will be effective. Our negative Z' value suggests it will be challenging to distinguish between the positive and negative control data; however, a student's t -test shows that the positive and negative controls are statistically distinguishable (**Figure 4.7A**). A more careful examination of the positive control data shows the data to be non-normal with extreme values. In non-normal datasets, the σ can be

significantly skewed by outliers. Log transformation of the data is often used to minimize the effect of outliers and better approximate a normal distribution. When using the log transformed control data for the macrocycle peak area, the $Z' = 0.23$. This is likely an underrepresentation of the Z' as the negative control data for macrocycle area was set at the minimum cutoff area for calling a peak in our MZmine2 data analysis workflow, which significantly overestimates the magnitude of this area thus decreasing Z' . Regardless a Z' between 0 and 0.5 indicates a functional, though marginal assay. When comparing our positive control to our hits, they were shown to be statistically indistinguishable (**Figure 4.7B**). While the screen was deemed functional, this suggests our screen struggles in identifying improved macrocyclization variants. This is likely a result of the non-normal distribution of our positive control dataset, which skews the Z' and the p value from the student's t -test.

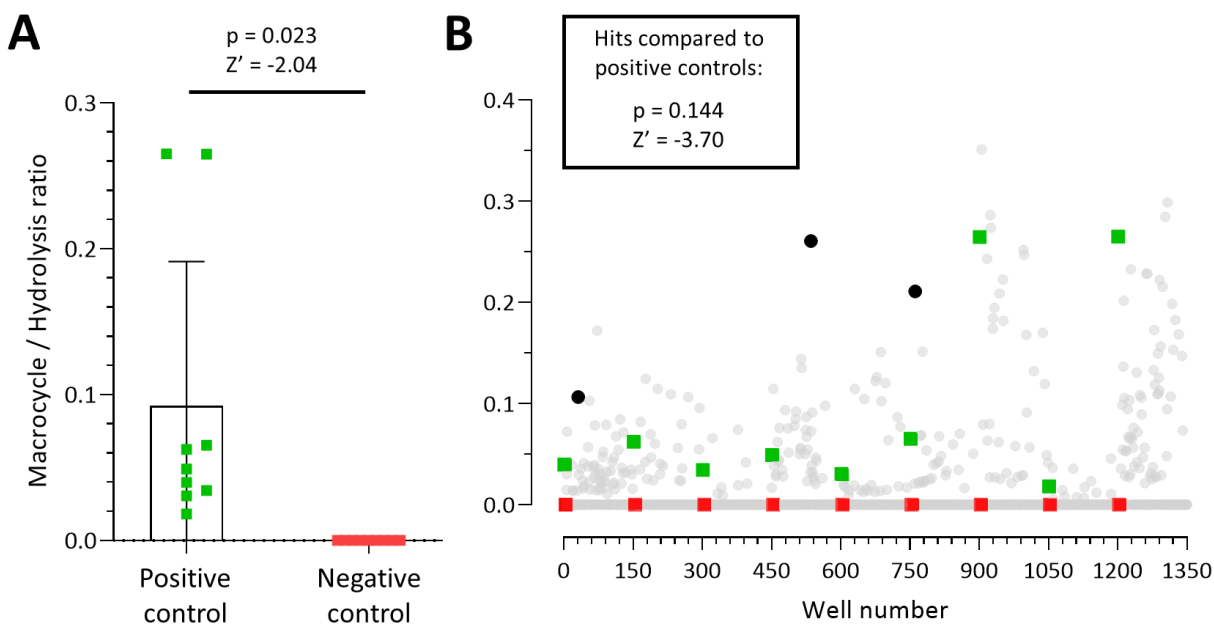


Figure 4.7. Statistical analysis of library screen. (A) Comparison of macrocycle to hydrolysis (M/H) ratio of the screen's positive and negative controls. (B) Scatter plot of M/H ratio for all mutants (grey dots) analyzed during the screen. Black dots represent the top macrocyclization mutant hits while the green and red squares represent the positive and negative controls respectively. In both panels, statistical significance was assessed and the Z' calculated according to equation [5] from Zhang *et al.*⁴¹

To increase Z' , it is necessary to reduce variability in the assay. The presence of two outliers in our positive controls and the observed scattering of WT clones (green dots) from the positive control (green square) at position A77 (**Figure 4.6B** and **4.7B**) suggests differences in total protein amount may contribute to variability. This phenomenon is well documented and has been shown to contribute to variability in well-plate-based screening assays due to differences in cell growth, efficiency in cell lysis^{43–45} and freeze-thaw cycling.⁴⁰ To mitigate this, we sought to normalize product AUC to protein amount by two distinct attempts. The first involved dot blot⁴⁶ normalization using an anti-His monoclonal antibody that binds to C-terminally hexa-His (6xHis) tagged mutants. The released chemiluminescence signal could then be normalized to a WT DEBS-TE control signal, similar to housekeeping protein normalization in Western Blotting.^{47,48} The second involved normalization of *E. coli* BL21(DE3)'s endogenously expressed β -galactosidase (*lacZ*) hydrolysis activity of *o*-nitrophenyl- β -galactoside (ONPG)⁴⁹ relative to a positive control. Unfortunately, both attempts proved unsuccessful due to an inadequate dynamic range.

While these techniques take into consideration the concentration of our protein of interest, it does not consider the concentration of endogenous *E. coli* α/β -hydrolases that can hydrolyze **4.1** in the background. Fluctuations in the levels of these background enzymes can change the M/H ratio due to changes in the total amount of **4.2** formed. These variations can be due to clonal variation and cell growth, which are inherently difficult to control. Unfortunately, there is no efficient method to date to reduce this variability in well-plated-based screens.

Given our efforts to reduce variability were unsuccessful, 10 mutants sufficiently distant from each other in the scatter plots (**Figure 4.6**) were selected for sequencing. From the sequenced variants, non-WT protein coding sequences that showed the greatest M/H ratio (ex: N180D, **Figure 4.6A**) were selected as hits for biochemical characterization. In addition, the top hydrolysis mutant was selected for

biochemical characterization. These mutants (2 clones per site) were expressed, purified, and biochemically characterized (mutant code s#, **Table S4.1**).

R32L, Q176M and N180D are better at macrocyclization than WT

To biochemically characterize hits, mutants were overexpressed in *E. coli* BL21(DE3) with 1 mM of IPTG at room temperature. All mutants were purified by metal-affinity chromatography to high purity (>95%, **Figure S4.1** and **S4.2**). Enzymatic assays with 5 μ M of enzyme and 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4) and 12% v/v DMSO were incubated at room temperature for turnover before LCMS analysis. Extracted ion chromatogram AUC ratios of product **4.3** over **4.2** (M/H) were normalized to the WT DEBS-TE (**Table S4.12**) to quantify fold improvement in macrocycle production and to account for differences in product ionization. Within the top macrocyclization group, mutants R32L, Q176M and N180D (subgroup m1) are shown to be better at macrocyclizing **4.1** with M/H ratios 1.7-, 2.0-, and 2.2-fold higher, thereby confirming the observations made during the screen (**Figure 4.8A**). While these mutants exhibit improved **4.3** production, substantial production of **4.2** is still observed (**Figure S4.4, S4.7** and **S4.9**). This outcome aligns with our proposed ICS model, suggesting a dynamic sampling of the acyl-TE intermediate and substantial competing hydrolysis reactivity. Mutants in subgroup m2 are in stark contrast, showing an unaltered or reduction in M/H ratios (**Figure 4.8A**). Unlike subgroup m1, variants in subgroup m2 were always proximal to the WT in the screen (**Figure 4.6B, S4.5, S4.8** and **S4.10-11**), which may account for the unaltered or reduced M/H ratios. Among the top hydrolysis group, none outperformed WT in forming **4.3** as expected. These observations further validate the screening strategy and suggest that even with the negative Z', the screen can help distinguish hits with enhanced macrocyclization activity over those with reduced activity.

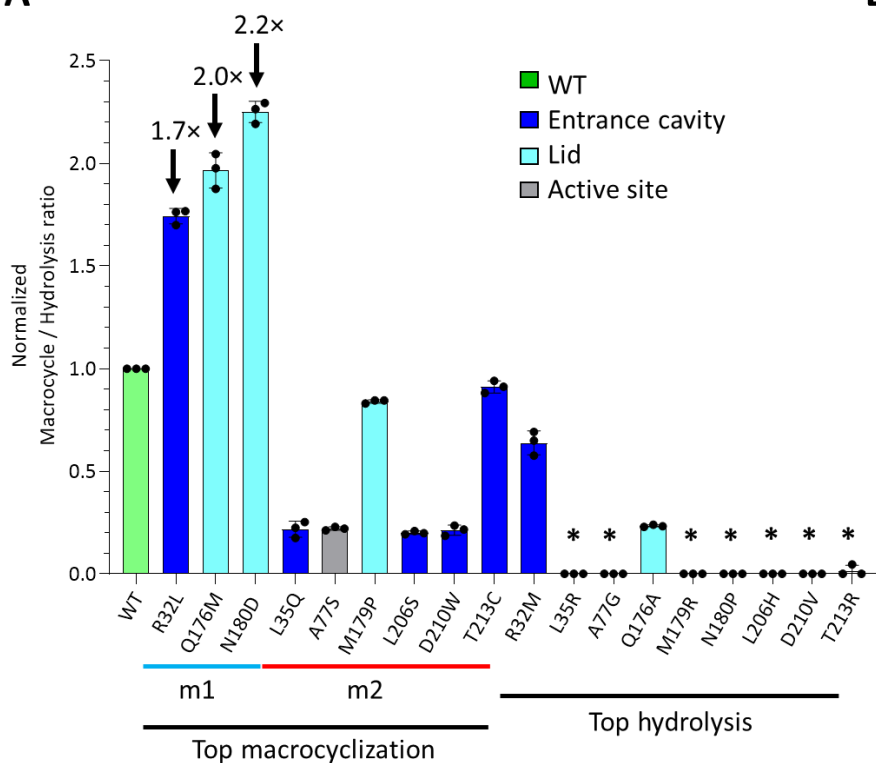
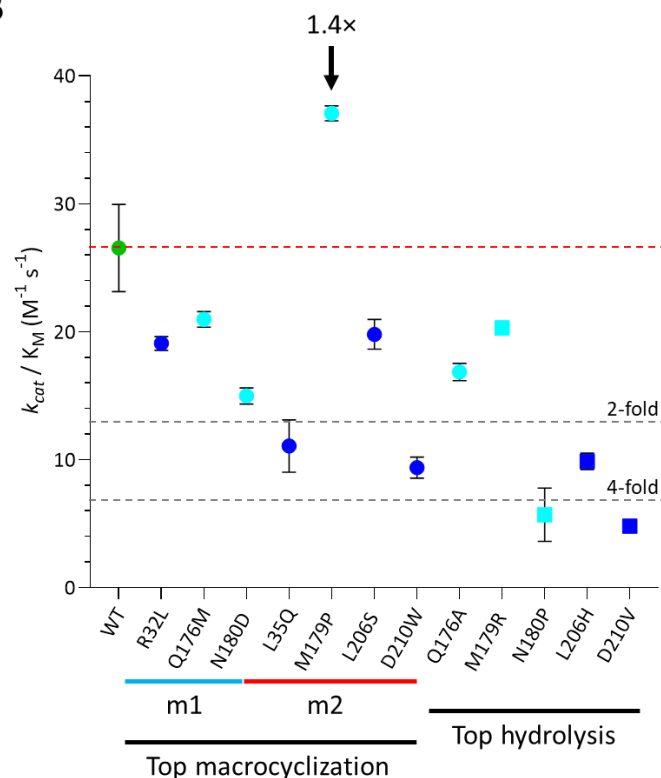
A**B**

Figure 4.8. Enzymatic and kinetic assay results for DEBS-TE single mutants. (A) Macrocycle **4.3** to hydrolysis **4.2** product ratios normalized to WT DEBS-TE product ratio (mean $\pm$ s.d., $n = 3$ technical replicates). Ratios were calculated from AUC of extracted ion chromatograms (see **Table S4.12**). Note mutants labelled with an asterisk are only hydrolytically active. Within the improved macrocyclization mutant group, subgroup m1 (blue line) are true mutants displaying improved macrocyclization activity and identified by arrows with the fold-increase shown. (B) Second order rate constants (k_{cat}/K_M) for WT DEBS-TE and mutants (mean $\pm$ s.d., $n = 3$ technical replicates). Grey dashed lines indicate the fold decrease in k_{cat}/K_M over WT. Mutants shown as squares are only hydrolytically active.

Kinetic characterization was performed for all variants and the WT DEBS-TE. For this, enzyme (5 μ M) was incubated with varying concentrations of **4.1** (0-2 mM) in potassium phosphate buffer (50 mM, pH 7.4) and DMSO (12% v/v). Initial velocities (v_0) were acquired by continuous spectrophotometric measurements at 412 nm with the addition of Ellman's reagent (DTNB), which tracks the release of SNAC free thiol upon substrate loading.^{14,16,28,50} Mutants R32M, L35R, A77G and T213R were excluded from

these experiments due to low observed activity in product distribution assays (**Figure S4.41**). In addition, efforts to kinetically characterize mutant T213C were unsuccessful. As DTNB reacts with free thiols, we believe DTNB forms an adduct with active site residue C213, inactivating the TE. Most mutants could not be saturated, even with 1 mM or more of substrate **4.1**, preventing independent determination of the first order rate constant k_{cat} and the Michaelis constant K_M . Consequently, second order rate constants k_{cat}/K_M were reported for a consistent and comparable analysis. The WT DEBS-TE is shown to turnover substrate **4.1** with a $k_{cat}/K_M = 27 \pm 3 \text{ M}^{-1} \text{ s}^{-1}$ (**Figure 4.9** and **S4.14**), which is 22-fold higher than with the previously tested phenyl derivative **4.5** (**Figure 4.9**)¹⁵ but comparable to the related PIK-TE domain.⁵¹ In almost all cases, the k_{cat}/K_M remains largely unchanged with a few being reduced ~2- or ~4-fold (**Figure 4.8B**). Acylation, or TE loading, is often rate determining *in vitro*,^{28,31} consequently k_{cat}/K_M is the second order rate constant describing TE loading. Because off-loading occurs after the rate-determining step, these values have limited relevance in the context of macrocyclization. Furthermore, our screen focused on enhancing macrocyclization activity without any consideration for the kinetics, thus these changes in k_{cat}/K_M are not surprising. The exception to this trend is mutant M179P which shows a 1.4-fold increase in k_{cat}/K_M (**Figure 4.8B**). Interestingly, variants in subgroup m1 have higher product to substrate ratios (P/S) despite having reduced k_{cat}/K_M values compared to WT (**Table S4.2**). This suggests these mutants are longer lived catalysts, enabling them to outperform the WT in product formation. Further investigation through enzyme stability experiments such as differential scanning calorimetry (DSC),⁵² isothermal titration calorimetry (ITC)⁵³ or circular dichroism (CD)⁵⁴ is warranted.

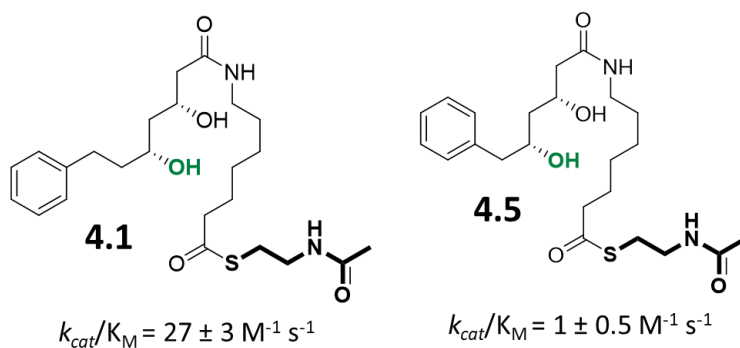


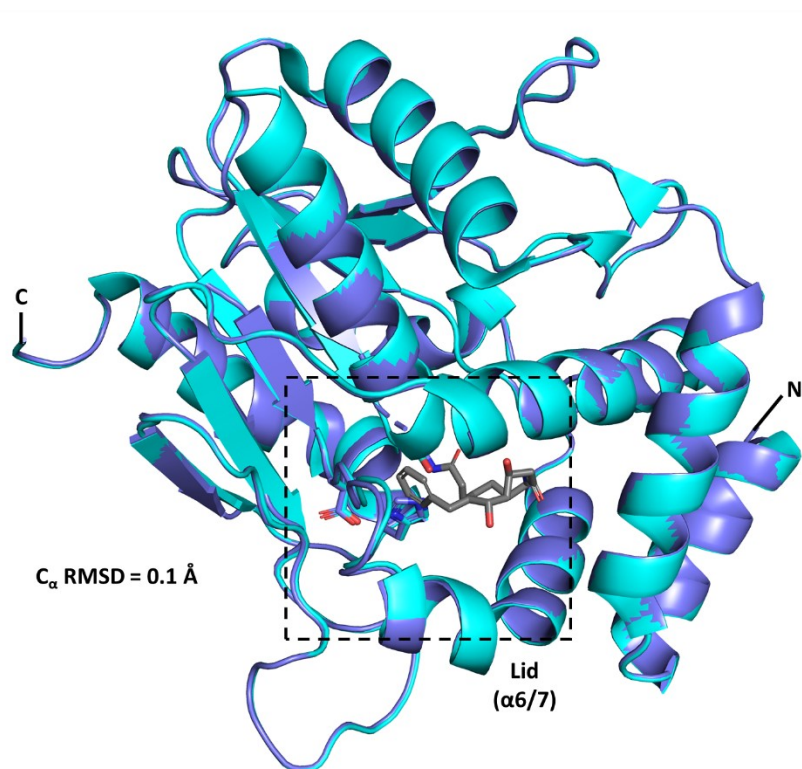
Figure 4.9. Comparison of substrate 4.1, 4.5 structures and WT DEBS-TE second order rate constants. Errors indicate s.d. of technical replicates ($n = 3$). Second order rate constant for substrate **4.5** was obtained from Pinto et al.¹⁵

WT DEBS-TE trapped acyl crystal structure

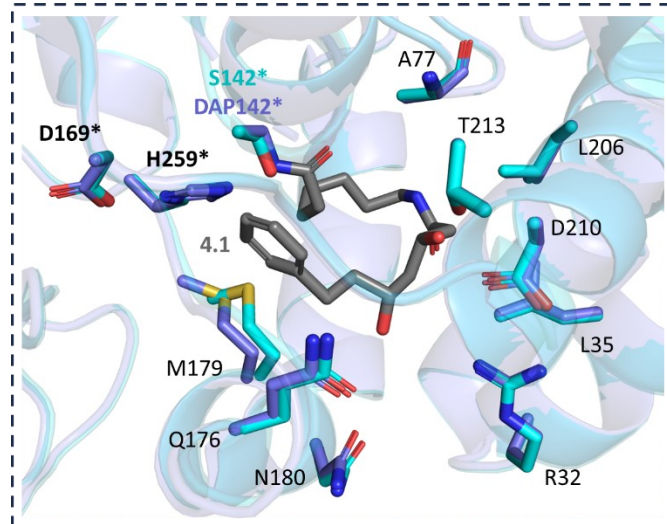
Our collaborators in Dr. Schmeing's lab were able to acquire a high-resolution (1.4 Å) crystal structure of the WT DEBS-TE with **4.1** trapped as an acyl-TE intermediate (**Figure 4.10**). Comparing this structure with the previously acquired high-resolution structure of the *apo* DEBS-TE (PDB: 5D3K, 1.7 Å resolution) show little difference in their conformations, including the lid region (C_α RMSD = 0.1 Å, **Figure 4.10A**). This is surprising as the lid is known to be a mobile region that interacts with TE substrates. Similarly, comparison of the chosen active site residues for mutagenesis shows little change in their rotameric configuration. Only residue M179's side chain is shifted 1.8 Å away from **4.1**, presumably due to repulsion between the benzyl moiety and the sulfur atom (in M179) electron clouds (**Figure 4.10B**). Although substrate **4.1** is shown to be in a cyclic conformation, it is not catalytically relevant for macrocyclization as the C13 alcohol is not optimally placed for ring closure. In addition, the proposed backbone NH group of A77, presumed to be part of the oxyanion hole, is not optimally placed for stabilization of the tetrahedral intermediate (**Figure 4.10C**). This could simply be an artifact of crystallization and warrants further investigation. Despite this, the contacts observed in this structure may

provide some useful insights on TE-substrate interactions that generate a cyclic conformation conducive to macrocyclization.

A



B



C

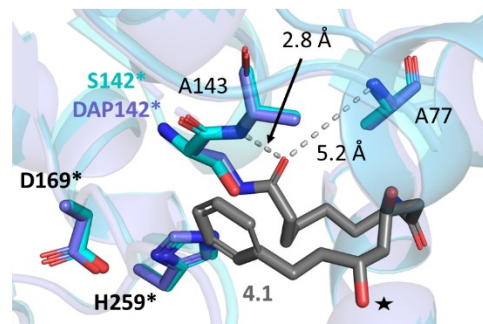


Figure 4.10. WT DEBS-TE trapped acyl crystal structure compared to *apo* DEBS-TE. (A) WT DEBS-TE in complex with **4.1** (1.4 Å resolution, purple cartoon) acquired by our collaborators in Dr. Schmeing's lab aligned with *apo* WT DEBS-TE (PDB: 5D3K, 1.7 Å resolution, cyan cartoon). Overall conformation of both TE structures is very similar (C $_{\alpha}$ RMSD = 0.1 Å). (B) Zoom-in of the active site, showing little difference in rotameric configuration of catalytic triad and our chosen sites for mutagenesis. M179 side chain is the only shown to shift by 1.8 Å in the **4.1** complexed structure (purple sticks). (C) Oxyanion hole H-bonds (grey dashed lines) stabilizing the tetrahedral intermediate with the distances shown. The C13 alcohol for ring closure is labelled with a black star. In all panels, catalytic triad is shown as sticks and labelled with an asterisk while **4.1** is shown as grey sticks.

Proposed mutation effects to rationalize changes in macrocyclization activity

Examining the impact of our single mutations on the DEBS-TE structure can provide crucial insight into DEBS-TE mediated macrocyclization. Currently, crystal structures of these mutants remain unavailable, although efforts are ongoing to obtain high-resolution structures of non-hydrolyzable acyl-TE intermediates of **4.1** with these mutants.⁵⁵ In the absence of these structures, we analyze our data based on the recently acquired trapped acyl-TE crystal structure of the WT DEBS-TE, generously provided by our collaborators in Dr. Schmeing's lab, to formulate hypotheses on the impact of these mutations.

R32 is shown to form two H-bonds with D210, positioning $\alpha 2$ relative to $\alpha 10$. In addition, a water mediated H-bonding interaction is formed with the C13 alcohol of **4.1** (**Figure 4.11**). Mutating to a neutral Leu likely breaks these H-bonds and allows for helices $\alpha 2$, $\alpha 10$ and $\alpha 6/7$ (i.e. the lid) to draw closer and maximize hydrophobic packing interactions. As these structural elements dictate the shape of the entrance cavity, this mutation can lead to improved sampling of closed conformations that can impede water's access to the active site through sterics (i.e. solvent gating), reducing hydrolysis. These changes can impact the active site's structure through the lid, which is known to shape the active site.^{9,10,12,13} This alteration can favour compact conformations of the acyl-TE intermediate, effectively biasing its sampling towards conformations conducive to macrocyclization. Moreover, the water mediated H-bonding interaction with **4.1** is also likely broken. This water mediated interaction likely stabilizes the unreactive cyclic conformation of **4.1**, thus breaking these interactions could promote dynamic sampling of the acyl-TE intermediate. All of these taken together can explain the improvement in M/H ratio.

The complementary mutation at R32 that led to a drop in macrocyclization activity was R32M, which similarly disrupts the interaction with D210 and with **4.1**. Its subtly lower hydrophobicity as compared to the Leu⁵⁶ mutant (R32L) presumably leads to a similar effect, though with less restriction on water's access to the active site and thus it maintains an M/H ratio on par to the WT.

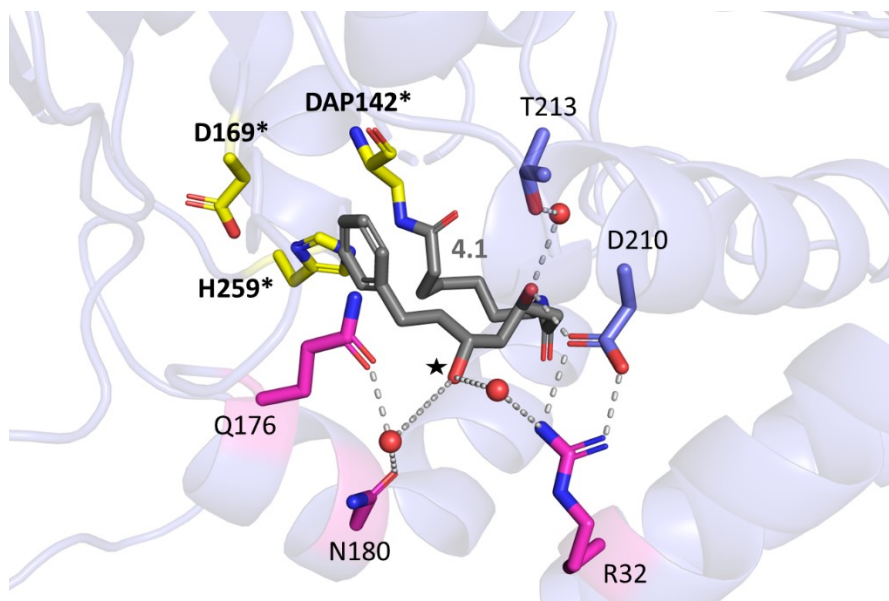


Figure 4.11. Interaction network of the DEBS-TE active site in complex with **4.1**. Zoom-in of the DEBS-TE (1.4 Å resolution, purple cartoon) active site in complex with **4.1** (grey sticks) with catalytic triad shown as yellow sticks and waters shown as red spheres. A series of H-bonding interactions (grey dashed lines) is formed with R32, Q176 and N180 residues (magenta sticks) with water, **4.1** and other active site residues (purple sticks). The C13 alcohol for ring closure is labelled with a black star. Structure was acquired by our collaborators in Dr. Schmeing's lab.

Mutation of R32's binding partner (D210) leads to a substantially different outcome. Both the better macrocyclizing D210 mutant, D210V, and the worst macrocyclizing mutant, D210W, exhibit some of the lowest M/H ratios. In the case of D210V, the mutation also resulted in a hydrolytic variant with the lowest k_{cat}/K_M . While disruption of the D210-R32 interaction via both D210V and D210W should enable helices α_2 , α_{10} and $\alpha_{6/7}$ to similarly restrict the entrance cavity, the substantial drop in M/H ratio and catalytic activity suggests a much more disruptive change to the active site. Further examination of the DEBS-TE non-hydrolyzable acyl-TE crystal structure shows D210 directly interacts with the distal C11 alcohol of **4.1** via its side chain carbonyl (**Figure 4.11**). We propose this interaction is required for the adoption of a cyclic conformation of the acyl-TE intermediate. This would be consistent with majority of the clones displaying very little macrocyclization activity in the screen (**Figure S4.11**).

Q176 forms a water mediated H-bonding interaction with the C13 alcohol of **4.1** (**Figure 4.11**). Disruption of this interaction can destabilize the unreactive cyclic conformation of **4.1**, providing a similar effect as R32L. In addition, we propose the large hydrophobic Met mutation provides an additional hydrophobic barrier to limit waters access to the active site. Consistent with this analysis is the observation that the Ala mutant has a substantially reduced M/H ratio, as would be expected from the smaller side chain.

Lastly, N180 also forms a water mediated H-bonding interaction as Q176 (**Figure 4.11**). The N180D mutation can potentially form a long-range charge-charge interaction⁵⁷ with R32 that can prompt both $\alpha 6/7$ and $\alpha 2$ to adopt a more closed conformation. This can displace the water molecules mediating H-bonding interactions between the nucleophilic OH groups with R32, Q176 and N180, thus resulting in an effect similarly described for the R32L and Q176M mutations. In contrast to the Asp mutation, the Pro substitution results in an exclusively hydrolytic variant. We propose that the destabilizing nature of Pro in α -helical structures likely induces a helix break or kink^{58–61} at position 180, shortening the lid helix which biases sampling towards open conformations that increase solvent access to the active site. This is consistent with the dramatic difference in M/H between the Asp and Pro mutants.

While the mutants investigated at the other 6 sites did not yield any improvement in macrocyclization activity, some useful information can still be inferred. For instance, the bulky side chains of L35 and L206 are shown facing the substrate and likely act as hydrophobic barriers, promoting the adoption of a curled acyl-TE intermediate conformation (**Figure 4.12**). This is consistent with many of the L35 and L206 clones having largely reduced macrocyclization activity (**Figure S4.5** and **S4.10**). The backbone NH of A77 is likely part of the oxyanion hole (**Figure 4.10C**) and thus would explain the little macrocyclization activity observed in the screen (**Figure 4.6B** and **S4.6**) and the little activity observed for the purified A77S/G variants (**Figure 4.8A** and **Table S4.12**). T213's β -OH group was shown to indirectly H-

bond to **4.1**'s distal OH group with water (**Figure 4.11**). As this distal OH group has a direct interaction with D210, breaking T213's interaction would presumably have little impact on the macrocyclization activity of the mutants. This is consistent with the retention in macrocyclization activity observed in the screen of T213 clones (**Figure 4.6B** and **S4.12**). Thus, we propose residues L35, A77 and L206 as crucial for the macrocyclization activity of the DEBS-TE while T213 is a tolerant position for mutagenesis.

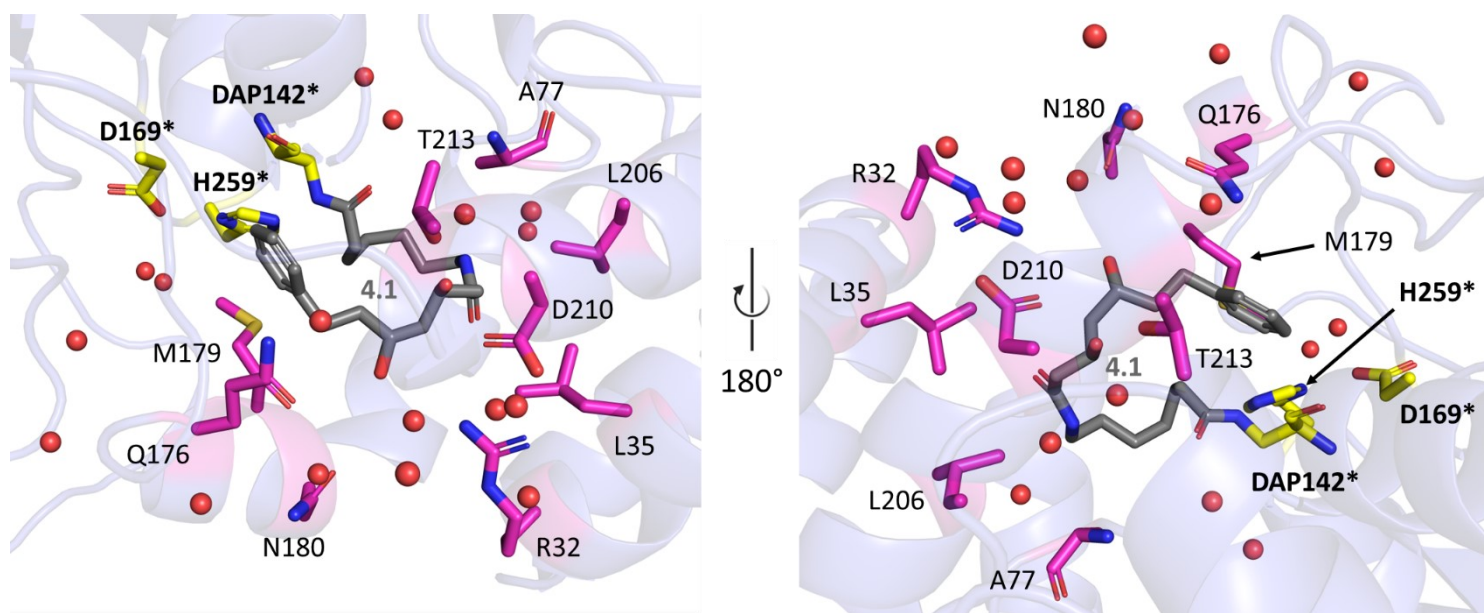


Figure 4.12. DEBS-TE active site in complex with substrate 4.1. Two views of the DEBS-TE (1.4 Å resolution, purple cartoon) active site in complex with **4.1** (grey sticks) with catalytic triad shown as yellow sticks, waters shown as red spheres and all mutated sites shown as magenta sticks. Structure was acquired by our collaborators in Dr. Schmeing's lab.

The hypotheses put forward to rationalize the changes in activity for the mutants have yet to be verified and necessitate further investigation either via molecular dynamic simulations (MDS) or via trapped acyl-TE crystal structures of these mutants using DAP incorporation technology.⁵⁵

Combination of single mutations do not exhibit synergistic effects

Having pinpointed R32L, Q176M and N180D as our top-performing macrocyclization variants, we investigated the potential synergy of these mutations by creating a combinatorial library. We also included the M179P mutation in our approach, due to its improved kinetics. Thus, we hoped to craft a better macrocyclization mutant with enhanced kinetics. We performed site directed mutagenesis³⁴ to generate all possible double, triple and quadruple mutants (11 total, mutant code c#, **Table S4.1**). All mutants were transformed and expressed in *E. coli* BL21(DE3) chemically competent cells and purified by metal-affinity chromatography to high purity (>95%, **Figure S4.1** and **S4.2**). Following purification, enzymes were incubated with **4.1** and analysed by LCMS exactly as described for the single mutants.

In general, the individual mutations are not synergistic and instead lead to a reduction in the M/H ratio (**Figure 4.13A**). This is unsurprising as combining single mutants rarely leads to improved variants, due to cryptic and complex interactions embedded in protein structures.^{62,63} Double mutant c46 (R32L/Q176M) retains a similar fold improvement in M/H ratio as its single mutant counterpart Q176M (2.4-fold increase), making it the sole exception to the previous trend. Because R32L is postulated to reduce the size of the entrance cavity, the addition of the Q176M mutation likely further accentuates this structural alteration by negating the potential electrostatic repulsion with D210 on $\alpha 10$. This could promote the adoption of closed conformations for effective solvent gating and can promote compact acyl-TE intermediate conformations. In addition, two water mediated H-bonding interactions with the C13 alcohol of **4.1** are likely broken, allowing for better sampling of conformations conducive to macrocyclization, thus increasing the M/H ratio. In all other double mutants, the reduction in M/H ratios suggests the structural alterations or changes in conformational dynamics induced by the individual mutations are incompatible when combined.

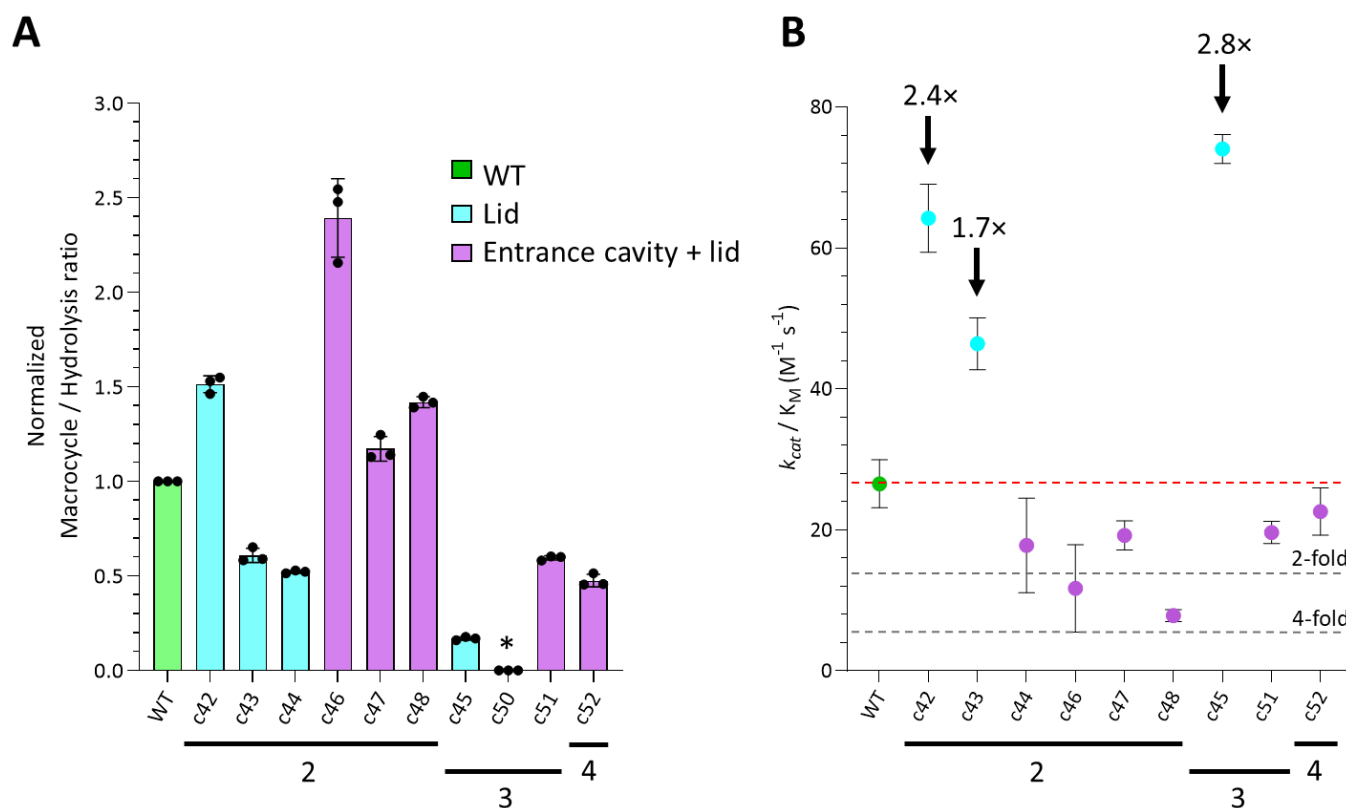


Figure 4.13. Enzymatic and kinetic assay results for DEBS-TE combinatorial mutants. Combinatorial mutants were generated by combining R32L, Q176M, M179P and N180D single mutations. The number shown below indicates the number of combined mutations. (A) Macrocycle **4.3** to hydrolysis **4.2** product ratios normalized to WT DEBS-TE product ratio (mean $\pm$ s.d., $n = 3$ technical replicates). Ratios were calculated from AUC of extracted ion chromatograms (see **Table S4.12**). Mutant labelled with an asterisk is only hydrolytically active. Only mutant c46 retains a M/H ratio fold improvement comparable to its single mutant counterparts (arrow, 2.4-fold). (B) Second order rate constants (k_{cat}/K_M) for WT DEBS-TE and combinatorial mutants (mean $\pm$ s.d., $n = 3$ technical replicates). Mutants with improved k_{cat}/K_M are identified by arrows with the fold-increase over WT shown. Grey dashed lines indicate the fold decrease in k_{cat}/K_M over WT.

In terms of kinetics, most mutants k_{cat}/K_M remain largely unchanged or reduced ~2-fold relative to the WT (**Figure 4.13B**). While the structural changes conferred by the individual mutations decrease the M/H ratio, their effects are minimal on the k_{cat}/K_M values. Out of all combinatorial mutants, c48 exhibited the slowest loading kinetics with a $k_{cat}/K_M = 8 \pm 0.8 \text{ M}^{-1} \text{ s}^{-1}$ (**Figure S4.32**). This suggests the structural alterations of mutations R32L and M179P affect the TE's catalytic rates. M179 is situated near the end of the lid helix, linked to a loop where an Y171-W242 H-bond which holds the loop containing the catalytic base (H259) in place.^{9,28} Introduction of a Pro at this position likely breaks the helix, enhancing loop flexibility and potentially displacing the catalytic base, thereby decreasing catalytic rates without much change in the M/H ratio. Surprisingly, the combinatorial mutants c43, c42 and c45 exhibited 1.7-, 2.4- and 2.8-fold improvement in k_{cat}/K_M over WT (**Figure S4.27-28** and **S4.33**). These mutants are double (c42-43) or triple (c45) mutants of the Q176M, M179P and N180D single mutations situated on the lid helix. We reasoned these mutations greatly alter the conformation of the lid that allows better substrate access to the active site and thus faster loading in addition to greater solvent access for a reduced M/H ratio relative to c46. Notably, our best performing double mutant (R32L/Q176M) exhibited reduced product to substrate ratio and a ~2.3-fold reduction in k_{cat}/K_M compared to WT (**Table S4.2**). This observation suggests the double mutation is destabilizing, yielding a much shorter-lived catalyst. Surprisingly, the quadruple mutant c52 retained a similar second order rate constant ($k_{cat}/K_M = 23 \pm 3 \text{ M}^{-1} \text{ s}^{-1}$, **Figure S4.35** and **Table S4.2**) to that of the WT. However, qualitative assessment of the progress curves during kinetic measurements with 1 mM of **4.1** showed it deviated from linearity much earlier than anticipated (**Figure S4.36**). This suggests irreversible inactivation of the mutant, presumably due to lower enzyme stability. Despite mutant c52 appearing less stable with loading kinetics similar to the WT, it demonstrated a 1.8-fold increase in product to substrate ratio over WT (**Table S4.2**). A likely explanation for this could be the mutant is much faster at off-loading than the WT and warrants further investigation.

Improved macrocyclization mutants are kinetically slower

When examining the M/H ratios of our top performing single mutants alongside their corresponding second order rate constants, an inverse Pearson and Spearman correlation became evident (**Figure 4.14**). This pattern is consistently observed with the combinatorial mutants, suggesting that better loading kinetics is detrimental to the production of **4.3**. Because solvent gating is likely a result of the mutation sites positioning at the entrance cavity, there is also a potential for substrate gating. For example, the R32L mutant induces a closed entrance cavity conformation through the approach of $\alpha 2$, $\alpha 10$ and the lid. This can gate both solvent and substrate, leading to a reduction in loading kinetics and a reduction in hydrolysis rates. Our proposed structural hypotheses would also suggest a more tightly packed active site of the *apo* TE. This would require more drastic structural rearrangements for binding and TE acylation, which can also contribute to the reduced kinetics. While these hypotheses are well in line with the data shown, the low R^2 in the Pearson correlation and the statistical insignificance of the Spearman correlation suggests additional complex factors contribute to the changes in both parameters.

Substrate binding with the DEBS-TE is believed to be influenced by hydrophobic packing interactions,^{9,28–30} potentially guiding the off-loading chemistry. While all hypotheses regarding the structural impacts the single mutations and their combinations support the above postulate and are consistent with the results presented herein, there is no structural data to support these. Our current efforts at acquiring trapped acyl-TE intermediate crystal structures⁵⁵ of our mutants with **4.1** will hopefully become available soon for further insights into TE off-loading chemistry selectivity and allow for the verification of our proposed hypotheses.

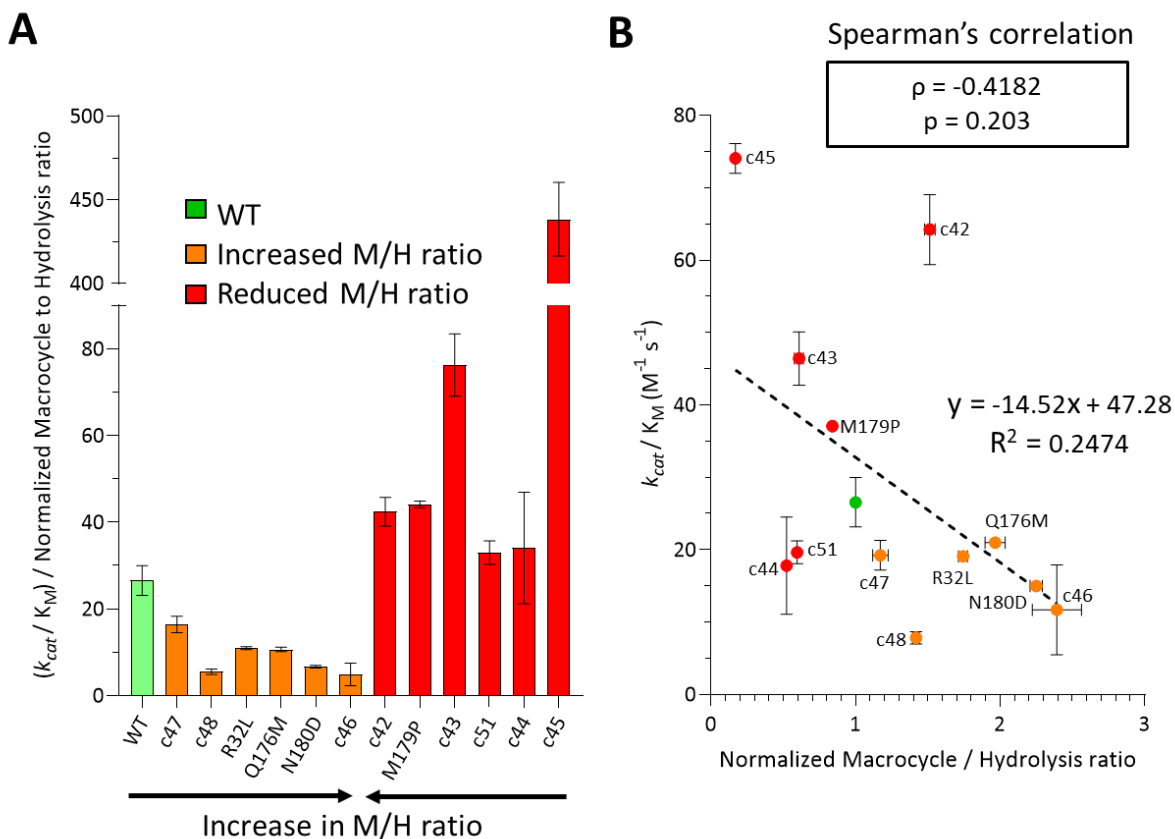


Figure 4.14. Relationship between k_{cat}/K_M and normalized M/H ratio. Relationship analysis used k_{cat}/K_M and normalized macrocycle to hydrolysis (M/H) ratio data of combinatorial variants, our top performing single variants (R32L, Q176M, M179P and N180D) and the WT. (A) Bar graph of k_{cat}/K_M over M/H ratio. Note the top performing macrocyclization variants have the smallest ratios. (B) Scatter plot of k_{cat}/K_M versus M/H ratio, showcasing a negative Pearson correlation (dashed line) and Spearman correlation (ρ).

4.4 Conclusions

To conclude, the work presented in this chapter demonstrated the ability to engineer the DEBS-TE for improved formation of macrocycle **4.3** when incubated with substrate **4.1**. We employed a rational protein design approach through site saturation mutagenesis of 9 residues that line the entrance cavity and the active site, yielding a comprehensive library of 1,323 clones. An LCMS based screen was performed, which revealed variants R32L, Q176M and N180D exhibited enhanced production of **4.3** over WT. The screen suffered from variability and a low Z' , presumably due to differences in total protein amount. In the future, generation of a joint DEBS-TE and superfolder GFP (sfGFP) construct would allow for protein normalization to minimize 96-well plate screening variability and improve the Z' .⁴³

Biochemical characterization of enzymes confirmed R32L, Q176M and N180D as the top macrocyclizing variants with 1.7-, 2.0- and 2.2-fold increase in **4.3** over **4.2** (M/H) ratio but with reduced loading kinetics relative to WT. Despite the pitfalls of the screen, these experiments confirm them as true positives, effectively corroborating the interpretations drawn from the screen. A combinatorial library was generated with the top performing macrocyclization variants in addition to mutation M179P, showing a 1.4-fold increase in kinetics over WT, for possible mutation synergy. None of the mutations in combination showed synergy. However, R32L/Q176M was identified as another top macrocyclization variant with a 2.4-fold increase in M/H ratio.

We currently hypothesize that the top performing mutations alter the structural landscape of the TE to favour conformations with a closed entrance cavity and altered active site that can (1) gate solvent's access to the active site and (2) bias acyl-TE intermediate conformations towards those conducive to macrocyclization while breaking key H-bonding interactions that stabilize an unreactive cyclic conformation of **4.1**, all of which is consistent with the data presented. These proposed hypotheses would need to be structurally verified with trapped acyl-TE crystal structures using DAP incorporation

technology⁵⁵ or investigated via MDS. While there was an apparent correlation between enhanced macrocyclization activity and slower loading kinetics, the statistical insignificance in these correlations suggests more complex factors are at play.

Finally, efforts in obtaining trapped acyl-TE crystal structures of our mutants are ongoing with substrate **4.1** and will be pivotal in verifying our proposed hypotheses. In addition, these will provide crucial insights into the discrete substrate-enzyme interactions that drive substrate binding and off-loading. With the acyl-TE crystal structure of the WT in hand, MDS can be performed to investigate possible sites important for macrocyclization activity along with the importance of solvent on hydrolysis rates. Regardless, this study deepened our mechanistic understanding in DEBS-TE catalyzed macrocyclization and provided the second engineered PKS TE with improved macrocyclization activity.

4.5 Materials and methods

Mutant library construction and storage

Construction and cloning of mutant library. Site saturation mutagenesis³⁴ was performed on the WT DEBS-TE sequence found within the pRSG33 plasmid.¹⁶ Primers shown in **Table 4.1** were used and reactions were set in duplicate for each primer pair. Each reaction contained the following: the correct forward and reverse primer pair (both 0.5 μ M), Q5 Reaction Buffer (1X), dNTPs (200 μ M), pRSG33 template (1 ng), Q5 High-Fidelity DNA Polymerase (0.5 units, #M0491 NEB). Negative controls were prepared identically except without the addition of polymerase. Full plasmid PCR was performed for every reaction as per the manufacturer's recommended conditions, with an annealing temperature of 74°C for 50 seconds.

