Directed Evolution of a Modular Polyketide Synthase Thioesterase for Generation of a Hybrid Macrocyclic Ring System

Maria L. Adrover-Castellano[a], Jennifer J. Schmidt[a], Carolyn A. Glasser[a], Fengrui Qu[b], and David H. Sherman[*a, c]

a. Life Sciences Institute, University of Michigan, 210 Washtenaw Avenue, Ann Arbor, MI 48109-2216 (USA)
b. Department of Chemistry, University of Michigan, Ann Arbor, MI 48109 (USA)
c. Life Sciences Institute, Departments of Medicinal Chemistry, Chemistry, and Microbiology & Immunology, University of Michigan (USA)

Abstract: Modular type I polyketide synthases (PKSs) comprise a family of enzymes that synthesize a diverse class of natural products with medicinal applications. The biochemical features of these systems include the extension and processing of polyketide chains in a stepwise, stereospecific manner, organized by a series of modules divided into distinct catalytic domains. Previous work revealed that a primary hurdle for utilizing PKS modules to create diverse macrolactones hinges on the selectivity of the thioesterase (TE) domain. Herein, we generated a novel hybrid 12-membered macrolactone/lactam ring system employing an unnatural amide hexaketide intermediate in conjunction with an engineered TE S148C mutant from the pikromycin (Pik) biosynthetic pathway. This unnatural macrocycle was initially formed in severely attenuated yields compared to the native product generated from the natural hexaketide substrate. A step-wise directed evolution campaign generated Pik TE variants with enhanced selectivity for macrocycle formation over hydrolysis. Over three rounds of evolution, a series of mutant Pik TE proteins were identified, and further combinations of beneficial mutations carried from each round produced a composite variant with six-fold enhanced isolated yield of the hybrid macrocycle compared to the parent TE enzyme. This study offers new insights into the range of amino acid residues, both proximal and distal to the active site that impart improved selectivity and yield against the unnatural polyketide substrate and overcoming a key PKS pathway gatekeeper.

Introduction

Macrocyclic scaffolds are found in many natural products from diverse sources. These commonly include twelve or more atoms in their cyclic framework.[1] Synthetic strategies have been developed for macrolactonization, and macrolactamization, which include transition metal catalyzed coupling reactions, click chemistry and ring-closing metathesis.[1, 2] A wide variety of these molecules have been developed as pharmaceuticals with over one hundred FDA approved drugs carrying a macrocyclic core (e.g. macrolide and cyclic peptide antibiotics, anthelmintics, anticancer agents, immunomodulators); showing their current medicinal value and continued promise for future drug discovery and development.[3, 4] Macrolide antibiotics, including erythromycin and its semi-synthetic derivatives (clarithromycin, azithromycin, among others) act against bacterial pathogens as ribosome inhibitors and have synergistic immunomodulatory activity in the human host.[5, 6] The macrolide chemical structure contains a macrolactone ring possessing one or more sugar moieties (amino sugars, deoxy sugars or both) and other functionalities such as hydroxyl groups important for their mechanism of action in the bacterial ribosome peptidyl transferase center.[7] Frequently, the structural complexity of large-ring macrolactones introduces significant synthetic challenges. Options for solving these include total synthesis, semi-synthesis and chemoenzymatic strategies.[8-17] For example, the first total synthesis of erythromycin[18-20] in 1981 reported over 50 synthetic steps and an overall yield of less than 1%. These large macrocycles exhibit stereochemical and structural diversity, which often results in laborious, low yielding reaction schemes, possessing a barrier for further diversification or pursuing structure-activity relationship studies. Alternatively, semi-synthesis has been employed to generate widely used antibiotics such as the erythromycin derivatives clarithromycin, azithromycin, telithromycin, and others.[5, 9, 12, 21] Despite these approaches to gain synthetic efficiencies and improve overall yields, modifications to the starting material scaffolds are often limited by reactive functional groups and selectivity challenges at several positions within these intricate ring structures.[2, 14, 21-25]

Chemoenzymatic synthesis encompasses assembly of intermediates from simple building blocks, combined with enzyme catalyzed reactions to obtain efficient production of novel molecules and pharmaceuticals using aqueous reaction conditions, while avoiding toxic reagents, solvents and protecting groups.[26-31] Although workable synthetic methodologies exist for macrolactonization, these are often met with unfavorable entropic and enthalpic factors relating to the energetics of intramolecular cyclization.[15-17, 32] To address the challenges of total and semi-synthesis, we have been investigating the pikromycin (Pik) biosynthetic pathway thioesterase (TE) domain as a key biocatalyst for macrolide antibiotic discovery.[26, 33] Earlier studies revealed how the Pik TE binds its linear substrate and prepares it for cyclization reactions. Moreover, it can accommodate hexa- and heptaketides in its active site through a hydrophobic chamber that include substrate-protein anchoring hydrogen bonds. Remarkably, the TE induces a curled conformation of the linear polyketide through the hydrophobicity of the surrounding residues, coupled with a hydrophilic barrier at the exit site of the enzyme channel. This barrier forces the substrate back into the active site, and positions the nucleophilic hydroxyl group for macrocyclization.[34, 35]

The Pik pathway is comprised of a modular type I PKS that selectively catalyzes key transformations on structurally distinct intermediates generating a 12-membered macrolactone ring, 10-deoxymethynolide (10-DML) and a 14-membered macrolactone ring, narbonolide (NBL). These megasynthases are comprised of modules divided into distinct catalytic domains with specific roles.[33, 36] Each round of polyketide elongation is performed by three main domains, the acyltransferase (AT) that selects an extender unit, an acyl carrier protein (ACP) and a ketosynthase (KS) that accepts the polyketide chain from an upstream ACP domain and catalyzes a decarboxylative Claisen condensation reaction. Modules may contain combinations of other domains including ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) to transform the β-keto group into hydroxyl, alkene or alkane functionalities, respectively.[26, 37, 38] At the end of the pathway, a terminal module, typically a TE domain, catalyzes formation and off-loading of the hydrolyzed or cyclized product.[38] Developing these enzymes as biocatalysts represents a complementary addition to synthetic chemistry for

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

achieving molecular diversification.[39-41] Thus, enhancing our understanding and improving Pik TE catalysis for macrocyclization represents a priority area to explore scalable applications of PKS systems.

Previous studies on PKS TEs from diverse organisms and biosynthetic pathways, as well as engineering efforts to understand hydrolysis and cyclization processes have been reported.[38, 42-62] In a broader context, classifications of TE enzyme families based on structure, function, substrate specificity, products generated, and catalytic mechanisms enable their use as biocatalysts for the synthesis of diverse molecules.[42, 44, 45, 48, 49, 63] Additionally, various TE evolutionary models for enzyme reactivity and selectivity have provided clues for altering TE catalysis to form a desired product.[42, 48] In some cases, TEs can process a wide variety of substrates, while in others, they show a high degree of selectivity for specific structural features, depending on the TE loading or release step considered during the overall off-loading mechanism.[42, 48] Thus, we envision that understanding TE function through the investigation of diverse unnatural substrates can shed light on the biocatalytic parameters required for generating a wider range of products with high selectivity and efficiency.

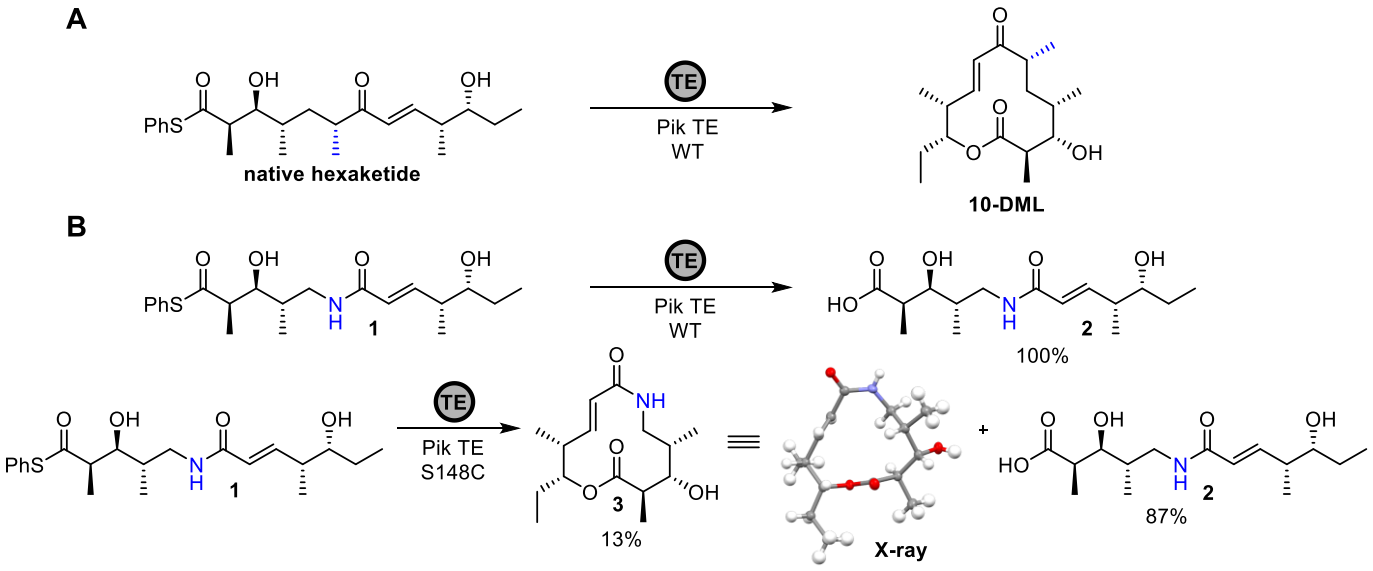
Previously, we demonstrated the ability to form 12-membered macrolactones utilizing unnatural pentaketides in conjunction with the PikAIII-TE fusion protein.[39] Depending on substrate structure, the cyclized products were obtained in 9%-66% isolated yields, with competing hexaketide hydrolysis and truncated byproducts formed when using PikAIII-TE WT[39] and the PikAIII-TE S148C mutant[40]. The results indicated that Pik TE has limited substrate flexibility and often functions as a gatekeeper in the processing of unnatural substrates.[39-41] Although diverse unnatural macrolactones of different ring sizes have been generated, the introduction of functional groups not typically found in polyketide chain elongation intermediates often results in failed or inefficient cyclization. Thus, we were motivated to deliberately engineer and expand the ability of Pik TE to process structurally variant polyketide substrates in the pursuit of understanding substrate-TE interactions. Protein engineering has been used widely to obtain enzymes with new and finely tuned properties,[64-67] but has not been applied to polyketide macrocyclization catalysts. In this work, we employed directed evolution utilizing an amide hexaketide (**1**) to create and assess a library of Pik TE variants leading to a decrease in hydrolysis product (**2**) and efficient formation of a hybrid macrolactone/lactam derivative (**3**) of 10-DML (Scheme 1).[33, 68-70] We identified new variants with enhanced selectivity for macrocyclization, which was demonstrated by increased total turnover numbers (TTNs), initial reaction rates and isolated yields.

