## Title

A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst



## Journal

Journal of the American Chemical Society, 139(38)

## ISSN

0002-7863

## Authors

Koch, Aaron A
Hansen, Douglas A
Shende, Vikram V
et al.

## Publication Date

2017-09-27

## DOI

10.1021/jacs.7b06436

Peer reviewed

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

# A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst

Aaron A. Koch,[†,‡] Douglas A. Hansen,[†,§] Vikram V. Shende,[†,⊥] Lawrence R. Furan,[†] K. N. Houk,*[,▽] Gonzalo Jiménez-Osés,*[,⊗] and David H. Sherman*[,†,‡,∥,#]

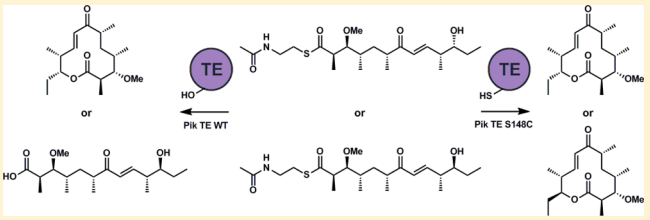
[†]Life Sciences Institute, [‡]Cancer Biology Graduate Program, [§]Department of Medicinal Chemistry, [⊥]Program in Chemical Biology, [∥]Department of Chemistry, and [#]Department of Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109, United States

[▽]Department of Chemistry and Biochemistry, University of California, Los Angeles, 607 Charles E. Young Drive East, Los Angeles, California 90095, United States

[⊗]Departamento de Química, Centro de Investigación en Síntesis Química, Universidad de La Rioja, 26006 Logroño, La Rioja, Spain

Supporting Information

**ABSTRACT:** Macrolactonization of natural product analogs presents a significant challenge to both biosynthetic assembly and synthetic chemistry. In the preceding paper, we identified a thioesterase (TE) domain catalytic bottleneck processing unnatural substrates in the pikromycin (Pik) system, preventing the formation of epimerized macrolactones. Here, we perform molecular dynamics simulations showing the epimerized hexaketide was accommodated within the Pik TE active site; however, intrinsic conformational preferences of the substrate resulted in predominately unproductive conformations, in agreement with the observed hydrolysis. Accordingly, we engineered the stereoselective Pik TE to yield a variant ($TE_{S148C}$) with improved reaction kinetics and gain-of-function processing of an unnatural, epimerized hexaketide. Quantum mechanical comparison of model $TE_{S148C}$ and $TE_{WT}$ reaction coordinate diagrams revealed a change in mechanism from a stepwise addition−elimination ($TE_{WT}$) to a lower energy concerted acyl substitution ($TE_{S148C}$), accounting for the gain-of-function and improved reaction kinetics. Finally, we introduced the S148C mutation into a polyketide synthase module (PikAIII-TE) to impart increased substrate flexibility, enabling the production of diastereomeric macrolactones.




## INTRODUCTION

Macrocycles are a common motif among natural product and natural product derived therapeutics, with >100 marketed drugs possessing a macrocyclic core.[1] The conformational preorganization of these large rings enables precise display of functionality to engage challenging targets such as protein−protein interactions, and, in the case of macrolides, ribosomal machinery.[2−4]

The macrolactone core of macrolide[5] natural products is formed through two steps: first, the linear intermediate is transferred from the upstream acyl carrier protein (ACP) to a catalytic serine of the TE via a transesterification to form an acyl-enzyme complex. Next, the acyl intermediate is offloaded via intermolecular water hydrolysis or intramolecular nucleophilic attack to release the product as a linear acid or macrolactone, respectively[6] (Figure 1).

Biochemical and structural studies of the Pik TE and related DEBS TE (erythromycin pathway) have provided initial insights into the mechanism of ring formation over hydrolysis.[7,8] While it is generally accepted that modular type I TEs have high substrate flexibility for the initial acylation step, the second, macrolactone forming release step is far more stringent.[6] Since the formation of the macrocyclic core is essential for downstream tailoring and biological activity of a natural product, aberrant hydrolysis limits access to new macrocyclic analogs.