Amplification of DNA by PCR was confirmed via gel electrophoresis and every reaction purified using the E.Z.N.A. Cycle-Pure Kit (#D6492, Omega Bio-tek) as described in their product manual followed by an overnight DpnI digestion (10 units, #R0176 NEB) at 37°C. Digested products were then purified once more using the E.Z.N.A. Cycle-Pure Kit as described in their product manual. Eluted products were phosphorylated for 1 hour at 37°C and ligated for 2 hours at room temperature before their transformation into *E. coli* BL21(DE3) chemically competent cells. Transformed cells were plated onto LB agar ampicillin (100 μ g/mL) plates and incubated overnight at 37 °C to obtain colonies the next day. Colonies were used for library screening in 96-well plates. Each phosphorylation reaction contained the following: 10 μ L of the eluted product, T4 DNA Ligase Buffer (1X, #B0202 NEB), T4 Polynucleotide Kinase (2.5 units, #M0201 NEB). T4 DNA Ligase (5 units, #M0202 NEB) was added to the same phosphorylation reactions prior to the 2-hour ligation incubation. A negative control was also prepared and treated similarly as described above except without the addition of T4 PNK and T4 DNA ligase to account for background colonies after transformation.

Table 4.1. Saturation mutagenesis primers used for library construction. Primer pairs were designed for full-plasmid PCR with a T_m of ~68-69°C and ordered from Integrated DNA Technologies with standard desalting. All primers are shown 5' to 3' with the degenerate codon underlined. In each row, the top sequence is the forward primer whilst the bottom sequence is the reverse primer.

Position	Oligos (Forward/Reverse)
R32	<u>NN</u> STCCTACCTCGACCTGCTGGCG GACCCTGCCCCGACACGCCCCG
L35	<u>NN</u> SGACCTGCTGGCGGGGCT GTAGGACCGGACCCTGCCCCGACACG
A77	<u>NN</u> SGCGATCTCCGGTCCGCAC CGTTCCCGCGCAGCAGATCACCG
Q176	<u>NN</u> SGACGCGATGAACGCCTGGCT GTGACCGGGCGGGTAGACGTCG
M179	<u>NN</u> SAACGCCTGGCTGGAGGAGCTGAC CGCGTCCTGGTGACCGGGC
N180	ATCTCGATCCCGCGAAATTAAT <u>S</u> NNCATCGCGTCCTGGTGACCGGGC
L206	<u>NN</u> SGGCGCCTACGACCGCCT GGCGGTGAGCCTGGTGTCGTCC
D210	<u>NN</u> SCGCCTCACCGGTCACTGGCG GTAGGCGCCCAGGGCGGTG
T213	<u>NN</u> SGGTCACTGGCGACCCCG GAGGCGGTCTAGGCGCCAG

Mutant library preparation and storage. Conical 96-well microtiter plates (#249944, ThermoFisher Scientific) were sterilized in an autoclave prior to use. A total of 147 (not including controls) mutants were used per position for >99% library coverage (~4.6× oversampling, 9 positions for total library size n = 1,323 clones).³⁵ Each well was filled with 200 µL of sterile LB media supplemented with 100 µg/mL of ampicillin and a single colony from the appropriate plate was then inoculated into 92 wells of one well plate (M1) and 55 into another well plate (M2). The first well of plate M1 was never inoculated to ascertain sterility. Separate pET-21c(+) (#69770, Millipore Sigma), pRSG33 and pJBH09 transformed *E. coli* BL21 (DE3) colonies were inoculated into the next 3 wells of plate M1 to serve as background, positive and negative

controls respectively. Plates were then covered with a sterile porous film (#60941-086, VWR) and incubated for 24 hours at 37°C with 180 rpm shaking. Wells were topped off with LB supplemented with 100 µg/mL of ampicillin and 30% w/v glycerol to a final volume of 300 µL then stored in a -80°C freezer as master plates for future hit identification.

Mutant library screen and data analysis

Mutant library expression and extraction. Cells were mixed and pin replicated from master plates to clean, and pre-sterilized 96 deep-well plates (#10755-248, VWR) containing 1.8 mL of sterile LB media supplemented with 100 µg/mL of ampicillin per well. Plates were covered with a sterile porous film and incubated for 24 hours at 37°C with 180 rpm shaking. After growth, protein expression was induced with sterile LB media supplemented with 100 µg/mL of ampicillin and 3 mM of IPTG. Wells were mixed with a multichannel pipettor and incubated with the same porous film at room temperature with 180 rpm shaking for 20 hours. Porous film was removed, replaced with sterile adhesive sealing foil (#89040-034, VWR) and spun down at 4000 g in a centrifuge for 20 minutes at 4°C. Supernatant was removed, cell pellets resuspended in a lysozyme lysis buffer (5 mM potassium phosphate buffer pH 7.4, 1 µg/mL of leupeptin and pepstatin A and 10 mg/mL of lysozyme) and crude cell lysate extracted by 2 freeze thaw cycles: a 10-minute incubation at -80°C followed by a 20-minute incubation in a 25 °C water bath.⁴⁰ Plates were then spun down at 4000 g in a centrifuge for 30 minutes at 4°C. Supernatant, corresponding to the crude cell lysate, were used for the library screen.

Mutant library screen. A 10 µL aliquot of supernatant was transferred to a clean microplate containing 1 volume of reaction master mix. Every well contained the following: potassium phosphate buffer (50 mM, pH 7.4) and substrate **4.1** (1 mM) in DMSO (4% v/v). The first well of every plate was prepared similarly as described except with purified WT DEBS-TE (5 µM) to serve as a second positive control. Plates were then

incubated on benchtop for 24 hours sealed with adhesive sealing foil. Reaction mixes were quenched with 3 volumes of LCMS Optima grade MeCN with 0.05% v/v LCMS Optima grade formic acid (FA) and spun down at 4000 g in a centrifuge for 30 minutes. An aliquot of 60 μ L from the top layer in each well was transferred to a new microplate containing 2 volumes of LCMS Optima grade H₂O with 0.05% v/v FA and mixed before analysis by LCMS (Shimadzu Prominence UFLC system with SPD-20A Prominence UV/VIS Detector, LCMS-2020 single quadrupole MS). Chromatographic separations were done using a Leapsil (3.5-micron, 100 $\times$ 2.1 mm) reverse phase column in ESI+ mode with a flow rate of 1 mL/min. The mobile phase was composed of LCMS Optima grade H₂O (A) and MeCN (B) both supplemented with 0.05% v/v FA. Samples were monitored at 210 and 254 nm with an eluting gradient of 10% B to 95% B over 7 minutes.

MZmine2 data preprocessing. Preprocessing of LCMS data was done in MZmine2⁶⁴ with the following options described in order of their execution. Masses were detected using the Exact Mass Detector and the noise threshold set for each plate according to the average noise levels in chromatograms of methanol blanks that ran alongside each plate. Chromatograms were generated for each mass detected using the ADAP Chromatogram Builder (MZmine2 v2.53).⁶⁵ The minimum group size in number of scans was set to 5 with an m/z tolerance of 0.5. Chromatogram deconvolution was performed using the Baseline Cut-Off algorithm. Isotopic Peak Grouper was used to remove masses corresponding to an isotopic peak pattern. The m/z tolerance was set to 1.0 with a retention time tolerance of 0.2 minutes, maximum charge value of 2 and the selection of the monotonic shape option. Gap filling of dataset was done using the Peak Finder (Multithreaded) with an intensity, m/z and retention time tolerance of 10%, 0.5 and 0.05 minutes respectively. Peaks corresponding to the masses shown in **Figure S4.3** were selected using the Custom Database Search alongside their ID, m/z , retention time and identity with an m/z tolerance of 0.5. The Join Aligner was used to align mass peaks using the same parameters as described in the “Isotopic Peak

Grouper” setting with a value of 1 for both the weight for m/z and RT options. Data was then exported into Excel by peak area as a csv file with each peaks corresponding m/z , retention time and identity.

Mutant hit identifications. For each plate preprocessed by Mzmine2, a scatter plot of the areas under the curve (AUC) for peaks corresponding to macrocycle **4.3** ($m/z = 348$, $[M+H]^+$) and hydrolysis **4.2** ($m/z = 366$, $[M+H]^+$) products were constructed. A total of 10 mutants per position were selected from each scatter plot, prioritizing mutants with improved hydrolysis and macrocycle yields. Wells were picked with sterilized toothpicks and grown overnight at 37 °C in 5 mL of LB media supplemented with ampicillin (100 $\mu\text{g/mL}$). Plasmid DNA was extracted from overnight cultures using an E.Z.N.A. Plasmid Mini Kit I (#D694202, Omega Bio-tek) and retransformed into *E. coli* XLI-blue chemically competent cells. Colonies were picked, grown and plasmid DNA extracted similarly as described above for high quality plasmid DNA isolation. The quality of plasmid DNA extracted were verified by gel electrophoresis and sequenced by Sanger sequencing (G  nome Qu  bec) to determine mutations. From these 10 mutants, 2 were chosen for further experimental characterization per position (total of 18 single mutants): one with improved macrocyclization activity and one with improved hydrolysis activity relative to the WT DEBS-TE control. A full list of all mutants selected for characterization is shown in **Table S4.1**.

Combinatorial mutant generation and cloning

Construction and cloning of combinatorial mutants. Site-directed mutagenesis³⁴ was performed using the primers shown in **Table 4.2** with the indicated template plasmids to generate the desired combinatorial mutants. Every reaction was set in duplicate and contained the following: the correct forward and reverse primer pair (both 0.5 μM), Q5 Reaction Buffer (1X), dNTPs (200 μM), template DNA (1 ng), Q5 High-Fidelity DNA Polymerase (0.5 units, #M0491 NEB). Negative controls were prepared identically except without the addition of polymerase. Full plasmid PCR was performed for every reaction

as per the manufacturer's recommended conditions, with an annealing temperature of 74°C for 50 seconds.

Amplification of DNA by PCR was confirmed via gel electrophoresis and every reaction purified using the E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek #D6492) as described in their product manual followed by an overnight DpnI digestion (10 units, NEB #R0176) at 37°C. Digested products were then purified once more using the E.Z.N.A. Cycle-Pure Kit (#D6492, Omega Bio-tek) as described in their product manual. Eluted products were then phosphorylated for 1 hour at 37°C and ligated for 2 hours at room temperature before their transformation into *E. coli* XL1-blue chemically competent cells. Each phosphorylation reaction contained the following: 10 µL of the eluted product, T4 DNA Ligase Buffer (1X, NEB #B0202), T4 Polynucleotide Kinase (2.5 units, NEB #M0201). T4 DNA Ligase (5 units, NEB #M0202) was added to the same phosphorylation reactions prior to the 2-hour ligation incubation. A negative control was also prepared and treated similarly as described above except without the addition of T4 PNK and T4 DNA ligase to account for background colonies after transformation.

Table 4.2. Primers for combinatorial mutant generation. Primer pairs were designed for full-plasmid PCR with a T_m of ~68-69°C and ordered from Integrated DNA Technologies with standard desalting. All primers are shown 5' to 3' with the mutagenesis codon underlined. In each row, the top sequence is the forward primer whilst the bottom sequence is the reverse primer.

Mutant code	Combinatorial mutation	Template plasmid	Oligos (Forward/Reverse)
c42	Q176M/N180D	pJBH31	<u>GACGCCTGGCTGGAGGAGCT</u> CATCGCGTCCATGTGACCGGGC
c43	Q176M/M179P	pJBH31	<u>CCCAACGCCTGGCTGGAGGAGCTGACC</u> CGCGTCCATGTGACCGGGCGGG
c44	M179P/N180D	pJBH34	<u>CCCGACGCCTGGCTGGAGGAGCTGACC</u> CGCGTCCTGGTGACCGGGCGGG
c46	R32L/Q176M	pJBH31	<u>CTGTCCTACCTCGACCTGCTGGCG</u> GACCCTGCCCCGACACGCCC
c47	R32L/N180D	pJBH34	<u>CTGTCCTACCTCGACCTGCTGGCG</u> GACCCTGCCCCGACACGCCC
c48	R32L/M179P	pJBH32	<u>CTGTCCTACCTCGACCTGCTGGCG</u> GACCCTGCCCCGACACGCCC
c45	Q176M/M179P/N180D	pJBH42	<u>CCCGACGCCTGGCTGGAGGAGCTGACC</u> CGCGTCCTGGTGACCGGGCGGG
c49	R32L/Q176M/N180D	pJBH42	<u>CTGTCCTACCTCGACCTGCTGGCG</u> GACCCTGCCCCGACACGCCC
c50	R32L/Q176M/M179P	pJBH46	<u>CCCAACGCCTGGCTGGAGGAGCTGACC</u> CGCGTCCATGTGACCGGGCGGG
c51	R32L/M179P/N180D	pJBH47	<u>CCCGACGCCTGGCTGGAGGAGCTGACC</u> CGCGTCCTGGTGACCGGGCGGG
c52	R32L/Q176M/M179P/N180D	pJBH45	<u>CTGTCCTACCTCGACCTGCTGGCG</u> GACCCTGCCCCGACACGCCC

Expression and purification of mutants

Expression. WT DEBS-TE and 29 DEBS-TE mutants were recombinantly expressed: 18 single mutants identified from the library screen and 11 combinatorial mutants of R32L, Q176M, M179P and N180D single mutations (**Table S4.1**). All constructs were confirmed by Sanger sequencing (G  nome Qu  bec). Expression vectors were transformed into *E. coli* BL21(DE3) chemically competent cells for protein

expression. A flask containing 400 mL of LB media supplemented with 100 µg/mL of ampicillin were inoculated with 0.5 % v/v of dense transformed overnight cultures. For mutants s28, s29 and c48, 100 mL of LB media was used instead. Cultures were grown at 37 °C with 200 rpm shaking to an OD₆₀₀ of 0.5-0.6. Protein expression was then induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cultures were then incubated at room temperature with 200 rpm shaking for 18 hours. Mutant c49 could not be expressed and was not pursued further in this study.

Protein purification. Entire procedure was performed at 4 °C and all plasticware used were pre-chilled. Cells were harvested by centrifugation at 4,000 g and resuspended in 10 mL of lysis buffer (100 mM sodium phosphate, 300 mM NaCl, 10% w/v glycerol, 1 mg/mL lysozyme, 1 µg/mL of pepstatin A and 1 µg/mL leupeptin, pH 8.0). Cells were lysed by sonication on ice and cell debris removed by centrifugation at 12,000 g for 1 hour. Resulting lysate was incubated with 800 µL of Ni-NTA resin slurry (#30210, QIAGEN) on a rocking platform shaker for 1.5 hours. Lysate was then loaded into a gravity flow column and resin allowed to settle by gravity. Purification buffer (100 mM Tris, 300 mM NaCl, pH 7.4) was used at the indicated imidazole concentrations for resin washing and protein elution. Resin was first washed once with 0 mM imidazole and once more with 20 mM imidazole. Protein was eluted twice with 100 mM imidazole and then twice more with 250 mM imidazole. Protein collected during elution was concentrated and the buffer exchanged for dialysis buffer (100 mM Tris, 300 mM NaCl, pH 7.4) by centrifugation at 4,000 g using an Amicon Ultra-15 10 kDa MWCO tube (#UFC9010, Millipore). An equal volume of 60% w/v glycerol was added to protein concentrate, flash frozen in liquid nitrogen then stored at -80 °C. Protein concentration was determined by Bradford assay (#5000006, Bio-Rad).

For mutants s25, s26, s27, s36 and s37: cells were harvested by centrifugation at 4,000 g and resuspended in 2.5 mL of lysis buffer (50 mM potassium phosphate, 1 mg/mL lysozyme, 1 µg/mL of pepstatin A and 1 µg/mL leupeptin, pH 7.4). Cells were lysed by sonication on ice and cell debris removed

by centrifugation at 12,000 g for 1 hour. Resulting lysate was incubated with 250 μ L of Ni-NTA resin slurry (#30210, QIAGEN) on a rocking platform shaker for 1.5 hours. Lysate was then loaded into a gravity flow column and resin allowed to settle by gravity. Purification buffer (50 mM potassium phosphate, pH 7.4) was used at the indicated imidazole concentrations for resin washing and protein elution. Resin was first washed once with 0 mM imidazole and once more with 20 mM imidazole. Protein was eluted twice with 100 mM imidazole and then twice more with 250 mM imidazole. Protein collected during elution was dialysed in a Slide-A-Lyzer 10 kDa MWCO dialysis cassette (#66810, ThermoFisher) as per the manufacturer's recommended protocol in 300 $\times$ sample volume of 50 mM potassium phosphate buffer (pH 7.4). Dialysed samples were concentrated once by centrifugation at 4,000 g using a Vivaspin-20 10 kDa MWCO centrifugal concentrator tube (#VS2002, Sartorius). An equal volume of 60% w/v glycerol was added to protein concentrate, flash frozen in liquid nitrogen then stored at -80 $^{\circ}$ C. Protein concentration was determined by Bradford assay (#5000006, Bio-Rad). Mutant c49 could not be expressed and purified recombinantly.

For mutants s28, s29 and c48: cells were harvested, lysed and purified similarly as described in the section above. Proteins collected during elution were dialysed in a Slide-A-Lyzer 10 kDa MWCO dialysis cassette (#66810, ThermoFisher) as per the manufacturer's recommended protocol in 300 $\times$ sample volume of 50 mM potassium phosphate buffer (pH 7.4). Protein concentration of dialysed mutants was determined by Bradford assay (#5000006, Bio-Rad) and biochemical assays performed the same day.

Biochemical assays of WT DEBS-TE and mutants

Michaelis-Menten kinetics. Continuous enzymatic assays using DTNB (#D8130, Sigma-Aldrich) to track consumption of SNAC coupled substrate was performed in a half-area 96-well plate (50 mM potassium phosphate buffer pH 7.4, 12% v/v DMSO, 4% v/v DTNB, 5 μ M enzyme) and varying concentrations of substrate **4.1** (0 to 2 mM) in triplicate. A 2X substrate master mix (containing potassium phosphate buffer,

DMSO and substrate) and enzyme master mix (containing potassium phosphate buffer, DTNB and enzyme) were prepared separately and arrayed into a half-area 96-well plate (#07-200-848, ThermoFisher). Enzyme master mix was mixed into each substrate master mix well and absorbance measurements were taken every 20 seconds for 1 hour at 412 nm and 25 °C in a microplate reader. An SNAC-TNB adduct standard curve (0 to 100 μ M) was generated to convert absorbance readings into concentrations (μ M). Initial velocities were plotted against substrate concentration and the data fitted to the Michaelis-Menten model using GraphPad Prism 8.

Mutants s25, s27, s28, s41 and c50 were not kinetically characterized due to low observed activity in product distribution assays. Mutant s40 was not kinetically characterized due to inactivation by DTNB.

Product distribution assays. Assays were carried out in eppendorf tubes in a 15 μ L final reaction volume (50 mM potassium phosphate buffer pH 7.4, 12% v/v DMSO, 1 mM substrate **4.1**, 5 μ M enzyme) in triplicate. A substrate and negative control were prepared similarly as described above with the addition of water and heat denatured enzyme (5 μ M) in place of functional enzyme respectively. Reactions were left to incubate on the benchtop at room temperature for turnover. Reactions were quenched after 22 hours in 6 volumes of LCMS Optima grade acetonitrile (MeCN) supplemented with 0.05% v/v LCMS Optima grade formic acid (FA). Quenched reactions were left on the benchtop for 15 minutes, then centrifuged at max speed for 10 minutes at room temperature before analysis by LCMS (Shimadzu Prominence UFLC system with SPD-20A Prominence UV/VIS Detector, LCMS-2020 single quadrupole MS). Chromatographic separations were done using an Agilent Eclipse Plus C18 (3.5-micron, 150 $\times$ 4.6 mm) reverse phase column in ESI+ mode with a flow rate of 1 mL/min. The mobile phase was composed of LCMS Optima grade H₂O (A) and MeCN (B) both supplemented with 0.05% v/v FA. Samples were monitored at 210 and 254 nm with an elution gradient of 10% B to 95% B over 7 minutes.

4.6 Supplementary information

Constructs used and protein purities

Table S4.1. List of constructs used in this study. All constructs were cloned into a pET-21c(+) expression vector. Mutants were generated via full plasmid PCR site-directed mutagenesis as previously described, except for pRSG33¹⁶ which was not generated in this study. Mutant approximate expression yields, codes, plasmid names and the mutations are identified in the table.

Mutant code	Plasmid	Mutation	Yield (mg per L of culture)
WT	pRSG33	Wild-type DEBS-TE	9.24
tDEBS	pJBH09	S112STOP (truncated)	-
s24	pJBH24	R32L	2.07
s25	pJBH25	R32M	1.00
s26	pJBH26	L35Q	1.12
s27	pJBH27	L35R	1.14
s28	pJBH28	A77G	2.46
s29	pJBH29	A77S	1.82
s30	pJBH30	Q176A	8.36
s31	pJBH31	Q176M	11.69
s32	pJBH32	M179P	5.28
s33	pJBH33	M179R	3.99
s34	pJBH34	N180D	3.95
s35	pJBH35	N180P	1.60
s36	pJBH36	L206S	2.18
s37	pJBH37	L206H	3.75
s38	pJBH38	D210V	5.61
s39	pJBH39	D210W	8.94
s40	pJBH40	T213C	2.73
s41	pJBH41	T213R	3.94
c42	pJBH42	Q176M/N180D	2.88
c43	pJBH43	Q176M/M179P	1.92
c44	pJBH44	M179P/N180D	0.55
c46	pJBH46	R32L/Q176M	4.89
c47	pJBH47	R32L/N180D	2.97
c48	pJBH48	R32L/M179P	3.70
c45	pJBH45	Q176M/N180D/M179P	4.53
c49	pJBH49	R32L/Q176M/N80D	-
c50	pJBH50	R32L/Q176M/M179P	1.97
c51	pJBH51	R32L/M179P/N180D	3.30
c52	pJBH52	R32L/Q176M/M179P/N180D	8.54

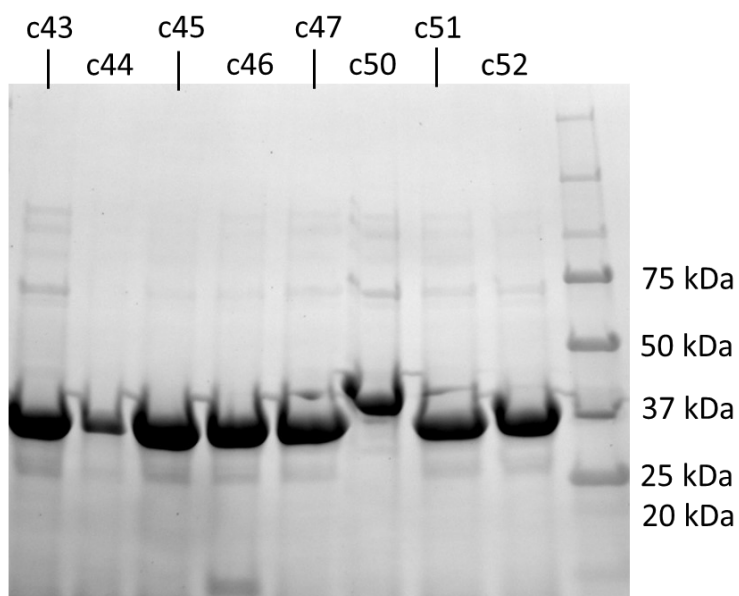
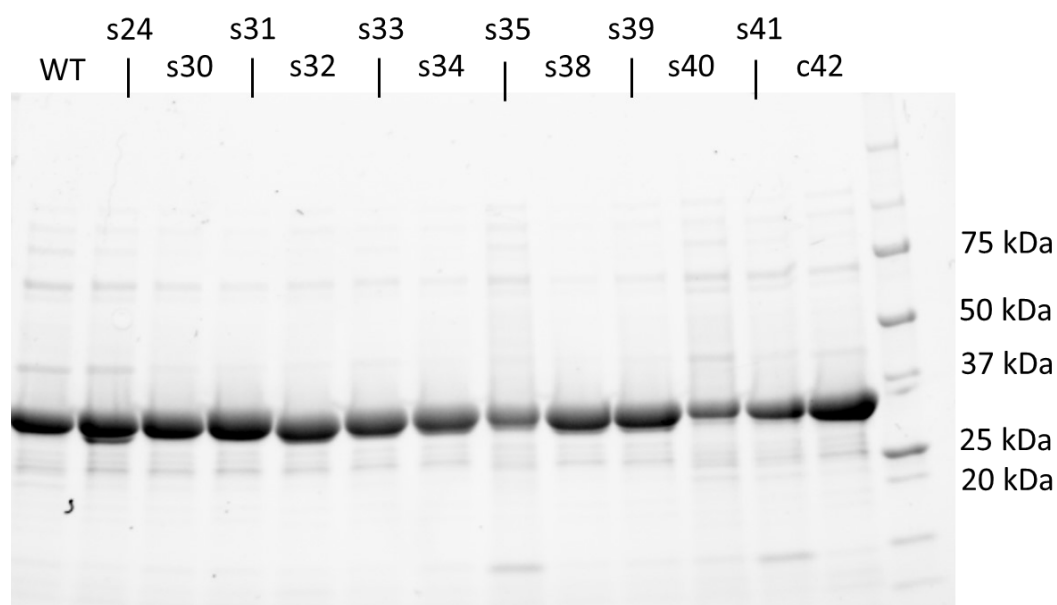


Figure S4.1. SDS-PAGE analysis of purified recombinant WT DEBS-TE and mutants. The expected molecular weight for WT DEBS-TE and mutants is approximately 32 kDa. The rightmost lane of gel corresponds to Precision Plus Protein Unstained Protein Standards (#1610363, Bio-Rad), all other lanes contain protein samples and are labeled on top. All Protein samples were purified and concentrated by ultrafiltration.

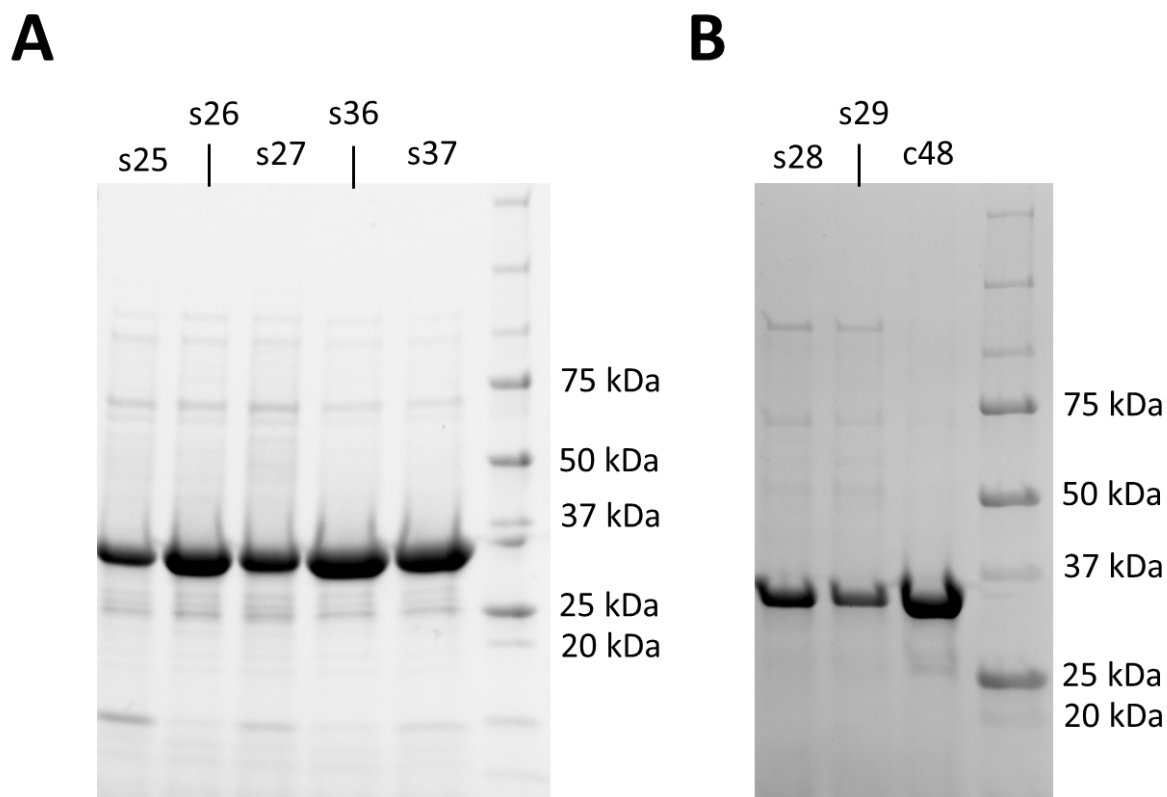


Figure S4.2. SDS-PAGE analysis of purified and dialysed recombinant DEBS-TE mutants. The expected molecular weight for WT DEBS-TE and mutants is approximately 32 kDa. The rightmost lane of gels in both panels corresponds to Precision Plus Protein Unstained Protein Standards (#1610363, Bio-Rad), all other lanes contain protein samples and are labeled on top. (A) Protein samples purified, dialysed and concentrated once by ultrafiltration. (B) Protein samples purified and dialysed (without ultrafiltration).

Mutant library LCMS screen data and statistics

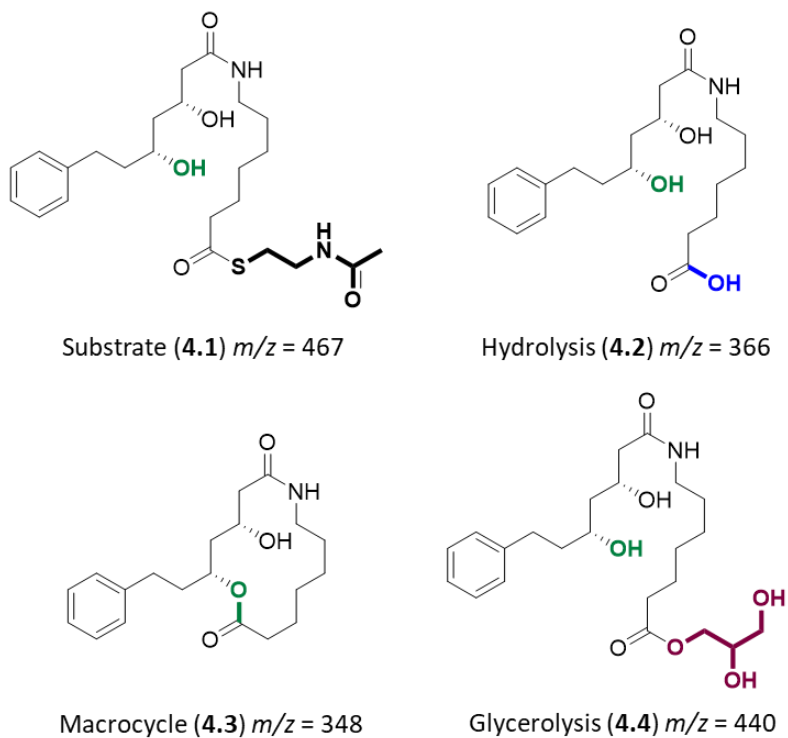


Figure S4.3. Structures of substrate and TE released products. Structures of substrate (4.1), hydrolysis (4.2), macrocycle (4.3) and glycerolysis (4.4) products. The m/z of each compound corresponding to an $[M + H^+]$ ion is shown and was used for Mzmine2 library screen data preprocessing.

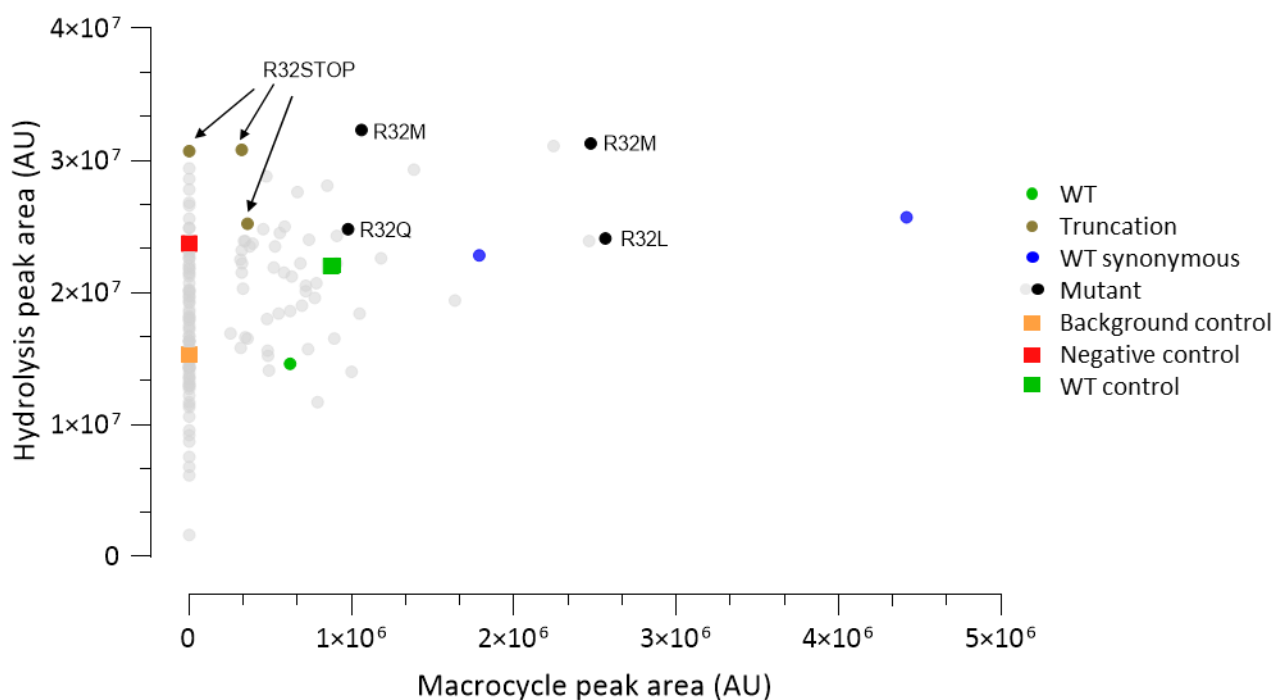


Figure S4.4. Screen data for DEBS R32 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position R32 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

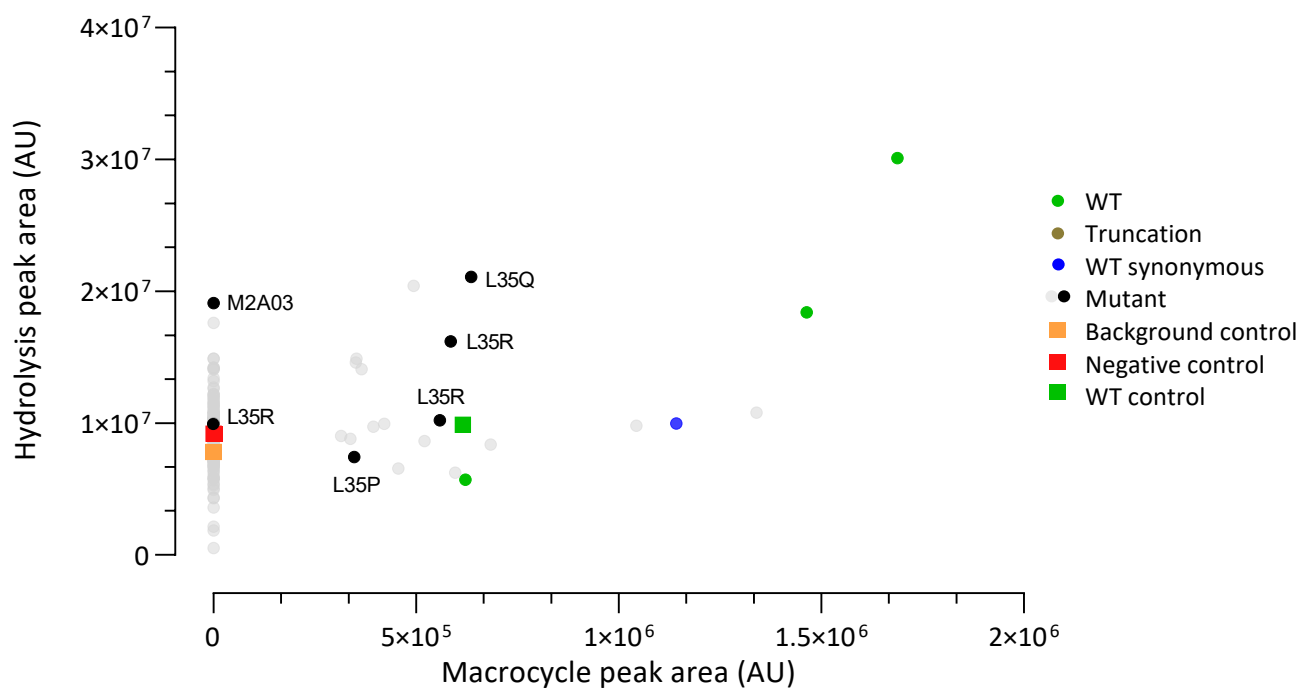


Figure S4.5. Screen data for DEBS L35 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position L35 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (**4.2**) and macrocycle (**4.3**) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

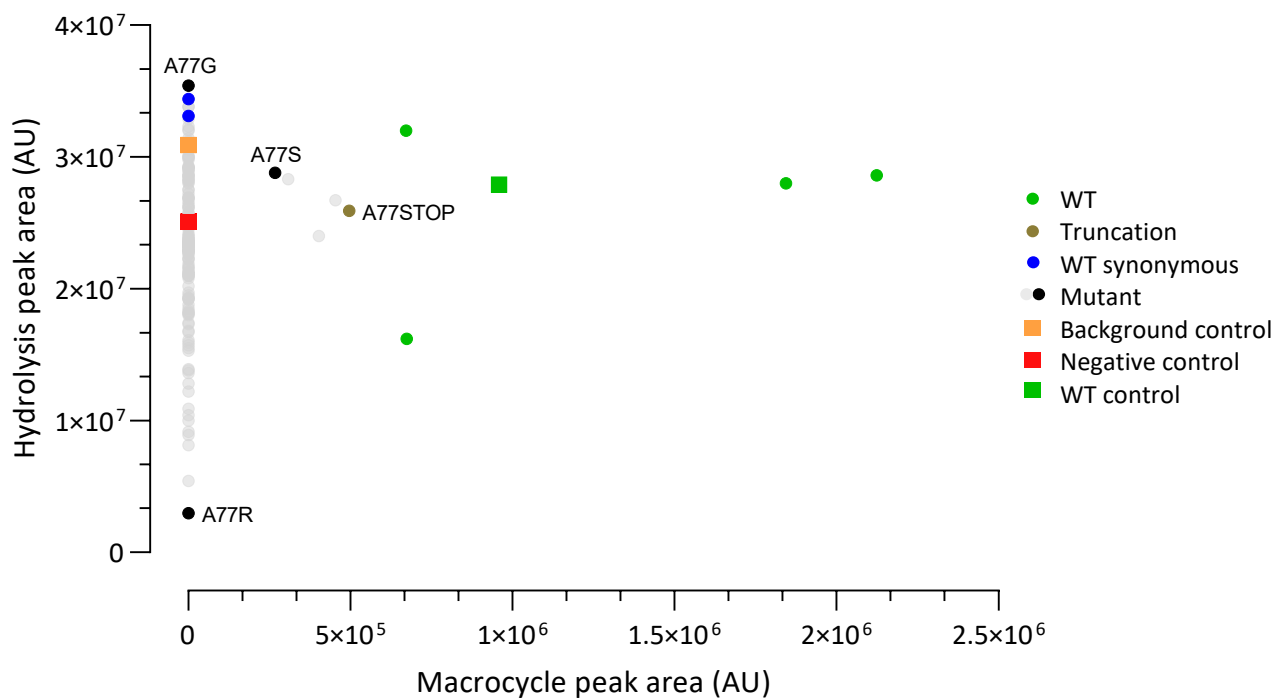


Figure S4.6. Screen data for DEBS A77 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position A77 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

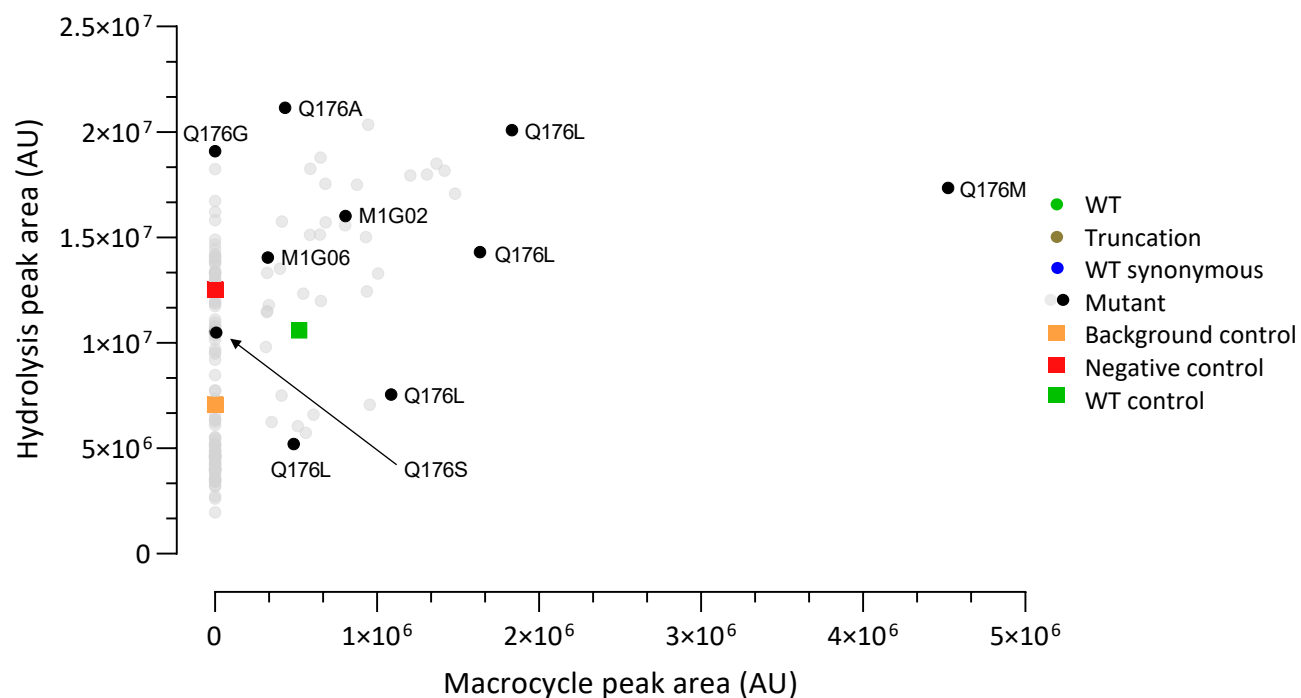


Figure S4.7. Screen data for DEBS Q176 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position Q176 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

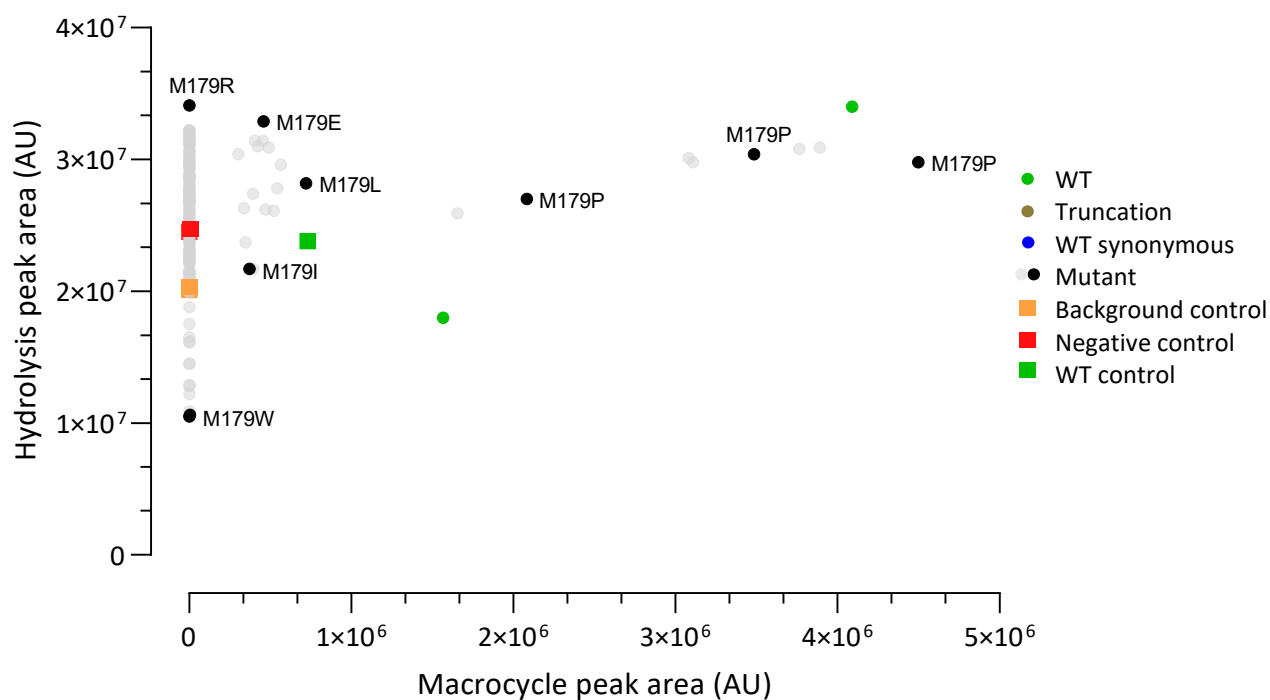


Figure S4.8. Screen data for DEBS M179 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position M179 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

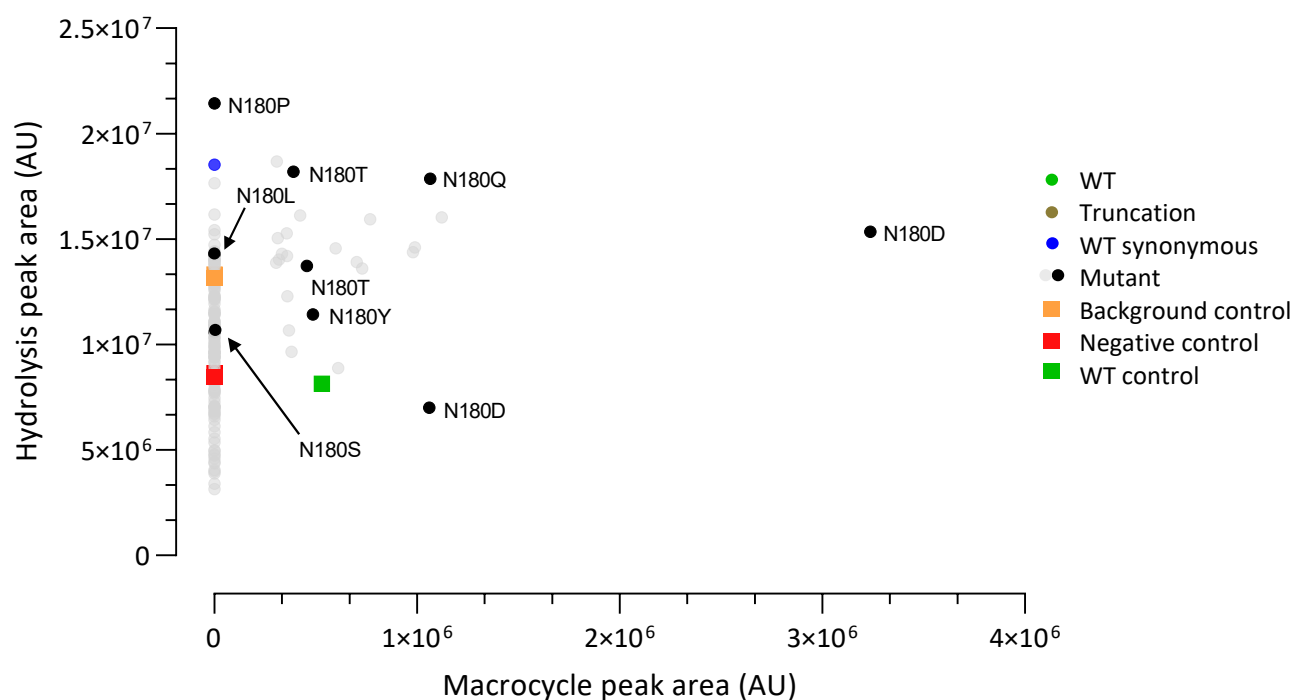


Figure S4.9. Screen data for DEBS N180 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position N180 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