Results

Chemical Synthesis of Amide Hexaketide Substrate

Our earlier studies with the native hexaketide demonstrated an effective chemoenzymatic strategy for the generation of 10-DML when utilizing Pik TE WT (Scheme 1A).[70] We then focused on simplifying the native hexaketide and altering the stereochemistry of the nucleophilic hydroxyl group.[39, 40] In choosing a new hexaketide target, we were motivated by amide-containing derivatives of erythromycin that show antibiotic activity comparable to azithromycin.[71-75] Thus, the impact of exchanging a central native C-Me functionality to N-H, creating an amide group in the chain-elongation intermediate (Scheme 1B) became a central objective. This choice was also made to assess the impact of a single heteroatom replacement in the chain on TE-mediated cyclization. Moreover amide functionality is present in over half of FDA approved pharmaceuticals, and editing the macrocycle to include a new H-bond donor was also a compelling objective for future bioactivity studies.[76-81] Robust methods are available for amide synthesis, including dehydrative condensations between carboxyl and amino groups, utilization of condensation reagents, Beckmann rearrangements, and Schmidt reaction mechanisms.[82, 83] Thus, to further pursue diversification of polyketide systems, we designed an amide-containing Pik hexaketide analog to explore TE-mediated cyclization toward a hybrid ring system (Scheme 2).

The target amide hexaketide (**1**) was synthesized starting with Evan's aldol condensation[84] between (*R*)-4-benzyl-3-propionyloxazolidin-2-one (**4**)[85, 86] and aldehyde **5** (generated from (*S*)-3-(Boc-amino)-2-methylpropionic acid) employing dibutylboron triflate to give **6**. Next, lithium hydroxide cleavage



Scheme 1 (A) Native hexaketide substrate (derived from an NBOM hydroxyl protected intermediate)[70] against Pik TE WT generates 10-DML. (B) Amide hexaketide (**1**) against Pik TE WT generates hydrolysis product (**2**) whereas Pik TE S148C catalyzes formation of the hybrid macrolactone/lactam (**3**) with major hydrolysis product (**2**).

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

of the Evan's auxiliary[11, 87] afforded **7**, followed by trimethylsilyl (TMS) diazomethane esterification to make **8**. Subsequent deprotection of *N*-Boc enabled coupling with **9** to yield amide **10**. Hydrolysis with lithium hydroxide (**11**) followed by thioesterification and final cleavage of the silyl ether[11] with hydrofluoric acid (HF) generated the amide hexaketide (**1**) (Scheme 2).

Isolation of Novel 12-membered Macrolactone/Lactam Ring

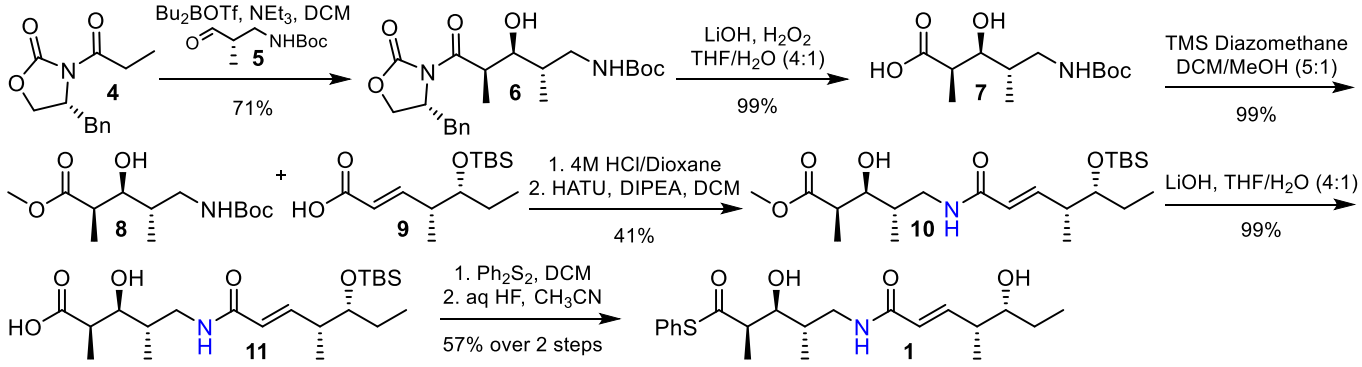
With the desired substrate **1** in hand, we tested it against Pik TE WT and the Pik TE S148C active site mutant.[40] Pilot scale reactions indicated that Pik TE WT catalyzed exclusive formation of the amide hydrolysis product **2** [M+H] = 302 m/z (Scheme 1B). Although reaction with Pik TE S148C also resulted in majority formation of the seco-acid hydrolysis product **2**, it was accompanied by the mass expected for the hybrid 12-membered macrolactone/lactam **3** [M+H] = 284 m/z (Scheme 1B), suggesting potential formation of the desired macrocycle (Fig. S1). Subsequent scale-up, high-performance liquid chromatography (HPLC) purification and isolation provided a 13% isolated yield of the macrocyclic compound **3** confirmed by NMR (Table S1) and X-ray crystallography (Scheme 1B), which verified the molecular architecture and stereochemistry of the product (Fig. S24, CCDC Deposition Number 2294580). The corresponding hydrolysis product **2** was isolated as a mixture of spontaneous degradation products[39] that were not further characterized. Thus, the Pik TE S148C active site mutant was able to catalyze formation of the desired 12-membered ring **3**, albeit in poor overall yield. As a next step, we decided to pursue directed evolution to create additional Pik TE S148C variants with enhanced selectivity for the desired macrocycle **3**.

Pik TE Directed Evolution Strategy

Our objective for reactions using Pik TE S148C with the amide hexaketide substrate **1** was to obtain incremental improvements in macrocycle formation with a corresponding decrease in hydrolysis products screened via liquid chromatography-mass spectrometry (LC-MS). To guide the first round of evolution, we examined the Pik TE WT X-ray crystal structure[34] with a covalent pentaketide affinity label (PDB: 2HFJ), and twenty-one resides were identified within 6 Å of the substrate in the active site region (Fig. 1). The crystal structure-substrate mimic complex likely represents a close approximation to the native hexaketide substrate covalently bound to Ser148 in the TE active site showing the Ser148, Asp176, His268 catalytic triad, and revealed the residues likely involved in binding and cyclization to generate macrolactone **3**. Next, site saturation mutagenesis was employed to create thousands of Pik TE (S148C) active site region mutants by changing each selected residue one by one to all twenty canonical amino acids.[34, 88, 89] Screening 2,016 protein isolates, we identified five single amino acid Pik TEs S148C variants (L29M, R160C, A217T, A217Y, G222V) with improved macrocycle yields. These variants were further confirmed by protein production, purification and enzymatic reactions. The Pik TE S148C with additional single mutations L29M, A217T and A217Y were identified at the expected 6 Å distance from the active site region. However, the corresponding G222V and R160C mutations were located 12 Å and 20 Å, respectively, distal to the active site as unexpected polymerase-generated polymerase chain reaction (PCR) non-silent mutations. Surprisingly, these amino acid changes resulted in enhanced macrolactone ring formation. Next, we explored combinations of beneficial mutations from round 1 (R1), and found that the Pik TE S148C, A217T, G222V variant showed 3-fold enhanced activity (39% macrocycle yield based on LC-MS) compared to the original Pik TE S148C enzyme (Fig. 2A). This triple-mutant form of Pik TE was generated by combining S148C, with mutation A217T within 6 Å and the unexpected non-silent mutation G222V located 12 Å from the active site region.

Next, we conducted a second round of evolution focused on improvement of the optimal first round combined variant Pik TE S148C, A217T, G222V. Based on the beneficial G222V we were motivated to pursue random mutagenesis in regions more distal to the active site. To this end, we employed an error-prone polymerase chain reaction (epPCR) strategy to create random mutations across the *pikAIV* TE gene at a higher mutation rate.[89, 90] In round 2 (R2), we screened 2,880 protein isolates and identified sixteen variants with improved production of hybrid macrolactone/lactam **3**. Assessing the individual mutations revealed that the distances of mutated amino acid residues relative to the active site region varied from 6 - 20 Å. This outcome demonstrated that critical residues impacting Pik TE selectivity and turnover can reside at relatively remote distances from the active site. The identified variants with improved selectivity included single, double, and triple mutations added to the original Pik TE R1 variant. These were selected for protein production, purification, and confirmation of enhanced biocatalytic activity, similar to our approach during R1. For R2, we generated the combination TE variant comprised of six mutations (Pik TE S148C, A217T, G222V, M271V, Y25C, L126V) with a 5.5-fold improvement (71% macrocycle yield based on LC-MS) (Fig. 2B).

We subsequently reiterated epPCR random mutagenesis utilizing the R2 variant for a third round of evolution, following the same parameters as the second round. After screening 1,536 protein isolates, six single amino acid variants (all ~20 Å from active site region) were identified with potentially improved ring formation. However, further scaleup revealed similar yields (71% macrocycle yield based on LC-MS) compared to R2. Although these hits and two combination mutants were generated (including the combination mutant S148C, A217T, G222V, M271V, Y25C, L126V, Q18K, L87Q, A9G, A120V,



Scheme 2 Chemical synthesis of amide hexaketide (**1**).