In an preceding paper,[9] we employed a series of synthetic Pik pentaketides to evaluate the substrate tolerance of PikAIII-TE. Characterization of the in vitro reaction products indicated a catalytic bottleneck localized to the Pik TE domain, preventing macrolactonization to form C-10 and C-11 epimers of 10-dml (**1**). Furthermore, this bottleneck was confirmed directly by incubation of **5** with Pik TE resulting exclusively in hydrolysis[9] (Figure 2). Given the importance of the TE domain in production of macrolactone analogs, we sought to engineer the excised Pik TE domain to expand its substrate scope and, potentially, improve reaction rates. In this study, we report the identification of a single active site mutation (S148C) in Pik TE that generates a more effective macrolactonization catalyst by increasing both the substrate flexibility and efficiency of the enzyme. Subsequent introduction of the engineered TE into PikAIII-TE results in a polyketide synthase with increased

**Received:** June 20, 2017
**Published:** August 24, 2017

DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

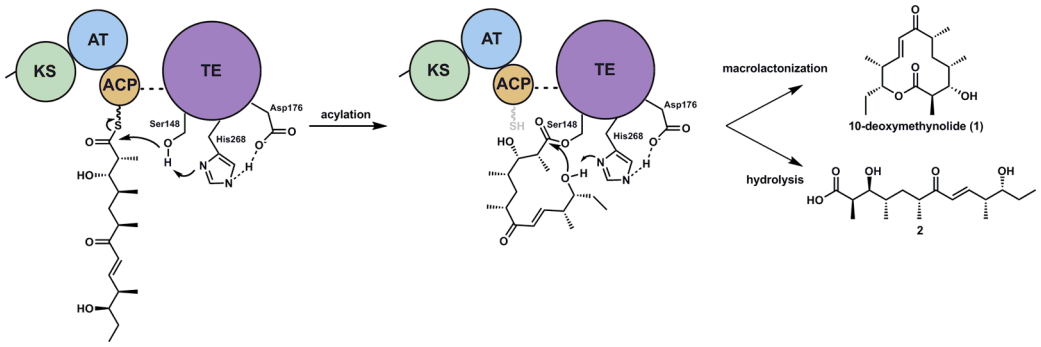



**Figure 1.** Macrolactonization or hydrolysis of an ACP-tethered polyketide intermediate.

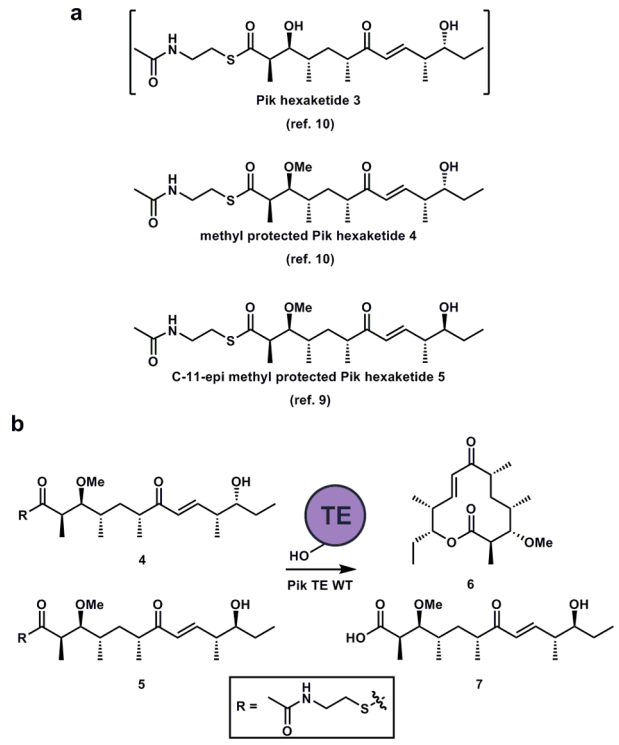



**Figure 2.** Pik TE displays a high level of substrate stereoselectivity. (a) Pik hexaketides used in this study to probe Pik TE substrate flexibility. **3** is generated in situ by photolysis of the 2-nitrobenzyloxymethyl ether (NBOM) protected native hexaketide.[10] (b) Incubation of C-11-epimerized Pik hexaketide 5 results exclusively in hydrolysis.[9]

substrate flexibility capable of producing epimeric macrolactones.