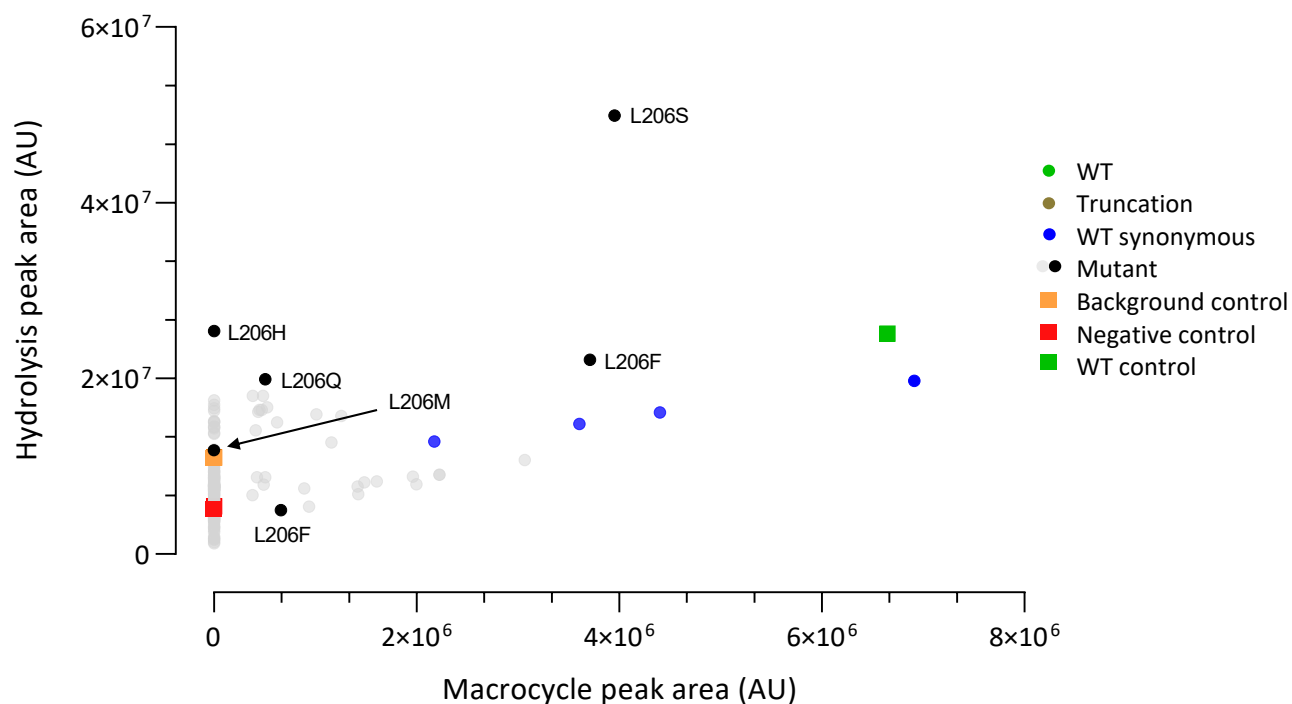


Figure S4.10. Screen data for DEBS L206 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position L206 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

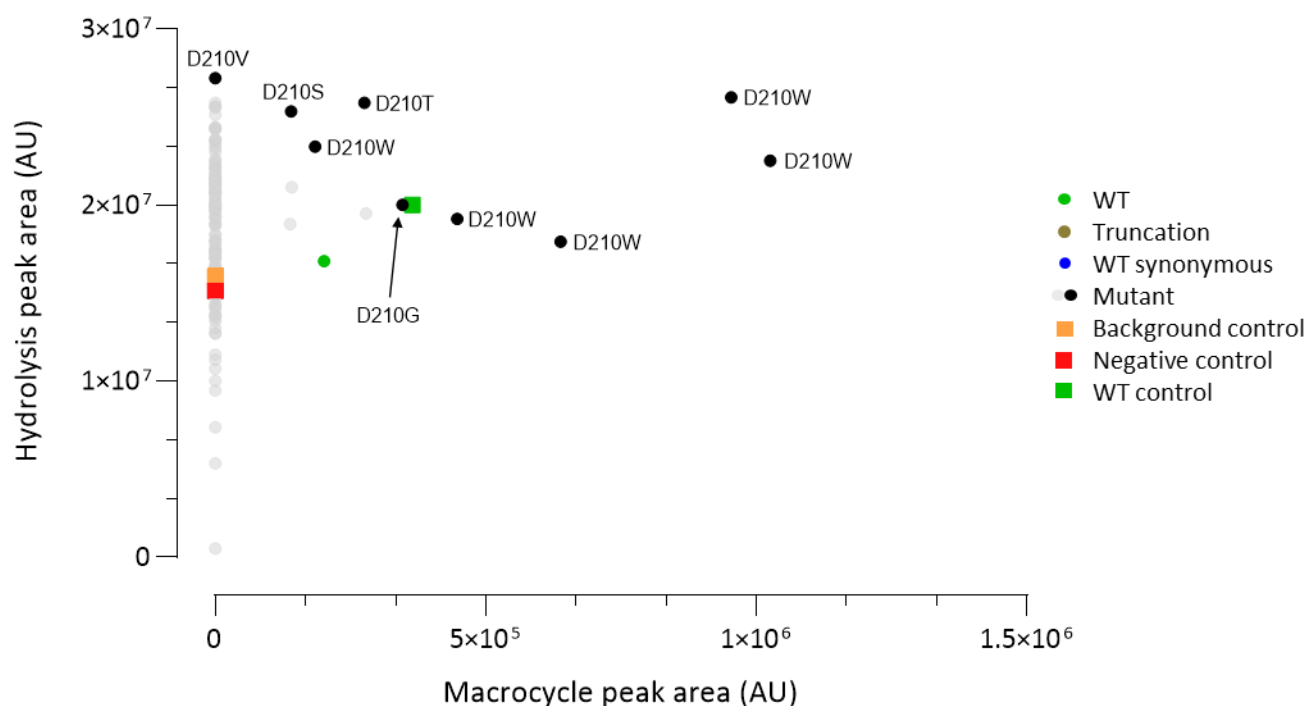


Figure S4.11. Screen data for DEBS D210 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position D210 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

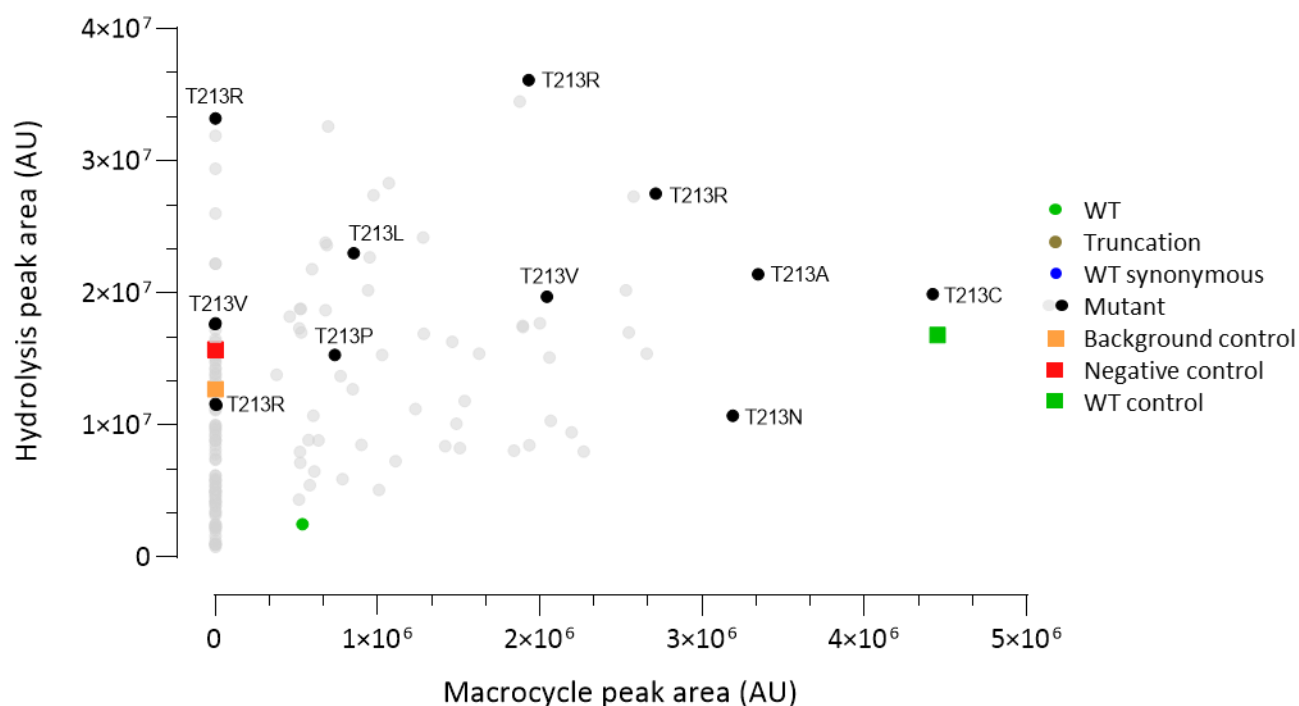


Figure S4.12. Screen data for DEBS T213 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position T213 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative, background and positive controls respectively (shown as squares). AUC of hydrolysis (4.2) and macrocycle (4.3) products for each clone (shown as dots) was extracted using Mzmine2 and plotted as a scatter plot. A total of 10 clones were selected for sequencing for mutant identification (shown as black dots). A symbol and colour legend is shown to the right of the graph.

Michaelis-Menten kinetics and example standard curve

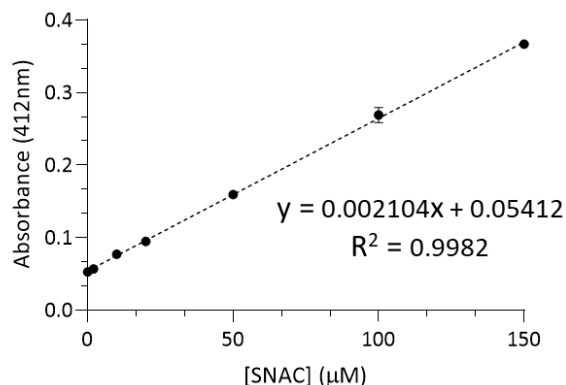


Figure S4.13. Example standard curve. Single absorbance measurements at 412 nm were taken in a half-area 96-well plate. Mixtures containing varying concentrations of SNAC thiolate (0 to 150 μM) were prepared in triplicate in 50 mM of potassium phosphate buffer (pH 7.4) and 12% v/v DMSO with 5 μM enzyme. 4% v/v DTNB reacts with SNAC thiolate forming a yellow coloured TNB anion that can be quantified by spectrophotometry.^{14,16,28,50} Absorbance readings were plotted against SNAC concentration to yield the above standard curve (mean ± s.d., n = 3 technical replicates).

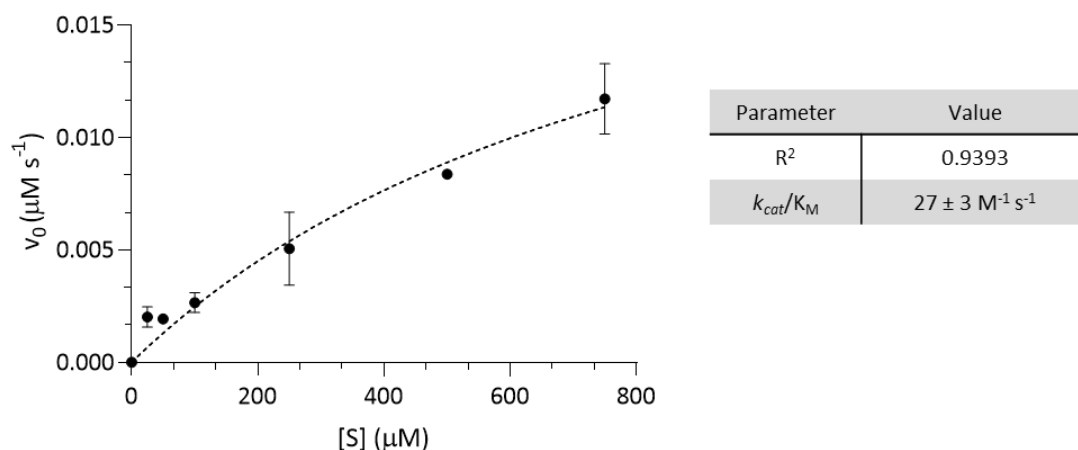


Figure S4.14. Michaelis-Menten kinetics of WT DEBS-TE. WT (5 μM) was incubated with varying concentrations of substrate **4.1** (0-800 μM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean ± s.d., n = 3 technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

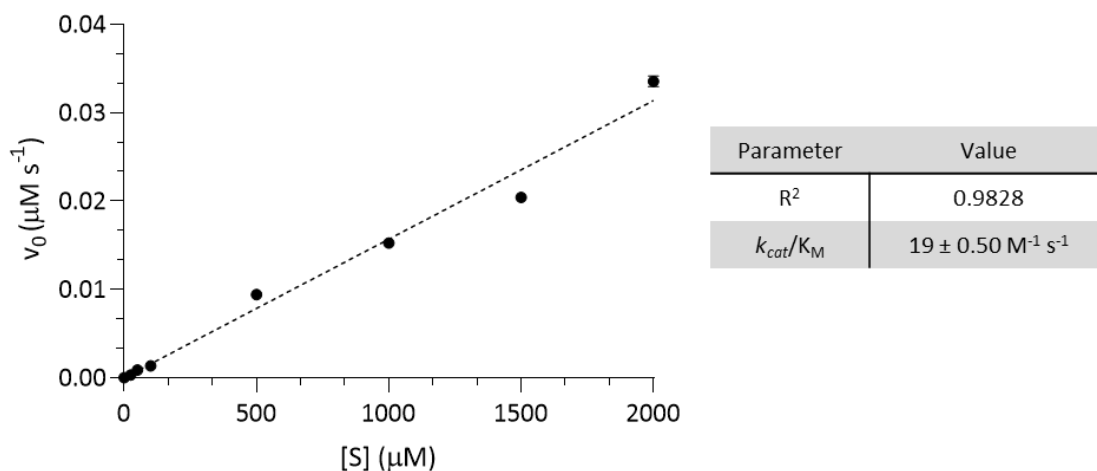


Figure S4.15. Michaelis-Menten kinetics of DEBS R32L mutant. Mutant s24 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-2 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

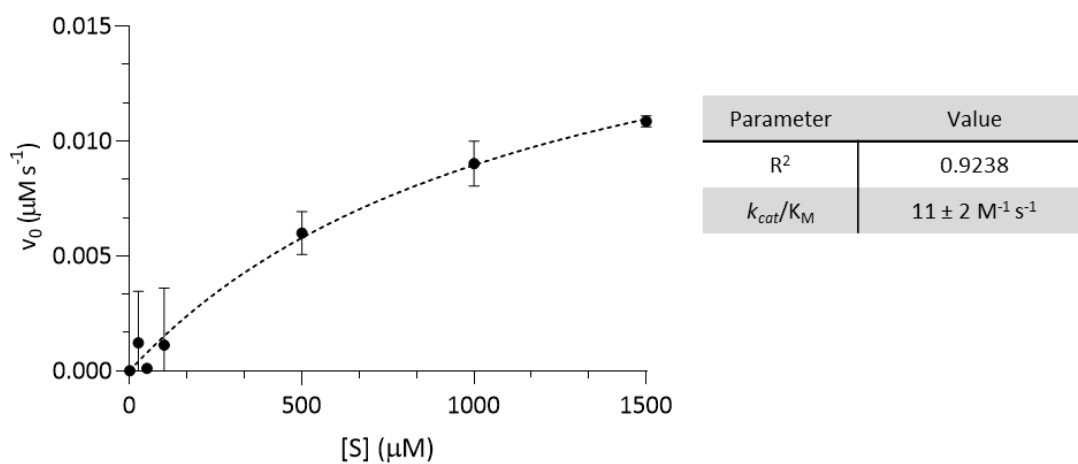


Figure S4.16. Michaelis-Menten kinetics of DEBS L35Q mutant. Mutant s26 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1.5 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

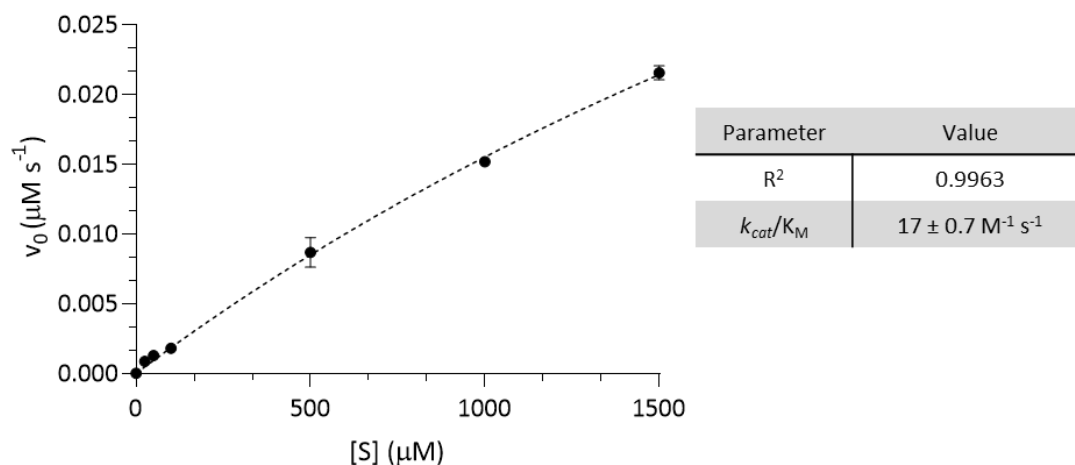


Figure S4.17. Michaelis-Menten kinetics of DEBS Q176A mutant. Mutant s30 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1.5 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

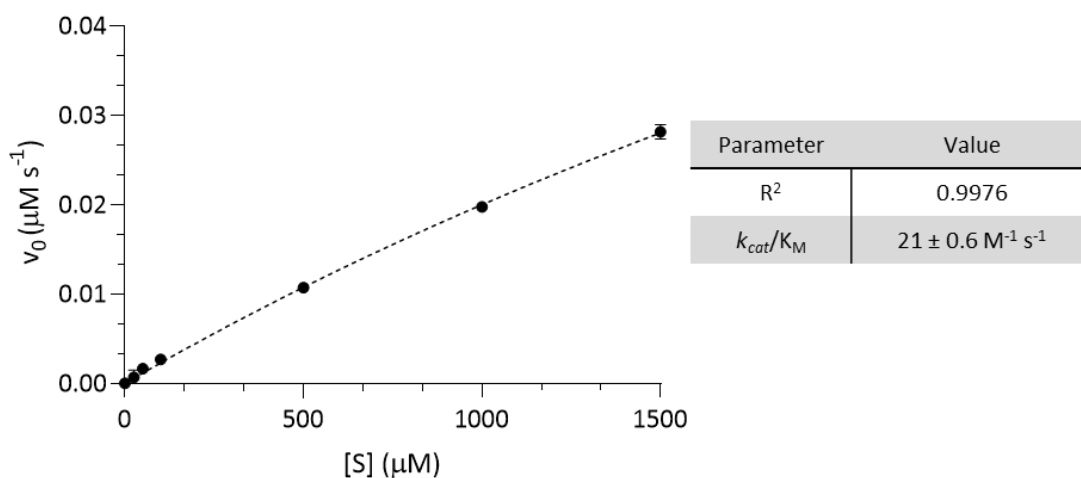


Figure S4.18. Michaelis-Menten kinetics of DEBS Q176M mutant. Mutant s31 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1.5 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

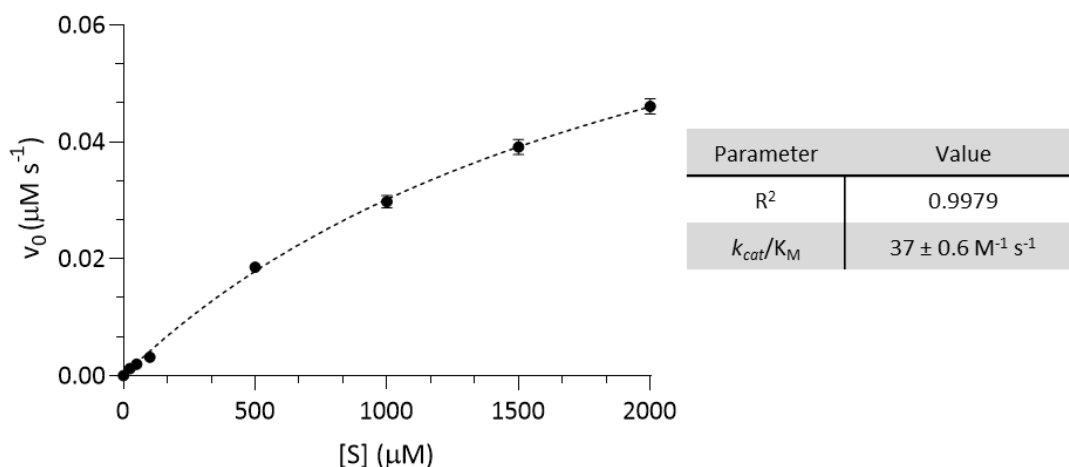


Figure S4.19. Michaelis-Menten kinetics of DEBS M179P mutant. Mutant s32 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-2 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

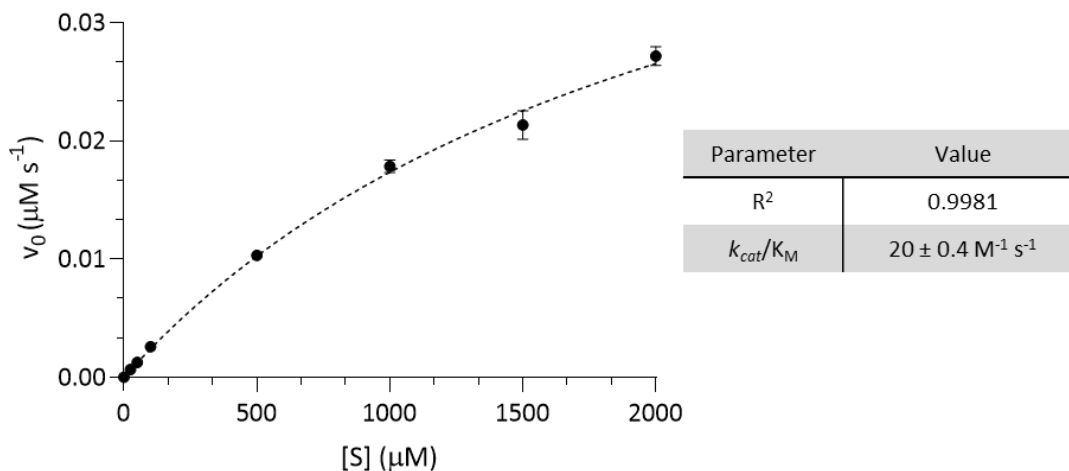


Figure S4.20. Michaelis-Menten kinetics of DEBS M179R mutant. Mutant s33 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-2 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

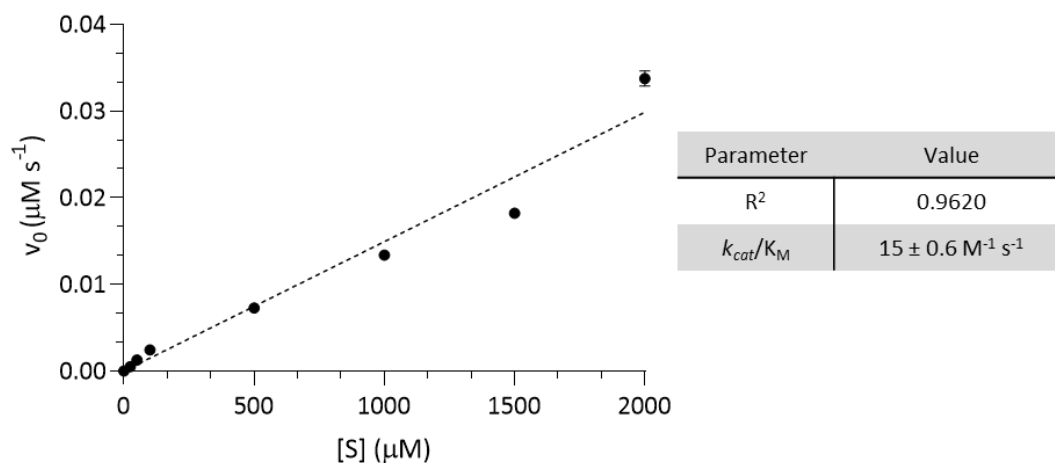


Figure S4.21. Michaelis-Menten kinetics of DEBS N180D mutant. Mutant s34 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-2 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

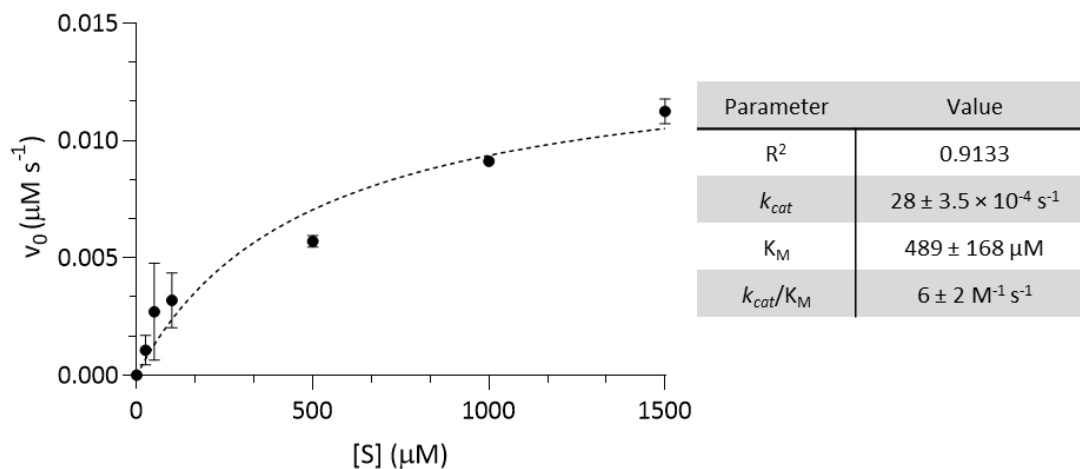


Figure S4.22. Michaelis-Menten kinetics of DEBS N180P mutant. Mutant s35 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1.5 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

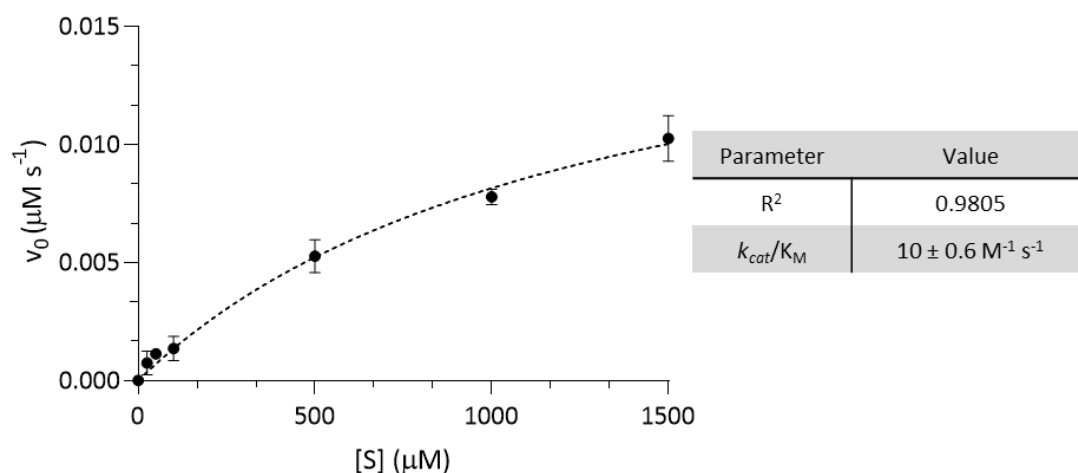


Figure S4.23. Michaelis-Menten kinetics of DEBS L206H mutant. Mutant s37 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1.5 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

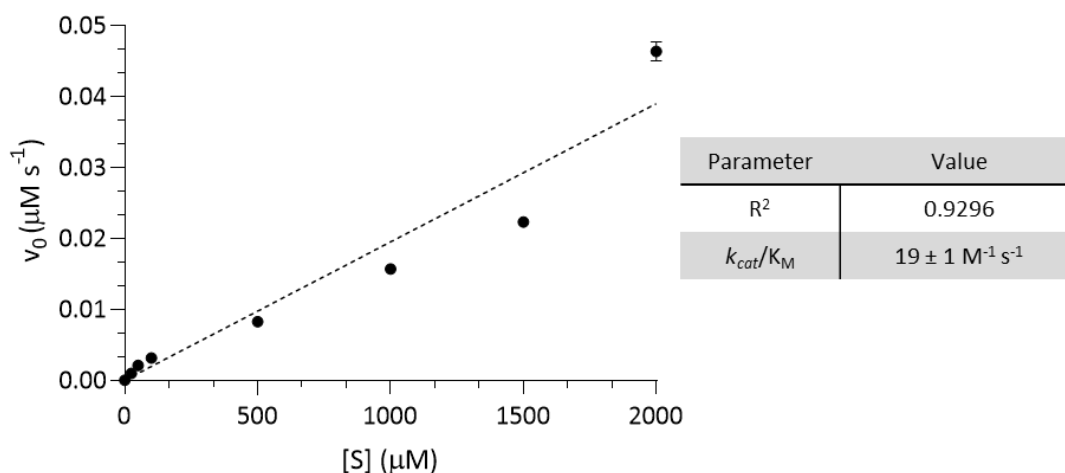


Figure S4.24. Michaelis-Menten kinetics of DEBS L206S mutant. Mutant s36 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-2 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

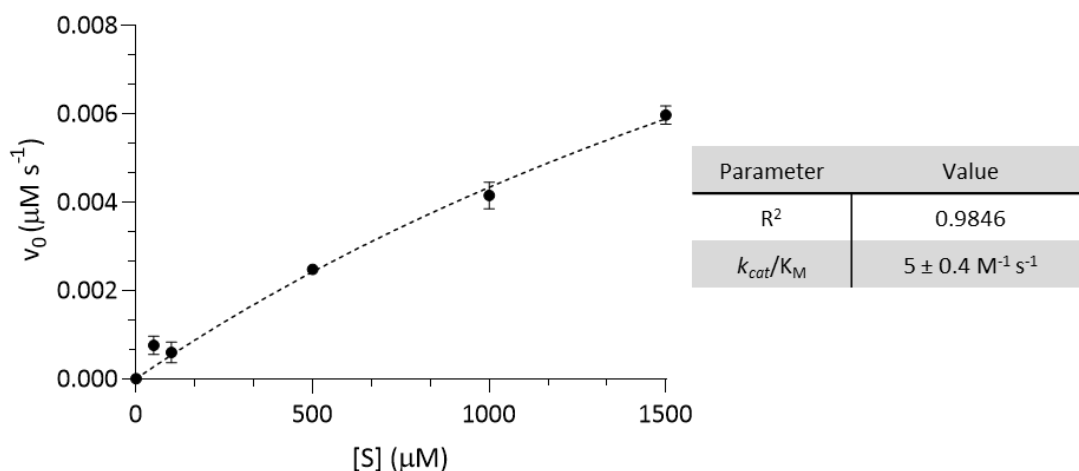


Figure S4.25. Michaelis-Menten kinetics of DEBS D210V mutant. Mutant s38 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1.5 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

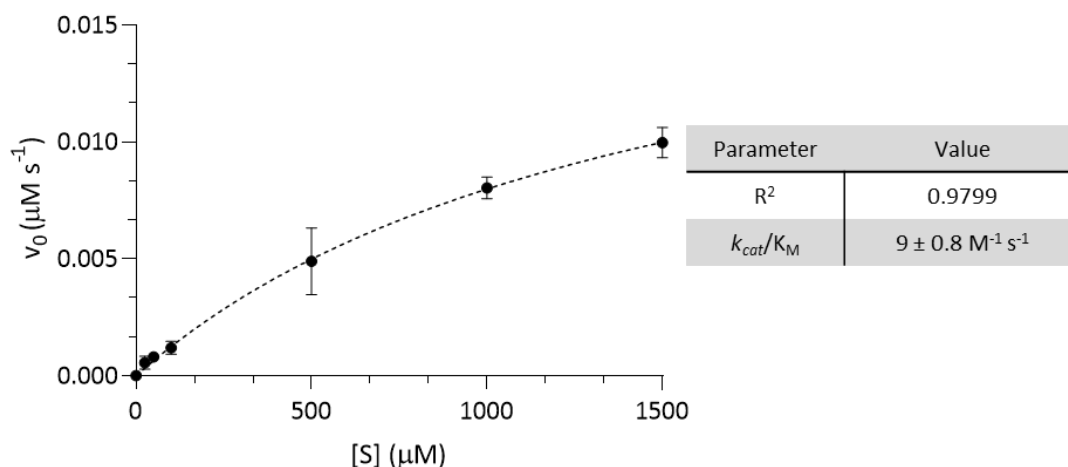


Figure S4.26. Michaelis-Menten kinetics of DEBS D210W mutant. Mutant s39 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1.5 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

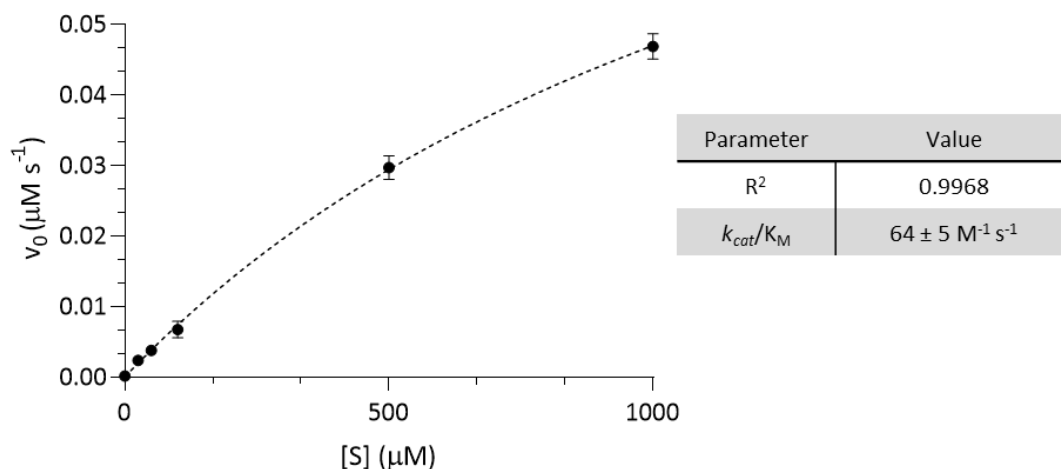


Figure S4.27. Michaelis-Menten kinetics of DEBS Q176M/N180D mutant. Mutant c42 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

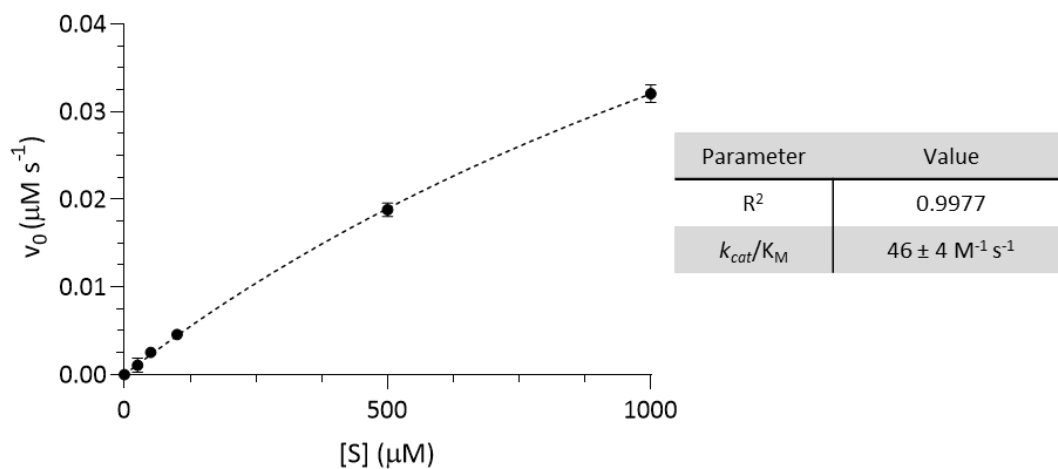


Figure S4.28. Michaelis-Menten kinetics of DEBS Q176M/M179P mutant. Mutant c43 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

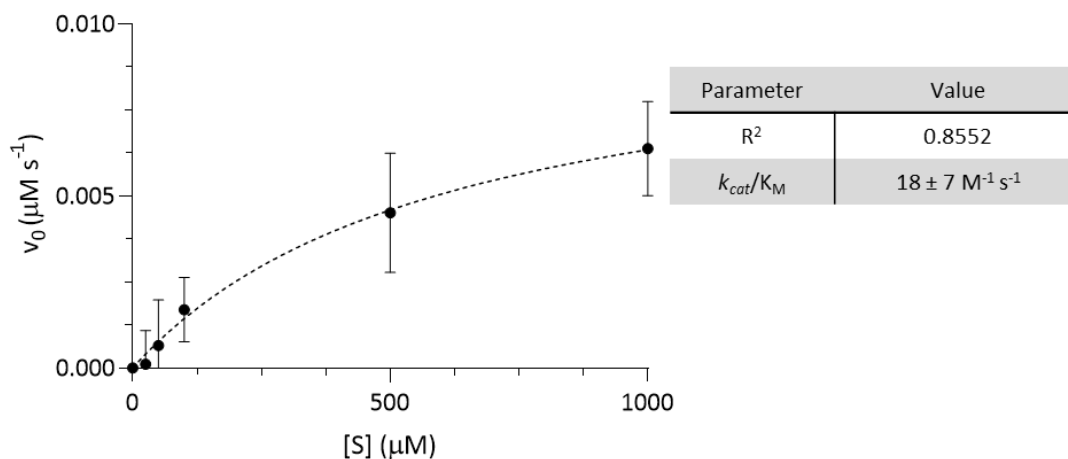


Figure S4.29. Michaelis-Menten kinetics of DEBS M179P/N180D mutant. Mutant c44 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

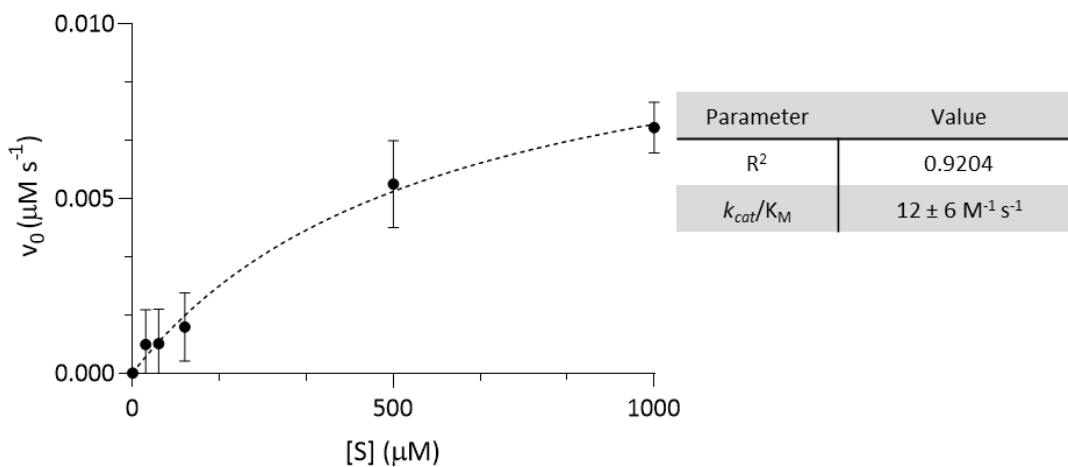


Figure S4.30. Michaelis-Menten kinetics of DEBS R32L/Q176M mutant. Mutant c46 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

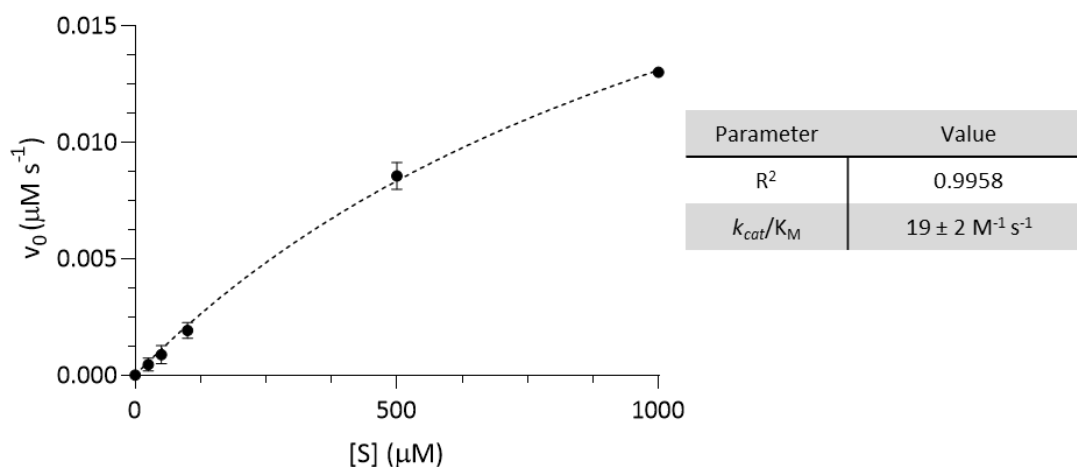


Figure S4.31. Michaelis-Menten kinetics of DEBS R32L/N180D mutant. Mutant c47 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 $^{\circ}\text{C}$ in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

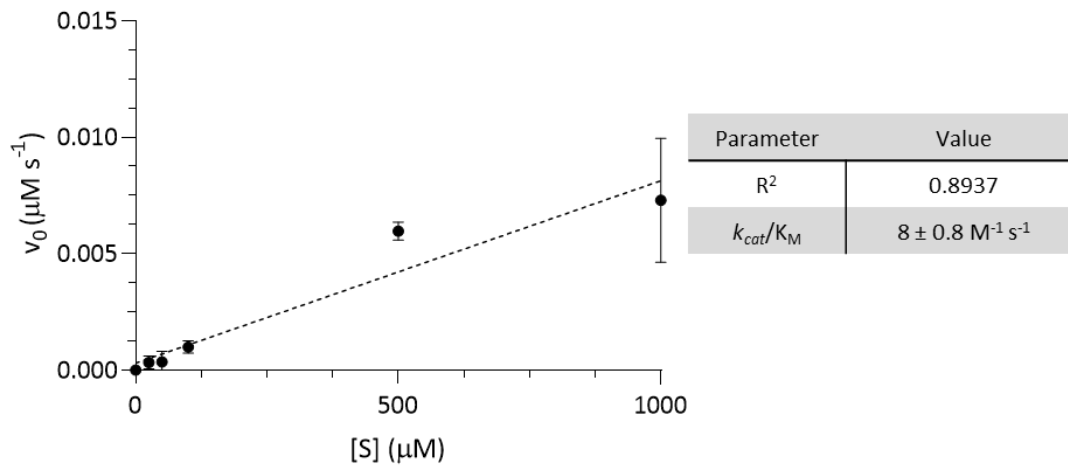


Figure S4.32. Michaelis-Menten kinetics of DEBS R32L/M179P mutant. Mutant c48 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 $^{\circ}\text{C}$ in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

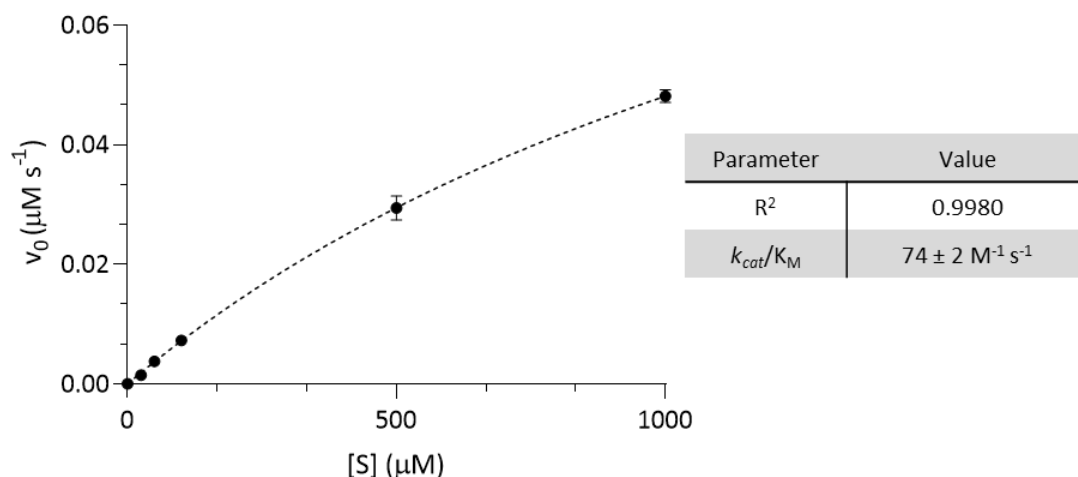


Figure S4.33. Michaelis-Menten kinetics of DEBS Q176M/M179P/N180D mutant. Mutant c45 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

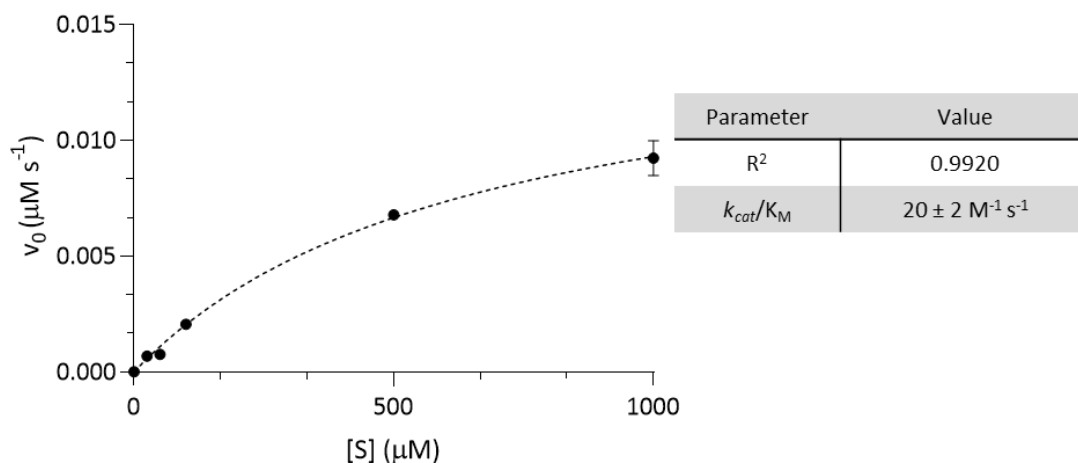


Figure S4.34. Michaelis-Menten kinetics of DEBS R32L/M179P/N180D mutant. Mutant c51 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 °C in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

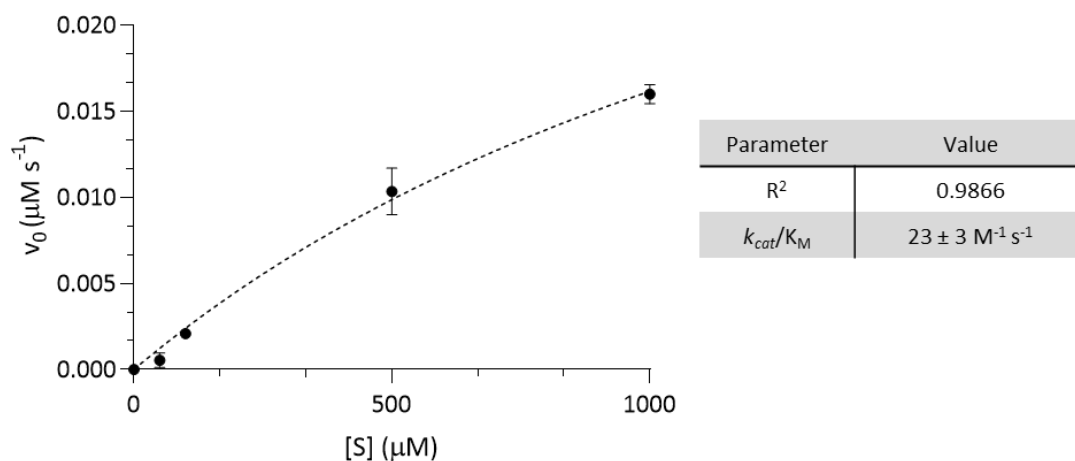


Figure S4.35. Michaelis-Menten kinetics of DEBS R32L/Q176M/M179P/N180D mutant. Mutant c52 (5 μM) was incubated with varying concentrations of substrate **4.1** (0-1 mM) and initial velocities (v_0) measured for 1 hour at 25 $^{\circ}\text{C}$ in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Initial velocities were plotted against substrate concentration and data fitted to the Michaelis-Menten model (dashed curve). Parameters of the fit are displayed in the table.