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

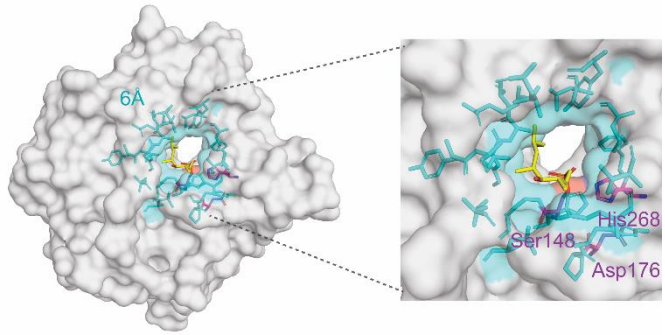


Fig. 1 Pik TE WT (Ser148, Asp176, His268) with covalent pentaketide affinity label (PDB: 2HFJ) (yellow). Twenty-one residues identified within 6 Å of the substrate in the active site region (cyan). Active site catalytic triad (purple).

E294G, G138W from round 3 (R3) containing 12 mutations by combining R2 and the six hits or mutations obtained in R3), none resulted in a yield increase and in some cases, a slight reduction on macrocyclization was observed (Fig. 2A).

To demonstrate productivity of the engineered Pik TE variants at a larger scale, we conducted reactions utilizing ~10 mg of the amide hexaketide substrate **1**. HPLC purification and isolation of the desired cyclic product **3** afforded 42% (Pik TE R1), 79% (Pik TE R2) and 69% (Pik TE R3) isolated yields (Fig. 2C).

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

Additional analysis of the Pik TE S148C, R1, R2 and R3 mutants, including TTNs, thermal shift assay melting temperatures (Tm's), and initial rates for the reactions, revealed individual variant's stability and catalytic performance.[91-93] TTN enzyme measurements enabled determination of the number of catalytic events performed by one biocatalyst active site during its lifespan or until its total decay.[91, 94, 95] Our results demonstrated an increase in TTN for cyclized product compared to Pik TE S148C (Fig. 3A). Notably, for the TTNs over the 18-hour reaction period and 0.2 mol % reduced biocatalyst loading, both Pik TE R2 and R3 retained the highest amount of starting material among the four enzymes, suggesting potential protein deactivation[91, 94, 95] (Fig. S8). Interestingly, Pik TE R1 and R3 exhibited highly similar TTN values, yet Pik TE R1 showed the lowest remaining substrate compared to R3. We then performed a thermal shift assay using SYPRO orange dye to determine protein decay and stability by thermal denaturation of the mutants.[92, 96] The results indicated that both Pik TE WT and the S148C mutant had highly similar Tm values, whereas Pik TE R1 showed a Tm increase of 0.6 ºC. Pik TE R2 and R3 showed a decrease in Tm by 5.0 ºC and 3.8 ºC, respectively, suggesting enzyme denaturation occurs at lower temperatures, which reflects lower protein stability compared to Pik TE WT, S148C and R1 (Fig. 3A). These data help explain the results obtained for TTN values, since R1 as the more stable protein had the highest TTN value due to greater stability during the 18-hour TTN reaction time and 0.2 mol % reduced enzyme loading. Pik TE R2 and R3 TTN reactions had considerable amounts of starting material remaining, consistent with the deactivation of the mutant proteins under the TTN experiment conditions (Fig. S8).

Additionally, a time course analysis was conducted with all four Pik TE mutants, illustrating the differences in selectivity for generating the hybrid macrolactone/lactam **3** and hydrolysis products over three rounds of directed evolution (Fig. 3B). Initial rates for substrate consumption were determined for all Pik TEs, revealing similar values across the mutants, suggesting that higher concentrations of protein (1.0 mol %) and substrate have an impact on the reaction rate compared to TTN values. Likewise, initial rates of ring formation were determined, with increasing numbers for the evolved variants (Fig. 3A). Compared to the initial rate for Pik TE S148C, Pik TE R1 has a 3.5-fold improvement, followed by Pik TE R2 and Pik TE R3 with a ~14-fold enhancement on initial rates for cyclization. These results show the increase in selectivity of our evolved Pik TE variants for macrocyclization (Fig. 3B). The data also indicates that despite the TTN results and thermal shift assay showing Pik TE R2 and R3 variants becoming more unstable over time, addition of 1.0 mol % of the enzyme hastens the reaction sufficiently to generate high quantities of the desired macrocyclic product without compromising the % yield.

Pik TE Variants Cysteine 148 to Serine Reversions

Finally, to query the importance of the Pik TE S148C active site mutation in the Pik TE R1 variant, it was reverted to the WT catalytic triad bearing Ser148. The enzymatic reaction with the amide hexaketide **1** and Pik TE R1 C148S$_{reversion}$ showed a >10-fold decrease in ring formation (39% to 3% yield) by LC-MS analysis. This result confirms that the S148C mutation is a critical starting point for improved activity in the Pik TE R1 biocatalyst. Moreover, combining the S148C, A217T and G222V variants within the Pik TE R1 revealed synergistic effects of the combined individual mutations. Similarly, we reverted the Pik TE R2 variant back to the C148S WT residue, generating Pik TE R2

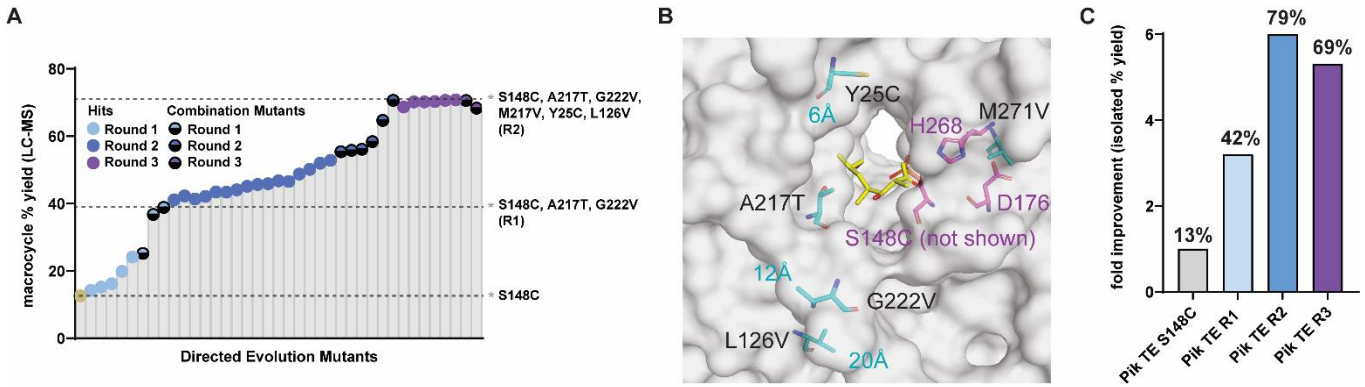


Fig. 2 (A) Pik TE variants from directed evolution confirmed by scale-up protein production, purification and enzymatic reactions analyzed in LC-MS. (B) Pik TE WT crystal structure (PDB: 2HFJ) with mutations in cyan depicting Pik TE R2 using PyMOL mutagenesis tool. Pentaketide mimic affinity label in yellow and the TE catalytic triad in purple. (C) Fold improvement (% isolated yield) in formation of macrocycle **3** from substrate **1** using parental Pik TE S148C and optimized mutants.

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

A

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

B

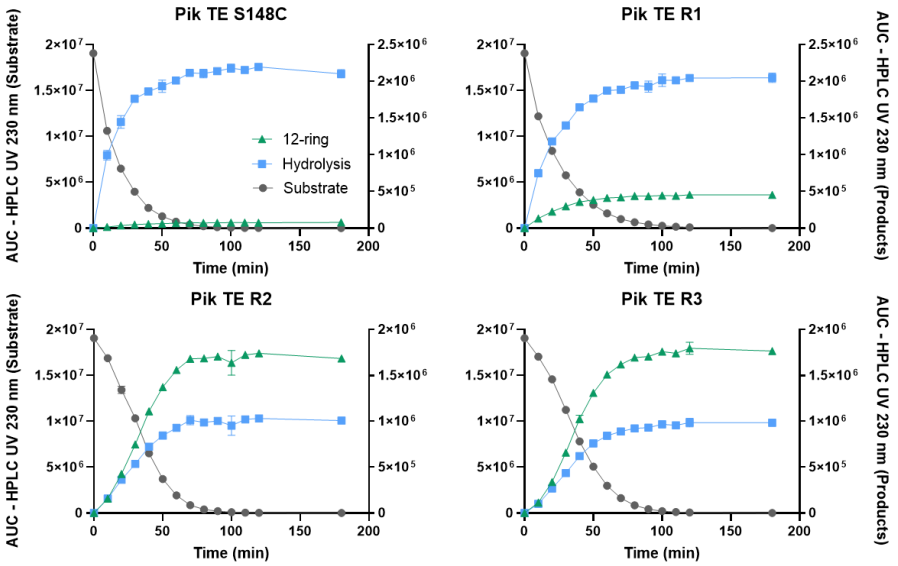


Fig. 3 (A) Analysis for total turnover numbers (TTNs), melting temperatures (Tm, °C), initial rates (μM/min) substrate consumption and macrolactone formation for select Pik TE variants. (B) Reaction time courses: substrate and product HPLC peaks shown as area under the curve (AUC) as a function of time (min) (see Supplemental Information (SI) data for additional details on AUC versus μM concentrations).

C148S$_{reversion}$. Testing amide hexaketide **1** against the Pik TE R2 reversion mutant resulted in a decrease from 71% to a 14% yield of the macrocycle (Fig. 4A). Despite the attenuated activity of the reversion mutants, which confirmed Pik TE S148C as foundational to success of the current strategy, these results demonstrate gain-of-function for the Pik TE WT (Ser148) and could be a potential future avenue to independently generate the hybrid macrocyclic product **3** through further directed evolution (Fig. 4B).

A

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

B

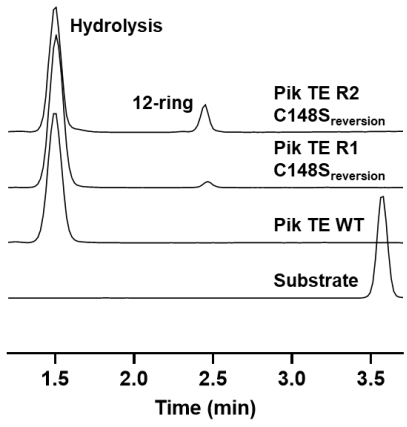


Fig. 4 (A) LC-MS % yields and (B) traces for Pik TE WT, Pik TE R1 C148S$_{reversion}$ and Pik TE R2 C148S$_{reversion}$ highlighting a gain of function for the WT enzyme.

