The Dissertation Committee for Shawn Kristen Piasecki Certifies that this is the approved version of the following dissertation:

INVESTIGATIONS INTO THE BIOCATALYTIC POTENTIAL OF MODULAR POLYKETIDE SYNTHASE KETOREDUCTASES.

Committee:

Adrian Keatinge-Clay, Supervisor

Walter Fast

David Hoffman

Hung-wen Liu

Dionicio Siegel

Investigations into the Biocatalytic Potential of Modular Polyketide Synthase Ketoreductases.

by

Shawn Kristen Piasecki, B.S. Chem.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2013

Dedication

To my family, D.S.P., S.J.P., K.L.P., and M.K.P., for their never-ending support of my busy life for the past five years. I never would have achieved this goal if it weren't for your constant encouragement, and most importantly, love. To my grandparents, that although you aren't here to see me reach this milestone, I'm always thinking of you and wish that you could have been here to experience this with me. Lastly, to my best friend and adopted family member, Otis: you have voyaged through my entire undergraduate and graduate career with me, never complaining when I would come home late from school or not have enough time to take you to the park; but you would always greet me with your smiling pug face, cuddly fat body, and unconditional love. So to my entire family for helping me achieve this goal, I dedicate this dissertation.

Acknowledgements

There have been numerous people who have influenced my life throughout my college career, all contributing in some way to getting me where I am today. First of all, if it weren't for my parents I would not even be here today; so thank you mom and dad for deciding to bring me into this world. My family and friends along the way have been wonderfully tolerant of my busy life throughout college, and for that I am eternally grateful. A decade-younger version of me would not believe that I would now be finishing my PhD. I entered into my undergraduate studies as an electrical engineering/dance major, after all.

During my undergraduate studies, multiple people had influenced me to decide on furthering my studies in graduate school. It did not click until I joined Dr. Andrew Ellington's lab. The main person I have to thank for this is Dr. John F. Stanton. If it weren't for his mentorship during my time in his physical chemistry class and recommendation to Dr. Ellington, I would not have had that research opportunity to shape my scientific interests to what they are today. Also I must thank Dr. Andy Ellington for taking a chance on a chemistry undergraduate student interested in biochemical research, and also for rescuing me (literally) from waiting tables at Chili's; for both, I am much appreciative. I will always have fond memories of my time in his lab, and forever treasure my "Andy sayings" quote book. Most of all, I must thank Dr. Bradley Hall for his mentorship in the Ellington lab. He definitely made my time there entertaining, and although as an undergraduate I heard how horrible graduate school was, he was always happy when it came to research and teaching students. He taught me so much of what I know today, and for that I am forever thankful.

Once I had been accepted to the ICMB PhD program, I desired research concerning chemical biology, yet it was not until I went to a seminar in the department given by an incoming faculty member that I decided exactly what my scientific passion was. Of course the incoming faculty was none other than Dr. Adrian Keatinge-Clay, and he shared his research interests on polyketide synthases. I was fascinated by the use of these biocatalysts for the synthesis of new medicines. Of course I asked him afterward to rotate in his lab, and later joined his group. Studying natural product catalysis and synthetic manipulation were hence forth my main research interests. I quickly learned that Adrian succeeds at pretty much anything he tries. Chess? He's a master. Science? He motivated our lab to publish over 15 papers in his first 5 years as a tenure-track professor. Oh, Adrian wants to join a running club with me? Guess what- he's faster than me, even though I run frequently. By far the thing he is best at is being a graduate student mentor. For the past 5 years, he has seen most of his graduate students break down in tears over stress or life. He has always been patient and compassionate during those hard times and helped us get through them; yet was also always there to jump for joy when we "won science" that day with a great result, passed preliminary exams, or had a paper accepted. His research and personal mentorship were vital to my success in graduate school, and will always be cherished. I plan to continually stay in contact with him- even if he gets sick of hearing from me.

Finally to my lab: if it weren't for you, I would have gone one of two ways: 1) crazy from not having any friends, or 2) gotten a lot more work done. Probably the latter is true. All joking aside, you were always there for me when I needed to vent/cry/yell/laugh, and I hope that I reciprocated the same to all of you. J.T.Z.- thanks for being my KR buddy! I enjoyed our combined efforts on projects, and I think Adrian did as well. A.D.H.- thank you for teaching me everything you know about the internals of an HPLC. You truly are a

walking scientific dictionary. D.C.G.- I'm happy that I had you to laugh with at all of our inside jokes, and my extra food in lab was happy that you were always hungry. C.D.F.- I'm so grateful that you were my desk neighbor in lab. Thanks for always listening to my constant re-coloring mouse clicks ("…shopping!") in Photoshop, and I will surely miss our vent-session lunches. E.A.I-M.- I'm sad that I was not able to spend my last summer in lab with you, but I'm so happy for you and your new family! Congratulations on such a beautiful baby and I hope to see you soon for a final "5 o'clock dance party." C.B.B.- thanks for always letting me and A.J.H. pick on you. We have to make sure that the "dominant females of the AKC lab" live on when we leave, so we're really just toughening you up because we like you. D.T.W.- your personality and unwillingness to give up is not only amusing, but also perfect to carry on the personality of the lab. I'm glad that our paths crossed, even for this short amount of time. A.J.H.- you were certainly my partner in all things fun and scientific for the past 5 years, and I will always remember the time we decided we would be best lab friends ("turn it off! …I can't seeeeee!"). It's sad to think we won't work together next, but I hope to again someday. Finally to D.I.L.- even though you weren't in the lab, we did everything else together. I'm so glad that we became best friends, neighbors, workout buddies, and dog moms! I hope we both continue our careers near each other.

Of course all this would not be possible without the help of my committee members. To Dr.'s Ben Liu, Dave Hoffman, Dio Siegel, and Walt Fast: countless thanks to all of you for your mentorship throughout my graduate career. Without your support, my research would not have been as successful. Thank you for everything.

Investigations into the Biocatalytic Potential of Modular Polyketide Synthase Ketoreductases.

Shawn Kristen Piasecki, PhD

The University of Texas at Austin, 2013

Supervisor: Adrian Keatinge-Clay

The production of new drugs as potential pharmaceutical targets is arguably one of the most important avenues of medicine, as existing diseases not only require treatment, but it is also certain that new diseases will appear in the future which will need treatment. Indeed, existing medicines such as antibiotics and immunosuppressants maintain their current activities in their respective realms, yet the molecular and stereochemical complexity of these compounds cause a burden on organic synthetic chemists that may prohibit the high yields required to manufacture a drug. The enzyme systems that naturally manufacture these compounds are incredibly efficient in doing so, and also do not use environmentally harmful solvents, chiral auxiliaries, or metals that are utilized in the current syntheses of these compounds; therefore utilizing these enzymes' machinery for the biocatalysis of new medicinally-relevant compounds, as researchers have in the past, is undeniably a rewarding endeavor.

In order to harness these systems' biocatalytic potential, we must understand the processes which they operate. This work focuses on ketoreductase domains, since they are responsible for setting most of the stereocenters found within these complex secondary metabolites. We have supplied a library of substrates to multiple ketoreductases to test their

limits of stereospecificity and found that, for the most part, they maintain their natural product stereospecificity seen in nature. We were even able to convert a previously non-stereospecific ketoreductase to a stereospecific catalyst. We have also developed a new technique to follow ketoreductase catalysis in real-time, which can also differentiate between which diastereomeric product is being produced. Finally, we have elucidated the structure of a ketoreductase that reduces non-canonically at the α- and β- position, and functionally characterized its activities on shortened substrate analogs. With the knowledge gained from this dissertation we hope that the use of ketoreductases as biocatalysts in the biosynthesis of new natural product-based medicines is a much nearer reality than before.

Table of Contents

List of Tables

Chapter 2

Chapter 4

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 1. Introduction

POLYKETIDE SYNTHASES AS NATURAL PRODUCT MACHINES

Modular polyketide synthases (PKSs) are enzymatic assembly lines that control the substitution and stereochemistry of complex polyketides in the biosyntheses of such compounds as the antibacterial erythromycin A, the antifungal amphotericin B, and the immunosuppressant rapamycin (Staunton and Weissman, 2001; Sherman and Smith, 2006; Khosla et al., 2007; Smith and Tsai, 2007) (**Figure 1.1**). Many synthetic organic chemists seek to emulate the reactions they catalyze and obtain the products they yield; however, if PKS enzymes themselves were harnessed as biocatalysts, coveted chiral building blocks and biologically active molecules would be more readily accessed.

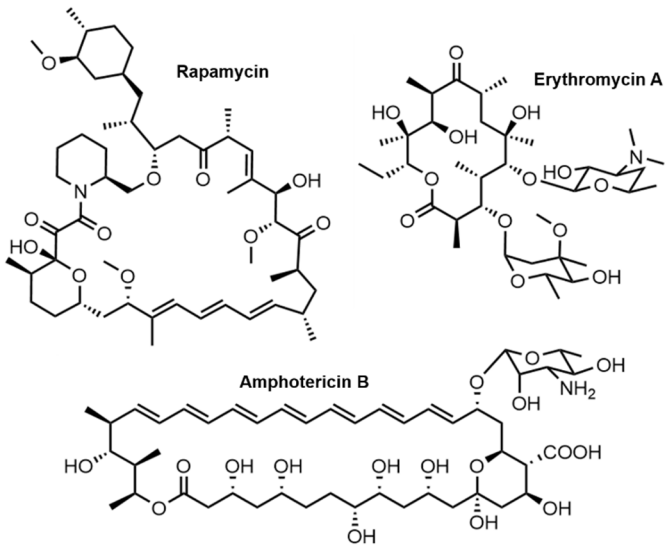

Figure 1.1. Polyketide synthases produce complex molecules with diverse functions and structures.

Libraries of chiral building blocks are highly desired by synthetic organic chemists to accelerate the synthesis of natural products and pharmaceuticals (Hilterhaus and Liese, 2007; Patel, 2008). Such compounds would be particularly helpful in the syntheses of complex polyketides because the required synthetic fragments often harbor multiple stereocenters and are challenging to prepare. The complexity of synthesizing reduced polyketides can be appreciated by examining the synthetic routes established for 6-deoxyerythronolide B (6-

dEB), the precursor of erythromycin antibiotics (Masamune et al., 1981; Myles et al., 1990; Evans et al., 1998; Crimmins and Slade, 2006; Stang and White, 2009; Gao et al., 2013). Although several groups over the last 32 years have worked on the synthesis of this archetypal macrolide core, the shortest is 14 steps with an overall yield of ~5% (Gao et al., 2013).

The soil bacterium *Saccharopolyspora erythraea* synthesizes 6-dEB with DEBS (6-dEB synthase), a PKS that contains diverse catalytic activities: acyltransferases (ATs) catalyze the transesterification of small organic extender units from acyl-CoAs onto the PKS, ketosynthases (KSs) catalyze decarboxylative condensations that create carbon-carbon bonds between these extender units and the growing polyketide chain, ketoreductases (KRs) catalyze stereospecific, NADPH-coupled reductions of β-keto groups formed through condensation, a dehydratase (DH) catalyzes an elimination reaction to yield an olefin, and an enoylreductase (ER) catalyzes the stereospecific, NADPH-coupled reduction of that olefin (Khosla et al., 2007) (**Figure 1.2**). Growing polyketides are passed along the PKS assembly line through *trans*-thioesterification reactions between acyl carrier proteins (ACPs) and KSs, and lastly to a thioesterase (TE) that catalyzes the cyclization of a heptaketide to yield 6-dEB. Tailoring enzymes then decorate 6-dEB with sugar moieties and hydroxyl groups to produce the bioactive antibiotic erythromycin. Thus, enzymes in this PKS cooperate to transform seven molecules of propionate into one molecule of 6-dEB under ambient, aqueous conditions. As the current syntheses of 6-dEB and related natural products rely on organic solvents, chiral auxiliaries, and metals that can be both expensive and environmentally harmful, employing PKS enzymes in such syntheses would be advantageous.

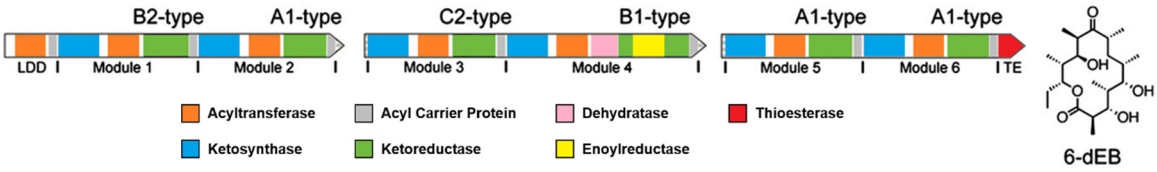

Figure 1.2. The DEBS PKS. This assembly line synthesizes 6-dEB; 9/10 stereocenters are set by KRs that are classified by the combination of stereocenters they produce (marked above their locations). LDD, loading didomain.

ENZYMOLOGY STUDIES OF DISCRETE PKS DOMAINS

Because the presence of specific domains within a modular PKS correspond to the final product of that PKS, and modules are used in the order listed, it is simple to predict products from gene organization (Cane et al., 1998; Katz, 2009). The co-linear nature of these biosynthetic machines allow for the facile synthesis of novel natural products. Utilizing PKS domains as biocatalysts for the production of new natural products is not an untouched field; in fact, experiments attempting to alter function have been in existence for over 2 decades (Strohl and Connors, 1992). The earliest attempts were to swap domains from one system for another of desired function. One study, for example, was successful in generating a hybrid module that accepts acetate starter units in place of propionate units (Oliynyk et al., 1996). In another, EryKR2 (which produces an L-hydroxy product) was replaced with RapKR2, from the rapamycin PKS pathway (which normally produces a D-hydroxy product), and the expected exchanged stereochemistry was found (Kao et al., 1998). Although other attempts were successful, some suffered in yield of product formed (Hans et al., 2003; Kellenberger et al., 2012). It is widely accepted that intramodular domain architectures are specific for each module, and that domain swapping possibly interrupts these important contacts (Tsuji et al., 2001; Kumar et al., 2003).

Another route toward modifying PKS systems to produce new natural products uses directed evolution (Dunn and Khosla, 2013). In one study, the substrate preference of EryAT6 was altered to accept an alkyne extender unit instead of its natural (2*S*)-methylmalonyl-CoA (Sundermann et al., 2013).

The most recent approach is instead to mutate residues within a domain in order to alter enzymatic function (Del Vecchio et al., 2003; Baerga-Ortiz et al., 2006; O'Hare et al., 2006; Kushnir et al., 2012). For example, we were able to transform an A1-type KR into an A2-type KR when tested in vitro (Zheng et al., 2010; Zheng et al., 2013). However, when the same mutations are tested in vivo, the enzymes don't often yield the same result (Kwan et al., 2011). Other possible mutation sites may be required to realize the full potential of mutagenesis studies. Most studies mentioned above used *N*-acetylcysteamine (*S*-NAC) thioester analogs in place of the ACP's ~18 Å phosphopantetheinyl arm which naturally delivers substrates (**Figure 1.3**), and has since proven to be an acceptable surrogate for many PKS enzymology studies.

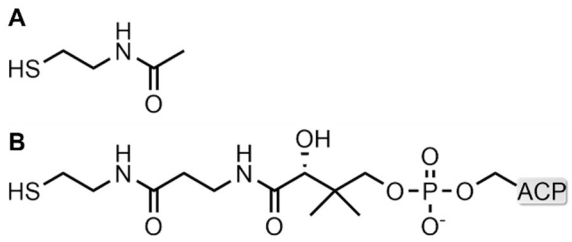

Figure 1.3. (A) The *S*-NAC arm (shown as a free thiol) is a shortened version of (B) the ACP's natural phosphopantetheinyl arm (also shown as a free thiol).

KETOREDUCTASES PLAY A LARGE ROLE IN THE STEREOCHEMICAL COMPLEXITY OF POLYKETIDES

Aldol reactions employed in organic syntheses to form carbon-carbon bonds and set stereocenters in the same reaction sometimes lack the desired level of stereocontrol. The equivalent process in PKSs is carefully regulated by two enzymes: KSs that mediate

carbon-carbon bond formation and KRs that guide the setting of stereocenters (Kwan and Schulz, 2011). The combination of stereocenters produced is dependent on the KR type. KRs that perform reductions yielding L-β-hydroxyl groups are referred to as ‘‘A-type,’’ and KRs that perform reductions yielding D-β-hydroxyl groups are referred to as ‘‘B-type’’ (reductase-incompetent KRs are referred to as ‘‘C-type’’) (Caffrey, 2003; Reid et al., 2003; Keatinge-Clay, 2007; Valenzano et al., 2009; Zheng et al., 2011). KRs that reduce D-α-substituted intermediates are denoted with a ‘‘1,’’ whereas KRs that reduce L-α-substituted intermediates are denoted with a ‘‘2’’ (Keatinge-Clay, 2007). Thus, an α-substituted, β-keto diketide intermediate can be reduced to the ‘‘2*R*,3*S*’’ product by an A1-type KR, the ‘‘2*S*,3*S*’’ product by an A2-type KR, the ‘‘2*R*,3*R*’’ product by a B1-type KR, or the ‘‘2*S*,3*R*’’ product by a B2-type KR (Keatinge-Clay and Stroud, 2006; Keatinge-Clay, 2007; Zheng et al., 2010) (**Figure 1.4**).

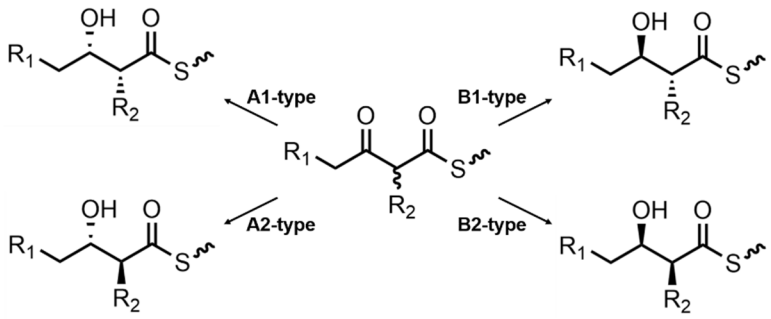

Figure 1.4. The four reductive KR types. Depending on KR type, four products are possible. A1- and B2-type KRs produce *syn* enantiomers, and A2- and B1-type KRs produce *anti* enantiomers. C-type KRs are reductase-incompetent and therefore not depicted.

Extender units also contribute to the complexity of the polyketide, adding either standard malonyl-CoA, (2*S*)-methylmalonyl-CoA, or (2*S*)-ethylmalonyl-CoA extender units (Marsden et al., 1994); or more unusual extenders such as (2*R*)-hydroxymalonyl-ACP, (2*R*)-methoxymalonyl-ACP, or (2*S*)-aminomalonyl-ACP units (Chan et al., 2009; Chan and Thomas, 2010). Although extender unit variety does contribute to the complexity of

the growing polyketide, they account for only a small percentage of the polyketide's stereocomplexity, since KRs can change the stereochemistry after extender unit condensation, if there is one present in the module. As such, this work focuses on the ketoreductase as the main gatekeeper of stereocomplexity.

OVERVIEW AND SCOPE OF THIS WORK

Since the extent to which various KRs stereospecifically reduce substrate analogs in vitro was not extensively tested, we sought to characterize their catalytic efficiency on a library of *S*-NAC surrogate substrates. Past studies providing (2*RS*)-methyl-3-oxopentanoyl-*S*-NAC to discrete KR domains found, generally, the anticipated reduced product (Holzbaur et al., 1999; Siskos et al., 2005). The KRs that normally reduce shorter substrates resembling the diketide-*S*-NACs had little difficulty reducing to the correct isomer. However, KRs from later modules (EryKR2, EryKR5, and EryKR6) which normally reduce substrates longer than the provided diketide-*S*-NACs were not as stereospecific (Holzbaur et al., 2001; Siskos et al., 2005). These KRs were, however, shown to become more stereospecific upon reduction of triketide-ACP substrates (Castonguay et al, 2007).

Chapter 2 of this dissertation examines the extent to which KRs from various PKSs maintain their natural stereocontrol when presented a panel of diketide-*S*-NAC substrates that differ at the α- and/or γ-position (**Figure 1.5**). An NADPH regeneration system (Wong et al., 1985) was employed to produce a constant supply of cofactor, and it was found that KRs generally maintain their natural stereocontrol, even on substrates that are one carbon longer or shorter at the α- and/or γ-position. We were also able to transform a previously non-stereospecific KR into a stereospecific KR in reactions with diketide attached instead

to a longer D-pantetheine arm mimic. It was also found that reactions performed with overexpressed KR cell lysate remained as stereospecific when compared to their purified assay counterpart. These biocatalytic reactions were performed on a large scale, capable of producing ~100 mg of stereospecifically-reduced diketide-*S*-NAC. Although KR reduction was monitored via HPLC, we were uncertain of the fates of substrates provided in our cell lysate reduction assay, since they contain all soluble materials produced by the overexpression host. We sought to develop a quantitative and uninvasive technique to continuously monitor substrate and product, in the event that an undesirable transformation would occur.

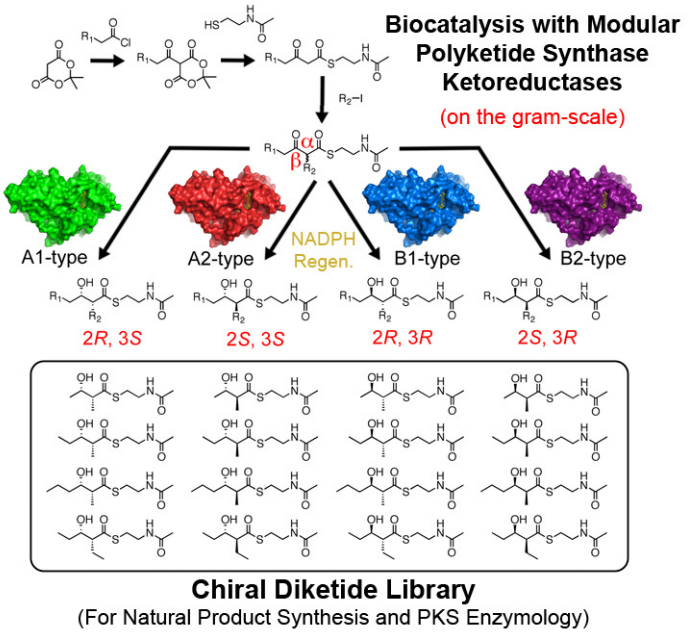

Figure 1.5. Chapter 2 overview. Synthetic substrates were used in reduction assays to produce a library of chirally-pure diketides.

Chapter 3 showcases a new technique which utilizes trifluorinated-*S*-NAC substrates and ^{19}F-NMR; allowing the monitoring of PKS biocatalytic transformations in real-time (**Figure 1.6**). The terminal thioesterase of the DEBS (EryTE) was used as a proof of principle discrete domain, and we were pleased to see that kinetic constants acquired from HPLC assays matched those of our new ^{19}F-NMR assay. After using our new technique with TylKR1 and an α-unsubstituted trifluorinated-diketide *S*-NAC we also saw similar results. Once provided with an α-substituted trifluorinated-diketide *S*-NAC, diastereomers

of the 4 KR types were able to be distinguished on a ^{19}F-NMR spectrum. Since we had extensively studied KRs that normally reduce a β-carbonyl to a β-hydroxy species, we were then interested in pursuing KRs that held unusual functions. The *trans*-AT PKSs (which do not house covalently-bound ATs, as *cis*-AT PKSs do) contain a realm of unusual domain functions (Piel, 2010). One interesting example is from the first PKS module of the bacillaene hybrid NRPS/PKS (Calderone et al., 2008). This KR is 1 of 3 known cases in which α-reduction occurs in a PKS (Hildebrand et al., 2004; Magarvey et al., 2006), and is the only known example in which β-reduction also occurs non-canonically.

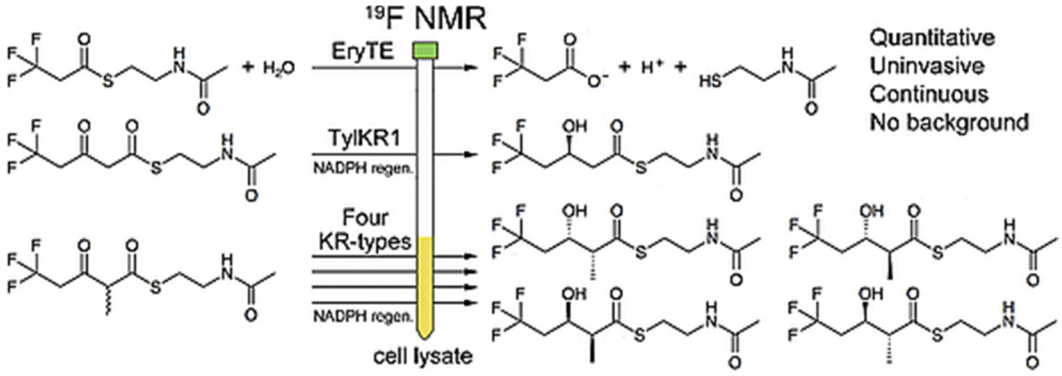

Figure 1.6. Chapter 3 overview. The catalytic turnover of fluorinated substrates was viewed in real-time by using ^{19}F-NMR.

Chapter 4 unveils the structure of this dual-function α- and β-reducing KR, as well as functionally characterizing its activity using *S*-NAC substrates (**Figure 1.7**). Overall, we found the structure to resemble the previously-solved structures of other KR types (Keatinge-Clay and Stroud, 2006; Keatinge-Clay, 2007; Zheng et al., 2010; Zheng and Keatinge-Clay, 2011; Zheng et al., 2013), yet sequence analysis revealed only a few conserved residues, although other KRs from *cis*-AT PKSs share specific conserved motifs, depending on the KR-type (Keatinge-Clay, 2007). Since we were interested in elucidating

a similar sequence-to-function set of rules, we performed sequence alignments from multiple *trans*-AT PKS KRs, and found a residue that may control β-hydroxy stereochemistry. We also present two possible substrate binding conformations which account for the two different stereochemistries produced from this KR at the α- and β-position. We hope to continue studying unusual domain functions, as this will enhance our repertoire of catalysts to choose from in the production of new natural products.

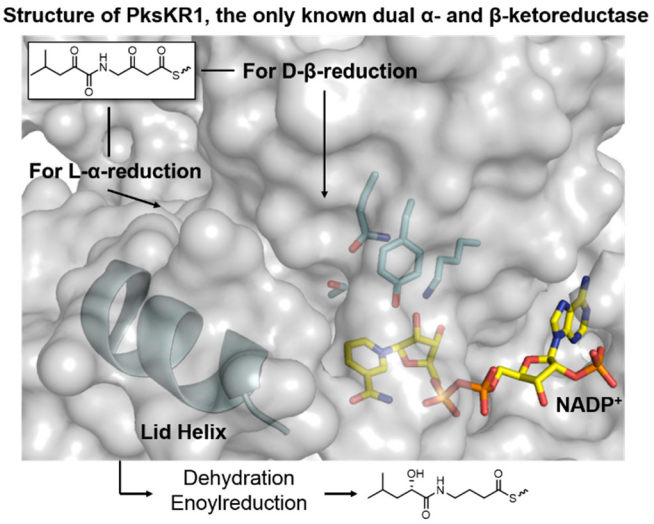

Figure 1.7. Chapter 4 overview. The structure of PksKR1 was solved and its function as an α- and β-ketoreductase was tested with *S*-NAC substrates. A sequence analysis suggests a residue important in determining stereochemistry of product. Two binding conformations are presented to explain this KRs dual-function.

The work outlined in this dissertation concentrates on KR domains as biocatalysts, in hopes to further understand the capability of this enzyme to produce chirally-pure compounds for downstream bioactive compounds. We show in **Chapter 2** that KRs are rather promiscuous (therefore allowing stereospecific reduction of even unnatural substrates), as well as robust biocatalysts (producing ~100 mg of chirally-pure product). **Chapter 3** presents a new method to follow KR catalysis in real-time, even allowing determination of the diastereomer being produced. Finally, **Chapter 4** exposes the structure of a dual-function KR, as well as its characterization with *S*-NAC analogs. This work contributes an

abundance of information for choosing discrete KR domains able to provide several desired stereochemistries, much like choosing a tool from a toolbox.

Chapter 2. Employing Modular Polyketide Synthase Ketoreductases as Biocatalysts in the Preparative Chemoenzymatic Syntheses of Diketide Chiral Building Blocks

SUMMARY

Chiral building blocks are valuable intermediates in the syntheses of natural products and pharmaceuticals. A scalable chemoenzymatic route to chiral diketides has been developed that includes the general synthesis of α-substituted, β-ketoacyl *N*-acetylcysteamine thioesters, followed by a biocatalytic cycle in which a glucose-fueled NADPH-regeneration system drives reductions catalyzed by isolated modular polyketide synthase (PKS) ketoreductases (KRs). To identify KRs that operate as active, stereospecific biocatalysts, 11 isolated KRs were incubated with 5 diketides and their products were analyzed by chiral chromatography. KRs that naturally reduce small polyketide intermediates were the most active and stereospecific toward the panel of diketides. Enhanced stereocontrol was observed when *N*-acetylcysteamine was replaced by the more natural D-pantetheine. Several biocatalytic reactions were scaled up to yield more than 100 mg of product. These syntheses demonstrate the ability of PKS enzymes to economically and greenly generate diverse chiral building blocks on a preparative scale.

INTRODUCTION

Because KRs can exert control over two stereocenters in one reduction reaction and, as a family, operate on diverse organic substrates, they possess extraordinary potential as biocatalysts (Patel, 2008). At least two studies of KRs as biocatalysts have been reported; however, the substrates examined were not very similar to those naturally encountered and

the stereochemical purities of the reaction products were not investigated (Bali et al., 2006; Bali and Weissman, 2006). Analytical-scale enzymological studies have focused on the substrate analog (2*RS*)-methyl-3-oxopentanoyl-*S*-*N*-acetylcysteamine (-SNAC), **2.3**. Reductions of **2.3** mediated by EryKR1, EryKR2, EryKR5, EryKR6, MycKRA, MycKRB, and TylKR1 revealed that EryKR1 (B2-type) and TylKR1 (B1-type) were completely stereospecific, whereas the other KRs produced a mixture of diastereomers (isolated domains are referred to by their PKS and module of origin, except the mycolactone KRs, which are identified by KR type) (Siskos et al., 2005; Bali and Weissman, 2006). That from the isolated KRs examined for their stereospecificity, two KRs were identified that catalyze stereospecific reduction suggested that other KRs that perform stereocontrolled reductions isolated from their PKSs could be discovered. If all KR types were available as biocatalysts, then the full stereochemical diversity naturally achieved by KRs could be accessed.

Thus, to demonstrate the utility of PKS enzymes in natural product and pharmaceutical synthesis, we aimed to develop KRs as biocatalysts and access a library of diketide chiral building blocks (**Figure 2.1C**). Required were both a facile synthetic route to α-substituted, β-keto diketide thioester analogs of the polyketide intermediates naturally reduced by KRs, as well as a scalable, economic biocatalytic process that does not require stoichiometric NADPH for KR-mediated reduction.

A general synthesis of α-substituted, β-ketoacyl *N*-acetylcysteamine thioesters was developed and used to generate diketide substrates **2.1**–**2.5**. An NADPH-regeneration system in which KR-mediated reductions are driven by glucose was designed. Eleven isolated KRs were incubated with **2.1**–**2.5** in biocatalytic reactions and the products were analyzed by chiral chromatography. Robust A1-type and A2-type KR biocatalysts were

identified. Compared to reactions with (2*RS*)-methyl-3-oxopentanoyl-SNAC (**2.3**), reactions with (2*RS*)-methyl-3-oxopentanoyl-D-pantetheine (**2.6**) were generally both more complete and more stereocontrolled. Several biocatalytic reactions were scaled up, both with purified KRs as well as KR-containing lysate, to readily provide 100 mg yields of stereopure chiral building blocks. In addition to illustrating that building blocks for the synthesis of natural products and pharmaceuticals can be accessed by PKS enzymes, the described biocatalytic systems provide PKS enzymologists access to polyketide substrates and enable experiments on PKS enzymes to be performed on a more readily analyzable scale.

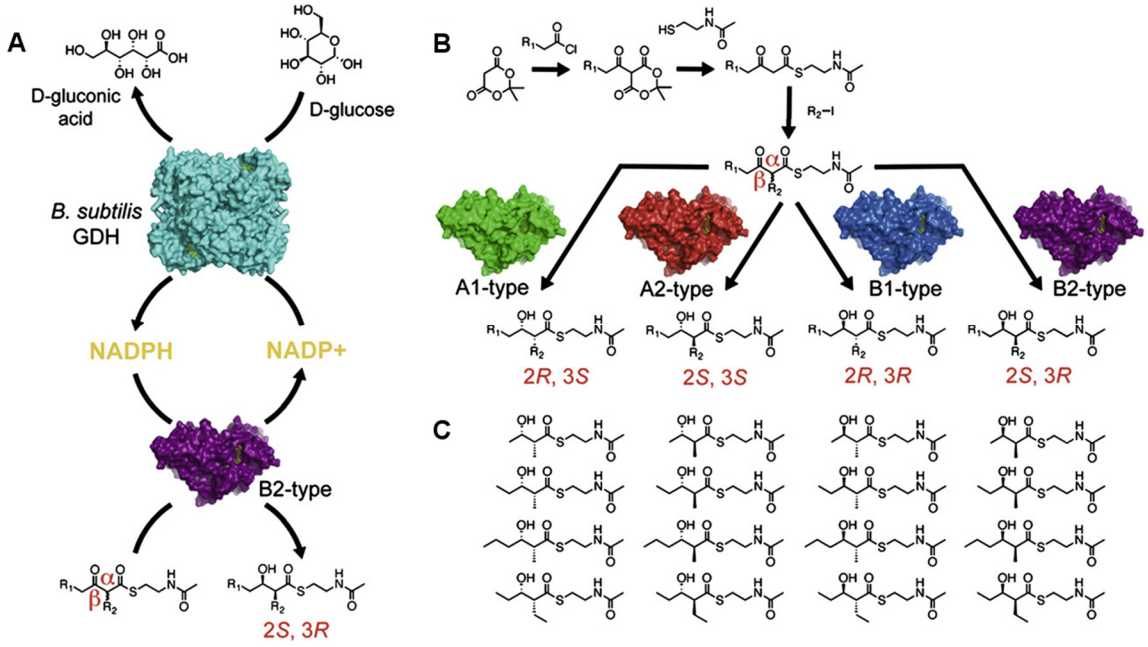

Figure 2.1. Employing PKS KRs to chemoenzymatically synthesize a library of chiral building blocks.
(A) GDH helps drive the reduction of diketide-SNACs, regenerating NADPH through the oxidation of D-glucose.
(B) A general synthesis of α-substituted, β-keto diketides was developed. Stereospecific KRs can convert these common intermediates into chiral diketide-SNACs.
(C) Representatives of the library of diketide chiral building blocks that can be accessed through KR biocatalytic reactions.