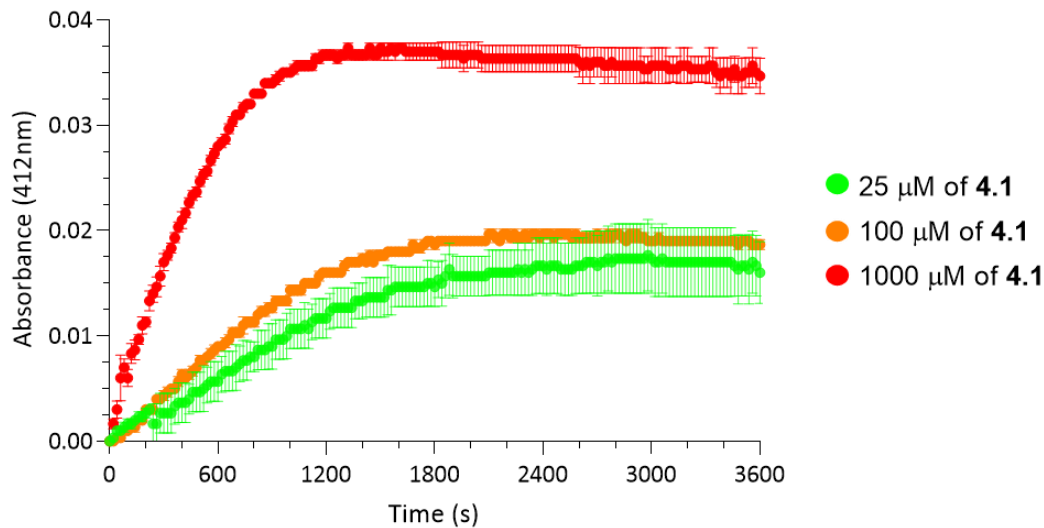


Figure S4.36. Example absorbance kinetic curves for DEBS R32L/Q176M/M179P/N180D mutant. Mutant c52 (5 μM) was incubated with 25, 100 and 1000 μM of **4.1** for measurement of initial velocities (v_0) for 1 hour at 25 $^{\circ}\text{C}$ in a microplate reader at 412 nm (mean $\pm$ s.d., $n = 3$ technical replicates). Note the progress curve with 1000 μM of **4.1** (red curve) qualitatively attains a plateau much earlier than anticipated, suggesting mutant is irreversibly inactivated.

Example LCMS chromatograms for library screen and biochemical assays

Table S4.2. Normalized product to substrate ratios and fold-change in k_{cat}/K_M . Product to substrate ratios were calculated by adding the AUC of macrocycle **4.3** and hydrolysis **4.2** products together before dividing by the substrate **4.1** AUC then normalized to WT. The fold decrease (-) or increase (+) in k_{cat}/K_M was calculated by dividing the larger value for each mutant compared to the WT. Mutants showing increased product to substrate (P/S) ratios with reduced k_{cat}/K_M are likely longer-lived catalysts and are colored red.

Mutant code	Mutation	Normalized product to substrate ratio (P/S)	Fold change in k_{cat}/K_M
s24	R32L	1.20	- 1.39
s25	R32M	0.23	-
s26	L35Q	0.43	- 2.39
s27	L35R	0.21	-
s28	A77G	0.07	-
s29	A77S	0.30	-
s30	Q176A	1.28	- 1.57
s31	Q176M	1.81	- 1.27
s32	M179P	3.34	+ 1.40
s33	M179R	1.26	- 1.31
s34	N180D	1.97	- 1.77
s35	N180P	1.14	- 4.65
s36	L206S	2.38	- 1.34
s37	L206H	0.73	- 2.69
s38	D210V	0.25	- 5.53
s39	D210W	0.56	- 2.83
s40	T213C	0.48	-
s41	T213R	0.11	-
c42	Q176M/N180D	6.94	+ 2.42
c43	Q176M/M179P	3.11	+ 1.75
c44	M179P/N180D	46.0	- 1.50
c46	R32L/Q176M	0.31	- 2.27
c47	R32L/N180D	0.60	- 1.38
c48	R32L/M179P	0.78	- 3.39
c45	Q176M/N180D/M179P	5.54	+ 2.79
c50	R32L/Q176M/M179P	0.08	-
c51	R32L/M179P/N180D	1.04	- 1.35
c52	R32L/Q176M/M179P/N180D	1.76	- 1.17

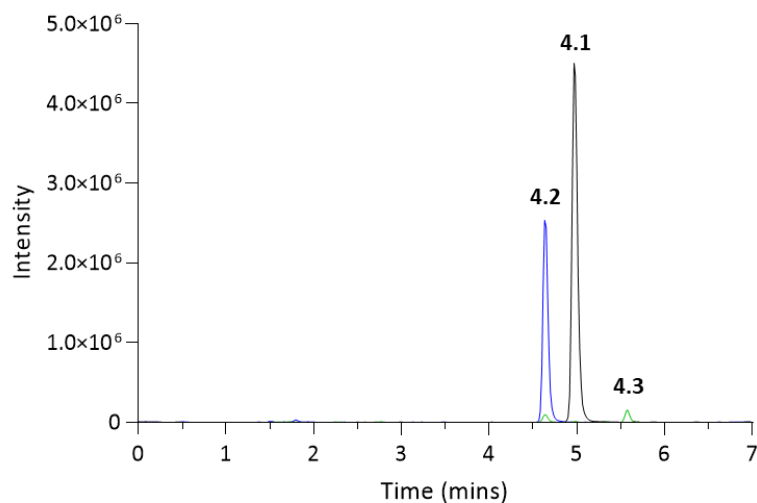
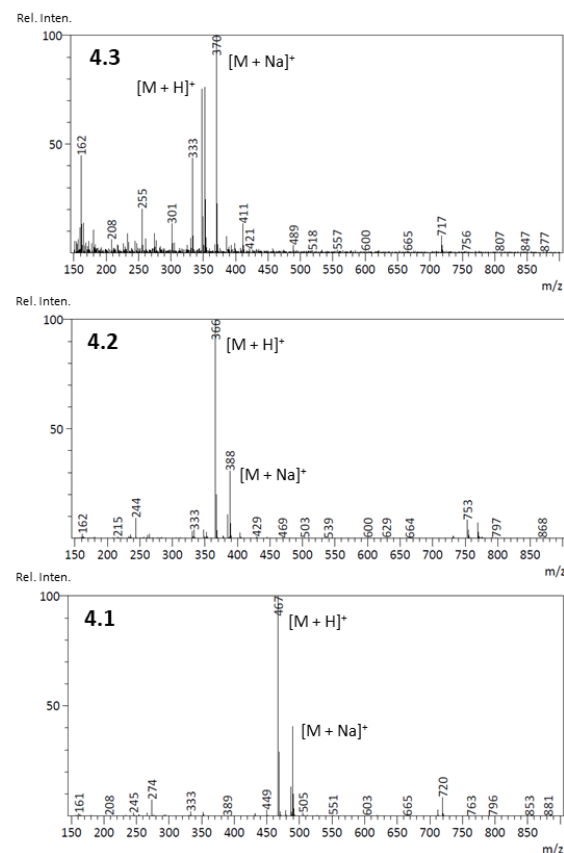
A**B**

Figure S4.37. Example chromatogram for library screen. (A) Extracted ion chromatogram for library Q176, well M1A02 (WT, positive control) *E. coli* lysate screen assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 348$ (**4.3**), 366 (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2** and macrocycle **4.3** products, mass adducts are identified in each spectra.

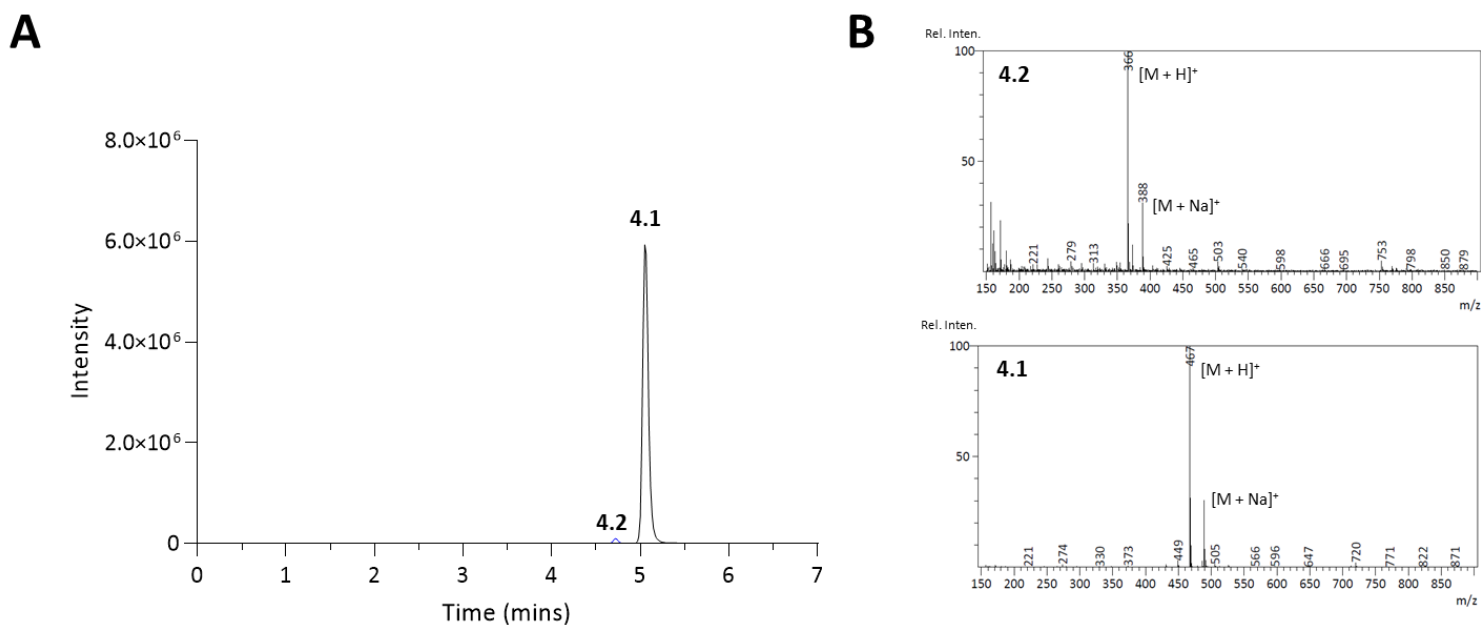


Figure S4.38. Substrate 4.1 only control 0-hour incubation. (A) Extracted ion chromatogram for substrate **4.1** (1 mM) only assay after 0 hours of incubation in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 366$ (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1** and hydrolysis **4.2** product, mass adducts are identified in each spectra. Presence of hydrolysis product **4.2** varied among samples due to freeze thaw cycles of substrate stock. All dead enzyme controls showed similar results after 0 and 24 hours of incubation.

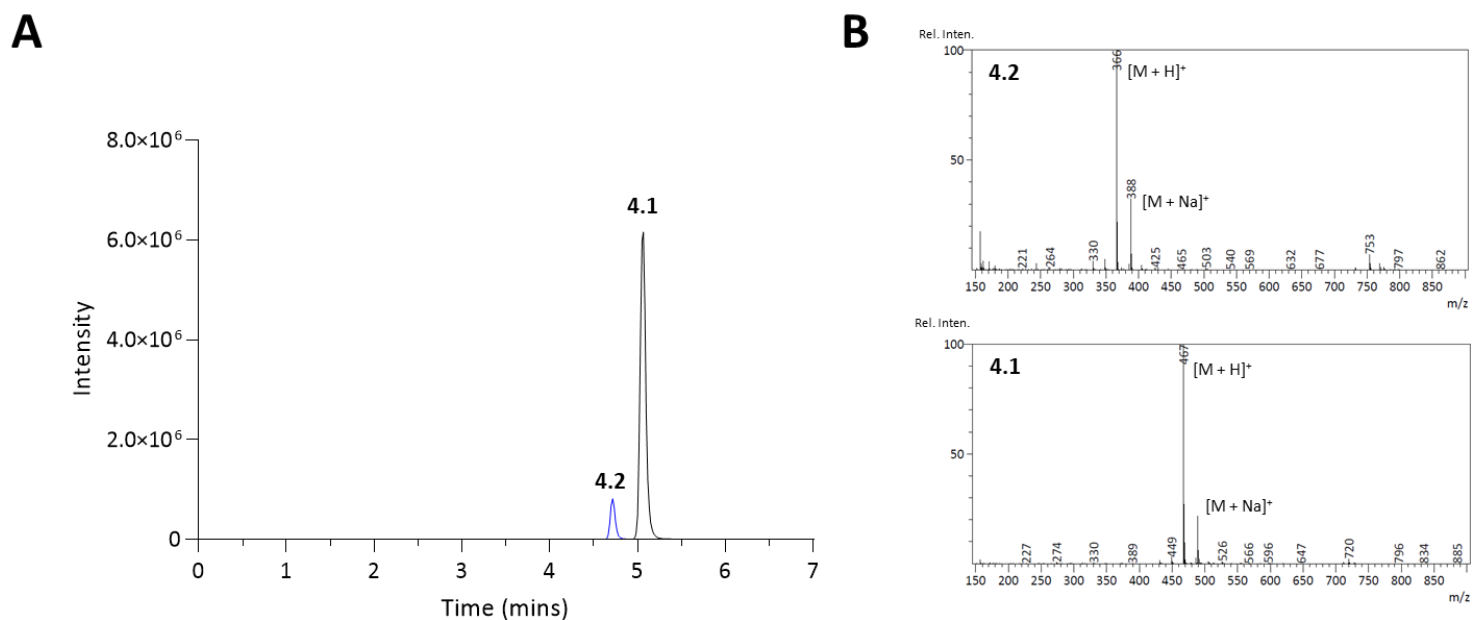


Figure S4.39. Substrate 4.1 only control 20-day incubation. (A) Extracted ion chromatogram for substrate **4.1** (1 mM) only assay after 20 days of incubation in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 366$ (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1** and hydrolysis **4.2** product, mass adducts are identified in each spectra. Presence of hydrolysis product **4.2** was due to background hydrolysis.

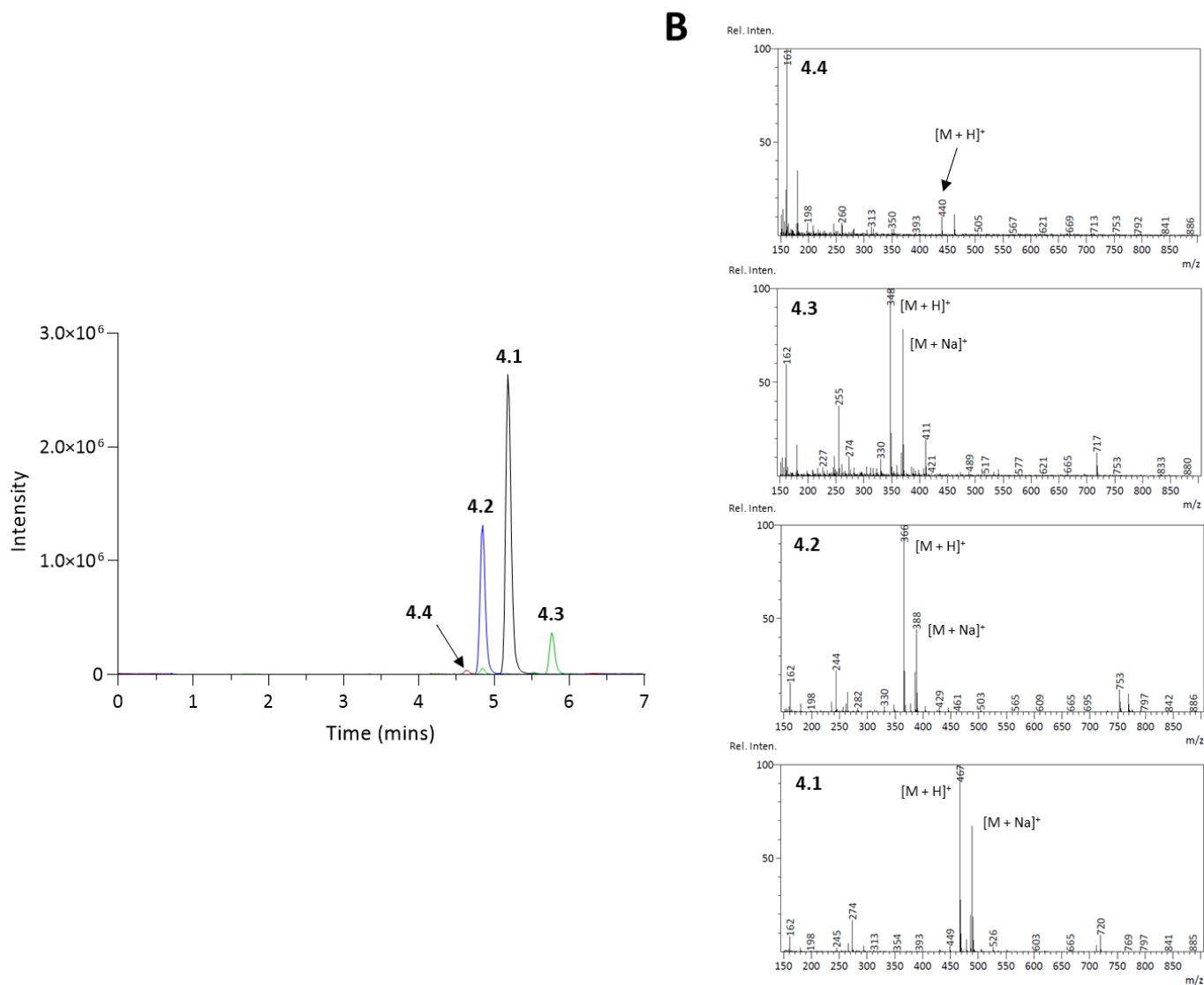


Figure S4.40. WT DEBS-TE 24-hour incubation with substrate 4.1. (A) Extracted ion chromatogram for WT (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 440$ (**4.4**), 348 (**4.3**), 366 (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2**, macrocycle **4.3** and glycerolysis **4.4** products, mass adducts are identified in each spectra.

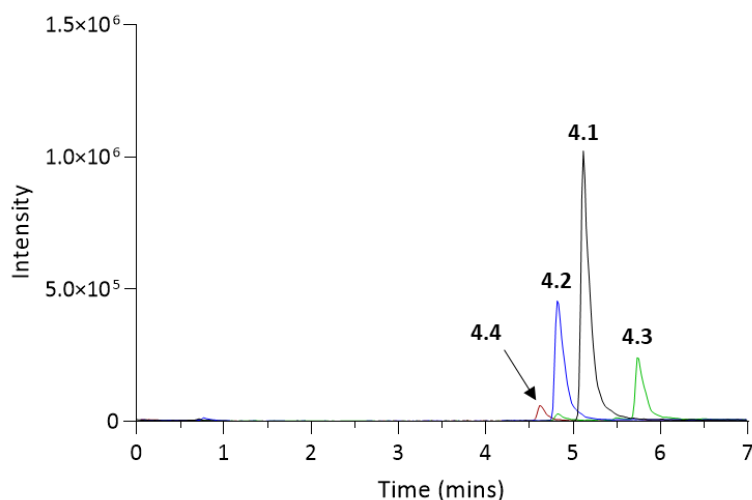
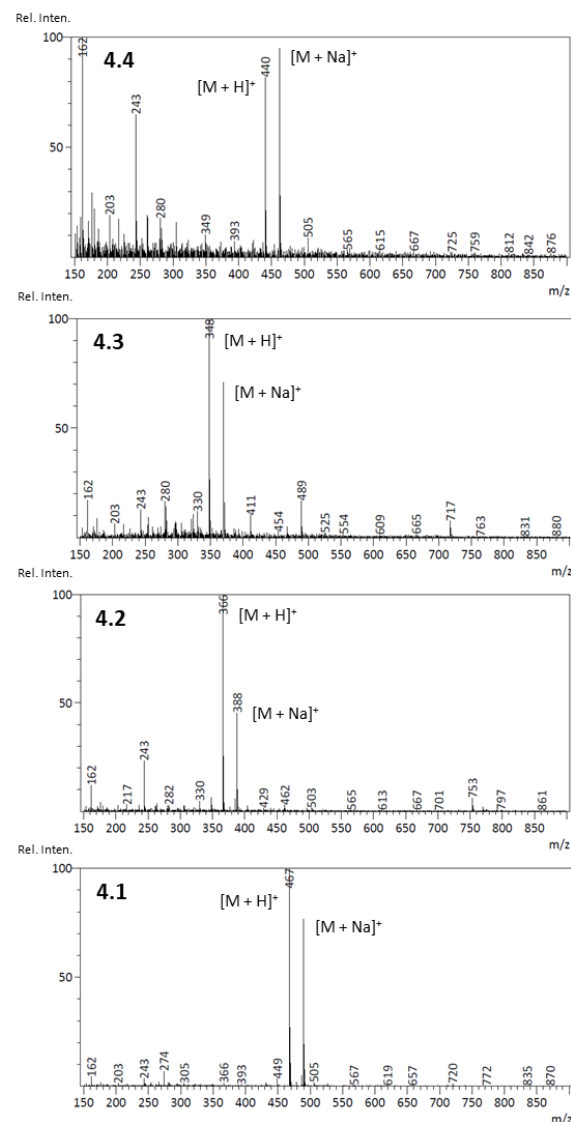
A**B**

Figure S4.41. DEBS R32L mutant 24-hour incubation with substrate 4.1. (A) Extracted ion chromatogram for mutant s24 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 440$ (**4.4**), 348 (**4.3**), 366 (**4.2**) and 467 (**4.1**). Mutant shows a ~ 1.7 -fold increase in M/H ratio over WT DEBS-TE (**Table S4.12**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2**, macrocycle **4.3** and glycerolysis **4.4** products, mass adducts are identified in each spectra.

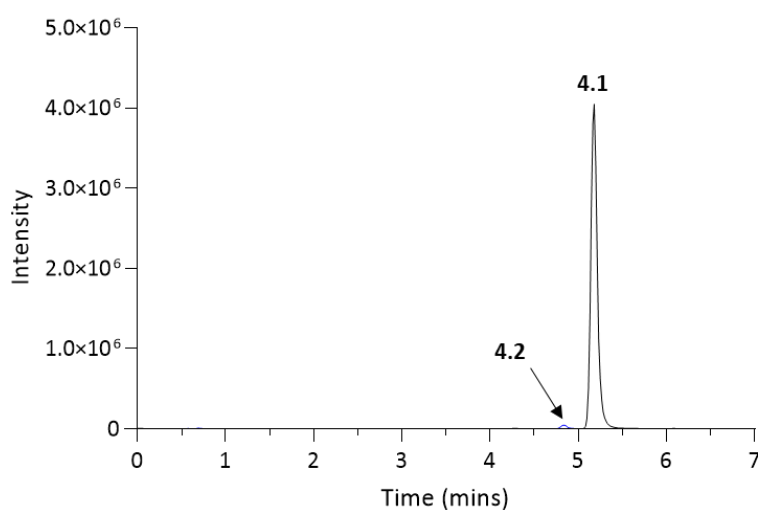
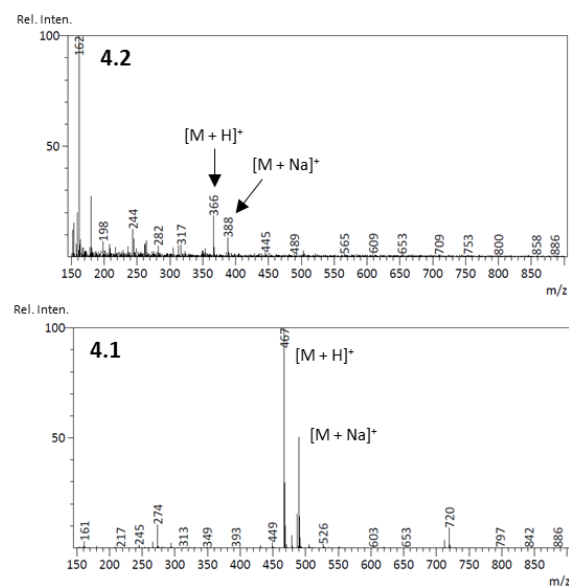
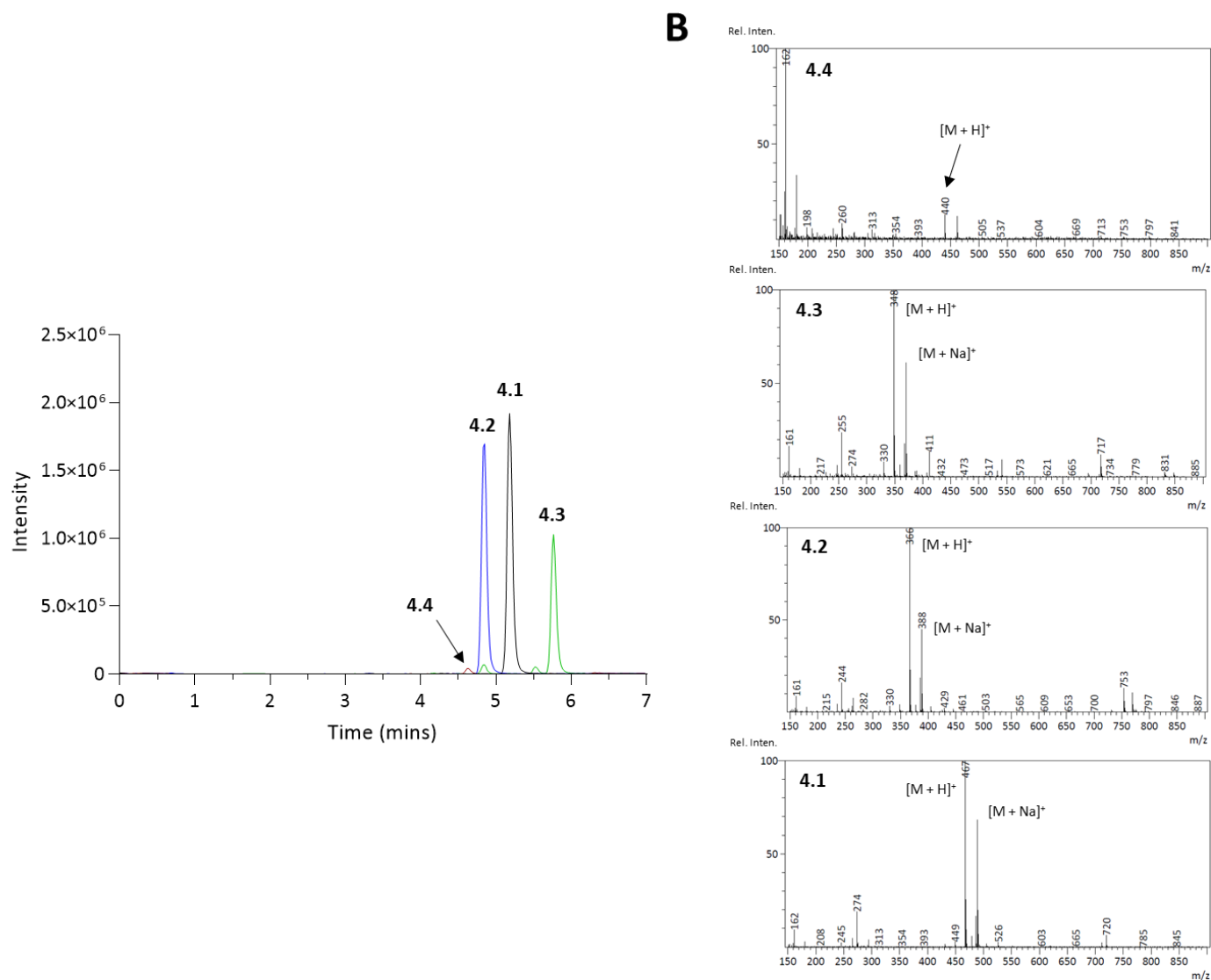
A**B**

Figure S4.42. DEBS L35R mutant 24-hour incubation with substrate 4.1. (A) Extracted ion chromatogram for mutant s27 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 366$ (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1** and hydrolysis **4.2** product, mass adducts are identified in each spectra. All dead enzyme controls and catalytically dead mutants showed similar results.



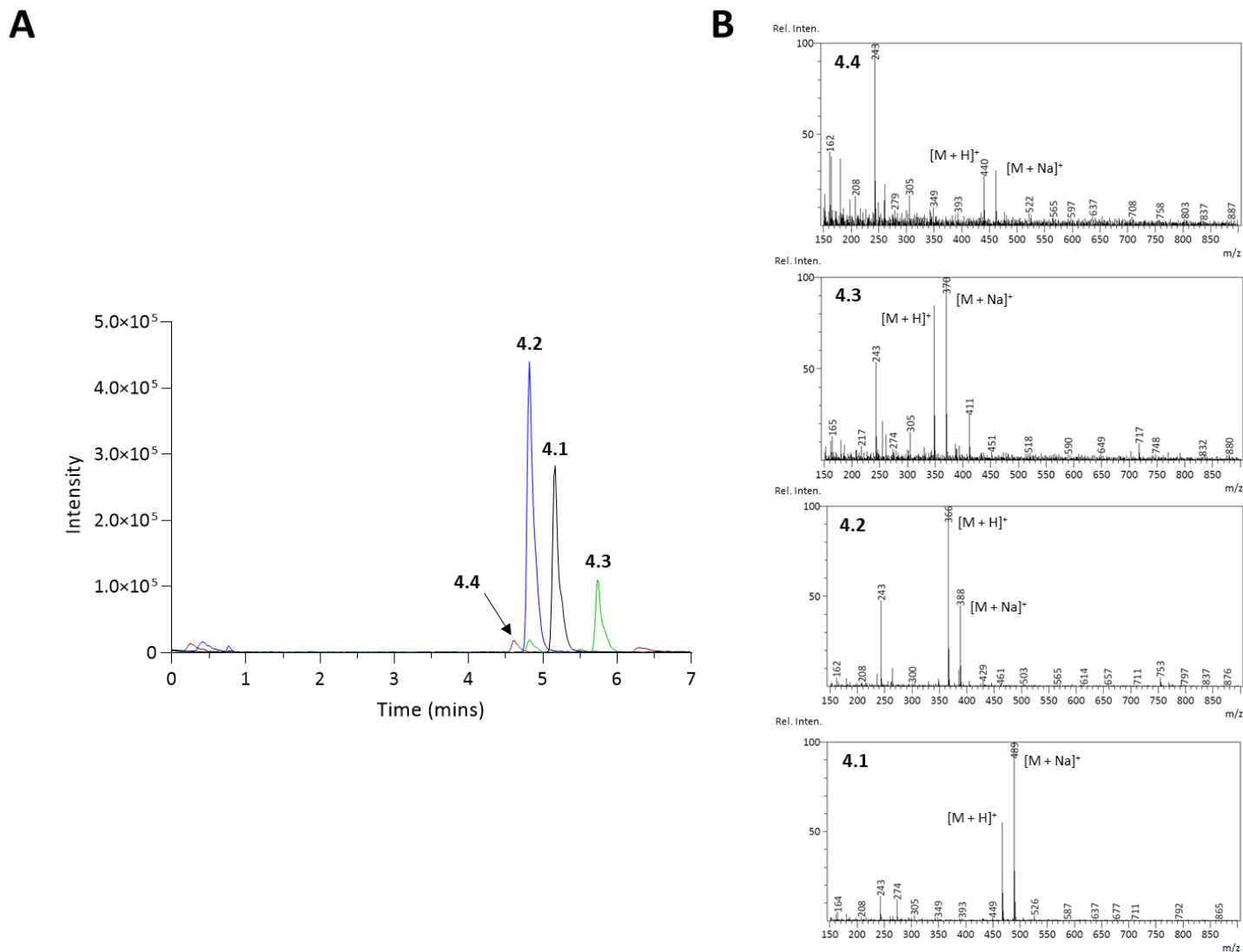


Figure S4.44. DEBS M179P mutant 24-hour incubation with substrate **4.1.** (A) Extracted ion chromatogram for mutant s32 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 440$ (**4.4**), 348 (**4.3**), 366 (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2**, macrocycle **4.3** and glycerolysis **4.4** products, mass adducts are identified in each spectra.

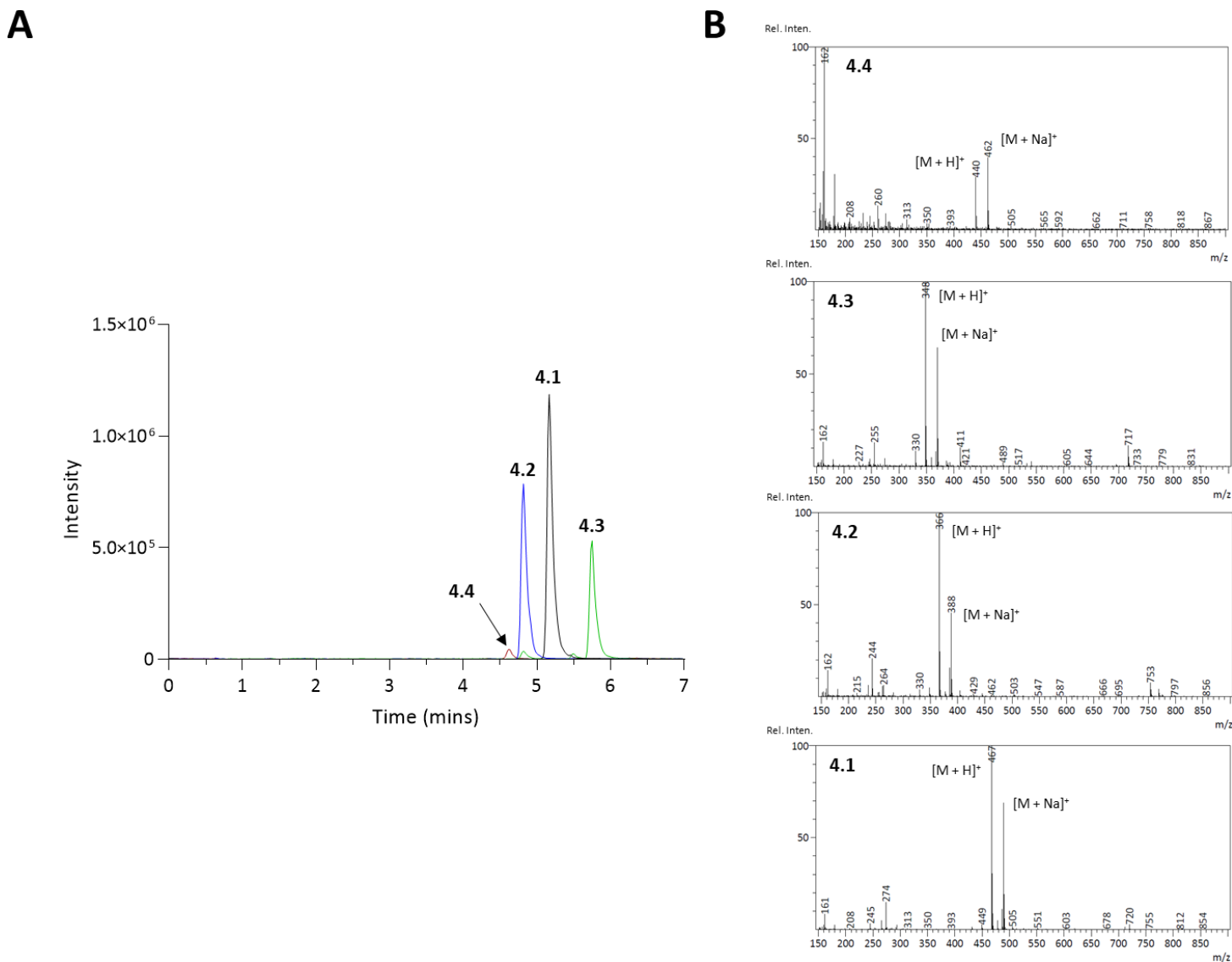


Figure S 4.45. DEBS N180D mutant 24-hour incubation with substrate 4.1. (A) Extracted ion chromatogram for mutant s34 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 440$ (**4.4**), 348 (**4.3**), 366 (**4.2**) and 467 (**4.1**). Mutant shows a ~ 2.2 -fold increase in M/H ratio over WT DEBS-TE (Table S4.12). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2**, macrocycle **4.3** and glycerolysis **4.4** products, mass adducts are identified in each spectra.

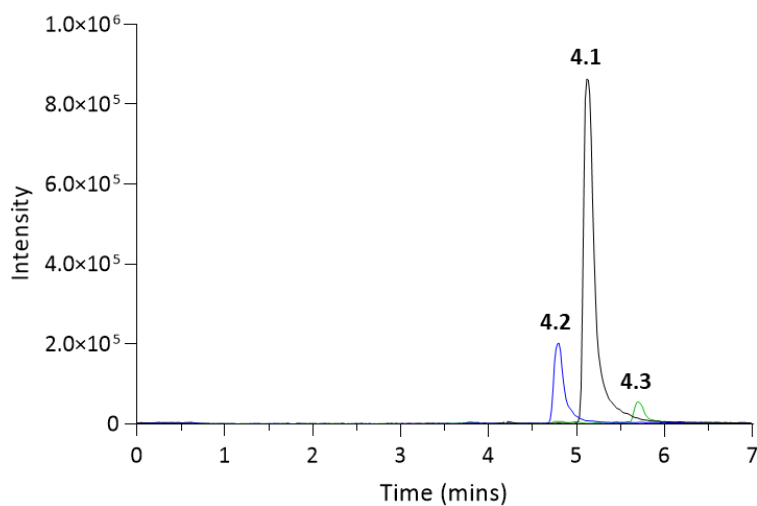
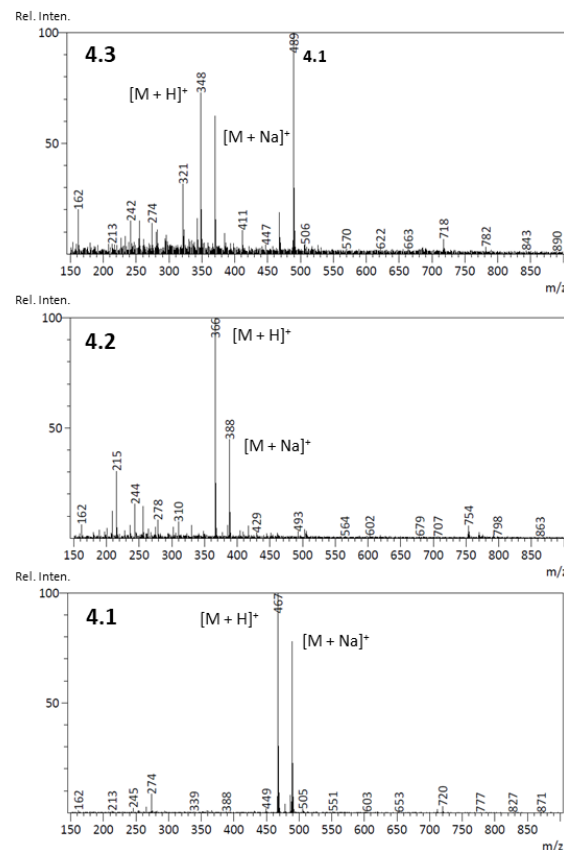
A**B**

Figure S4.46. DEBS R32L/Q176M mutant 24-hour incubation with substrate **4.1.** (A) Extracted ion chromatogram for mutant c46 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for [M + H]⁺ ions with m/z = 348 (**4.3**), 366 (**4.2**) and 467 (**4.1**). Mutant shows a ~2.4-fold increase in M/H ratio over WT DEBS-TE (**Table S4.12**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2** and macrocycle **4.3** products, mass adducts are identified in each spectra.

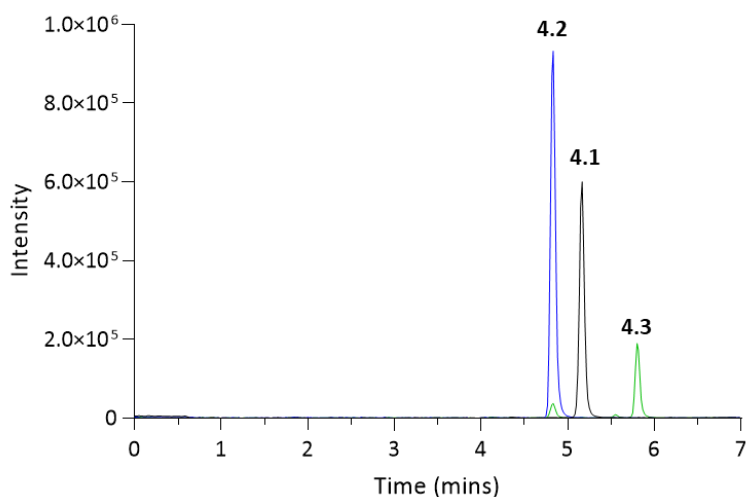
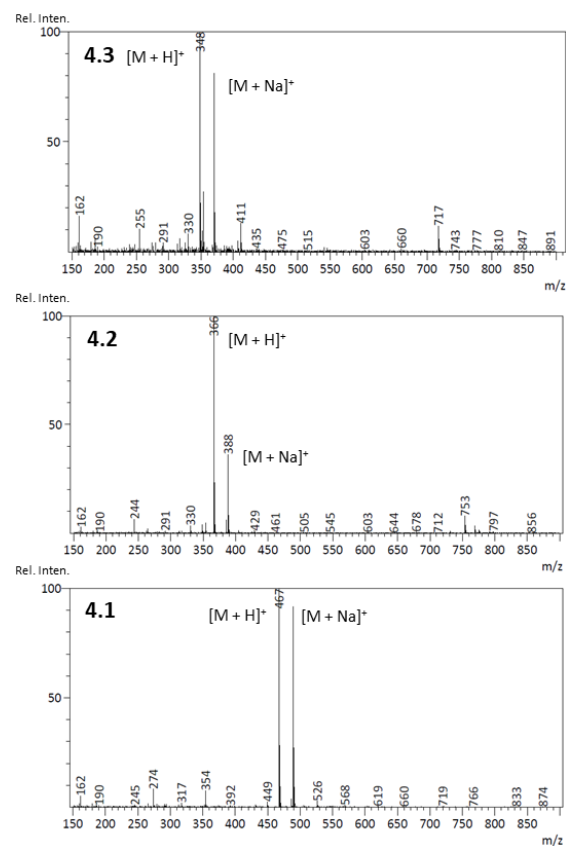
A**B**

Figure S4.47. DEBS Q176M/M179P mutant 24-hour incubation with substrate 4.1. (A) Extracted ion chromatogram for mutant c43 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 348$ (**4.3**), 366 (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2** and macrocycle **4.3** products, mass adducts are identified in each spectra.

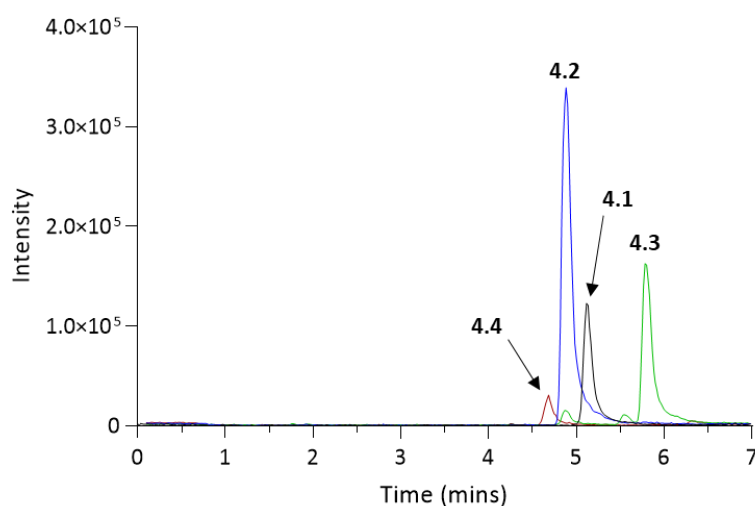
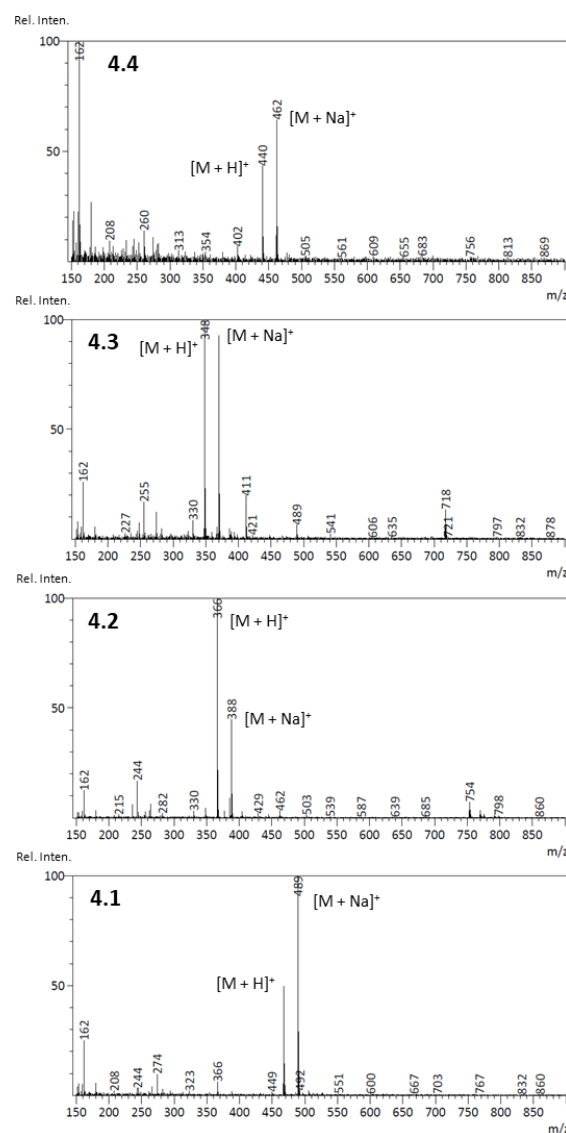
A**B**

Figure S4.48. DEBS Q176M/N180D mutant 24-hour incubation with substrate 4.1. (A) Extracted ion chromatogram for mutant c42 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with $m/z = 440$ (**4.4**), 348 (**4.3**), 366 (**4.2**) and 467 (**4.1**). Mutant shows a ~ 1.5 -fold increase in M/H ratio over WT DEBS-TE (**Table S4.12**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2**, macrocycle **4.3** and glycerolysis **4.4** products, mass adducts are identified in each spectra.

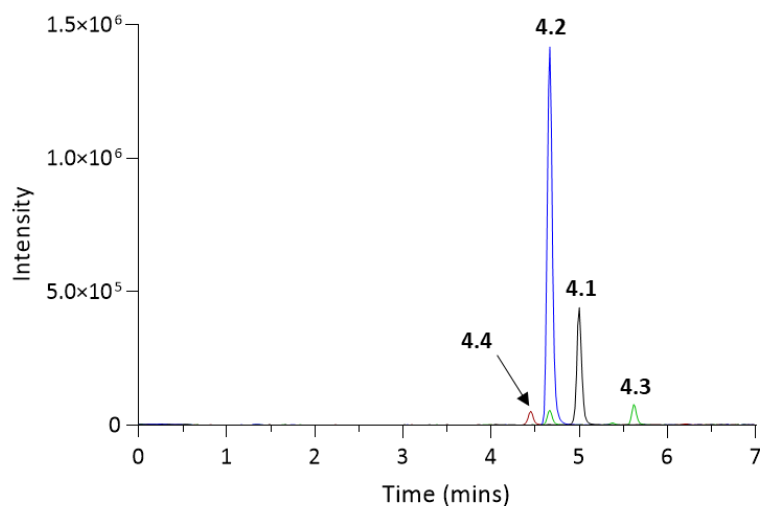
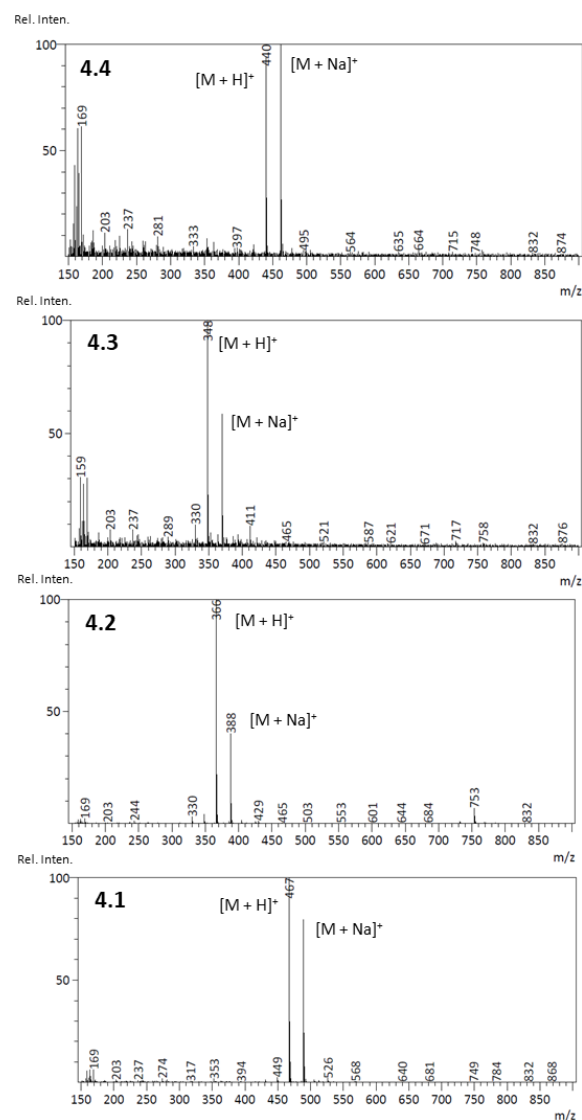
A**B**

Figure S4.49. DEBS Q176M/M179P/N180D mutant 24-hour incubation with substrate 4.1. (A) Extracted ion chromatogram for mutant c45 (5 μ M) assay after 24 hours of incubation with 1 mM of **4.1** in potassium phosphate buffer (50 mM, pH 7.4). Trace is an ion extraction for $[M + H]^+$ ions with m/z = 440 (**4.4**), 348 (**4.3**), 366 (**4.2**) and 467 (**4.1**). (B) Mass spectra for substrate **4.1**, hydrolysis **4.2**, macrocycle **4.3** and glycerolysis **4.4** products, mass adducts are identified in each spectra.