Discussion

Macrolactonization represents a challenging transformation in organic synthesis, including unfavorable energetics of intramolecular reactions to form rings with more than six atoms.[15-17] Although the generation of macrolactones and macrolides by total and semi-synthesis is feasible, limitations to both approaches including the large number of synthetic steps, low yields, difficulties for selective addition of diverse functional groups and the inherent reactivity of large macrocycles represents a compelling area to explore with chemoenzymatic synthesis.[2, 4, 9, 12, 21-23] The discovery of modular type I PKSs that assemble complex natural products like erythromycin and pikromycin allowed us to identify pathway enzymes that catalyze macrolactonization.[33, 39, 40, 70, 97] In these biosynthetic systems, the PKS TE domain controls the formation of 12- and 14-membered ring structures.[38] In earlier studies, Pik TE was shown to be a catalytic bottleneck for the desired ring closing reaction and was identified as a gatekeeper for processing unnatural substrates, hence impeding structural diversification and production of new macrocycles.[39-41] This work demonstrates the value of TE directed evolution for macrocycle bioengineering to create and enhance the formation of new macrocyclic compounds.

Previously, the Pik TE S148C mutation revealed enhanced reaction kinetics and gain of function for the processing of an unnatural hexaketide diastereomer, which led to a shift in reaction mechanism between the WT and S148C variant as determined by quantum mechanical (QM) methods.[40] In our case, the S148C mutation is foundational for the generation of the hybrid 12-membered macrolactone/lactam ring **3**, whereas the Pik TE WT leads to exclusive formation of the seco acid amide hexaketide. With only two rounds of evolution, the Pik TE R2 provided a 6-fold improvement in isolated % yield compared to the Ser148Cys starting-point enzyme, enabling construction

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

of a novel cyclic ring system with high productivity and selectivity (Fig. 2C). The discrepancy of Pik TE R1 and R2 between macrocycle LC-MS % yields (Fig. 2A) and isolated % yields (Fig. 2C) could be attributed to molecule ionization dependence in LC-MS whereas isolated products are purified and quantified by weighted mass. For Pik TE R3, the variability in isolated % yield could reflect loss of product **3** during reaction workup or HPLC purification.

Our work benefitted from the Pik TE/substrate mimic crystal structure complex,[34] which guided us to identify residues within the active site region, where site saturation mutagenesis was employed in a first round of evolution (Fig. 1). Pik TE R1 contains the A217T mutation (6 Å from active site region) that was shown in the Pik TE structure with two well-ordered water molecules coordinated between side chains Gln183 and backbone carbonyl of A217, and the C5-hydroxyl group of the pentaketide substrate mimic. This region was predicted to form a 'hydrophilic barrier' to direct the acyl chain back into the TE substrate channel hydrophobic environment and guide the nucleophilic hydroxyl group toward the active site serine-ester bond.[34] We reasoned that the A217T could have an additional coordination site with the threonine side chain via hydrogen bonding interactions with the nucleophilic hydroxyl group. Although the G222V mutation is 12 Å from the active site region, increasing non-polar or hydrophobic effects of the more distal enzyme regions could induce conformational changes near the active site that promote cyclization and stability (Fig. 2B). Other studies have similarly shown that mutating residues relatively distant from the active site can improve enzymatic activity.[67, 98-100]

The epPCR random mutagenesis strategy enabled us to investigate residues from the entire Pik TE sequence. Over the course of this study, sixteen mutants were identified that improved macrocycle product yields compared to variants generated during the first round of directed evolution. Additional synergistic effects were observed when combining three of the mutations M271V, Y25C, L126V with Pik TE R1 (Fig. 2A). Residue Y25 was shown in the Pik TE crystal structure to engage in direct interactions with the 12-membered cyclized product 10-DML.[34] Therefore, we propose that Cys25 might promote polar contacts with the unnatural substrate that favor product formation. Mutation M271V is positioned behind the Asp176 and His268 catalytic triad residues 12 Å distal to the structure bound substrate in the active site region, whereas L126V is 20 Å away, and to the opposite side of the triad (Fig. 2B).[34] These valine mutations in Pik TE R2 could have hydrophobic interactions that create a more hospitable active site pocket and nearby regions of the catalytic site, allowing macrocyclization to occur more readily and excluding water that alleviates competing TE mediated hydrolysis.[101]

Regarding R3 mutations, all resided 20 Å distal to the active site region located at the surface of the TE and failed to provide higher production of **3**, thus revealing the evolutionary limits of our current screen. Our most highly optimized mutant Pik TE R2 (79% isolated yield), contains three valines, two cysteines and one threonine mutations added compared to the Pik TE WT (Fig. 2B). Enzyme performance limits with directed evolution have been reported previously.[102-104] Although predicting the potential benefits of additional mutations or rounds of evolution can be difficult, exploring existing mutations and understanding their impact on enzyme active site conformation, substrate binding, reaction mechanism, including the effects of individual mutations and combinations for overall protein stability could be informative for future studies.[105-107] In addition, complementary or parallel trajectories of individual variants can be pursued to maximally improve enzyme activities.[102, 103, 108, 109] Another option is ultra-high-throughput droplet-based microfluidic screening to identify further optimized enzymes.[110] TTNs were chosen as these experiments show the timescale of the enzyme activity over the course of its deactivation.[91] TTN values for the Pik TE variants were experimentally shown to be Pik TE R1 > Pik TE R3 > Pik TE R2 > Pik TE S148C (Fig. 3A). We hypothesized that this trend may be attributed to enzymatic deactivation with increased ring selectivity for R3, and resulting elevated TTN values comparable to R1. Thus, although Pik TE R1 has a lower selectivity for macrocyclization, it remains active for a longer period.[111, 112] Our observation of the trend observed between TTN and Tm values for the Pik TE variants could be rationalized from other studies showing enzymes possessing high thermal stability also exhibit superior TTN numbers.[91] Improved thermostability typically indicates that the protein retains activity over a longer period of time.[113, 114] A study where TTNs were estimated using the half-life of glucose dehydrogenase from *Bacillus subtilis* and its mutants showcased that in some cases the activity of an enzyme may be increased at the expense of another such as the half-life. One of the glucose dehydrogenase mutant forms was 19% less active than the WT and contained an 80% longer half-life, which resulted in a TTN 49% higher than the WT. Although this might not always be the case, longer half-lives often resulted in higher biocatalyst TTNs.[91] The trade-off observed between adding novel or improved enzyme functionalities and encountering stability issues has been observed frequently in protein engineering studies.[105, 107, 111, 115] Alternatives to improve the stability of the biocatalysts include rounds of evolution targeted to increasing the thermostability[113, 116-118] and protein immobilization strategies.[117] Our thermal shift assay analysis revealed that Pik TE % yield improvement for macrocyclization occurred at the cost of protein stability due to lower Tm values from R2 and R3 in comparison to Pik TE WT and the S148C mutant (Fig. 3A). Although some degree of enzyme deactivation was observed in the thermal shift assay (Fig. 3A), it did not significantly impact the enzymatic reactions at both analytical and scale-up levels. This was mainly due to employing higher enzyme concentrations (1.0 mol %) resulting in complete conversion of the substrate within one hour, overcoming possible stability challenges present in longer reaction times (Fig. 3B).

In comparison, the initial rates for substrate consumption by the Pik TE S148C, R1, R2, and R3 mutants showed similar μM/min values, indicating that an increase in protein concentration and substrate had an impact on the rate. Moreover, initial rates for compound **3** ring formation demonstrated an increase in enzymatic activity and shift in selectivity for the engineered TEs (Fig. 3A). In the reaction time courses, Pik TE R2 and R3 showed an initial increase in macrocycle and hydrolysis products at the onset of the reaction and then macrolactone formation surpassed seco-acid formation, which could relate to the enzyme mechanism parameters that influence hydrolysis or cyclization (Fig. 3B). In a previous study, we showed that Pik TE WT followed a stepwise addition-elimination mechanism with transient formation of a tetrahedral intermediate. By contrast, Pik TE S148C followed a concerted acyl substitution process, which is lower in energy.[40] The new mutations from this directed evolution campaign could have an impact on the biocatalyst energetic barriers in comparison to both Pik TE WT and S148C.[34, 35, 101, 119-125]

The Pik TE S148C catalytic triad mutation was also identified as fundamental in Pik TE R1 and Pik TE R2 variants as shown in reversion to the WT Ser148. Restoration of the Ser active site led to a >10-fold decrease in macrocycle formation (39% to 3%

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

LC-MS yield) in R1 and 5-fold decrease for R2 (71% to 14% LC-MS yield) (Fig. 4). Previously, the S148C mutation was shown to possess improved reaction kinetics for processing and accepting a hexaketide substrate with non-native hydroxyl group stereochemistry.[40]

Overall, the current study represents a proof-of-concept approach for relieving the PKS TE bottleneck in the formation of new macrocycles. The results demonstrate that directed evolution improved selectivity and productivity in the catalysis of an unnatural amide-containing hexaketide substrate **1** and offers significant promise for expanding to other substrates with diverse functionalities, aiming to generate novel macrolactones, macrolactams, depsipeptides, and other TE-mediated biosynthetic systems. Moreover, future computational analysis through QM, molecular dynamics (MD) calculations and machine learning (ML) approaches[126-131] could provide new mechanistic insights relating to mutations identified in our of biocatalyst variants. These approaches could also predict additional amino acids that modulate selectivity in TEs beyond the amide hexaketide substrate **1**.

Conclusions

In summary, we have assessed the ability of Pik TE and a library of variants to catalyze formation of a hybrid 12-membered macrolactone/lactam ring from an unnatural amide hexaketide substrate. We employed directed evolution to engineer the Pik TE, starting from the previously reported Pik TE S148C mutant[40] for an increase in macrocyclization from 13% up to 79% isolated yield. With three rounds of evolution, we generated and screened >6,000 mutants through site saturation mutagenesis focused on the active site region (R1), and random mutagenesis covering the entire protein sequence (R2, R3). We initially identified residues of potential importance for cyclization using the previously reported Pik TE WT crystal structure with a covalently bound pentaketide mimic,[34] which enabled target residues of interest to be identified for directed evolution. We expect that applying computational and ML-based approaches[40, 126-129] will deepen our mechanistic understanding of the cyclization process and the impact of select mutations on catalytic productivity. Moreover, our Pik TE variant library will provide an ongoing resource to identify efficient biocatalysts for varied hexaketide/heptaketide substrates with the goal of obtaining novel bioactive macrolides and related molecules.