RESULTS AND DISCUSSION

Chemoenzymatic Design

NADPH-Regeneration Scheme

To drive KR-mediated reduction reactions, an NADPH-regeneration scheme was employed (**Figure 2.1A**). Glucose dehydrogenase (GDH) cloned from *Bacillus subtilis* was used to couple the oxidation of glucose to the reduction of diketides. GDH generates NADPH from $NADP^+$ as it oxidizes glucose to gluconolactone (Wong et al., 1985; Kataoka et al., 2003). A high concentration of glucose establishes a large NADPH:$NADP^+$ ratio that drives the reduction reaction toward completion. At the pH of the biocatalytic reactions (pH 7.5), most of the gluconolactone produced is hydrolyzed to gluconic acid, which deprotonates to gluconate. Because decreases in pH result in GDH and KRs precipitating out of solution, excess buffering strength is supplied.

General Synthetic Route to α-Substituted, β-Keto N-Acetylcysteamine Thioesters

The route most commonly employed to generate **2.3** requires a costly chiral auxiliary to generate an α-substituted, β-ketoimide that is subsequently reacted with excess *N*-acetylcysteamine (even more expensive than the chiral auxiliary) (Gilbert et al., 1995). Instead, we access α-substituted, β-keto *N*-acetylcysteamine thioesters through an efficient and cost-effective extension of a described route to β-keto *N*-acetylcysteamine thioesters (Gilbert et al., 1995) (**Figure 2.1B**). After the desired β-keto *N*-acetylcysteamine thioester has been obtained from Meldrum's acid, an acyl chloride, and *N*-acetylcysteamine (generated from cysteamine and acetic anhydride), an α substituent is installed via an alkylation reaction. Thus, the γ substituent of a diketide-SNAC is selected through the acyl

chloride, and the α-substituent is selected through the alkyl halide. A one-pot synthesis was also developed to produce (2*RS*)-methyl-3-oxopentanoyl-D-pantetheine (**2.6**).

That the syntheses yield racemic mixtures of α-substituted, β-ketoacyl-SNACs is a nonissue because the enantiomers rapidly interconvert in the aqueous environment of the subsequent biocatalytic reaction.

For isolated KRs to be considered robust biocatalysts, they should catalyze the reduction of a variety of substrates. Thus, a panel of five diketide substrates was synthesized to test the KRs: 3-oxopentanoyl-SNAC (**2.1**), (2*RS*)-methyl-3-oxobutanoyl-SNAC (**2.2**), (2*RS*)-methyl-3-oxopentanoyl-SNAC (**2.3**), (2*RS*)-methyl-3-oxohexanoyl-SNAC (**2.4**), and (2*RS*)-ethyl-3-oxopentanoyl-SNAC (**2.5**) (**Figure 2.2**). Compared to **2.3**, each of the other diketides is one methylene shorter or longer at the α- or γ-position.

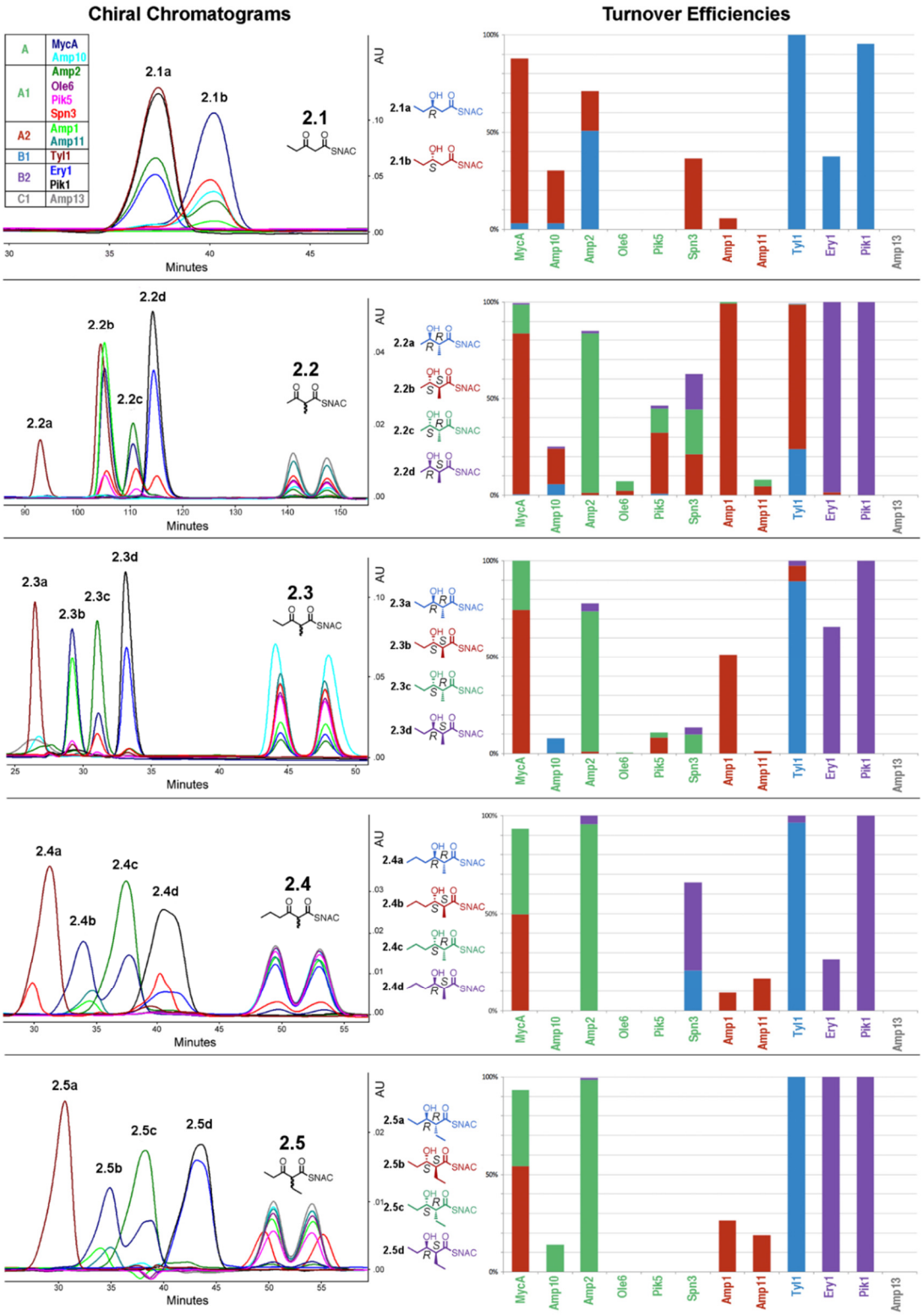
Chiral Chromatograms
Turnover Efficiencies
A
MycA
Amp10
A1
Amp2
Ole6
Pik5
Spn3
A2
Amp1
Amp11
B1
Tyl1
B2
Ery1
Pik1
C1
Amp13
2.1a
2.1b
2.1
2.2a
2.2b
2.2c
2.2d
2.2
2.3a
2.3b
2.3c
2.3d
2.3
2.4a
2.4b
2.4c
2.4d
2.4
2.5a
2.5b
2.5c
2.5d
2.5
SNAC
AU
Minutes
100%
50%
0%
MycA
Amp10
Amp2
Ole6
Pik5
Spn3
Amp1
Amp11
Tyl1
Ery1
Pik1
Amp13

Figure 2.2. KR Biocatalytic Assays (on previous page).
Chromatograms show the reactions of tested KRs with each of the diketide substrates **2.1**–**2.5** on the left. Chiral building block products are displayed in the middle. The turnover efficiencies of each KR for each substrate are graphed on the right ("turnover efficiency" is defined as the quantity of a product divided by quantities of the products and remaining substrate). The heights of the bars can be interpreted as relative activities, whereas the ratio of colors within a bar can be interpreted as a measure of the stereocontrol exerted by a KR on a substrate. Diverse KR types were used in these assays: two A-types that naturally operate on α-unsubstituted intermediates, four A1-types, two A2-types, one B1-type, two B2-types and one C1-type (as a negative control).

Selection of KRs

To identify KRs capable of stereospecifically reducing a range of diketide substrates, 11 diverse reductase-competent KRs were selected (see "Cloning" in Experimental Procedures). They were isolated from modules that do not possess DHs or ERs because KRs that cooperate with these enzymes are usually B1-type (Valenzano et al., 2010). MycKRA (A-type), TylKR1 (B1-type), and EryKR1 (B2-type) were chosen, in part, to serve as positive controls, as how they process **2.3** has already been determined (Siskos et al., 2005; Bali and Weissman, 2006). The selection of the other eight KRs was biased toward A-type KRs, both because isolated A1- or A2-type KRs that stereospecifically reduce β-ketoacyl-SNAC substrates had previously not been identified, and because A1-type KRs are the most common in modules without DHs or ERs. All but two of the KRs selected naturally operate on α-substituted intermediates because we aimed to set two stereocenters through a single KR-mediated reduction reaction. KRs were chosen from diverse PKSs (amphotericin, erythromycin, mycolactone, oleandomycin, pikromycin, spinosyn, and tylosin) and from varying locations within PKS assembly lines (**Figure 2.6**). Thus, the selection of reductase-competent KRs includes two A-type KRs that naturally reduce α-unsubstituted intermediates, four A1-type KRs, two A2-type KRs, one B1-type KR, and two B2-type KRs. A reductase-incompetent C1-type KR was also selected as a negative control (**Figure 2.2**).

Chiral Chromatography and Establishing Elution Orders

To determine the stereocontrol exerted by KRs when reducing **2.3**, Leadlay and coworkers synthesized standards that established the **2.3a**–**2.3d** elution order from a chiral column (Holzbaur et al., 1999). We sought to establish the elution order of **2.2a**–**2.2d** and **2.4a**–

2.4d on the same chiral column to identify the reduced diketides produced by the 11 KRs in the biocatalytic reactions. Thus, well-established diastereospecific aldol reactions were performed to obtain synthetic "2*R*,3*R*," "2*R*,3*S*," and "2*S*,3*R*" diketide-SNACs and the elution orders of **2.2a**–**2.2d** and **2.4a**–**2.4d** were determined (Evans et al., 1981; Raimundo and Heathcock, 1995) (**Figure 2.3**). The products of **2.2**, **2.3**, and **2.4** elute in the same order- first the two *anti*-products, then the two *syn*-products: "2*R*,3*R*" (B1-type product), "2*S*,3*S*" (A2-type product), "2*R*,3*S*" (A1-type product), and "2*S*,3*R*" (B2-type product) (Holzbaur et al., 1999).

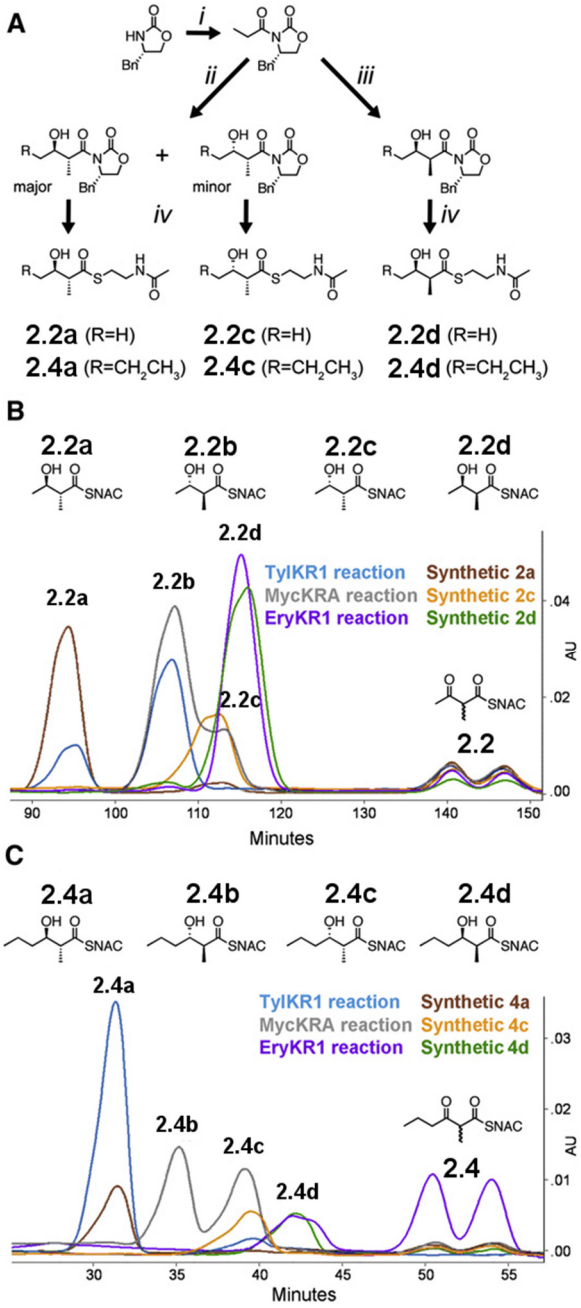

Figure 2.3. Determining Elution Orders. (A) Aldol reactions helped synthesize chiral chromatography standards. After (*i*) the propionyl group was added to the chiral auxiliary, then either (*ii*) the Heathcock route using two equivalents of dibutylboron triflate, or (*iii*) the standard Evans route using one equivalent yielded diketide oxazolidinones that were (*iv*) hydrolyzed and coupled to NAC.
(B) and (C) Standards were compared to products from MycKRA (A-type), TylKR1 (B1-type), and EryKR1 (B2-type) to establish the elution order of products from KR reactions with **2.2** and **2.4**. See also Table 2.2.

Syntheses of reduced diketide-SNACs were initiated through aldol reactions between (4*S*)-4-benzyl-3-propionyl-2-oxazolidinone and the appropriate aldehyde (**Figure 2.3A**). In keeping with the classic Evans *syn*-selective reaction protocol, one equivalent of dibutylboron triflate was used to generate the "2*S*,3*R*" product (Evans et al., 1981). Heathcock and coworkers demonstrated that if two equivalents of dibutylboron triflate are supplied to this reaction, the *anti*-product with the α-stereocenter opposite in configuration to the Evans *syn*-product as well as some quantity of anti-Evans *syn*-product are generated (Raimundo and Heathcock, 1995). We utilized this protocol to generate a mixture of "2*R*,3*R*" and "2*R*,3*S*" products. Chemistry detailed by Boddy and coworkers was applied to each of the diketide-oxazolidinones to hydrolyze the diketides from their chiral auxiliaries and couple them to NAC (Sharma and Boddy, 2007). The "2*R*,3*R*" and "2*R*,3*S*" diketide-SNACs were separated by reverse-phase HPLC.

Analyzing KR Reactions

After reactions in which each of the 11 reductase-competent KRs was paired with each of **2.1**–**2.5** had been incubated for 24 hr at 22°C, they were extracted with ethyl acetate and analyzed by chiral chromatography (**Figure 2.2**). Some reactions may not have gone to completion due to substrate incompatibility, slow kinetics, protein stability, or many other factors. Because the overall fitness of KRs as biocatalysts was of interest, a metric of "turnover efficiency" was constructed, defined as the quantity of a product formed divided by the quantities of all products and unreacted substrate. Relative quantities were measured by integrating peaks from the 235 nm absorbance traces of chiral chromatography runs (for data, see Appendix). Because all reactions were performed in duplicate, the reported turnover efficiencies are averages from two runs (in no case was there a deviation larger

than 10%). As **2.1** could not be accurately quantified (it migrates as an extremely broad peak), the turnover efficiencies of **2.1a** and **2.1b** were calculated relative to the TylKR1 reaction, which produced the largest product peak. The elution order for **2.1a** and **2.1b** was assigned as first the "3*R*" product then the "3*S*" product, because reactions of active B-type KRs generated the first peak and reactions of all of the active A-type KRs (with the exception of AmpKR2) generated the second peak. The elution order of **2.5a**–**2.5d** was assigned in the same order as the products of **2.2**–**2.4**. Added validation for this assignment comes from the observation that the *syn*- and *anti*-products of **2.2**–**2.4** have different maximal absorbances (the *anti*-products maximally absorb at ~232 nm, whereas the *syn*-products maximally absorb at ~233 nm): the first two biocatalytic products of **2.5** that elute from the chiral column absorb as *anti*-products, whereas the last two absorb as *syn*-products.

Reactions with the diketide-pantetheine substrate **2.6** were analyzed by reverse-phase chromatography (**Figure 2.4**). Stereochemistries of the two product peaks were assigned using AmpKR2 (A1-type), MycKRA (produces both A1-type and A2-type products), TylKR1 (B1-type), and EryKR1 (B2-type). While the products are all diastereomers, the third chiral center on the D-pantetheine moiety is distant from the other two on the diketide. Thus, on the reverse-phase column the A1-type and B2-type (*syn*) products possessed equivalent retention times, as did the A2-type and B1-type (*anti*) products. The A1-type and B2-type products migrated faster than the A2-type and B1-type products. That most KRs produce the same stereochemistries on the pantetheine-bound substrate **2.6** as with the SNAC-bound substrate **2.3** is supported by the equivalent ratios of A1-type and A2-type products made by MycKRA.

KR Biocatalysis Results

KRs Usually Maintain Natural Stereocontrol

Active KRs usually maintained the stereocontrol anticipated from their KR types (**Figure 2.2**). Most of these KRs even maintained stereocontrol toward substrate **2.5**, which harbors an uncommon α-ethyl group. AmpKR2 was identified as a general A1-type KR biocatalyst, and AmpKR1 was identified as a general A2-type KR biocatalyst. Because TylKR1 proved to be a general B1-type KR biocatalyst and EryKR1 proved to be a general B2-type KR biocatalyst (PikKR1 is even more robust), the desired spectrum of KR activities was realized. That the KR reactions are as stereocontrolled as they were determined to be is noteworthy, because the natural substrates of these enzymes are polyketide intermediates bound to the pantetheinyl arm prosthetic group of holo-ACPs.

KRs That Naturally Reduce Smaller Intermediates Are Most Active Toward Diketide-SNACs

In general, KRs that reduce small polyketides within their native PKSs, such as AmpKR1, AmpKR2, EryKR1, PikKR1, and TylKR1, were most active on the panel of diketides. KRs that reduce larger polyketides within their native PKSs, such as PikKR5, OleKR6, AmpKR10, and AmpKR11, were nearly inactive toward **2.1**–**2.5**. This could indicate that KRs that naturally reduce short-chain polyketide intermediates within their native PKSs obtain binding energy for catalysis from interactions with the polyketide.

MycKRA was one of the most active KRs in these studies. Three modules of the mycolactone PKS contain 99% identical MycKRA enzymes as a result of duplication events; thus, within the synthase MycKRA enzymes reduce a diketide, a triketide, and a

hexaketide (Stinear et al., 2004). That both A1-type and A2-type products were produced when MycKRA was incubated with α-substituted diketides **2.2**–**2.5** is understandable in that this enzyme naturally operates on α-unsubstituted intermediates. MycKRA did not produce B-type products. Broad in its substrate selectivity, MycKRA is a robust, stereocontrolled A-type KR biocatalyst.

Deviations from Expected

A few KRs were sensitive to slight differences within the panel of diketides. TylKR1, which is in general a robust, active B1-type KR, produced mainly the A2-type product **2.2b** when incubated with **2.2**, which is one methylene shorter than its natural substrate. AmpKR2, an A1-type KR, converted the α-unsubstituted diketide **2.1** mainly to a B-type product. Therefore, we sought to determine whether KRs are more active and/or stereocontrolled toward pantetheine-bound diketides than NAC-bound diketides.

A Pantetheinyl Attachment Enhances Activity and Stereocontrol

The minimal substrate to achieve stereocontrolled KR reactions has yet to be determined. It has been shown that EryKR2 (A1-type) reduces **2.3** to all four products, while EryKR1 produces exclusively one product. Thus, some KRs may require more of the phosphopantetheinyl arm or even their cognate ACP to achieve stereocontrol than others. As some KRs were observed to lose stereocontrol towards the SNAC-bound diketides **2.1**-**2.5**, we hypothesized that the SNAC moiety is sometimes a suboptimal mimic of the natural phosphopantetheinyl arm, and that a more extended mimic may result in increased stereocontrol. Thus, **2.6** was synthesized and incubated with EryKR2 and several robust KRs (**Figure 2.4**).

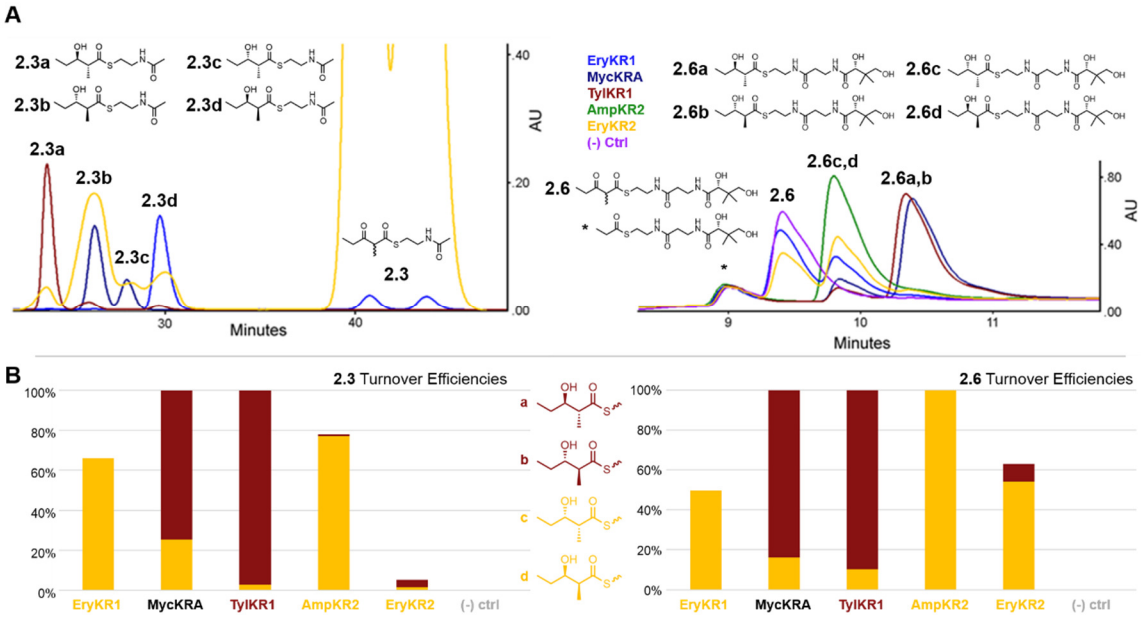

Figure 2.4. The Pantetheinyl Moiety Generally Enhances KR Activity and Stereocontrol. (A) Reactions show that EryKR2 is barely active and does not exert stereocontrol when incubated with **2.3** (the EryKR2 chromatogram was expanded to show product detail). Chromatogram for EryKR2 reaction with **2.3** was overlaid with representative KR types from Figure 2.2. Reactions using **2.6** were run using reverse-phase HPLC, and showed an increase in the correct isomer produced for EryKR2. The purple trace is from an incubation with no KR (negative control). The starred compound is a byproduct from synthesis of **2.6**. (B) Turnover efficiencies were calculated exactly as for Figure 2.2. Peak areas from reactions yielding *syn*-products are shown in gold, whereas peak areas from reactions yielding *anti*-products are shown in brown. When incubated with **2.6**, EryKR2 was both more active and more stereocontrolled (compare the production of mostly A2-type product **2.3b** from the incubation with **2.3** to the production of the A1-type (or B2-type) product (**2.6c** & **d**) when incubated with **2.6**.

The longer pantetheinyl attachment resulted in increases both in the activity and in the stereoselectivity of EryKR2. The ratios of the products from the incubation of EryKR2 with **2.3** agree with previous studies with EryKR2 (all four are produced, with the A2-type product dominating) (Baerga-Ortiz et al., 2006). When the same diketide was attached to the pantetheine moiety (**2.6**), an increase of stereoselectivity was observed (little to no A2-product was produced, and a peak that represents either the A1-type or B2-type product

was produced). An increase in activity was also observed, as the majority of **2.6** was reduced (in contrast to the few percent of **2.3** that was reduced). Interestingly, AmpKR2 (another A1-type KR) was also more active toward **2.6** than **2.3**, converting all substrate to product. Indeed, the pantetheine moiety increased the activity and stereocontrol of EryKR2. However, it is possible that the activity and stereoselectivity of EryKR2 may be further optimized through interactions with the phosphoryl moiety of the phosphopantetheinyl arm.

Scaled-Up Reactions

Because KR biocatalysts representing each KR type were identified that could generate the desired library of diketide chiral building blocks, we sought to demonstrate that useful quantities of chiral building blocks could be obtained. AmpKR2 (A1-type), AmpKR1 (A2-type), TylKR1 (B1-type), and EryKR1 (B2-type) were each incubated in scaled-up reactions with **2.3**. Over 100 mg each of **2.3a**, **2.3b**, **2.3c**, and **2.3d** was obtained from these scaled-up reactions. Because the reactions with lower turnover efficiencies required significant KR concentrations and the nickel column purification step only captures a small percentage of KR in the lysate from overexpressing cells, reactions were also attempted using lysate. Reactions using the lysate were complete within 6 hr and also provided more than 100 mg of chiral product. Both for reactions using pure KRs and KR-containing lysate, KR reactions were as stereocontrolled as in the smaller-scale reactions.

Building Block Libraries, PKS Biocatalysis, and PKS Enzymology

As well as containing valuable combinations of stereocenters, the described building blocks harbor a synthetic handle in the form of a thioester. This moiety can be converted to a carboxylic acid or an aldehyde, or used to couple the diketide to other fragments

through an ester, amide, or ketone linkage, making thioester-containing building blocks very versatile in the syntheses of natural products (Bruice et al., 1963; Tokuyama et al., 1998). The installation of a second synthetic handle on the other end of the diketide chain is also desirable. Because the biocatalytic assays reported here revealed that variations at the γ-position of the diketide substrates are tolerated by many KRs, it may be possible to incorporate synthetic handles such as a terminal chlorine, alcohol, double bond, or triple bond.

Additional building block libraries may be accessed using other PKS enzymes as biocatalysts. Because many DHs form trans double bonds from (2*R*,3*R*)-α-alkyl-β-hydroxy intermediates, preparative quantities of α-substituted, α/β-unsaturated diketides may be accessed by incubating α-substituted, β-ketodiketide-SNACs with appropriate KRs and DHs (Keatinge-Clay, 2008). As ERs stereospecifically reduce such double bonds, α-substituted, β-keto-diketide-SNACs could be converted, through the driving force of an NADPH-regeneration system, to chiral α-substituted, β-methylene-diketide-SNACs by employing a combination of KR, DH, and ER biocatalysts (Kwan et al., 2008). Diketide-SNACs can also be enzymatically converted to triketide lactones containing four contiguous stereocenters (**Figure 2.5**). Such molecules have been generated by Module+TEs (PKS modules C-terminally fused to EryTE) in quantities analyzable by radio-TLC or GC-MS as readout of enzymological studies (Chen et al., 2006). Recently, we have developed an economical route to supply extender units accepted by ATs to in vitro PKS reactions (Hughes and Keatinge-Clay, 2011). Thus, the scaled-up incubation of

reduced, chiral diketide-SNACs with diverse Module+TEs in the presence of NADPH may yield libraries of triketide building blocks.

Currently, most experiments on modular PKS enzymes are performed either in vivo, by expressing PKSs heterologously and extracting polyketides from the growth medium, or in vitro, by combining the appropriate enzymes, cofactors, and substrates and then analyzing the reaction products by radio-TLC, GC-MS, or LC-MS (reduced polyketides are usually invisible on UV absorbance traces). An experimental platform in which PKS enzymes can perform catalysis in a controlled setting to produce quantities of polyketides analyzable by NMR or crystallography is highly desired (the stereocenters of products would be much more readily determined, for example). Expense has been the principal deterrent of larger in vitro reactions because polyketide extender units such as methylmalonyl-CoA and cofactors like NADPH are costly. However, such biocatalytic processes as the MatB-catalyzed production of extender units and GDH-catalyzed regeneration of NADPH should enable PKS enzymologists to perform in vitro reactions more affordably under controlled conditions, generating multimilligram quantities of polyketide products.

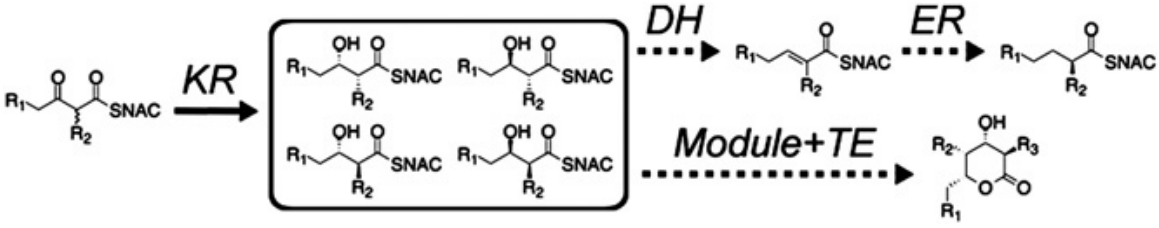

Figure 2.5. Accessing Diverse Building Block Libraries through PKS Biocatalysis. Biocatalytic reactions using DHs, ERs, and Module+TEs could enable the generation of libraries of other building blocks useful in the syntheses of natural products, such as α-substituted diketides and triketide lactones containing up to four contiguous stereocenters (representative molecules are shown).

SIGNIFICANCE

The scalable production of reduced diketide-SNACs (-*S*-*N*-acetylcysteamines) has been realized through the described general chemoenzymatic route. A screen of 11 diverse, isolated ketoreductases (KRs) toward diketides **2.1**–**2.5** revealed that KRs isolated from modules that naturally reduce shorter polyketide intermediates are most active on diketide-SNACs. AmpKR2 and AmpKR1 were identified as robust A1- and A2-type KR biocatalysts, and TylKR1 and EryKR1 were shown to be robust B1- and B2-type KR biocatalysts. EryKR2 showed greater activity and stereocontrol when presented the pantetheinyl-bound substrate **2.6** over the SNAC-bound substrate **2.3**, suggesting that binding energy is obtained from interactions between EryKR2 and the pantetheinyl group. From the structure of AmpKR2, another A1-type KR, the phosphopantetheinyl arm has the opportunity to make contacts with the active site groove (Zheng et al., 2010). More than 100 mg of each diketide **2.3a**–**2.3d** was produced to demonstrate that quantities of chiral diketide building blocks useful for natural product and pharmaceutical syntheses could be generated economically and greenly. Other polyketide synthase (PKS) enzymes such as dehydratases, enoylreductases, and PKS modules C-terminally fused to EryTE can now be investigated as biocatalysts in the generation of an even greater diversity of diketide and triketide building blocks. The described biocatalysis will also benefit PKS enzymology because it is no longer cost-prohibitive to perform in vitro reactions that yield significant quantities of polyketide products as readout.

EXPERIMENTAL PROCEDURES

Cloning

All KRs, with the exception of EryKR1 and GDH, were cloned from genomic DNA (gDNA) (**Table 2.1**). Regions of DNA encoding KRs from amphotericin (Amp), oleandomycin (Ole), pikromycin (Pik), spinosyn (Spn), and tylosin (Tyl) were amplified from *Streptomyces nodosus*, *Streptomyces antibioticus*, *Streptomyces venezuelae*, *Saccharopolyspora spinosa*, and *Streptomyces fradiae* gDNA, respectively. N- and C-terminal boundaries were chosen using the boundaries determined from the structures of EryKR1 and TylKR1 (Keatinge-Clay and Stroud, 2006; Keatinge-Clay, 2007). The region encoding the KR from the mycolactone (Myc) PKS was amplified from a bacterial artificial chromosome containing DNA from pMUM001 (Stinear et al., 2004). The regions encoding the KRs from the erythromycin (Ery) PKS were amplified from a synthetic gene (Kodumal et al., 2004). GDH was cloned from *B. subtilis* gDNA. NdeI and EcoRI restriction sites were designed to flank the amplicons encoding AmpKR2, AmpKR10, AmpKR11, AmpKR13, EryKR2, OleKR6, PikKR1, PikKR5, and SpnKR3; NdeI and XhoI restriction sites were designed to flank the amplicons encoding AmpKR1, EryKR1, TylKR1, and GDH; NheI and NotI restriction sites were designed to flank the amplicon encoding MycKRA. All restriction sites were engineered so that histidine tags encoded by pET28b (Novagen) would be N-terminally fused to the proteins, facilitating purification.

Table 2.1. Cloning Primers.
For each protein, forward and reverse primers were designed. Underlined sequences indicate restriction sites.