## RESULTS

**Molecular Dynamics (MD) Simulations of Pik TE Macrolactonization.** Intrigued by the stark catalytic divergence from macrolactonization to hydrolysis in Pik TE reactions containing hexaketide **5** bearing the epimerized C-11 hydroxyl group (Figure 2), we envisioned that comparing MD simulations of Pik $TE_{WT}$ acylated with hexaketides **4** and **5** (denoted Pik TE-**4** and Pik TE-**5**, respectively, Figure 3) could provide structural insight into these experimental findings. Accordingly, we performed MD simulations of each catalytic system modeled as acyl-enzyme intermediates bound to the C-3 methyl-protected hexaketides (Figure 3), using the conformation of the derivatized products[10] as the starting arrangement for the individual substrates. MD simulations (see Computational Details in the Supporting Information) were initiated with the hexaketide C-11 alcohol constrained in a reactive conformation by imposing a maximum distance restraint of 2.3 Å between the hexaketide nucleophilic hydroxyl hydrogen and the Nε nitrogen of His268. After 50 ns, the distance restraint was removed and the simulations were allowed to continue for 500 ns.

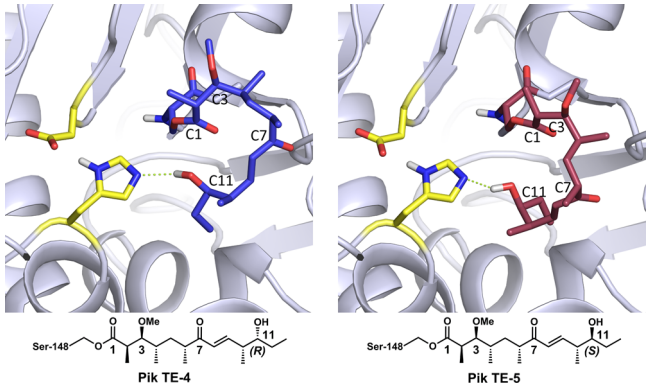



**Figure 3.** Acyl-enzyme starting structures for the MD simulations of **4** (Pik TE-**4**) and C-11-epimerized **5** (Pik TE-**5**).

Simulations of Pik TE-**4** revealed that while the initial catalytic restraint was in place, the native hexaketide readily adopts two main conformations favorable toward macrolactonization (denoted as **I** and **II**, Figure 4a). At the initial stages of the simulation, conformation **I** predominates, accompanied by a high level of shape complementarity with the TE active site. The main interactions constituting conformation **I** are hydrophobic contacts between the hexaketide and the side chains of residues lining the active site. Additionally, a hydrogen bond between the hexaketide C-7 carbonyl and the side chain of Thr77 facilitates stabilization of the cyclic, reactive conformation (Figure S13). After ~30 ns the hexaketide evolves to conformation **II** through a substrate tail rotation that places the C-11 alcohol 0.2 and 0.7 Å closer to His268 and the C-1 carbonyl, respectively. This orientation resembles the macrolactonization transition state (*vide infra*) even closer than conformation **I** and is likely catalytically productive.