Conflicts of interest

The authors declare no competing financial interest.

Acknowledgements

We are grateful for support from NIH grant R35 GM118101, a Diversity Supplement to R35 GM118101S, a F31 Fellowship GM143769 (to M.L.A.C.) and the Hans W. Vahlteich Professorship (to D.H.S.). We also thank the UM Rackham Graduate Student Research Grant program. We thank Filipa Pereira for advice and assistance in data analysis, interpretation and graphic design, also to Sean Newmister, Matthew S. Sigman, Austin LeSueur and Hanna D. Clements for helpful scientific discussions. The authors acknowledge Rajani Arora, UM LSI Multimedia and Social Media Specialist, for her contribution to the TOC figure.

References

1. I. Saridakis, D. Kaiser and N. Maulide, Unconventional Macrocyclizations in Natural Product Synthesis, *ACS Cent. Sci.*, 2020, **6**, 1869-1889.
2. X. Yu and D. Sun, Macrocyclic Drugs and Synthetic Methodologies Toward Macrocycles, *Molecules*, 2013, **18**, 6230-6268.
3. E. M. Driggers, S. P. Hale, J. Lee and N. K. Terrett, The Exploration of Macrocycles for Drug Discovery - An Underexploited Structural Class, *Nat. Rev. Drug Discov.*, 2008, **7**, 608-624.
4. M. D. Cummings and S. Sekharan, Structure-Based Macrocycle Design in Small-Molecule Drug Discovery and Simple Metrics To Identify Opportunities for Macrocyclization of Small-Molecule Ligands, *J. Med. Chem.*, 2019, **62**, 6843-6853.
5. H. A. Kirst, Semi-Synthetic Derivatives of Erythromycin, *Prog. Med. Chem.*, 1993, **30**, 57-88.
6. M. J. Parnham, V. E. Haber, E. J. Giamarellos-Bourboulis, G. Perletti, G. M. Verleden and R. Vos, Azithromycin: Mechanisms of Action and their Relevance for Clinical Applications, *Pharmacol. Ther.*, 2014, **143**, 225-245.
7. E. Breiner-Goldstein, Z. Eyal, D. Matzov, Y. Halfon, G. Cimicata, M. Baum, A. Rokney, A. V. Ezernitchi, A. N. Lowell, J. J. Schmidt, H. Rozenberg, E. Zimmerman, A. Bashan, L. Valinsky, Y. Anzai, D. H. Sherman and A. Yonath, Ribosome-Binding and Anti-Microbial Studies of the Mycinamicins, 16-Membered Macrolide Antibiotics from Micromonospora griseorubida, *Nucleic Acids Res.*, 2021, **49**, 9560-9573.
8. C. N. S. S. P. Kumar, in *Organic Synthesis - A Nascent Relook*, ed. B. P. Nandeshwarappa, 2019, DOI: 10.5772/intechopen.87898.
9. P. Fernandes, E. Martens and D. Pereira, Nature Nurtures the Design of New Semi-Synthetic Macrolide Antibiotics, *J. Antibiot. (Tokyo)*, 2017, **70**, 527-533.
10. P. B. Shinde, H. S. Oh, H. Choi, K. Rathwell, Y. H. Ban, E. J. Kim, I. Yang, D. G. Lee, D. H. Sherman, H. Y. Kang and Y. J. Yoon, Chemoenzymatic Synthesis of Glycosylated Macrolactam Analogues of the Macrolide Antibiotic YC-17, *Adv. Synth. Catal.*, 2015, **357**, 2697-2711.
11. A. N. Lowell, M. D. DeMars, S. T. Slocum, F. A. Yu, K. Anand, J. A. Chemler, N. Korakavi, J. K. Priessnitz, S. R. Park, A. A. Koch, P. J. Schultz and D. H. Sherman, Chemoenzymatic Total Synthesis and Structural Diversification of Tylactone-Based Macrolide Antibiotics through Late-Stage Polyketide Assembly, Tailoring, and C-H Functionalization, *J. Am. Chem. Soc.*, 2017, **139**, 7913-7920.
12. D. Jelić and R. Antolović, From Erythromycin to Azithromycin and New Potential Ribosome-Binding Antimicrobials, *Antibiotics (Basel)*, 2016, **5**, 29.
13. K. Tatsuta, Total Synthesis of the Big Four Antibiotics and Related Antibiotics, *J. Antibiot. (Tokyo)*, 2013, **66**, 107-129.
14. P. C. Hogan, C. L. Chen, K. M. Mulvihill, J. F. Lawrence, E. Moorhead, J. Rickmeier and A. G. Myers, Large-Scale Preparation of Key Building Blocks for the Manufacture of Fully Synthetic Macrolide Antibiotics, *J. Antibiot. (Tokyo)*, 2018, **71**, 318-325.
15. V. Marti-Centelles, M. D. Pandey, M. I. Burguete and S. V. Luis, Macrocyclization Reactions: The Importance of Conformational, Configurational, and Template-Induced Preorganization, *Chem. Rev.*, 2015, **115**, 8736-8834.

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

16. A. Parenty, X. Moreau, G. Niel and J. M. Campagne, Update 1 of: Macrolactonizations in the Total Synthesis of Natural Products, *Chem. Rev.*, 2013, **113**, PR1-PR40.
17. A. Fürstner, Lessons from Natural Product Total Synthesis: Macrocyclization and Postcyclization Strategies, *Acc. Chem. Res.*, 2021, **54**, 861-874.
18. R. B. Woodward, E. Logusch, K. P. Nambiar, K. Sakan, D. E. Ward, B. W. Auyeung, P. Balaram, L. J. Browne, P. J. Card, C. H. Chen, R. B. Chenevert, A. Fliri, K. Frobel, H. J. Gais, D. G. Garratt, K. Hayakawa, W. Heggie, D. P. Hesson, D. Hoppe, I. Hoppe, J. A. Hyatt, D. Ikeda, P. A. Jacobi, K. S. Kim, Y. Kobuke, K. Kojima, K. Krowicki, V. J. Lee, T. Leutert, S. Malchenko, J. Martens, R. S. Matthews, B. S. Ong, J. B. Press, T. V. Rajanbabu, G. Rousseau, H. M. Sauter, M. Suzuki, K. Tatsuta, L. M. Tolbert, E. A. Truesdale, I. Uchida, Y. Ueda, T. Uyehara, A. T. Vasella, W. C. Vladuchick, P. A. Wade, R. M. Williams and H. N. C. Wong, Asymmetric Total Synthesis of Erythromycin. 1. Synthesis of an Erythronolide A Seco Acid-Derivative via Asymmetric Induction, *J. Am. Chem. Soc.*, 1981, **103**, 3210-3213.
19. R. B. Woodward, E. Logusch, K. P. Nambiar, K. Sakan, D. E. Ward, B. W. Auyeung, P. Balaram, L. J. Browne, P. J. Card, C. H. Chen, R. B. Chenevert, A. Fliri, K. Frobel, H. J. Gais, D. G. Garratt, K. Hayakawa, W. Heggie, D. P. Hesson, D. Hoppe, I. Hoppe, J. A. Hyatt, D. Ikeda, P. A. Jacobi, K. S. Kim, Y. Kobuke, K. Kojima, K. Krowicki, V. J. Lee, T. Leutert, S. Malchenko, J. Martens, R. S. Matthews, B. S. Ong, J. B. Press, T. V. Rajanbabu, G. Rousseau, H. M. Sauter, M. Suzuki, K. Tatsuta, L. M. Tolbert, E. A. Truesdale, I. Uchida, Y. Ueda, T. Uyehara, A. T. Vasella, W. C. Vladuchick, P. A. Wade, R. M. Williams and H. N. C. Wong, Asymmetric Total Synthesis of Erythromycin. 2. Synthesis of an Erythronolide A Lactone System, *J. Am. Chem. Soc.*, 1981, **103**, 3213-3215.
20. R. B. Woodward, E. Logusch, K. P. Nambiar, K. Sakan, D. E. Ward, B. W. Auyeung, P. Balaram, L. J. Browne, P. J. Card, C. H. Chen, R. B. Chenevert, A. Fliri, K. Frobel, H. J. Gais, D. G. Garratt, K. Hayakawa, W. Heggie, D. P. Hesson, D. Hoppe, I. Hoppe, J. A. Hyatt, D. Ikeda, P. A. Jacobi, K. S. Kim, Y. Kobuke, K. Kojima, K. Krowicki, V. J. Lee, T. Leutert, S. Malchenko, J. Martens, R. S. Matthews, B. S. Ong, J. B. Press, T. V. Rajanbabu, G. Rousseau, H. M. Sauter, M. Suzuki, K. Tatsuta, L. M. Tolbert, E. A. Truesdale, I. Uchida, Y. Ueda, T. Uyehara, A. T. Vasella, W. C. Vladuchick, P. A. Wade, R. M. Williams and H. N. C. Wong, Asymmetric Total Synthesis of Erythromycin. 3. Total Synthesis of Erythromycin, *J. Am. Chem. Soc.*, 1981, **103**, 3215-3217.
21. I. B. Seiple, Z. Zhang, P. Jakubec, A. Langlois-Mercier, P. M. Wright, D. T. Hog, K. Yabu, S. R. Allu, T. Fukuzaki, P. N. Carlsen, Y. Kitamura, X. Zhou, M. L. Condakes, F. T. Szczypinski, W. D. Green and A. G. Myers, A Platform for the Discovery of New Macrolide Antibiotics, *Nature*, 2016, **533**, 338-345.
22. G. Force, A. Perfetto, R. J. Mayer, I. Ciofini and D. Leboeuf, Macrolactonization Reactions Driven by a Pentafluorobenzoyl Group, *Angew. Chem. Int. Ed.*, 2021, **60**, 19843-19851.
23. E. Marsault and M. L. Peterson, Macrocycles are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery, *J. Med. Chem.*, 2011, **54**, 1961-2004.
24. M. Yang, X. W. Wang and J. F. Zhao, Ynamide-Mediated Macrolactonization, *ACS Catal.*, 2020, **10**, 5230-5235.
25. S. Sengupta and G. Mehta, Macrocyclization via C-H Functionalization: A New Paradigm in Macrocycle Synthesis, *Org. Biomol. Chem.*, 2020, **18**, 1851-1876.
26. J. D. Mortison and D. H. Sherman, Frontiers and Opportunities in Chemoenzymatic Synthesis, *J. Org. Chem.*, 2010, **75**, 7041-7051.
27. J. Li, A. Amatuni and H. Renata, Recent Advances in the Chemoenzymatic Synthesis of Bioactive Natural Products, *Curr. Opin. Chem. Biol.*, 2020, **55**, 111-118.
28. C. N. Stout and H. Renata, Reinvigorating the Chiral Pool: Chemoenzymatic Approaches to Complex Peptides and Terpenoids, *Acc. Chem. Res.*, 2021, **54**, 1143-1156.
29. L. C. Yang, H. P. Deng and H. Renata, Recent Progress and Developments in Chemoenzymatic and Biocatalytic Dynamic Kinetic Resolution, *Org. Process Res. Dev.*, 2022, **26**, 1925-1943.
30. E. P. Vanable, L. G. Habgood and J. D. Patrone, Current Progress in the Chemoenzymatic Synthesis of Natural Products, *Molecules*, 2022, **27**, 6373.
31. E. M. M. Abdelraheem, H. Busch, U. Hanefeld and F. Tonin, Biocatalysis Explained: From Pharmaceutical to Bulk Chemical Production, *React. Chem. Eng.*, 2019, **4**, 1878-1894.
32. S. W. Sisco, B. M. Larson and J. S. Moore, Relaxing Conformational Constraints in Dynamic Macrocycle Synthesis, *Macromolecules*, 2014, **47**, 3829-3836.
33. D. A. Hansen, C. M. Rath, E. B. Eisman, A. R. Narayan, J. D. Kittendorf, J. D. Mortison, Y. J. Yoon and D. H. Sherman, Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin, *J. Am. Chem. Soc.*, 2013, **135**, 11232-11238.
34. D. L. Akey, J. D. Kittendorf, J. W. Giraldes, R. A. Fecik, D. H. Sherman and J. L. Smith, Structural Basis for Macrolactonization by the Pikromycin Thioesterase, *Nat. Chem. Biol.*, 2006, **2**, 537-542.
35. J. W. Giraldes, D. L. Akey, J. D. Kittendorf, D. H. Sherman, J. L. Smith and R. A. Fecik, Structural and Mechanistic Insights into Polyketide Macrolactonization from Polyketide-Based Affinity Labels, *Nat. Chem. Biol.*, 2006, **2**, 531-536.
36. K. J. Weissman, Uncovering the Structures of Modular Polyketide Synthases, *Nat. Prod. Rep.*, 2015, **32**, 436-453.
37. K. J. Weissman, Genetic Engineering of Modular PKSs: From Combinatorial Biosynthesis to Synthetic Biology, *Nat. Prod. Rep.*, 2016, **33**, 203-230.
38. M. L. Adrover-Castellano, J. J. Schmidt and D. H. Sherman, Biosynthetic Cyclization Catalysts for the Assembly of Peptide and Polyketide Natural Products, *ChemCatChem*, 2021, **13**, 2095-2116.
39. D. A. Hansen, A. A. Koch and D. H. Sherman, Identification of a Thioesterase Bottleneck in the Pikromycin Pathway through Full-Module Processing of Unnatural Pentaketides, *J. Am. Chem. Soc.*, 2017, **139**, 13450-13455.
40. A. A. Koch, D. A. Hansen, V. V. Shende, L. R. Furan, K. N. Houk, G. Jimenez-Oses and D. H. Sherman, A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst, *J. Am. Chem. Soc.*, 2017, **139**, 13456-13465.
41. A. A. Koch, J. J. Schmidt, A. N. Lowell, D. A. Hansen, K. M. Coburn, J. A. Chemler and D. H. Sherman, Probing Selectivity and Creating Structural Diversity