Mutant library screen AUC raw data from MZmine2

Table S4.3. Raw AUC for DEBS R32 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position R32 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	5.58	4.63	4.97
Row Identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	6980333.5	18965205	17627019.5
M1A02	875022	22011440.5	21844368.5
M1A03	0	15336617	34359051.5
M1A04	0	23716931.5	19395017.5
M1A05	0	20718279.5	27830848
M1A06	327195.5	22171812.5	28349355
M1A07	850343	28115771	17335168.5
M1A08	2245134	31052349.5	11355635
M1A09	0	21508107.5	25444324.5
M1A10	0	19330438	31191916.5
M1A11	0	21241640.5	27993767.5
M1A12	0	24851826	27164871.5
M1B01	0	16253247.5	33284038
M1B02	0	15488844.5	35411957.5
M1B03	0	9564320	38357031
M1B04	0	16435381.5	32406383.5
M1B05	0	14336597	35477105.5
M1B06	773191.5	19627877.5	29494419.5
M1B07	0	25579234	18896216
M1B08	338682	23881770.5	20292731.5
M1B09	324053.5	23210125	17680934
M1B10	0	19725230.5	26021706
M1B11	0	18533228	26156387.5
M1B12	0	26620867.5	24950348.5
M1C01	0	16741288.5	34439477.5
M1C02	0	12223447.5	37151094.5

M1C03	0	15465880.5	34868074
M1C04	0	18192363.5	29783174.5
M1C05	0	17516557.5	33601007
M1C06	317197	22539896	24540262
M1C07	0	23828395	20624856
M1C08	2564950	24095761.5	8473652
M1C09	632365	21202471.5	17594123
M1C10	0	22550066.5	24721428.5
M1C11	0	11660136.5	36754529
M1C12	0	20076352.5	31079315.5
M1D01	0	20234510.5	29584541.5
M1D02	718897	20128267	29387098
M1D03	0	11458517	35193875
M1D04	0	16652198	33169730.5
M1D05	0	12956336.5	33164045
M1D06	520874.5	21913378.5	27285525
M1D07	331558.5	20266420	12500980
M1D08	907283	22158571	4642209.5
M1D09	456225	24830615.5	15009882
M1D10	0	22823287	22383150
M1D11	0	19470835	27167481.5
M1D12	0	23184056.5	26371100.5
M1E01	0	6791171.5	42749027
M1E02	0	20236217	28885026
M1E03	0	6128165.5	39241788.5
M1E04	0	21409449.5	30897276.5
M1E05	1050838.5	18436401.5	37728232
M1E06	783005	20709097	24884045
M1E07	528433	23465952	21306669
M1E08	2464262.5	23928116.5	10531209.5
M1E09	324978.5	21547391	19289320
M1E10	0	19955470	23307459
M1E11	0	19234624	26167715.5
M1E12	0	22027945	25401361
M1F01	0	11322037.5	41107865.5
M1F02	0	14970942.5	33284090
M1F03	0	16292485	30542531.5
M1F04	0	12896477.5	35486967
M1F05	909459.5	24301772.5	20422987.5
M1F06	585319.5	21538577	28537065
M1F07	978936	24823379.5	19241125.5
M1F08	1788198.5	22785526	18403998.5

M1F09	684310.5	22185287.5	20896055
M1F10	345623.5	23921369.5	18520085
M1F11	0	23981974.5	16193328
M1F12	0	22107854	29536835
M1G01	4418932	25668856	13125406
M1G02	0	29360943.5	10695548.5
M1G03	0	21814177.5	25053884.5
M1G04	719624.5	20545485	26503074
M1G05	667785.5	27594538.5	12998897.5
M1G06	360279	25179343.5	19388109.5
M1G07	392326.5	23710882	24536077
M1G08	737722.5	24026391	20536236
M1G09	589456	24984423.5	21411217
M1G10	0	24887626.5	18909975.5
M1G11	0	23766324	17143567.5
M1G12	0	27759692	21093721
M1H01	1385293	29280584	6564949
M1H02	1181751.5	22602992.5	25705598.5
M1H03	0	13307216.5	36701105
M1H04	480626	18002464	32881855.5
M1H05	558547	24520129	27447351
M1H06	476897.5	28755462.5	12238426
M1H07	0	21807011	32531804
M1H08	2472551.5	31262842.5	11496411
M1H09	0	28582700	28167082
M1H10	1061901.5	32348147.5	11203349
M1H11	325235	30755707.5	20402704.5
M1H12	0	30730004	36623893.5
M2A01	0	1634114.5	46658691.5
M2A02	0	7547967.5	43517713.5
M2A03	0	16311402.5	39609519.5
M2A04	0	14592016	40890196
M2A05	0	14816452.5	40102217.5
M2A06	622713	18594860.5	39116036
M2A07	0	15474560.5	36824324.5
M2B01	0	17320191	32619312
M2B02	0	18063653	29800528
M2B03	622130	14647493	35341405
M2B04	317488	15819595	30435238.5
M2B05	0	14151594.5	35865397
M2B06	731805	15714362.5	32383467.5
M2B07	359232.5	16499585.5	32962277

M2C01	255046	16909384.5	33779013
M2C02	0	15296983	33793845.5
M2C03	0	8699746	37042326.5
M2C04	791070	11674572.5	33460315
M2C05	0	10605444	34431997.5
M2C06	0	13648022	31766396.5
M2C07	0	12745237.5	33246886.5
M2D01	0	17882294	33434418.5
M2D02	893516	16499699.5	31547935.5
M2D03	0	18742897	27406751
M2D04	0	15867864	37001918.5
M2D05	0	9195545	37828667.5
M2D06	1637423	19387369	29147162
M2D07	0	15056118	38863158
M2E01	0	16262931.5	32789837.5
M2E02	346873	16583960.5	30147876.5
M2E03	694614.5	19003909	30665823
M2E04	0	18053318	37112373
M2E05	484192.5	15576502	36785557.5
M2E06	1000965	13994903	35537786
M2E07	0	15658552	35851745.5
M2F01	0	17161531.5	32239781
M2F02	0	13461973.5	34770253.5
M2F03	485450.5	15168998.5	27714789.5
M2F04	0	14259198	31503528.5
M2F05	0	19689225	22991869
M2F06	0	22640771.5	21964733.5
M2F07	0	14337153.5	34370915.5
M2G01	0	13082870.5	34220685
M2G02	0	14338827.5	33700375
M2G03	0	17852608	27870427
M2G04	0	14463632	33383325.5
M2G05	0	22994631	17814531
M2G06	0	26833275	6801199
M2G07	0	13892731	35693979.5
M2H01	0	13577281	37744518.5
M2H02	491112.5	14093492.5	35848011
M2H03	551519	18379753	31276757.5
M2H04	0	12883248.5	37586348
M2H05	372849	23480177.5	24813068.5
M2H06	0	23889155.5	25971566

Table S4.4. Raw AUC for DEBS L35 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position L35 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	5.57	4.64	4.97
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	3790573.5	9614055	2597341.5
M1A02	615415.5	9862671	20165410.5
M1A03	0	7823559	19889938
M1A04	0	9193017.5	20109977.5
M1A05	0	8984946.5	19320734
M1A06	0	11940264.5	14817544
M1A07	0	9161823.5	23990061.5
M1A08	0	10296933.5	27620678
M1A09	0	14934370.5	29918997
M1A10	0	9035782	24253393
M1A11	0	11577519.5	23529787.5
M1A12	0	10484256	20095490.5
M1B01	0	7940301	13637227
M1B02	0	8207423	16441624.5
M1B03	0	7349718.5	18539484
M1B04	314988.5	9023500.5	16678958.5
M1B05	684061	8359406	16582385
M1B06	0	10284610.5	14007805
M1B07	0	14095592	11187877.5
M1B08	0	14163659.5	10252815.5
M1B09	0	9910391	14526613
M1B10	0	12198869.5	12622505.5
M1B11	0	8720957	17274705.5
M1B12	0	12609494	10338986.5
M1C01	0	10348646.5	9999083
M1C02	0	9716207.5	15623153.5
M1C03	421570.5	9941856	14660475
M1C04	1340485	10762476	13866379
M1C05	0	7801591	15858889.5

M1C06	0	8935512	13833233
M1C07	365196	14142903	5093072
M1C08	0	10055419.5	12916029.5
M1C09	0	12187949	7736623
M1C10	0	12131206.5	7888204
M1C11	0	14920468	7458013.5
M1C12	0	11250417	14805483
M1D01	0	8096456	13551499
M1D02	0	8286971	16766504
M1D03	456075.5	6556982.5	12089955.5
M1D04	347183	7425486	11207932.5
M1D05	0	5055638	12990140.5
M1D06	0	7803964.5	12498695
M1D07	0	11061405	7004892
M1D08	0	7536949	14488524.5
M1D09	351004.5	14589102	3416292.5
M1D10	0	10763081	10189512
M1D11	0	9994203.5	16465405
M1D12	520986	8631810	16748418
M1E01	0	9450075	11757959
M1E02	0	9348262	16023579
M1E03	0	7101340.5	12888629
M1E04	0	7519142.5	13444816.5
M1E05	0	7293181.5	12999047.5
M1E06	1142629.5	9973567.5	10829080.5
M1E07	0	9101673	12897541
M1E08	0	6096700	17956668.5
M1E09	0	8303519	14534626.5
M1E10	0	11388444.5	8466124.5
M1E11	636316	21123508	11480891
M1E12	394380.5	9726155.5	16268072.5
M1F01	0	7685507	12155387.5
M1F02	0	11537861.5	10774344.5
M1F03	0	8547945	10626392
M1F04	0	9581113.5	10398395
M1F05	0	7927582.5	12780809
M1F06	338133.5	8811365	12707995
M1F07	0	5834263	15659911
M1F08	0	6728201.5	13952894
M1F09	0	4937192	15323118
M1F10	0	5823658.5	16840938
M1F11	0	6939203	19896107

M1F12	0	7485151	20199075
M1G01	0	5382490.5	18422173.5
M1G02	0	7342172.5	16967262
M1G03	0	9607299.5	12393125.5
M1G04	0	10907346	12028833.5
M1G05	0	7148383	16193431.5
M1G06	0	10728456.5	14588135
M1G07	0	8906691.5	17280937
M1G08	0	7674868.5	16736473.5
M1G09	0	12657768.5	11944524.5
M1G10	0	6406140.5	21173293
M1G11	621990.5	5704999.5	21855066.5
M1G12	0	8821773	18693672
M1H01	0	5661660	16533286.5
M1H02	0	7842024.5	20684485
M1H03	0	7861618.5	17197370.5
M1H04	0	13222429	11085130
M1H05	0	13383905	8574943
M1H06	0	7679815	21581496
M1H07	0	11747002.5	16495745.5
M1H08	0	9192440.5	19595182
M1H09	0	7157797.5	21172755
M1H10	0	6822210.5	22092143
M1H11	0	6631473	23801516.5
M1H12	0	6944385.5	22184169.5
M2A01	0	9359122	17867786
M2A02	1687901.5	30107834	55023497
M2A03	0	19134707.5	40013435
M2A04	0	8810793	17197637.5
M2A05	494028.5	20448102	33907030
M2A06	0	14201240.5	29964602
M2A07	0	7956812	19563923.5
M2B01	0	17603374.5	26201968
M2B02	353067.5	14887378.5	23658924
M2B03	1464096.5	18426173	32974413.5
M2B04	585506.5	16234701.5	24188357.5
M2B05	0	14077675.5	23699496
M2B06	0	9202050	28399599.5
M2B07	0	1846658.5	6784781.5
M2C01	0	10681189.5	23031090.5
M2C02	0	10429753	20354406
M2C03	0	7685308	22009095.5

M2C04	0	8773134	18798541.5
M2C05	0	9925234.5	17868679.5
M2C06	0	7672059.5	19675444.5
M2C07	0	5300050	13855573.5
M2D01	0	11191489.5	29538928
M2D02	0	11855995	28050153
M2D03	1043704.5	9807184	17750731.5
M2D04	0	8568879.5	18464078
M2D05	0	10045538.5	16873938.5
M2D06	0	8589552.5	20817459.5
M2D07	0	3605534	7370688
M2E01	0	10942954	19065528
M2E02	0	10461607.5	17808346
M2E03	0	8618784.5	12774042
M2E04	0	7192641.5	13937109
M2E05	0	7421752	14189988
M2E06	0	5845784	18318607
M2E07	0	2137511	8119743
M2F01	0	9678671.5	22348041
M2F02	0	10714849	19363788
M2F03	0	7196956.5	10858548
M2F04	0	6334090.5	19827910.5
M2F05	0	7525958	18773708.5
M2F06	0	6689878.5	18434050
M2F07	0	4350100.5	11067040
M2G01	0	10423852	29417660
M2G02	0	8179358.5	25020143.5
M2G03	0	9962497.5	27531853
M2G04	0	7578460	16837109.5
M2G05	0	10142522.5	25400462.5
M2G06	558373.5	10192935	23967853
M2G07	596409	6236225.5	15570696
M2H01	0	4301221	13104726
M2H02	0	9377856.5	27821026.5
M2H03	0	10262933.5	27266426.5
M2H04	0	8756448.5	25474487.5
M2H05	0	9852615.5	22247301.5
M2H06	0	513702.5	26126772

Table S4.5. Raw AUC for DEBS A77 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position A77 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	5.59	4.65	4.98
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	6778661.5	16086805.5	19067426.5
M1A02	959148.5	27916285.5	28088272
M1A03	0	30870069.5	23151615.5
M1A04	0	25104048	30295940
M1A05	0	35364211.5	47877905
M1A06	0	21321350.5	34913381
M1A07	0	28872514.5	64719040.5
M1A08	0	29095818	39934684
M1A09	0	32044737.5	15504721
M1A10	0	32260499	12562634
M1A11	0	29835294.5	21061771
M1A12	0	22632064.5	35140128.5
M1B01	0	18164169.5	27725677
M1B02	0	23194269	31815900
M1B03	0	29202593.5	27485784
M1B04	0	23310594.5	32827693.5
M1B05	0	27188146	8724239
M1B06	0	22802723.5	35792569
M1B07	0	32860937.5	9633724
M1B08	0	21221053	35729236.5
M1B09	0	31241693.5	11857141
M1B10	0	33841647.5	16489992
M1B11	1844153	28034683	21954019.5
M1B12	0	18255275	37184654.5
M1C01	0	13765289.5	34417101
M1C02	0	19672027.5	37350655.5
M1C03	0	26089580	31070174
M1C04	0	18108401	41288745.5
M1C05	0	25401926	20569967

M1C06	0	29217679.5	28782654.5
M1C07	0	26915510.5	30759007
M1C08	0	30019448	19141361
M1C09	0	25186788	32786001
M1C10	307927	28261472	23550189.5
M1C11	0	19339749	36888070.5
M1C12	0	17992329.5	37058891.5
M1D01	0	15949523	32039797.5
M1D02	0	23527168.5	33359886.5
M1D03	0	21901561	39741808
M1D04	0	23506193.5	35560681
M1D05	0	26207899	13296917.5
M1D06	453500	26661931	28222259.5
M1D07	0	25371455.5	37290446
M1D08	0	15549181	38476134
M1D09	268094	28762573.5	26800747
M1D10	0	23590128.5	34110983.5
M1D11	0	17363604.5	38512085
M1D12	0	20150347.5	35735190.5
M1E01	0	22930305	24529913.5
M1E02	0	28278522.5	20195249.5
M1E03	0	29892440.5	27933413
M1E04	0	26589211.5	23472137
M1E05	0	23484223	12674690.5
M1E06	0	31361268.5	15502858.5
M1E07	0	24486743.5	32200068.5
M1E08	0	12155577	40350209
M1E09	0	23204469.5	34433641.5
M1E10	0	22911897	34784415
M1E11	0	22254112.5	31895932.5
M1E12	0	23544188.5	33372382.5
M1F01	0	24174214	20643876
M1F02	0	28962212.5	17269629.5
M1F03	0	23953084	37433846.5
M1F04	0	22879863	33001929
M1F05	0	23676901	13795432.5
M1F06	0	29267145	22233233.5
M1F07	0	23497268.5	30817939.5
M1F08	0	21098969.5	36414518.5
M1F09	0	21154827	36125175
M1F10	0	23966457.5	34948732.5
M1F11	0	20751765.5	37154081

M1F12	0	15280209.5	41525812
M1G01	0	16067820	34621047
M1G02	0	25472361	28666037
M1G03	0	30110407	26490326
M1G04	0	23682797.5	32107920.5
M1G05	0	23325341	18482626.5
M1G06	2124066	28612515.5	13246600
M1G07	496401.5	25906353.5	27537858.5
M1G08	0	23246849.5	30650121.5
M1G09	0	23771771	32614749
M1G10	0	18827857	39224366
M1G11	0	19341518	37585009.5
M1G12	0	21601454	35312526
M1H01	0	22259891.5	28225459
M1H02	0	20989557.5	34830144
M1H03	0	28556475.5	17012343.5
M1H04	0	18509614	36001690
M1H05	0	10003169	877412.5
M1H06	0	27972947.5	22087948.5
M1H07	0	21000787.5	29113376.5
M1H08	0	24658460.5	26174961
M1H09	0	28626957	9924998.5
M1H10	0	22758249.5	30007872.5
M1H11	674137.5	16154236.5	38401643
M1H12	0	28175238	19732804
M2A01	0	32089692.5	49935166
M2A02	672204	32025250.5	51733790
M2A03	0	26904808	52568888.5
M2A04	0	26903165	48655690
M2A05	0	33786969.5	27641790.5
M2A06	0	22197433	49833249
M2A07	0	17314108.5	47247110
M2B01	0	24981183	38383217
M2B02	0	26389744.5	36431162
M2B03	0	27510312	37691454.5
M2B04	0	34376468.5	14417248.5
M2B05	0	33076794	4721857.5
M2B06	0	19247654.5	38647967.5
M2B07	0	16759861.5	43122135.5
M2C01	0	25238370	36297542
M2C02	0	22988563.5	34978777
M2C03	0	29208839.5	28679113

M2C04	0	23835183.5	31120898.5
M2C05	0	21499913.5	36388094
M2C06	0	22688309.5	35137767.5
M2C07	0	8129699.5	47842265
M2D01	0	25039335	34909792
M2D02	0	10904556.5	45243277
M2D03	0	21729677.5	36948867
M2D04	0	10372041.5	45563380.5
M2D05	0	25827537.5	10081586.5
M2D06	0	23043578.5	33089893
M2D07	0	2965645	48172721.5
M2E01	0	24759290.5	31151398.5
M2E02	0	16700959	41537600.5
M2E03	0	13612896.5	39539391
M2E04	0	28341053	27261385.5
M2E05	0	26202232.5	24801913
M2E06	0	29990668	24543244
M2E07	0	5405713	43490323.5
M2F01	0	28057078	20490733.5
M2F02	0	12769799.5	42074895
M2F03	0	21010891.5	32601443.5
M2F04	0	9146008	43599792
M2F05	0	28636735.5	16904765.5
M2F06	0	30620585.5	11687610.5
M2F07	0	8895199	44696790
M2G01	0	28464405.5	20064817.5
M2G02	0	25033928	28515356.5
M2G03	0	26779958	28817352.5
M2G04	0	19248054	35436021
M2G05	0	19541611.5	38419608
M2G06	0	27511725.5	28920585.5
M2G07	0	15704774.5	42477862
M2H01	0	30682896	19357906
M2H02	0	18100824.5	41753091
M2H03	0	13884674.5	43372264
M2H04	0	25307982	17286842.5
M2H05	0	25700525	32460841.5
M2H06	402865.5	24017272	35052143

Table S4.6. Raw AUC for DEBS Q176 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position Q176 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	5.57	4.66	5.00
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	2254998	6232770.5	3954820.5
M1A02	520099	10599505.5	20560738
M1A03	0	7074821.5	22563973
M1A04	0	12538838.5	39275982.5
M1A05	1636173.5	14295215	34172196.5
M1A06	585887	15116553	42182778
M1A07	0	11099466.5	35895567
M1A08	0	9483199	36775933.5
M1A09	0	12968281.5	41141003
M1A10	400832.5	13509568	46058105.5
M1A11	332247	11784906.5	41012349.5
M1A12	0	13096472	40188749
M1B01	0	6252456.5	14417556.5
M1B02	931721.5	15006093.5	34409208
M1B03	648430.5	15131124	41418088
M1B04	0	10571389	42991857.5
M1B05	545836.5	12336964	41588956.5
M1B06	1006935.5	13290211	39122474.5
M1B07	0	18243051	36196933
M1B08	0	13840547	38910095.5
M1B09	0	11874978.5	38575296
M1B10	0	13977757	40878330.5
M1B11	0	10571768	37988497.5
M1B12	650778.5	18788958	30336820.5
M1C01	0	6384208.5	18450440
M1C02	322860.5	11502934.5	34758934.5
M1C03	0	5500401.5	18841593
M1C04	0	4286708	20589168
M1C05	0	4484930.5	18562856.5

M1C06	485590	5196012.5	19432791.5
M1C07	0	4671253.5	19124388
M1C08	0	3561339.5	15026254
M1C09	0	8471659	35107824.5
M1C10	0	4625899	19651793.5
M1C11	0	10516833.5	33545721
M1C12	937789	12436913.5	33669523.5
M1D01	0	6608981.5	21528125.5
M1D02	802542	15576527	37040017
M1D03	0	5151028.5	19903470
M1D04	0	4623124	22666758
M1D05	0	4389885	22220353
M1D06	0	4575901	20987266.5
M1D07	0	6087628.5	22312104.5
M1D08	0	3886698	15101851.5
M1D09	0	3755536	16552425
M1D10	0	4026853.5	18803440.5
M1D11	0	9559410.5	36130524.5
M1D12	0	7723117.5	36942547.5
M1E01	0	4035615.5	23422386.5
M1E02	1416401.5	18170358.5	33265018.5
M1E03	0	6364058.5	17587897.5
M1E04	0	4865019	20777022
M1E05	511367.5	6042273	17475610.5
M1E06	607733.5	6588103	18884605
M1E07	0	5112045.5	21321520
M1E08	0	2611566	19339354.5
M1E09	0	4179970.5	18865938.5
M1E10	0	4666927.5	15507254.5
M1E11	0	11735514	37125349
M1E12	0	10463317.5	28064856.5
M1F01	0	7350871	24605703.5
M1F02	876926.5	17497002.5	31534098.5
M1F03	411555.5	7510043	17582343.5
M1F04	349777	6246649	22031692.5
M1F05	1086757	7543551	15979400
M1F06	955613.5	7060717.5	18067305
M1F07	0	4592849	22186546.5
M1F08	0	4660050.5	21516909
M1F09	0	5516344.5	19470818
M1F10	0	4544611.5	15230690.5
M1F11	0	12890999	38756133

M1F12	0	10151466.5	33803532
M1G01	0	6980234.5	24751064.5
M1G02	805969	16017521	35345620.5
M1G03	1366903	18502413	37371926
M1G04	1481693	17076449	36890575.5
M1G05	1206154.5	17937363.5	38913845.5
M1G06	326732.5	14044336.5	39358075.5
M1G07	652861.5	11981475	41666378.5
M1G08	0	10355136	41795744.5
M1G09	0	13831594.5	37445331.5
M1G10	0	10751227	39953337.5
M1G11	0	7739143.5	43107213.5
M1G12	0	9204981	30252864.5
M1H01	315339	9803779	21279104
M1H02	4524291.5	17349668.5	34821564
M1H03	1309049	17981574	39200464.5
M1H04	680326.5	17539655	39846868.5
M1H05	1831887	20090051.5	41478861.5
M1H06	587688	18253977.5	42185921.5
M1H07	0	12892796	43756198.5
M1H08	320060	11450004.5	39627084
M1H09	323012.5	13313102.5	39393081
M1H10	0	13724616	38113571.5
M1H11	0	12540438	41474306
M1H12	433459.5	21140134	50280063
M2A01	944527.5	20339488.5	55853148
M2A02	682686	15715068	53481481
M2A03	0	16208951.5	54997996
M2A04	0	13369560	50286388.5
M2A05	0	14088812.5	50527326.5
M2A06	0	14670829	48892730.5
M2A07	0	12885852	47616909
M2B01	0	13270940	44703881.5
M2B02	0	11876462.5	41984210
M2B03	0	10957832	43103398.5
M2B04	0	9707942.5	43530411
M2B05	0	1957649.5	10170096
M2B06	0	4666162.5	18999664.5
M2B07	0	4746534	18349141.5
M2C01	412546	15753494.5	41731716.5
M2C02	0	10803291.5	41013393
M2C03	0	4021727	20070160.5

M2C04	0	4134134	20822201.5
M2C05	0	3535020	14786007.5
M2C06	0	5189343.5	18321808.5
M2C07	560991	5737538.5	16308328
M2D01	0	13394790.5	45887328.5
M2D02	0	10678942	42397633.5
M2D03	0	3188364.5	16615127.5
M2D04	0	3503214	20187227
M2D05	0	3434348.5	16804811.5
M2D06	0	3967812.5	18130289
M2D07	0	5205159.5	18995131
M2E01	0	14434122.5	44540326.5
M2E02	0	12929150.5	40787070.5
M2E03	0	2726966.5	16022058.5
M2E04	0	3226758	18817900.5
M2E05	0	3974384.5	18901310
M2E06	0	4947381.5	17956108.5
M2E07	0	4592579.5	19279073
M2F01	0	14102712.5	41523301.5
M2F02	0	12011793.5	40905186
M2F03	0	3436412.5	18045087
M2F04	0	3998164	18956558
M2F05	0	4040745	18033688.5
M2F06	0	3445336.5	13553105.5
M2F07	0	5395508.5	20374148.5
M2G01	0	14896825	42793847
M2G02	0	13292750	42360075.5
M2G03	0	4216860.5	18340378
M2G04	0	4888272	20056622
M2G05	0	4337078	19204636.5
M2G06	0	4804432.5	20198722.5
M2G07	0	3948189	21225343.5
M2H01	0	19093037	46971946.5
M2H02	0	16720538.5	44466879.5
M2H03	0	12334891	38787580.5
M2H04	0	13281234.5	38565517
M2H05	0	15808986.5	42545699
M2H06	0	14200657	41812117.5

Table S4.7. Raw AUC for DEBS M179 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position M179 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	5.59	4.63	4.97
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	6972535.5	17737249	20826630
M1A02	729602.5	23836104	25949240
M1A03	0	20105937.5	33511636
M1A04	0	24493903	27499543
M1A05	0	24773510	26087291.5
M1A06	0	22647082.5	33193164.5
M1A07	520814	26119400.5	19869098
M1A08	0	22372146.5	25306692.5
M1A09	470324.5	26199035.5	22725574.5
M1A10	0	27651836.5	27425315
M1A11	721267	28232677.5	20220925.5
M1A12	0	31193099	9100640.5
M1B01	0	28423795.5	5275847.5
M1B02	0	28144254.5	17845570.5
M1B03	0	29351026.5	11393263
M1B04	0	28681762	10912621
M1B05	0	31379060	4469827
M1B06	302104	30361428	5589327
M1B07	455237.5	31402247	6561068
M1B08	0	31160909	1144164.5
M1B09	0	24149373	29445002.5
M1B10	0	26948243.5	23725709
M1B11	0	24422850.5	28708250.5
M1B12	0	27643251.5	18262645
M1C01	0	23783451.5	23745635
M1C02	0	26755200.5	20771191.5
M1C03	0	28773677	24272431

M1C04	0	24430465.5	30722430
M1C05	0	21337370	34241157.5
M1C06	563962	29649015.5	14089581.5
M1C07	3484769	30398576.5	9442278
M1C08	0	28705274.5	22427483.5
M1C09	0	29483655.5	14111831
M1C10	0	21456345.5	32108226
M1C11	0	23780197	30385044.5
M1C12	0	28410782	17102493.5
M1D01	0	22175949.5	27782874.5
M1D02	0	27838328.5	17069135.5
M1D03	0	24173016	28834758
M1D04	0	30113719	16801977
M1D05	0	23176209.5	35519561.5
M1D06	0	32011118.5	4049516.5
M1D07	393529	27378222.5	27224110.5
M1D08	0	26758412	29136209
M1D09	0	29803266.5	14487750
M1D10	0	19884354	35309388.5
M1D11	0	26296058	25442976.5
M1D12	0	28679445	18800366
M1E01	0	24625138	24907070
M1E02	338449.5	26271788	20177073.5
M1E03	0	16217744.5	40257988
M1E04	3109092	29761091	8494937
M1E05	0	27681105	27994488.5
M1E06	0	21062679.5	32914399
M1E07	0	21447417	34772607.5
M1E08	489329	30906293.5	12965983.5
M1E09	0	30493465.5	15272411
M1E10	0	31228948	14241777
M1E11	0	31219617.5	7284998.5
M1E12	458680	32877012	4228213.5
M1F01	0	28634860	8009496
M1F02	0	27206825.5	16187800
M1F03	0	28560952.5	19015290.5
M1F04	0	29883163.5	13754979
M1F05	0	31907605	4744090
M1F06	0	26583721.5	26372921
M1F07	542690	27846857.5	19712675
M1F08	0	27168325.5	22161084.5
M1F09	0	24466596	30381320

M1F10	0	27750460	17807723.5
M1F11	0	30236988.5	10251273.5
M1F12	0	27373114.5	24135523.5
M1G01	0	29938073.5	5204166
M1G02	0	25615781.5	28005405
M1G03	405697	31449673.5	6618227
M1G04	3764646.5	30780438	2117038
M1G05	0	29412815.5	16815740.5
M1G06	0	31840931.5	6219914.5
M1G07	3890920	30864619.5	1049559
M1G08	0	32248321	3670959.5
M1G09	0	28814209	16437510
M1G10	0	29006260.5	22370650.5
M1G11	0	31289479.5	11579881.5
M1G12	0	31647922.5	3770387
M1H01	0	20215892	11297300.5
M1H02	422060.5	31040421	9174696
M1H03	0	29391784.5	5913203
M1H04	4497462.5	29819924.5	963508.5
M1H05	3079950	30092406	3057716.5
M1H06	0	30668255.5	10453030.5
M1H07	0	31355480.5	10639900
M1H08	0	32178872.5	5129814
M1H09	0	31687395.5	749136.5
M1H10	0	23247963	30722662
M1H11	0	29707643	12957653.5
M1H12	0	31214388	7846796
M2A01	0	26118771	43448071.5
M2A02	4090811	34049207	34297775.5
M2A03	0	20854415.5	41107167
M2A04	0	21584211.5	42537252.5
M2A05	348371	23743956	44835248
M2A06	0	22681759	44573636.5
M2A07	0	10472347.5	44972242
M2B01	1655295.5	25910272.5	25735877
M2B02	0	22610577	32972177
M2B03	0	27416206.5	22757485.5
M2B04	0	26900636	23230129
M2B05	0	26627461.5	27248216.5
M2B06	0	14512767	42383957
M2B07	0	12207455	43397637.5
M2C01	0	34142170	3349232

M2C02	0	27487923	26996350
M2C03	0	22055114.5	32167394
M2C04	0	23663781.5	32038138.5
M2C05	0	22237662	32942003.5
M2C06	0	19989769	35433417
M2C07	0	12754157.5	42218154
M2D01	0	23036932	34520847.5
M2D02	0	24205313.5	32041810.5
M2D03	0	23593143	27579644
M2D04	0	21333882	34669606.5
M2D05	0	24438152	29799642.5
M2D06	0	25288270	23804181.5
M2D07	1566037.5	18034950	38999262.5
M2E01	0	22973477.5	31529461
M2E02	395483.5	21739236.5	34661781.5
M2E03	0	26784352.5	22562011
M2E04	0	16520696.5	34407092
M2E05	0	25975591.5	19428378
M2E06	2083062.5	27023775.5	17338806.5
M2E07	0	20963947	33728450.5
M2F01	0	22964057	34679029.5
M2F02	0	25150819.5	22222847
M2F03	0	22891345.5	30554585.5
M2F04	0	17512658	36395341.5
M2F05	0	25572964.5	25916684
M2F06	0	25440908.5	21736549
M2F07	0	16099901.5	39739213
M2G01	0	25861463.5	26675465.5
M2G02	0	12903006	40145062
M2G03	0	32170004	5107331
M2G04	0	30477730.5	7160372.5
M2G05	0	30741499.5	8231400.5
M2G06	371614	21657279	32063180
M2G07	0	14505909	42771726
M2H01	0	22237244	35499502.5
M2H02	0	27042558	25095116
M2H03	0	25777645	23776198.5
M2H04	0	18819755.5	36007444.5
M2H05	0	27956301	23027413.5
M2H06	0	10940300	42497313

Table S4.8. Raw AUC for DEBS N180 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position N180 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	5.57	4.63	4.96
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	3745152	9625547	7588472
M1A02	530240	8137592	37592471
M1A03	0	13323647	37038112
M1A04	0	8634466	35126441
M1A05	0	13118429	33308943.5
M1A06	0	8961966	39461104.5
M1A07	0	7079538	38462580
M1A08	0	6601955	36186643.5
M1A09	0	4989894	41067301.5
M1A10	0	13114245	46580760
M1A11	312376	15041673	48570269
M1A12	3236767	15347107	44699086
M1B01	0	5528598	19113929
M1B02	0	3876853	40641018.5
M1B03	0	8103578	38116875.5
M1B04	0	4728840	41786157.5
M1B05	0	9589926	42454945
M1B06	0	6702325	44663086
M1B07	0	12875000	43255279
M1B08	389731	18196026	37871106.5
M1B09	0	9633584	38517908
M1B10	0	12185880	41766581
M1B11	318647.5	14024701	41877569.5
M1B12	980763	14381354	40631118
M1C01	610930.5	8886187	17389826.5
M1C02	0	4937233	40048964.5
M1C03	0	8179426	43302899.5
M1C04	1060384	7008169	42591629
M1C05	0	11069268	34658808

M1C06	304514.5	13882193	27519503
M1C07	333119.5	14300429	37396252
M1C08	357761	15272978	35431448
M1C09	0	7037164	41791894.5
M1C10	0	9625359	42123343.5
M1C11	0	10863659	40696470.5
M1C12	729320.5	13603157	42945684
M1D01	0	6737839	28692774.5
M1D02	0	7777622	36376749
M1D03	0	11455594	35991404
M1D04	0	7676439	42180122.5
M1D05	0	10966088	33373081
M1D06	0	14274101	39525778.5
M1D07	358299	14190142	37490715.5
M1D08	0	21439488	32972202
M1D09	0	10588038	39275965.5
M1D10	0	6117593	44820243
M1D11	0	8269723	42935632
M1D12	422753	16122517	38280101.5
M1E01	0	7272951	29055277
M1E02	0	7451780	32111601
M1E03	0	9401499	35641641.5
M1E04	380953.5	9654528	35322884
M1E05	0	13854367	43280626
M1E06	0	12568096	39356664
M1E07	0	12660239	37768443.5
M1E08	0	14721513	35054556.5
M1E09	0	12154665	37504962
M1E10	598066	14559930	37023632.5
M1E11	0	12988437	38272211.5
M1E12	0	16161082	41413932.5
M1F01	0	3135953	34157174.5
M1F02	0	10382996	31088829
M1F03	486723	11426349	31039037
M1F04	989674.5	14607258	43614291
M1F05	0	9610542	39465978
M1F06	0	10945770	39895009
M1F07	0	11552821	38166975
M1F08	768934.5	15936790	34292965.5
M1F09	0	9132059	39382276.5
M1F10	0	10877040	42680430
M1F11	0	9401497	41943677.5

M1F12	0	12821707	41107737
M1G01	0	4600012	34788915.5
M1G02	0	12334871	49179065.5
M1G03	0	3394018	38061291
M1G04	0	4434718	36012381
M1G05	0	6894098	36196500
M1G06	359942	12289252	44014999.5
M1G07	0	11119999	42280322
M1G08	457036.5	13724192	36311135
M1G09	0	8794567	39352904
M1G10	0	9710935	39745906.5
M1G11	0	14368210	37098201.5
M1G12	0	14007099	41532413
M1H01	0	13971975	31859585
M1H02	0	15417043	49481759
M1H03	0	7750859	35285316
M1H04	0	10561518	43217483.5
M1H05	0	12112069	43518366
M1H06	0	15220812	47524803
M1H07	0	12270782	45115092
M1H08	0	13991969	42608906
M1H09	0	13655314	41508482.5
M1H10	1065929	17862893	45032428.5
M1H11	0	13096657	38239548.5
M1H12	0	14417689	39387580.5
M2A01	0	7840119	36819849.5
M2A02	0	4044491	29872906.5
M2A03	0	7067352	39781329
M2A04	0	13237394	46840396.5
M2A05	0	18526903	41220649.5
M2A06	0	13386929	40220630.5
M2A07	0	17643612	25474589
M2B01	0	14255930	49511816.5
M2B02	0	7014580	46776356.5
M2B03	368380.5	10660741	38653396.5
M2B04	0	10252919	41409882.5
M2B05	0	3967339	34673694.5
M2B06	1121794	16020764	37101801.5
M2B07	0	12914703	43362778.5
M2C01	0	13331377	44767199
M2C02	0	8611022	40277993.5
M2C03	0	11556930	39229060

M2C04	0	8463766	38217902
M2C05	0	5352076	41602972.5
M2C06	0	4788503	26272210
M2C07	0	10695060	39736512.5
M2D01	0	6826341	33332528
M2D02	308167	18671336	36548754.5
M2D03	0	8435262	39887426
M2D04	0	8310237	40272438.5
M2D05	0	9443976	38525462
M2D06	0	13825911	35003311.5
M2D07	0	12272579	40030897
M2E01	0	13749909	48846963.5
M2E02	0	10828314	39583416.5
M2E03	0	9931972	38347523
M2E04	0	8406638	40967829.5
M2E05	0	9833884	37227996
M2E06	0	9215381	36200527
M2E07	0	11934781	39031532
M2F01	0	10112145	46164571.5
M2F02	0	10040949	41395719.5
M2F03	0	6431003	44251523
M2F04	0	5837619	42766892.5
M2F05	0	11671010	36978035
M2F06	0	9409412	38350792.5
M2F07	702276	13915040	36079381
M2G01	0	4371125	27072506
M2G02	0	14349400	42228858
M2G03	0	11074836	39062836
M2G04	0	8636331	42314160
M2G05	0	0	0
M2G06	0	14070290	36221928
M2G07	0	12716161	39492293
M2H01	0	11459360	47017717.5
M2H02	0	7891184	42403886
M2H03	0	14058419	43032476.5
M2H04	0	11398888	43044332.5
M2H05	0	10341060	44832287.5
M2H06	0	9903592	42902102.5

Table S4.9. Raw AUC for DEBS L206 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position L206 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (4.3), hydrolysis (4.2) products and substrate (4.1) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	3.66	2.94	3.22
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	1977389.5	5940917	4372841.5
M1A02	6648298.5	25115070	6286253.5
M1A03	0	11013266	22271049.5
M1A04	0	5427809	7244282
M1A05	0	15145266	13649646
M1A06	6912137.5	19686575	7935142.5
M1A07	0	15117937.5	13923524
M1A08	0	14267595	22985067
M1A09	3953505	49936041	3683603.5
M1A10	0	10536634.5	31500534
M1A11	0	25428958	8543418.5
M1A12	0	16966783.5	12804005.5
M1B01	0	5490821.5	3761366
M1B02	0	8284199.5	0
M1B03	0	13840262.5	807935.5
M1B04	0	5566634	6596677
M1B05	0	7475130.5	377419
M1B06	3608555.5	14839200.5	1363380
M1B07	1009519.5	15906884.5	1584864.5
M1B08	0	16595567	2730691.5
M1B09	0	10308828	14639137
M1B10	0	16289614.5	8967932
M1B11	1256103.5	15741379	0
M1B12	437792.5	16194561.5	438426.5
M1C01	3069094	10724057.5	1710134.5
M1C02	4403538	16081996	0
M1C03	0	4691545	6422195.5
M1C04	0	8851905.5	413719
M1C05	0	5406048	1811292.5

M1C06	937038.5	5387040.5	0
M1C07	1416687.5	7692856.5	581253.5
M1C08	0	5536565	7098307.5
M1C09	0	4566462	8537977
M1C10	1607023.5	8256271.5	936197.5
M1C11	485370	18017743.5	0
M1C12	0	6775383.5	586711.5
M1D01	0	5024012.5	2276987
M1D02	0	7758341.5	0
M1D03	0	8693632	1331381.5
M1D04	0	9445918.5	806300
M1D05	0	5085257.5	7208461
M1D06	0	6139803.5	6098574
M1D07	489124.5	7904367	979207.5
M1D08	0	8462614	0
M1D09	1424681	6825022.5	486175.5
M1D10	0	4471146	1787915.5
M1D11	505337.5	19945760	0
M1D12	0	15019794	0
M1E01	0	6244963	18348898.5
M1E02	0	11832298.5	677339.5
M1E03	0	7647180	844078
M1E04	1962463.5	8809797.5	370767
M1E05	1483716	8165764	1929422
M1E06	0	3805381	3651005
M1E07	0	3994618	825245.5
M1E08	0	2473158	2389975
M1E09	0	4772325.5	0
M1E10	0	7756327.5	4941932.5
M1E11	0	7099987.5	18336647.5
M1E12	524365	16707940	1360071
M1F01	0	10340849.5	12046229.5
M1F02	0	10246688.5	10355863
M1F03	0	7171728.5	5735111.5
M1F04	0	10424670.5	937506.5
M1F05	0	8166346.5	3694438.5
M1F06	0	8998078.5	1896811.5
M1F07	0	7937552	3865022.5
M1F08	0	6816957.5	5829735
M1F09	0	8707074	843152
M1F10	0	5665133	584947.5
M1F11	384273	18001761.5	638342.5

M1F12	0	14455040	9328021.5
M1G01	0	4361743.5	7022103
M1G02	0	17508991	15595592
M1G03	0	6610886	8757572
M1G04	0	4383994.5	7992989.5
M1G05	0	7636258.5	1726291
M1G06	0	5884984.5	0
M1G07	507118	8740680	2318497.5
M1G08	0	9341826.5	0
M1G09	0	9724063	8420996
M1G10	0	4206552.5	4415751
M1G11	449337	16409235	0
M1G12	0	8640484	16567681
M1H01	0	9887480	1097394.5
M1H02	0	5516904.5	6000949.5
M1H03	0	4753074.5	9208178.5
M1H04	0	1599125.5	10725308.5
M1H05	0	7981999	450742.5
M1H06	0	5280706	6553772.5
M1H07	0	11538483.5	4962280
M1H08	471609	16403254	0
M1H09	0	5707113	5554003.5
M1H10	0	9031951	13664448.5
M1H11	0	13735818	1804905
M1H12	412707	14102688.5	0
M2A01	1998509	7933028	2247253.5
M2A02	2226814	9015035	5003931
M2A03	0	4984074.5	12458317.5
M2A04	0	7721852.5	26413711.5
M2A05	0	2693233.5	12542049
M2A06	0	14527088	20291226.5
M2A07	3710989.5	22121444	9973814
M2B01	1160823	12732575	4240774.5
M2B02	0	4288523	7267119.5
M2B03	0	5038272	8654534
M2B04	0	1356071	5777479
M2B05	0	4861229	5590713
M2B06	0	7359628	3832004.5
M2B07	380178	6716336.5	6066245.5
M2C01	0	11587401	11360198
M2C02	0	3013438	5539981
M2C03	0	5863938.5	7255479.5

M2C04	0	4173688	8664865
M2C05	0	7165649	3554846.5
M2C06	0	6363509.5	5288880.5
M2C07	0	3179398	2181238
M2D01	0	8573167	11426455
M2D02	661573.5	5006944	4366302
M2D03	0	9345719.5	5477981
M2D04	0	7415890	5413597.5
M2D05	0	7699239.5	1304164.5
M2D06	0	6056627.5	7103185
M2D07	0	7821794	3297323.5
M2E01	0	10753316.5	15986090.5
M2E02	0	7790116.5	638027
M2E03	0	1222151	1458384.5
M2E04	0	3494099	1664444.5
M2E05	0	1848251	3305300.5
M2E06	0	3788308	1376321
M2E07	0	4363542.5	5185955.5
M2F01	0	5941888	6037385
M2F02	0	7517024.5	2839628.5
M2F03	0	1989179.5	2261397
M2F04	0	1687445.5	1049660
M2F05	0	2996834	7889182.5
M2F06	0	7342801	3506965
M2F07	0	5429998	9627084.5
M2G01	2173976.5	12793875.5	5013111
M2G02	0	7406805.5	11082598
M2G03	0	4032941.5	8680366.5
M2G04	0	3757614.5	6054435.5
M2G05	890033	7452004.5	2702180.5
M2G06	0	6836695.5	1643623.5
M2G07	0	7914620	4990894.5
M2H01	0	5176620.5	6582368
M2H02	425728	8726092.5	2249946.5
M2H03	0	11899204.5	17271731.5
M2H04	0	11155795.5	389619.5
M2H05	622909.5	15014182.5	13643206
M2H06	0	4134373.5	7727648.5

Table S4.10. Raw AUC for DEBS D210 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position D210 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	5.75	4.83	5.17
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	6420470	18889927.5	31257912
M1A02	364819.5	19965894	60423194
M1A03	0	15868396.5	50320130
M1A04	0	15121345	53127853
M1A05	0	25635616	43718475
M1A06	0	22110347.5	45617291
M1A07	0	23689870.5	23292220
M1A08	0	23457855.5	44721923
M1A09	0	27225125	29452917
M1A10	953946	26066406	39339559.5
M1A11	140586.5	25314133	41533083.5
M1A12	0	23006889	43748048.5
M1B01	0	14775040	12618038.5
M1B02	0	20821637	19903358
M1B03	0	17892393.5	25242816
M1B04	0	21498717.5	17334632.5
M1B05	0	22590231	15352393
M1B06	0	18003296	28323817
M1B07	0	17861322	9733683.5
M1B08	0	20062495.5	14206432
M1B09	0	25605254	1656074.5
M1B10	0	20550846.5	18043598
M1B11	0	16204458	25878901.5
M1B12	0	15272692.5	22244931
M1C01	0	16457794	6940729
M1C02	0	17151377	21672657.5
M1C03	0	20743865.5	16442513

M1C04	0	21672405.5	13390781
M1C05	0	22489625	11790378
M1C06	0	21500199	17655758.5
M1C07	278695.5	19489696.5	4032369
M1C08	0	19673537	17441513
M1C09	0	25798671.5	3900212.5
M1C10	0	20101592.5	19875854.5
M1C11	0	24281428	5822535
M1C12	184562	23288384	3614716
M1D01	0	15172857	9348219
M1D02	0	17462698.5	15372324.5
M1D03	0	21485076.5	11157747
M1D04	0	19873554	12855669
M1D05	0	24300676.5	9677616.5
M1D06	0	20315819.5	13490787
M1D07	0	21206660.5	8646977.5
M1D08	0	19730809.5	16619749.5
M1D09	0	24342888.5	5021790
M1D10	0	20728819.5	13300740.5
M1D11	0	25052173	1920219.5
M1D12	0	21015303	11321959.5
M1E01	0	13425039	11902248
M1E02	0	16467794	16584754
M1E03	0	17283018	14403686
M1E04	0	21096516.5	8474843
M1E05	0	20958742.5	15202219
M1E06	0	21501693	11325229.5
M1E07	638649	17856030	0
M1E08	0	19903520	13280066
M1E09	0	21342581.5	14675069
M1E10	0	22435693	10117908.5
M1E11	0	21245410	10187788.5
M1E12	0	18918195.5	14593669
M1F01	0	15134814	6643804.5
M1F02	138998	18857976.5	9993107.5
M1F03	0	18283085	14423806
M1F04	0	22060802.5	8684158.5
M1F05	0	20700534	11035166.5
M1F06	0	22149965.5	8254920
M1F07	0	20286633	6077890
M1F08	0	19933407	8355261
M1F09	0	18796729	15453413

M1F10	0	19199653.5	16325400.5
M1F11	0	21269797	7328254
M1F12	0	23175528	3900821
M1G01	0	17515811	2926844
M1G02	0	19298828.5	8803120
M1G03	0	19255986	9373149
M1G04	0	21430705	7768170.5
M1G05	0	20651690.5	7700382
M1G06	141753	20995091.5	6347772.5
M1G07	346527	19996715	4337214
M1G08	0	18929122	15820646.5
M1G09	0	20772474.5	8214315
M1G10	0	19989145	10013489
M1G11	0	21630429.5	6415782
M1G12	0	25524695.5	3205877
M1H01	0	16107209.5	7428123.5
M1H02	0	19975008.5	7180195.5
M1H03	0	22085497	5358620.5
M1H04	0	21861676	6851136.5
M1H05	0	23613122	2451824
M1H06	0	21150381	4160566.5
M1H07	0	17046796	1124635
M1H08	0	21806206.5	9487784
M1H09	447602	19171433.5	1440188
M1H10	0	17033700	3734792
M1H11	0	17330845.5	327971
M1H12	0	14316553	285225.5
M2A01	0	484504.5	47157658
M2A02	0	15086933.5	24031635.5
M2A03	0	15417738.5	31173160
M2A04	0	23729765	16968190
M2A05	0	16972847.5	39186666.5
M2A06	0	16375050.5	28328208
M2A07	0	15707417.5	30258258
M2B01	0	21112271.5	18073766.5
M2B02	201501	16776215.5	23863422
M2B03	0	20715364.5	18260396
M2B04	0	17697002	22507876.5
M2B05	0	19808705	19504447
M2B06	0	17077628	24273951
M2B07	0	13794642	30738533.5
M2C01	0	14145013	30522381

M2C02	0	19727195	19757649.5
M2C03	0	20033717.5	20373436
M2C04	0	17434427.5	24586360.5
M2C05	0	19968185	17904347
M2C06	0	11518309.5	34225232
M2C07	0	13621465.5	33020095
M2D01	0	19509394	20048604.5
M2D02	0	22253607.5	13186587.5
M2D03	0	12691409.5	31508514
M2D04	0	12690290	30171685
M2D05	0	23677791.5	5429814.5
M2D06	0	14227034.5	28345675.5
M2D07	0	10720438	36274636
M2E01	0	18920079.5	22397283
M2E02	0	15963615.5	24998627
M2E03	0	10029882	34926175
M2E04	0	24406831.5	6246979
M2E05	0	24442649	8724820.5
M2E06	0	14563992.5	29616162.5
M2E07	0	9454690.5	36744144.5
M2F01	0	5318602	42362789
M2F02	0	15562244	26267688
M2F03	0	16057787.5	26077236
M2F04	0	13025048.5	31218305.5
M2F05	1025879	22476115.5	3685671
M2F06	276003.5	25778446	6762761.5
M2F07	0	7373435	42400467
M2G01	0	16945576.5	27563113
M2G02	0	11223909	34423069
M2G03	0	15303962	30211588.5
M2G04	0	18020428	25343859
M2G05	0	18289370	24640514
M2G06	0	15794696	28446342.5
M2G07	0	13674774.5	34117050
M2H01	0	13729356	32760030
M2H02	0	17408752.5	23414170
M2H03	0	20458396.5	20383685
M2H04	0	14304628	30527634.5
M2H05	0	16629558	29059501
M2H06	0	23261902	15727296

Table S4.11. Raw AUC for DEBS T213 mutants. A library of 147 clones was generated by saturation mutagenesis using the NNS degenerate codon at position T213 of the DEBS-TE. Assays were carried out using *E. coli* lysate and screened by LCMS with the expressed protein products from pJBH09, pET-21c(+) and pRSG33 acting as negative (M1A04), background (M1A03) and positive controls (M1A02) respectively. M1A01 contains recombinantly expressed and purified WT DEBS-TE (5 μ M) incubated under the same conditions (not shown in scatter plots). Raw AUC of macrocycle (**4.3**), hydrolysis (**4.2**) products and substrate (**4.1**) for each mutant (rows) was extracted using MZmine2.