Primer Name	Primer Sequence (5'-3')
AmpKR1-F	GGAGATATACATATGGGCGAGCGCTCCACCGTCGAC
AmpKR1-R	CGTAATGCTCGAGTCAGTCGGCCAGCATCCGCGCCAC
AmpKR2-F	ATCGTAATCCATATGGAGCCGCTGGCCCAGATCC
AmpKR2-R	TGATTCGATGAATTCACTCGTCGCTGAGCGCCGAGACG
AmpKR10-F	ATCGTAATCCATATGCTCCTCGACACGCTGCGCTAC
AmpKR10-R	TGATTCGATGAATTCAGGTGTCGGCGGGGGTGTTCTC
AmpKR11-F	ATCGTAATCCATATGCGCTACCGCCCGACCTGGAA
AmpKR11-R	TGATTCGATGAATTCACTGCACCTCGGGGACCTCGT
AmpKR13-F	ATCGTAATCCATATGGACTGGCGCTACCACGAGAG
AmpKR13-R	TGATTCGATGAATTCAGCCGTCGAAGAGTGCGGTC
EryKR1-F	GGAGATATACATATGTCTACCGAGGTTGATGAAGTC
EryKR1-R	GTGGTGCTCGAGTCATCAGCGTGGCTCAGCAGCGGCCTGC
MycKRA-F	TATGTCCACGCTAGCGCCCAGAACGTTGAATCCCG
MycKRA-R	ACAGGTTCCAGCGGCCGCTCACGACTCGCGGTAAGCCATTACC
OleKR6-F	ATCGTAATCCATATGCATCGACGGCTTCAGGACTGG
OleKR6-R	TGATTCGATGAATTCAGGTACGGACCTCCGGGAGGT
PikKR1-F	ATCGTAATCCATATGGCAACTGGTGATGATT
PikKR1-R	TGATTCGATGAATTCAGTCACGGGCATCGATGATAC
PikKR5-F	ATCGTAATCCATATGGACGACTGGCGTTACCGCGT
PikKR5-R	TGATTCGATGAATTCACGGGACGCCGTCGAGCAGAA
SpnKR3-F	ATCGTAATCCATATGAGGGTCTTGGCAGACGGTTG
SpnKR3-R	TGATTCGATCTCGAGTCAATTCGCGGGCATCGCCTTTC
TylKR1-F	GCAGATATACATATGAGCCCCACCGATGCCTGGCGC
TylKR1-R	GTGGTGCTCGAGTCATCAGGCTGCCGTCAGGGCCTCCCG
GDH-F	GCAGATATACATATGTATCCGGATTTAAAAGG
GDH-R	GTGGTGCTCGAGTTAACCGCGGCCTGCCTGG

Protein Expression and Purification

All plasmids were transformed into *Escherichia coli* BL21(DE3). The cells were grown in Luria broth with 25 mg/L kanamycin at 37°C until they reached an OD_{600} of 0.5. They were

then cooled to 15°C, induced with 1 mM IPTG, and grown for an additional 16 hr. Cells were spun down, resuspended in lysis buffer (10% [w/v] glycerol, 0.5 M NaCl, 30 mM Tris [pH 7.5]), and lysed by sonication. After centrifugation at 30,000 x g for 30 min, lysate was poured over nickel-NTA resin (Qiagen) equilibrated with lysis buffer. Bound protein was washed with 15 mM imidazole in lysis buffer and eluted with 150 mM imidazole in lysis buffer. Proteins were then polished using a Superdex 200 gel filtration column that was equilibrated with 10% (w/v) glycerol, 150 mM NaCl, 10 mM Tris (pH 7.5), and concentrated to 200 mM. Protein purity was evaluated by Coomassie-stained SDS-PAGE gels (**Figure 2.6**).

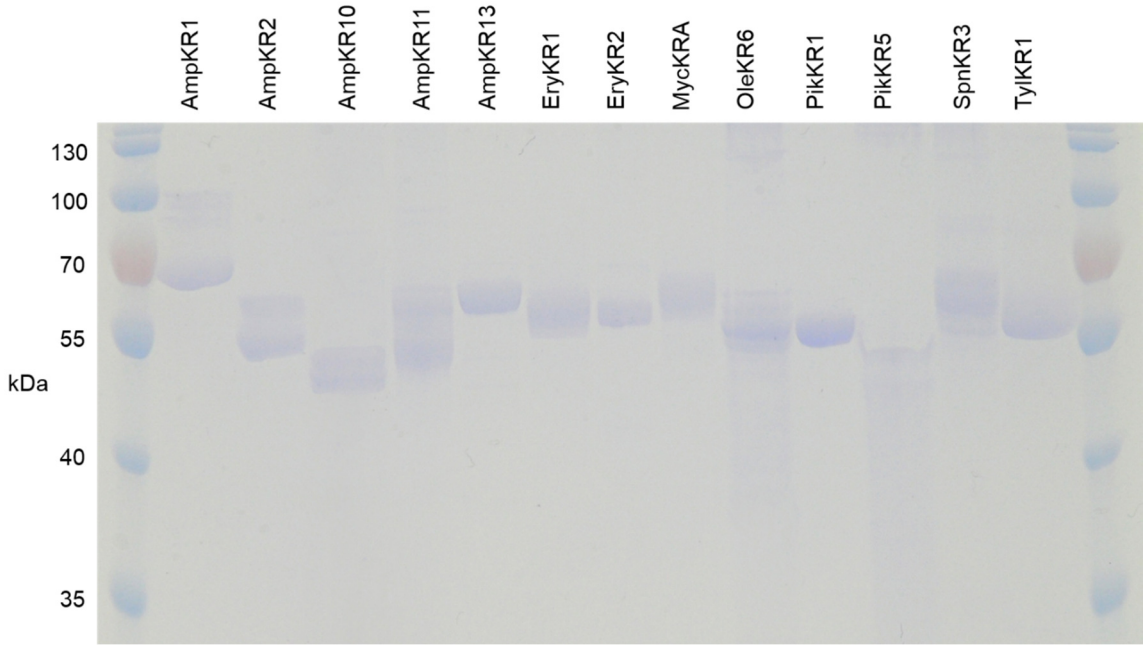

Figure 2.6. Protein Purity.
A Coomassie-stained 12% SDS-PAGE gel shows the level of purity of the KRs.

KR Reductions

Biocatalytic Screen

Reactions were performed on 10 mM **2.1**–**2.5** in 10% (w/v) glycerol, 100 mM NaCl, 150 mM HEPES (pH 7.5), 200 mM D-glucose, 100 mM $NADP^+$, 1 mM GDH, and 5 mM KR in a total volume of 200 mL. Following incubation at 22°C for 1 day, reactions were extracted with 2 x 500 mL ethyl acetate (EtOAc) and evaporated. The extracts were resuspended in ethanol prior to analysis by chiral chromatography. The pantetheine diketide reactions were incubated under the same conditions; however, they were not extracted prior to injection for analysis by reverse-phase chromatography.

Scaled-Up Reactions

Reactions were performed on 50 mM **2.3** in 10% (w/v) glycerol, 300 mM HEPES (pH 7.5), 100 mM NaCl, 300 mM D-glucose, 100 mM $NADP^+$, 1 mM GDH, and 50 mM purified KR in a total volume of 15 mL. All reactions were incubated at 22°C for 1 day. Each was determined to be complete by reverse-phase HPLC. After EtOAc extraction and silica gel chromatography (EtOAc), yields of ~75% were obtained. Reactions were also performed using lysate (50 mL from 6 L culture) that was not poured over a nickel column, but rather dialyzed overnight at 4°C in 10 kDa MWCO cellophane dialysis tubing against 10% (w/v) glycerol, 250 mM NaCl, and 30 mM HEPES (pH 7.5). In these reactions, 50 mM **2.3** was incubated in 10% (w/v) glycerol, 300 mM HEPES (pH 7.5), 100 mM NaCl, 300 mM D-glucose, 100 mM $NADP^+$, and 1 mM GDH in a total volume of 15 mL, but with 5 mL of the reaction volume being supplied from dialyzed lysate. All reactions were incubated at 22°C for 6 hr. Each was determined to be complete by reverse-phase HPLC. After EtOAc

extraction and silica gel chromatography (EtOAc), yields of ~85% were obtained. Background thioester hydrolysis occurs with a half-time of ~3 days for both reactions with purified KRs and reactions with KR-containing lysate.

Chiral Chromatography

Extracts of reactions from the biocatalytic screens were separated using a ChiralCel OC-H column (250 x 4.6 mm) on a Beckman Coulter HPLC system with a 20 µL loop. Solvent systems and flow rates were optimized for each set of reactions with a particular diketide-SNAC (**Table 2.2**). Substrates and products were observed at a wavelength of 235 nm. Reaction products from incubations with **2.6** were separated using a C18 column (Vydac, 250 x 4.6 mm) on a Waters 1525/2998 system with a 50 µL loop. A gradient of 20-70% B over 15 minutes was used, where the mobile phases were H_2O with 0.1% TFA (A) and MeOH with 0.1% TFA (B) pumped at a flow rate of 1 mL/min. Absorbance was followed at 235 nm.

Table 2.2. Chiral Chromatography Parameters. Solvent systems and flow rates were individually optimized to analyze the reactions with substrates **2.1**-**2.5**.

Substrate	EtOH:Hexanes	Isocratic Flow Rate (mL/min)
1 - 3-oxopentanoyl-N-acetylcysteamine thioester	7:93	0.8
2 - (2*RS*)-methyl-3-oxobutanoyl-N-acetylcysteamine thioester	4:96	0.6
3 - (2*RS*)-methyl-3-oxopentanoyl-N-acetylcysteamine thioester	7:93	0.8
4 - (2*RS*)-methyl-3-oxohexanoyl-N-acetylcysteamine thioester	5:95	0.8
5 - (2*RS*)-ethyl-3-oxopentanoyl-N-acetylcysteamine thioester	5:95	0.8

Synthetic Protocols

General Synthetic Route to α-Substituted, β-Keto Diketide-SNAC Substrates

N-Acetylcysteamine

Cysteamine hydrochloride (3.05 g, 53.6 mmol, 1 eq.), potassium hydroxide (1.50 g, 53.6 mmol, 1 eq.), and sodium bicarbonate (6.75 g, 160.8 mmol, 3 eq.) were added to 50 mL H_2O. Acetic anhydride (2.28 mL, 48.2 mmol, 0.9 eq.) was then added dropwise over 5 min, and the solution was allowed to stir at 22°C for 15 min. The reaction was quenched with 1 M HCl, and the product was extracted with ethyl acetate (3 x 50 mL), dried with $MgSO_4$, filtered, and concentrated in vacuo, yielding a clear, viscous liquid.

Meldrum's Acid Derivatives

Pyridine (2 eq.) was added to Meldrum's acid (1 eq.) in 40 mL dry dichloromethane (DCM) at 0°C. Next, the acyl chloride (1 eq.) was added dropwise over 15 min, turning the solution a dark orange color, and the reaction was stirred overnight at 22°C. The reaction was washed with 0.1 M HCl (3 x 50 mL) and the organic layer was dried with $MgSO_4$, filtered, and concentrated in vacuo.

α-Unsubstituted Diketide-SNACs

The Meldrum's acid derivative (1 eq.) and *N*-acetylcysteamine (1 eq.) were refluxed in dry toluene for 5 hr. The resulting solution was then concentrated in vacuo.

α-Substituted Diketide-SNACs

The α-unsubstituted diketide-SNAC (1 eq.) was dissolved in 50 mL dry THF at 0°C. Potassium *tert*-butoxide (1.2 eq.) was added. After the alkyl iodide (5 eq.) was injected, the reaction was allowed to warm to 22°C and stir overnight. After the reaction was quenched with 0.1 M HCl, the product was extracted with ethyl acetate (3 x 50 mL), dried with $MgSO_4$, filtered, and concentrated in vacuo, resulting in a clear, viscous liquid.

Methylketene Dimer

Propionyl chloride (9.43 mL, 108 mmol, 1 eq.) was added dropwise to TEA (14.93 mL, 108 mmol, 1 eq.) in 250 mL dichloromethane at 0°C. The light orange reaction was stirred overnight at 22°C. The solution was concentrated twofold before 250 mL hexanes was added to cause triethylammonium chloride to crash out of solution. After the salt was filtered away, the solution was dried with $MgSO_4$ and the solvent was evaporated, yielding a viscous liquid. The byproduct propionic anhydride comprises ~15 mol% of the product.

(2RS)-Methyl-3-Oxopentanoate-D-Pantetheine Thioester (**2.6**)

Sodium borohydride (80.0 mg, 2.12 mmol, 16 eq.) was added to oxidized D-pantethine (94.0 mg, 0.17 mmol, 1 eq.) dissolved in 5 mL of 80:20 MeOH:0.25 M $NaHCO_3$ (aq.) over 20 min at 22°C. After 1 hr, the reaction was quenched by adding glacial acetic acid dropwise until bubbling ceased and was then buffered with 250 mM HEPES (pH 7.5) (final). Methylketene dimer (11.2 mg, 0.1 mmol, 1.4 eq.) was added to the solution at 22°C. After 1 hr, the reaction was evaporated. HEPES was removed via silica gel column purification using 2% MeOH:dichloromethane.

Syntheses of α-Methyl, β-Hydroxyl Diketide-SNAC Standards

(4S)-4-Benzyl-3-Propionyl-2-Oxazolidinone (**2.7**)

(4*S*)-Benzyl-2-oxazolidinone (250 mg, 1.41 mmol, 1 eq.) was dissolved in 5 mL dry THF at -78°C under nitrogen. *n*-butyllithium (0.64 mL of a 2.5 M stock in hexanes, 1.59 mmol, 1.13 eq.) was added dropwise to the cooled solution by syringe and stirred for 5 min. Propionyl chloride (0.158 mL, 1.81 mmol, 1.28 eq.) was added dropwise by syringe, and the solution was stirred at -78°C for 30 min. The reaction was warmed to room temperature

and stirred for an additional 30 min before being quenched with 3 mL saturated NH_4Cl (aq.). The product was extracted with DCM (2 x 20 mL), washed with 10 mL saturated $NaHCO_3$ (aq.) and then 10 mL brine, dried with $MgSO_4$, filtered, and concentrated in vacuo to produce a clear oil (260 mg, 80%). R_f 0.40 (20% EtOAc in hexanes).

(4S,2'S,3'R)-3-(3'-Hydroxy-2'-Methylbutanoyl)-4-Benzyl-2-Oxazolidinone (**2.8**)

(4*S*)-4-Benzyl-3-propionyl-2-oxazolidinone (**2.7**) (560 mg, 2.40 mmol, 1 eq.) was added to 10 mL dry DCM under nitrogen and cooled to 0°C. Dibutylboron triflate (Bu_2BOTf) (2.83 mL of a 1 M solution in DCM, 2.83 mmol, 1.18 eq.) was supplied via syringe. Diisopropylethylamine (0.56 mL, 3.19 mmol, 1.33 eq.) was then added dropwise over 2 min, and the solution was cooled to -78°C. After acetaldehyde (0.151 mL, 270 mmol, 1.12 eq.) was added, the solution was stirred at -78°C for 30 min, allowed to warm to 0°C, and further stirred for 2 hr. A solution of 3.3 mL Na_2HPO_4 (1 M, pH 7.4) and 8.3 mL MeOH was added to the reaction, followed by the dropwise addition of 8.5 mL 2:1 MeOH:30% H_2O_2 (aq.). After 1.5 hr, organic solvents were evaporated in vacuo and the remaining material was resuspended in 10 mL 5% $NaHCO_3$ (aq.). The product was extracted in DCM (3 x 20 mL), washed with 20 mL 5% $NaHCO_3$ (aq.) and then 20 mL brine, dried with $MgSO_4$, filtered, and concentrated in vacuo to yield a clear oil (350 mg, 53% after flash chromatography). Flash chromatography was performed with a 15%–30% EtOAc in hexanes gradient. R_f 0.38 (50% EtOAc in hexanes).

(4S,2'S,3'R)-3-(3'-Hydroxy-2'-Methylhexanoyl)-4-Benzyl-2-Oxazolidinone (**2.9**)

The synthesis for **2.8** was followed with the exception that butyrylaldehyde (0.213 mL, 2.41 mmol, 1.12 eq.) was substituted for acetaldehyde. Flash chromatography was

performed in the same manner, yielding 241 mg (33%) of the desired product. R_f 0.67 (50% EtOAc in hexanes).

(4S,2'R,3'R)-3-(3'-Hydroxy-2'-Methylbutanoyl)-4-Benzyl-2-Oxazolidinone (**2.10**) and *(4S,2'R,3'S)-3-(3'-Hydroxy-2'-Methylbutanoyl)-4-Benzyl-2-Oxazolidinone* (**2.11**)

Synthesis followed that of **2.8**, with the exception that two equivalents of Bu_2BOTf were used (2.4 mL of a 1 M solution in DCM, 2.4 mmol, 2 eq.). R_f 0.56 and 0.60 (50% EtOAc in hexanes).

(4S,2'R,3'R)-3-(3'-Hydroxy-2'-Methylhexanoyl)-4-Benzyl-2-Oxazolidinone (**2.12**) and *(4S,2'R,3'S)-3-(3'-Hydroxy-2'-Methylhexanoyl)-4-Benzyl-2-Oxazolidinone* (**2.13**)

Synthesis followed that of **2.9**, with the exception that two equivalents of Bu_2BOTf were added (4.32 mL of a 1 M solution in DCM, 4.32 mmol, 2 eq.). R_f 0.73 and 0.77 (50% EtOAc in hexanes).

(2S,3R)-3-Hydroxy-2-Methylbutanoyl-N-Acetylcysteamine Thioester (**2.14**)

Aldol product **2.8** (187 mg, 0.68 mmol) was dissolved in 3 mL of a 4:1 mixture of THF:H_2O and cooled to 0°C. Slowly, 3 mL of 30% H_2O_2 was added, followed by 1 mL saturated LiOH (0.5 g LiOH in 1 mL H_2O). The reaction was left to stir at 0°C for 2 hr, at which point Na_2SO_3 was added until H_2O_2 was quenched (bubbling ceased). The aqueous phase was washed with DCM (2 x 15 mL), and the DCM layers were back-extracted with H_2O (1 x 15 mL). The combined aqueous layers were cooled to 0°C and the pH was lowered to 1 with 6 M HCl. The aldol acid was extracted into EtOAc (3 x 20 mL), dried with $MgSO_4$, filtered, and concentrated in vacuo (as the aldol acid is somewhat volatile, the solution was

only concentrated to 1 mL). 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (211.1 mg, 1.36 mmol, 2 eq.) was added to the solution under nitrogen at 0°C. NAC (96.8 mg, 0.81 mmol, 1.2 eq.) and a few crystals of 4-dimethylaminopyridine were added. After 4 hr at 0°C, the reaction was quenched with 5 mL H_2O. Product was extracted into EtOAc (3 x 15 mL), washed with 5 mL brine, dried with $MgSO_4$, and filtered. A yellowish oil was produced after concentration in vacuo (50 mg, 47%). R_f 0.20 (EtOAc).

(2S,3R)-3-Hydroxy-2-Methylhexanoyl-N-Acetylcysteamine Thioester (**2.15**)

2.15 was synthesized in the same manner as **2.14** from **2.9** (81%). R_f 0.35 (in EtOAc).

(2R,3R)-3-Hydroxy-2-Methylbutanoyl-N-Acetylcysteamine Thioester (**2.16**) and

(2R,3S)-3-Hydroxy-2-Methylbutanoyl-N-Acetylcysteamine Thioester (**2.17**)

2.16 and **2.17** were synthesized in the same manner as **2.14** from a mixture of **2.10** and **2.11** (33%, **2.16** + **2.17**). They were then separated by reverse-phase HPLC on a semipreparative C18 column (BondClone, 300 x 7.8 mm; Phenomenex) connected to a Waters 1525 HPLC system through a gradient of 15%-100% B over 30 min, where the mobile phases were (A) H_2O with 0.1% TFA and (B) MeOH with 0.1% TFA, pumped at a flow rate of 2 mL/min. R_f 0.15 and 0.20 (in EtOAc).

(2R,3R)-3-Hydroxy-2-Methylhexanoyl-N-Acetylcysteamine Thioester (**2.18**) and

(2R,3S)-3-Hydroxy-2-Methylhexanoyl-N-Acetylcysteamine Thioester (**2.19**)

2.18 and **2.19** were synthesized in the same manner as **2.14** from a mixture of **2.12** and **2.13** (85%, **2.18** + **2.19**). They were then separated as described for **2.16** and **2.17**. R_f 0.30 and 0.35 (in EtOAc).

Characterization via NMR and LC-MS

NMR was performed either on a Varian Mercury 400 MHz or a Varian INOVA 500 MHz instrument. LC-MS was performed on an Agilent Technologies 1200 Series HPLC with a Gemini C18 column (5 μm, 2 x 50 mm; Phenomenex) coupled to an Agilent Technologies 6130 quadrupole mass spectrometer system equipped with an electrospray-ionization source. A 5%-95% B gradient over 12 min at a flow rate of 0.7 mL/min was run in which the mobile phases were (A) H_2O with 0.1% formic acid and (B) acetonitrile with 0.1% formic acid.

3-Oxopentanoyl-N-Acetylcysteamine Thioester (**2.1**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.08 (t, 3H, J = 7.4 Hz, CH_3), 1.96 (s, 3H, CH_3-C=O), 2.55 (q, 2H, J = 7.4 Hz, C-CH_2-C=O), 3.00–3.10 (m, 2H, S-CH_2), 3.4–3.5 (m, 2H, N-CH_2), 3.7 (s, 2H, O=C-CH_2-C=O), 5.88 (br s, 1H, NH). ESI-MS expected mass: 218.3; observed mass: 218.2.

3-Oxobutanoyl-N-Acetylcysteamine Thioester

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.92 (s, 3H, CH_3-C(=O)-N), 2.20 (s, 3H, CH_3-C=O), 2.93–3.16 (m, 2H, S-CH_2), 3.36–3.42 (m, 2H, N-CH_2), 3.62 (s, 2H, O=C-CH_2-C=O), 5.99 (br s, 1H, NH).

(2RS)-Methyl-3-Oxobutanoyl-N-Acetylcysteamine Thioester (**2.2**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.33 (d, 3H, J = 7.5 Hz, CH_3-C-$(C=O)_2$), 1.90 (s, 3H, CH_3-C(=O)-N), 2.19 (s, 3H, CH_3-C=O), 2.95–3.08 (m, 2H, S-CH_2), 3.30–3.46 (m, 2H, N-CH_2),

3.71 (q, 1H, J = 7.5 Hz, O=C-CH-C=O), 5.87 (br s, 1H, NH). ESI-MS expected mass: 218.3; observed mass: 218.4.

(2RS)-Methyl-3-Oxopentanoyl-N-Acetylcysteamine Thioester (**2.3**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.08 (t, 3H, J = 7.5 Hz, CH_3), 1.38 (d, 3H, J = 7.5 Hz, CH_3-C-$(C=O)_2$), 1.96 (s, 3H, CH_3-C=O), 2.48–2.68 (m, 2H, C-CH_2-C=O), 3.00–3.14 (m, 2H, S-CH_2), 3.36–3.52 (m, 2H, N-CH_2), 3.79 (q, 1H, J = 7.5 Hz, O=C-CH-C=O), 5.92 (br s, 1H, NH). ESI-MS expected mass: 232.3; observed mass: 232.4.

3-Oxohexanoyl-N-Acetylcysteamine Thioester

^{1}H NMR (400 MHz, $CDCl_3$) δ 0.90–1.00 (m, 3H, CH_3), 1.55–1.70 (m, 2H, C-CH_2-C), 1.96 (s, 3H, CH_3-C=O), 2.48–2.60 (m, 2H, C-CH_2-C=O), 3.00–3.14 (m, 2H, S-CH_2), 3.40–3.52 (m, 2H, N-CH_2), 3.70 (s, 2H, O=C-CH_2-C=O), 5.92 (br s, 1H, NH).

(2RS)-Methyl-3-Oxohexanoyl-N-Acetylcysteamine Thioester (**2.4**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 0.85 (t, 3H, J = 7.6 Hz, CH_3), 1.33 (d, 3H, J = 7.9 Hz, CH_3-C-$(C=O)_2$), 1.50–1.60 (m, 2H, C-CH_2-C), 1.98 (s, 3H, CH_3-C=O), 2.45 (m, 2H, C-CH_2-C=O), 2.95–3.12 (m, 2H, S-CH_2), 3.34–3.52 (m, 2H, N-CH_2), 3.74 (q, 1H, J = 7.9 Hz, O=C-CH-C=O), 5.92 (br s, 1H, NH). ESI-MS expected mass: 246.3; observed mass: 246.4.

(2RS)-Ethyl-3-Oxopentanoyl-N-Acetylcysteamine Thioester (**2.5**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 0.94 (t, 3H, J = 7.3 Hz, CH_3), 1.17 (t, 3H, J = 8.4 Hz, CH_3-C-C=O), 1.88–1.98 (m, 2H, C-CH_2-C-$(C=O)_2$), 2.15 (s, 3H, CH_3-C=O), 2.48–2.66 (m, 2H,

C-CH_2-C=O), 3.01–3.20 (m, 2H, S-CH_2), 3.41–3.58 (m, 2H, N-CH_2), 3.68–3.74 (m, 1H, O=C-CH-C=O), 6.27 (br s, 1H, NH). ESI-MS expected mass: 246.3; observed mass: 246.4.

(*2RS)-Methyl-3-Oxopentanoate-D-Pantetheine Thioester* (**2.6**)

^{1}H NMR (400 MHz, D_2O) δ 0.70-0.78 (d, 6H, J = 16 Hz, $(CH_3)_2$-C), 0.81-0.88 (t, 3H, J = 8 Hz, CH_3-C), 1.20-1.25 (d, 3H, J = 8 Hz, CH_3-C-$(C=O)_2$), 2.28-2.33 (m, 2H, CH_2-C(=O)-N) 2.51-2.58 (q, 2H, J = 8 Hz, C-CH_2-C=O), 2.90-2.97 (t, 2H, J = 7.2 Hz, CH_2-S), 3.20-3.26 (m, 2H, S-C-CH_2), 3.30-3.40 (m, 2H, N-(C=O)-C-CH_2), 3.40-3.45 (m, 2H, O-CH_2), 3.81 (s, 1H, O=C-CH-O), 4.00-4.05 (q, 1H, J = 8 Hz, $(O=C)_2$-CH). ESI-MS expected mass: 391.5; observed mass: 391.2.

(4S)-4-Benzyl-3-Propionyl-2-Oxazolidinone (**2.7**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.22 (t, 3H, J = 8.0 Hz, CH_3), 2.76–2.82 (m, 1H, one of CH_2-Ph), 2.90–3.05 (m, 2H, CH_2-C=O), 3.31 (dd, 1H, J = 15.2 Hz, J = 5.5 Hz, one of CH_2-Ph), 4.16–4.23 (m, 2H, CH_2-O), 4.66–4.70 (m, 1H, CH-N).

(4S,2'S,3'R)-3-(3'-Hydroxy-2'-Methylbutanoyl)-4-Benzyl-2-Oxazolidinone (**2.8**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.23 (d, 3H, J = 7.4 Hz, CH_3), 1.28 (d, 3H, J = 6.1 Hz, CH_3-C-C=O), 2.76–2.82 (m, 1H, one of CH_2-Ph), 2.93 (br s, 1H, OH), 3.26 (dd, 1H, J = 12.3 Hz, J = 3.3 Hz, one of CH_2-Ph), 3.75 (dq, 1H, J = 9.8 Hz, J = 3.1 Hz, CH), 4.16–4.23 (m, 2H, CH_2-O), 4.65–4.69 (m, 1H, CH-N).

(4S,2'S,3'R)-3-(3'-Hydroxy-2'-Methylhexanoyl)-4-Benzyl-2-Oxazolidinone (**2.9**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 0.94 (t, 3H, J = 7.2 Hz, CH_3), 1.30 (d, 3H, CH_3-C-C=O), 1.35–1.45 (m, 1H, one of C-CH_2-C), 1.45–1.58 (m, 1H one of C-CH_2-C), 2.75–2.81 (m, 1H, one of CH_2-Ph), 2.84 (d, 1H, OH), 3.26 (dd, 1H, J = 13.7 Hz, J = 3.1 Hz, one of CH_2-Ph), 3.76 (dq, 1H, J = 7.2 Hz, J = 2.6 Hz, CH), 3.95–4.00 (m, 1H, CH-C=O), 4.17–4.27 (m, 2H, CH_2-O), 4.68–4.74 (m, 1H, CH-N).

(4S,2'R,3'R)-3-(3'-Hydroxy-2'-Methylbutanoyl)-4-Benzyl-2-Oxazolidinone (**2.10**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.22 (d, 3H, J = 5.6 Hz, CH_3), 1.32 (d, 3H, J = 6.2 Hz, CH_3-C-C=O), 2.56 (d, 1H, J = 7.4 Hz one of CH_2-Ph), 2.82 (br s, 1H, OH), 3.28–3.35 (m, 1H, one of CH_2-Ph), 3.90–3.95 (m, 1H, CH), 3.96–4.00 (m, 1H, CH-C=O), 4.15–4.25 (m, 2H, CH_2-O), 4.67–4.72 (m, 1H, CH-N).

(4S,2'R,3'R)-3-(3'-Hydroxy-2'-Methylhexanoyl)-4-Benzyl-2-Oxazolidinone (**2.12**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 0.94 (t, 3H, CH_3), 1.20 (d, 3H, J = 7.2 Hz, CH_3-C-C=O), 1.35–1.45 (m, 1H, one of C-CH_2-C), 1.45–1.58 (m, 1H one of C-CH_2-C), 2.75–2.82 (m, 1H, one of CH_2-Ph), 2.91 (d, 1H, J = 7.2 Hz, OH), 3.31 (dd, 1H, J = 7.9 Hz, J = 3.1 Hz, one of CH_2-Ph), 3.85 (dq, 1H, J = 7.2 Hz, J = 2.7 Hz, CH), 3.95–4.04 (m, 1H, CH-C=O), 4.15–4.26 (m, 2H, CH_2-O), 4.66–4.74 (m, 1H, CH-N).

(2S,3R)-3-Hydroxy-2-Methylbutanoyl-N-Acetylcysteamine Thioester (**2.14**)

^{1}H NMR (500 MHz, $CDCl_3$) δ 1.22 (d, 3H, J = 7.6 Hz, CH_3), 1.24 (d, 3H, CH_3-C-C=O), 2.01 (s, 3H, CH_3-C=O), 2.70–2.75 (m, 1H, CH-C=O), 3.00–3.10 (m, 2H, S-CH_2), 3.42–3.53 (m, 2H, N-CH_2), 3.70–3.79 (m, 1H, CH), 5.93 (br s, 1H, NH). ESI-MS expected mass: 220.3; observed mass: 220.2.

(2S,3R)-3-Hydroxy-2-Methylhexanoyl-N-Acetylcysteamine Thioester (**2.15**)

^{1}H NMR (500 MHz, $CDCl_3$) δ 0.94 (t, 3H, J = 7.7 Hz, CH_3), 1.22 (d, 3H, J = 7.7 Hz, CH_3-C-C=O), 1.50–1.60 (m, 1H, one of C-CH_2-C), 1.60–1.70 (m, 1H one of C-CH_2-C), 1.95 (s, 3H, CH_3-C=O), 2.70–2.75 (m, 1H, CH-C=O), 3.00–3.08 (m, 2H, S-CH_2), 3.40–3.52 (m, 2H, N-CH_2), 3.92–3.98 (m, 1H, J = 5.4 Hz, CH), 5.86 (br s, 1H, NH). ESI-MS expected mass: 248.3; observed mass: 248.4.

(2R,3R)-3-Hydroxy-2-Methylbutanoyl-N-Acetylcysteamine Thioester (**2.16**)

^{1}H NMR (500 MHz, $CDCl_3$) δ 1.22 (d, 3H, J = 7.7 Hz, CH_3), 1.24 (d, 3H, CH_3-C-C=O), 1.96 (s, 3H, CH_3-C=O), 2.70–2.76 (m, 1H, CH-C=O), 3.01–3.14 (m, 2H, S-CH_2), 3.47–3.58 (m, 2H, N-CH_2), 3.79–3.83 (m, 1H, CH), 6.26 (br s, 1H, NH). ESI-MS expected mass: 220.3; observed mass: 220.2.

(2R,3R)-3-Hydroxy-2-Methylhexanoyl-N-Acetylcysteamine Thioester (**2.18**)

^{1}H NMR (500 MHz, $CDCl_3$) δ 0.94 (t, 3H, J = 7.2 Hz, CH_3), 1.22 (d, 3H, J = 6.9 Hz, CH_3-C-C=O), 1.38–1.46 (m, 1H, one of C-CH_2-C), 1.46–1.56 (m, 1H one of C-CH_2-C), 2.01 (s, 3H, CH_3-C=O), 2.72–2.79 (m, 1H, CH-C=O), 3.01–3.10 (m, 2H, S-CH_2), 3.43–3.55 (m, 2H, N-CH_2), 3.72–3.78 (m, 1H, CH), 6.07 (br s, 1H, NH). ESI-MS expected mass: 248.3; observed mass: 248.4.

ACKNOWLEDGEMENTS

The authors thank the Robert A. Welch Foundation for funding (F-1712 for A.T.K.-C. and F-1694 for D.R.S.).

Chapter 3. Monitoring Biocatalytic Transformations Mediated by Polyketide Synthase Enzymes in Cell Lysate via Fluorine NMR

SUMMARY

The biocatalytic employment of modular polyketide synthase enzymes in cell lysate has become a viable route to preparative quantities of synthetically valuable polyketide fragments. We report the quantitative, uninvasive, and continuous monitoring of such biocatalytic reactions by observing trifluoromethyl-bearing substrates via ^{19}F NMR spectroscopic analysis. To demonstrate the utility of this technique, we followed reactions catalyzed by a thioesterase and several ketoreductases.

INTRODUCTION

Advances in employing isolated PKS enzymes as biocatalysts have enabled access to preparative quantities of polyketide fragments (Piasecki et al., 2011; Hughes and Keatinge-Clay, 2012). These biocatalytic reactions are performed in the cell lysate of the overexpression host (e.g., *Escherichia coli*), thus maximizing the quantity of enzyme and avoiding resource-intensive protein purification. Because cell lysate is a complex mixture of all the soluble biomolecules produced by the overexpression host, other fates are possible for a substrate entered into such a system in addition to the desired transformation. To evaluate such biocatalytic reactions we sought a quantitative and uninvasive technique that could continuously monitor both substrates and products.

To these ends, NMR spectroscopy would be useful; however, observing the conversion of substrate into product via ^{1}H NMR spectroscopic analysis is complicated by the high background generated by the hydrogen-containing components of biocatalytic reactions

(e.g., biomolecules, co-substrates, buffering agent, glycerol, DMSO) and by significant noise from water even when suppression methods are employed. However, if the substrates contained a trifluoromethyl group and were observed by ^{19}F NMR spectroscopic analysis, no background would interfere with measurements because *E. coli* does not incorporate fluorine into any of its biomolecules (few organisms do) (Murphy et al., 2003). Furthermore, the ^{19}F nucleus is (i) 83% as intrinsically sensitive as ^{1}H, (ii) 100% naturally abundant, (iii) isosteric with hydrogen, and (iv) very responsive to its electronic environment (Murphy, 2007). These properties have been exploited in drug discovery efforts; one technique, called FABS (Fluorine Atoms for Biochemical Screening), identifies inhibitors by detecting a decrease in the rate of conversion from a fluorinated substrate into its product by ^{19}F NMR spectroscopic analysis (Dalvit et al., 2003).