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

through Hybrid Polyketide Synthases, *Angew. Chem. Int. Ed.*, 2020, **59**, 13575-13580.
42. M. E. Horsman, T. P. Hari and C. N. Boddy, Polyketide Synthase and Non-Ribosomal Peptide Synthetase Thioesterase Selectivity: Logic Gate or a Victim of Fate?, *Nat. Prod. Rep.*, 2016, **33**, 183-202.
43. L. Yuan, T. A. Voelker and D. J. Hawkins, Modification of the Substrate Specificity of an Acyl-Acyl Carrier Protein Thioesterase by Protein Engineering, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 10639-10643.
44. D. C. Cantu, Y. F. Chen and P. J. Reilly, Thioesterases: A New Perspective Based on their Primary and Tertiary Structures, *Protein Sci.*, 2010, **19**, 1281-1295.
45. B. T. Caswell, C. C. de Carvalho, H. Nguyen, M. Roy, T. Nguyen and D. C. Cantu, Thioesterase Enzyme Families: Functions, Structures, and Mechanisms, *Protein Sci.*, 2022, **31**, 652-676.
46. K. M. Mayer and J. Shanklin, A Structural Model of the Plant Acyl-Acyl Carrier Protein Thioesterase FatB Comprises Two Helix/4-Stranded Sheet Domains, the N-terminal Domain Containing Residues that Affect Specificity and the C-Terminal Domain Containing Catalytic Residues, *J. Biol. Chem.*, 2005, **280**, 3621-3627.
47. M. Wang and C. N. Boddy, Examining the Role of Hydrogen Bonding Interactions in the Substrate Specificity for the Loading Step of Polyketide Synthase Thioesterase Domains, *Biochemistry*, 2008, **47**, 11793-11803.
48. T. P. Hari, P. Labana, M. Boileau and C. N. Boddy, An Evolutionary Model Encompassing Substrate Specificity and Reactivity of Type I Polyketide Synthase Thioesterases, *ChemBioChem*, 2014, **15**, 2656-2661.
49. P. Argyropoulos, F. Bergeret, C. Pardin, J. M. Reimer, A. Pinto, C. N. Boddy and T. M. Schmeing, Towards a Characterization of the Structural Determinants of Specificity in the Macrocyclizing Thioesterase for Deoxyerythronolide B Biosynthesis, *Biochim. Biophys. Acta*, 2016, **1860**, 486-497.
50. F. Jing, M. D. Yandeau-Nelson and B. J. Nikolau, Identification of Active Site Residues Implies a Two-Step Catalytic Mechanism for Acyl-ACP Thioesterase, *Biochem. J.*, 2018, **475**, 3861-3873.
51. F. Jing, L. Zhao, M. D. Yandeau-Nelson and B. J. Nikolau, Two Distinct Domains Contribute to the Substrate Acyl Chain Length Selectivity of Plant Acyl-ACP Thioesterase, *Nat. Commun.*, 2018, **9**, 860.
52. R. F. Wang, W. T. Tao, L. Liu, C. Li, L. Q. Bai, Y. L. Zhao and T. Shi, Insights into Specificity and Catalytic Mechanism of Amphotericin B/Nystatin Thioesterase, *Proteins*, 2021, **89**, 558-568.
53. D. Jiang, Y. Li, W. Wu, H. Zhang, R. Xu, H. Xu, R. Zhan and L. Sun, Identification and Engineering on the Nonconserved Residues of Metallo-Beta-Lactamase-Type Thioesterase to Improve the Enzymatic Activity, *Biotechnol. Bioeng.*, 2021, **118**, 4623-4634.
54. R. F. Little and C. Hertweck, Chain Release Mechanisms in Polyketide and Non-Ribosomal Peptide Biosynthesis, *Nat. Prod. Rep.*, 2022, **39**, 163-205.
55. R. S. Gokhale, D. Hunziker, D. E. Cane and C. Khosla, Mechanism and Specificity of the Terminal Thioesterase Domain from the Erythromycin Polyketide Synthase, *Chem. Biol.*, 1999, **6**, 117-125.
56. J. J. Schmidt, Y. Khatri, S. I. Brody, C. Zhu, H. Pietraszkiewicz, F. A. Valeriote and D. H. Sherman, A Versatile Chemoenzymatic Synthesis for the Discovery of Potent Cryptophycin Analogs, *ACS Chem. Biol.*, 2020, **15**, 524-532.
57. S. C. Tsai, L. J. Miercke, J. Krucinski, R. Gokhale, J. C. Chen, P. G. Foster, D. E. Cane, C. Khosla and R. M. Stroud, Crystal Structure of the Macrocycle-Forming Thioesterase Domain of the Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**, 14808-14813.
58. S. C. Tsai, H. X. Lu, D. E. Cane, C. Khosla and R. M. Stroud, Insights into Channel Architecture and Substrate Specificity from Crystal Structures of Two Macrocycle-Forming Thioesterases of Modular Polyketide Synthases, *Biochemistry*, 2002, **41**, 12598-12606.
59. C. M. Kao, G. L. Luo, L. Katz, D. E. Cane and C. Khosla, Manipulation of Macrolide Ring Size by Directed Mutagenesis of a Modular Polyketide Synthase, *J. Am. Chem. Soc.*, 1995, **117**, 9105-9106.
60. H. Lu, S. C. Tsai, C. Khosla and D. E. Cane, Expression, Site-Directed Mutagenesis, and Steady State Kinetic Analysis of the Terminal Thioesterase Domain of the Methymycin/Picromycin Polyketide Synthase, *Biochemistry*, 2002, **41**, 12590-12597.
61. A. Pinto, M. Wang, M. Horsman and C. N. Boddy, 6-Deoxyerythronolide B Synthase Thioesterase-Catalyzed Macrocyclization Is Highly Stereoselective, *Org. Lett.*, 2012, **14**, 2278-2281.
62. R. M. Kohli and C. T. Walsh, Enzymology of Acyl Chain Macrocyclization in Natural Product Biosynthesis, *Chem. Commun.*, 2003, DOI: 10.1039/b208333g, 297-307.
63. C. M. D. Swarbrick, J. D. Nanson, E. I. Patterson and J. K. Forwood, Structure, Function, and Regulation of Thioesterases, *Prog. Lipid Res.*, 2020, **79**, 101036.
64. J. Kaur and R. Sharma, Directed Evolution: An Approach to Engineer Enzymes, *Crit. Rev. Biotechnol.*, 2006, **26**, 165-199.
65. D. P. Nannemann, W. R. Birmingham, R. A. Scism and B. O. Bachmann, Assessing Directed Evolution Methods for the Generation of Biosynthetic Enzymes with Potential in Drug Biosynthesis, *Future Med. Chem.*, 2011, **3**, 809-819.
66. P. A. Dalby, Strategy and Success for the Directed Evolution of Enzymes, *Curr. Opin. Struct. Biol.*, 2011, **21**, 473-480.
67. F. H. Arnold, Directed Evolution: Bringing New Chemistry to Life, *Angew. Chem. Int. Ed.*, 2018, **57**, 4143-4148.
68. H. Xiao, Z. Bao and H. Zhao, High Throughput Screening and Selection Methods for Directed Enzyme Evolution, *Ind. Eng. Chem. Res.*, 2015, **54**, 4011-4020.
69. L. Ye, C. Yang and H. Yu, From Molecular Engineering to Process Engineering: Development of High-Throughput Screening Methods in Enzyme Directed Evolution, *Appl. Microbiol. Biotechnol.*, 2018, **102**, 559-567.
70. D. A. Hansen, A. A. Koch and D. H. Sherman, Substrate Controlled Divergence in Polyketide Synthase Catalysis, *J. Am. Chem. Soc.*, 2015, **137**, 3735-3738.
71. D. Pavlovic, A. Fajdetic and S. Mutak, Novel Hybrids of 15-Membered 8a- and 9a-Azahomoerythromycin A Ketolides and Quinolones as Potent Antibacterials, *Bioorg. Med. Chem.*, 2010, **18**, 8566-8582.
72. V. Stimac, S. Alihodzic, G. Lazarevski, S. Mutak, Z. M. Istuk, A. Fajdetic, I. Palej, H. C. Paljetak, N. Marsic, J. Padovan, B. Tavcar and V. Erakovic Haber, Synthesis and Biological Properties of 4"-O-Acyl Derivatives of 8a-Aza-8a-Homoerythromycin, *J. Antibiot. (Tokyo)*, 2009, **62**, 133-144.