Row <i>m/z</i>	348	366	467
Row retention time (mins)	3.62	2.94	3.20
Row identity	Macrocycle, 4.3	Hydrolysis, 4.2	Substrate, 4.1
M1A01	1612008.5	4812580.5	2161177
M1A02	4451061	16799691	4617255.5
M1A03	0	12662501	25144738
M1A04	0	15683753	31100693.5
M1A05	0	15327236.5	40323482.5
M1A06	0	33219417.5	24670434.5
M1A07	0	31918354	13636625.5
M1A08	975409.5	27429769	32213537.5
M1A09	0	29350711	35175533
M1A10	0	25967523	35308362.5
M1A11	2713944.5	27517897.5	2484241
M1A12	1876156.5	34484023	18186206.5
M1B01	2059893.5	15087773.5	0
M1B02	2067141	10250027.5	0
M1B03	1625979	15431553.5	4138383
M1B04	0	22182414	7285001
M1B05	678559.5	23777467	8799749.5
M1B06	1280255.5	24220867.5	8389462
M1B07	1932123	36076416	28911704
M1B08	0	16464053.5	22050572.5
M1B09	0	13614001	14724050.5
M1B10	0	15138730.5	24339092
M1B11	2577075.5	27270124	3256959.5
M1B12	1028922.5	15256611	17576522.5
M1C01	377186.5	13770280	0
M1C02	2043645	19720389	699415.5
M1C03	574615	8846166.5	941202.5

M1C04	518336	17286539	0
M1C05	2195176.5	9430548.5	0
M1C06	847356.5	12654073	3022199
M1C07	0	7341413	1280628.5
M1C08	521541	7956670.5	922230.5
M1C09	0	9173103	10059753.5
M1C10	603513	10678078.5	3909214
M1C11	1285127	16919355	9897209
M1C12	1069884.5	28254470	582603
M1D01	0	5508343.5	5342953.5
M1D02	0	5381433.5	7907293.5
M1D03	0	11052382.5	0
M1D04	0	9848556.5	4025559
M1D05	0	22214526	12794566.5
M1D06	1537413.5	11764177.5	635698.5
M1D07	951268.5	22651141.5	3526904.5
M1D08	634938	8834074.5	0
M1D09	0	14863928	24201412.5
M1D10	0	17742652	21700040.5
M1D11	0	14197240.5	17286002.5
M1D12	679907	18744170.5	14433972
M1E01	1894336.5	17504560	0
M1E02	0	5770055.5	7722274
M1E03	0	13931943.5	11700892.5
M1E04	899316.5	8482325.5	527902.5
M1E05	1459691.5	16337619	0
M1E06	0	11600973.5	0
M1E07	0	9962668	490077
M1E08	941262	20209641	2288080
M1E09	0	8834780.5	10348843.5
M1E10	0	4884409.5	7492163
M1E11	771480	13657746.5	13272372.5
M1E12	694376.5	32555771.5	666618.5
M1F01	0	9227947	1296108
M1F02	0	11172206.5	0
M1F03	1935208.5	8471381.5	605047
M1F04	0	4153716	7454698.5
M1F05	1840438	8062631.5	0
M1F06	0	8817590.5	574204.5
M1F07	0	8461917.5	6583411.5
M1F08	0	4972656.5	9599574
M1F09	0	13271558	13982303

M1F10	0	12973190	0
M1F11	522069.5	18803218.5	0
M1F12	0	14328004	21295078
M1G01	529988	17013495.5	0
M1G02	0	8846377.5	832242
M1G03	458470	18212603	0
M1G04	0	10002007.5	880549.5
M1G05	853681.5	22980894	0
M1G06	0	9753926	5470500.5
M1G07	783704.5	5890980	5126484
M1G08	2658605.5	15388811.5	1514061.5
M1G09	515071	4342632.5	1090032.5
M1G10	0	900967.5	2197495
M1G11	526699.5	18768347.5	0
M1G12	597152	21828704.5	0
M1H01	0	17236946	4918452
M1H02	0	16298745	6346572.5
M1H03	1896150.5	17361506.5	4730008.5
M1H04	2547968.5	17028235	0
M1H05	1233883.5	11209741.5	14473833.5
M1H06	2530890	20166137.5	445683.5
M1H07	1999754.5	17718188.5	0
M1H08	4421321.5	19894620.5	5492129
M1H09	0	13786049	20479933.5
M1H10	3344357	21377542	4640640
M1H11	687878.5	23595449	0
M1H12	0	12542652.5	29596447.5
M2A01	0	2189798.5	1354860.5
M2A02	537267.5	2493633.5	1536310
M2A03	0	14749926	25940392.5
M2A04	0	943928.5	1349312
M2A05	0	1040822	0
M2A06	0	1212426.5	4437625.5
M2A07	0	953162	2903592
M2B01	2271229	7981686.5	575656.5
M2B02	0	3893510	416239.5
M2B03	0	8113340	4928716
M2B04	736388	15345477.5	6535535.5
M2B05	610662	6475438	2631828
M2B06	3188178.5	10662696.5	4831291
M2B07	0	4602933.5	11423260.5
M2C01	0	9647639.5	20733201.5

M2C02	0	1988287.5	2809658.5
M2C03	0	6176065.5	5612894.5
M2C04	0	6174968.5	0
M2C05	0	1585379.5	1446055.5
M2C06	582391	5444623.5	2096305.5
M2C07	0	5752981	4468573.5
M2D01	0	2467200	0
M2D02	1008755	5081001.5	3945292
M2D03	0	5058412.5	9333151
M2D04	1110898	7264272	3642889
M2D05	0	4073497.5	9271936.5
M2D06	0	2369200	2768248.5
M2D07	0	3479787	2092206
M2E01	0	2314289	4108067.5
M2E02	1506557	8247698.5	2626152.5
M2E03	0	4172696.5	4957874
M2E04	0	5007579	8021937
M2E05	0	4231152	7092620
M2E06	0	2281498	3547690
M2E07	0	5931636.5	6653338
M2F01	0	989056	3632920
M2F02	1415685.5	8404609.5	4783096
M2F03	0	7441139	1505746.5
M2F04	0	5818414	5412806
M2F05	0	7753644	3922545
M2F06	0	4648819	7889774
M2F07	0	3392351	8816154
M2G01	0	3278019.5	0
M2G02	1484930.5	10094487.5	11839032.5
M2G03	0	9478738.5	0
M2G04	524048.5	7144066.5	875808
M2G05	0	3855552.5	5517504.5
M2G06	0	4959965	1524901
M2G07	0	2163399	4847736.5
M2H01	0	2517082.5	2596252
M2H02	0	4802992	3148176
M2H03	0	762460.5	1242792
M2H04	0	3169286.5	0
M2H05	0	4392494	10411569
M2H06	0	5314006	7989912

Biochemical assays raw AUC and product ratios data

Table S4.12. Raw AUC and product ratios of biochemical assays. A total of 29 DEBS-TE mutants were recombinantly expressed and purified for biochemical assays. Each mutant was incubated with 1 mM of substrate **4.1** in potassium phosphate buffer (50 mM, pH 7.4) for 22 hours (triplicate dataset). WT DEBS-TE, substrate **4.1** only and heat denatured enzyme assays were also prepared similarly to act as positive, background and negative controls respectively. Assays were analyzed by LCMS and AUC calculated from ion extracted chromatograms for $[M + H]^+$ ions with $m/z = 348$ (**4.3**), 366 (**4.2**) and 467 (**4.1**). Macrocycle to hydrolysis ratios were calculated by dividing the AUC of **4.3** to **4.2** then normalized to the WT DEBS-TE ratio. Mutants highlighted in red showed little activity compared to their respective dead enzyme controls.

Sample	AUC														
	Hydrolysis, 4.2			Macrocycle, 4.3			Substrate, 4.1			Macrocycle/Hydrolysis ratio			Normalized Macrocycle/Hydrolysis ratio		
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
4.1 control	202010	164679	101953	0	0	0	20510894	16127260	11434044	0.00	0.00	0.00	0.00	0.00	0.00
WT	6444885	5434030	2233914	1908527	1659204	714513	13432447	11367805	5297629	0.30	0.31	0.32	1.00	1.00	1.00
dead WT	192134	228734	117423	0	0	0	19462272	21756153	10886032	0.00	0.00	0.00	0.00	0.00	0.00
s24	3620085	3609582	3621538	1959517	1883913	1965540	7753166	7556198	7729499	0.54	0.52	0.54	1.83	1.71	1.70
dead s24	177623	116522	177623	0	0	0	16472422	16519980	16472422	0.00	0.00	0.00	0.00	0.00	0.00
s25	774538	794575	722273	154394	140725	153388	6888447	6517493	6122442	0.20	0.18	0.21	0.67	0.58	0.66
dead s25	73208	101814	94349	0	0	0	8096477	7986447	8474191	0.00	0.00	0.00	0.00	0.00	0.00
s26	2838035	3142147	3345226	152242	217372	259618	12079401	13167620	13386917	0.05	0.07	0.08	0.18	0.23	0.24
dead s26	246980	236868	226779	0	0	0	17122088	17482411	16699316	0.00	0.00	0.00	0.00	0.00	0.00
s27	896143	855210	861345	0	0	0	6787795	7117356	6967317	0.00	0.00	0.00	0.00	0.00	0.00
dead s27	76511	85372	78468	0	0	0	8650423	8401650	7971529	0.00	0.00	0.00	0.00	0.00	0.00
s28	761101	723992	708314	0	0	0	16580594	16479031	16185370	0.00	0.00	0.00	0.00	0.00	0.00
dead s28	297329	287780	288925	0	0	0	18850522	19399663	18349794	0.00	0.00	0.00	0.00	0.00	0.00
s29	2410758	2587922	2671309	169281	175081	173874	14828540	15123269	15179313	0.07	0.07	0.07	0.24	0.22	0.20
dead s29	550030	557413	561004	0	0	0	16633079	17133701	17145834	0.00	0.00	0.00	0.00	0.00	0.00
s30	6454400	7469261	3798709	474508	532991	267823	8518069	9707576	6001493	0.07	0.07	0.07	0.25	0.23	0.22
dead s30	220444	206946	252221	0	0	0	20320930	19552611	11373354	0.00	0.00	0.00	0.00	0.00	0.00
s31	4998660	4510355	4592355	2880182	2833663	2786949	7224575	6819895	6773142	0.58	0.63	0.61	1.95	2.06	1.90
dead s31	194366	212056	211638	0	0	0	16922905	17367961	17356910	0.00	0.00	0.00	0.00	0.00	0.00
s32	1451067	2123851	2820777	376669	550922	718989	911535	1321626	1770791	0.26	0.26	0.25	0.88	0.85	0.80
dead s32	97599	157608	160935	0	0	0	11170881	17402047	17439706	0.00	0.00	0.00	0.00	0.00	0.00
s33	5772724	5653562	6037694	0	0	0	7750189	7570876	7674811	0.00	0.00	0.00	0.00	0.00	0.00

dead s33	150549	137533	157830	0	0	0	16933074	16081551	17069412	0.00	0.00	0.00	0.00	0.00	0.00
s34	4427531	4609066	4582005	3078600	3103963	3228069	6354303	6280478	6886187	0.70	0.67	0.70	2.35	2.21	2.20
dead s34	412505	449602	417275	0	0	0	17126536	16756552	16393458	0.00	0.00	0.00	0.00	0.00	0.00
s35	5180874	5363407	5797120	0	0	0	7550842	7730365	8582568	0.00	0.00	0.00	0.00	0.00	0.00
dead s35	305562	305115	245416	0	0	0	18478935	17866650	18153841	0.00	0.00	0.00	0.00	0.00	0.00
s36	6762696	6770427	7131288	412853	435674	423460	4959790	5041269	5318647	0.06	0.06	0.06	0.21	0.21	0.19
dead s36	328573	316798	318745	0	0	0	17306130	18002687	17815584	0.00	0.00	0.00	0.00	0.00	0.00
s37	4367845	4611720	4966286	0	0	0	9951879	10634045	11377572	0.00	0.00	0.00	0.00	0.00	0.00
dead s37	283943	285357	299985	0	0	0	16877573	177296198	17544564	0.00	0.00	0.00	0.00	0.00	0.00
s38	1649059	1668687	1574055	0	0	0	11175687	11452528	10384573	0.00	0.00	0.00	0.00	0.00	0.00
dead s38	139351	136446	124598	0	0	0	14060917	14500520	14139975	0.00	0.00	0.00	0.00	0.00	0.00
s39	3105157	3025829	2838214	177288	218521	188501	9984287	8981430	9345425	0.06	0.07	0.07	0.19	0.24	0.21
dead s39	142451	95137	143436	0	0	0	14412247	14819416	14989217	0.00	0.00	0.00	0.00	0.00	0.00
s40	2695502	2387466	1992419	754454	689337	539146	12264154	11421919	8092469	0.28	0.29	0.27	0.95	0.95	0.85
dead s40	133429	145782	107559	0	0	0	15657857	17299634	11732152	0.00	0.00	0.00	0.00	0.00	0.00
s41	982588	929567	757772	0	13141	0	16327863	15210437	10210850	0.00	0.01	0.00	0.00	0.05	0.00
dead s41	137513	154719	114276	0	0	0	16998178	17804644	11114068	0.00	0.00	0.00	0.00	0.00	0.00
c42	2972689	3094618	2997202	1414087	1390743	1408557	889280	1399281	1008056	0.48	0.45	0.47	1.61	1.47	1.47
dead c42	152163	243173	189027	0	0	0	7470896	7533351	6738234	0.00	0.00	0.00	0.00	0.00	0.00
c43	3816343	3810956	4246217	764596	690443	760470	2509103	2416663	2605688	0.20	0.18	0.18	0.68	0.59	0.56
dead c43	518513	529982	538464	0	0	0	9851115	9619332	9751386	0.00	0.00	0.00	0.00	0.00	0.00
c44	5983166	6299223	6498259	973201	995315	1041876	259543	262570	266267	0.16	0.16	0.16	0.55	0.52	0.50
dead c44	1843625	1817539	1865865	0	0	0	10845519	10384537	10839687	0.00	0.00	0.00	0.00	0.00	0.00
c45	5821501	5896015	6194566	302785	321119	304367	1760479	1901456	2011657	0.05	0.05	0.05	0.18	0.18	0.15
dead c45	902512	931161	955611	0	0	0	12672206	12418986	12996987	0.00	0.00	0.00	0.00	0.00	0.00
c46	652686	614804	784226	496720	480565	519317	6004403	5735082	7210434	0.76	0.78	0.66	2.57	2.56	2.07
dead c46	143746	147733	76476	0	0	0	7950181	643235	5472606	0.00	0.00	0.00	0.00	0.00	0.00
c47	1365078	1408306	1200618	477931	488287	459888	5151281	5038247	4740425	0.35	0.35	0.38	1.18	1.14	1.20
dead c47	134492	111745	97946	0	0	0	8526390	8278331	7985843	0.00	0.00	0.00	0.00	0.00	0.00
c48	1763815	1859540	1792667	784538	809155	765775	5581926	5410204	5673717	0.44	0.44	0.43	1.50	1.43	1.34
dead c48	419365	394679	405901	0	0	0	10664106	10111819	10001192	0.00	0.00	0.00	0.00	0.00	0.00
c50	457474	474650	462155	0	0	0	10230161	10124677	10143732	0.00	0.00	0.00	0.00	0.00	0.00
dead c50	194001	138820	131338	0	0	0	11029736	10992442	11123481	0.00	0.00	0.00	0.00	0.00	0.00
c51	3845483	3949824	3906306	712074	729170	698667	7002902	7573704	7522148	0.19	0.18	0.18	0.63	0.60	0.56
dead c51	191595	174969	164009	0	0	0	13850312	13838741	13987553	0.00	0.00	0.00	0.00	0.00	0.00
c52	3140438	3198935	3033100	441227	447078	478497	3444130	3392527	3330752	0.14	0.14	0.16	0.47	0.46	0.49
dead c52	101521	100988	97125	0	0	0	9527859	9230617	9477511	0.00	0.00	0.00	0.00	0.00	0.00

4.7 References

- (1) Lussier, F.; Colatrigano, D.; Wiltshire, Z.; Page, J. E.; Martin, V. J. J. Engineering Microbes for Plant Polyketide Biosynthesis Abstract : Polyketides Are an Important Group of Secondary Metabolites , Many of Which Have Important Industrial Applications in the Food and Pharmaceutical Industries . Polyketides Are Synthesized Fr. *Comput. Struct. Biotechnol. J.* **2012**, 3 (4), e201210020.
- (2) Ridley, C. P.; Lee, H. Y.; Khosla, C. Evolution of Polyketide Synthases in Bacteria. *Proc. Natl. Acad. Sci.* **2008**, 105 (12), 4595–4600. <https://doi.org/10.1073/pnas.0710107105>.
- (3) Minami, A. Predicting the Chemical Space of Fungal Polyketides by Phylogeny - Based Bioinformatics Analysis of Polyketide Synthase - Nonribosomal Peptide Synthetase and Its Modification Enzymes. *Sci. Rep.* **2020**, 10 (Article 13556), 1–13. <https://doi.org/10.1038/s41598-020-70177-w>.
- (4) Bilyk, O.; Samborsky, M.; Leadlay, P. F. The Biosynthetic Pathway to Ossamycin, a Macrocyclic Polyketide Bearing a Spiroacetal Moiety. *PLoS One* **2019**, 14 (4), 1–17. <https://doi.org/10.1371/journal.pone.0215958>.
- (5) Tenson, T.; Lovmar, M.; Ehrenberg, M. The Mechanism of Action of Macrolides, Lincosamides and Streptogramin B Reveals the Nascent Peptide Exit Path in the Ribosome. *J. Mol. Biol.* **2003**, 330 (5), 1005–1014. [https://doi.org/10.1016/S0022-2836\(03\)00662-4](https://doi.org/10.1016/S0022-2836(03)00662-4).
- (6) Goodin, S.; Kane, M. P.; Rubin, E. H. Epothilones: Mechanism of Action and Biologic Activity. *J. Clin. Oncol.* **2004**, 22 (10), 2015–2025. <https://doi.org/10.1200/JCO.2004.12.001>.
- (7) Sigal, M. V.; Wiley, P. F. Erythromycin I. Properties and Degradation. *J. Am. Chem. Soc.* **1954**, 76 (5), 3121–3130.
- (8) Gale, R.; Pettinga, C. W.; Gale, R. Erythromycin. XII. The Isolation, Properties and Partial Structure of Erythromycin C. *J. Am. Chem. Soc.* **1957**, 79 (22), 6074–6077.
- (9) Argyropoulos, P.; Bergeret, F.; Pardin, C.; Reimer, J. M.; Pinto, A.; Boddy, C. N.; Schmeing, T. M. Towards a Characterization of the Structural Determinants of Specificity in the Macrocyclizing Thioesterase for Deoxyerythronolide B Biosynthesis. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, 1860 (3), 486–497. <https://doi.org/10.1016/j.bbagen.2015.11.007>.
- (10) He, W.; Wu, J.; Khosla, C.; Cane, D. E. Macrolactonization to 10-Deoxymethynolide Catalyzed by the Recombinant Thioesterase of the Picromycin/Methymycin Polyketide Synthase. *Bioorganic Med. Chem. Lett.* **2006**, 16 (2), 391–394. <https://doi.org/10.1016/j.bmcl.2005.09.077>.
- (11) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. Manipulation of Macrolide Ring Size by Directed Mutagenesis of a Modular Polyketide Synthase Department of Chemical Engineering. *J. Am. Chem. Soc.* **1995**, 117 (35), 9105–9106.
- (12) Tsai, S. C.; Lu, H.; Cane, D. E.; Khosla, C.; Stroud, R. M. Insights into Channel Architecture and Substrate Specificity from Crystal Structures of Two Macrocyclic-Forming Thioesterases of Modular Polyketide Synthases. *Biochemistry* **2002**, 41 (42), 12598–12606. <https://doi.org/10.1021/bi0260177>.
- (13) Tsai, S. C.; Miercke, L. J. W.; Krucinski, J.; Gokhale, R.; Chen, J. C. H.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. Crystal Structure of the Macrocyclic-Forming Thioesterase Domain of the Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98 (26), 14808–14813. <https://doi.org/10.1073/pnas.011399198>.
- (14) Hari, T. P. A.; Labana, P.; Boileau, M.; Boddy, C. N. An Evolutionary Model Encompassing Substrate Specificity and Reactivity of Type I Polyketide Synthase Thioesterases. *ChemBioChem* **2014**, 15 (18), 2656–2661. <https://doi.org/10.1002/cbic.201402475>.

- (15) Pinto, A.; Wang, M.; Horsman, M.; Boddy, C. N. 6-Deoxyerythronolide B Synthase Thioesterase-Catalyzed Macrocyclization Is Highly Stereoselective. *Org. Lett.* **2012**, *14* (9), 2278–2281. <https://doi.org/10.1021/ol300707j>.
- (16) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. Mechanism and Specificity of the Terminal Thioesterase Domain from the Erythromycin Polyketide Synthase. *Chem. Biol.* **1999**, *6* (2), 117–125. [https://doi.org/10.1016/S1074-5521\(99\)80008-8](https://doi.org/10.1016/S1074-5521(99)80008-8).
- (17) Kao, C. M.; Mcpherson, M.; Mcdaniel, R. N.; Fu, H.; Cane, D. E.; Khosla, C. Gain of Function Mutagenesis of the Erythromycin Polyketide Synthase . 2 . Engineered Biosynthesis of Departments of Chemical Engineering , Chemistry and Thesis of Polyketides , Natural Product Molecules with Remarkable Organization of These Multifunction. *J. Am. Chem. Soc.* **1997**, *119* (46), 11339–11340.
- (18) Jacobsen, J. R.; Hutchinson, C. R.; Cane, D. E.; Khosla, C. Precursor-Directed Biosynthesis of Erythromycin Analogs by an Engineered Polyketide Synthase. *Sci. Reports* **1997**, *277* (5324), 367–370.
- (19) Kao, C. M.; Lun, G.; Katz, L.; Kane, D. E.; Khosla, C. Engineered Biosynthesis of Structurally Diverse Tetraketides by a Trimodular Polyketide Synthase. *J. Am. Chem. Soc.* **1996**, *118* (38), 9184–9185.
- (20) Weber, J. M.; Leung, J. O.; Maine, G. T.; Potenz, R. H. B.; Paulus, T. J.; DeWitt, J. P. Organization of a Cluster of Erythromycin Genes in *Saccharopolyspora Erythraea*. *J. Bacteriol.* **1990**, *172* (5), 2372–2383. <https://doi.org/10.1128/jb.172.5.2372-2383.1990>.
- (21) Reeves, N. R.; English, R. S.; Lampel, J. S.; Post, D. A.; Boom, T. J. V. Transcriptional Organization of the Erythromycin Biosynthetic Gene Cluster of *Saccharopolyspora Erythraea*. *J. Bacteriol.* **1999**, *181* (22), 7098–7106. <https://doi.org/10.1128/jb.181.22.7098-7106.1999>.
- (22) Du, L.; Lou, L. PKS and NRPS Release Mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255–278. <https://doi.org/10.1039/b912037h>.
- (23) Mortison, J. D.; Sherman, D. H. Frontiers and Opportunities in Chemoenzymatic Synthesis. *J. Org. Chem.* **2010**, *75* (21), 7041–7051. <https://doi.org/10.1021/jo101124n>.
- (24) Parenty, A.; Moreau, X.; Niel, G.; Campagne, J. M. Update 1 of: Macrolactonizations in the Total Synthesis of Natural Products. *Chem. Rev.* **2013**, *113* (1), PR1–PR40. <https://doi.org/10.1021/cr300129n>.
- (25) Garnes-Portolés, F.; Leyva-Pérez, A. Macrocyclization Reactions at High Concentration ($\geq 0.2\text{M}$): The Role of Catalysis. *ACS Catal.* **2023**, *13* (14), 9415–9426. <https://doi.org/10.1021/acscatal.3c02006>.
- (26) Woodward, R. B.; Au-Yeung, B. W.; Balaram, P.; Browne, L. J.; Ward, D. E.; Card, P. J.; Chen, C. H.; Chenevert, R. B.; Fliri, A.; Frobel, K.; Gais, H.-J.; Garratt, D. G.; Hayakawa, K.; Heggie, W.; Hesson, D. P.; Hoppe, D.; Hoppe, I.; Hyatt, J. A.; Ikeda, D.; Jacobi, P. A.; Kim, K. S.; Kobuke, Y.; Kojima, K.; Krowicki, K.; Lee, V. J.; Leutert, T.; Malchenko, S.; Martens, J.; Matthews, R. S.; Ong, B. S.; Press, J. B.; Rajan Babu, T. V.; Rousseau, G.; Sauter, H. M.; Suzuki, M.; Tatsua, K.; Tolbert, L. M.; Truesdale, E. A.; Uchida, I.; Ueda, Y.; Uyehara, T.; Vasella, A. T.; Vladuchick, W. C.; Wade, P. A.; Williams, R. M.; Wong, H. N.-C. Asymmetric Total Synthesis of Erythromycin. 2. Synthesis of an Erythronolide A Lactone System. *J. Am. Chem. Soc.* **1981**, *130* (7), 3213–3215.
- (27) Khosla, C.; Gokhale, R. S.; Jacobsen, J. R.; Cane, D. E. Tolerance and Specificity of Polyketide Synthases. *Annu. Rev. Biochem.* **1999**, *68*, 219–253.
- (28) Wang, M.; Boddy, C. N. Examining the Role of Hydrogen Bonding Interactions in the Substrate Specificity for the Loading Step of Polyketide Synthase Thioesterase Domains. *Biochemistry* **2008**, *47* (45), 11793–11803. <https://doi.org/10.1021/bi800963y>.
- (29) Akey, D. L.; Kittendorf, J. D.; Giraldes, J. W.; Fecik, R. A.; Sherman, D. H.; Smith, J. L. Structural

- Basis for Macrolactonization by the Pikromycin Thioesterase. *Nat. Chem. Biol.* **2006**, 2 (10), 537–542. <https://doi.org/10.1038/nchembio824>.
- (30) Giraldes, J. W.; Akey, D. L.; Kittendorf, J. D.; Sherman, D. H.; Smith, J. L.; Fecik, R. A. Structural and Mechanistic Insights into Polyketide Macrolactonization from Polyketide-Based Affinity Labels. *Nat. Chem. Biol.* **2006**, 2 (10), 531–536. <https://doi.org/10.1038/nchembio822>.
 - (31) Horsman, M. E.; Hari, T. P. A. A.; Boddy, C. N. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016**, 33 (2), 183–202. <https://doi.org/10.1039/c4np00148f>.
 - (32) Lu, C.; Shen, F.; Wang, S.; Wang, Y.; Liu, J.; Bai, W. J.; Wang, X. An Engineered Self-Sufficient Biocatalyst Enables Scalable Production of Linear α -Olefins from Carboxylic Acids. *ACS Catal.* **2018**, 8 (7), 5794–5798. <https://doi.org/10.1021/acscatal.8b01313>.
 - (33) Hartley, C. J.; Williams, C. C.; Scoble, J. A.; Churches, Q. I.; North, A.; French, N. G.; Nebl, T.; Coia, G.; Warden, A. C.; Simpson, G.; Frazer, A. R.; Jensen, C. N.; Turner, N. J.; Scott, C. Engineered Enzymes That Retain and Regenerate Their Cofactors Enable Continuous-Flow Biocatalysis. *Nat. Catal.* **2019**, 2 (11), 1006–1015. <https://doi.org/10.1038/s41929-019-0353-0>.
 - (34) Fisher, C. L.; Pei, G. K. Modification of a PCR-Based Site-Directed Mutagenesis Method. *Biotechniques* **1997**, 23 (4), 570–574.
 - (35) Bosley, A. D.; Ostermeier, M. Mathematical Expressions Useful in the Construction, Description and Evaluation of Protein Libraries. *Biomol. Eng.* **2005**, 22 (1–3), 57–61. <https://doi.org/10.1016/j.bioeng.2004.11.002>.
 - (36) Clarke, L.; Carbon, J. Selection of Specific Clones from Colony Banks by Suppression or Complementation Tests. *Methods Enzymol.* **1979**, 68 (27), 396–408.
 - (37) Xiao, H.; Bao, Z.; Zhao, H. High Throughput Screening and Selection Methods for Directed Enzyme Evolution. *Ind. Eng. Chem. Anal.* **2015**, 54 (16), 4011–4020.
 - (38) Ellefson, J. W.; Meyer, A. J.; Hughes, R. A.; Cannon, J. R.; Brodbelt, J. S.; Ellington, A. D. Directed Evolution of Genetic Parts and Circuits by Compartmentalized Partnered Replication. *Nat. Biotechnol.* **2014**, 32 (1), 97–101. <https://doi.org/10.1038/nbt.2714>.
 - (39) Mcfadden, B.; Heitzman-powell, L. A System for the Continuous Directed Evolution of Biomolecules PACE. *Nature* **2011**, 472 (7344), 499–503. <https://doi.org/10.1038/nature09929.A>.
 - (40) Robinson, A. S.; Hines, V.; Wittrup, K. D. Recombinant Proteins Can Be Isolated from E. Coli Cells by Repeated Cycles of Freezing and Thawing. *Nat. Biotechnol.* **1994**, 12, 1357–1360.
 - (41) Zhang, J. H.; Chung, T. D.; Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **1999**, 4 (2), 67–73.
 - (42) Birmingham, A.; Selfors, L. M.; Forster, T.; Wrobel, D.; Kennedy, C. J.; Shanks, E.; Santoyo-lopez, J.; Dunican, D. J.; Long, A.; Kelleher, D.; Smith, Q.; Beijersbergen, R. L.; Ghazal, P.; Shamu, C. E. Statistical Methods for Analysis of High- Throughput RNA Interference Screens. *Nat. Methods* **2009**, 6 (8), 569–575. <https://doi.org/10.1038/nmeth.1351>.
 - (43) Santos-aberturas, J.; Do, M.; Waldo, G. S.; Uwe, T.; Waldo, G. S.; Bornscheuer, U. T. In-Depth High-Throughput Screening of Protein Engineering Libraries by Split-GFP Direct Crude Cell Extract Data Normalization Resource In-Depth High-Throughput Screening of Protein Engineering Libraries by Split-GFP Direct Crude Cell Extract Data Normaliz. *Chem. Biol.* **2015**, 22 (10), 1406–1414. <https://doi.org/10.1016/j.chembiol.2015.08.014>.
 - (44) Taverna, D. M.; Goldstein, R. A. Why Are Proteins Marginally Stable ? *Proteins Struct. Funct. Genet.* **2002**, 46, 105–109. <https://doi.org/10.1002/prot.10016>.
 - (45) Mansoury, M.; Hamed, M.; Karmustaji, R.; Hannan, F. Al; Safrany, S. T. The Edge Effect : A Global Problem. The Trouble with Culturing Cells in 96-Well Plates. *Biochem. Biophys. Reports* **2021**, 26

- (Article 100987), 1–5. <https://doi.org/10.1016/j.bbrep.2021.100987>.
- (46) Stott, D. I. Immunoblotting and Dot Blotting. *J. Immunol. Methods* **1989**, *119* (2), 153–187.
 - (47) Tian, G.; Tang, F.; Yang, C.; Zhang, W.; Bergquist, J.; Wang, B.; Mi, J.; Zhang, J. Quantitative Dot Blot Analysis (QDB), a Versatile High Throughput Immunoblot Method. *Oncotarget* **2017**, *8* (35), 58553–58562. <https://doi.org/10.18632/oncotarget.17236>.
 - (48) Taylor, S. C.; Berkelman, T.; Yadav, G.; Hammond, M. A Defined Methodology for Reliable Quantification of Western Blot Data. *Mol. Biotechnol.* **2013**, *55* (3), 217–226. <https://doi.org/10.1007/s12033-013-9672-6>.
 - (49) Schaefer, J.; Jovanovic, G.; Kotta-Loizou, I.; Buck, M. Single-Step Method for β -Galactosidase Assays in Escherichia Coli Using a 96-Well Microplate Reader. *Anal. Biochem.* **2016**, *503*, 56–57. <https://doi.org/10.1016/j.ab.2016.03.017>.
 - (50) Brazeau-Henrie, J. T.; Paquette, A. R.; Boddy, C. N. In Vitro Biochemical Characterization of Excised Macrocyclizing Thioesterase Domains from Non-Ribosomal Peptide Synthetases. In *Non-Ribosomal Peptide Biosynthesis and Engineering*; 2023; Vol. 2670, pp 101–125.
 - (51) Koch, A. A.; Hansen, D. A.; Shende, V. V.; Furan, L. R.; Houk, K. N.; Jiménez-Osés, G.; Sherman, D. H. A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst. *J. Am. Chem. Soc.* **2017**, *139* (38), 13456–13465. <https://doi.org/10.1021/jacs.7b06436>.
 - (52) Johnson, C. M. Differential Scanning Calorimetry as a Tool for Protein Folding and Stability. *Arch. Biochem. Biophys.* **2013**, *531* (1–2), 100–109. <https://doi.org/10.1016/j.abb.2012.09.008>.
 - (53) Baranauskienė, L.; Kuo, T. C.; Chen, W. Y.; Matulis, D. Isothermal Titration Calorimetry for Characterization of Recombinant Proteins. *Curr. Opin. Biotechnol.* **2019**, *55*, 9–15. <https://doi.org/10.1016/j.copbio.2018.06.003>.
 - (54) Norma J. Greenfield. Using Circular Dichroism Collected as a Function of Temperature to Determine the Thermodynamics of Protein Unfolding and Binding Interactions. *Nat. Protoc.* **2009**, *1* (6), 2527–2535. <https://doi.org/10.1038/nprot.2006.204>.Using.
 - (55) Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dornan, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W. Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature* **2019**, *565* (7737), 112–117. <https://doi.org/10.1038/s41586-018-0781-z>.
 - (56) Zhu, C.; Gao, Y.; Li, H.; Meng, S.; Li, L.; Francisco, J. S.; Zeng, X. C. Characterizing Hydrophobicity of Amino Acid Side Chains in a Protein Environment via Measuring Contact Angle of a Water Nanodroplet on Planarpeptide Network. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (46), 12946–12951. <https://doi.org/10.1073/pnas.1616138113>.
 - (57) Zhou, H.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2019**, *118* (4), 1691–1741. <https://doi.org/10.1021/acs.chemrev.7b00305>.Electrostatic.
 - (58) Cao, Z.; Bowie, J. U. Shifting Hydrogen Bonds May Produce Flexible Transmembrane Helices. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (21), 8121–8126. <https://doi.org/10.1073/pnas.1201298109>.
 - (59) Barlow, D. J.; Thornton, J. M. Helix Geometry in Proteins. *J. Mol. Biol.* **1988**, *201* (3), 601–619. [https://doi.org/10.1016/0022-2836\(88\)90641-9](https://doi.org/10.1016/0022-2836(88)90641-9).
 - (60) Langelaan, D. N.; Wieczorek, M.; Blouin, C.; Rainey, J. K. Improved Helix and Kink Characterization in Membrane Proteins Allows Evaluation of Kink Sequence Predictors. *J. Chem. Inf. Model.* **2010**, *50*, 2213–2220.
 - (61) Hall, S. E.; Roberts, K.; Vaidehi, N. Position of Helical Kinks in Membrane Protein Crystal

- Structures and the Accuracy of Computational Prediction. *J. Mol. Graph. Model.* **2009**, *27* (8), 944–950. <https://doi.org/10.1016/j.jmglm.2009.02.004>.
- (62) Tyukhtenko, S.; Rajarshi, G.; Karageorgos, I.; Zvonok, N.; Gallagher, E. S.; Huang, H.; Vemuri, K.; Hudgens, J. W.; Ma, X.; Nasr, M. L.; Pavlopoulos, S.; Makriyannis, A. Effects of Distal Mutations on the Structure, Dynamics and Catalysis of Human Monoacylglycerol Lipase. *Sci. Rep.* **2018**, *8* (1), 1–17. <https://doi.org/10.1038/s41598-017-19135-7>.
- (63) Brodtkin, H. R.; Delateur, N. A.; Somarowthu, S.; Mills, C. L.; Novak, W. R.; Beuning, P. J.; Ringe, D.; Ondrechen, M. J. Prediction of Distal Residue Participation in Enzyme Catalysis. *Protein Sci.* **2015**, *24* (5), 762–778. <https://doi.org/10.1002/pro.2648>.
- (64) Pluskal, T.; Castillo, S.; Villar-briones, A.; Ore, M. MZmine 2 : Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry- Based Molecular Profile Data. *BMC Bioinformatics* **2010**, *11*, 395.
- (65) Myers, O. D.; Sumner, S. J.; Li, S.; Barnes, S.; Du, X. One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks. *Anal. Chem.* **2017**, *89* (17), 8696–8703. <https://doi.org/10.1021/acs.analchem.7b00947>.

CHAPTER 5 : Conclusions and future directions

5.1 Preface

As demonstrated in the literature review of **Chapter 1**, macrocycles are challenging to access,^{1,2} spurring the development of many synthetic macrocyclization methodologies.^{2,3} Researchers have started exploring the use of TEs as macrocyclization biocatalysts, showing they can be quite useful.^{4–20} However, other competing reactivities catalyzed by the TE can impact macrocyclization yields and currently there is no effective method to predict this reactivity.^{21–23} Two models were proposed to offer a rational framework for understanding TE-mediated macrocyclization: the ICS and CS models.

While TEs offer significant advantages for macrocycle synthesis, there are a few limitations such as competing hydrolysis reactivity and their inherent selectivity that limit our access to certain macrocycles. Characterizing new TEs can expand our access to macrocycles and has thus far been successful in providing an extensive repertoire of useful TE biocatalysts. The issue pertaining to competing hydrolysis off-loading would require protein engineering to optimize reactivity; however, this remains largely unaddressed. The work in this thesis aims to expand our current repertoire of TE biocatalysts, showcase their use in macrocyclization and address the lack of protein engineering efforts on TEs. **Chapter 2** characterizes two new macrocyclizing NRPS TEs, showcasing their application in the chemoenzymatic synthesis of two macrocyclic depsipeptides, thereby expanding our current repertoire of biocatalysts. **Chapter 3** delves into the characterization of a tandem NRPS TE domain (TE1-TE2), revealing TE1 was largely responsible for its activity as well as showing its surprising substrate tolerance for macrocyclization.

Finally, **Chapter 4** provides the first extensive effort in engineering a PKS TE to optimize its macrocyclization activity towards a non-native substrate.

5.2 Perspective on ICS and CS models

In **Chapter 1**, two archetypal models were proposed to act as a rational framework for TE-mediated macrocyclization. In the ICS model common to PKS TEs, the TE active site is acylated with a conformationally disordered substrate, from which the acyl-TE intermediate samples conformations until one conducive to macrocycle formation is achieved. Because this sampling occurs in the TE active site, structural changes in the protein, like the lid,¹² can contribute to biasing the conformational landscape towards a product-like conformation. However, water can diffuse in the active site during sampling to generate hydrolysis product. In the CS model, more common to NRPS TEs, the substrate pre-organizes into a product-like conformation, which is then selected by the TE for loading and macrocyclization.^{13,24} Minimal structural rearrangements are required for macrocyclization, making it macrocyclize faster and thus limiting water's ability to hydrolyze these intermediates.

The macrocycle to hydrolysis (M/H) product ratios serve as a key indicator to determine the adherence of one or both models. For instance, the substantial hydrolysis reactivity observed in the DEBS and PIK PKS TEs^{10,20,25–27} (resulting in low M/H ratios) is indicative of the ICS model, which predicts substantial competing hydrolysis reactivity. Conversely, a higher M/H ratio points towards the CS model. While M/H ratio is a great indicator, other metrics are also useful for applying these models. Instances of macrocycle product inhibition and TE-mediate macrocycle hydrolysis (ex: Epo and Srf TEs)^{9,15} suggest the application of the CS model. Such cases likely involve similar binding interactions of the TE with both product and substrate, indicating the substrate adopts a product-like conformation as stipulated in our CS model. Moreover, spontaneous macrocyclization of a substrate hint at the application of our CS model. Such a process would only be feasible if the entropic cost associated with end-to-end contact was met

through substrate pre-organization, as demonstrated with patellamide A and tyrocidine (Tyc).^{28–30} In contrast, clear conformational shifts in the TE that sterically control the acyl/peptidyl-TE intermediate conformation (ex: Vlm-TE)¹² point towards the ICS model as this is indicative of a dynamic sampling of the acyl/peptidyl-TE intermediate. In addition, the ability for an ICS TE to bind a conformationally disordered substrate can potentially affect the rate of binding, resulting in an increased K_M . Thus, TE's exhibiting larger K_M are likely to adhere to our ICS model and is consistent with the large K_M values obtained for the DEBS and PIK TEs.^{10,20,25–27}

Applying these models to TE-mediated macrocyclization provides a rational framework for understanding the observed TE reactivities reported in the literature. As more TEs are characterized, further support for these models is anticipated. We utilize the indicators discussed earlier (summarized in **Table 5.1**) to apply these models and rationalize the activities of the TEs characterized in this thesis, which are discussed further in this chapter.

Table 5.1. Summary of key indicators for the ICS and CS models with examples from the literature. Shown is a summary of the key indicators used to apply the models to TE examples from the literature. Abbreviations; DEBS, 6-deoxyerythronolide B; PIK, pikromycin; Vlm, valinomycin; Tyc, tyrocidine; Grs, gramicidin; Sgd, seongsanamide; Srf, surfactin; Epo, epothilone; LybB, lysobactin; Rzm, rhizomide.

Model	Indicator	TE examples	Citations
ICS	High K_M	DEBS, PIK	10,20,25–27
	Low M/H ratio (i.e. substantial hydrolysis)	DEBS, PIK	10,20,25–27
	Steric control of acyl/peptidyl-TE intermediate via TE conformational changes	Vlm	12
CS	Substrate pre-organization	Tyc, Grs, patellamide A	14,28–34
	Macrocycle ring opening	Sgd, Srf, Epo	4,9,15
	Spontaneous macrocyclization	Tyc, patellamide A	28–30
	Macrocycle product inhibition	-	-
	High M/H ratio (i.e. little hydrolysis)	Tyc, Grs, Srf, LybB, Sgd, Rzm	4,5,13–15,19,24,33,35

5.3 Perspective on NRPS TEs

The work presented in **Chapter 2** establishes the ability for both seongsanamide (Sgd) and rhizomide (Rzm) TEs to catalyze formation of 5- and 6-residue depsipeptide macrolactones, respectively, using *N*-acetylcysteamine (SNAC) activated substrate mimics with little competing hydrolysis reactivity (**Figure 5.1**).^{4,5} Compared to other previously characterized NRPS TEs, Sgd is the only known TE to generate a 5-residue macrolactone.⁴ The only other TE known to form a 6-residue macrolactone is the streptogramin forming TE (SnbDE-TE). However, it produces substantial amounts of hydrolysis product while the Rzm-TE does not.^{5,16} These results highlight the benefits of adding new TEs in our repertoire to access macrocycles of the desired ring size. While all attempts to synthetically macrolactonize failed,⁴ the Sgd and Rzm TEs were able to form the desired macrolactones without much issue, highlighting the power of TEs as macrolactonization biocatalysts.

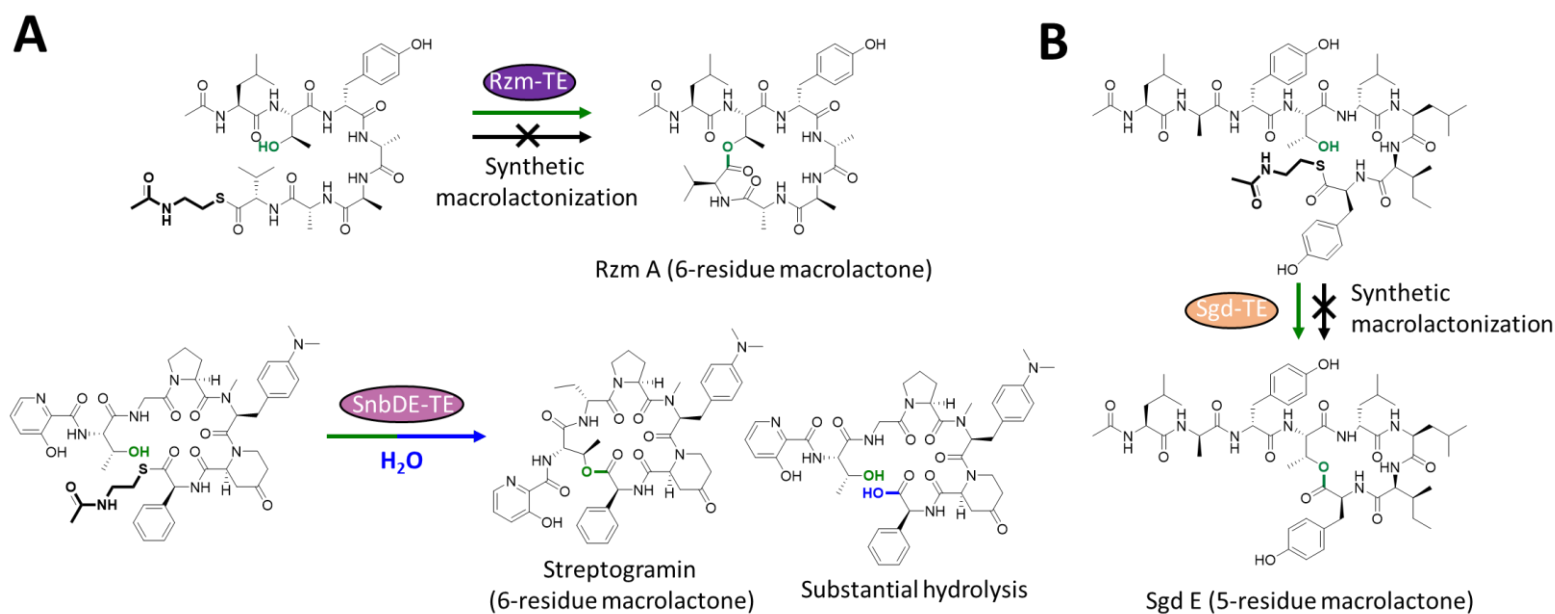


Figure 5.1. General catalytic schemes for Rzm, SnbDE and Sgd TEs. (A) Formation of 6-residue macrolactones done by Rzm-TE and streptogramin forming TE (SnbDE-TE). SnbDE-TE also forms free acid as a substantial side-product. (B) Formation of 5-residue macrolactone done by Sgd-TE. These are the only TEs in the literature shown to form 5- and 6-residue macrolactones. Crosse out black arrows indicate failed synthetic macrolactonization attempts.

The Sgd-TE has shown promise as a macrocyclization biocatalyst, being the only known TE to catalyze formation of a 5-residue macrolactone with little competing hydrolysis reactivity. However, it was also shown to catalyze ring-opening of the macrocycle, a rare activity only seen for the epothilone (Epo) and surfactin (Srf) TEs.^{9,15} Given the lactone ester linkage is sterically hindered and Sgd-TE's ability to hydrolyze the macrocycle, it suggests discrete binding of Sgd E in such a way that allows for nucleophilic attack of the lactone carbonyl for TE acylation. This would also suggest the Sgd-TE binds to and stabilizes a product-like conformation of the substrate (i.e. pre-organized state), aligning with our proposed CS model. According to this model, binding of this pre-organized state would drive macrocyclization and thus limit water's ability to hydrolyze the acyl-TE intermediate. This hypothesis is consistent with the limited production of seco acid observed during the first time points of the Sgd assay. Similarly, in the Rzm assays, very little seco acid was detected during the incubations, aligning well with the expectations set forth in our proposed CS model. However, whether the Sgd and Rzm peptide substrates pre-organize has yet to be seen and warrants further investigation.