RESULTS AND DISCUSSION

EryTE Hydrolysis Can be Observed in Real-Time by ^{19}F NMR

First, we observed a hydrolysis reaction mediated by the erythromycin thioesterase (EryTE), which is known to catalyze the hydrolysis of acyl thioesters (Gokhale et al., 1999; Sharma and Boddy, 2007; Hughes and Keatinge-Clay, 2012). We sought to determine how accurately the kinetics of EryTE-mediated hydrolysis could be measured within the cell lysate by ^{19}F NMR spectroscopic analysis compared to the more traditional technique of employing HPLC and a UV detector. Thus, 3,3,3-trifluoropropionyl-*S*-*N*-acetylcysteamine (NAC; **3.1**) was incubated in EryTE-containing cell lysate both in an NMR tube and in a separate vessel so that for every spectrum acquired, a sample was also quenched for later HPLC analysis. ^{19}F NMR spectroscopic analysis (without proton decoupling) yielded a

triplet for each species due to splitting of the fluorine resonance by the adjacent methylene hydrogens. The triplet of substrate **3.1** appeared at δ = -63.00 ppm, while the 3,3,3-trifluoropropionate product (**3.2**) showed triplets at δ = -63.55 and -63.70 ppm, possibly due to two different interactions with counterions, or formation of a glycerolysis adduct (Hughes et al., 2013) (**Figure 3.1**). Kinetic characterization was performed by measuring the change in concentration of **3.1** by both ^{19}F NMR spectroscopic analysis and reverse-phase HPLC. The determined k_{cat} and K_m parameters agreed within the error limits. Thus, ^{19}F NMR spectroscopic analysis accurately measured the kinetics of EryTE on **3.1** in cell lysate (k_{cat} = 0.077 ± 0.010 s^{-1}, K_m = 39.3 ± 4.5 mM, k_{cat}/K_m = 1.97 ± 0.14 $M^{-1}s^{-1}$). Previous kinetic analysis of EryTE towards similar, unfluorinated thioester substrate analogs gave comparable values- one unreduced diketide *S*-NAC had a k_{cat}/K_m value of 4.2 ±

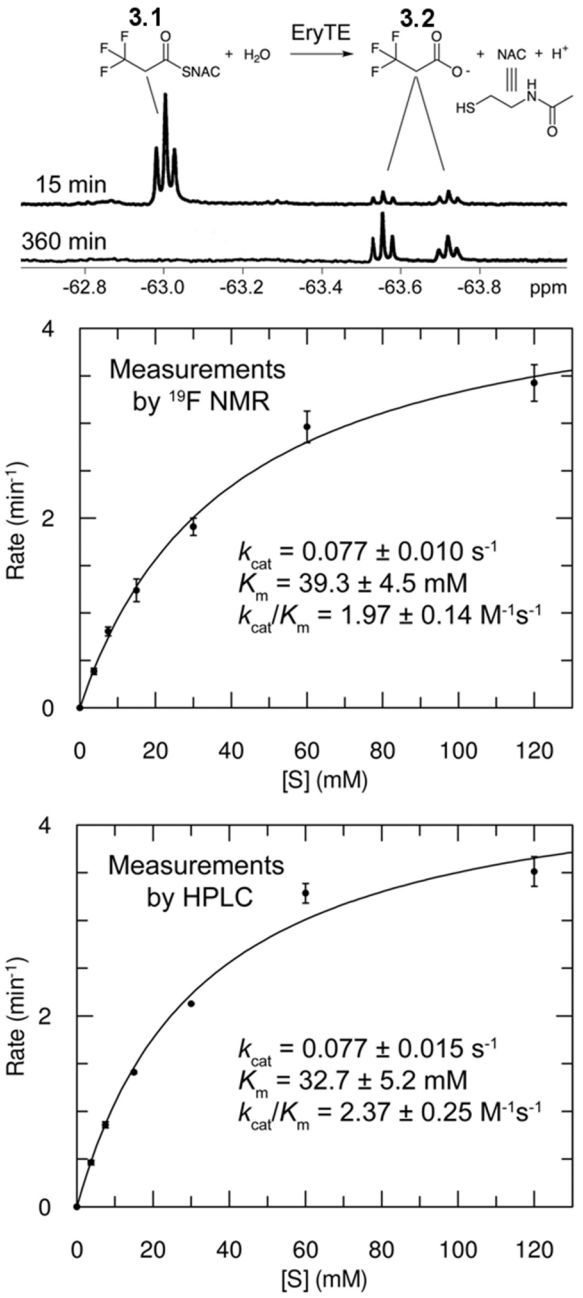

Figure 3.1. Accurate kinetic characterization of EryTE-mediated hydrolysis of fluorinated **3.1** in cell lysate via ^{19}F NMR spectroscopic analysis.

0.4 $M^{-1}s^{-1}$ (Gokhale et al., 1999; Sharma et al., 2007). Kinetic analysis is not only more facile by ^{19}F NMR spectroscopic analysis than by HPLC, but also more informative: generation of the nonchromophoric product **3.2** was observed in the EryTE reaction.

TylKR1 Reduction Can be Observed in Real-Time by ^{19}F NMR

We next sought to monitor a more advanced biocatalytic transformation driven by an NADPH-regeneration system comprised of glucose dehydrogenase, $NADP^+$, and glucose (Siskos et al., 2005; Keatinge-Clay, 2007; Valenzano et al., 2009; Piasecki et al., 2011). Thus, the conversion of 3-oxo-5,5,5-trifluoropentanoyl-*S*-NAC (**3.3**) into (3*R*)-hydroxy-5,5,5-trifluoro-pentanoyl-*S*-NAC (**3.4**) by the KR from the first module of the tylosin PKS (TylKR1) was observed by ^{19}F NMR spectroscopic analysis (**Figure 3.2**). The triplet from substrate **3.3** appeared at δ = -63.18 ppm, while the triplet from product **3.4** (confirmed by mass spectrometry and NMR analysis) appeared at δ = -63.60 ppm. The rate of the reaction measured by ^{19}F NMR spectroscopic analysis concurred with the rate measured by HPLC (at 30 mM of **3.3**, V_o was calculated to be 0.20 ± 0.02 mM·min^{-1} by ^{19}F NMR spectroscopic analysis and 0.23 ± 0.05 mM·min^{-1} by HPLC), approximately matching the rate anticipated from previous kinetic studies of TylKR1 on (2*RS*)-methyl-3-oxopentanoyl-*S*-NAC (0.81 mM·min^{-1}) (Siskos et al., 2005) (for data, see Appendix).

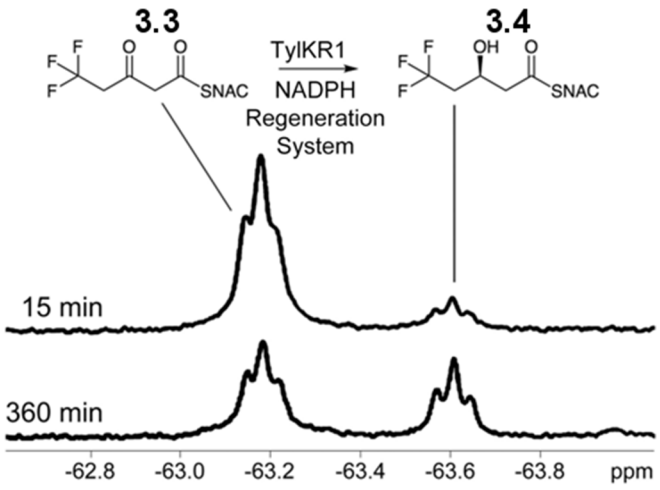

Figure 3.2. Monitoring biocatalysis mediated by TylKR1 in cell lysate.

Diastereomers from Ketoreduction Can be Differentiated in Real-Time by ^{19}F NMR

We then sought to determine whether the terminal fluorine atoms in an ε-trifluoro, α-substituted β-ketothioester were sensitive enough to distinguish stereochemical differences generated at the α- and β-carbon atoms by stereocontrolled, KR-mediated reductions. Thus, 2-methyl-3-oxo-5,5,5-trifluoropentanoyl-*S*-NAC (**3.5**) was incubated with each of the four KR types: AmpKR2 (the KR from the second module of the amphotericin PKS), RifKR7 (the KR from the seventh module of the rifamycin PKS), TylKR1, and EryKR1 (the KR from the first module of the erythromycin PKS), which are known to set (2*R*,3*S*), (2*S*,3*S*), (2*R*,3*R*), and (2*S*,3*R*) stereochemistries, respectively, when reducing 2-methyl-3-oxopentanoyl-*S*-NAC (Siskos et al., 2005; Keatinge-Clay, 2007; Valenzano et al., 2009 and 2010; Piasecki et al., 2011) (**Figure 3.3**). The triplet for substrate **3.5** appeared at δ = -63.23 ppm. Notably, the triplets for the *syn*- and *anti*-products were well-resolved (0.34 ppm); the triplet for the *syn*-products **3.6** and **3.9** appeared at δ = -63.68 ppm, and the triplet for *anti*- products **3.7** and

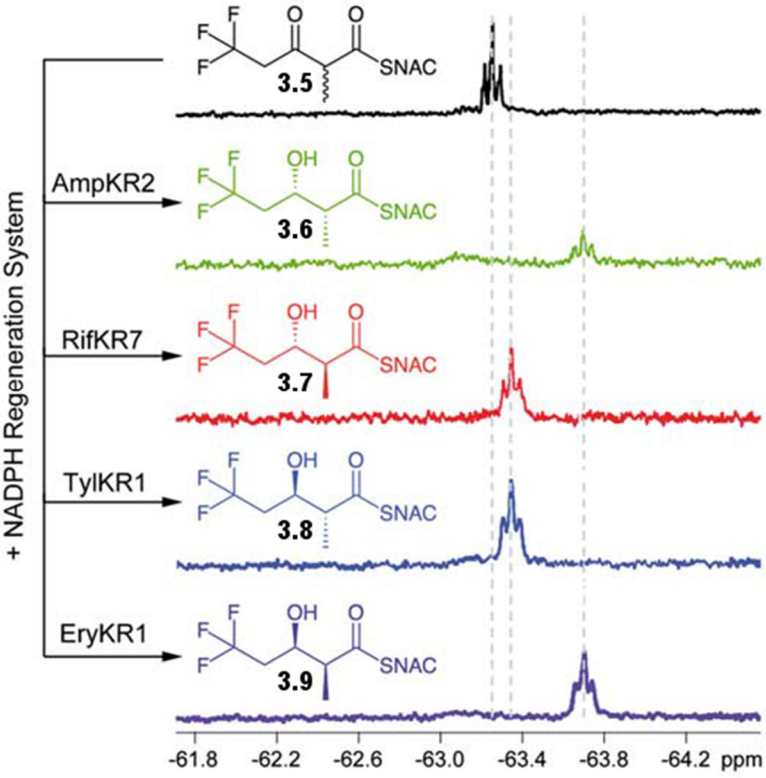

Figure 3.3. ^{19}F NMR resolution of diastereomers generated by stereocontrolled, KR-mediated biocatalysis in cell lysate.

3.8 appeared at δ = -63.34 ppm. That the diastereomers could be readily distinguished within cell lysate highlights the sensitivity of this technique.

SIGNIFICANCE

Whether the trifluoromethyl-bearing substrate analogs would be processed in the same manner as the nonfluorinated substrates normally entered into these biocatalytic reactions was initially unclear; however, EryTE and the KRs were shown to catalyze the desired transformations, stereoselectively in the case of the KRs (Piasecki et al., 2011). Studies reporting the acceptance of monofluorinated thioesters by the erythromycin PKS and the studies presented here suggest that PKS enzymes are generally tolerant towards fluorinated substrates and that biocatalytic syntheses of polyketides can be reliably evaluated by ^{19}F NMR spectroscopic analysis (Goss and Hong, 2005; Ashley et al., 2006; Ward et al., 2007). The ability to follow nonchromophoric substrates by ^{19}F NMR spectroscopic analysis enables monitoring of substrates and products in biocatalytic reactions generating reduced polyketides such as triketide lactone chiral building blocks. The rapid detection of thioester hydrolysis or other undesired transformations will facilitate the optimization of biocatalytic reactions performed in cell lysate.

We also seek to observe and optimize biocatalytic transformations mediated by other PKS enzymes such as dehydratases (DHs), which convert β-hydroxy intermediates into α,β-unsaturated intermediates, enoylreductases (ERs), which stereoselectively reduce α,β-unsaturated intermediates, ketosynthases (KSs), which form a carbon-carbon bond to elongate an intermediate, and PKS modules, which both elongate and process intermediates. The monitoring of whole-cell biocatalysis in which fluorinated precursors are fed to microbes expressing PKS enzymes may also be possible through in-cell NMR

spectroscopic analysis (Stevens et al., 1984; Serber et al., 2001). In summary, we have established a facile and powerful technique to monitor reactions in complex media that will aid in the optimization of biocatalytic transformations that generate preparative quantities of complex polyketides.

EXPERIMENTAL PROCEDURES

Synthetic Protocols

Fluorinated Substrates 3.1, 3.3, and 3.5

N-Acetylcysteamine

N-Acetylcysteamine was prepared by combining cysteamine hydrochloride (2 g, 17.6 mmol, 1 eq.), sodium bicarbonate (4.4 g, 52.8 mmol, 3 eq.), and potassium hydroxide (988 mg, 17.6 mmol, 1 eq.) in that order to 100 mL H_2O at 22°C. Acetic anhydride (832 µL, 8.8 mmol, 0.5 eq.) was added dropwise over 5 min, and the reaction was stirred for 10 min. The reaction was then acidified to pH 7. The product was extracted with EtOAc (3 x 200 mL), dried over $MgSO_4$, filtered, and dried in vacuo to yield a clear oil.

3,3,3-Trifluoropropionyl-N-Acetylcysteamine Thioester (**3.1**)

N-Acetylcysteamine (522 mg, 4.4 mmol, 1 eq.) was dissolved in 10 mL dry DCM at 0°C under N_2. 3,3,3-Trifluoropropionyl chloride (452 µL, 4.4 mmol, 1 eq.) and triethylamine (731 µL, 5.3 mmol, 1.2 eq.) were added, and the reaction was left to warm to room temperature while stirring overnight. The reaction was washed with brine (3 x 20 mL) and the brine was back-extracted with DCM (3 x 20 mL). The organic layers were combined

and dried in vacuo to yield a brown oil (800 mg, 79.5%). TLC analysis of the fractions was carried out in 100% EtOAc (R_f 0.40).

3-Oxo-5,5,5-Trifluoropentanoyl-N-Acetylcysteamine Thioester (**3.3**)

Meldrum's acid (1.0 g, 6.9 mmol, 1 eq.) was dissolved in 10 mL dry DCM under N_2. Pyridine (1.2 mL, 14.2 mmol, 2.05 eq.) was added, and the reaction was cooled to 0°C. 3,3,3-Trifluoropropionyl chloride (860 µL, 8.3 mmol, 1.2 eq.) was dissolved in 2 mL dry DCM and added dropwise to the Meldrum's acid solution over 1 hr. The reaction was stirred at 0°C for an additional hour and warmed to room temperature overnight. The reaction was diluted with 10 mL DCM, washed with 10% aqueous HCl (3 x 50 mL), dried with $MgSO_4$, filtered, and evaporated to yield a dark orange oil (1.2 g, 72.5%). The Meldrum's acid derivative (1.6 g, 6.3 mmol, 1 eq.) was dissolved in 100 mL dry toluene under N_2. *N*-Acetylcysteamine (750 mg, 6.3 mmol, 1 eq.) was added, and the reaction was refluxed overnight. The reaction was evaporated to yield a dark orange oil. The product was purified via flash chromatography with a ~3 cm Cu^{2+}-impregnated silica gel layer on top with a 50-100% EtOAc:hexanes gradient to yield an orange powder after evaporation (572 mg, 33%). TLC analysis of the fractions was carried out in 100% EtOAc (R_f 0.35).

(2RS)-Methyl-3-Oxo-5,5,5-Trifluoropentanoyl-N-Acetylcysteamine Thioester (**3.5**)

Thioester **3.3** (200 mg, 0.74 mmol, 1 eq.) was dissolved in 40 mL dry THF and cooled to 0°C. Potassium *tert*-butoxide (99.4 mg, 0.89 mmol, 1.2 eq.) was added, and the flask was placed under N_2. Iodomethane (230.3 µL, 3.7 mmol, 5 eq.) was added dropwise over 5 min, and the reaction was warmed to room temperature while stirring overnight. The pH of the reaction was neutralized, H_2O was added to create an aqueous layer, and the product was

extracted with EtOAc (3 x 200 mL). The organic layer was dried with $MgSO_4$, filtered, and evaporated to yield a dark orange oil. The product was purified over a silica gel column using a 50-100% EtOAc:hexanes gradient to yield a dark orange oil (135 mg, 64%). TLC analysis of the fractions was performed in 100% EtOAc (R_f 0.37).

Characterization by ^{1}H NMR, ^{13}C NMR, ^{19}F NMR, and LC-MS

NMR characterization was performed on a Varian Mercury 400 MHz instrument. LC-MS was performed on an Agilent Technologies 1200 HPLC with a Gemini C18 column (5 μm, 2 x 50 mm; Phenomenex). This was coupled to an Agilent Technologies 6130 quadrupole mass spectrometer system containing an ESI source run in positive mode. The aqueous mobile phase was (A) H_2O with 0.1% formic acid, and the organic mobile phase was (B) acetonitrile with 0.1% formic acid. A 5-95% B gradient was run over 12 min with a flow rate of 0.7 mL/min.

3,3,3-Trifluoropropanoyl-N-Acetylcysteamine Thioester (**3.1**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.97 (s, 3H, CH_3-C=O), 3.10-3.15 (m, 2H, S-CH_2), 3.39 (q, 2H, J = 9.97 Hz, CF_3-CH_2), 3.45-3.50 (m, 2H, N-CH_2), 5.84 (br s, 1H, NH).

^{13}C NMR (400 MHz, $CDCl_3$) δ 22.10 (s, $\underline{C}H_3$-C=O), 29.10 (s, S-CH_2), 38.95 (s, N-CH_2), 45.49-46.00 (m, F_3C-$\underline{C}H_2$), 122.97 (q, J = 277.53 Hz, CF_3), 170.85 (s, CH_3-$\underline{C}$=O), 189.15 (s, S-C=O).

^{19}F NMR (400 MHz, $CDCl_3$) δ -62.79 (t, 3F, J = 9.95 Hz, CF_3). See also **Figures 3.4A and 3.5A**. ESI-MS expected mass: 230.2; observed mass: 230.2.

3-Oxo-5,5,5-Trifluoropentanoyl-N-Acetylcysteamine Thioester (**3.3**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.98 (s, 3H, CH_3-C=O), 2.99 (q, 2H, J = 6.77 Hz, CF_3-CH_2, keto), 3.10-3.15 (m, 2H, S-CH_2), 3.40-3.45 (m, 2H, CF_3-CH_2, enol), 3.45-3.50 (m, 2H, N-CH_2), 3.80 (s, 2H, O=C=CH_2-C=O), 5.62 (s, 1H, HO-C=CH-C=O, enol), 5.84 (br s, 1H, NH).

^{13}C NMR (400 MHz, $CDCl_3$) δ 22.12 (s, $\underline{C}H_3$-C=O), 27.10 (s, S-CH_2), 38.42 (s, N-CH_2), 38.50-38.55 (m, F_3C-$\underline{C}H_2$, enol), 45.35-45.40 (m, F_3C-$\underline{C}H_2$, keto), 56.19 (s, O=C-$\underline{C}H_2$-C=O, keto), 101.53 (s, HOC=$\underline{C}$H-C=O, enol), 122.76 (q, J = 277.53 Hz, CF_3, enol), 122.86 (q, J = 277.53 Hz, CF_3, keto), 169.68 (s, N-C=O), 190.47 (s, CH_2-$\underline{C}$OH=CH, enol), 191.14 (s, S-C=O), 193.96 (s, CH_2-$\underline{C}$O-CH_2, keto).

^{19}F NMR (400 MHz, $CDCl_3$) δ -63.57 (t, J = 10.03 Hz, CF_3, enol), -62.45 (t, 3F, J = 10.12 Hz, CF_3, keto). See also **Figures 3.4B and 3.5B**. ESI-MS expected mass: 272.3; observed mass: 272.0.

(3R)-Hydroxy-5,5,5-Trifluoropentanoyl-N-Acetylcysteamine Thioester (**3.4**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.98 (s, 3H, CH_3-C=O), 2.34 (m, 2H, CF_3-CH_2), 2.81 (m, 2H, J = 4.06 Hz, HO-C-CH_2-C=O), 3.02-3.08 (m, 2H, S-CH_2), 3.43-3.50 (m, 2H, N-CH_2), 4.45 (br s, 1H, OH-CH), 5.83 (br s, 1H, NH).

^{19}F-NMR (400 MHz, DMSO-d_6) δ -63.6 (t, J = 10.71 Hz, CF_3). ESI-MS expected mass: 274.3; observed mass: 274.2.

(2RS)-Methyl-3-Oxo-5,5,5-Trifluoropentanoyl-N-Acetylcysteamine Thioester (**3.5**)

^{1}H NMR (400 MHz, $CDCl_3$) δ 1.44 (d, 3H, J = 7.06 Hz, O=C-C-CH_3, keto), 1.97 (s, 3H, O=C-CH_3), 3.11 (q, 2H, J = 6.77 Hz, CF_3-CH_2), 3.10-3.15 (m, 2H, S-CH_2), 3.40-3.45 (m,

2H, N-CH_2), 3.47 (s, 3H, HO-C=C-CH_3, enol), 3.85 (q, 1H, J = 7.2, O=C-CH, keto), 6.20 (br s, 1H, NH).

^{13}C NMR (400 MHz, $CDCl_3$) δ 12.29 (s, O=C-CH-$\underline{C}H_3$), 22.06 (s, $\underline{C}H_3$-C=O), 28.13 (s, S-$\underline{C}H_2$), 36.55-36.60 (m, F_3C-$\underline{C}H_2$, enol), 37.95 (s, N-$\underline{C}H_2$), 43.60-43.65 (m, F_3C-$\underline{C}H_2$, keto), 60.41 (s, O=C-$\underline{C}$H-C=O, keto), 107.11 (s, HO-C=$\underline{C}$-C=O), enol), 117.90 (q, J = 277.53 Hz, $\underline{C}F_3$, enol), 127.41 (q, J = 277.53 Hz, $\underline{C}F_3$, keto), 169.77 (s, HN-$\underline{C}$=O), 193.76 (s, HO-$\underline{C}$, enol), 195.37 (s, S-$\underline{C}$=O), 197.46 (s, CH_2-$\underline{C}$O-C , keto).

^{19}F NMR (400 MHz, $CDCl_3$) δ -62.91 (t, 3F, J = 10.12 Hz, CF_3, enol), -62.54 (t, 3F, J = 10.03 Hz, CF_3, keto). See also **Figures 3.4C and 3.5C**. ESI-MS expected mass: 286.3; observed mass: 286.0.

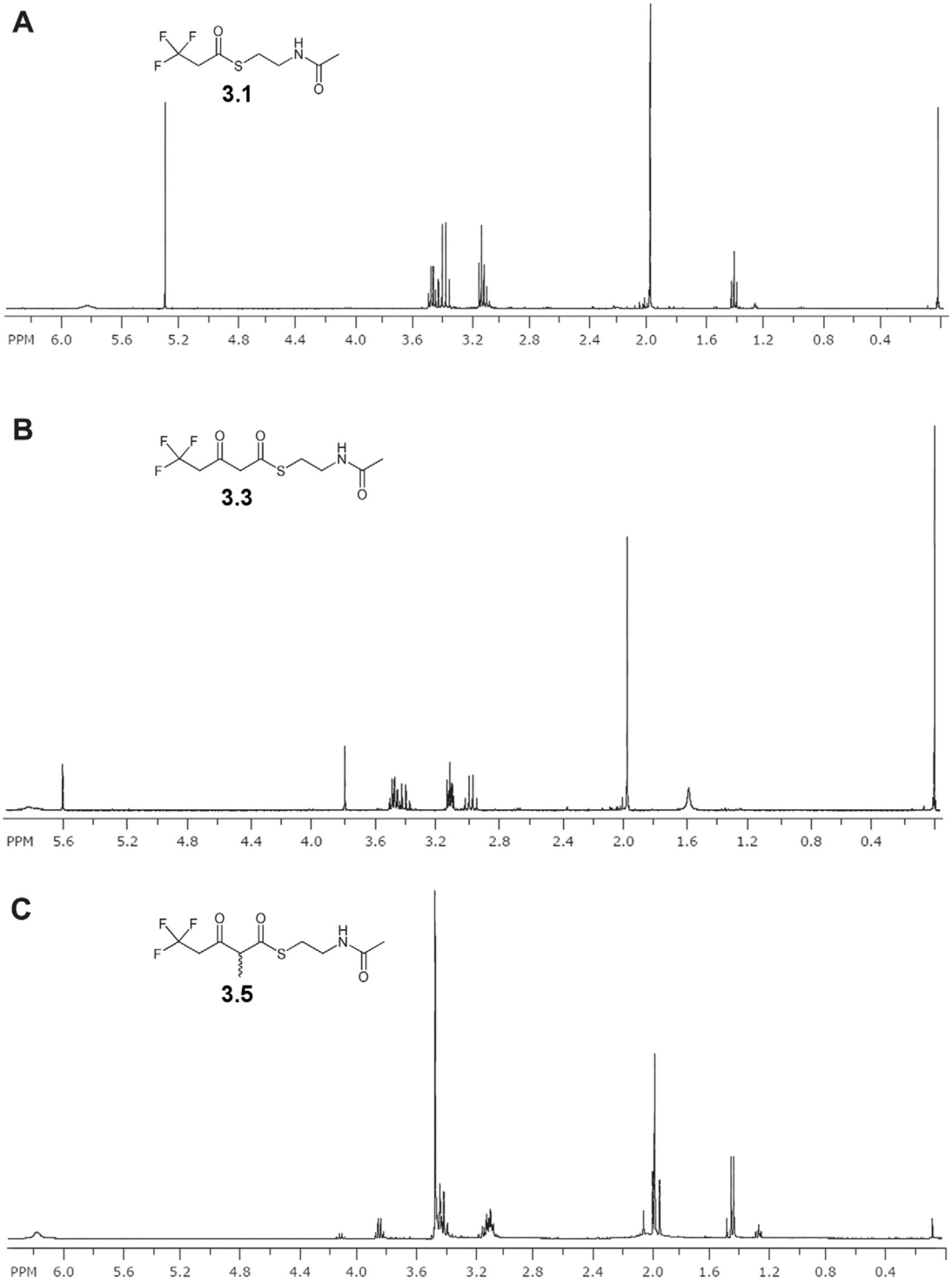

Figure 3.4. ^{1}H NMR spectra of (A) substrate **3.1**, (B) substrate **3.3**, and (C) substrate **3.5**.

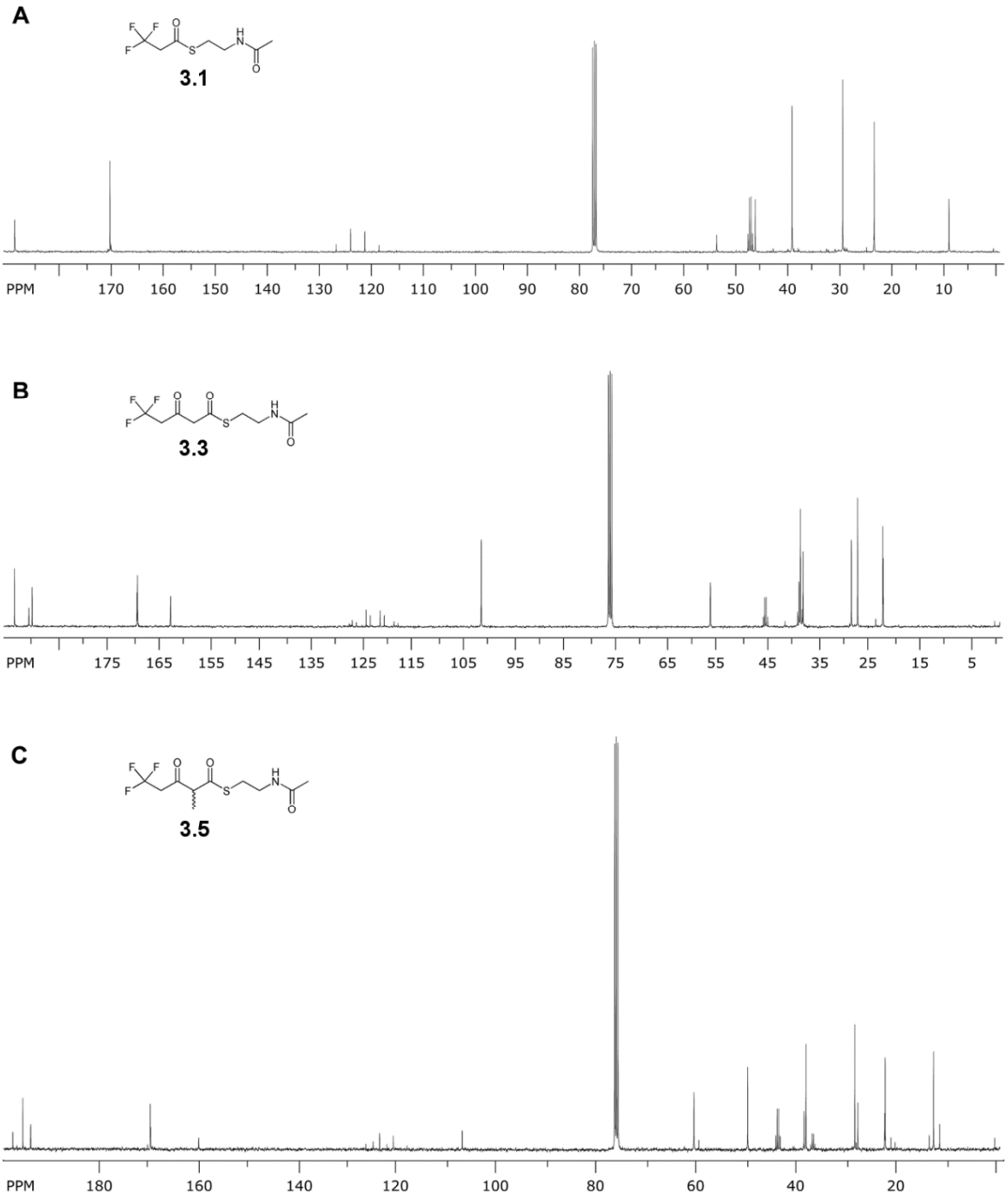

Figure 3.5. ^{13}C NMR spectra of (A) substrate **3.1**, (B) substrate **3.3**, and (C) substrate **3.5**.

Protein Expression

The construction of the expression plasmids is detailed in Piasecki et al., 2011. Expression plasmids were transformed into *Escherichia coli* BL21(DE3) cells. Overnight cultures were inoculated into Luria broth containing 25 mg/L kanamycin and grown at 37°C to an OD_{600} of 0.6. Cells were then cooled to 15°C, induced with 1 mM IPTG, and left to grow for an additional 16 hr. Cells were spun down at 4,000 x g for 30 min at 4°C, and pellets were resuspended in lysis buffer (10% v/v glycerol, 300 mM NaCl, 50 mM Tris [pH 8.0]). Cells were lysed via sonication and centrifuged at 30,000 x g for 30 min at 4°C. Cell lysate was dialyzed twice into dialysis buffer (10% v/v glycerol, 100 mM NaCl, 30 mM HEPES, [pH 7.5]) using 10 kDa MWCO cellophane dialysis tubing. Dilutions of each protein were either compared to purified enzyme of known concentration, or BSA as a standard to determine their concentrations within the lysate (**Figure 3.6**).

Kinetic Assays Analyzed by Reverse-Phase HPLC

HPLC analysis was performed at 235 nm on a Waters 1525 binary pump attached to a 2998 PDA detector on a Microsorb C18 column (300 μm, 250 x 4.6 mm, Varian). A gradient of 5%-100% B was employed (solvent A: H_2O containing 0.1% TFA; solvent B: MeOH containing 0.1% TFA).

EryTE and Substrate 3.1

Assays were set up using 115 μM EryTE cell lysate with either 3.75, 7.5, 15, 30, 60, or 120 mM of **3.1** at a final volume of 1.5 mL. The substrate was dissolved in 4% v/v DMSO-d_6 prior to mixing into 10% v/v glycerol, 200 mM HEPES, pH 7.5 (final concentration). 100 μL timepoints were taken every 10 min for 1 hr by quenching with 500 μL EtOAc. Each

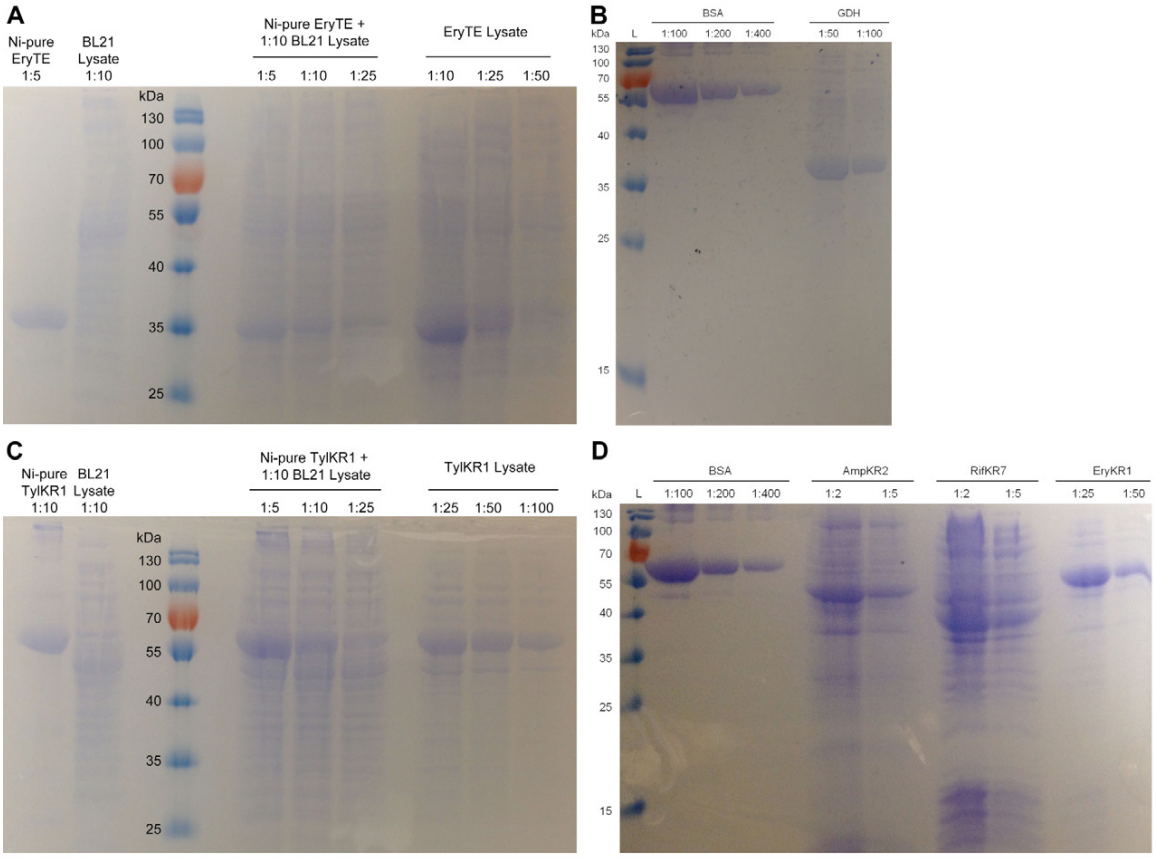

Figure 3.6. 12% SDS-PAGE gels helped to estimate the concentrations of overexpressed PKS enzymes in the cell lysate through either a comparison with purified enzyme, or a 1 mM BSA standard. Enzymes used in kinetic assays (EryTE and TylKR1) were compared to purified enzyme to more accurately determine their concentrations within lysate. Standards were made by mixing known concentrations of purified enzyme with *E. coli* BL21(DE3) cell lysate.
(A) The intensity of the EryTE band from a 1:50 dilution of EryTE-containing lysate equates to a 1:25 dilution of purified EryTE + cell lysate. Since the concentration of undiluted, purified EryTE was 95 μM, the concentration of EryTE in this cell lysate prep was ~190 μM.
(B) The concentration of GDH in this cell lysate prep was estimated at ~480 μM through a comparison with BSA standards
(C) The concentration of TylKR1 in this cell lysate prep was estimated at ~325 μM through a comparison with TylKR1 standards (undiluted, purified TylKR1 was 163 μM).
(D) The concentrations of AmpKR2, RifKR7, and EryKR1 in their cell lysate preps were estimated to be ~16 μM, ~19 μM, and ~200 μM, respectively, through a comparison to BSA standards. "L" denotes a Fermentas Page Ruler ladder.

timepoint was extracted with EtOAc (3 x 500 μL) and dried in vacuo. Prior to injection on the HPLC each was dissolved in MeOH. The substrate **3.1** peak was integrated in each chromatogram and compared to a standard curve in order to determine initial rates from the linear portion of each graph. The standard curve was made by injecting known concentrations of **3.1** onto the HPLC, plotting the resulting peak area for each substrate concentration, and fitting to a straight line. Rates were plotted against substrate concentration using GraFit, and curves were fit to the Michealis equation to extract kinetic constants (**Figure 3.1**). Each assay was performed in duplicate.