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

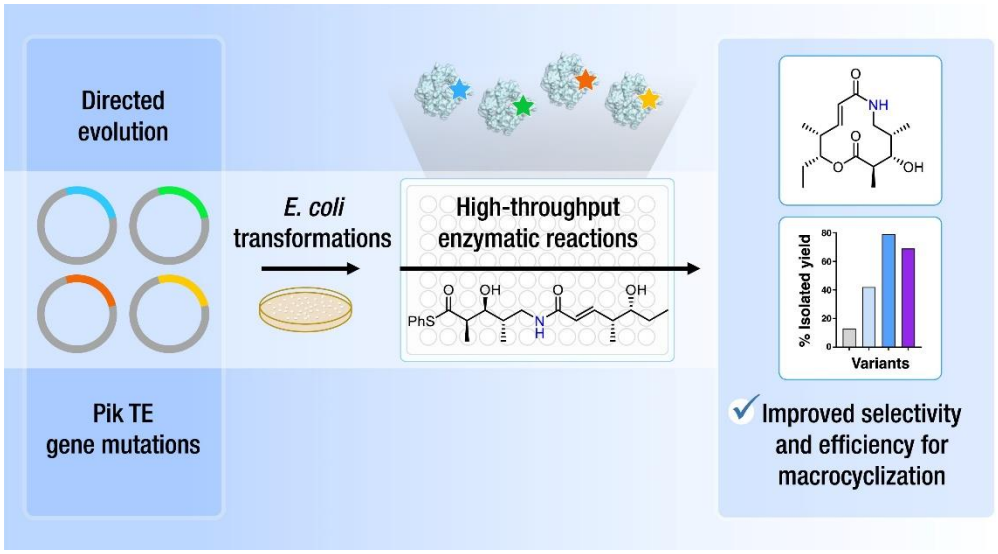
73. A. Hutinec, M. Derek, G. Lazarevski, V. Sunjic, H. C. Paljetak, S. Alihodzic, V. Erakovic Haber, M. Dumic, N. Marsic and S. Mutak, Novel 8a-Aza-8a-Homoerythromycin-4"-(3-Substituted-Amino)Propionates with Broad Spectrum Antibacterial Activity, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 3244-3249.
74. D. Pavlovic and S. Mutak, Synthesis and Structure-Activity Relationships of Novel 8a-Aza-8a-Homoerythromycin A Ketolides, *J. Med. Chem.*, 2010, **53**, 5868-5880.
75. A. Janas and P. Przybylski, 14-and 15-Membered Lactone Macrolides and their Analogues and Hybrids: Structure, Molecular Mechanism of Action and Biological Activity, *Eur. J. Med. Chem.*, 2019, **182**, 111662.
76. M. Lubberink, W. Finnigan, C. Schnepel, C. R. Baldwin, N. J. Turner and S. L. Flitsch, One-Step Biocatalytic Synthesis of Sustainable Surfactants by Selective Amide Bond Formation, *Angew. Chem. Int. Ed.*, 2022, **61**, e202205054.
77. D. G. Brown and J. Bostrom, Analysis of Past and Present Synthetic Methodologies on Medicinal Chemistry: Where Have all the New Reactions Gone?, *J. Med. Chem.*, 2016, **59**, 4443-4458.
78. E. Massolo, M. Pirola and M. Benaglia, Amide Bond Formation Strategies: Latest Advances on a Dateless Transformation, *Eur. J. Org. Chem.*, 2020, 4641-4651.
79. J. Boström, D. G. Brown, R. J. Young and G. M. Keseru, Expanding the Medicinal Chemistry Synthetic Toolbox, *Nat. Rev. Drug Discov.*, 2018, **17**, 709-727.
80. D. B. Tiz, L. Bagnoli, O. Rosati, F. Marini, C. Santi and L. Sancineto, FDA-Approved Small Molecules in 2022: Clinical Uses and Their Synthesis, *Pharmaceutics*, 2022, **14**, 2538.
81. D. G. Jimenez, V. Poongavanam and J. Kihlberg, Macrocycles in Drug Discovery-Learning from the Past for the Future, *J. Med. Chem.*, 2023, **66**, 5377-5396.
82. W. Muramatsu, T. Hattori and H. Yamamoto, Amide Bond Formation: Beyond the Dilemma Between Activation and Racemisation, *Chem. Commun.*, 2021, **57**, 6346-6359.
83. L. Bering, E. J. Craven, S. A. S. Thomas, S. A. Shepherd and J. Micklefield, Merging Enzymes with Chemocatalysis for Amide Bond Synthesis, *Nat. Commun.*, 2022, **13**, 380.
84. D. A. Evans, J. Bartroli and T. L. Shih, Enantioselective Aldol Condensations. 2. Erythro-Selective Chiral Alol Condensations via Boron Enolates, *J. Am. Chem. Soc.*, 1981, **103**, 2127-2129.
85. J. R. Gage and D. A. Evans, Diastereoselective Aldol Condensation Using a Chiral Oxazolidinone Auxiliary - (2S*, 3S*)-3-Hydroxy-3-Phenyl-2-Methylpropanoic Acid, *Org. Synth.*, 1990, **68**, 83-91.
86. J. Willwacher, N. Kausch-Busies and A. Furstner, Divergent Total Synthesis of the Antimitotic Agent Leiodermatolide, *Angew. Chem. Int. Ed.*, 2012, **51**, 12041-12046.
87. G. L. Beutner, B. M. Cohen, A. J. DelMonte, D. D. Dixon, K. J. Fraunhoffer, A. W. Glace, E. Lo, J. M. Stevens, D. Vanyo and C. Wilbert, Revisiting the Cleavage of Evans Oxazolidinones with LiOH/H2O2, *Org. Process Res. Dev.*, 2019, **23**, 1378-1385.
88. O. F. Brandenberg, K. Chen and F. H. Arnold, Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds, *J. Am. Chem. Soc.*, 2019, **141**, 8989-8995.
89. D. K. Romney, J. Murciano-Calles, J. E. Wehrmuller and F. H. Arnold, Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues, *J. Am. Chem. Soc.*, 2017, **139**, 10769-10776.
90. R. S. McIsaac, M. K. Engqvist, T. Wannier, A. Z. Rosenthal, L. Herwig, N. C. Flytzanis, E. S. Imasheva, J. K. Lanyi, S. P. Balashov, V. Gradinaru and F. H. Arnold, Directed Evolution of a Far-Red Fluorescent Rhodopsin, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 13034-13039.
91. T. A. Rogers and A. S. Bommarius, Utilizing Simple Biochemical Measurements to Predict Lifetime Output of Biocatalysts in Continuous Isothermal Processes, *Chem. Eng. Sci.*, 2010, **65**, 2118-2124.
92. K. Huynh and C. L. Partch, Analysis of Protein Stability and Ligand Interactions by Thermal Shift Assay, *Curr. Protoc. Protein Sci.*, 2015, **79**, 28.29.21-28.29.14.
93. M. D. Gomes and J. M. Woodley, Considerations when Measuring Biocatalyst Performance, *Molecules*, 2019, **24**, 3573.
94. A. S. Bommarius, Total Turnover Number - A Key Criterion for Process Evaluation, *Chem. Ing. Tech.*, 2023, **95**, 491-497.
95. S. Kozuch and J. M. L. Martin, "Turning Over" Definitions in Catalytic Cycles, *ACS Catal.*, 2012, **2**, 2787-2794.
96. K. M. Polizzi, A. S. Bommarius, J. M. Broering and J. F. Chaparro-Riggers, Stability of Biocatalysts, *Curr. Opin. Chem. Biol.*, 2007, **11**, 220-225.
97. J. D. Mortison, J. D. Kittendorf and D. H. Sherman, Synthesis and Biochemical Analysis of Complex Chain-Elongation Intermediates for Interrogation of Molecular Specificity in the Erythromycin and Pikromycin Polyketide Synthases, *J. Am. Chem. Soc.*, 2009, **131**, 15784-15793.
98. J. T. Payne, C. B. Poor and J. C. Lewis, Directed Evolution of RebH for Site-Selective Halogenation of Large Biologically Active Molecules, *Angew. Chem. Int. Ed.*, 2015, **54**, 4226-4230.
99. L. M. Mendonca and S. R. Marana, Single Mutations Outside the Active Site Affect the Substrate Specificity in a β-Glycosidase, *Biochim. Biophys. Acta*, 2011, **1814**, 1616-1623.
100. G. Rix, E. J. Watkins-Dulaney, P. J. Almhjell, C. E. Boville, F. H. Arnold and C. C. Liu, Scalable Continuous Evolution for the Generation of Diverse Enzyme Variants Encompassing Promiscuous Activities, *Nat. Commun.*, 2020, **11**, 5644.
101. L. Liu, W. T. Tao, L. Q. Bai, E. S. Kim, Y. L. Zhao and T. Shi, Why Does Tautomycetin Thioesterase Prefer Hydrolysis to Macrocyclization? Theoretical Study on its Catalytic Mechanism, *Cata. Sci. Technol.*, 2019, **9**, 6391-6403.
102. M. Goldsmith, N. Aggarwal, Y. Ashani, H. Jubran, P. J. Greisen, S. Ovchinnikov, H. Leader, D. Baker, J. L. Sussman, A. Goldenzweig, S. J. Fleishman and D. S. Tawfik, Overcoming an Optimization Plateau in the Directed Evolution of Highly Efficient Nerve Agent Bioscavengers, *Protein Eng. Des. Sel.*, 2017, **30**, 333-345.
103. F. Baier, N. Hong, G. Yang, A. Pabis, C. M. Miton, A. Barrozo, P. D. Carr, S. C. L. Kamerlin, C. J. Jackson and N. Tokuriki, Cryptic Genetic Variation Shapes the Adaptive Evolutionary Potential of Enzymes, *Elife*, 2019, **8**, e40789.
104. H. Li, Q. Q. Bao, J. F. Zhao, Y. B. Xu, S. Y. Yang, W. S. Xue, Y. Sun and Y. P. Liu, Directed Evolution Engineering to Improve Activity of Glucose Dehydrogenase by Increasing Pocket Hydrophobicity, *Front. Microbiol.*, 2022, **13**, 1044226.