Because Sgd- and Rzm-TE likely adhere to our proposed CS model, they might exhibit expanded substrate scopes similar to the Tyc-TE due to substrate pre-organization.^{13,24} Systematically substituting each residue in the macrolactone to Ala (i.e. Ala scan) of the peptide substrates could provide useful insight on residue selectivity for Sgd- and Rzm-TE macrocyclization activity. This experiment would offer useful information for a more guided substrate scope study, testing macrocyclization activity against peptides with changes in internal nucleophile its stereo-chemical configuration and amino acid identity. Additionally, regio-selectivity can be probed by changing the internal nucleophile position to form macrocycles of varying sizes. Utilizing the LCMS-based screen established in **Chapter 4**, substrate scope can be extensively explored by employing a large peptide library (on the order of 10^3 peptides, **Figure 5.2**) in a timely manner.^{36,37} Although gramicidin (Grs) and Srf-TE ring size scope^{14,38} and the Tyc-TE substrate

scope^{13,33,35,39} have been investigated, the literature on NRPS TE substrate scope is still underdeveloped. These two proposed future studies on the Sgd and Rzm TEs hold the potential to provide crucial insights on this underdeveloped topic. Moreover, they would push the boundaries of Sgd- and Rzm-TE as general peptide macrocyclization biocatalysts.

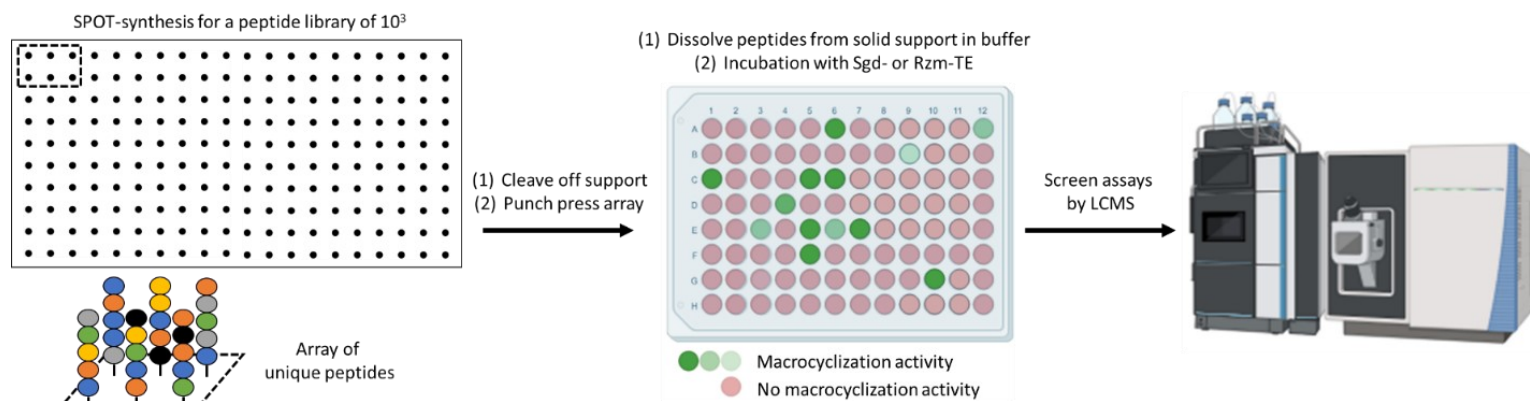


Figure 5.2. Overview of proposed peptide library screen.

The work on Sgd also opens the door to study the bioinformatically predicted P450 (SgdC) involved in the formation of the biaryl ether linkage found in Sgd A-D.⁴⁰ Biaryl ether linkages have been found in the macrocyclic natural products vancomycin, catalyzed by P450 embedded in its biosynthetic gene cluster.⁴¹ Unlike the P450 in vancomycin (ex: OxyB) which requires docking onto X domains for oxidative modification (**Figure 5.3A**), SgdC would act on the released natural product post assembly as the gene cluster does not contain any X domains (**Figure 5.3B**). While the P450 responsible for the biaryl coupling (C-C bond) in arylomycin (AryC) also works post assembly,⁴² SgdC is unique as it catalyzes a biaryl ether linkage (C-O-C). Its study would be a great addition to the current repertoire of P450 catalysts for oxidative modifications of natural products *in vitro*.

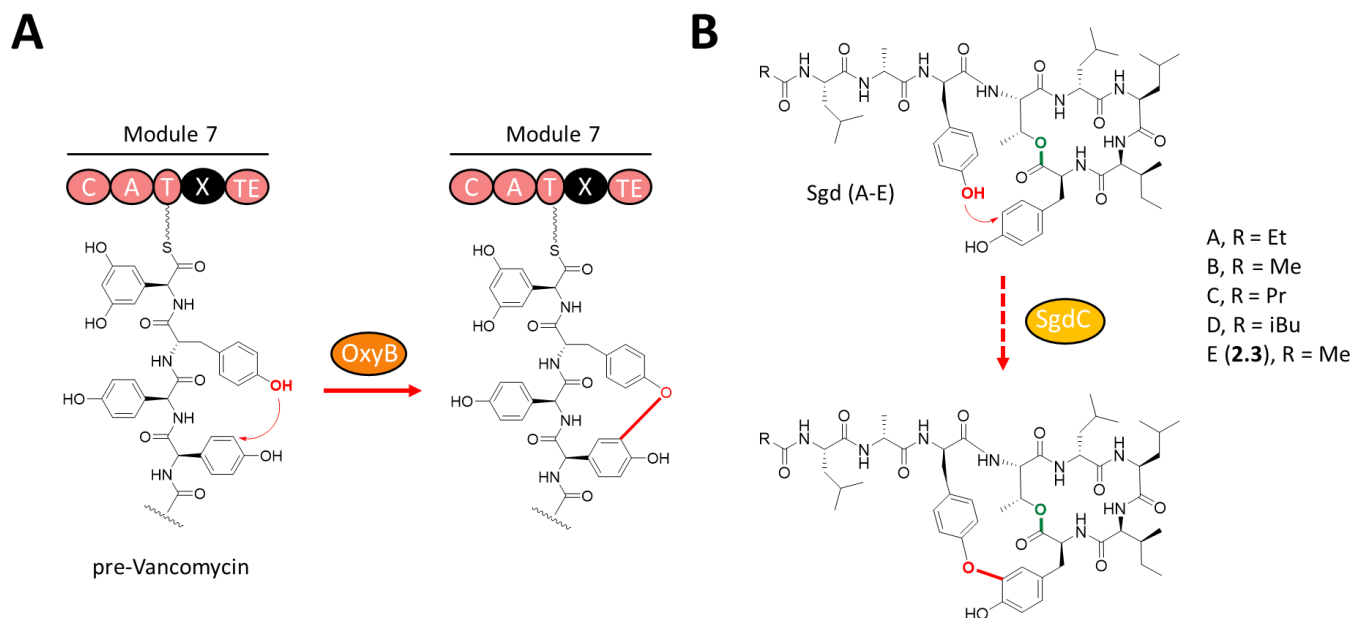


Figure 5.3. Biaryl ether coupling in vancomycin (A) and seongsanamide (B) biosynthesis.

The preliminary work on the gacamide A (Gac A) tandem TE domain (TE1-TE2) presented in **Chapter 3** highlights its potential as a versatile peptide macrocyclization biocatalyst. It demonstrated catalytic activity on two native-like peptide substrates with varying N-terminal groups (*N*-Cbz protecting group and *N*-acetyl), suggesting it was tolerant to N-terminal peptide modifications (**Figure 5.4A**). Additionally, it was shown to turnover two *de novo* peptides, generating non-native 5- and 7-residue macrolactones using an internal nucleophile that differed in terms of identity and stereochemical configuration (L-Thr) compared to the native D-Ser in the Gac A peptide (**Figure 5.4B**). In particular, the formation of a 7-residue macrolactone using a non-native nucleophile distinguishes it as a unique TE biocatalyst.^{4,5,11,12,15–19} Future studies on the tandem TE's substrate scope via a similar peptide library workflow presented in **Figure 5.2** would be a great addition to the literature. In addition, experimentally confirming the observed results in **Chapter 3** with purified *de novo* peptides is essential. Furthermore, the literature on tandem TE domains often yield conflicting results regarding the activities of the individual TEs (i.e. TE1 and TE2),^{18,19,43} elucidating the specific role of TE2 remains a critical aspect. Conducting

experiments using recombinantly purified TE2 as a stand-alone enzyme would provide further insight into this elusive TE2 domain.

Remarkably, the Gac-TE1-TE2 exhibits characteristics aligning with both ICS and CS models, as exemplified by its ability to process both native and non-native peptide substrates. The *N*-Cbz protected peptide substrate was shown to spontaneously macrocyclize. As both Tyc and patellamide A underwent spontaneous macrocyclization due to substrate pre-organization,^{28–30} the Gac *N*-Cbz peptide likely pre-organizes, suggesting Gac-TE1-TE2 operates within the framework of our proposed CS model. Conversely, non-native 5- and 7-residue macrolactones of *de novo* peptide substrates were only formed under enzymatic conditions. In addition, macrocycle ring opening and spontaneous macrocyclization were not observed with these peptides. This suggests that substrate pre-organization is not the driving force for macrocyclization catalysis, effectively ruling out the strict application of the CS model. Instead, the observed behaviour aligns more closely with our ICS model, wherein the TE biases the conformational landscape of a dynamic peptidyl-TE intermediate towards those conducive to macrocyclization. The substantial levels of seco acid observed in these assays further support the application of the ICS model.

This dual behaviour observed in Gac-TE1-TE2 underscores its inherent flexibility in macrocyclization. In addition, it serves as a compelling illustration of how the CS model aptly explains scenarios involving substrate pre-organization, while the ICS model offers a broader framework that can accommodate scenarios where substrate pre-organization is not the driving force. This coexistence and potential interplay between these models provide valuable insights into the diverse strategies employed by TEs in macrocyclization catalysis.

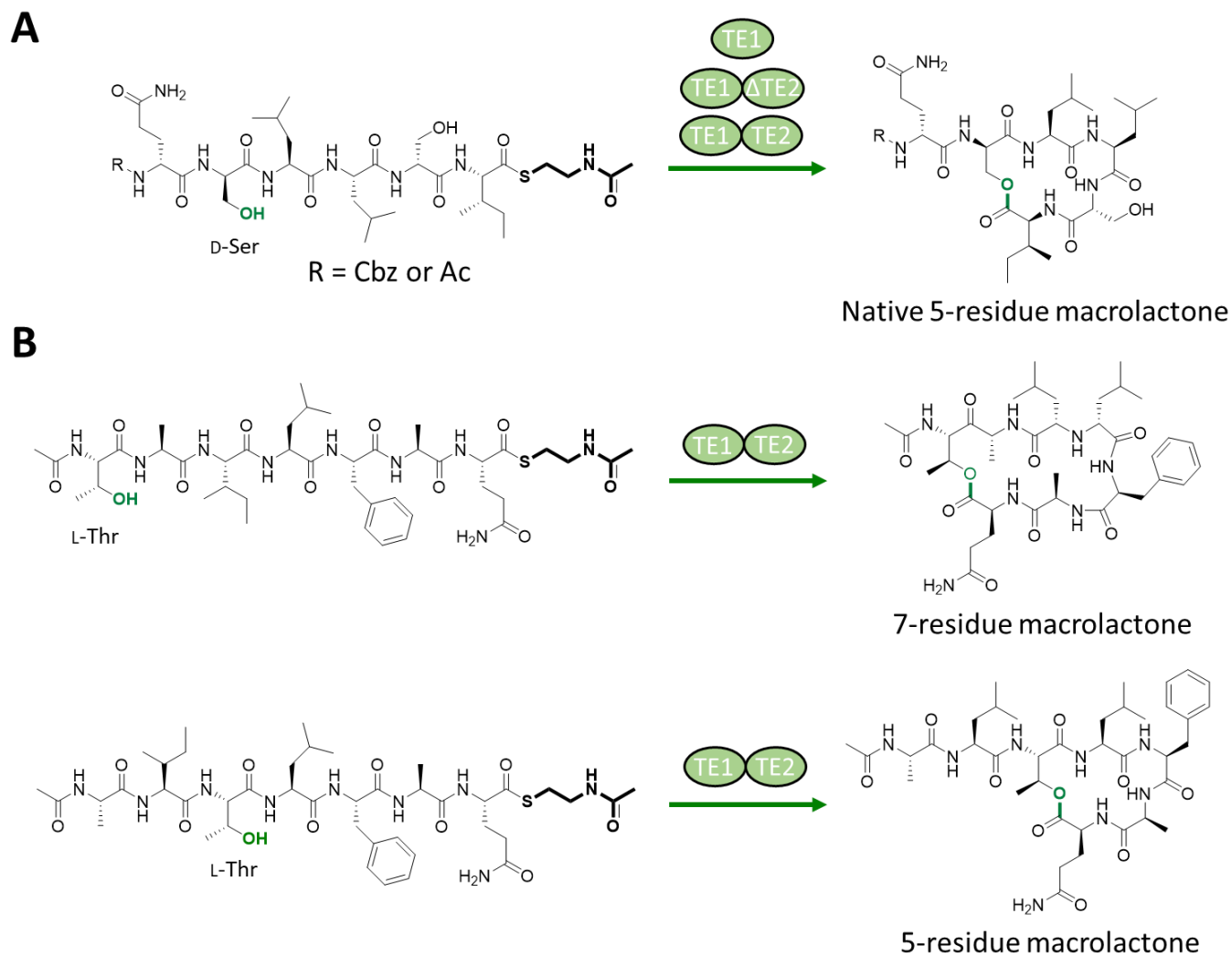


Figure 5.4. Macrocyclization activity of Gac-tandem domain on various substrates. (A) Formation of the native 5-residue macrolactone using a D-Ser nucleophile with varying N-terminal protecting group on the peptide substrate. (B) Formation of non-native 7- and 5-residue macrolactones using an L-Thr nucleophile of *de novo* peptide substrates.

5.4 Outlook on engineering TEs

The work on engineering the PKS 6-deoxyerythronolide B (DEBS) TE presented in **Chapter 4** describes the first extensive protein engineering study on a TE to date using an LCMS-based screen of a comprehensive 1,323 clone library. We showed that through rational protein design, single mutations within the active site are sufficient to yield improvements in macrocyclization activity over the WT. Mutations were not synergistic when combined; however, a double mutant retained a similar fold-improvement in activity compared to its single mutant counterparts. It was proposed that these mutations affected the dynamics of the TE by favouring the adoption of “closed” conformations that bias acyl-TE sampling towards compact conformers while limiting water’s access to the active site. In addition, key H-bonding interactions with the substrate’s nucleophilic alcohol, presumed to stabilize the observed unreactive cyclic conformation, were likely broken. This likely increases its ability to dynamically sample conformations, all of which favour formation of the macrocycle. Using 2,3-diaminopropionic acid (DAP) incorporation technology, acyl-TE intermediates can be trapped via formation of a stable amide bond without the need for maturation as seen in acylphosphonate inhibitors.¹² Thus, the structural determinants for off-loading in WT DEBS-TE and the improved mutants can be elucidated through X-ray crystallography. This would serve to confirm the above proposition and is currently being explored by our collaborators in Dr. Schmeing’s Lab. Furthermore, computational studies can provide useful insight on TE-mediated macrocyclization. Molecular dynamic simulation (MDS) and quantum mechanics/molecular mechanics (QM/MM) studies performed on the DEBS, PIK and pimaricin TEs underscore the importance of hydrophobic packing interactions in macrocyclization,^{27,44,45} highlighting the value of computational studies. With the WT DEBS-TE crystal structure in complex with a substrate mimic shown in **Chapter 4**, this provides a starting point for future computational studies.

While the screen could distinguish between active and inactive macrocyclization variants, the significant variation in macrocyclization activity of the positive control posed challenges in identifying mutants with improved activity over WT. This large variation in the positive control activity suggests variations in total protein amount contributes to this, a well documented and long-standing challenge in LCMS-based screens.^{46–48} Protein normalization using a joint superfolder GFP (sfGFP) and DEBS-TE construct may minimize variation and improve the screen's ability to differentiate improved variants from WT.⁴⁶ Reducing screen variability will improve this versatile screening platform, and can serve as a springboard for a second-generation DEBS-TE engineering project, focusing on iterative saturation mutagenesis of our 9 chosen sites and exploration of distal residues from the active site. Given that ICS TEs, such as DEBS, do not rely on substrate driven pre-organization for macrocyclization, they are ideal candidates for protein engineering to generate a more general and macrocyclization specific biocatalyst. In the coming years, we should see examples of other TEs being engineered through LCMS-based screens.

A substitution of the active site Ser to Cys (S→C) in the stereo-selective pikromycin (PIK) TE domain was enough to yield a variant with improved macrocyclization activity against an epimerized pentaketide substrate.²⁷ As both DEBS and PIK TEs are closely related, it would be worth investigating the S→C substitution in DEBS and the improved mutants from **Chapter 4** to see if a similar activity enhancement and stereo tolerance is observed.

5.5 Conclusions

To conclude, the development and use of TEs as macrocyclization biocatalysts is an ever-expanding field that has been met with great progress. **Chapter 2** characterizes and showcases the application of two new NRPS TEs as stand-alone macrocyclization biocatalysts in the chemoenzymatic synthesis of two depsipeptide macrolactones. **Chapter 3** examines the macrocyclic activity and surprising substrate tolerance of a tandem NRPS TE domain. **Chapter 4** engineered a PKS TE for improved macrocyclization activity through rational protein design and an LCMS-based screen. These studies have expanded our current repertoire of TE biocatalysts and have provided a versatile platform for improving TE activity via protein engineering.

5.6 References

- (1) Mortison, J. D.; Sherman, D. H. Frontiers and Opportunities in Chemoenzymatic Synthesis. *J. Org. Chem.* **2010**, *75* (21), 7041–7051. <https://doi.org/10.1021/jo101124n>.
- (2) Parenty, A.; Moreau, X.; Niel, G.; Campagne, J. M. Update 1 of: Macrolactonizations in the Total Synthesis of Natural Products. *Chem. Rev.* **2013**, *113* (1), PR1–PR40. <https://doi.org/10.1021/cr300129n>.
- (3) Fürstner, A. Lessons from Natural Product Total Synthesis: Macrocyclization and Postcyclization Strategies. *Acc. Chem. Res.* **2021**, *54* (4), 861–874. <https://doi.org/10.1021/acs.accounts.0c00759>.
- (4) Brazeau-Henrie, J. T.; Paquette, A. R.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N.; O'Rourke, A. Q.; Darnowski, M. G.; Boddy, C. N. Total and Chemoenzymatic Synthesis of Seongsanamide E. *Org. Lett.* **2022**, *24* (35), 6369–6373. <https://doi.org/10.1021/acs.orglett.2c02271>.
- (5) Lepetit, C. A.; Paquette, A. R.; Brazeau-Henrie, J. T.; Boddy, C. N. Total and Chemoenzymatic Synthesis of the Lipodepsipeptide Rhizomide A. *Bioorg. Med. Chem. Lett.* **2023**, *96* (Article 129506), 1–5. <https://doi.org/10.1016/j.bmcl.2023.129506>.
- (6) Heberlig, G. W.; Brown, J. T. C.; Simard, R. D.; Wirz, M.; Zhang, W.; Wang, M.; Susser, L. I.; Horsman, M. E.; Boddy, C. N. Chemoenzymatic Macrocyclic Synthesis Using Resorcylic Acid Lactone Thioesterase Domains. *Org. Biomol. Chem.* **2018**, *16* (32), 5771–5779. <https://doi.org/10.1039/c8ob01512k>.
- (7) Schmidt, J. J.; Khatri, Y.; Brody, S. I.; Zhu, C.; Pietraszkiewicz, H.; Valeriote, F. A.; Sherman, D. H. A Versatile Chemoenzymatic Synthesis for the Discovery of Potent Cryptophycin Analogs. *ACS Chem. Biol.* **2020**, *15* (2), 524–532. <https://doi.org/10.1021/acscchembio.9b00998>.
- (8) Hansen, D. A.; Rath, C. M.; Eisman, E. B.; Narayan, A. R. H.; Kittendorf, J. D.; Mortison, J. D.; Yoon, Y. J.; Sherman, D. H. Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin. *J. Am. Chem. Soc.* **2013**, *135* (30), 11232–11238. <https://doi.org/10.1021/ja404134f>.
- (9) Boddy, C. N.; Schneider, T. L.; Hotta, K.; Walsh, C. T.; Khosla, C. Epothilone C Macrolactonization and Hydrolysis Are Catalyzed by the Isolated Thioesterase Domain of Epothilone Polyketide Synthase. *J. Am. Chem. Soc.* **2003**, *125* (12), 3428–3429. <https://doi.org/10.1021/ja0298646>.
- (10) He, W.; Wu, J.; Khosla, C.; Cane, D. E. Macrolactonization to 10-Deoxymethynolide Catalyzed by the Recombinant Thioesterase of the Picromycin/Methymycin Polyketide Synthase. *Bioorganic Med. Chem. Lett.* **2006**, *16* (2), 391–394. <https://doi.org/10.1016/j.bmcl.2005.09.077>.
- (11) Heberlig, G. W.; Boddy, C. N. Thioesterase from Cereulide Biosynthesis Is Responsible for Oligomerization and Macrocyclization of a Linear Tetradepsipeptide. *J. Nat. Prod.* **2020**, *83* (6), 1990–1997. <https://doi.org/10.1021/acs.jnatprod.0c00333>.
- (12) Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dornan, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W. Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature* **2019**, *565* (7737), 112–117. <https://doi.org/10.1038/s41586-018-0781-z>.
- (13) Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. Peptide Cyclization Catalysed by the Thioesterase Domain of Tyrocidine Synthetase. *Nature* **2000**, *407* (6801), 215–218. <https://doi.org/10.1038/35025116>.
- (14) Hoyer, K. M.; Mahlert, C.; Marahiel, M. A. The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chem. Biol.* **2007**, *14* (1), 13–22. <https://doi.org/10.1016/j.chembiol.2006.10.011>.

- (15) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. Characterization of the Surfactin Synthetase C-Terminal Thioesterase Domain as a Cyclic Depsipeptide Synthase. *Biochemistry* **2002**, *41* (45), 13350–13359. <https://doi.org/10.1021/bi026592a>.
- (16) Mahlert, C.; Sieber, S. A.; Grünewald, J.; Marahiel, M. A. Chemoenzymatic Approach to Enantiopure Streptogramin B Variants: Characterization of Stereoselective Pristinamycin I Cyclase from *Streptomyces Pristinaespiralis*. *J. Am. Chem. Soc.* **2005**, *127* (26), 9571–9580. <https://doi.org/10.1021/ja051254t>.
- (17) Grünewald, J.; Sieber, S. A.; Marahiel, M. A. Chemo- and Regioselective Peptide Cyclization Triggered by the N-Terminal Fatty Acid Chain Length: The Recombinant Cyclase of the Calcium-Dependent Antibiotic from *Streptomyces Coelicolor*. *Biochemistry* **2004**, *43* (10), 2915–2925. <https://doi.org/10.1021/bi036140d>.
- (18) Mandalapu, D.; Ji, X.; Chen, J.; Guo, C.; Liu, W. Q.; Ding, W.; Zhou, J.; Zhang, Q. Thioesterase-Mediated Synthesis of Teixobactin Analogues: Mechanism and Substrate Specificity. *J. Org. Chem.* **2018**, *83* (13), 7271–7275. <https://doi.org/10.1021/acs.joc.7b02462>.
- (19) Hou, J.; Robbel, L.; Marahiel, M. A. Identification and Characterization of the Lysobactin Biosynthetic Gene Cluster Reveals Mechanistic Insights into an Unusual Termination Module Architecture. *Chem. Biol.* **2011**, *18* (5), 655–664. <https://doi.org/10.1016/j.chembiol.2011.02.012>.
- (20) Aldrich, C. C.; Venkatraman, L.; Sherman, D. H.; Fecik, R. A. Chemoenzymatic Synthesis of the Polyketide Macrolactone 10-Deoxymethynolide. *J. Am. Chem. Soc.* **2005**, *127* (25), 8910–8911. <https://doi.org/10.1021/ja0504340>.
- (21) Conway, K. R.; Boddy, C. N. ClusterMine360: A Database of Microbial PKS/NRPS Biosynthesis. *Nucleic Acids Res.* **2013**, *41* (D1), 402–407. <https://doi.org/10.1093/nar/gks993>.
- (22) Horsman, M. E.; Hari, T. P. A. A.; Boddy, C. N. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016**, *33* (2), 183–202. <https://doi.org/10.1039/c4np00148f>.
- (23) Hari, T. P. A.; Labana, P.; Boileau, M.; Boddy, C. N. An Evolutionary Model Encompassing Substrate Specificity and Reactivity of Type I Polyketide Synthase Thioesterases. *ChemBioChem* **2014**, *15* (18), 2656–2661. <https://doi.org/10.1002/cbic.201402475>.
- (24) Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. Generality of Peptide Cyclization Catalyzed by Isolated Thioesterase Domains of Nonribosomal Peptide Synthetases. *Biochemistry* **2001**, *40* (24), 7099–7108. <https://doi.org/10.1021/bi010036j>.
- (25) Argyropoulos, P.; Bergeret, F.; Pardin, C.; Reimer, J. M.; Pinto, A.; Boddy, C. N.; Schmeing, T. M. Towards a Characterization of the Structural Determinants of Specificity in the Macrocyclizing Thioesterase for Deoxyerythronolide B Biosynthesis. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (3), 486–497. <https://doi.org/10.1016/j.bbagen.2015.11.007>.
- (26) Akey, D. L.; Kittendorf, J. D.; Giraldez, J. W.; Fecik, R. A.; Sherman, D. H.; Smith, J. L. Structural Basis for Macrolactonization by the Pikromycin Thioesterase. *Nat. Chem. Biol.* **2006**, *2* (10), 537–542. <https://doi.org/10.1038/nchembio824>.
- (27) Koch, A. A.; Hansen, D. A.; Shende, V. V.; Furan, L. R.; Houk, K. N.; Jiménez-Osés, G.; Sherman, D. H. A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst. *J. Am. Chem. Soc.* **2017**, *139* (38), 13456–13465. <https://doi.org/10.1021/jacs.7b06436>.
- (28) Koehnke, J.; Bent, A. F.; Houssen, W. E.; Mann, G.; Jaspars, M.; Naismith, J. H. The Structural Biology of Patellamide Biosynthesis. *Curr. Opin. Struct. Biol.* **2014**, *29*, 112–121. <https://doi.org/10.1016/j.sbi.2014.10.006>.
- (29) Milne, B. F.; Long, P. F.; Starcevic, A.; Hranueli, D.; Jaspars, M. Spontaneity in the Patellamide

- Biosynthetic Pathway. *Org. Biomol. Chem.* **2006**, *4* (4), 631–638.
<https://doi.org/10.1039/b515938e>.
- (30) Bu, X.; Wu, X.; Xie, G.; Guo, Z. Synthesis of Tyrocidine A and Its Analogues by Spontaneous Cyclization in Aqueous Solution. *Org. Lett.* **2002**, *4* (17), 2893–2895.
<https://doi.org/10.1021/ol0263191>.
 - (31) Brady, S. F.; Varga, S. L.; Freidinger, R. M.; Schwenk, D. A.; Mendlowski, M.; Holly, F. W.; Veber, D. F. Practical Synthesis of Cyclic Peptides, with an Example of Dependence of Cyclization Yield upon Linear Sequence. *J. Org. Chem.* **1979**, *44* (18), 3101–3105. <https://doi.org/10.1021/jo01332a003>.
 - (32) Kopple, K. D. Synthesis of Cyclic Peptides. *J. Pharm. Sci.* **1972**, *61* (9), 1345–1356.
<https://doi.org/10.1002/jps.2600610902>.
 - (33) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry* **2001**, *40* (24), 7092–7098. <https://doi.org/10.1021/bi010035r>.
 - (34) Gibbs, C. A.; Kondejewski, H. L.; Gronwald, W.; Nip, M. A.; Hodges, S. R.; Sykes, D. B.; Wishart, S. D. Unusual β -Sheet Periodicity in Small Cyclic Peptides. *Nat. Struct. Biol.* **1998**, *5* (4), 284–288.
 - (35) Yeh, E.; Lin, H.; L. Clugston, S.; M. Kohli, R.; T. Walsh, C. Enhanced Macrocyclizing Activity of the Thioesterase from Tyrocidine Synthetase in Presence of Nonionic Detergent. *Chem. Biol.* **2004**, *11* (11), 1573–1582. <https://doi.org/10.1016/j.chembiol.2004.09.003>.
 - (36) Szymczak, L. C.; Kuo, H. Y.; Mrksich, M. Peptide Arrays: Development and Application. *Anal. Chem.* **2018**, *90* (1), 266–282. <https://doi.org/10.1021/acs.analchem.7b04380>.
 - (37) Volkmer, R. Synthesis and Application of Peptide Arrays: Quo Vadis SPOT Technology. *ChemBioChem* **2009**, *10* (9), 1431–1442. <https://doi.org/10.1002/cbic.200900078>.
 - (38) Mootz, H. D.; Kessler, N.; Linne, U.; Eppelmann, K.; Schwarzer, D.; Marahiel, M. A. Decreasing the Ring Size of a Cyclic Nonribosomal Peptide Antibiotic by In-Frame Module Deletion in the Biosynthetic Genes. *J. Am. Chem. Soc.* **2002**, *124* (37), 10980–10981.
<https://doi.org/10.1021/ja027276m>.
 - (39) Kohli, R. M.; Burke, M. D.; Tao, J.; Walsh, C. T. Chemoenzymatic Route to Macrocyclic Hybrid Peptide/Polyketide-like Molecules. *J. Am. Chem. Soc.* **2003**, *125* (24), 7160–7161.
<https://doi.org/10.1021/ja0352202>.
 - (40) Kim, G. J.; Li, X.; Kim, S. H.; Yang, I.; Hahn, D.; Chin, J.; Nam, S. J.; Nam, J. W.; Nam, D. H.; Oh, D. C.; Chang, H. W.; Choi, H. Seongsanamides A-D: Antiallergic Bicyclic Peptides from *Bacillus Safensis* KCTC 12796BP. *Org. Lett.* **2018**, *20* (23), 7539–7543.
<https://doi.org/10.1021/acs.orglett.8b03293>.
 - (41) Woihte, K.; Geib, N.; Zerbe, K.; Dong, B. L.; Heck, M.; Fournier-Rousset, S.; Meyer, O.; Vitali, F.; Matoba, N.; Abou-Hadeed, K.; Robinson, J. A. Oxidative Phenol Coupling Reactions Catalyzed by OxyB: A Cytochrome P450 from the Vancomycin Producing Organism. Implications for Vancomycin Biosynthesis. *J. Am. Chem. Soc.* **2007**, *129* (21), 6887–6895.
<https://doi.org/10.1021/ja071038f>.
 - (42) Aldemir, H.; Shu, S.; Schaefer, F.; Hong, H.; Richarz, R.; Gulder, T. A. M. Carrier Protein-Free Enzymatic Biaryl Coupling in Arylomycin A2 Assembly and Structure of the Cytochrome P450 AryC. *Chem. - A Eur. J.* **2021**, *28* (2), e2021103389. <https://doi.org/10.1002/chem.202103389>.
 - (43) Roongsawang, N.; Washio, K.; Morikawa, M. In Vivo Characterization of Tandem C-Terminal Thioesterase Domains in Arthrofactin Synthetase. *ChemBioChem* **2007**, *8* (5), 501–512.
<https://doi.org/10.1002/cbic.200600465>.
 - (44) Fan, S.; Wang, R.; Li, C.; Bai, L.; Zhao, Y. L.; Shi, T. Insight into Structural Characteristics of Protein-Substrate Interaction in Pimaricin Thioesterase. *Int. J. Mol. Sci.* **2019**, *20* (4).

- <https://doi.org/10.3390/ijms20040877>.
- (45) Chen, X. P.; Shi, T.; Wang, X. L.; Wang, J.; Chen, Q.; Bai, L.; Zhao, Y. L. Theoretical Studies on the Mechanism of Thioesterase-Catalyzed Macrocyclization in Erythromycin Biosynthesis. *ACS Catal.* **2016**, *6* (7), 4369–4378. <https://doi.org/10.1021/acscatal.6b01154>.
 - (46) Santos-aberturas, J.; Do, M.; Waldo, G. S.; Uwe, T.; Waldo, G. S.; Bornscheuer, U. T. In-Depth High-Throughput Screening of Protein Engineering Libraries by Split-GFP Direct Crude Cell Extract Data Normalization Resource In-Depth High-Throughput Screening of Protein Engineering Libraries by Split-GFP Direct Crude Cell Extract Data Normaliz. *Chem. Biol.* **2015**, *22* (10), 1406–1414. <https://doi.org/10.1016/j.chembiol.2015.08.014>.
 - (47) Taverna, D. M.; Goldstein, R. A. Why Are Proteins Marginally Stable ? *Proteins Struct. Funct. Genet.* **2002**, *46*, 105–109. <https://doi.org/10.1002/prot.10016>.
 - (48) Mansoury, M.; Hamed, M.; Karmustaji, R.; Hannan, F. Al; Safrany, S. T. The Edge Effect : A Global Problem. The Trouble with Culturing Cells in 96-Well Plates. *Biochem. Biophys. Reports* **2021**, *26* (Article 100987), 1–5. <https://doi.org/10.1016/j.bbrep.2021.100987>.

APPENDIX I: Permissions and manuscripts

Total and Chemoenzymatic Synthesis of Seongsanamide E

Author: Jordan T. Brazeau-Henrie, André R. Paquette, Allison Q. O'Rourke, et al

Publication: Organic Letters

Publisher: American Chemical Society

Date: Sep 1, 2022

Copyright © 2022, American Chemical Society



PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK

CLOSE WINDOW



Total and chemoenzymatic synthesis of the lipodepsipeptide rhizomide A

Author: Corinne A. Lepetit, André R. Paquette, Jordan T. Brazeau-Henrie, Christopher N. Boddy

Publication: Bioorganic & Medicinal Chemistry Letters

Publisher: Elsevier

Date: 15 November 2023

© 2023 Elsevier Ltd. All rights reserved.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

Total and Chemoenzymatic Synthesis of Seongsanamide E

Jordan T. Brazeau-Henrie, André R. Paquette, Allison Q. O'Rourke, Michael G. Darnowski, and Christopher N. Boddy*



Cite This: *Org. Lett.* 2022, 24, 6369–6373



Read Online

ACCESS |



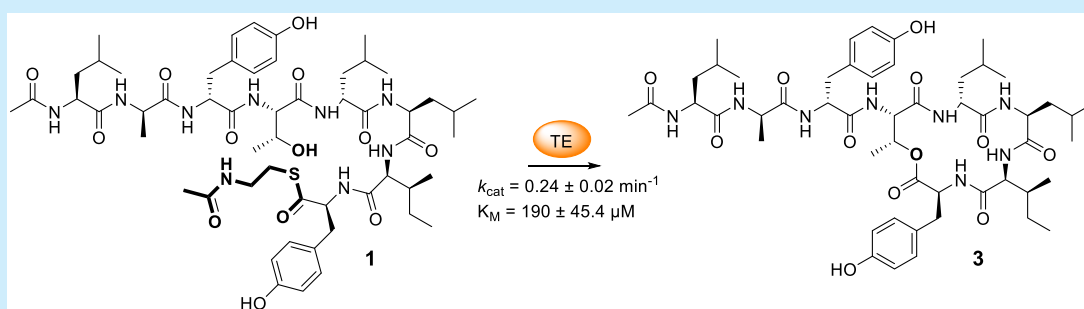
Metrics & More



Article Recommendations



Supporting Information



ABSTRACT: The total and chemoenzymatic synthesis of the depsipeptide natural product seongsanamide E, **3**, is described. The synthetic C-terminal *N*-acetylcysteamine thioester of linear natural product **1** was macrolactonized by the excised recombinant purified seongsanamide thioesterase (Sgd-TE) domain, generating **3**. Sgd-TE also effects the ring opening of **3**. Chemical synthesis provided **3** through a macrolactamization strategy. This work confirms the biosynthesis of **3** and demonstrates the power of Sgd-TE as a biocatalyst.

The seongsanamides (Sgds) make up a family of cyclic nonribosomal depsipeptides isolated from the marine sponge-associated bacterium *Bacillus safensis* KCTC 12796BP.¹ There are six known seongsanamides, A–F. A–E are *N*-acylated octapeptides with a lactone linkage between the terminal carboxylate of Tyr₈ and the side chain hydroxyl of Thr₄. All five have the L-Leu-D-Ala-D-Tyr-L-Thr-D-Leu-L-Leu-L-Ile-L-Tyr peptide sequence. Sequencing of the *B. safensis* genome identified a nonribosomal peptide synthetase (NRPS) gene cluster^{1,2} that encoded two NRPS genes possessing A domain selectivity,^{3–6} consistent with these amino acids and E domains appropriately placed to generate the three D-configured residues.

On the basis of this putative gene cluster, it is proposed that a linear NRPS-bound octapeptide is generated and macrocyclized seongsanamide E, **3**, is then released¹ by the C-terminal thioesterase (TE) domain (Figure 1).⁷ Ultimately, **3** is proposed to be further cyclized by a P450,⁸ forming a biaryl ether linkage between Tyr₃ and Tyr₈ via oxidative phenol coupling, generating seongsanamide B (Figure S1). In this study, we characterized the ability of the TE to macrocyclize the *N*-acetylcysteamine (SNAC) derivative of the native linear peptide substrate. We show the recombinant purified excised Sgd-TE effectively forms **3**, experimentally confirming that the biosynthetic gene cluster identified by Kim et al.¹ is responsible for seongsanamide biosynthesis. We also show that chemical macrolactonization does not occur from the linear free acid substrate but that an authentic macrocycle can be synthesized

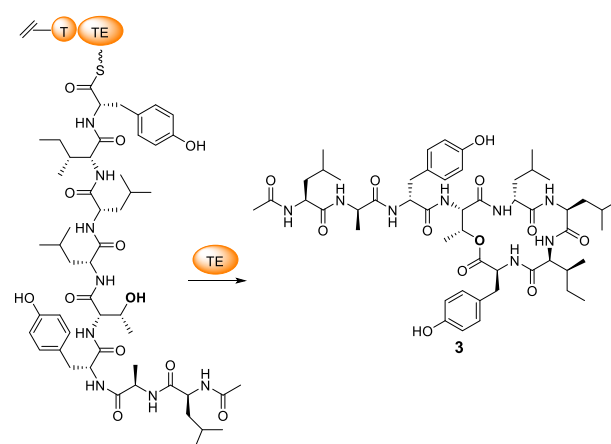


Figure 1. Seongsanamide E TE-mediated formation.

via lactamization, successfully completing a total synthesis of **3** and demonstrating the macrolactonizing power of the Sgd-TE. Lastly, we show that the TE can effect hydrolytic ring opening

Received: July 5, 2022

Published: August 25, 2022

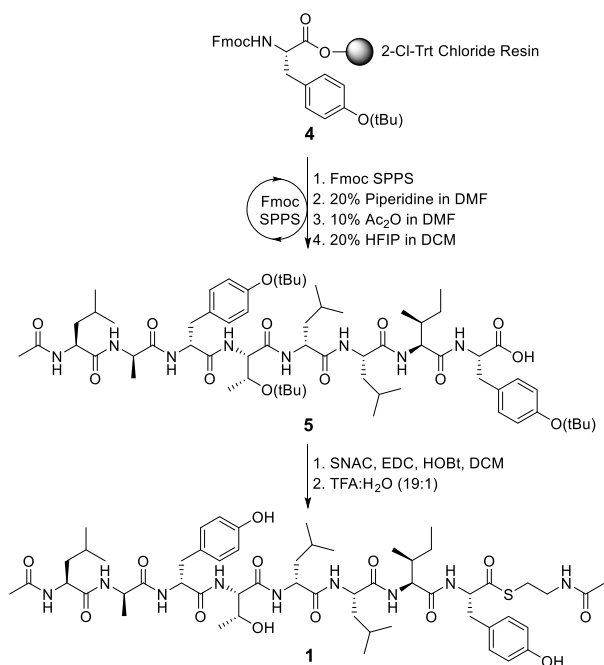


of **3**, a rare enzymatic activity seen for only the epothilone (Epo) TE.⁹

Characterization of the excised Sgd-TE domain required access to an appropriate substrate, an authentic standard of the natural product, and the recombinant purified TE domain. Recently, seongsanamides A,¹⁰ B,¹¹ and E¹² have been successfully synthesized using a combination of solution and solid-phase peptide synthesis. We envisioned a similar strategy for accessing SNAC-activated substrate **1** and authentic natural product **3**.

C-Terminal thioester peptides like **1** are synthetically challenging.¹³ While linear unactivated peptides can be routinely synthesized with automated Fmoc solid-phase peptide synthesis (SPPS), C-terminal thioesters require specialized resins^{14,15} or synthetic modifications.^{16,17} In our hands, the most effective approach for accessing the peptidyl-SNAC thioesters is to install the thioester using solution-phase chemistry on a protected linear free acid peptide.^{18,19} Thus, we generated the on-resin linear protected peptide from hypersensitive 2-Cl-Trt chloride resin²⁰ loaded with the C-terminal Fmoc-Tyr(OtBu)-OH **4**, which was subjected to SPPS under HBTU/DIEA conditions, subsequent N-terminal Fmoc deprotection, and acetylation (Scheme 1). Cleavage from the

Scheme 1. SPPS of Linear SNAC Substrate 1



resin under mild acidic conditions yielded free acid **5**, which was activated as the SNAC thioester. The side chain tBu protecting groups were removed with TFA, and the substrate, **1**, was purified to >95% purity using RP-HPLC [65% yield from **5** (Figure S41)].

Synthesis of authentic **3** was not possible via a macrolactonization strategy. All attempts to generate a macrocycle from a linear free acid with an unmasked Thr β -hydroxyl group failed.²¹ We thus turned our attention to a common alternative approach for depsipeptide cyclization,^{22–24} macrolactamization. We focused on cyclization through the Ile-Tyr amide bond. Wang resin loaded with Fmoc-Ile-OH **6** underwent SPPS under DIC/Oxyma Pure conditions to extend the linear peptide sequence, incorporating Fmoc-Thr-OH **7** with an

unprotected β -alcohol. On-resin peptide ester **8** was formed through standard esterification conditions with Boc-Tyr-(OtBu)-OH, followed by N-terminal deprotection and acetylation to complete the linear sequence. The peptide was fully deprotected and cleaved from the resin using TFA to afford the linear peptide ester. Macrolactamization under dilute HATU/HOAt coupling conditions generated **3**.^{11,25} Upon LCMS analysis, two peaks with the same mass were observed eluting close together. The isomers were separated by RP-HPLC, and ¹H NMR (Figures S43 and S44) confirmed the first eluting isomer was the natural product seongsanamide E, **3**. Coupling conditions known to reduce the level of epimerization at the Ile α -carbon were screened, though no difference in product distribution was observed.²⁵ Regardless, authentic **3** could be isolated in >95% purity by RP-HPLC [32% yield (Figure S43)].

To obtain recombinant purified Sgd-TE, a synthetic DNA construct encoding the excised Sgd-TE from the *B. safensis* genome was generated.¹ This gene was cloned into an *Escherichia coli* expression vector under the control of an inducible T7 promoter and fused to an N-terminal hexa-His tag. Sgd-TE was overexpressed and purified by metal-affinity chromatography to high purity [>95% (see Figure S3)]. Enzymatic assays with 1 mM **1** and recombinant purified Sgd-TE (5 μ M) in phosphate buffer (50 mM, pH 7.4) and 7.5% (v/v) DMSO were performed at room temperature. Aliquots were taken at varying time points and analyzed by LCMS to evaluate the product distribution.

Incubation of **1** with Sgd-TE showed formation of macrocycle **3**, confirming that the stand-alone TE is active (Figure 2). No formation of **3** was observed in the no enzyme or boiled enzyme control incubations. **3** was readily detected at the 30 min time point with Sgd-TE (Figure S11). At 1.5 h, formation of hydrolysis product **2** could be detected above background in the Sgd-TE incubation (Figure 2B) but not in the control incubations. At 20 h, significantly more hydrolysis product was detected in the Sgd-TE incubations (Figure 2C). Because both **3** and **2** are generated from the same peptidyl-TE intermediate, the relative product ratios for macrolactamization and hydrolysis should not change over the course of the reaction. The observation that the level of hydrolysis increases with prolonged incubation thus suggested additional reactivity was occurring (Scheme 2).

As the Epo-TE has been shown to effect ring opening of epothilone C,⁹ we hypothesized that Sgd-TE could be effecting ring opening of **3**, ultimately leading to additional hydrolysis product **2**. To test this, 0.5 mM **3** was incubated with Sgd-TE (5 μ M) in phosphate buffer (50 mM, pH 7.4) and 7.5% (v/v) DMSO at room temperature. LCMS analysis showed the formation of **2**, confirming that Sgd-TE can hydrolyze **3** (Figure 2F).

Sgd-TE consumption of **1** was kinetically characterized. Sgd-TE (5 μ M) was incubated with varying concentrations of **1** (0–750 μ M) in phosphate buffer (50 mM, pH 7.4) and 7.5% (v/v) DMSO. DTNB was also added to each incubation to continuously track the release of SNAC at 412 nm as substrate **1** is consumed. Incubations were performed for 1 h at room temperature. Initial velocities were all linear over the time course of the incubations, and <10% of **1** was consumed. Fitting these data to the Michaelis–Menten model showed Sgd-TE to be able to turn over **1** in the presence of DTNB with a k_{cat} of $0.24 \pm 0.02 \text{ min}^{-1}$ and a K_{M} of $190 \pm 45.4 \mu\text{M}$ [$k_{\text{cat}}/K_{\text{M}} = 21.0 \pm 5.3 \text{ M}^{-1} \text{ s}^{-1}$ (Figure S5)]. These kinetic

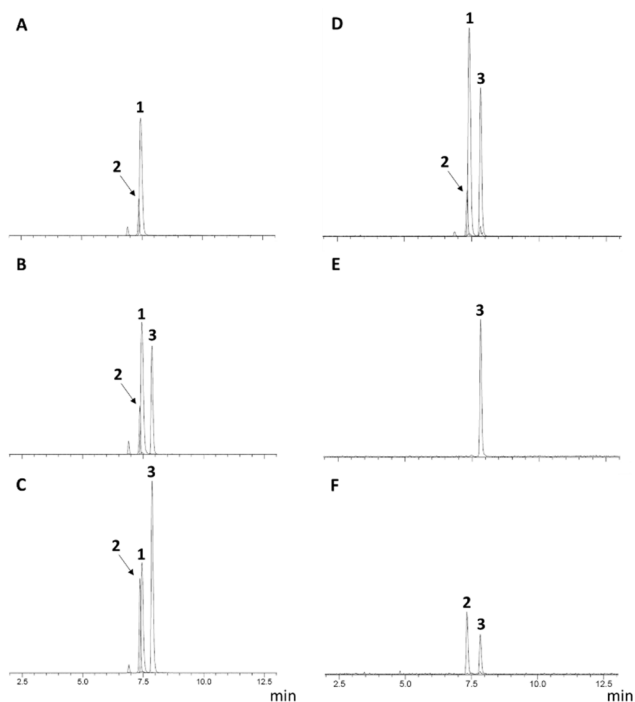


Figure 2. Extracted ion chromatograms of biochemical assays. (A) Substrate **1** only control. The boiled TE control showed a similar result (see the Supporting Information). Sgd-TE assays after incubation for (B) 1.5 and (C) 20 h with 1 mM **1**. (D) Co-injection of the Sgd-TE 3 h incubation assay (1 mM **1**) with macrocycle standard **3**. (E) Macrocycle **3** only control. The boiled TE control showed a similar result (see the Supporting Information). (F) Sgd-TE assay after 20 h with 0.5 mM **3**.

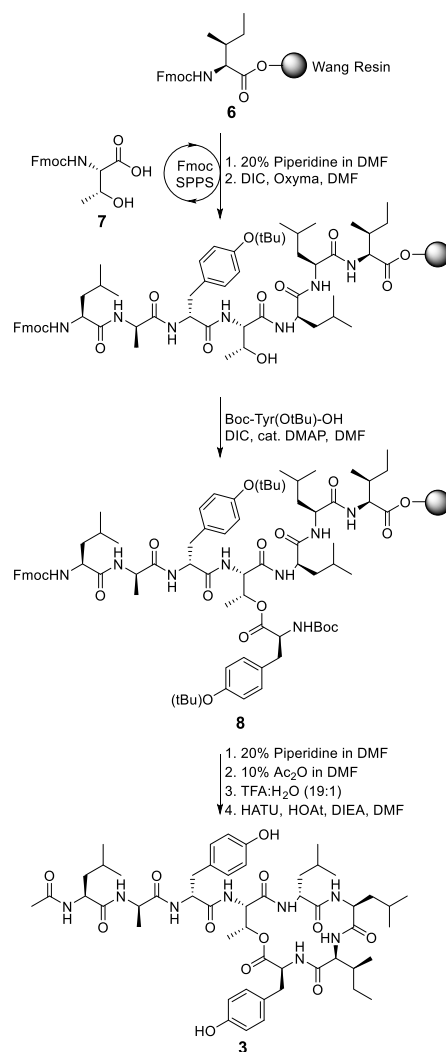
parameters are comparable to those obtained from other previously characterized NRPS TEs.^{19,26–32}

Macrocyclization is a demanding type of reactivity. Formation of a macrolactone from a thioester substrate does not occur spontaneously as demonstrated by the no enzyme and boiled enzyme negative controls (Figures S18 and S19). Furthermore, even substrate tolerant macrocyclizing TEs, like the DEBS-TE,^{33–39} are unable to macrocyclize **1** (Figure S23). Thus, the ability of the Sgd-TE to convert the SNAC-activated analogue of its native substrate **1** into natural product **3** experimentally confirms that the gene cluster identified through bioinformatic analysis by Kim et al.¹ is responsible for the biosynthesis of the seongsanamide family of natural products.