Following formation of conformation **II**, the hexaketide displays a large amount of conformational freedom within the TE active site. In the most prevalent conformation (**III**, Figure 4a) hydrophobic packing is largely reduced and nonproductive hydrogen bonds between the C-11 alcohol and Tyr25/Leu193 are formed (Figure S13). Loss of the key hydrophobic


DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

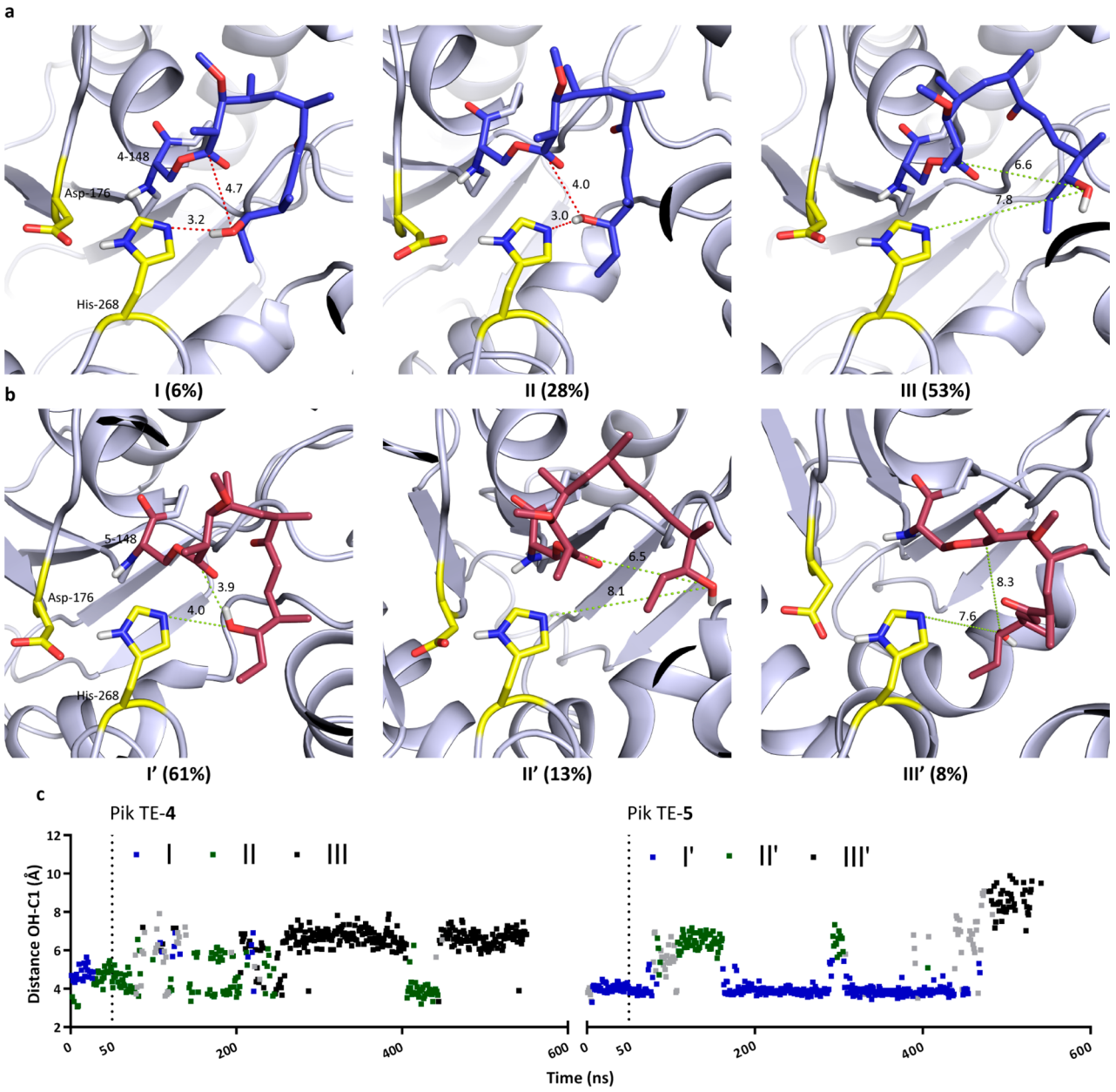



**Figure 4.** Comparison of the reactive conformations for each acyl-enzyme intermediate obtained from clustering analysis of MD simulations with Pik $TE_{WT}$: (a) Pik TE-**4** and (b) Pik TE-**5**. Pik TE-**4** conformations **I** and **II** contain a hexaketide orientation most conducive to macrolactonization with the C-11 OH in close proximity to both His268 and the C-1 carbonyl. The corresponding conformation (cluster **I′**) in the Pik TE-**5** simulation likely represents a larger barrier to macrolactonization as the distance between the C-11 OH and His268 has increased and the resulting geometry hinders deprotonation. The Pik TE-**5** hexaketide continues to evolve toward a linear conformation until the final cluster **III′** is reached which places the C-11 OH distal to both His268 and the C-1 carbonyl and in an orientation susceptible to hydrolysis. The catalytic triad His268 and Asp176 residues are colored yellow. For each conformation the distance in angstroms from the nucleophilic hydroxyl oxygen to the N$\varepsilon$ nitrogen of His268 and the ester C-1 carbonyl is displayed above the dashed lines. Clusters containing catalytically productive conformations contain red dashed lines. (c) The distance of the nucleophilic hydroxyl oxygen and the ester C-1 carbonyl plotted for each frame of the MD simulation with each data point colored according to the corresponding clustered conformation. The vertical dashed line at 50 ns indicates when the distance constraints were released.