TylKR1 and Substrate 3.3

30 mM substrate **3.3** was set up in duplicate with 81 μM TylKR1 cell lysate at a final volume of 1 mL. The substrate was dissolved in 4% v/v DMSO-d_6 prior to mixing into 300 mM D-glucose, 100 μM $NADP^+$, 1 μM GDH, 10% v/v glycerol, 80 mM NaCl, and 150 mM Tris, pH 7.5 (final concentration). 100 μL timepoints were taken every 10 min for 1 hr by quenching with 500 μL EtOAc. Each timepoint was extracted with EtOAc (3 x 500 μL) and dried in vacuo. Prior to injection on the HPLC each was dissolved in MeOH. Peak areas were compared to a standard curve in order to extract initial velocities. The standard curve was made by injecting known concentrations of **3.3** onto the HPLC, plotting the resulting peak area for each substrate concentration, and fitting to a straight line.

Kinetic Assays Analyzed by ^{19}F NMR

Kinetic assays were performed on a Varian Inova 500 MHz instrument in 10 mm tubes. Prior to adding the enzyme, the magnetic field was shimmed on the FID from DMSO-d_6

used to solubilize the substrate. Only expected reactions occurred, as no other peaks arose after a 24 hr period (**Figure 3.10**).

EryTE and Substrate 3.1

The reactions were set up the same as reported in the previous HPLC section. The first timepoint was taken 15 min after enzyme addition since it takes time to set up the instrument (this delay was kept constant in all assays). The peaks were integrated using VNMR software, and the areas for each timepoint were compared to a standard curve in order to determine concentration from peak area. The standard curve was prepared by adding a known amount (2 mg, 10.4 mM final) of 3,3,3-trifluoropropionic acid to solutions of substrate **3.1** (final concentrations of 3.75, 7.5, 15, 30, 60, and 120 mM) and comparing the peak areas to determine actual concentrations of **3.1** in each array. Arrays were collected at 470 MHz at a probe temperature of 27°C, sweepwidth = 9400 Hz, 64K data points, pulse angle = 81°, receiver gain = 40, 8 FIDs per timepoint, repetition time of 300 sec. for 21,600 sec. total. 3,3,3-trifluoropropionic acid was also used as a standard to verify product formation (**Figure 3.7**).

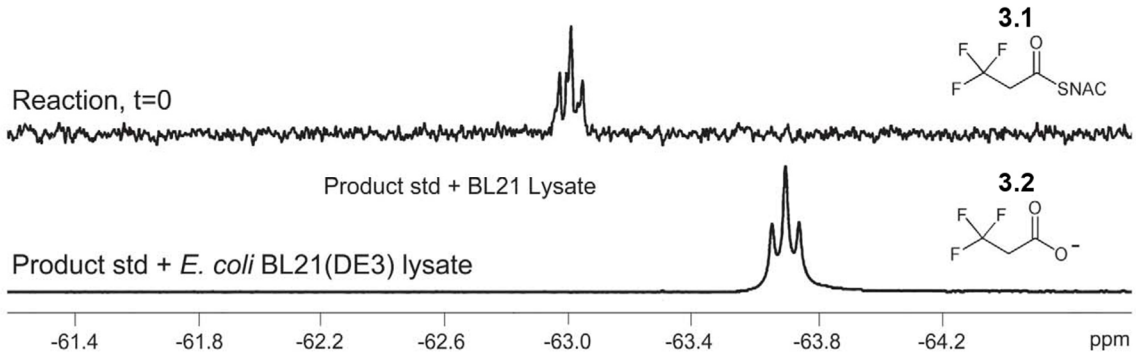

Figure 3.7. 3,3,3-Trifluoropropionic acid was used as a product standard to confirm EryTE product formation. The standard was added to *E. coli* BL21(DE3) cell lysate. A shift of -63.71 ppm aligns with a peak observed in the EryTE hydrolysis assays (Figure 3.1).

TylKR1 and Substrate 3.3

The reactions were set up the same as reported in the previous HPLC section. A standard curve was performed in the same manner as with substrate **3.1** but with 7.5, 15, 30, 60, and 120 mM of **3.3**. All acquisition and processing parameters were as reported above. The first timepoint was taken 15 min after enzyme addition since it takes time to set up the instrument. ESI-MS was performed on the EtOAc-extracted reaction (3 x 500 μL) after data collection, and an observed ESI-MS mass of 274.2 matched the expected mass (274.3) for the reduced product. ^{19}F and ^{1}H NMR were also used to verify product formation (**Figure 3.8**).

Stereocontrolled Reduction of Substrate 3.5

Reduction Assays with AmpKR2, RifKR7, TylKR1, and EryKR1

KR reduction assays were performed similar to the reduction assays of TylKR1 on substrate **3.3**. 20 mM of **3.5** was dissolved into 4% v/v DMSO-d_6 prior to mixing into 300 mM D-glucose, 100 μM $NADP^+$, 80 mM NaCl, 10% v/v glycerol, 150 mM Tris, pH 7.5, as well as 1 μM GDH, and 5 μM of KR (final concentrations) in an 800 μL volume. Reactions were incubated at room temperature, and ^{19}F NMR spectra were acquired on a Varian Mercury 400 MHz instrument after 24 hr by moving the reaction into a 10 mm NMR tube. Since the *anti*- diastereomers only shift slightly from the unreduced substrate **3.3**, an additional ^{19}F NMR spectrum is included (**Figure 3.9**) showing the TylKR1 reaction progress after 1 hr in which the substrate and product triplets partially overlap. Once NMR analysis was complete for each of the KR types, reactions were extracted with EtOAc (3 x

500 μL) and analyzed by ESI-MS. An observed mass of 288.2 verified that the reduced product (expected mass of 288.3) was being generated in each KR reaction.

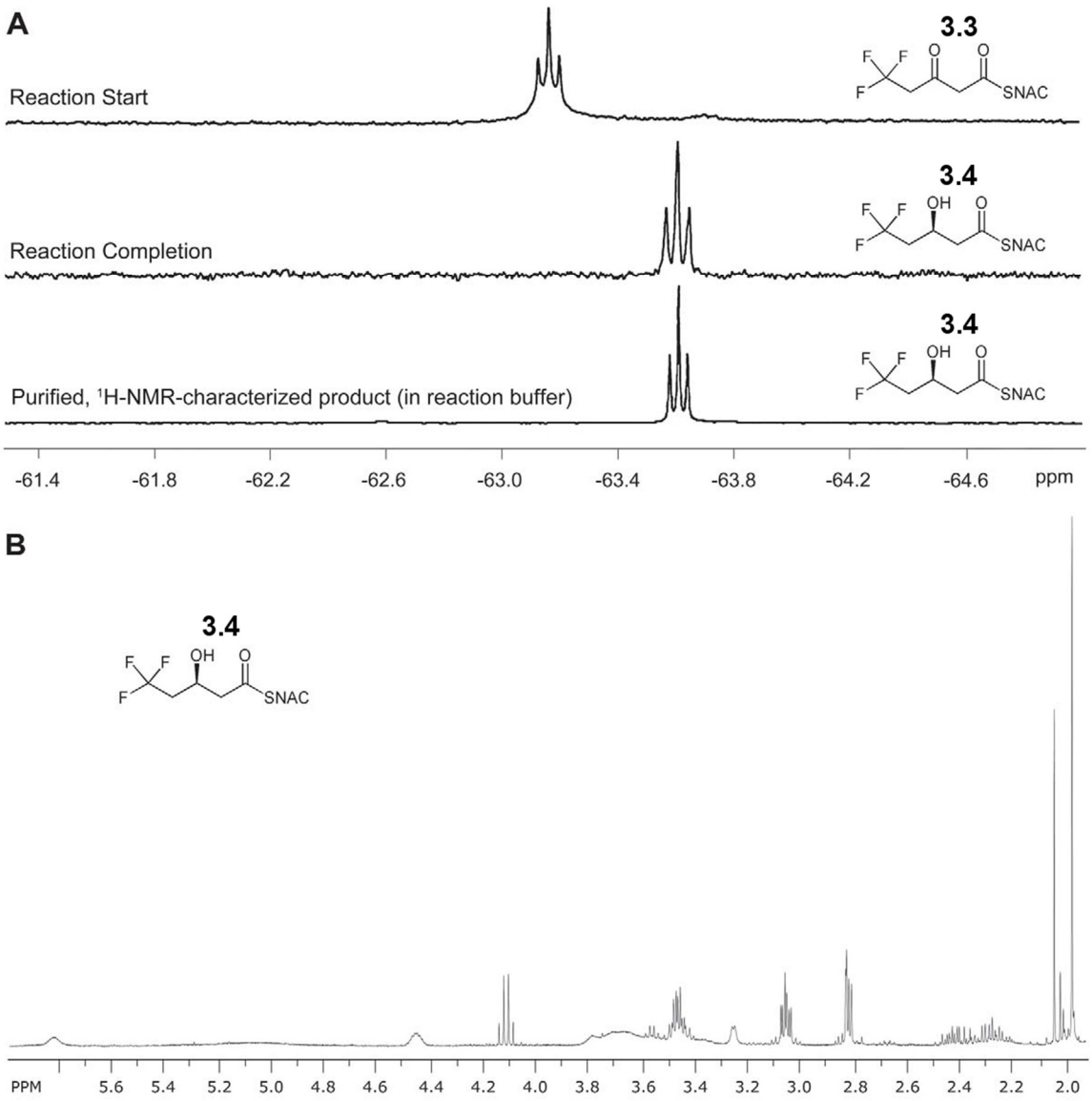

Figure 3.8. In addition to ESI-MS, ^{19}F and ^{1}H NMR were used to verify formation of **3.4** by TylKR1. (A) A ^{19}F NMR spectrum taken at the completion of the reaction shows a triplet at -63.6 ppm. (B) The product was purified and resuspended in $CDCl_3$ to obtain a ^{1}H NMR spectrum. After the product was again isolated and added to the reaction solution, the same NMR triplet was observed by ^{19}F NMR (A).

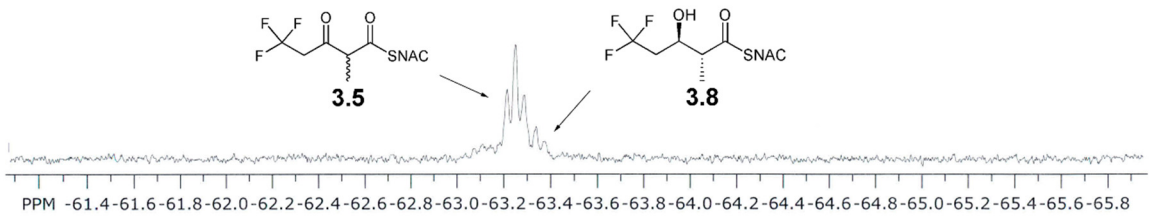

Figure 3.9. A ^{19}F NMR spectrum showing the transformation of substrate **3.5** to product **3.8** by TylKR1. The reaction progress was observed 1 hr after enzyme was added. The substrate and product peaks partially overlap.

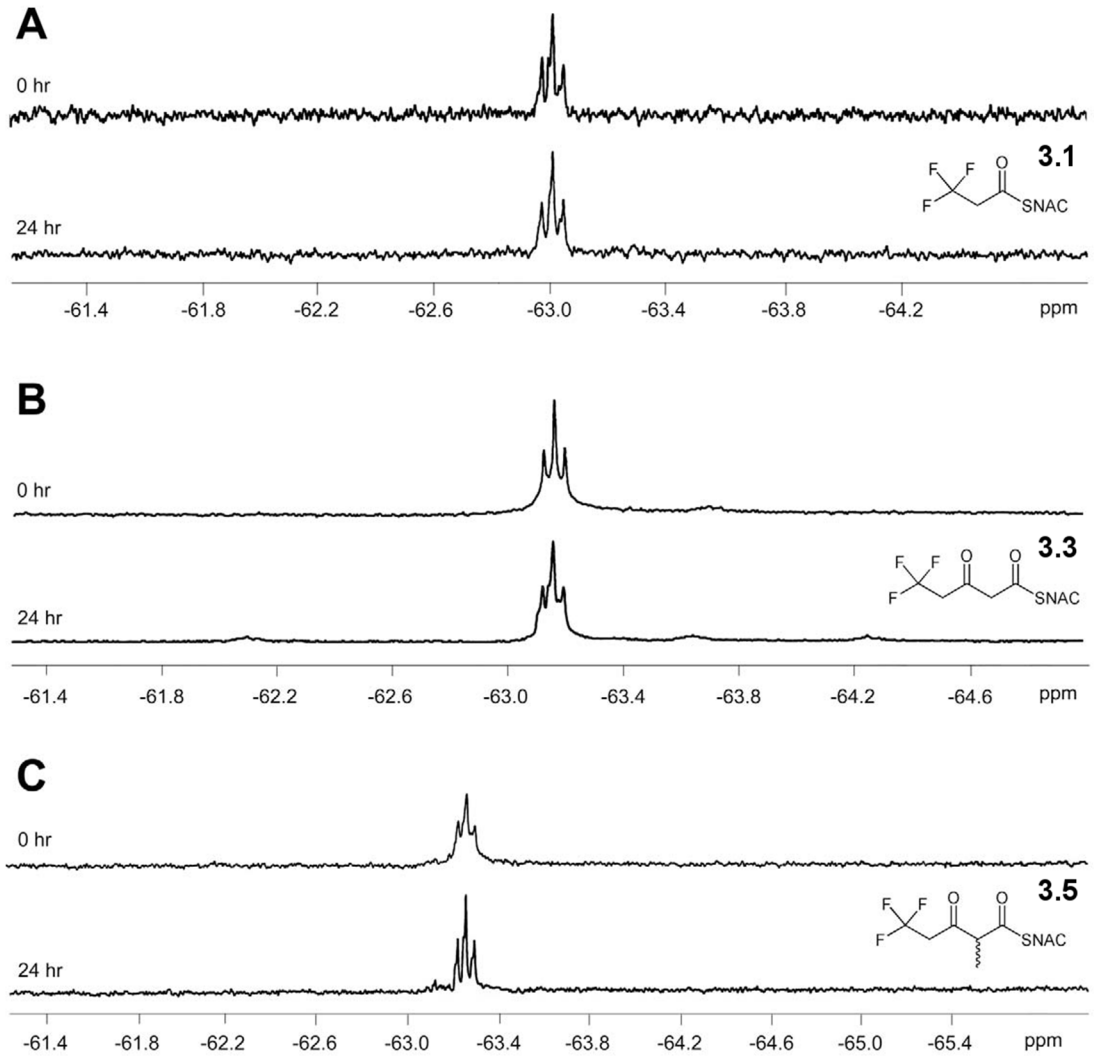

Figure 3.10. No detectable reaction occurs to the substrates of this study when incubated in *E. coli* BL21(DE3) cell lysate lacking PKS biocatalysts.

ACKNOWLEDGEMENTS

This research was funded by the Robert A. Welch Foundation (Grant F-1712). We thank Joshua Detelich and Madeline Detelich for help synthesizing fluorinated substrates. We also thank Angela Spangenberg and the UT Austin NMR core facility for help collecting kinetic data.

Chapter 4. Structural and Functional Studies of a *trans*-Acyltransferase Polyketide Synthase Ketoreductase That Reduces α- and β-Keto Groups

SUMMARY

While all the domains from *cis*-acyltransferase-containing modular type I Polyketide Synthases (PKSs) have been studied via X-ray crystallography, there are few solved domain structures from *trans*-acyltransferase (AT) variants. The Bacillaene nonribosomal peptide synthase (NRPS)/PKS hybrid (PksX) is an example which recruits *trans*-ATs for extender unit delivery during biosynthesis. Here we examine the 2.35 Å structure of the first ketoreductase from the Bacillaene NRPS/PKS hybrid (PksKR1) in the presence and absence of its cofactor, $NADP^+$. The interesting nature of this enzyme's dual-function as an α- and β-ketoreductase is also characterized using *N*-acetylcysteamine thioester substrate analogs in ketoreduction assays. Activity was observed separately for both α- and β-reduction, and incredibly, a native β-keto reducing KR was able to reduce an α-keto group stereospecifically. We propose two models for distinct binding conformations for both α- and β-reduction, and present a sequence analysis which reveals a residue acting as a stereoselective gate-keeper for β-reduction.

INTRODUCTION

Type I PKSs do not always carry covalently-attached AT domains; *trans*-AT PKSs instead utilize discrete AT domains (Cheng et al., 2009). The model PKS studied here is the Bacillaene PksX synthase from *Bacillus subtilis*, which is comprised of 5 smaller subunits (PksJ, L, M, N, R), two *trans*-AT domains (PksC & D), a *trans*-AT-ER (PksE), and a cytochrome P450 (PksS), which after catalysis results in the metabolite bacillaene

(Reddick et al., 2007; Calderone et al., 2008; Piel, 2010) (**Figure 4.1A**). This metabolite was originally isolated from fermentation broths of *B. subtilis*, (Patel et al., 1995), and also discovered through bioinformatic studies of *B. amyloliquefaciens* FZB 42 (Chen et al., 2006), and small molecule screening studies from soil bacteria (Straight et al., 2007). Although multiple groups updated the gene annotation of PksX (Calderone et al., 2006; Chen et al., 2006; Butcher et al., 2007; Straight et al., 2007), and the structure of bacillaene was known (Butcher et al., 2007), it was a decade after bacillaene was discovered when 13 of the intermediates of this pathway were elucidated. These intermediates were released as shunt products (detected by LC-MS) after deletion of the terminal thioesterase (TE) (Moldenhauer et al., 2007). Until recently, most research has concentrated on the characterization of intermediates and final products in these *trans*-AT systems within PksX, thus much is still unknown about the structural architecture and substrate preferences of this megasynthase.

Perhaps most surprising is the unusual substrate preference of PksJ. The starter unit for the first NRPS module of PksX (A-ACP loading didomain) was originally thought to be α-hydroxyisocaproate (α-HIC), which would be condensed with glycine (from C-A-ACP, NRPS module) and passed to the next module (Butcher et al., 2007). Instead it was found that the first ketoreductase of PksJ (PksKR1) actually performs α-reduction on an initially-loaded α-ketoisocaproate (α-KIC) group, as well as its β-reductase activity (**Figure 4.1B**) (Calderone et al., 2008). Therefore, although PksKR1 acts as a β-ketoreductase, interestingly it also reduces non-canonically at the α-position on its natural substrate. Ketoreductases performing α-reduction are quite rare in the literature, two examples being 1) an A domain-embedded KR within the cereulide NRPS (domain structure A-KR-PCP) that reduces an α-KIC starter unit to D-α-HIC in CesA and L-α-HIC in CesB (Magarvey et

al., 2006), and 2) the bryostatin loading module (BryA, proposed to have domain structure DH-KR-FkbH-ACP) wherein the KR is hypothesized to perform α-reduction on a 2-oxopropanoyl-*S*-ACP thioester after the processing of 1,3-bisphosphoglycerate by FkbH and the DH (**Figure 4.1C**) (Hildebrand et al., 2004; Piel, 2010). Among the few examples of α-ketoreduction known, there is no precedent to a ketoreductase that is able to perform both α- and β-reduction. The closest example of a dual-function KR is from a fungal PKS which is able to produce differing β-hydroxy stereochemistry, based on the length of substrate encountered (Zhou et al., 2012). Studies have attempted to provide an α-keto acid to a β-reducing KR, yet no reduction was detected by any of the KRs tested (Bali and Weissman, 2006; Bali et al., 2006). This may have been due to the KR's inability to process a carboxylic acid, because when the same KRs were tested with 1,2-cyclo-hexanedione, slight activity was observed, albeit with reductions in kinetic rates when compared to a natural substrate mimic, (2*RS*)-methyl-3-oxopentanoyl-*N*-acetylcysteamine thioester (Bali and Weissman, 2006; Bali et al., 2006). Thus we were interested in characterizing this dual-function ketoreductase with *S*-NAC diketide thioesters, which had not been completed for any *trans*-AT ketoreductases.

The *cis*-AT PKSs containing ketoreductases, however, have been well characterized and in fact, the product of ketoreduction is easily predicted from multiple sequence motifs (Keatinge-Clay, 2007; Keatinge-Clay, 2012). There has yet to be such a system in which product stereochemistry can be predicted from *trans*-AT PKS KRs. Consequently not only would biochemical characterization of these domains aid in elucidating the stereoselectivity and processivity of these enzymes, but structural information would also provide insight where sequence cannot, enabling a better understanding of KR stereoselectivity.

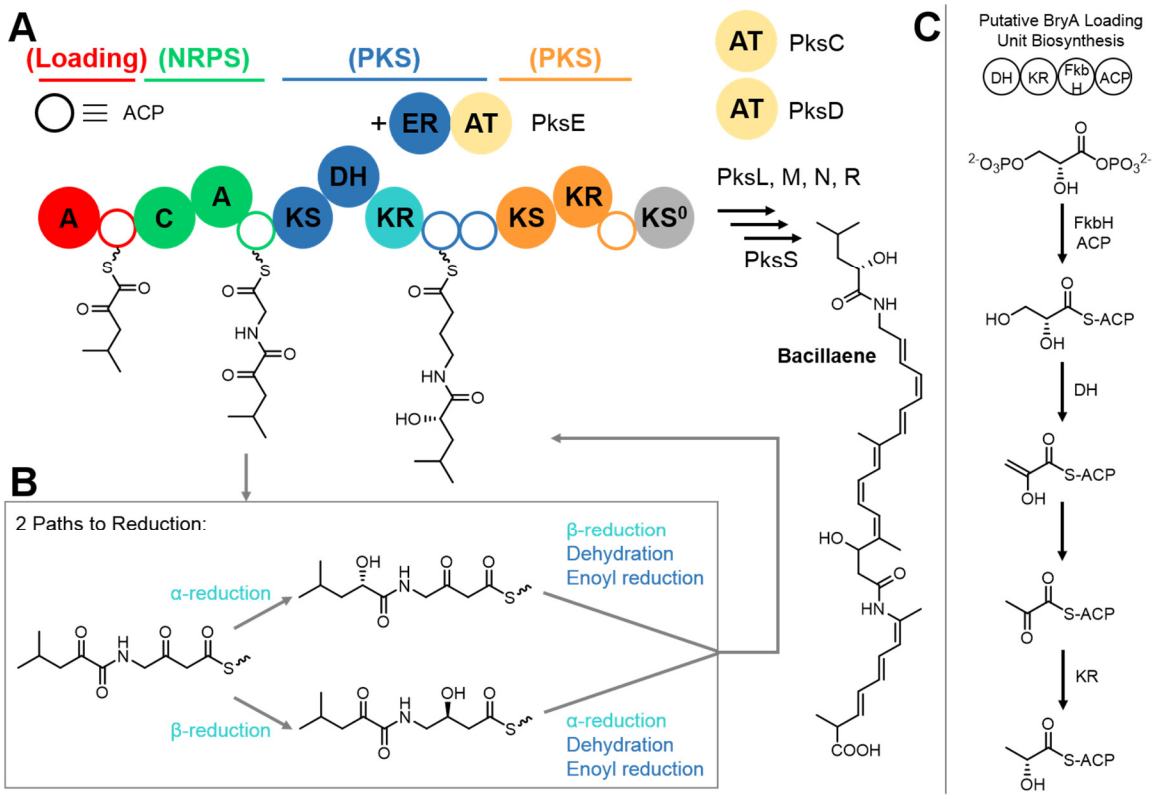

Figure 4.1. The Bacillaene PKS (PksX).
(A) Three *trans*-AT domains supplement PksX: PksC, D, and E (an AT-ER didomain). After further processing by PksL, M, N, and R, dihydrobacillaene is oxidized by a cytochrome P450 (PksS) to bacillaene.
(B) The order of α- and β-ketoreduction is still unclear, hence either route is possible toward the saturated product of enoylreduction.
(C) An example of α-ketoreduction lies in the proposed starter unit biosynthesis by BryA with domain order of DH-KR-FkbH-ACP. Bisphosphoglycerate is dephosphorylated by FkbH, dehydrated by the DH, and α-reduction is performed by the KR after rearrangement.

Only two crystal structures from *trans*-AT PKSs have been reported to date: an AT from the disorazol PKS (PDB: 3RGI and 3SBM) and an ACP didomain from the mupirocin PKS (PDB: 2L22) (Wong et al., 2011; Dong et al., unpublished). A crystal structure was obtained for PksC from *B. subtilis*, one of the *trans*-ATs from PksX, yet no structural information or catalytic characterization has followed (Cuskin et al., 2011). The atypical architectures of modules within *trans*-AT PKSs result in an array of product possibilities, unlike the *cis*-AT PKSs, which are capable of a mere eight module types (Piel, 2010). This leads to difficulty in the prediction of intermediates or final products, which is exacerbated by the lack of existing *trans*-AT PKS sequences. Analyzing structural information of these megasynthases seems to be the only answer to understanding how the systems operate, thus an increase in known structures from *trans*-AT systems is desired. Here, we present the structure of the first ketoreductase of PksX (PksKR1) with hopes to reveal any clues regarding the enzyme's ability to catalyze both α- and β-reduction on the same substrate. A 2.35 Å structure of PksKR1 in complex with $NADP^+$ is presented, alongside sequence analysis of other *trans*-AT PKS KRs which reveals a residue important for β-reduction. Reduction assays were also carried out using *S*-NAC thioester substrate mimics, and two binding modes are proposed to describe the dual-function of this KR.

RESULTS AND DISCUSSION

Crystallization of PksKR1 domain

The gene for PksKR1 was cloned from *Bacillus subtilis* genomic DNA and inserted into a pET28b backbone. Expressed protein containing an N-terminal 6x His tag was purified via nickel affinity chromatography then gel filtration for crystallization trials.

Selenomethionine-labeled protein was also expressed and its crystals used to solve the phases via single wavelength anomalous dispersion (SAD). Two structures resulted: the apo structure (3.0 Å), and complexed structure (2.35 Å) containing electron density for the cofactor after co-crystallizing with 5 mM $NADP^+$ (**Table 4.1**). The ~50 kDa PksKR1 protein consists of two Rossmann-like folds, one of which is structural and the other catalytic. The structural subdomain does not contain the NAD(P)H binding motif that is included in the catalytic subdomain (TGGAGSLG in this structure); therefore, the structural portion of the KR is thought to stabilize the catalytic portion during catalysis. In PksKR1, NADPH and NADH have both been shown as active hydride donors in biocatalytic assays (Calderone et al., 2008). As expected, the PksKR1 active site contains the same catalytic residues (Y330, K293, S317) as *cis*-AT PKS KRs with no major conformational changes once cofactor binds (**Figure 4.2A and 4.2B**). The $NADP^+$-complexed structure exhibits full electron density around most of the cofactor (**Figure 4.2C**), and is positioned to deliver a hydride to the substrate β-carbonyl (also see **Figure 4.3**).

Table 4.1. Data Collection and Refinement Statistics.

Data Collection			
Dataset	Protein	Protein-$NADP^+$	Seleno-Protein
Wavelength (Å)	0.9765	0.9795	0.9765
Space group	$P6_1$	$P6_1$	$P6_1$
Cell dimensions			
a, *b*, *c* (Å)	59.6, 59.6, 221.1	59.5, 59.5, 220.8	59.5, 59.5, 221.5
Resolution (Å)	50-3.0(3.05-3.0)	50-2.35(2.39-2.35)	50-2.6(2.64-2.60)
R_{merge}	0.134(0.502)	0.069(0.464)	0.14(1.0)
$I / \sigma(I)$	18.4(4.6)	25.8(5.2)	24(2.8)
Completeness (%)	99.9(100)	99.4(100)	100(100)
Redundancy	9.5(9.6)	10.9(11.1)	22.5(19.7)
Refinement			
Resolution (Å)	50-3.0	50-2.35	
No. reflections	8353	17260	
R_{work} / R_{free}	0.258/0.290	0.232/0.253	
No. atoms			
Protein	3330	3330	
$NADP^+$		48	
B-factors ($Å^2$)			
Protein	67.441	73.199	
$NADP^+$		31.824	
R.m.s. deviations			
Bond lengths (Å)	0.011	0.008	
Bond angles (°)	0.865	1.091	

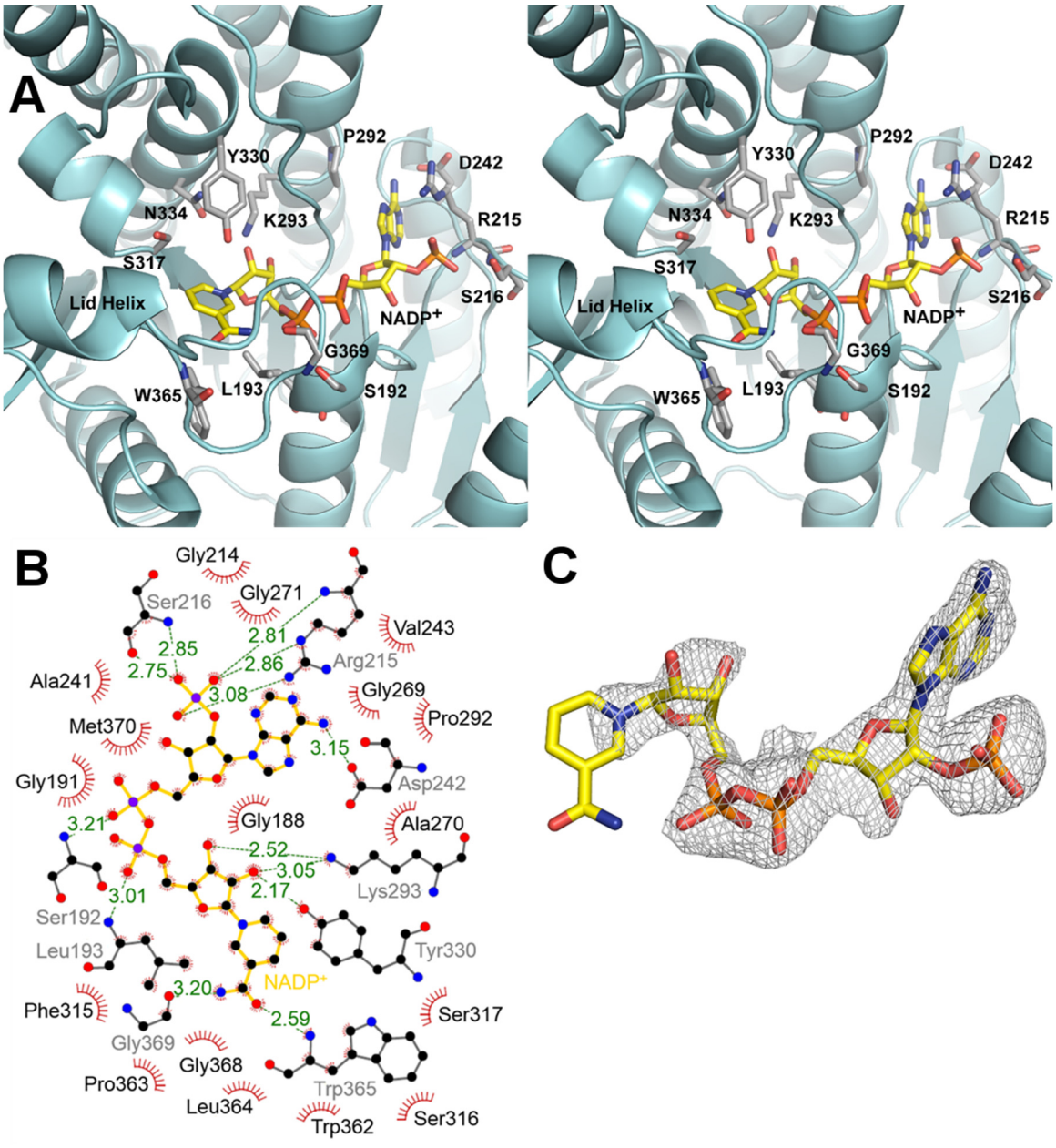

Figure 4.2. The active site of PksKR1.
(A) A stereoview of NADP$^+$ (yellow) within the active site of PksKR1. Residues shown are those that interact with cofactor, based on Ligplot prediction.
(B) A Ligplot representation showing interactions between NADP$^+$ (yellow) and PksKR1 residues.
(C) The $|F_o-F_c|$ omit map of NADP$^+$ against the final structure (contoured at 3.0 σ) illustrates the extent to which the cofactor fits within electron density.