Sgd-TE also demonstrates potential as a potent biocatalyst. While we were ultimately able to perform the total synthesis of **3**, macrocyclization proved to be challenging. Macrolactonization was not possible under non-enzymatic conditions, a problem common in depsipeptide total syntheses.^{11,40–42} Ultimately, we relied on macrolactamization to access authentic **3** in 32% yield, though the macrocyclization suffered from unwanted isomerization, presumably due to epimerization of the activated ester.²⁵ Sgd-TE, on the contrary, readily generated the depsipeptide product in an unoptimized 27% yield through macrolactonization and did not suffer from epimerization issues, thus proving itself as a powerful biocatalyst and a promising alternative strategy for depsipeptide synthesis.

Lastly, we show that Sgd-TE can ring open its cognate natural product **3**. While the Epo-TE has been shown to ring

Scheme 2. SPPS of Seongsanamide E, **3**



open epothilone **C**,⁹ other characterized TEs do not demonstrate this activity.⁷ In the case of epothilone, the ester bond is sterically very accessible and is known to be highly sensitive to hydrolysis, a problem that limits the serum half-life of epothilone in clinical settings.^{43,44} In the case of seongsanamide, the ester is far more hindered and thus more stable. Ring opening by Sgd-TE suggests a discrete binding of **3**, consistent with enzyme–substrate interactions specifically stabilizing a late transition state for macrocyclization, thus binding the product in a conformation that enables attack of the active site Ser on the lactone carbonyl. This is supported by the low Michaelis constant, which is atypical of macrocyclizing PKS TEs and more similar to those of macrocyclizing NRPS TEs,^{28,45–47} which, at least in the case of the tyrocidine (Tyc) TE, bind their substrate in a very product-like conformation.^{47–49} Interestingly, product inhibition of macrocyclization might be anticipated from this model; however, our kinetic characterization fits a Michaelis–Menten model without recourse to including a product inhibition step.

In summary, using SPPS we have completed the total synthesis of seongsanamide **E 3** through a macrolactamization strategy. In a complementary biocatalytic strategy using recombinant purified Sgd-TE, we have effectively macrolactonized the SNAC-activated biosynthetic precursor to **3**.

This result confirms the involvement of the previously identified biosynthetic gene cluster in seongsanamide biosynthesis and shows the power of TEs in accessing macrocycles. Lastly, we have shown that Sgd-TE can effect ring opening of the natural product, providing new insights into potential enzyme substrate interactions driving macrocyclization.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.orglett.2c02271>.

NMR data, LCMS chromatograms, mass spectra, UV traces, kinetic data, and synthetic and experimental procedures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Christopher N. Boddy – Department of Chemistry and Biomolecular Sciences, Centre for Chemical and Synthetic Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada; orcid.org/0000-0002-3259-3184; Email: cboddy@uottawa.ca

Authors

Jordan T. Brazeau-Henrie – Department of Chemistry and Biomolecular Sciences, Centre for Chemical and Synthetic Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada

André R. Paquette – Department of Chemistry and Biomolecular Sciences, Centre for Chemical and Synthetic Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada

Allison Q. O'Rourke – Department of Chemistry and Biomolecular Sciences, Centre for Chemical and Synthetic Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada

Michael G. Darnowski – Department of Chemistry and Biomolecular Sciences, Centre for Chemical and Synthetic Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.orglett.2c02271>

Author Contributions

J.T.B.-H. and A.R.P. contributed equally to this work. J.T.B.-H. performed cloning, protein expression, purification, kinetic analysis, enzymatic assays, and LCMS analysis of the biochemical assays. A.R.P. synthesized and characterized all compounds and performed LCMS analysis of the biochemical assays. A.Q.O. performed enzymatic assays and HPLC analysis of the biochemical assays. M.G.D. performed compound characterization.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) (RGPIN 2019-06859). The authors thank the John L. Holmes Mass Spectrometry Facility for assistance in acquiring mass spectrometry data and

the University of Ottawa Faculty of Science NMR facility for assistance in collecting NMR spectra.

■ REFERENCES

- (1) Kim, G. J.; Li, X.; Kim, S. H.; Yang, I.; Hahn, D.; Chin, J.; Nam, S. J.; Nam, J. W.; Nam, D. H.; Oh, D. C.; et al. Seongsanamides A-D: Antiallergic Bicyclic Peptides from *Bacillus Safensis* KCTC 12796BP. *Org. Lett.* **2018**, *20* (23), 7539–7543.
- (2) Hanh, N. P. K.; Kim, S. H.; Kim, G. J.; Choi, H.; Nam, D. H. The Complete Genome Sequence of a Marine Sponge-Associated Bacteria, *Bacillus Safensis* KCTC 12796BP, Which Produces the Anti-Allergic Compounds. *Misaengmul Hakhoehi* **2018**, *54* (4), 448–452.
- (3) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The Specificity-Confering Code of Adenylation Domains in Nonribosomal Peptide Synthetases. *Chem. Biol.* **1999**, *6* (8), 493–505.
- (4) Kries, H.; Niquille, D. L.; Hilvert, D. A Subdomain Swap Strategy for Reengineering Nonribosomal Peptides. *Chem. Biol.* **2015**, *22* (5), 640–648.
- (5) Crüsemann, M.; Kohlhaas, C.; Piel, J. Evolution-Guided Engineering of Nonribosomal Peptide Synthetase Adenylation Domains. *Chem. Sci.* **2013**, *4* (3), 1041–1045.
- (6) Challis, G. L.; Ravel, J.; Townsend, C. A. Predictive, Structure-Based Model of Amino Acid Recognition by Nonribosomal Peptide Synthetase Adenylation Domains. *Chem. Biol.* **2000**, *7* (3), 211–224.
- (7) Horsman, M. E.; Hari, T. P. A. A.; Boddy, C. N. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat. Prod. Rep.* **2016**, *33* (2), 183–202.
- (8) Podust, L. M.; Sherman, D. H. Diversity of P450 Enzymes in the Biosynthesis of Natural Products. *Nat. Prod. Rep.* **2012**, *29* (10), 1251–1266.
- (9) Boddy, C. N.; Schneider, T. L.; Hotta, K.; Walsh, C. T.; Khosla, C. Epithilone C Macrolactonization and Hydrolysis Are Catalyzed by the Isolated Thioesterase Domain of Epithilone Polyketide Synthase. *J. Am. Chem. Soc.* **2003**, *125* (12), 3428–3429.
- (10) Han, F.; Guo, Y.; Ye, T. Total Synthesis of Antiallergic Bicyclic Peptide Seongsanamide A. *Org. Chem. Front.* **2020**, *7* (13), 1658–1662.
- (11) Shabani, S.; Hutton, C. A. Total Synthesis of Seongsanamide B. *Org. Lett.* **2020**, *22* (11), 4557–4561.
- (12) Shabani, S.; Hutton, C. A. Depsipeptide Synthesis Using a Late-Stage Ag(I)-Promoted Macrolactonisation of Peptide Thioamides. *Chem. Commun.* **2021**, *57* (16), 2081–2084.
- (13) Dawson, P. E.; Kent, S. B. H. Synthesis of Native Proteins by Chemical Ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
- (14) Blanco-Canosa, J. B.; Dawson, P. E. An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation. *Angew. Chemie - Int. Ed.* **2008**, *47* (36), 6851–6855.
- (15) Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. Solid Phase Synthesis of Peptide C-Terminal Thioesters by Fmoc/t-Bu Chemistry. *J. Am. Chem. Soc.* **1999**, *121* (49), 11369–11374.
- (16) Zheng, J. S.; Tang, S.; Qi, Y. K.; Wang, Z. P.; Liu, L. Chemical Synthesis of Proteins Using Peptide Hydrazides as Thioester Surrogates. *Nat. Protoc.* **2013**, *8* (12), 2483–2495.
- (17) Elashal, H. E.; Sim, Y. E.; Raj, M. Serine Promoted Synthesis of Peptide Thioester-Precursor on Solid Support for Native Chemical Ligation. *Chem. Sci.* **2017**, *8* (1), 117–123.
- (18) Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dorman, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W. Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature* **2019**, *565* (7737), 112–117.
- (19) Heberlig, G. W.; Boddy, C. N. Thioesterase from Cereulide Biosynthesis Is Responsible for Oligomerization and Macrocyclization of a Linear Tetradepsipeptide. *J. Nat. Prod.* **2020**, *83* (6), 1990–1997.
- (20) Von Eggelkraut-Gottanka, R.; Klose, A.; Beck-Sickinger, A. G.; Beyermann, M. Peptide Athioester Formation Using Standard Fmoc-Chemistry. *Tetrahedron Lett.* **2003**, *44* (17), 3551–3554.

- (21) Chandra, J.; Manne, S. R.; Mondal, S.; Mandal, B. (E)-Ethyl-2-Cyano-2-(((2,4,6-Trichlorobenzoyl)Oxy)Imino)Acetate: A Modified Yamaguchi Reagent for Enantioselective Esterification, Thioesterification, Amidation, and Peptide Synthesis. *ACS Omega* **2018**, *3* (6), 6120–6133.
- (22) Gimenez, D.; Phelan, A.; Murphy, C. D.; Cobb, S. L. Fengycin A Analogues with Enhanced Chemical Stability and Antifungal Properties. *Org. Lett.* **2021**, *23* (12), 4672–4676.
- (23) Hawkins, P. M. E.; Giltrap, A. M.; Nagalingam, G.; Britton, W. J.; Payne, R. J. Total Synthesis of Ecumicin. *Org. Lett.* **2018**, *20* (4), 1019–1022.
- (24) Moreira, R.; Wolfe, J.; Taylor, S. D. A High-Yielding Solid-Phase Total Synthesis of Daptomycin Using a Fmoc SPPS Stable Kynurenine Synthron. *Org. Biomol. Chem.* **2021**, *19* (14), 3144–3153.
- (25) Kaur, H.; Harris, P. W. R.; Little, P. J.; Brimble, M. A. Total Synthesis of the Cyclic Depsipeptide YM-280193, a Platelet Aggregation Inhibitor. *Org. Lett.* **2015**, *17* (3), 492–495.
- (26) Lin, H.; Thayer, D. A.; Wong, C. H.; Walsh, C. T. Macrolactamization of Glycosylated Peptide Thioesters by the Thioesterase Domain of Tyrocidine Synthetase. *Chem. Biol.* **2004**, *11* (12), 1635–1642.
- (27) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. Characterization of the Surfactin Synthetase C-Terminal Thioesterase Domain as a Cyclic Depsipeptide Synthase. *Biochemistry* **2002**, *41* (45), 13350–13359.
- (28) Hoyer, K. M.; Mahlert, C.; Marahiel, M. A. The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chem. Biol.* **2007**, *14* (1), 13–22.
- (29) Samel, S. A.; Wagner, B.; Marahiel, M. A.; Essen, L. O. The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-Ribosomal Lipopeptide. *J. Mol. Biol.* **2006**, *359* (4), 876–889.
- (30) Mulyani, S.; Egel, E.; Kittel, C.; Turkanovic, S.; Wohlleben, W.; Süßmuth, R. D.; Van Pée, K. H. The Thioesterase Bhp Is Involved in the Formation of β -Hydroxytyrosine during Balhimycin Biosynthesis in Amycolatopsis Balhimycina. *ChemBioChem.* **2010**, *11* (2), 266–271.
- (31) Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A. Regeneration of Misprimed Nonribosomal Peptide Synthetases by Type II Thioesterases. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 14083.
- (32) Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. Generality of Peptide Cyclization Catalyzed by Isolated Thioesterase Domains of Nonribosomal Peptide Synthetases. *Biochemistry* **2001**, *40* (24), 7099–7108.
- (33) Argyropoulos, P.; Bergeret, F.; Pardin, C.; Reimer, J. M.; Pinto, A.; Boddy, C. N.; Schmeing, T. M. Towards a Characterization of the Structural Determinants of Specificity in the Macrocyclizing Thioesterase for Deoxyerythronolide B Biosynthesis. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (3), 486–497.
- (34) Lu, H.; Tsai, S. C.; Khosla, C.; Cane, D. E. Expression, Site-Directed Mutagenesis, and Steady State Kinetic Analysis of the Terminal Thioesterase Domain of the Methymycin/Picromycin Polyketide Synthase. *Biochemistry* **2002**, *41* (42), 12590–12597.
- (35) Tsai, S. C.; Miercke, L. J. W.; Krucinski, J.; Gokhale, R.; Chen, J. C. H.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. Crystal Structure of the Macrocyclic-Forming Thioesterase Domain of the Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (26), 14808–14813.
- (36) Pinto, A.; Wang, M.; Horsman, M.; Boddy, C. N. 6-Deoxyerythronolide B Synthase Thioesterase-Catalyzed Macrocyclization Is Highly Stereoselective. *Org. Lett.* **2012**, *14* (9), 2278–2281.
- (37) He, W.; Wu, J.; Khosla, C.; Cane, D. E. Macrolactonization to 10-Deoxymethynolide Catalyzed by the Recombinant Thioesterase of the Picromycin/Methymycin Polyketide Synthase. *Bioorg. Med. Chem. Lett.* **2006**, *16* (2), 391–394.
- (38) Wang, M.; Boddy, C. N. Examining the Role of Hydrogen Bonding Interactions in the Substrate Specificity for the Loading Step of Polyketide Synthase Thioesterase Domains. *Biochemistry* **2008**, *47* (45), 11793–11803.
- (39) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. Mechanism and Specificity of the Terminal Thioesterase Domain from the Erythromycin Polyketide Synthase. *Chem. Biol.* **1999**, *6* (2), 117–125.
- (40) Kaneda, M.; Sueyoshi, K.; Teruya, T.; Ohno, H.; Fujii, N.; Oishi, S. Total Synthesis of Odoamide, a Novel Cyclic Depsipeptide, from an Okinawan Marine Cyanobacterium. *Org. Biomol. Chem.* **2016**, *14* (38), 9093–9104.
- (41) Grab, H. A.; Kirsch, V. C.; Sieber, S. A.; Bach, T. Total Synthesis of the Cyclic Depsipeptide Vioprolide D via Its (Z)-Diastereoisomer. *Angew. Chem., Int. Ed.* **2020**, *59* (30), 12357–12361.
- (42) Wang, D.; Song, S.; Tian, Y.; Xu, Y.; Miao, Z.; Zhang, A. Total Synthesis of the Marine Cyclic Depsipeptide Viequeamide A. *J. Nat. Prod.* **2013**, *76* (5), 974–978.
- (43) Chou, T. C.; O'Connor, O. A.; Tong, W. P.; Guan, Y.; Zhang, Z. G.; Stachel, S. J.; Lee, C.; Danishefsky, S. J. The Synthesis, Discovery, and Development of a Highly Promising Class of Microtubule Stabilization Agents: Curative Effects of Desoxyepothilones B and F against Human Tumor Xenografts in Nude Mice. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (14), 8113–8118.
- (44) Blum, W.; Aichholz, R.; Ramstein, P.; Kühnöl, J.; Brügggen, J.; O'Reilly, T.; Flörshheimer, A. In Vivo Metabolism of Epothilone B in Tumor-Bearing Nude Mice: Identification of Three New Epothilone B Metabolites by Capillary High-Pressure Liquid Chromatography/Mass Spectrometry/Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15* (1), 41–49.
- (45) Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. Peptide Cyclization Catalysed by the Thioesterase Domain of Tyrocidine Synthetase. *Nature* **2000**, *407* (6801), 215–218.
- (46) Bruner, S. D.; Weber, T.; Kohli, R. M.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T.; Stubbs, M. T. Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE. *Structure* **2002**, *10* (3), 301–310.
- (47) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry* **2001**, *40* (24), 7092–7098.
- (48) Bu, X.; Wu, X.; Xie, G.; Guo, Z. Synthesis of Tyrocidine A and Its Analogues by Spontaneous Cyclization in Aqueous Solution. *Org. Lett.* **2002**, *4* (17), 2893–2895.
- (49) Kuo, M. C.; Gibbons, W. A. Determination of Individual Side-Chain Conformations, Tertiary Conformations, and Molecular Topography of Tyrocidine A from Scalar Coupling Constants and Chemical Shifts. *Biochemistry* **1979**, *18* (26), 5855–5867.



Total and chemoenzymatic synthesis of the lipodepsipeptide rhizomide A

Corinne A. Lepetit^{a,b}, André R. Paquette^a, Jordan T. Brazeau-Henrie^a, Christopher N. Boddy^{a,*}^a Department of Chemistry and Biomolecular Sciences, Centre for Chemical and Synthetic Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada^b Cergy Paris Université, 5 Mail Gay Lussac, 95000 Cergy, France

ARTICLE INFO

Keywords:

Depsipeptide
Chemoenzymatic synthesis
Total synthesis
Biocatalysis
Peptide chemistry
Enzymatic assays

ABSTRACT

Rhizomides are a family of depsipeptide macrolactones synthesized by a non-ribosomal peptide synthetase (NRPS) encoded in the genome of *Paraburkholderia rhizoxinica* str. HKI 454. In this study, the total and chemoenzymatic synthesis of the depsipeptide rhizomide A is described. Rhizomide A was generated through macrolactamization while the linear C-terminal *N*-acetylcysteamine (SNAC) thioester substrate was synthesized through a C-terminal thioesterification strategy. It was shown that the rhizomide A thioesterase (RzmA-TE) is an active macrocyclization catalyst, allowing the chemoenzymatic synthesis of rhizomide A. This work further showcases the biocatalytic power of TEs in accessing complex macrocyclic natural products.

Rhizomides (Rzms) are a class of cyclic lipo-depsipeptides encoded in the genome of the β -proteobacterium *Paraburkholderia rhizoxinica* str. HKI 454.^{1,2} The strain-specific megaplasmid pBRH01 contains the nonribosomal peptide biosynthetic gene cluster (BGC) responsible for the biosynthesis of the rhizomides,^{1–3} however, under typical cultivation conditions this BGC is silent.¹ Installing the constitutive *apr* promoter along with the apramycin resistance gene directly upstream of the BGC in the native host provided a 70-fold increase in rhizomide production, enabling its isolation and structure determination.¹ The rhizomides show weak antitumor, antifungal, and antibacterial activity, suggesting they may have a highly strain specific growth inhibitory activity or play a role in intra- or inter-species communication similar to other cyclic depsipeptides.⁴

The rhizomides are biosynthesized by a single nonribosomal peptide synthetase (NRPS) complex (RzmA) that contains 7 modules. It begins with a starter condensation (C_S) domain responsible for the condensation of acetyl-CoA to a thiolation domain (T) bound L-Leu, yielding the peptidyl-enzyme intermediate *N*-acetyl-L-Leu-T.^{5–7} Chain extension then proceeds with the incorporation of L-Thr, D-Tyr, D-Ala, L-Ala, D-Ala and L-Val giving the linear rhizomide A precursor (Fig. 1). Based on the structures of rhizomide A-C, the adenylation (A) domains of modules 5 and 6 are likely promiscuous and can activate the non-native substrates L-Ser, to yield the minor products rhizomide B and C (Fig. 1).¹ The linear peptide chain is then passed to a terminal thioesterase (TE) domain,⁸ which is hypothesized to be responsible for the cyclization and release of rhizomides A-C through macrolactonization of the L-Thr₂ side-chain alcohol with the backbone carboxylate of L-Val₇ (Fig. 1).

The synthesis of depsipeptides however can be challenging.^{9–11} While synthesis of the linear peptidyl chain is typically straight forward, macrolactonization has been a longstanding synthetic challenge. Effective macrolactonization reactions, such as the Corey-Nicolau and Yamaguchi esterifications routinely struggle to generate cyclic depsipeptides^{12–14} or lead to epimerization of the C-terminal α -carbon generating diastereomeric products.^{15–17} To address this problem, alternative sites of ring closure via macrolactamization can be used, however, this requires more elaborate branched precursor synthesis.^{18–20} The TEs responsible for macrolactonizing natural products during their biosynthesis circumvents these synthetic challenges. We have successfully harnessed TEs as biocatalysts for the chemoenzymatic synthesis of natural products and new to nature compounds,^{20–23} highlighting the power of these enzymes as macrocyclization biocatalysts.

Herein we characterize the ability of the rhizomide TE (RzmA-TE) to macrocyclize the linear *N*-acetylcysteamine (SNAC) derivative of rhizomide A, **1**. We show the recombinant purified RzmA-TE effectively forms the natural product **2**, completing the first chemoenzymatic synthesis through macrolactonization of the linear SNAC peptide. We also synthesized the macrocyclic authentic standard, **2**, successfully completing the first total synthesis via macrolactamization.

Characterization of the ability of the excised TE from RzmA to macrocyclize a linear substrate into rhizomide A required synthesis of both the linear substrate **1** and authentic standard of the natural product, **2**. While linear peptides with C-terminal carboxylates or carboxamides are routinely synthesized via standard Fmoc solid phase peptide

* Corresponding author.

E-mail address: cboddy@uottawa.ca (C.N. Boddy).<https://doi.org/10.1016/j.bmcl.2023.129506>

Received 31 July 2023; Received in revised form 27 September 2023; Accepted 7 October 2023

Available online 12 October 2023

0960-894X/© 2023 Elsevier Ltd. All rights reserved.

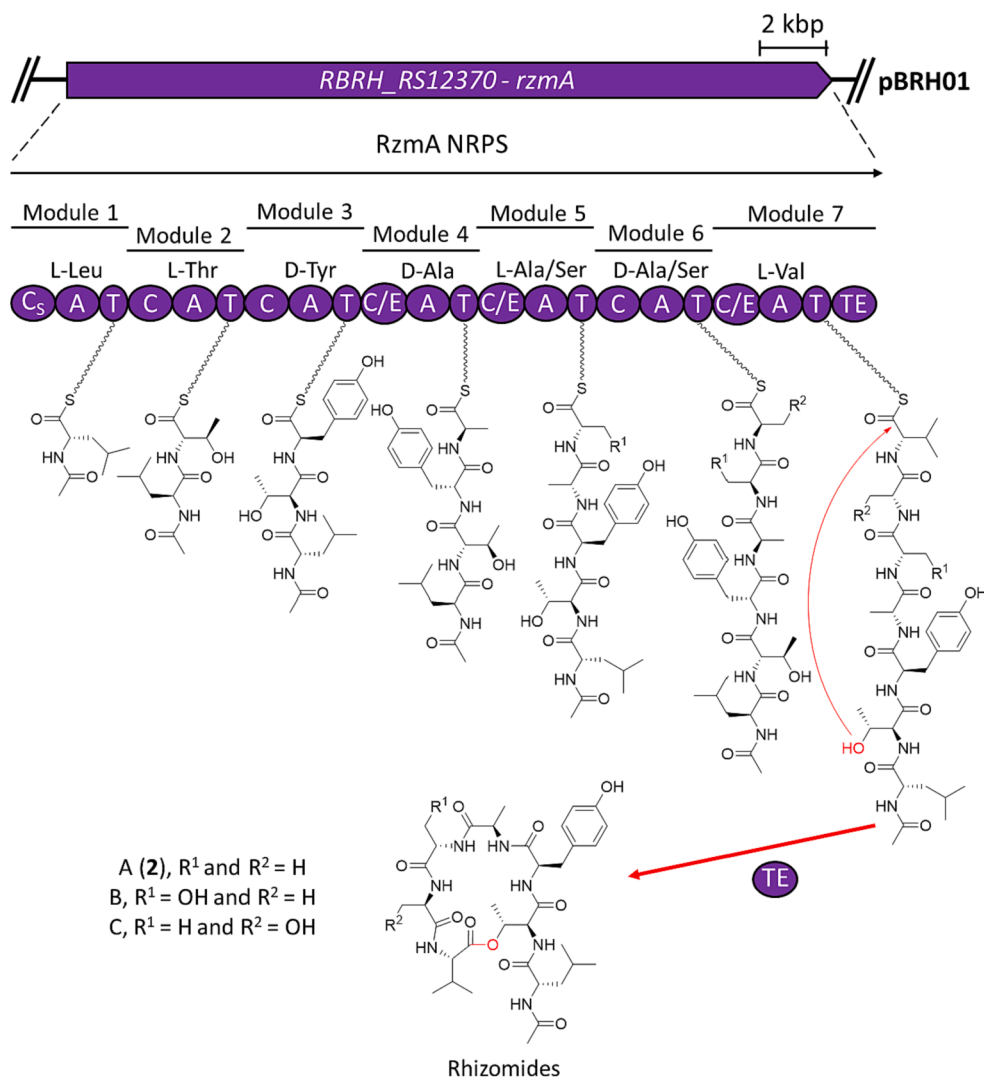


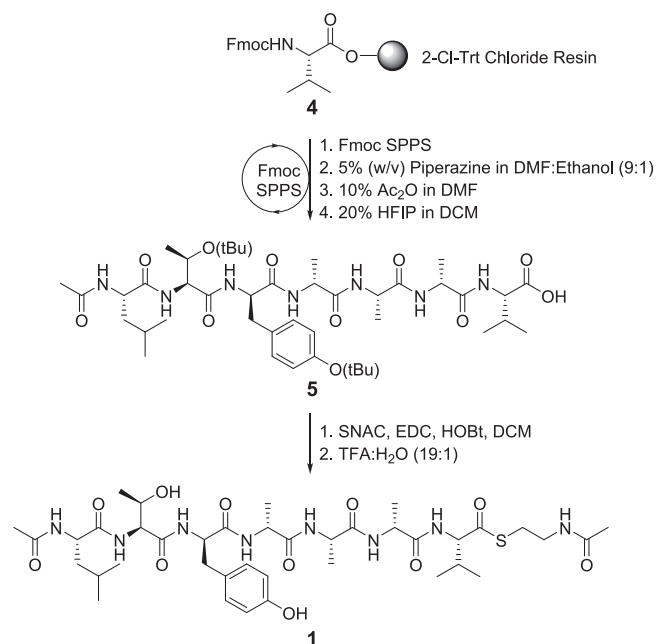
Fig. 1. Biosynthesis of rhizomides A-C (Rzm). The single reading frame **RBRH_RS12370** (*rzmA*) encodes for the singular NRPS complex, RzmA. RzmA contains 7 modules, with A domain selectivity consistent with the biosynthesis of rhizomides A-C. Dual C/E domains are appropriately positioned (modules 4, 5 and 7) for conversion of loaded L-amino acids to their D-amino acid enantiomers. Module 7 encodes a C-terminal TE domain responsible for macrocyclic release of the final natural products. No resistance, transport and tailoring genes have been identified beyond the cluster. Abbreviations; C_s, starter condensation domain; C, condensation domain; A, adenylation domain; T, thiolation or peptidyl carrier domain; C/E, dual epimerization-condensation domain; TE, thioesterase domain.

synthesis (SPPS), C-terminal thioester peptides typically require specialized strategies.^{24–26} To access **1** (Scheme 1), Fmoc SPPS using HBTU/DIEA conditions was performed on the hypersensitive 2-Cl-Trt chloride resin and the full length, *N*-acetylated linear peptide was cleaved from the resin under mildly acidic conditions²⁰ to generate the protected peptide with a free carboxylic acid at the C-terminus, **5**. Solution phase chemistry was then used to generate the *N*-acetyl cysteamine (SNAC) thioester peptide. Lastly removal of the sidechain *t*Bu protecting groups using TFA followed by HPLC purification afforded substrate **1** (5 % yield from **5**) (Scheme 2).

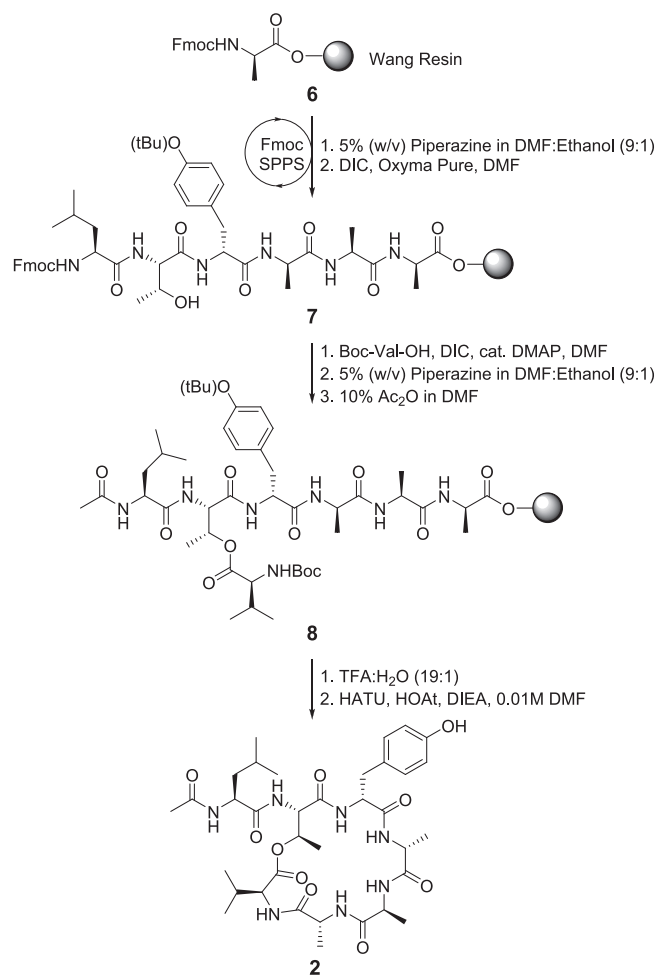
As macrolactonization of cyclic depsipeptides is typically challenging, an authentic standard of rhizomide A, **2**, was synthesized via macrolactamization between L-Val₇ and D-Ala₆. Branched peptide **7** was synthesized using SPPS on Wang resin, incorporating unprotected sidechain Fmoc-Thr-OH. On-resin esterification of L-Thr with Boc-Val-OH under standard DIC/DMAP conditions afforded the ester in quantitative conversion based on LCMS analysis of microcleavages. Subsequent *N*-terminus Fmoc deprotection and acetylation afforded peptide **8**, which was then cleaved off resin and fully deprotected. The crude peptide ester underwent macrolactamization under high dilution with HATU/HOAt coupling conditions to afford **2** following RP-HPLC

purification (30 % yield from **8**). Comparison of our synthetic material with the reported isolation NMR spectra confirmed that our product was indistinguishable from rhizomide A (Fig. S21, Table S3), providing a total synthesis of rhizomide A and unambiguous confirmation of its chemical structure.

To test the macrocyclization capability of the TE domain from RzmA, we expressed and purified the TE domain as a stand-alone protein (RzmA-TE). A codon optimized synthetic gene construct encoding the C-terminal TE domain from *rzmA* (Fig. S1) was cloned into a pET-28a(+) expression vector under an inducible T7 promoter and fused to an N-terminal hexa-His tag. RzmA-TE was overexpressed in *Escherichia coli* BL21(DE3) and purified by metal affinity chromatography to high purity (>95 %, Fig. S2). To assay for macrocyclization activity, RzmA-TE (25 μ M) was incubated with **1** (1 mM) in 50 mM of phosphate buffer (pH 7.4) and analyzed by LCMS for formation of **2**. After 3 h of incubation at room temperature, **2** was detected (4.7 %, Fig. 2D and S4) and by 24 h 15 % of **1** was converted to **2** (Fig. 2C and S5). Both the thermally denatured enzyme and substrate only controls showed no formation of **2** within 24 h (Fig. 2E and S3). These data are consistent with a TE-dependent rate enhancement for macrocyclization, albeit a modest enhancement compared to other characterized NRPS TEs such as for



Scheme 1. Synthesis of SNAC substrate 1.



Scheme 2. Synthesis of rhizomide A, 2.

tyrocidine and seongsanamide.^{20,27} In addition, no other products, such as hydrolysis, 3, or higher order oligomers, were detected at levels greater than the boiled enzyme control, demonstrating the selectivity of this macrocyclization biocatalyst. These data show that the RzmA-TE is an active macrocyclization biocatalyst and demonstrate the first chemoenzymatic total synthesis of rhizomide A.

Kinetic characterization of the RzmA-TE with 1 was attempted. Ellman's reagent (DTNB), which tracks the release of free thiol during loading of 1 on to the enzyme active site Ser, was included in the assay to enable spectrophotometric quantification of TE kinetics.²⁸ Unfortunately, addition of DTNB into reaction mixtures lead to rapid and irreversible RzmA-TE inactivation. As DTNB reacts with thiols and RzmA-TE has six Cys residues (Fig. S1), we hypothesize that the DTNB forms adducts with one or more of these Cys and inactivates the TE. As such, kinetic characterization could not be performed, however, qualitative assessment of the extent of product formation in the product distribution assays point towards RzmA-TE being kinetically slow compared to other characterized depsipeptide forming TEs where a typical specificity constant (k_{cat}/K_M) is approximately 1–5 mM⁻¹min⁻¹.^{20,29,30} This is not uncommon however as the calcium dependent antibiotic forming and the cereulide forming TE domains are both substantially slower ($k_{cat}/K_M \sim 0.1$ mM⁻¹min⁻¹ and 0.02 mM⁻¹min⁻¹, respectively).^{22,31}

Macrocyclization is a chemically challenging reaction due to the large entropic barrier associated with end-to-end contact.^{32,33} In addition, it is made more difficult with alcohol nucleophiles, as these require activation to become sufficiently nucleophilic for cyclization, particularly in an aqueous environment where they must outcompete water. Yet, despite the large entropic barrier and the vast excess of water, only macrocycle 2 is formed by the RzmA-TE, highlighting the power of TEs as specific macrocyclization catalysts.

In addition to the RzmA-TE, eight other macrolactonizing TEs have been used in the chemoenzymatic synthesis of depsipeptides. These TEs are from the teixobactin (Txo),³⁴ seongsanamide,²⁰ streptogramin (SnbDE),²⁹ surfactin (Srf),³⁰ lysobactin (LydB),³⁵ calcium dependent antibiotic,³¹ valinomycin,²¹ and cereulide²² biosynthetic pathways. Like the RzmA-TE, most TEs can accept SNAC substrates, though the LydB-TE requires the more reactive thiophenol thioester.³⁵ Interestingly some are also able to process ester substrates such as the Txo-³⁴ and the Srf-TE.³⁰ These nine TEs generate 4-,³⁴ 5-,²⁰ 6-,²⁹ 8-,³⁰ 9-,³⁵ 10-,³¹ and 12-residue^{21,22} depsipeptide products. While both the SnbDE-TE and RzmA-TE generate 6-residue macrolactones, only the RzmA-TE catalyzes macrocyclization without substantial hydrolysis of the substrate. This highlights the benefit of characterizing multiple TE homologs for macrolactonization as their inherent reactivity's differ.

In summary, we have completed the first total synthesis of rhizomide A, 2, through a macrolactamization strategy from a branched SPPS derived peptide. In addition, rhizomide A was synthesized chemoenzymatically via macrolactonization of the thioester activated precursor 1 with the excised RzmA-TE. This work unambiguously confirms the structure of rhizomide A and adds to the growing body of work demonstrating the utility of TEs as biocatalysts to access complex macrocyclic natural products.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Christopher N. Boddy reports financial support was provided by Natural Sciences and Engineering Research Council of Canada.

Data availability

Data will be made available on request.

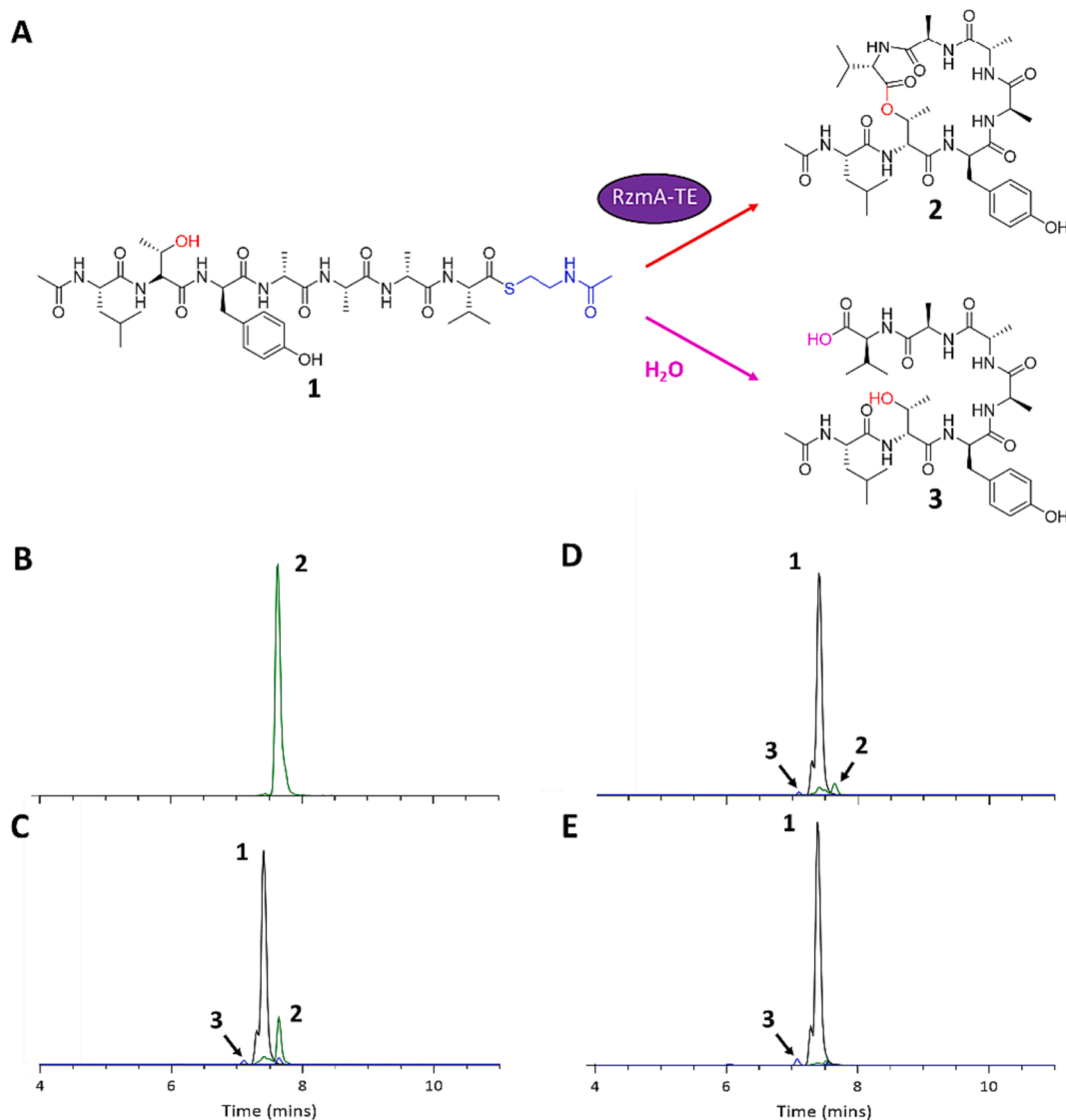


Fig. 2. Product distribution assays of RzmA-TE. (A) Schematic representation of 1 macrocyclization by RzmA-TE to form 2. (B) Macrocycle 2 (1 mM) only control in 50 mM of phosphate buffer (pH 7.4). (C) RzmA-TE (25 μ M) with 1 mM of 1 after 24 h of incubation showing 15 % conversion into 2 (Table S2). (D) RzmA-TE (25 μ M) with 1 mM of 1 after 3 h of incubation showing 4.7 % conversion into 2 (Table S2). (E) Substrate 1 only control after 24 h of incubation. All substrate and thermally denatured enzyme controls showed similar results (Fig. S5). All traces show ion extractions for the $[M + H]^+$ ions of the substrate 1 ($m/z = 852$, black), rhizomide A 2 ($m/z = 733$, green), and the hydrolysis product 3 ($m/z = 750$, blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bmcl.2023.129506>.

References

- Wang X, Zhou H, Chen H, et al. Discovery of Recombinases Enables Genome Mining of Cryptic Biosynthetic Gene Clusters in Burkholderia Species. *PNAS*. 2018;115. <https://doi.org/10.1073/pnas.1720941115>.
- Lackner G, Moebius N, Partida-Martinez LP, Boland S, Hertweck C. Evolution of an Endofungal Lifestyle: Deductions from the Burkholderia Rhizoxinica Genome. *BMC Genom*. 2011;12. <https://doi.org/10.1186/1471-2164-12-210>.
- Lackner G, Moebius N, Partida-Martinez L, Hertweck C. Complete Genome Sequence of Burkholderia Rhizoxinica, an Endosymbiont of Rhizopus Microsporus. *J Bacteriol*. 2011;193:783–784. <https://doi.org/10.1128/JB.01318-10>.
- Alonzo DA, Schmeing TM. Biosynthesis of Depsipeptides, or Depsi: The Peptides with Varied Generations. *Protein Sci*. 2020;29:2316–2347. <https://doi.org/10.1002/pro.3979>.
- Chooi YH, Tang Y. Adding the Lipo to Lipopeptides: Do More with Less. *Chem Biol*. 2010;17:791–793. <https://doi.org/10.1016/j.chembiol.2010.08.001>.
- Zhong L, Diao X, Zhang N, et al. Engineering and Elucidation of the Lipoinitiation Process in Nonribosomal Peptide Biosynthesis. *Nat Commun*. 2021;12:296. <https://doi.org/10.1038/s41467-020-20548-8>.
- Imker HJ, Krahn D, Clerc J, Kaiser M, Walsh CT. N-Acylation during Glidobactin Biosynthesis by the Tridomain Nonribosomal Peptide Synthetase Module GIBF. *Chem Biol*. 2010;17:1077–1083. <https://doi.org/10.1016/j.chembiol.2010.08.007>.
- Horsman ME, Hari TPAA, Boddy CN. Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate? *Nat Prod Rep*. 2016;33:183–202. <https://doi.org/10.1039/c4np00148f>.
- Shabani S, Hutton CA. Depsipeptide Synthesis Using a Late-Stage Ag(I)-Promoted Macrolactonisation of Peptide Thioamides. *Chem Commun*. 2021;57:2081–2084. <https://doi.org/10.1039/d0cc07747j>.
- Kaur H, Harris PWR, Little PJ, Brimble MA. Total Synthesis of the Cyclic Depsipeptide YM-280193, a Platelet Aggregation Inhibitor. *Org Lett*. 2015;17:492–495. <https://doi.org/10.1021/ol503507g>.
- Hawkins PME, Tran W, Nagalingam G, et al. Total Synthesis and Antimycobacterial Activity of Ohmyungsamycin A, Deoxyecumicin, and Ecumicin. *Chem – A Eur J*. 2020;26:15200–15205. <https://doi.org/10.1002/chem.202002408>.
- Shabani S, Hutton CA. Total Synthesis of Seongsanamide B. *Org Lett*. 2020;22:4557–4561. <https://doi.org/10.1021/acs.orglett.0c01642>.
- Han F, Guo Y, Ye T. Total Synthesis of Antiallergic Bicyclic Peptide Seongsanamide A. *Org Chem Front*. 2020;7:1658–1662. <https://doi.org/10.1039/d0qo00531b>.
- Yudin AK. Macrocycles: Lessons from the Distant Past, Recent Developments, and Future Directions. *Chem Sci*. 2015;6:30–49. <https://doi.org/10.1039/c4sc03089c>.

- 15 Ghosh AK, Xu CX. A Convergent Synthesis of the Proposed Structure of Antitumor Depsipeptide Stereocalpin A. *Org Lett*. 2009;11:1963–1966. <https://doi.org/10.1021/ol900412u>.
- 16 Cochran JR, McErlean CSP, Jolliffe KA. Total Synthesis and Assignment of the Side Chain Stereochemistry of LI-F04a: An Antimicrobial Cyclic Depsipeptide. *Org Lett*. 2010;12:3394–3397. <https://doi.org/10.1021/ol101254m>.
- 17 Kopple KD. Synthesis of Cyclic Peptides. *J Pharm Sci*. 1972;61:1345–1356. <https://doi.org/10.1002/jps.2600610902>.
- 18 Kuranaga T, Enomoto A, Tan H, Fujita K, Wakimoto T. Total Synthesis of Theonellapeptolide Id. *Org Lett*. 2017;19:1366–1369. <https://doi.org/10.1021/acs.orglett.7b00249>.
- 19 Bando Y, Hou Y, Seyfarth L, et al. Total Synthesis and Structure Correction of the Cyclic Lipodepsipeptide Orfamide A. *Chem – A Eur J*. 2022;28:2–6. <https://doi.org/10.1002/chem.202104417>.
- 20 Brazeau-Henrie JT, Paquette AR, O'Rourke AQ, et al. Total and Chemoenzymatic Synthesis of Seongsanamide E. *Org Lett*. 2022;24:6369–6373. <https://doi.org/10.1021/acs.orglett.2c02271>.
- 21 Huguénin-Dezot N, Alonzo DA, Heberlig GW, et al. Trapping Biosynthetic Acyl-Enzyme Intermediates with Encoded 2,3-Diaminopropionic Acid. *Nature*. 2019;565:112–117. <https://doi.org/10.1038/s41586-018-0781-z>.
- 22 Heberlig GW, Boddy CN. Thioesterase from Cereulide Biosynthesis Is Responsible for Oligomerization and Macrocyclization of a Linear Tetradepsipeptide. *J Nat Prod*. 2020;83:1990–1997. <https://doi.org/10.1021/acs.jnatprod.0c00333>.
- 23 Heberlig GW, Brown JTC, Simard RD, et al. Chemoenzymatic Macrocyclic Synthesis Using Resorcylic Acid Lactone Thioesterase Domains. *Org Biomol Chem*. 2018;16:5771–5779. <https://doi.org/10.1039/c8ob01512k>.
- 24 Ingenito R, Bianchi E, Fattori D, Pessi A. Solid Phase Synthesis of Peptide C-Terminal Thioesters by Fmoc/t-Bu Chemistry. *J Am Chem Soc*. 1999;121:11369–11374. <https://doi.org/10.1021/ja992668n>.
- 25 Zheng JS, Chang HN, Wang FL, Liu L. Fmoc Synthesis of Peptide Thioesters without Post-Chain-Assembly Manipulation. *J Am Chem Soc*. 2011;133:11080–11083. <https://doi.org/10.1021/ja204088a>.
- 26 Blanco-Canosa JB, Dawson PE. An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation. *Angew Chemie – Int Ed*. 2008;47:6851–6855. <https://doi.org/10.1002/anie.200705471>.
- 27 Trauger JW, Kohli RM, Walsh CT. Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry*. 2001;40:7092–7098. <https://doi.org/10.1021/bi010035r>.
- 28 Brazeau-Henrie JT, Paquette AR, Boddy CN. In Vitro Biochemical Characterization of Excised Macrocyclizing Thioesterase Domains from Non-Ribosomal Peptide Synthetases. *Non-Ribosomal Peptide Biosynthesis and Engineering*. 2023;2670:101–125.
- 29 Mahler C, Sieber SA, Grünwald J, Marahiel MA. Chemoenzymatic Approach to Enantiopure Streptogramin B Variants: Characterization of Stereoselective Pristinamycin I Cyclase from *Streptomyces Pristinaespiralis*. *J Am Chem Soc*. 2005;127:9571–9580. <https://doi.org/10.1021/ja051254t>.
- 30 Tseng CC, Bruner SD, Kohli RM, Marahiel MA, Walsh CT, Sieber SA. Characterization of the Surfactin Synthetase C-Terminal Thioesterase Domain as a Cyclic Depsipeptide Synthase. *Biochemistry*. 2002;41:13350–13359. <https://doi.org/10.1021/bi026592a>.
- 31 Grünwald J, Sieber SA, Marahiel MA. Chemo- and Regioselective Peptide Cyclization Triggered by the N-Terminal Fatty Acid Chain Length: The Recombinant Cyclase of the Calcium-Dependent Antibiotic from *Streptomyces Coelicolor*. *Biochemistry*. 2004;43:2915–2925. <https://doi.org/10.1021/bi036140d>.
- 32 Parenty A, Moreau X, Campagne JM. Macrolactonizations in the Total Synthesis of Natural Products. *Chem Rev*. 2006;106:911–939. <https://doi.org/10.1021/cr0301402>.
- 33 Mortison JD, Sherman DH. Frontiers and Opportunities in Chemoenzymatic Synthesis. *J Org Chem*. 2010;75:7041–7051. <https://doi.org/10.1021/jo101124n>.
- 34 Mandalapu D, Ji X, Chen J, et al. Thioesterase-Mediated Synthesis of Teixobactin Analogues: Mechanism and Substrate Specificity. *J Org Chem*. 2018;83:7271–7275. <https://doi.org/10.1021/acs.joc.7b02462>.
- 35 Hou J, Robbel L, Marahiel MA. Identification and Characterization of the Lysobactin Biosynthetic Gene Cluster Reveals Mechanistic Insights into an Unusual Termination Module Architecture. *Chem Biol*. 2011;18:655–664. <https://doi.org/10.1016/j.chembiol.2011.02.012>.