interactions is detrimental to macrolactonization as it increases the conformational space accessible to the hexaketide, which adopts a more linear, unreactive conformation. Although conformation **III** accounts for the majority of the Pik TE-**4** simulation, conformation **II** developed again later in the trajectory (406−443 ns, Figure 4c), suggesting that productive conformations for macrocyclization are frequently sampled despite the high substrate flexibility.

The simulations of Pik TE-**5** involving the C-11-epimerized hexaketide revealed a significantly different scenario. While the initial distance restraint was in place, the hexaketide adopted a cyclic conformation stabilized by hydrophobic packing with the TE active site (**I′**, Figure 4b). However, in contrast with the native hexaketide, the C-11 alcohol of the epimerized hexaketide is prone to intramolecular hydrogen bonding with the C-1 carbonyl oxygen. Despite maintaining the substrate in a cyclic conformation with the nucleophilic hydroxyl in close proximity to the C-1 carbonyl, the resulting geometry impedes macrolactonization. Furthermore, His268 remains 4.0 Å away and positioned at an angle suboptimal for deprotonation,


DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

making this conformation catalytically unproductive. After ~80 ns, unrestricted Pik TE-**5** transitions to the second most prevalent conformation that is primarily defined by interruption of the intramolecular hydrogen-bond and rotation of the hexaketide tail away from His268 resulting in poor shape complementarity with the TE active site and a more linear conformation of the hexaketide (**II′**, Figure 4b). These two unproductive conformations are alternately sampled for the majority of the simulation (~470 ns). Afterwards, the hexaketide chain evolves toward even more extended conformations (**III′**, Figure 4b) characterized by a gradual loss of hydrophobic interactions with the protein, and formation of unproductive hydrogen bonds between the C-11 alcohol and Glu85 (Figure S13). This conformation impedes macrolactonization and increases the likelihood competing hydrolysis of the acyl-enzyme intermediate will occur.