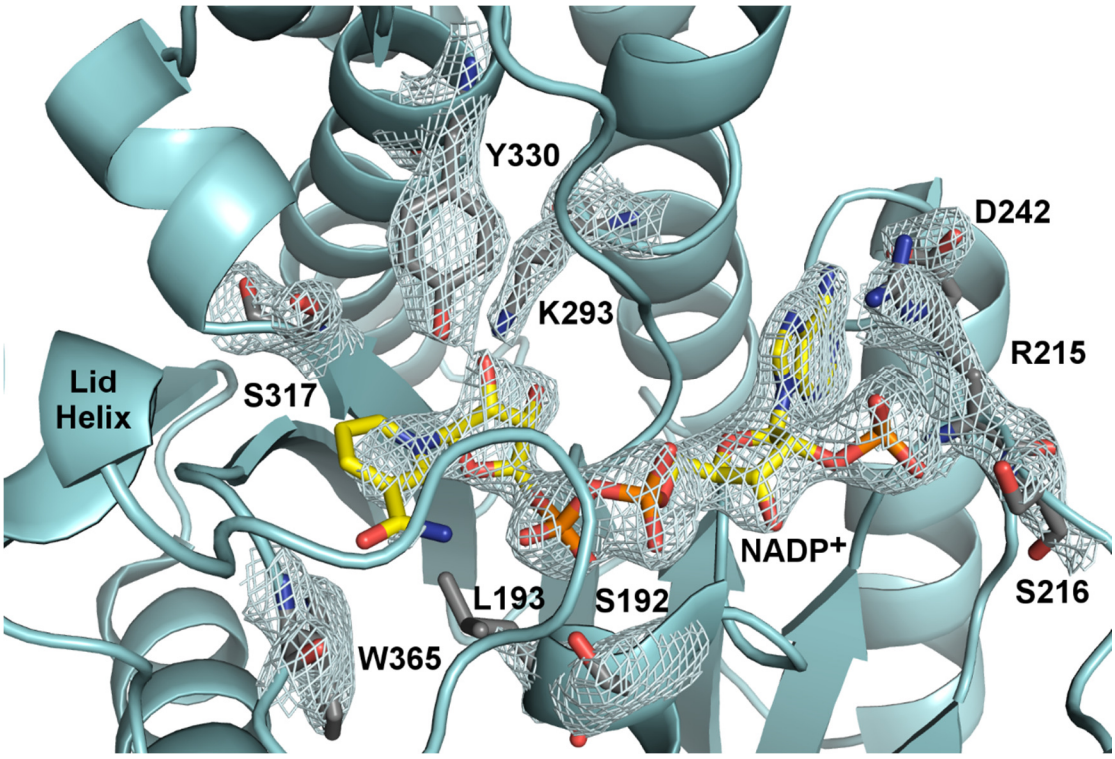

Figure 4.3. The $|2F_o-F_c|$ map. Map of PksKR1 active site residues (contoured at 2.0 σ) illustrate the extent to which these residues and cofactor fit within electron density. Map was hidden for clarity of residues not shown. Active site residues and cofactor maintain the same orientation as in other solved KR structures.

Comparing KRs from *cis*-AT and *trans*-AT PKSs

We were interested in the crystallographic elucidation of this ketoreductase in order to uncover structural evidence for its dual α- and β-reduction functionality. Upon assessment of the PksKR1 sequence, it was obvious that a sequence-to-function analysis (which is possible with *cis*-AT PKS KRs) was not possible. The most obvious difference is the lack of the LD portion of the LDD motif as found in B-type β-reducing KRs (highlighted grey in **Figure 4.4**). Usually, if a β-reducing KR lacks an LDD motif, it is deemed A-type and

there is a conserved Trp that instead interacts with the substrate, guiding entrance in from the left, behind the lid helix; yet that residue is missing from PksKR1 as well. Instead, Leu is found at that position (highlighted orange in **Figure 4.4**), which is sometimes seen for β-reducing KRs that are B1-type. Interestingly, 3 residues N-terminal to the active site Tyr resides a very well-conserved Gln which is normally seen in A1- and B1-type KRs (highlighted green in **Figure 4.4**). The lid helix in the crystal structure of PksKR1 does sit further away from the active site than the lid helix in the EryKR1 structure (**Figure 4.5**), which may provide a larger active site for the longer substrate which PksKR1 encounters. The lid helix is also slightly longer than in KRs that are not neighboring other processing enzymes, which is also seen in the Spn(KR+ER)2 didomain structure (Zheng et al., 2012). Because we were unable to characterize this ketoreductase using the established *cis*-AT PKS KR motifs, we were interested in finding conserved residues in *trans*-AT PKS KRs that could possibly form similar rules.

KR sequences from 10 PKS pathways were chosen for sequence analysis to predict whether a substrate would be reduced to a D- or L-hydroxy product. KRs from these PKS pathways were chosen if they resided alongside a dehydratase (DH) domain. In most cases, DHs from *cis*-AT PKSs will install a trans double bond if the hydroxyl group is of "D" stereochemistry (Wu et al., 2005; Guo et al., 2010; Valenzano et al., 2010; Keatinge-Clay, 2012) and a cis double bond if the hydroxyl group instead is of "L" stereochemistry (Reid et al., 2003; Keatinge-Clay, 2008; Akey et al., 2010; Bonnett et al., 2013). Accordingly, KR sequences were further separated into two categories based on their known or predicted isomerism of the olefin intermediate (Piel, 2010): 1) KRs that are predicted to yield L-hydroxy intermediates (a cis double bond, or an "A-type" KR), and 2) KRs that are predicted to yield D-hydroxy intermediates (a trans double bond, or a "B-type" KR). Once

separated, the catalytic portion of those KR sequences were then aligned using ClustalW. Motifs and conserved residues normally found in *cis*-AT PKS KRs are highlighted in **Figure 4.4**.

The Lys, Ser, and Tyr catalytic triad (colored in fuchsia) is conserved in all sequences, except the KR from module 12 of the chivosazol PKS, which contains an Ala in place of the Ser. A conserved Gln (in green) is present in most of these KRs, which is also seen in *cis*-AT PKS KRs, except for virginiamycin KR1, bryostatin KR5, and disorazol KRs 5 and 7. The LDD motif (in grey) normally found in *cis*-AT PKS KRs, which predicts reduction to be of D-configuration, is present in only a few of the *trans*-AT PKS KRs, yet those missing the motif do not have the Trp normally found (in orange) in KRs which yield L-hydroxy stereochemistry. In most KRs from *cis*-AT PKSs there is usually an invariant Asn 4 residues C-terminal to the catalytic Tyr (in yellow), yet PksKR1 (signified by a star) and one third of the other *trans*-AT PKS KRs analyzed in this study instead have a Ser or Cys in its place. When comparing "A-type" to "B-type" KRs from *trans*-AT PKSs, one significant difference is the presence (or absence) of the final Asp in the LDD motif. We believe that this residue sets these two KR types apart, so if a *trans*-AT PKS KR contains this final Asp of the motif, it will normally produce a D-hydroxy group, and if the final Asp is replaced by another residue, the KR will normally yield an L-hydroxy group.

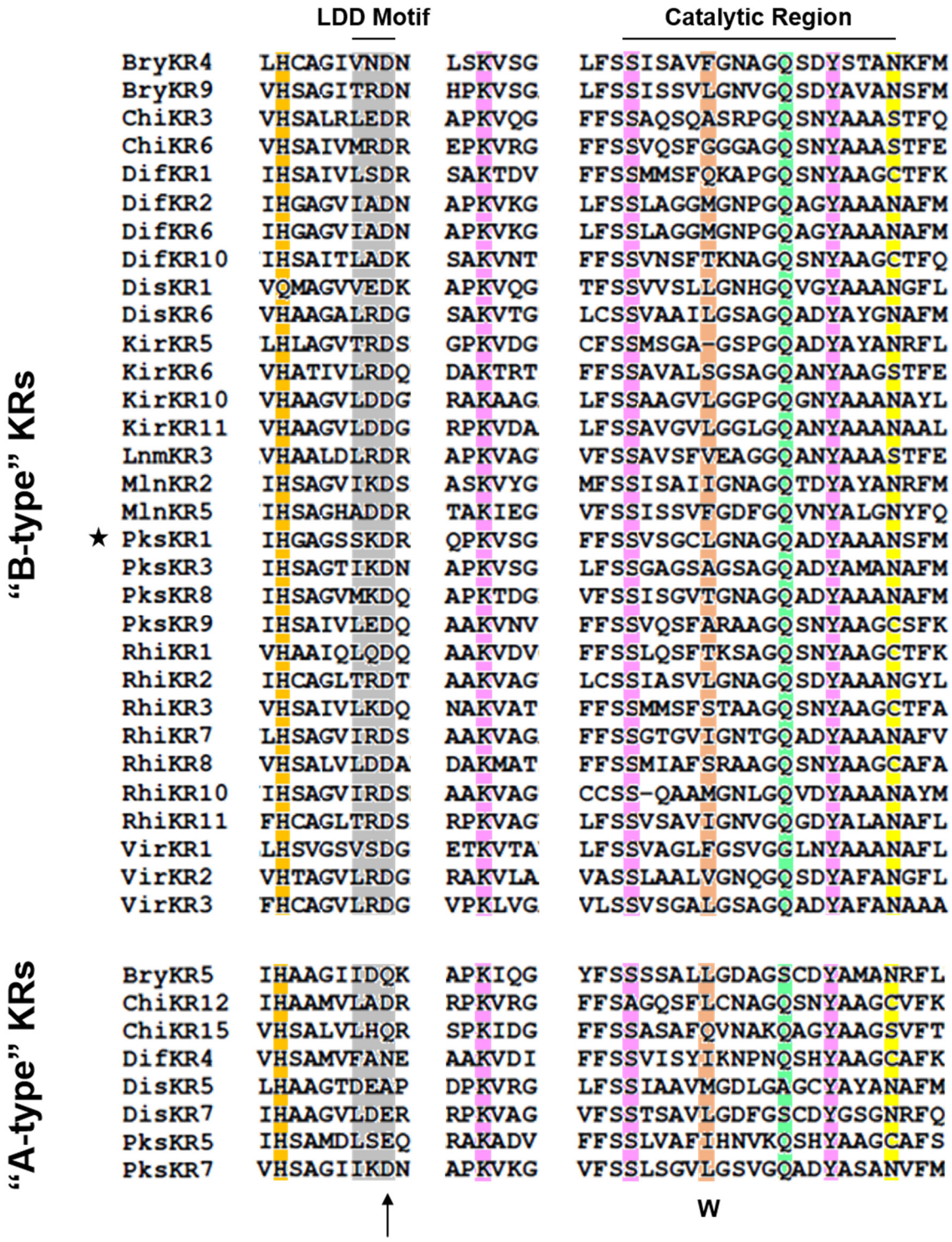
LDD Motif
Catalytic Region
"B-type" KRs
"A-type" KRs
BryKR4 LHCAGIVNDN LSKVSG LFSSISAVFGNAGQSDYSTANKFM
BryKR9 VHSAGITRDN HPKVSG LFSSISSVLGNVGQSDYAVANSFM
ChiKR3 VHSALRLEDR APKVQG FFSSAQSQASRPGQSNYAAASTFQ
ChiKR6 VHSAIVMRDR EPKVRG FFSSVQSFGGGAGQSNYAAASTFE
DifKR1 IHSAIVLSDR SAKTDV FFSSMMSFQKAPGQSNYAAGCTFK
DifKR2 IHGAGVIADN APKVKG LFSSLAGGMGNPGQAGYAAANAFM
DifKR6 IHGAGVIADN APKVKG LFSSLAGGMGNPGQAGYAAANAFM
DifKR10 IHSAITLADK SAKVNT FFSSVNSFTKNAGQSNYAAGCTFQ
DisKR1 VQMAGVVEDK APKVQG TFSSVVSLLGNHGQVGYAAANGFL
DisKR6 VHAAGALRDG SAKVTG LCSSVAAILGSAGQADYAYGNAFM
KirKR5 LHLAGVTRDS GPKVDG CFSSMSGA-GSPGQADYAYANRFL
KirKR6 VHATIVLRDQ DAKTRT FFSSAVALSGSAGQANYAAGSTFE
KirKR10 VHAAGVLDDG RAKAAG LFSSAAGVLGGPGQGNYAAANAYL
KirKR11 VHAAGVLDDG RPKVDA LFSSAVGVLGGLGQANYAAANAAL
LnmKR3 VHAALDLRDR APKVAG VFSSAVSFVEAGGQANYAAASTFE
MlnKR2 IHSAGVIKDS ASKVYG MFSSISAIIGNAGQTDYAYANRFM
MlnKR5 IHSAGHADDR TAKIEG VFSSISSVFGDFGQVNYALGNYFQ
★ PksKR1 IHGAGSSKDR QPKVSG FFSSVSGCLGNAGQADYAAANSFM
PksKR3 IHSAGTIKDN APKVSG LFSSGAGSAGSAGQADYAMANAFM
PksKR8 IHSAGVMKDQ APKTDG VFSSISGVTGNAGQADYAAANAFM
PksKR9 IHSAIVLEDQ AAKVNV FFSSVQSFARAAGQSNYAAGCSFK
RhiKR1 VHAAIQLQDQ AAKVDV FFSSLQSFTKSAGQSNYAAGCTFK
RhiKR2 IHCAGLTRDT AAKVAG LCSSIASVLGNAGQSDYAAANGYL
RhiKR3 VHSAIVLKDQ NAKVAT FFSSMMSFSTAAGQSNYAAGCTFA
RhiKR7 LHSAGVIRDS AAKVAG FFSSGTGVIGNTGQADYAAANAFV
RhiKR8 VHSALVLDDA DAKMAT FFSSMIAFSRAAGQSNYAAGCAFA
RhiKR10 IHSAGVIRDS AAKVAG CCSS-QAAMGNLGQVDYAAANAYM
RhiKR11 FHCAGLTRDS RPKVAG LFSSVSAVIGNVGQGDYALANAFL
VirKR1 LHSVGSVSDG ETKVTA LFSSVAGLFGSVGGLNYAAANAFL
VirKR2 VHTAGVLRDG RAKVLA VASSLAALVGNQGQSDYAFANGFL
VirKR3 FHCAGVLRDG VPKLVG VLSSVSGALGSAGQADYAFANAAA
BryKR5 IHAAGIIDQK APKIQG YFSSSSALLGDAGSCDYAMANRFL
ChiKR12 IHAAMVLADR RPKVRG FFSAGQSFLCNAGQSNYAAGCVFK
ChiKR15 VHSALVLHQR SPKIDG FFSSASAFQVNAKQAGYAAGSVFT
DifKR4 VHSAMVFANE AAKVDI FFSSVISYIKNPNQSHYAAGCAFK
DisKR5 LHAAGTDEAP DPKVRG LFSSIAAVMGDLGAGCYAYANAFM
DisKR7 IHAAGVLDER RPKVAG VFSSTSAVLGDFGSCDYGSGNRFQ
PksKR5 IHSAMDLSEQ RAKADV FFSSLVAFIHNVKQSHYAAGCAFS
PksKR7 VHSAGIIKDN APKVKG VFSSLSGVLGSVGQADYASANVFM
W

Figure 4.4. Sequence analysis reveals a residue important for determining *trans*-AT PKS KR stereochemistry (on previous page).
A sequence analysis displays motifs shared between *trans*-AT PKS KRs and *cis*-AT PKS KRs. Sequences were taken from modules which contain a neighboring DH. PksKR1 is labeled with a star. The LDD motif is highlighted in grey, with the last Asp denoted with an arrow. This residue is conserved in all *trans*-AT PKS KRs that are hypothesized to produce D-hydroxy stereochemistry. The placement of the conserved Trp in all A-type *cis*-AT PKS KRs is marked with a W and highlighted in orange.

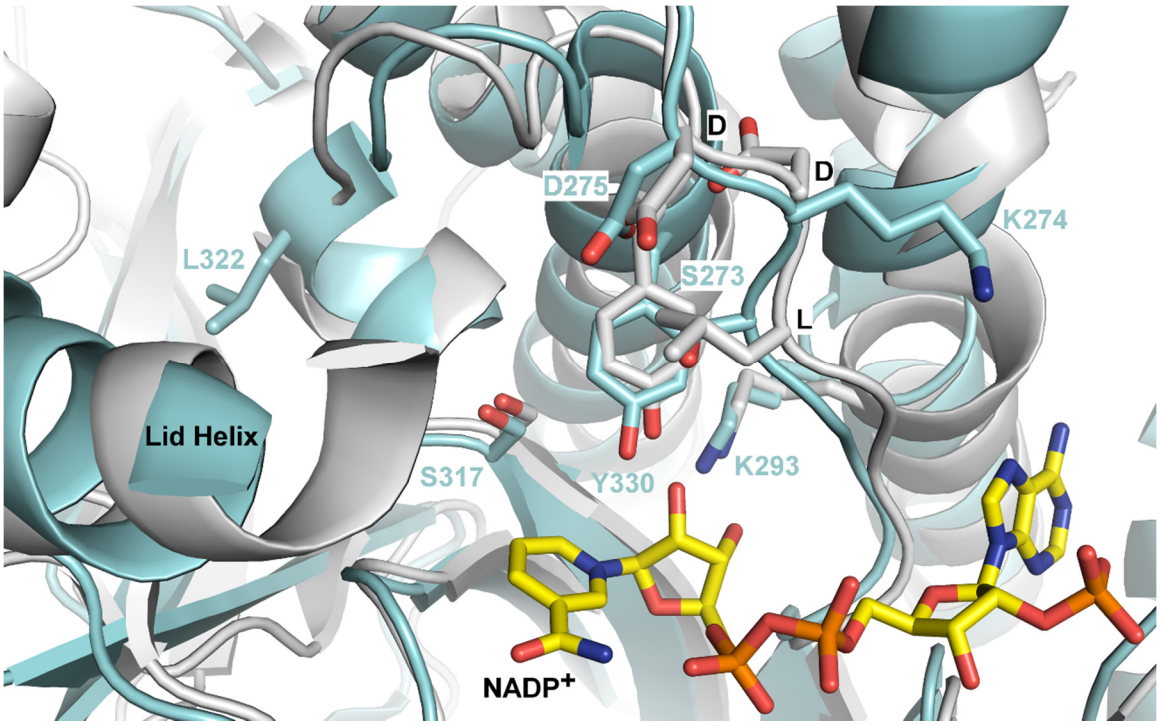

Figure 4.5. PksKR1 was overlaid with EryKR1 (PDB: 2FR0) to show structural differences between the two enzymes. Differences arise in the active site Tyr and lid helix (some of which was removed for clarity), yet the LDD loop remains static in both structures, although the bacillaene KR only contains the last Asp in that motif (S273 and K274 replace the first two residues).

Stereochemical Assays

Reduction assays were performed in order to determine whether PksKR1 is able to maintain natural stereocontrol on substrates linked to a truncated phosphopantetheinyl arm mimic (*N*-acetylcysteamine). Compounds **4.1**-**4.5** were synthesized (see Experimental Procedures and **Figure 4.8**) and utilized in ketoreduction assays employing a NADPH regeneration scheme (Piasecki et al., 2011). Glucose dehydrogenase (GDH, from *B. subtilis*) was used to regenerate the reduced cofactor NADPH, providing the ketoreductase with a continual supply for catalysis. The substrates α-ketoisocaproyl-*N*-acetylcysteamine

thioester (α-KIC-*S*-NAC, **4.1**), α-ketoisocaproyl-γ-aminobutyryl-*N*-acetylcysteamine thioester (α-KIC-GABA-*S*-NAC, **4.2**), 3-oxopentanoyl-*N*-acetylcysteamine thioester (**4.4**) and (2*RS*)-methyl-3-oxopentanoyl-*N*-acetylcysteamine thioester (**4.5**) were provided to PksKR1 and products were analyzed via chiral HPLC (**Figure 4.6**). Standard (4*S*)-α-hydroxyisocaproyl-γ-aminobutyryl-*N*-acetylcysteamine thioester ((4*S*)-α-HIC-GABA-*S*-NAC, **4.3**) was used to verify stereochemistry of the PksKR1 product for α-reduction.

We found that not only is the KR able to reduce stereospecifically at the α-position on its natural substrate mimic **4.2** (**Figure 4.6A**), which aligned with reduced standard **4.3**, but it is also able to reduce at the β-position on **4.4** (aligning with other KRs that produce "D" stereochemistry) (**Figure 4.6B**). Since **4.4** elutes as an extremely broad peak, it was not shown in the chromatogram. Remarkably, EryKR1 (a "B-type" β-reducing KR in nature) was also able to reduce **4.2** stereospecifically to the "B-type" product, yet TylKR1 (also a "B-type" β-reducing KR in nature) did not show any activity on this substrate. Reduction of **4.1** was not observed via HPLC or LC-MS for any of the KRs tested. It may be that PksKR1 gains binding energy from interactions with longer substrates, which was previously observed for *cis*-AT PKS KRs of downstream modules, reducing poorly on shorter *S*-NAC substrates (Piasecki et al., 2011). As expected, the resulting stereochemistry for the reduction of **4.4** was the "3D" isomer (**Figure 4.6B**), as observed in the reduction of the natural substrate; however, we were unsure which stereoisomer would be observed when PksKR1 was given the opportunity to be stereoselective at the α-position; hence, the reduction of **4.5** was attempted. No reduction was detected via HPLC or LC-MS when supplying the PksKR1 with **4.5** (data not shown). Although this KR was not observed to be an active reductase for substrates containing an α-methyl group, it still functions well as an α- and β-ketoreducing KR if no α-substituent is present, other than an α-keto group. The

fact that PksKR1 is a robust α- and β-reducing KR with substrates attached to shorter phosphopantetheinyl arm mimics is in itself a remarkable result. Since we controlled which type of reduction PksKR1 was to perform with the substrate provided, the order of α- and β-reduction in nature is still unknown.

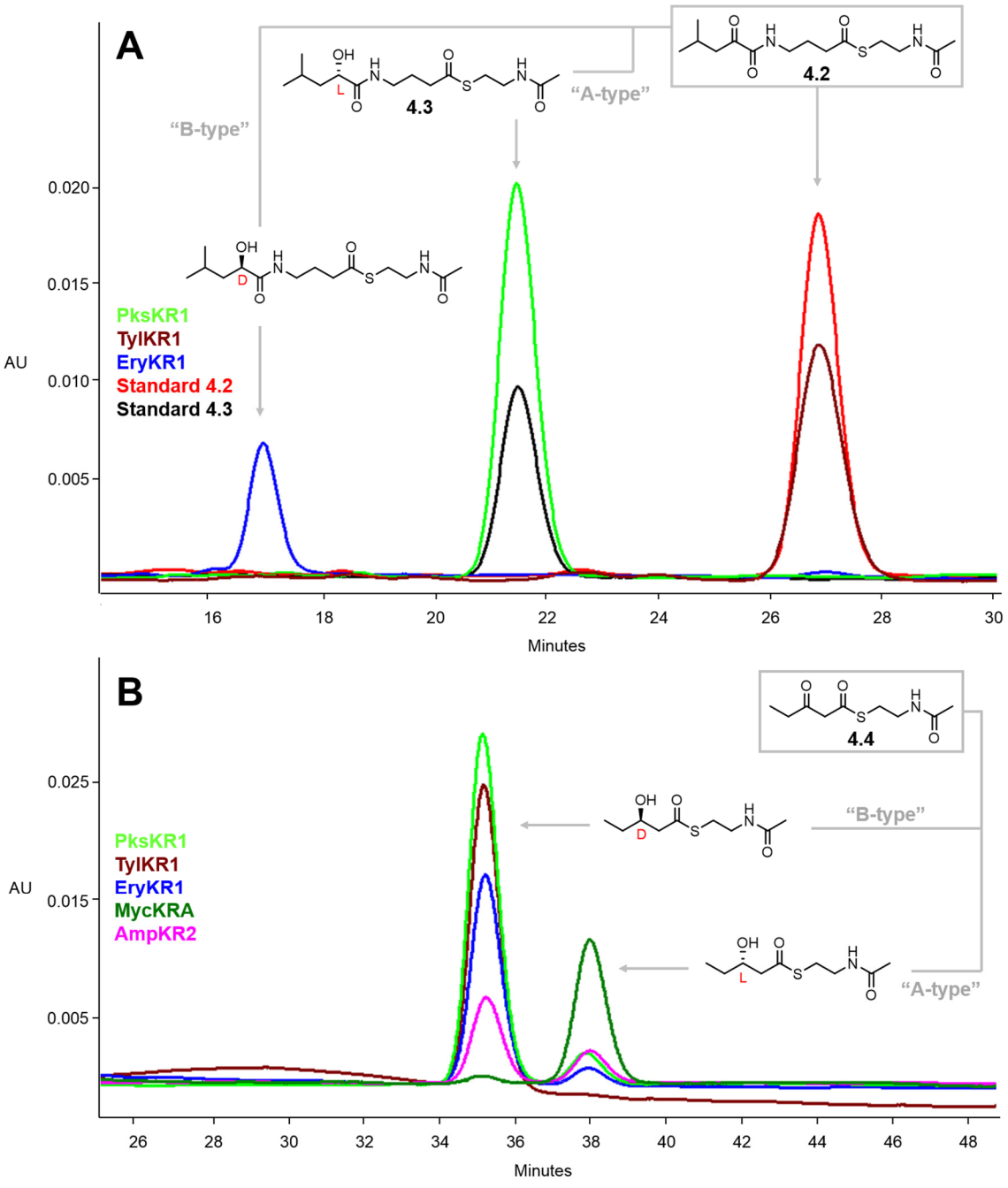
A
4.3
4.2
"A-type"
"B-type"
PksKR1
TylKR1
EryKR1
Standard 4.2
Standard 4.3
AU
Minutes
B
4.4
"B-type"
"A-type"
PksKR1
TylKR1
EryKR1
MycKRA
AmpKR2
AU
Minutes

Figure 4.6. Chiral HPLC analysis of PksKR1 α- and β-reduction (on previous page).
(A) PksKR1, TylKR1, and EryKR1 reduction of **4.2**. TylKR1 is inactive toward this substrate, whereas EryKR1 and PksKR1 reduce to D- and L-stereochemistry, respectively. Standard **4.3** verifies the stereochemistry of PksKR1 reduction.
(B) PksKR1, TylKR1, EryKR1, MycKRA, and AmpKR2 reduction of **4.4**. All KRs were able to reduce **4.4** with expected natural stereochemistry. Substrate **4.4** elutes as an extremely broad peak and is therefore not shown for clarity. All product peaks were confirmed by LC-MS.

Two Binding Conformation Models Describe the Dual-Function of PksKR1

Two models are proposed in **Figure 4.7** to account for the differing stereochemical outcome of PksKR1's α- and β-reduction. The natural substrate PksKR1 encounters is depicted as a SNAC thioester mimic and was modeled into the active site using Coot. For α-reduction that yields "L" stereochemistry, the substrate was modeled in using an A-type KR as a guide (**Figure 4.7A**). For β-reduction that yields "D" stereochemistry, the substrate was instead modeled in using a B-type KR as a guide (**Figure 4.7B**). (Zheng et al., 2013). The A-type KRs position the *re*-face of the keto group toward the NADPH hydride to yield "L" stereochemistry, whereas the B-type KRs position the *si*-face of the keto group toward the hydride to be delivered and therefore yield "D" stereochemistry. We hypothesize that the final Asp in the LDD motif plays a critical role in guiding the substrate into the active site from the right side, behind the lid helix in the *trans*-AT PKS KRs that are predicted to yield the "B-type" product of "D" stereochemistry. There does not appear to be sufficient area for the substrate to enter in from the left side behind the lid helix to perform β-reduction yielding "L" stereochemistry, as the terminal methyl groups would cause steric clashes with residues within the active site. This could explain why PksKR1 does not perform L-β-reduction. Although we contribute both models, only a crystal structure of an enzyme-substrate complex will accurately reveal how KRs from these *trans*-AT PKS pathways guide substrates into the active site for stereospecific reduction.

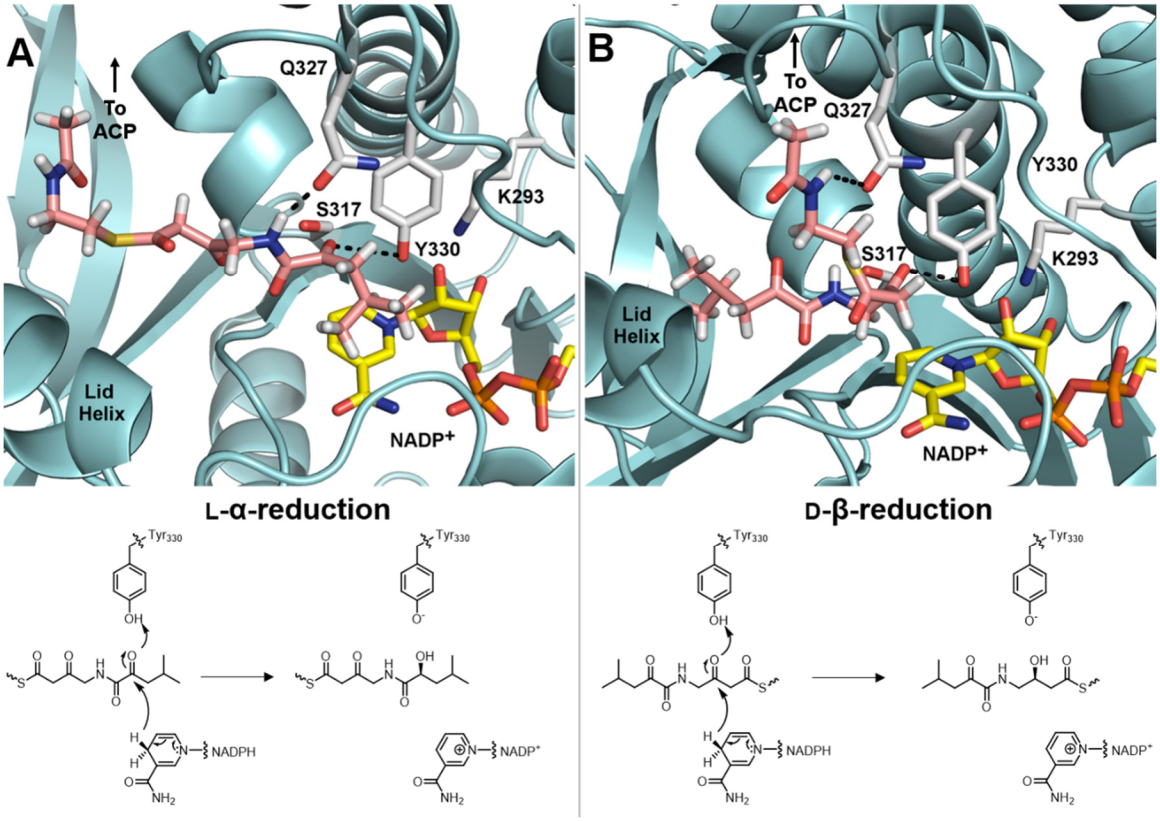

Figure 4.7. Two Proposed Models for Substrate Binding. The substrate is depicted as a SNAC thioester for clarity. Direction of ACP arm is therefore noted in the figure.
(A) α-reduction: the substrate enters the active site from the left side behind the lid helix, positioning the *re*-face of the α-keto group toward the NADPH hydride to yield "L" stereochemistry.
(B) β-reduction: the substrate enters the active site from the right side behind the lid helix, presenting the *si*-face of the β-keto group toward the NADPH hydride to yield "D" stereochemistry.

SIGNIFICANCE

We have presented for the first time the structure of an α-reducing KR, which also is the only known KR to reduce both α- and β-keto groups in nature. As a standalone enzyme, PksKR1 is a remarkable model to study: its unprecedented dual-function reductase activity has been shown in this study to translate to shorter substrate mimics, and the proposed

binding modes presented here showcase a structural basis for dual-catalysis. Since harnessing KR catalytic function is key to producing new biologically-relevant drug targets, this bifunctional KR is not only a valuable tool for biochemists, but also organic chemists wishing to incorporate unique stereopure compounds into retrosynthetic syntheses.

EXPERIMENTAL PROCEDURES

Cloning, Protein Expression, and Purification

The gene for PksKR1 was amplified from *B. subtilis* gDNA with primers 5-ATCGTAATCcatatgGAACGCTTAATGCTTGAACCGGTGT-3 and 5-TGATTCGATgaattcATCCTTGATCCTGATCCGCCTTTCTC-3, resulting in restriction sites NdeI and EcoRI (shown in lowercase). These sites were used to clone the fragment into pET28b plasmid, which was subsequently transformed into BL21(DE3) *Escherichia coli* cells. The cells were grown to an OD_{600} of 0.5 in Luria broth containing 50 mg/L kanamycin at 37°C, and the temperature was then dropped to 15°C prior to inducing protein expression with 0.5 mM IPTG, and grown for an additional 16 hours. Cells were collected via centrifugation at 4,000 x g for 30 min, resuspended in lysis buffer containing 500 mM NaCl and 30 mM HEPES pH 7.5, and lysed by sonication. Cell debris was removed by centrifugation at 30,000 x g for 30 min and the cell lysate was then poured over Ni-NTA resin (Qiagen) which was equilibrated in lysis buffer. Bound protein was washed with lysis buffer containing 15 mM imidazole, and eluted with lysis buffer containing 150 mM imidazole. A Superdex 200 gel filtration column equilibrated with 150 mM NaCl and 10 mM HEPES pH 7.5 was used to polish prior to crystallization trials. Buffer exchange into

25 mM NaCl, 10 mM HEPES pH 7.5, and 1 mM DTT was performed using protein concentrators to a final concentration of ~50 mg/mL.

Crystallization and Structure Determination

Crystals of PksKR1 grew over one week by sitting drop vapor diffusion at 22°C. Selenomethionine-labeled crystals were formed by mixing 2 μL protein solution (23.5 mg/mL PksKR1 in 25 mM NaCl, 1 mM DTT, and 10 mM HEPES [pH 7.5]) with 1 μL crystallization buffer (30% Jeffamine ED-2001 and 0.1 M HEPES [pH 7.0]). Native crystals were formed by mixing equal volumes of protein solution and crystallization buffer (25% w/v PEG 1000 and 0.1 M Tris [pH 8.5]). A final concentration of 5 mM $NADP^+$ was added to the native protein solution to obtain crystals of the PksKR1/$NADP^+$ complex, which were formed by mixing equal volumes of protein solution and crystallization buffer (30% w/v PEG 400 and 0.1 M Tris [pH 8.5]). Crystals were frozen in liquid nitrogen. The selenomethionine-labeled data sets and the native data sets were collected at ALS Beamlines 5.0.3 and 5.0.2, respectively. All data set were processed in HKL2000. The structure was solved by single wavelength anomalous dispersion phasing using the program Phenix. All models were iteratively built and refined by using the programs Coot and Refmac5.