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

105. N. Tokuriki and D. S. Tawfik, Stability Effects of Mutations and Protein Evolvability, *Curr. Opin. Struct. Biol.*, 2009, **19**, 596-604.
106. R. A. Studer, P. A. Christin, M. A. Williams and C. A. Orengo, Stability-Activity Tradeoffs Constrain the Adaptive Evolution of RubisCO, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 2223-2228.
107. P. A. Romero and F. H. Arnold, Exploring Protein Fitness Landscapes by Directed Evolution, *Nat. Rev. Mol. Cell. Biol.*, 2009, **10**, 866-876.
108. G. R. Yang, C. M. Miton and N. Tokuriki, A Mechanistic View of Enzyme Evolution, *Protein Sci.*, 2020, **29**, 1724-1747.
109. G. P. Pinto, M. Corbella, A. O. Demkiv and S. C. L. Kamerlin, Exploiting Enzyme Evolution for Computational Protein Design, *Trends Biochem. Sci.*, 2022, **47**, 375-389.
110. R. Obexer, A. Godina, X. Garrabou, P. R. Mittl, D. Baker, A. D. Griffiths and D. Hilvert, Emergence of a Catalytic Tetrad During Evolution of a Highly Active Artificial Aldolase, *Nat. Chem.* , 2017, **9**, 50-56.
111. S. D. Stimple, M. D. Smith and P. M. Tessier, Directed Evolution Methods for Overcoming Trade-Offs Between Protein Activity and Stability, *AIChE J.*, 2020, **66**, e16814.
112. T. Yang, Z. X. Ye and M. D. Lynch, "Multiagent" Screening Improves Directed Enzyme Evolution by Identifying Epistatic Mutations, *ACS Synth. Biol.*, 2022, **11**, 1971-1983.
113. O. Salazar, P. C. Cirino and F. H. Arnold, Thermostabilization of a Cytochrome p450 Peroxygenase, *ChemBioChem*, 2003, **4**, 891-893.
114. Y. Wang, W. Huang, N. Sathitsuksanoh, Z. Zhu and Y. H. Zhang, Biohydrogenation from Biomass Sugar Mediated by In Vitro Synthetic Enzymatic Pathways, *Chem. Biol.*, 2011, **18**, 372-380.
115. J. D. Bloom, S. T. Labthavikul, C. R. Otey and F. H. Arnold, Protein Stability Promotes Evolvability, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 5869-5874.
116. C. B. Poor, M. C. Andorfer and J. C. Lewis, Improving the Stability and Catalyst Lifetime of the Halogenase RebH by Directed Evolution, *ChemBioChem*, 2014, **15**, 1286-1289.
117. C. K. Prier, K. C. Soto, J. H. Forstater, N. Kuhl, J. T. Kuethe, W. L. Cheung-Lee, M. J. Di Maso, C. M. Eberle, S. T. Grosser, H. I. Ho, E. Hoyt, A. Maguire, K. M. Maloney, A. Makarewicz, J. P. McMullen, J. C. Moore, G. S. Murphy, K. Narsimhan, W. L. Pan, N. R. Rivera, A. Saha-Shah, D. A. Thaisrivongs, D. Verma, A. Wyatt and D. Zewge, Amination of a Green Solvent via Immobilized Biocatalysis for the Synthesis of Nemtabrutinib, *ACS Catal.*, 2023, **13**, 7707-7714.
118. S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius and U. T. Bornscheuer, Biocatalysis: Enzymatic Synthesis for Industrial Applications, *Angew. Chem. Int. Ed.*, 2021, **60**, 88-119.
119. A. Peracchi, Enzyme Catalysis: Removing Chemically 'Essential' Residues by Site-Directed Mutagenesis, *Trends Biochem. Sci.*, 2001, **26**, 497-503.
120. X. P. Chen, T. Shi, X. L. Wang, J. T. Wang, Q. H. Chen, L. Q. Bai and Y. L. Zhao, Theoretical Studies on the Mechanism of Thioesterase-Catalyzed Macrocyclization in Erythromycin Biosynthesis, *ACS Catal.*, 2016, **6**, 4369-4378.
121. P. Kokkonen, M. Slanska, V. Dockalova, G. P. Pinto, E. M. Sanchez-Carnerero, J. Damborsky, P. Klan, Z. Prokop and D. Bednar, The Impact of Tunnel Mutations on Enzymatic Catalysis Depends on the Tunnel-Substrate Complementarity and the Rate-Limiting Step, *Comput. Struct. Biotechnol. J.* , 2020, **18**, 805-813.
122. J. W. Schafer, I. Zoi, D. Antoniou and S. D. Schwartz, Optimization of the Turnover in Artificial Enzymes via Directed Evolution Results in the Coupling of Protein Dynamics to Chemistry, *J. Am. Chem. Soc.*, 2019, **141**, 10431-10439.
123. K. D. Patel, F. B. d'Andrea, N. M. Gaudelli, A. R. Buller, C. A. Townsend and A. M. Gulick, Structure of a Bound Peptide Phosphonate Reveals the Mechanism of Nocardicin Bifunctional Thioesterase Epimerase-Hydrolase Half-Reactions, *Nat. Commun.*, 2019, **10**, 3868.
124. T. Shi, L. X. Liu, W. T. Tao, S. G. Luo, S. B. Fan, X. L. Wang, L. Q. Bai and Y. L. Zhao, Theoretical Studies on the Catalytic Mechanism and Substrate Diversity for Macrocyclization of Pikromycin Thioesterase, *ACS Catal.*, 2018, **8**, 4323-4332.
125. Y. C. Zhou, W. T. Tao, Z. Qi, J. H. Wei, T. Shi, Q. J. Kang, J. T. Zheng, Y. L. Zhao and L. Q. Bai, Structural and Mechanistic Insights into Chain Release of the Polyene PKS Thioesterase Domain, *ACS Catal.*, 2022, **12**, 762-776.
126. S. P. Kelly, V. V. Shende, A. R. Flynn, Q. Dan, Y. Ye, J. L. Smith, S. Tsukamoto, M. S. Sigman and D. H. Sherman, Data Science-Driven Analysis of Substrate-Permissive Diketopiperazine Reverse Prenyltransferase NotF: Applications in Protein Engineering and Cascade Biocatalytic Synthesis of (-)-Eurotiumin A, *J. Am. Chem. Soc.*, 2022, **144**, 19326-19336.
127. S. Yang, M. D. DeMars, J. M. Grandner, N. M. Olson, Y. Anzai, D. H. Sherman and K. N. Houk, Computational-Based Mechanistic Study and Engineering of Cytochrome P450 MycG for Selective Oxidation of 16-Membered Macrolide Antibiotics, *J. Am. Chem. Soc.*, 2020, **142**, 17981-17988.
128. V. V. Shende, N. R. Harris, J. N. Sanders, S. A. Newmister, Y. Khatri, M. Movassaghi, K. N. Houk and D. H. Sherman, Molecular Dynamics Simulations Guide Chimeragenesis and Engineered Control of Chemoselectivity in Diketopiperazine Dimerases, *Angew. Chem. Int. Ed.*, 2023, **62**, e2022102.
129. H. D. Clements, A. R. Flynn, B. T. Nicholls, D. Grosheva, S. J. Lefave, M. T. Merriman, T. K. Hyster and M. S. Sigman, Using Data Science for Mechanistic Insights and Selectivity Predictions in a Non-Natural Biocatalytic Reaction, *J. Am. Chem. Soc.*, 2023, **145**, 17656-17664.
130. M. C. Andorfer, D. Evans, S. Yang, C. Q. He, A. M. Girlich, J. Vergara-Coll, N. Sukumar, K. N. Houk and J. C. Lewis, Analysis of Laboratory-Evolved Flavin-Dependent Halogenases Affords a Computational Model for Predicting Halogenase Site Selectivity, *Chem. Catal.*, 2022, **2**, 2658-2674.
131. B. J. Wittmann, K. E. Johnston, Z. Wu and F. H. Arnold, Advances in Machine Learning for Directed Evolution, *Curr. Opin. Struct. Biol.*, 2021, **69**, 11-18.

https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Table of Contents Artwork:



https://doi.org/10.26434/chemrxiv-2024-02cj6 **ORCID:** https://orcid.org/0000-0001-8334-3647 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0