We next compared the frequency of productive conformations sampled in the MD simulations (Figure 5) by calculating

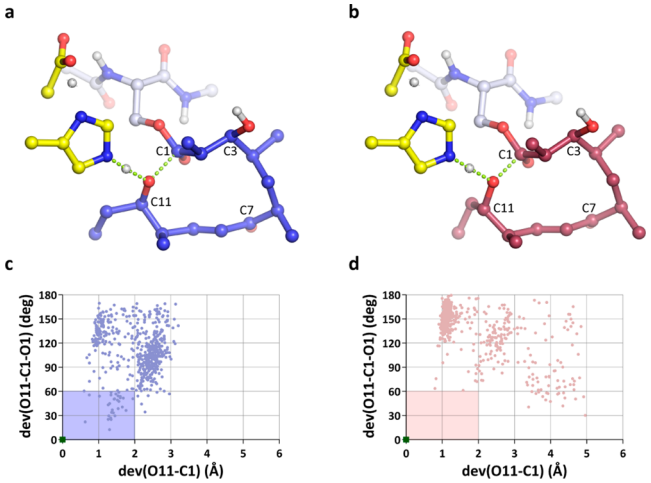
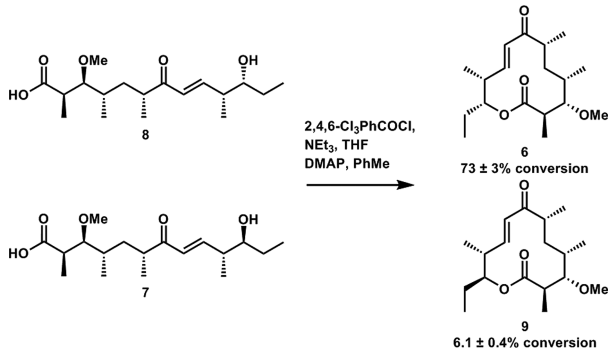



**Figure 5.** Procatalytic sampling of Pik TE during MD simulations. (a,b) Low-energy QM optimized transition states for macrolactonization of Pik hexaketides (a) Pik TE-**4** and (b) Pik TE-**5**. Nonpolar hydrogens have been removed for clarity. (c,d) Deviations of the key catalytic distances (*x* axis) and angles (*y* axis) in the MD simulations of (c) Pik TE-**4** and (d) Pik TE-**5** from their respective optimized transition structure (green square at the origin of coordinates). Each point represents a single frame from the 550 ns simulation, while the shaded rectangles represent frames from the MD that are likely in a catalytically productive state.

quantum mechanically (QM) optimized macrolactonization transition structures (TS) for each hexaketide (*vide infra*). We used these models to describe the precise angle of nucleophilic attack (O11−C1−O1) and distance (O11−C1) for macrolactonization of each hexaketide. The resulting geometric values were then compared to those extracted from each frame of the corresponding MD simulations and the entire trajectories were plotted according to their deviation from the ideal TS values (Figure 5). Consistent with our clustering analysis of the hexaketide conformations, comparison of the geometric deviations from ideal within each simulation revealed that Pik TE-**4** contained more frames with the hexaketide in a position favorable to macrolactonization (Figure 5c). In contrast, Pik TE-**5** presented a larger distribution of geometries with very few catalytically productive structures (Figure 5d). Taken together, the MD results on the covalent acyl-enzyme intermediates indicate that hydrophobic interactions accompanying a high level of substrate-TE shape complementarity are critical for maintaining each hexaketide in a catalytically productive conformation. These contacts are, to a lesser extent, maintained in simulations containing the C-11-epimerized hexaketide **5**, indicating that the Pik TE active site has sufficient flexibility to accommodate both epimers. The intrinsic conformational preferences of each hexaketide further influence catalysis. The acyl intermediates must reach conformations that are matched for deprotonation by His268 and subsequent nucleophilic attack in order to achieve macrolactonization. When these structural preferences are perturbed, as in the case with Pik TE-**5**, the energetic barrier to macrolactonization is increased above that of competing hydrolysis.