Synthetic Protocols

Compounds 4.1-4.5 for Ketoreduction Assays

Also see **Figure 4.8**.

N-Acetylcysteamine

Synthesis followed as previously described (Piasecki et al., 2011).

α-Ketoisocaproyl-N-Acetylcysteamine Thioester (α-KIC-SNAC) **(4.1)**

To a stirred solution of 4-methyl-2-oxopentanoic acid (260 mg, 2.0 mmol, 1.0 eq.) in DMF (8 mL) at 23°C was added *N*,*N*-diisopropylethylamine (0.89 mL, 5.0 mmol, 2.5 eq.) followed by HBTU (830 mg, 2.2 mmol, 1.1 eq.). After 5 min. *N*-acetylcysteamine (0.23 mL, 2.2 mmol, 1.1 eq.) was added in one portion. At 12 hr, the reaction was diluted with 1:1 hexanes:EtOAc (50 mL) and washed with sat. aqueous LiCl (3 x 50 mL). The organic layer was dried over sodium sulfate and concentrated in vacuo to furnish **4.1** (38 mg, 8%) as a pale oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.75 (bs, 1H), 3.47 (q, J = 6.2 Hz, 2H), 3.10 (t, J = 6.7 Hz, 2H), 2.70 (d, J = 6.9 Hz, 2H), 2.14-2.23 (m, 1H), 1.98 (s, 3H), 0.97 (d, J = 6.7 Hz, 6H). ESI-MS expected mass: 232.3; observed mass: 232.2.

Methyl-4-(4-Methyl-2-Oxopentanamido)Butanoate **(4.2a)**

To a stirred solution of 4-methyl-2-oxopentanoic acid (150 mg, 1.15 mmol, 1.0 eq.) in DMF (5.8 mL) at 23°C was added *N*,*N*-diisopropylethylamine (0.51 mL, 4.0 mmol, 3.5 eq.) followed by HBTU (493 mg, 1.30 mmol, 1.1 eq.). The solution was stirred for 5 min and γ-aminobutyric acid methylester hydrochloride (199 mg, 1.30 mmol, 1.1 eq.) was added in one portion. After 12 hr, the reaction was diluted with 1:1 hexanes:EtOAc (100 mL) and washed with sat. aqueous LiCl (3 x 50 mL). The organic layer was dried over sodium sulfate and concentrated in vacuo to furnish **4.2a** (72 mg, 27%) as a pale oil, which

did not require chromatographic purification. For analytical purposes, the material was chromatographed on silica gel eluting with 3:1 hexanes:EtOAc.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.08 (bs, 1H), 3.68 (s, 3H), 3.34 (q, J = 6.8 Hz, 2H), 2.79 (d, J = 6.8 Hz, 2H), 2.37 (t, J = 7.2 Hz, 2H), 2.16 (m, 1H), 1.89 (p, J = 7.2 Hz, 2H), 0.95 (d, J = 6.5 Hz, 6H).

4-(4-Methyl-2-Oxopentanamido)Butanoic Acid **(4.2b)**

To a stirred solution of **4.2a** (72 mg, 0.31 mmol, 1.0 eq.) in THF (1.6 mL) at 0°C was slowly added a solution of 1 M aqueous NaOH (0.47 mL, 0.47 mmol, 1.5 eq.). The reaction was allowed to warm to 23°C and at 5 hr the reaction was cooled to 0°C and a solution of 1 M aqueous HCl (0.5 mL) was added. The reaction was diluted with EtOAc (50 mL) and the aqueous layer was separated. The organic layer was dried over sodium sulfate and concentrated in vacuo. Crude **4.2b** was used immediately without further purification.

α-Ketoisocaproyl-γ-Aminobutyryl-N-Acetylcysteamine Thioester (α-KIC-GABA-SNAC) **(4.2)**

To a stirred solution of crude **4.2b** (160 mg, 0.74 mmol, 1.0 eq.) in DMF (3.7 mL) at 23°C was added *N*,*N*-diisopropylethylamine (0.33 mL, 1.85 mmol, 2.5 eq.) followed by HBTU (340 mg, 0.90 mmol, 1.2 eq.). After 5 min *N*-acetylcysteamine (0.09 mL, 0.90 mmol, 1.2 eq.) was added in one portion. After 12 hr the reaction was diluted with 1:1 hexanes:EtOAc (50 mL) and was washed with sat. aqueous LiCl (3 x 50 mL). The organic layer was dried over sodium sulfate and concentrated in vacuo. The crude material was chromatographed on silica gel eluting with 1:2 hexanes:EtOAc to yield **4.2** (58 mg, 24%) as a pale yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.10 (bs, 1H), 5.79 (bs, 1H), 3.45 (q, J = 12 Hz, J = 5.8 Hz, 2H), 3.33 (qd, J = 6.8 Hz, J = 2.7 Hz, 2H), 3.03 (m, 2H), 2.63 (t, J = 6.8 Hz, 2H), 1.97 (s, 3H), 1.92 (dt, J = 6.8 Hz, J = 2.7 Hz, 1H), 1.84 (m, 1H), 1.66 (m, 2H), 1.53 (m, 2H), 0.97 (d, J = 6.6 Hz, 3H), 0.96 (d, J = 6.6 Hz, 3H). ESI-MS expected mass: 317.4; observed mass: 317.3.

(S)-2-((tert-Butyldimethylsilyl)oxy)-4-Methylpentanoic Acid **(4.3a)**

To a stirred solution of (*S*)-(-)-2-hydroxyisocaproic acid (500 mg, 3.8 mmol, 1.0 eq.) in DMF (2.0 mL) at 23°C was added imidazole (1.24 g, 18.2 mmol, 4.8 eq.) followed by TBS-Cl (1.37 g, 9.1 mmol, 2.4 eq.). The pale yellow solution was stirred for 18 hr and diluted with 100 mL of 1:1 hexanes:EtOAc. The solution was washed with 10% aqueous citric acid (30 mL), saturated $NaHCO_3$ (50 mL), H_2O (50 mL), and then brine (50 mL). The organic layer was dried over sodium sulfate and concentrated in vacuo. The residue was dissolved in MeOH (40 mL), cooled to 0°C, and K_2CO_3 (1.3 g, 9.5 mmol, 2.5 eq.) and water (12 mL) were added. The resulting mixture was stirred for 4 hr at 23°C and the solvent was removed in vacuo. The residue was dissolved in H_2O (20 mL), cooled to 0°C and acidified to pH 4 by addition of 10% aqueous citric acid solution. The aqueous solution was washed with EtOAc (50 mL) and the organic layer was separated. The aqueous layer was washed with EtOAc (50 mL) and the combined organic layers were dried over sodium sulfate and concentrated in vacuo. The crude product was chromatographed on silica gel eluting with 2:1 EtOAc:hexanes to afford **4.3a** (68%, 640 mg) as a pale clear oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.27 (dd, J = 7.2 Hz, J = 5.1 Hz, 1H), 1.82 (m, 1H), 1.67-1.56 (m, 2H), 0.95 (d, J = 6.6 Hz, 3H), 0.94 (s, 9H), 0.94 (d, J = 6.6 Hz, 3H), 0.14 (s, 3H), 0.13 (s, 3H).

Methyl-(S)-4-(2-((tert-Butyldimethylsilyl)oxy)-4-Methylpentanamido)Butanoate **(4.3b)**

To a stirred solution of **4.3a** (140 mg, 0.57 mmol, 1.0 eq.) in DMF (2.9 mL) at 23°C was added *N*,*N*-diisopropylethylamine (0.36 mL, 2 mmol, 3.5 eq.) followed by HBTU (259 mg, 0.68 mmol, 1.2 eq.). After 5 min γ-aminobutyric acid methylester hydrochloride (107 mg, 0.68 mol, 1.2 eq.) was added in one portion. After 12 hr, the reaction was diluted with 1:1 hexanes:EtOAc (100 mL) and washed with sat. aqueous LiCl solution (3 x 50 mL). The organic layer was dried over sodium sulfate and concentrated in vacuo to furnish **4.3b** (150 mg, 76%) as a tan oil which did not require further purification. For analytical purposes, the material was chromatographed on silica gel eluting with 1:1 hexanes:EtOAc (R_f 0.5, 2:1 hexanes:EtOAc).

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.64 (bs, 1H), 4.12 (dd, J = 7.2 Hz, J = 4.8 Hz, 1H), 3.68 (s, 3H), 3.29 (qd, J = 7.2 Hz, J = 3.1 Hz, 2H), 2.36 (t, J = 7.5 Hz, J = 7.2 Hz, 2H), 1.88-1.74 (m, 3H), 1.56 (m, 2H), 0.93 (s, 9H), 0.92 (d, J = 6.5 Hz, 3H), 0.91 (d, J = 6.8 Hz, 3H), 0.09 (s, 3H), 0.06 (s, 3H).

(S)-4-(2-((tert-Butyldimethylsilyl)oxy)-4-Methylpentanamido)Butanoic Acid **(4.3c)**

To a stirred solution of **4.3b** (150 mg, 0.43 mmol, 1.0 eq.) in THF (2.1 mL) at 0°C was added a solution of aqueous NaOH (1.0 M, 1.29 mL, 1.29 mmol, 2.0 eq.) and upon complete addition the reaction was allowed to warm to 23°C. At 5 hr, the reaction was cooled to 0°C and was treated with 1 M HCl (aqueous, 1.4 mL). The reaction was diluted with EtOAc (100 mL) and the aqueous layer was separated. The organic layer was dried over sodium sulfate and concentrated in vacuo to furnish **4.3c** (125 mg, 88%) as a white soft solid. The crude material was used without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.74 (bs, 1H), 4.15 (dd, J = 7.2 Hz, J = 4.8 Hz, 1H), 3.35 (m, 2H), 2.39 (t, J = 7.2 Hz, 2H), 1.91-1.74 (m, 3H), 1.56 (m, 2H), 0.93 (s, 9H), 0.93 (d, J = 6.5 Hz, 3H), 0.91 (d, J = 6.8 Hz, 3H), 0.11 (s, 3H), 0.07 (s, 3H).

S-(2-Acetamidoethyl)-(S)-4-(2-((tert-Butyldimethylsilyl)oxy)-4-Methylpentanamido) Butanethioate **(4.3d)**

To a stirred solution of **4.3c** (162 mg, 0.49 mmol, 1.0 eq.) in DMF (2.5 mL) at 23°C was added *N*,*N*-diisopropylethylamine (0.36 mL, 2.0 mmol, 4.0 eq.) followed by HBTU (224 mg, 0.59 mmol, 1.2 eq.). After 5 min, *N*-acetylcysteamine (0.06 mL, 0.59 mmol, 1.2 eq.) was added in one portion. After 12 hr, the reaction was diluted with 1:1 hexanes:EtOAc and washed with sat. aqueous LiCl (3 x 50 mL). The organic layer was dried over sodium sulfate and concentrated in vacuo. The crude material was chromatographed on silica gel eluting with 1:2 hexanes:EtOAc to afford **4.3d** (128 mg, 60%) as a pale oil (R_f 0.19, 1:2 hexanes:EtOAc).

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.63 (bs, 1H), 6.33 (bs, 1H), 4.13 (dd, J = 6.8 Hz, J = 4.8 Hz, 1H), 3.44 (q, J = 12.3 Hz, J = 6.2 Hz, 2H), 3.29 (qd, J = 6.8 Hz, J = 2.1 Hz, 2H), 3.04 (t, J = 6.2 Hz, J = 5.8 Hz, 2H), 2.61 (t, J = 7.2 Hz, 2H), 1.97 (s, 3H), 1.87 (p, J = 7.2 Hz, 2H), 1.78 (m, 1H), 1.56 (m, 1H), 0.94 (s, 9H), 0.92 (d, J = 6.5 Hz, 3H), 0.91 (d, J = 6.8 Hz, 3H), 0.11 (s, 3H), 0.08 (s, 3H).

(4S)-α-Hydroxyisocaproyl-γ-Aminobutyryl-N-Acetylcysteamine Thioester ((4S)-α-HIC-GABA-SNAC) **(4.3)**

To a stirred solution of **4.3d** (153 mg, 0.37 mmol, 1.0 eq.) in DCM (3.7 mL) at 23°C was added neat TFA (84 μL, 1.1 mmol, 3.0 eq.) in one portion. At 6 hr, the reaction was

concentrated in vacuo. The crude material was chromatographed on silica gel eluting with 3% MeOH in DCM to 4:1 DCM:MeOH to afford **4.3** (93 mg, 82%) as a clear oil (R_f 0.25, 3% MeOH in DCM).

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.67 (bs, 1H), 6.20 (bs, 1H), 4.14 (dd, J = 9.9 Hz, J = 3.1 Hz, 1H), 3.45 (q, J = 12 Hz, J = 5.8 Hz, 2H), 3.33 (qd, J = 6.8 Hz, J = 2.7 Hz, 2H), 3.03 (m, 2H), 2.63 (t, J = 6.8 Hz, 2H), 1.97 (s, 3H), 1.92 (dt, J = 6.8 Hz, J = 2.7 Hz, 1H), 1.84 (m, 1H), 1.66 (m, 2H), 1.53 (m, 2H), 0.97 (d, J = 6.5 Hz, 3H), 0.96 (d, J = 6.5 Hz, 3H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 199.2, 175.2, 170.9, 70.6, 43.7, 40.8, 39.1, 37.7, 28.8, 25.2, 24.4, 23.5, 23.0, 21.3. ESI-MS expected mass: 319.2; observed mass: 341.2 $[M+Na]^+$.

3-Oxopentanoyl-N-Acetylcysteamine Thioester **(4.4)**

Synthesis followed as previously described (Piasecki et al., 2011).

(2RS)-Methyl-3-Oxopentanoyl-N-Acetylcysteamine Thioester **(4.5)**

Synthesis followed as previously described (Piasecki et al., 2011).

Figure 4.8. Synthetic routes to compounds **4.1**, **4.2**, and **4.3**.

Stereocontrol Assays

Ketoreduction assays were set up as previously described (Piasecki et al., 2011) with 1.5 mM substrate dissolved in 10% v/v DMSO, 200 mM HEPES pH 7.5, 100 mM NaCl, 200 mM D-Glucose, 10% v/v Glycerol, 100 μM $NADP^+$, 1 μM GDH (from *B. subtilis*), and 5 μM KR to a total volume of 500 μL and were left at 22°C overnight. Reactions were ethyl acetate extracted (3 x 500 μL) and dried in a speedvac. All samples were resuspended in ethanol prior to chiral chromatographic analysis.

Chiral Chromatography

A ChiralCel OC-H column (250 x 4.6 mm) was used to separate products of ketoreduction on a Beckman Coulter System Gold 126 pump and System Gold 166 PDA detector with a 20 μL loop. Absorbance was tracked at 235 nm. The solvent system and flow rate was optimized to be 14% Ethanol in Hexanes at 1 mL/min for **4.1**, **4.2**, and **4.3**, and 7% Ethanol in Hexanes at 0.8 mL/min for **4.4** and **4.5**.

Accession Numbers

Structure Coordinates were deposited in the Protein Data Bank under accession codes 4J1S (apo) and 4J1Q (with $NADP^+$).

ACKNOWLEDGEMENTS

Instrumentation and technical assistance for crystallographic work were provided by Dr. Art Monzingo and the Macromolecular Crystallography Facility, with financial support from the College of Natural Sciences, the Office of the Executive Vice President and Provost, and the Institute for Cellular and Molecular Biology at the University of Texas at

Austin. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We thank the National Institutes of Health (GM106112) and the Welch Foundation (F-1712) for supporting this research (A.T.K.-C.).

Conclusions and Future Prospects

Employing the catalytic power of PKS enzymes to generate polyketides possessing desired substituents and stereochemistries has long been a goal in biosynthetic engineering. Not only would it allow for the production of new medicines, but it would also aid synthetic organic chemists to access chirally-pure building blocks for more difficult syntheses. In **Chapter 2** we showed that ketoreductase domains are quite promiscuous, with most retaining their natural stereospecificity when given the opportunity to reduce substrate analogs. In **Chapter 3** we presented a new technique for monitoring biocatalytic assays in cell lysate using ^{19}F NMR. In **Chapter 4** we studied the structure of a dual-function ketoreductase able to reduce α- and β-carbonyl groups. These studies have provided the PKS community with the tools to produce a large library of chirally-pure molecules that can be used for further studies. Although this dissertation provides a considerable amount of information on the ketoreductase domain as a robust biocatalyst, it is by no means exhaustive. Future studies will be required to fully understand the full potential of these powerful catalysts. The most promising avenue seems to lie in the *trans*-AT PKS systems, as these synthases perform the most unusual chemistry within their domains. Studying these non-canonical systems will certainly reveal more biosynthetic opportunities for biochemists and chemists, alike.

Appendix

The following pages contain supplemental information for the designated chapter.

CHAPTER 2 SUPPLEMENTAL

Substrate **2.1** data

Run #1 Peak Integrations

Substrate 1	"3R"	"3S"	Total Product
MycKRA	24141	709959	734100
AmpKR10	12507	140313	152820
AmpKR2	283464	116483	399947
OleKR6	0	0	0
PikKR5	0	11228	11228
SpnKR3	0	293154	293154
AmpKR1	0	27829	27829
AmpKR11	0	0	0
TylKR1	407349	0	407349
EryKR1	116266	0	116266
PikKR1	632841	0	632841
AmpKR13	0	0	0

Run #2 Peak Integrations

Substrate 1	"3R"	"3S"	Total Product
MycKRA	49696	1298402	1348098
AmpKR10	56135	409744	465879
AmpKR2	780571	311702	1092273
OleKR6	0	0	0
PikKR5	0	0	0
SpnKR3	0	560493	560493
AmpKR1	0	86956	86956
AmpKR11	0	0	0
TylKR1	1534848	0	1534848
EryKR1	575846	0	575846
PikKR1	1465378	0	1465378
AmpKR13	0	0	0

Product Ratios

Substrate 1	"3R"	"3S"
MycKRA	0.03	0.97
AmpKR10	0.08	0.92
AmpKR2	0.71	0.29
OleKR6	0.00	0.00
PikKR5	0.00	1.00
SpnKR3	0.00	1.00
AmpKR1	0.00	1.00
AmpKR11	0.00	0.00
TylKR1	1.00	0.00
EryKR1	1.00	0.00
PikKR1	1.00	0.00
AmpKR13	0.00	0.00

Product Ratios

Substrate 1	"3R"	"3S"
MycKRA	0.04	0.96
AmpKR10	0.12	0.88
AmpKR2	0.71	0.29
OleKR6	0.00	0.00
PikKR5	0.00	0.00
SpnKR3	0.00	1.00
AmpKR1	0.00	1.00
AmpKR11	0.00	0.00
TylKR1	1.00	0.00
EryKR1	1.00	0.00
PikKR1	1.00	0.00
AmpKR13	0.00	0.00

Product Ratio Average

Substrate 1	"3R"	"3S"
MycKRA	3.5%	96.5%
AmpKR10	10.1%	89.9%
AmpKR2	71.2%	28.8%
OleKR6	0.0%	0.0%
PikKR5	0.0%	50.0%
SpnKR3	0.0%	100.0%
AmpKR1	0.0%	100.0%
AmpKR11	0.0%	0.0%
TylKR1	100.0%	0.0%
EryKR1	100.0%	0.0%
PikKR1	100.0%	0.0%
AmpKR13	0.0%	0.0%

Turnover Efficiency

Substrate 1	"3R"	"3S"	Fraction Reduced
MycKRA	3.1%	84.8%	0.88
AmpKR10	3.1%	27.3%	0.30
AmpKR2	50.6%	20.5%	0.71
OleKR6	0.0%	0.0%	0.00
PikKR5	0.0%	0.0%	0.00
SpnKR3	0.0%	36.5%	0.37
AmpKR1	0.0%	5.7%	0.06
AmpKR11	0.0%	0.0%	0.00
TylKR1	100.0%	0.0%	1.00
EryKR1	37.5%	0.0%	0.38
PikKR1	95.5%	0.0%	0.95
AmpKR13	0.0%	0.0%	0.00

Substrate **2.2** data

Run #1

Peak Integrations

Substrate 2	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product
MycKR4	34455	5967855	67348	0	6069658
AmpKR10	0	531141	0	0	531141
AmpKR2	0	12224	4039081	12154	4063459
OleKR6	0	28934	60388	0	89322
PikKR5	15689	849839	295145	23734	1184407
SpnKR3	9976	834603	908530	726267	2479376
AmpKR1	0	6184237	0	0	6184237
AmpKR11	0	33253	0	0	33253
TylKR1	2642501	8421921	0	28062	11092484
EryKR1	0	108205	0	8256894	8365099
PikKR1	0	0	0	7305498	7305498
AmpKR13	0	0	0	0	0

Product Ratios

Substrate 2	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKR4	0.01	0.98	0.01	0.00
AmpKR10	0.00	1.00	0.00	0.00
AmpKR2	0.00	0.00	0.99	0.00
OleKR6	0.00	0.32	0.68	0.00
PikKR5	0.01	0.72	0.25	0.02
SpnKR3	0.00	0.34	0.37	0.29
AmpKR1	0.00	1.00	0.00	0.00
AmpKR11	0.00	1.00	0.00	0.00
TylKR1	0.24	0.76	0.00	0.00
EryKR1	0.00	0.01	0.00	0.99
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Product Ratio Average

Substrate 2	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKR4	0.6%	83.3%	15.2%	0.7%
AmpKR10	22.5%	73.8%	0.0%	3.7%
AmpKR2	0.0%	1.6%	96.8%	1.6%
OleKR6	0.0%	31.8%	68.2%	0.0%
PikKR5	1.4%	68.7%	26.4%	3.4%
SpnKR3	0.6%	33.4%	36.8%	29.2%
AmpKR1	0.0%	99.2%	0.8%	0.0%
AmpKR11	0.0%	60.0%	40.0%	0.0%
TylKR1	24.0%	75.6%	0.1%	0.3%
EryKR1	0.2%	1.3%	0.0%	98.4%
PikKR1	0.1%	0.0%	0.0%	99.9%
AmpKR13	0.0%	0.0%	0.0%	0.0%

Run #2

Peak Integrations

Substrate 2	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product	Total Substrate	Total Peak Integration
MycKR4	48205	4796457	2054757	94677	6994096	45328	7039424
AmpKR10	219428	232331	0	35778	487537	1463924	1951461
AmpKR2	0	89294	2951251	90468	3131013	562403	3693416
OleKR6	0	40700	89855	0	130555	1659793	1790348
PikKR5	19166	822768	349507	60096	1251537	1458324	2709861
SpnKR3	24442	1014314	1134557	895574	3068887	1840616	4909503
AmpKR1	0	5906246	95930	0	6002176	0	6002176
AmpKR11	0	57882	231025	0	288907	3351607	3640514
TylKR1	1905978	5936454	9197	27284	7878913	68376	7947289
EryKR1	26572	71501	0	5217603	5315676	0	5315676
PikKR1	11810	0	0	7840200	7852010	0	7852010
AmpKR13	0	0	0	0	0	4091597	4091597

Product Ratios

Substrate 2	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKR4	0.01	0.69	0.29	0.01
AmpKR10	0.45	0.48	0.00	0.07
AmpKR2	0.00	0.03	0.94	0.03
OleKR6	0.00	0.31	0.69	0.00
PikKR5	0.02	0.66	0.28	0.05
SpnKR3	0.01	0.33	0.37	0.29
AmpKR1	0.00	0.98	0.02	0.00
AmpKR11	0.00	0.20	0.80	0.00
TylKR1	0.24	0.75	0.00	0.00
EryKR1	0.00	0.01	0.00	0.98
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Turnover Efficiency

Substrate 2	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Fraction Reduced
MycKR4	0.6%	82.9%	15.1%	0.7%	0.99
AmpKR10	5.6%	18.4%	0.0%	0.9%	0.25
AmpKR2	0.0%	1.3%	82.1%	1.4%	0.85
OleKR6	0.0%	2.3%	5.0%	0.0%	0.07
PikKR5	0.7%	31.8%	12.2%	1.6%	0.46
SpnKR3	0.4%	20.9%	23.0%	18.3%	0.63
AmpKR1	0.0%	99.2%	0.8%	0.0%	1.00
AmpKR11	0.0%	4.8%	3.2%	0.0%	0.08
TylKR1	23.8%	75.0%	0.1%	0.3%	0.99
EryKR1	0.2%	1.3%	0.0%	98.4%	1.00
PikKR1	0.1%	0.0%	0.0%	99.9%	1.00
AmpKR13	0.0%	0.0%	0.0%	0.0%	0.00

Substrate **2.3** data

Run #1

Peak Integrations

Substrate 3	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product
MycKRA	0	6738819	2243599	0	8982418
AmpKR10	217714	0	0	0	217714
AmpKR2	0	0	2825945	193990	3019935
OleKR6	0	0	0	0	0
PikKR5	0	206439	72207	0	278646
SpnKR3	0	0	570820	222449	793269
AmpKR1	0	3202475	0	0	3202475
AmpKR11	0	34239	0	0	34239
TylKR1	3570056	335796	0	99594	4005446
EryKR1	0	0	0	2492606	2492606
PikKR1	0	0	0	26590810	26590810
AmpKR13	0	0	0	0	0

Product Ratios

Substrate 3	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKRA	0.00	0.75	0.25	0.00
AmpKR10	1.00	0.00	0.00	0.00
AmpKR2	0.00	0.00	0.94	0.06
OleKR6	0.00	0.00	0.00	0.00
PikKR5	0.00	0.74	0.26	0.00
SpnKR3	0.00	0.00	0.72	0.28
AmpKR1	0.00	1.00	0.00	0.00
AmpKR11	0.00	1.00	0.00	0.00
TylKR1	0.89	0.08	0.00	0.02
EryKR1	0.00	0.00	0.00	1.00
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Product Ratio Average

Substrate 3	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKRA	0.0%	74.6%	25.4%	0.0%
AmpKR10	97.1%	1.4%	1.4%	0.0%
AmpKR2	0.0%	1.3%	93.2%	5.3%
OleKR6	0.0%	25.7%	24.3%	0.0%
PikKR5	0.0%	75.7%	24.3%	0.0%
SpnKR3	0.0%	0.0%	73.0%	27.0%
AmpKR1	0.0%	100.0%	0.0%	0.0%
AmpKR11	0.0%	100.0%	0.0%	0.0%
TylKR1	89.5%	7.9%	0.0%	2.7%
EryKR1	0.0%	0.0%	0.0%	100.0%
PikKR1	0.0%	0.0%	0.0%	100.0%
AmpKR13	0.0%	0.0%	0.0%	0.0%

Run #2

Peak Integrations

Substrate 3	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product	Total Substrate	Total Peak Integration
MycKRA	0	4086966	1421019	0	5507985	0	5507985
AmpKR10	893707	27401	26766	0	947874	11085156	12033030
AmpKR2	0	118752	4209643	207032	4535427	1274610	5810037
OleKR6	0	24353	23048	0	47401	5214147	5261548
PikKR5	0	477394	139720	0	617114	5059336	5676450
SpnKR3	0	0	681534	238761	920295	5913389	6833684
AmpKR1	0	3008667	0	0	3008667	2850400	5859067
AmpKR11	0	94269	0	0	94269	6736761	6830030
TylKR1	4153381	339881	0	132597	4625859	0	4625859
EryKR1	0	0	0	3719661	3719661	1923999	5643660
PikKR1	0	0	0	6471169	6471169	0	6471169
AmpKR13	0	0	0	0	0	6004328	6004328

Product Ratios

Substrate 3	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKRA	0.00	0.74	0.26	0.00
AmpKR10	0.94	0.03	0.03	0.00
AmpKR2	0.00	0.03	0.93	0.05
OleKR6	0.00	0.51	0.49	0.00
PikKR5	0.00	0.77	0.23	0.00
SpnKR3	0.00	0.00	0.74	0.26
AmpKR1	0.00	1.00	0.00	0.00
AmpKR11	0.00	1.00	0.00	0.00
TylKR1	0.90	0.07	0.00	0.03
EryKR1	0.00	0.00	0.00	1.00
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Turnover Efficiency

Substrate 3	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Fraction Reduced
MycKRA	0.0%	74.6%	25.4%	0.0%	1.00
AmpKR10	7.7%	0.1%	0.1%	0.0%	0.08
AmpKR2	0.0%	1.0%	72.8%	4.3%	0.78
OleKR6	0.0%	0.2%	0.2%	0.0%	0.01
PikKR5	0.0%	8.2%	2.6%	0.0%	0.11
SpnKR3	0.0%	0.0%	9.8%	3.6%	0.13
AmpKR1	0.0%	51.4%	0.0%	0.0%	0.51
AmpKR11	0.0%	1.4%	0.0%	0.0%	0.01
TylKR1	89.5%	7.9%	0.0%	2.7%	1.00
EryKR1	0.0%	0.0%	0.0%	65.9%	0.66
PikKR1	0.0%	0.0%	0.0%	100.0%	1.00
AmpKR13	0.0%	0.0%	0.0%	0.0%	0.00

Substrate **2.4** data

Run #1

Peak Integrations

Substrate 4	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product
MycKR4	0	1741559	1559402	0	3300961
AmpKR10	0	0	0	0	0
AmpKR2	0	0	4630157	212884	4843041
OleKR6	0	0	0	0	0
PikKR5	0	0	0	0	0
SpnKR3	658039	0	0	1534602	2192641
AmpKR1	0	307379	0	0	307379
AmpKR11	0	530683	0	0	530683
TylKR1	4317365	0	0	161670	4479035
EryKR1	0	0	0	942111	942111
PikKR1	0	0	0	4915578	4915578
AmpKR13	0	0	0	0	0

Product Ratios

Substrate 4	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKR4	0.00	0.53	0.47	0.00
AmpKR10	0.00	0.00	0.00	0.00
AmpKR2	0.00	0.00	0.96	0.04
OleKR6	0.00	0.00	0.00	0.00
PikKR5	0.00	0.00	0.00	0.00
SpnKR3	0.30	0.00	0.00	0.70
AmpKR1	0.00	1.00	0.00	0.00
AmpKR11	0.00	1.00	0.00	0.00
TylKR1	0.96	0.00	0.00	0.04
EryKR1	0.00	0.00	0.00	1.00
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Product Ratio Average

Substrate 4	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKR4	0.0%	53.3%	46.7%	0.0%
AmpKR10	0.0%	0.0%	0.0%	0.0%
AmpKR2	0.0%	0.0%	95.7%	4.3%
OleKR6	0.0%	0.0%	0.0%	0.0%
PikKR5	0.0%	0.0%	0.0%	0.0%
SpnKR3	31.6%	0.0%	0.0%	68.4%
AmpKR1	0.0%	100.0%	0.0%	0.0%
AmpKR11	0.0%	100.0%	0.0%	0.0%
TylKR1	96.5%	0.0%	0.0%	3.5%
EryKR1	0.0%	0.0%	0.0%	100.0%
PikKR1	0.0%	0.0%	0.0%	100.0%
AmpKR13	0.0%	0.0%	0.0%	0.0%

Run #2

Peak Integrations

Substrate 4	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product	Total Substrate	Total Peak Integration
MycKR4	0	2277407	1946219	0	4223626	306942	4530568
AmpKR10	0	0	0	0	0	3864355	3864355
AmpKR2	0	0	4474913	197504	4672417	0	4672417
OleKR6	0	0	0	0	0	3867559	3867559
PikKR5	0	0	0	0	0	3669583	3669583
SpnKR3	648667	0	0	1302047	1950714	1012238	2962952
AmpKR1	0	348022	0	0	348022	3382896	3730918
AmpKR11	0	674669	0	0	674669	3352816	4027485
TylKR1	4356809	0	0	156734	4512543	0	4512543
EryKR1	0	0	0	1053067	1053067	2890307	3943374
PikKR1	0	0	0	4806501	4806501	0	4806501
AmpKR13	0	0	0	0	0	3913983	3913983

Product Ratios

Substrate 4	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKR4	0.00	0.54	0.46	0.00
AmpKR10	0.00	0.00	0.00	0.00
AmpKR2	0.00	0.00	0.96	0.04
OleKR6	0.00	0.00	0.00	0.00
PikKR5	0.00	0.00	0.00	0.00
SpnKR3	0.33	0.00	0.00	0.67
AmpKR1	0.00	1.00	0.00	0.00
AmpKR11	0.00	1.00	0.00	0.00
TylKR1	0.97	0.00	0.00	0.03
EryKR1	0.00	0.00	0.00	1.00
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Turnover Efficiency

Substrate 4	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Fraction Reduced
MycKR4	0.0%	49.7%	43.5%	0.0%	0.93
AmpKR10	0.0%	0.0%	0.0%	0.0%	0.00
AmpKR2	0.0%	0.0%	95.7%	4.3%	1.00
OleKR6	0.0%	0.0%	0.0%	0.0%	0.00
PikKR5	0.0%	0.0%	0.0%	0.0%	0.00
SpnKR3	20.8%	0.0%	0.0%	45.0%	0.66
AmpKR1	0.0%	9.3%	0.0%	0.0%	0.09
AmpKR11	0.0%	16.8%	0.0%	0.0%	0.17
TylKR1	96.5%	0.0%	0.0%	3.5%	1.00
EryKR1	0.0%	0.0%	0.0%	26.7%	0.27
PikKR1	0.0%	0.0%	0.0%	100.0%	1.00
AmpKR13	0.0%	0.0%	0.0%	0.0%	0.00

Substrate **2.5** data

Run #1

Peak Integrations

Substrate 5	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product
MycKRA	0	1580902	1088660	0	2669562
AmpKR10	0	0	65433	0	65433
AmpKR2	0	0	2659330	27145	2686475
OleKR6	0	0	0	0	0
PikKR5	0	0	0	0	0
SpnKR3	0	0	0	0	0
AmpKR1	0	374427	0	0	374427
AmpKR11	0	395274	0	0	395274
TylKR1	3182085	0	0	0	3182085
EryKR1	0	0	2819669	0	2819669
PikKR1	0	0	3410703	0	3410703
AmpKR13	0	0	0	0	0

Product Ratios

Substrate 5	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKRA	0.00	0.59	0.41	0.00
AmpKR10	0.00	0.00	1.00	0.00
AmpKR2	0.00	0.00	0.99	0.01
OleKR6	0.00	0.00	0.00	0.00
PikKR5	0.00	0.00	0.00	0.00
SpnKR3	0.00	0.00	0.00	0.00
AmpKR1	0.00	1.00	0.00	0.00
AmpKR11	0.00	1.00	0.00	0.00
TylKR1	1.00	0.00	0.00	0.00
EryKR1	0.00	0.00	0.00	1.00
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Product Ratio Average

Substrate 5	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKRA	0.0%	58.2%	41.8%	0.0%
AmpKR10	0.0%	0.0%	100.0%	0.0%
AmpKR2	0.0%	0.0%	99.1%	0.9%
OleKR6	0.0%	0.0%	0.0%	0.0%
PikKR5	0.0%	0.0%	0.0%	0.0%
SpnKR3	0.0%	0.0%	0.0%	0.0%
AmpKR1	0.0%	100.0%	0.0%	0.0%
AmpKR11	0.0%	100.0%	0.0%	0.0%
TylKR1	100.0%	0.0%	0.0%	0.0%
EryKR1	0.0%	0.0%	0.0%	100.0%
PikKR1	0.0%	0.0%	0.0%	100.0%
AmpKR13	0.0%	0.0%	0.0%	0.0%

Run #2

Peak Integrations

Substrate 5	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Total Product	Total Substrate	Total Peak Integration
MycKRA	0	1158562	868676	0	2027238	148466	2175704
AmpKR10	0	0	279207	0	279207	1703579	1982786
AmpKR2	0	0	2673960	21212	2695172	14326	2709498
OleKR6	0	0	0	0	0	1398257	1398257
PikKR5	0	0	0	0	0	450836	450836
SpnKR3	0	335963	0	0	335963	6907801	6907801
AmpKR1	0	395425	0	0	395425	942681	1278644
AmpKR11	0	395425	0	0	395425	1702465	2097890
TylKR1	2944647	0	0	0	2944647	0	2944647
EryKR1	0	0	0	3066996	3066996	0	3066996
PikKR1	0	0	0	3116490	3116490	0	3116490
AmpKR13	0	0	0	0	0	2063477	2063477

Product Ratios

Substrate 5	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"
MycKRA	0.00	0.57	0.43	0.00
AmpKR10	0.00	0.00	1.00	0.00
AmpKR2	0.00	0.00	0.99	0.01
OleKR6	0.00	0.00	0.00	0.00
PikKR5	0.00	0.00	0.00	0.00
SpnKR3	0.00	0.00	0.00	0.00
AmpKR1	0.00	1.00	0.00	0.00
AmpKR11	0.00	1.00	0.00	0.00
TylKR1	1.00	0.00	0.00	0.00
EryKR1	0.00	0.00	0.00	1.00
PikKR1	0.00	0.00	0.00	1.00
AmpKR13	0.00	0.00	0.00	0.00

Turnover Efficiency

Substrate 5	"2R,3R"	"2S,3S"	"2R,3S"	"2S,3R"	Fraction Reduced
MycKRA	0.0%	54.2%	39.0%	0.0%	0.93
AmpKR10	0.0%	0.0%	14.1%	0.0%	0.14
AmpKR2	0.0%	0.0%	98.6%	0.9%	0.99
OleKR6	0.0%	0.0%	0.0%	0.0%	0.00
PikKR5	0.0%	0.0%	0.0%	0.0%	0.00
SpnKR3	0.0%	0.0%	0.0%	0.0%	0.00
AmpKR1	0.0%	26.3%	0.0%	0.0%	0.26
AmpKR11	0.0%	18.8%	0.0%	0.0%	0.19
TylKR1	100.0%	0.0%	0.0%	0.0%	1.00
EryKR1	0.0%	0.0%	0.0%	100.0%	1.00
PikKR1	0.0%	0.0%	0.0%	100.0%	1.00
AmpKR13	0.0%	0.0%	0.0%	0.0%	0.00

CHAPTER 3 SUPPLEMENTAL

Kinetic data for EryTE

Substrate **3.1** (3.75mM)

Trial 1

^{19}F-NMR data	
Time (min)	Substrate left (mM)
15	3.05
20	2.92
25	2.52
30	2.31
35	1.88
40	1.79
45	1.65
50	1.52
55	1.22
60	1.20

HPLC data	
Time (min)	Substrate left (mM)
0	3.75
10	3.31
20	2.47
30	1.62
40	1.51
50	1.41
60	1.15

Trial 2

Time (min)	Substrate left (mM)
15	2.98
20	2.48
25	2.47
30	2.13
35	1.99
40	1.81
45	1.60
50	1.42
55	1.51
60	1.10

Time (min)	Substrate left (mM)
0	3.75
10	2.91
20	2.51
30	2.10
40	1.39
50	1.47
60	1.08

	^{19}F-NMR		HPLC	
	Trial 1	Trial 2	Trial 1	Trial 2
v_o (mM/min)	0.047	0.042	0.051	0.055
Rate (min^{-1})	0.408	0.361	0.446	0.482
Avg Rate (min^{-1})	**0.384**		**0.464**	
St Deva	**0.033**		**0.025**	

[EryTE] 0.115 mM
[S] 3.75 mM

Substrate **3.1** (7.5mM)

Trial 1

^{19}F-NMR data	
Time (min)	Substrate left (mM)
15	6.16
20	5.47
25	5.05
30	4.55
35	3.91
40	3.61
45	3.14
50	2.76
55	2.42
60	2.27

HPLC data	
Time (min)	Substrate left (mM)
0	7.50
10	5.17
20	4.32
30	3.14
40	2.58
50	2.20
60	1.74

Trial 2

Time (min)	Substrate left (mM)
15	6.07
20	5.87
25	5.27
30	4.72
35	4.37
40	3.90
45	3.54
50	3.58
55	3.21
60	2.94

Time (min)	Substrate left (mM)
0	7.50
10	6.33
20	4.80
30	4.25
40	3.02
50	2.86
60	2.41

	^{19}F-NMR		HPLC	
	Trial 1	Trial 2	Trial 1	Trial 2
v_0 (mM/min)	0.096	0.089	0.101	0.096
Rate (min^{-1})	0.837	0.772	0.880	0.837
Avg Rate (min^{-1})	**0.805**		**0.858**	
St Deva	**0.046**		**0.031**	

[E*pyTE*] 0.115 mM
[S] 7.5 mM

Substrate **3.1** (15mM)

Trial 1

^{19}F-NMR data	
Time (min)	Substrate left (mM)
15	12.56
20	12.53
25	11.20
30	10.53
35	9.80
40	8.85
45	8.12
50	7.64
55	6.88
60	6.22

HPLC data	
Time (min)	Substrate left (mM)
0	15.00
10	13.01
20	10.01
30	8.49
40	6.82
50	6.36
60	5.48

Trial 2

Time (min)	Substrate left (mM)
15	12.76
20	11.92
25	11.12
30	10.47
35	9.84
40	8.90
45	8.37
50	7.98
55	7.24
60	6.78

Time (min)	Substrate left (mM)
0	15.00
10	13.06
20	10.95
30	9.50
40	7.15
50	7.42
60	6.67

[Ery7E] 0.115 mM
[S] 15 mM

	^{19}F-NMR		HPLC	
	Trial 1	Trial 2	Trial 1	Trial 2
v_o (mM/min)	0.152	0.133	0.161	0.163
Rate (min^{-1})	1.323	1.154	1.399	1.417
Avg Rate (min^{-1})	**1.238**		**1.408**	
St Deva	**0.119**		**0.013**	

Substrate **3.1** (30mM)

Trial 1

^{19}F-NMR data	
Time (min)	Substrate left (mM)
15	26.64
20	25.37
25	23.98
30	22.71
35	21.39
40	20.57
45	19.78
50	18.67
55	17.83
60	16.83

Trial 2

Time (min)	Substrate left (mM)
15	26.78
20	27.88
25	26.56
30	25.29
35	24.02
40	23.03
45	22.00
50	21.07
55	20.10
60	19.44

HPLC data	
Time (min)	Substrate left (mM)
0	30.00
10	28.19
20	22.73
30	19.45
40	18.05
50	17.70
60	15.78

Time (min)	Substrate left (mM)
0	30.00
10	32.70
20	29.19
30	26.34
40	20.19
50	24.03
60	19.89

	^{19}F-NMR		HPLC	
	Trial 1	Trial 2	Trial 1	Trial 2
v_0 (mM/min)	0.227	0.212	0.244	0.245
Rate (min^{-1})	1.972	1.844	2.122	2.130
Avg Rate (min^{-1})	**1.908**		**2.126**	
St Deva	**0.090**		**0.006**	

[EryTE] 0.115 mM
[S] 30 mM

Substrate **3.1** (60mM)

Trial 1

^{19}F-NMR data	
Time (min)	Substrate left (mM)
15	55.66
20	57.48
25	55.14
30	53.58
35	51.76
40	49.68
45	48.12
50	46.56
55	45.00
60	43.96

HPLC data	
Time (min)	Substrate left (mM)
0	60.00
10	55.93
20	53.28
30	43.49
40	43.01
50	35.67
60	39.58

Trial 2

Time (min)	Substrate left (mM)
15	57.28
20	55.03
25	53.53
30	51.78
35	50.03
40	48.53
45	47.52
50	46.52
55	45.27
60	44.27

Time (min)	Substrate left (mM)
0	60.00
10	53.94
20	41.09
30	43.62
40	39.32
50	42.40
60	38.41

	^{19}F-NMR		HPLC	
	Trial 1	Trial 2	Trial 1	Trial 2
v_o (mM/min)	0.354	0.327	0.386	0.370
Rate (min^{-1})	3.080	2.843	3.356	3.213
Avg Rate (min^{-1})	**2.962**		**3.284**	
St Deva	**0.167**		**0.101**	

[EryTE] 0.115 mM
[S] 60 mM

Substrate **3.1** (120mM)

Trial 1

^{19}F-NMR data	
Time (min)	Substrate left (mM)
15	112.80
20	109.53
25	107.57
30	105.28
35	102.99
40	101.36
45	99.39
50	98.09
55	95.80
60	94.49

HPLC data	
Time (min)	Substrate left (mM)
0	120.00
10	116.33
20	115.48
30	104.85
40	103.21
50	99.69
60	96.30

Trial 2

Time (min)	Substrate left (mM)
15	118.30
20	124.94
25	122.53
30	119.51
35	117.09
40	116.49
45	113.77
50	111.96
55	109.85
60	108.34

Time (min)	Substrate left (mM)
0	120.00
10	117.48
20	115.91
30	111.71
40	106.31
50	100.22
60	98.19

	^{19}F-NMR		HPLC	
	Trial 1	Trial 2	Trial 1	Trial 2
v_o (mM/min)	0.378	0.409	0.417	0.391
Rate (min^{-1})	3.289	3.560	3.623	3.402
Avg Rate (min^{-1})	**3.424**		**3.512**	
St Deva	**0.192**		**0.156**	

[Ery'E] 0.115 mM
[S] 120 mM

Kinetic data

for TylKR1

Substrate **3.3**

(30mM)

Trial 1

^{19}F-NMR data	
Time (min)	Substrate left (mM)
15	26.89
20	25.56
25	24.40
30	23.41
35	22.47
40	21.70
45	20.82
50	20.87
55	20.15
60	19.82

HPLC data	
Time (min)	Substrate left (mM)
0	30.00
10	25.05
20	23.36
30	21.98
40	20.40
50	20.86
60	20.45

Trial 2

Time (min)	Substrate left (mM)
15	28.37
20	27.29
25	26.39
30	25.71
35	25.48
40	25.21
45	24.90
50	24.40
55	24.13
60	24.18

*Performed at 1/2[E]

Time (min)	Substrate left (mM)
0	30.00
10	23.40
20	21.54
30	20.83
40	19.07
50	20.00
60	19.29

	^{19}F-NMR		HPLC	
	Trial 1	Trial 2*	Trial 1	Trial 2
V_o (mM/min)	0.199	0.104	0.223	0.244
St Deva	0.008	0.012	0.040	0.063
Avg v_o (mM/min)	**0.203**		**0.234**	
Avg St Deva	**0.016**		**0.051**	
Rate (min^{-1})	2.45	2.55	2.74	3.00
Avg Rate (min^{-1})	2.50		2.87	
stdeva	0.07		0.19	

[TylKR1] 0.081 mM
*[TylKR1] 0.041 mM
[S] 30 mM

References

Akey, D.L., Razelun, J.R., Tehranisa, J., Sherman, D.H., Gerwick, W.H., and Smith, J.L. (2010) Crystal structures of dehydratase domains from the curacin polyketide biosynthetic pathway. Structure *18*, 94-105.

Ashley, G.W., Burlingame, M., Desai, R., Fu, H., Leaf, T., Licari, P.J., Tran, C., Abbanat, D., Bush, K., and Macielag, M. (2006) Preparation of erythromycin analogs having functional groups at C-15. J. Antibiot. (Tokyo) *59*, 392.

Baerga-Ortiz, A., Popovic, B., Siskos, A.P., O'Hare, H.M., Spiteller, D., Williams, M.G., Campillo, N., Spencer, J.B., and Leadlay, P.F. (2006) Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase. Chem. Biol. *13*, 277-285.

Bali, S. and Weissman, K.J. (2006) Ketoreduction in mycolactone biosynthesis: insight into substrate specificity and stereocontrol from studies of discrete ketoreductase domains in vitro. ChemBioChem *7*, 1935-1942.

Bali, S., O'Hare, H.M., and Weissman, K.J. (2006) Broad substrate specificity of ketoreductases derived from modular polyketide synthases. ChemBioChem *7*, 478-484.

Bonnet, S.A., Whicher, J.R., Papireddy, K., Florova, G., Smith, J.L., and Reynolds, K.A. (2013) Structural and stereochemical analysis of a modular polyketide synthase ketoreductase domain required for the generation of a cis-alkene. Chem. Biol. *20*, 772-783.

Bruice, T.C., Bruno, J.J., and Chou, W. (1963) Nucleophilic displacement reactions at the thiol-ester bond of δ-thiolvalerolactone. J. Am. Chem. Soc. *85*, 1659–1669.

Butcher, R.A., Schroeder, F.C., Fischbach, M.A., Straight, P.D., Kolter, R., Walsh, C.T., and Clardy, J. (2007) The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *104*, 1506-1509.

Caffrey, P. (2003) Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. ChemBioChem *4*, 654–657.

Calderone, C.T., Kowtoniuk, W.E., Kelleher, N.L., Walsh, C.T., and Dorrestein, P.C. (2006) Convergence of isoprene and polyketide biosynthetic machinery: Isoprenyl-S-carrier proteins in the pksX pathway of *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *103*, 8977-8982.

Calderone, C.T., Bumpus, S.B., Kelleher, N.L., Walsh, C.T., and Magarvey, N.A. (2008) A ketoreductase domain in the PksJ protein of the bacillaene assembly line carries out both α- and β-ketone reduction during chain growth. Proc. Natl. Acad. Sci. USA *105*, 12809-12814.

Cane, D.E., Walsh, C.T., and Khosla, C. (1998) Harnessing the biosynthetic code: combination, permutations, and mutations. Science *282*, 63-68.

Castonguay, R., He, W., Chen, A.Y., Khosla, C., and Cane, D.E. (2007) Stereospecificity of ketoreductase domains of the 6-deoxyerythronolide B synthase. J. Am. Chem. Soc. *129*, 13758-13769.

Chan, Y.A., Podevels, A.M., Kevany, B.M., and Thomas, M.G. (2009) Biosynthesis of polyketide synthase extender units. Nat. Prod. Rep. *26*, 90-114.

Chan, Y.A. and Thomas, M.G. (2010) Recognition of (2*S*)-aminomalonyl-acyl carrier protein (ACP) and (2*R*)-hydroxymalonyl-ACP by acyltransferases in zwittermicin A biosynthesis. Biochemistry *49*, 3667-3677.

Chen, A.Y., Schnarr, N.A., Kim, C.Y., Cane, D.E., and Khosla, C. (2006) Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-deoxyerythronolide B synthase. J. Am. Chem. Soc. *128*, 3067–3074.

Chen, X.H., Vater, J., Piel, J., Franke, P., Scholz, R., Schneider, K., Koumoutsi, A., Hitzeroth, G., Grammel, N., Strittmatter, A.W., Gottschalk, G., Süssmuth, R.D., and Borriss, R. (2006) Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB 42. J. Bacteriol. 1*88*, 4024-4036.

Cheng, Y.Q., Coughlin, J.M., Lim, S.K., and Shen, B. (2009) Type I polyketide synthases that require discrete acyltransferases. Methods Enzymol. *459*, 165-186.

Crimmins, M.T. and Slade, D.J. (2006) Formal synthesis of 6-deoxyerythronolide B. Org. Lett. *8*, 2191–2194.

Cuskin, F., Solovyova, A.S., Lewis, R.J., and Race, P.R. (2011) Crystallization and preliminary X-ray analysis of the bacillaene synthase *trans*-acting acyltransferase PksC. Acta. Cryst. *67*, 464-466.

Dalvit, C., Ardini, E., Flocco, M., Fogliatto, G. P., Mongelli, N., and Veronesi, M. (2003) A general NMR method for rapid, efficient, and reliable biochemical screening. J. Am. Chem. Soc. *125*, 14620-14625.

Del Vecchio, F., Petkovic, H., Kendrew, S.G., Low, L., Wilkinson, B., Lill, R., Cortés, J., Rudd, B.A.M., Staunton, J., and Leadlay, P.F. (2003) Active-site residue, domain and module swaps in modular polyketide synthases. J. Ind. Microbiol. Biotechnol. *30*, 489-494.

Dong, X., Williams, C., Wattana-amorn, P., and Crump, M.P. (unpublished) Structural studies on Mupirocin didomain acyl carrier protein.

Dunn, B.J. and Khosla, C. (2013) Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. J. R. Soc. Interface *6*, (in print).

Evans, D.A., Bartroli, J., and Shih, T.L. (1981) Enantioselective aldol condensations. 2. Erythro-selective chiral aldol condensations via boron enolates. J. Am. Chem. Soc. *103*, 2127–2129.

Evans, D.A., Kim, A.S., Metternich, R., and Novack, V.J. (1998) General strategies toward the syntheses of macrolide antibiotics. The total syntheses of 6- deoxyerythronolide B and oleandolide. J. Am. Chem. Soc. *120*, 5921–5942.

Gao, X., Woo, S.K., and Krische, M.J. (2013) Total synthesis of 6-deoxyerythronolide B via C-C bond-forming transfer hydrogenation. J. Am. Chem. Soc. *135*, 4223-4226.

Gilbert, I.H., Ginty, M., O'Neill, J.A., Simpson, T.J., Staunton, J., and Willis, C.L. (1995) Synthesis of β-keto and α,β-unsaturated *N*-acetylcysteamine thioesters. Bioorg. Med. Chem. *5*, 1587–1590.

Gokhale, R.S., Hunziker, D., Cane, D.E., and Khosla, C. (1999) Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase. Chem. Biol. *6*, 117-125.

Goss, R. J. M. and Hong, H. (2005) A novel fluorinated erythromycin antibiotic. Chem. Commun. *31*, 3983-3985.

Guo, X., Liu, T., Valenzano, C.R., Deng, Z., and Cane, D.E. (2010) Mechanism and stereospecificity of a fully saturating polyketide synthase module: nanchangmycin synthase module 2 and its dehydratase domain. J. Am. Chem. Soc. *132*, 14694-14696.

Hans, M., Hornung, A., Dziarnowski, A., Cane, D.E., and Khosla, C. (2003) Mechanistic analysis of acyl transferase domain exchange in polyketide synthase modules. J. Am. Chem. Soc. *125*, 5366-5374.

Hildebrand, M., Waggoner, L.E., Liu, H., Sudek, S., Allen, S., Anderson, C., Sherman, D.H., and Haygood, M. (2004) bryA: An unusual modular polyketide synthase gene from the uncultivated bacterial symbiont of the marine bryozoan *Bugula neritina*. Chem. Biol. *11*, 1543-1552.

Hilterhaus, L. and Liese, A. (2007) Building blocks. Adv. Biochem. Eng. Biotechnol. *105*, 133–173.

Holzbaur, I.E., Harris, R.C., Bycroft, M., Cortes, J., Bisang, C., Staunton, J., Rudd, B.A., and Leadlay, P.F. (1999) Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase. Chem. Biol. *6*, 189-195.

Holzbaur, I.E., Ranganathan, A., Thomas, I.P., Kearney, D.J., Reather, J.A., Rudd, B.A., Staunton, J., and Leadlay, P.F. (2001) Molecular basis of Celmer's rules: role of the

ketosynthases domain in epimerization and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. Chem. Biol. *8*, 329-340.

Hughes, A.J. and Keatinge-Clay, A.T. (2011) Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. Chem. Biol. *18*, 165–176.

Hughes, A.J. and Keatinge-Clay, A.T. (2012) Employing a polyketide synthase module and thioesterase in the semipreparative biocatalysis of diverse triketide pyrones. Med. Chem. Commun. *3*, 956-959.

Hughes, A.J., Tibby, M.R., Brantley, J.N., and Keatinge-Clay, A.T. (2013) Investigating the reactivities of a polyketide synthase module through fluorescent click chemistry. *Submitted*.

Kao, C.M., McPherson, M., McDaniel, R.N., Fu, H., Cane, D.E., and Khosla, C. (1998) Alcohol stereochemistry in polyketide backbones is controlled by the β-ketoreductase domains of modular polyketide synthases. J. Am. Chem. Soc. *120*, 2478-2479.

Kataoka, M., Kita, K., Wada, M., Yasohara, Y., Hasegawa, J., and Shimizu, S. (2003) Novel bioreduction system for the production of chiral alcohols. Appl. Microbiol. Biotechnol. *62*, 437–445.

Katz, L. (2009) The DEBS paradigm for type I modular polyketide synthases and beyond. Methods Enzymol. *459*, 113-142.

Keatinge-Clay, A.T. and Stroud, R.M. (2006) The structure of a ketoreductase determines the organization of the b-carbon processing enzymes of modular polyketide synthases. Structure *14*, 737–748.

Keatinge-Clay, A.T. (2007) A tylosin ketoreductase reveals how chirality is determined in polyketides. Chem. Biol. *14*, 898–908.

Keatinge-Clay, A.T. (2008) Crystal structure of the erythromycin polyketide synthase dehydratase. J. Mol. Biol. *384*, 941–953.

Keatinge-Clay, A.T. (2012) The structures of type I polyketide synthases. Nat. Prod. Rep. *29*, 1050-1073.

Kellenberger, L., Galloway, I.S., Sauter, G., Bohm, G., Hanefeld, U., Cortés, J., Staunton, J., and Leadlay, P.F. (2008) A polylinker approach to reductive loop swaps in modular polyketide synthases. ChemBioChem. *9*, 2740-2749.

Khosla, C., Tang, Y., Chen, A.Y., Schnarr, N.A., and Cane, D.E. (2007) Structure and mechanism of the 6-deoxyerythronolide B synthase. Annu. Rev. Biochem. *76*, 195–221.

Kodumal, S.J., Patel, K.G., Reid, R., Menzella, H.G., Welch, M., and Santi, D.V. (2004) Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. Proc. Natl. Acad. Sci. USA *101*, 15573–15578.

Kumar, P., Li, Q., Cane, D.E., and Khosla, C. (2003) Intermodular communication in modular polyketide synthases: Structural and mutational analysis of linker mediated protein-protein recognition. J. Am. Chem. Soc. *125*, 4097-4102.

Kushnir, S., Sundermann, U., Yahiaoui, S., Brockmeyer, A., Janning, P., and Schulz, F. (2012) Minimally invasive mutagenesis gives rise to a biosynthetic polyketide library. Angew. Chem. Int. Ed. *51*, 10664-10669.

Kwan, D.H., Sun, Y., Schulz, F., Hong, H., Popovic, B., Sim-Stark, J.C.C., Haydock, S.F., and Leadlay, P.F. (2008) Prediction and manipulation of the stereochemistry of enoylreduction in modular polyketide synthases. Chem. Biol. *15*, 1231–1240.

Kwan, D.H. and Schulz, F. (2011) The stereochemistry of complex polyketide biosynthesis by modular polyketide synthases. Molecules *16*, 6092-6115.

Kwan, D.H., Tosin, M., Schläger, N., Schulz, F., and Leadlay, P.F. (2011) Insights into the stereospecificity of ketoreduction in a modular polyketide synthase. Org. Biomol. Chem. *9*, 2053-2056.

Magarvey, N.A., Ehling-Schulz, M., and Walsh, C.T. (2006) Characterization of the cereulide NRPS α-hydroxy acid specifying modules: Activation of α-keto acids and chiral reduction on the assembly line. J. Am. Chem. Soc. *128*, 10698-10699.

Marsden, A.F., Caffrey, P., Aparicio, J.F., Loughran, M.S., Staunton, J., and Leadlay, P.F. (1994) Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. Science *263*, 378-380.

Masamune, S., Hirama, M., Mori, S., Ali, S.A., and Garvey, D.S. (1981) Total synthesis of 6-deoxyerythronolide B. J. Am. Chem. Soc. *103*, 1568–1571.

Moldenhauer, J., Chen, X.H., Borriss, R., and Piel, J. (2007) Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the *trans*-AT family. Angew. Chem. *46*, 8195-8197.

Murphy, C.D., Schaffrath, C., and O'Hagan, D. (2003) Fluorinated natural products: the biosynthesis of fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*. Chemosphere *52*, 455-461.

Murphy, C.D. (2007) The application of ^{19}F nuclear magnetic resonance to investigate microbial biotransformations of organofluorine compounds. OMICS *11*, 314-324.

Myles, D.C., Danishefsky, S.J., and Schulte, G. (1990) Development of a fully synthetic stereoselective route to 6-deoxyerythronolide B by reiterative applications of the Lewis

acid catalyzed diene aldehyde cyclocondensation reaction: a remarkable instance of diastereofacial selectivity. J. Org. Chem. *55*, 1636–1648.

O'Hare, H.M., Baerga-Ortiz, A., Popovic, B., Spencer, J.B., and Leadlay, P.F. (2006) High-throughput mutagenesis to evaluate models of stereochemical control in ketoreductase domains from the erythromycin polyketide synthase. Chem. Biol. *13*, 287-296.

Oliynyk, M., Brown, M.J.B., Cortés, J., Staunton, J., and Leadlay, P.F. (1996) A hybrid modular polyketide synthase obtained by domain swapping. Chem. Biol. *3*, 833-839.

Patel, P.S., Huang, S., Fisher, S., Pirnik, D., Aklonis, C., Dean, L., Meyers, E., Fernandes, P., and Mayerl, F. (1995) Bacillaene, a novel inhibitor of procaryotic protein synthesis produced by *Bacillus subtilis*: production, taxonomy, isolation, physico-chemical characterization and biological activity. J. Antibiot. *48*, 997-1003.

Patel, R.N. (2008) Synthesis of chiral pharmaceutical intermediates by biocatalysis. Coord. Chem. Rev. *252*, 659–701.

Piel, J. (2010) Biosynthesis of polyketides by *trans*-AT polyketide synthases. Nat. Prod. Rep. *27*, 996-1047.

Piasecki, S.K., Taylor, C.A., Detelich, J.F., Liu, J., Zheng, J., Komsoukaniants, A., Siegel, D.R., and Keatinge-Clay, A.T. (2011) Employing Modular Polyketide Synthase

Ketoreductases as Biocatalysts in the Preparative Chemoenzymatic Syntheses of Diketide Chiral Building Blocks. Chem. Biol. *18*, 1331-1340.

Raimundo, B.C. and Heathcock, C.H. (1995) Further studies on the *anti*-selective aldol reaction of chiral imides. Synlett *12*, 1213–1214.

Reddick, J.J., Antolak, S.S., and Raner, G.M. (2007) PksS from *Bacillus subtilis* is a cytochrome P450 involved in bacillaene metabolism. Biochem. Biophys. Res. Commun. *358*, 363-367.

Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D.V., Hutchinson, C.R., and McDaniel, R. (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. Biochemistry *42*, 72-79.

Serber, Z., Keatinge-Clay, A.T., Ledwidge, R., Kelly, A.E., Miller, S.M., and Dötsch, V. (2001) High-resolution macromolecular NMR spectroscopy inside living cells. J. Am. Chem. Soc. *123*, 2446-2447.

Sharma, K.K. and Boddy, C.N. (2007) The thioesterase domain from the pimaricin and erythromycin biosynthetic pathways can catalyze hydrolysis of simple thioester substrates. Bioorg. Med. Chem. Lett. *17*, 3034–3037.

Sherman, D.H. and Smith, J.L. (2006) Clearing the skies over modular polyketide synthases. ACS Chem. Biol. *1*, 505–509.

Siskos, A.P., Baerga-Ortiz, A., Bali, S., Stein, V., Mamdani, H., Spiteller, D., Popovic, B., Spencer, J.B., Staunton, J., Weissman, K.J., and Leadlay, P.F. (2005) Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. Chem. Biol. *12*, 1145–1153.

Smith, S. and Tsai, S.C. (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases. Nat. Prod. Rep. *24*, 1041–1072.

Stang, E.M. and White, M.C. (2009) Total synthesis and study of 6-deoxyerythronolide B by late-stage C-H oxidation. Nat. Chem. *1*, 547–551.

Staunton, J. and Weissman, K.J. (2001) Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.

Stinear, T.P., Mve-Oblang, A., Small, P.L., Frigui, W., Pryor, M.J., Brosch, R., Jenkin, G.A., Johnson, P.D., Davies, J.K., Lee, R.E., et al. (2004) Giant plasmid-encoded polyketide synthases produce the macrolide toxin of Mycobacterium ulcerans. Proc. Natl. Acad. Sci. USA *101*, 1345–1349.

Stevens, A.N., Morris, P.G., Iles, R.A., Sheldon, P.W., and Griffiths, J.R. (1984) 5-fluorouracil metabolism monitored in vivo by ^{19}F NMR. Br. J. Cancer *50*, 113-117.

Straight, P.D., Fischback, M.A., Walsh, C.T., Rudner, D.Z., and Kolter, R. (2007) A singular enzymatic megacomplex from *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *104*, 305-310.

Strohl, W.R. and Connors, N.C. (1992) Significance of anthraquinone formation resulting from the cloning of actinorhodin genes in heterologous streptomycetes. Mol. Microbiol. *6*, 147-152.

Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., and Schulz, F. (2013) Enzyme-directed mutasynthesis: A combined experimental and theoretical approach to substrate recognition of a polyketide synthase. ACS Chem. Biol. *8*, 443-450.

Tokuyama, H., Yokoshima, S., Yamashita, T., Lin, S., Li, L., and Fukuyama, T. (1998) Facile palladium-mediated conversion of ethanethiol esters to aldehydes and ketones. J. Braz. Chem. Soc. *9*, 381–387.

Tsuji, S.Y., Cane, D.E., and Khosla, C. (2001) Selective protein-protein interactions direct channeling of intermediates between polyketide synthase modules. Biochemistry *40*, 2326-2331.

Valenzano, C.R., Lawson, R.J., Chen, A.Y., Khosla, C., and Cane, D.E. (2009) The biochemical basis for stereochemical control in polyketide biosynthesis. J. Am. Chem. Soc. *131*, 18501–18511.

Valenzano, C.R., You, Y.O., Garg, A., Keatinge-Clay, A., Khosla, C., and Cane, D.E. (2010) Stereospecificity of the dehydratase domain of the erythromycin polyketide synthase. J. Am. Chem. Soc. *132*, 14697–14699.

Ward, S.L., Desai, R.P., Hu, Z.H., Gramajo, H., and Katz, L. (2007) Precursor-directed biosynthesis of 6-deoxyerythronolide B analogues is improved by removal of the initial catalytic sites of the polyketide synthase. J. Ind. Microbiol. Biotechnol. *34*, 9-15.

Wong, C., Drueckhammer, D., and Sweers, H.M. (1985) Enzymatic vs. fermentative synthesis: thermostable glucose dehydrogenase catalyzed regeneration of NAD(P)H for use in enzymatic synthesis. J. Am. Chem. Soc. *107*, 4028–4031.

Wong, F.T., Jin, X., Matthews, II, Cane, D.E., and Khosla, C. (2011) Structure and mechanism of the *trans*-acting acyltransferase from the disorazole synthase. Biochemistry *50*, 6539-6548.

Wu, J., Zaleski, T.J., Valenzano, C., Khosla, C., and Cane, D.E. (2005) Polyketide double bond biosynthesis. Mechanistic analysis of the dehydratase-containing module 2 of the picromycin/methymycin polyketide synthase. J. Am. Chem. Soc. *127*, 17393-17404.

Zheng, J., Taylor, C.A., Piasecki, S.K., and Keatinge-Clay, A.T. (2010) Structural and functional analysis of A-type ketoreductases from the amphotericin modular polyketide synthase. Structure *18*, 913–922.

Zheng, J. and Keatinge-Clay, A.T. (2011) Structural and functional analysis of C2-type ketoreductases from modular polyketide synthases. J. Mol. Biol. *410*, 105-117.

Zheng, J., Gay, D.C., Demeler, B., White, M.A. and Keatinge-Clay, A.T. (2012) Divergence of multimodular polyketide synthases revealed by a didomain structure. Nat. Chem. Biol. *8*, 615-621.

Zheng, J., Piasecki, S.K., and Keatinge-Clay, A.T. (2013) Structural studies of an A2-type modular polyketide synthase ketoreductase reveal features controlling α-substituent stereochemistry. ACS Chem. Biol. *In press*.

Zhou, H., Gao, Z., Qiao, K., Wang, J., Vederas, J.C., and Tang, Y. (2012) A fungal ketoreductase domain that displays substrate-dependent stereospecificity. Nat. Chem. Biol. *8*, 331-333.

Vita

Shawn Kristen Piasecki was born in NJ then resided in south Austin, attending Montessori school before moving to Dripping Springs, TX for middle and high school. Shawn began her undergraduate studies at UT Austin in 2003, which resulted in majoring in chemistry and secondary education certification from the UTeach program. In order to combine teaching and science, she began working for the Freshman Research Initiative, aiding Dr. Bradley Hall in teaching students how to perform independent research. She then began research in Dr. Andrew Ellington's lab under the supervision of Dr. Bradley Hall prior to his graduation. In 2007, Shawn graduated with her B.S. in Chemistry and worked in the Ellington lab as a technician until the start of graduate school at UT Austin in August 2008. She has since worked under the supervision of Dr. Adrian T. Keatinge-Clay, concentrating on understanding the natural product potential of polyketide synthase ketoreductases. She plans to use her expertise in natural product biosynthesis and analysis in her new career.

Contact: shawn.piasecki@utexas.edu

This dissertation was typed by the author.