**Chemical Lactonizations.** As the MD simulations containing covalently bound hexaketides **4** and **5** revealed that both substrates are accommodated by the Pik TE active site, we next investigated if macrolactonization of the C-11-epimerized **5** suffered in bulk solution outside of enzymatic constraints. Starting with conditions employed for similar seco-acids,[11,12] we screened contemporary lactonization methodologies capable of cyclizing **8** to **6**. Optimization of Yamaguchi conditions originally reported from the total synthesis of spinosyn analogs,[13] enabled direct comparison of lactonization efficiency of **8** and **7**. Employing identical reaction conditions resulted in 73% and 6% conversions to generate **6** and **9**, respectively (Scheme 1). This result confirmed the intrinsically poor reactivity of **7** toward macrocyclization compared to **8**, corroborating the MD findings that the lack of macrolactone formation with the C-11-epimerized hexaketide is not due to steric clashes within the Pik TE active site.

**Scheme 1. Yamaguchi Macrolactonization of Methyl-Protected Hexaketides**[a]



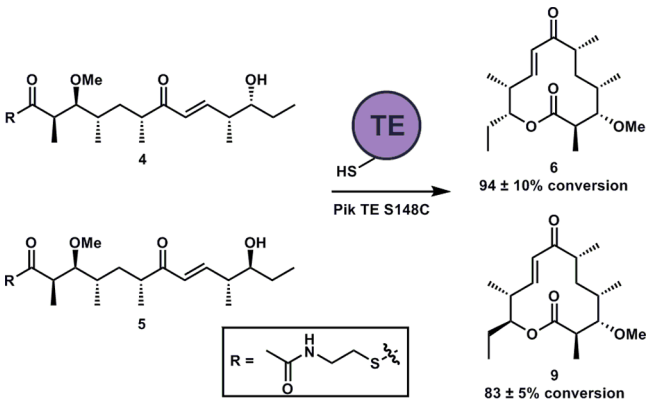

[a]Conversion of **8** to **6** and **7** to **9** was monitored by HPLC, with data represented as the mean ± standard deviation, where $n = 3$.

**Pik TE$_{S148C}$ Macrolactonization of epi-Hexaketide 5.** As the strict substrate stereoselectivity of Pik TE$_{WT}$ does not arise from inherent steric clashes within the TE active site, we considered the excised Pik TE domain to be a prime engineering target to generate a catalyst capable of cyclizing **5**. We envisioned that altering substrate specificity would likely require a protracted directed evolution campaign and, if successful, would potentially yield a variant incapable of cyclizing the starting, native substrate **3**. Prior to library generation, we performed preliminary site directed mutagenesis on select residues in the active site, including a serine to cysteine mutation of the serine-histidine-aspartate catalytic triad.[14,15] Remarkably, this variant was able to efficiently cyclize


DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

5 to 9 while also displaying improved native functionality (Scheme 2).

**Scheme 2. Evaluation of Pik $TE_{S148C}$ with Methyl-Protected Hexaketides 4 and 5**[a]

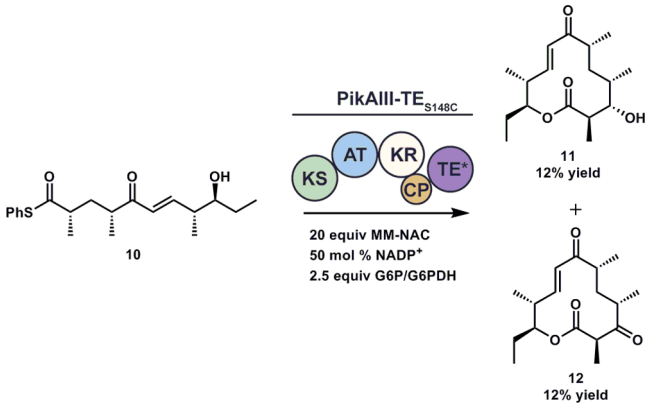



[a]Enzymatic reaction conditions: 1 mM hexaketide, 8 mM 2-vinylpyridine, purified Pik $TE_{S148C}$ (10 $\mu$M), 4 h, stationary, RT. Conversion of **4** to **6** and **5** to **9** was monitored (HPLC), with data represented as the mean ± standard deviation, where $n$ = 3.

**Kinetic Analysis of Pik TE WT and S148C.** In addition to the gain-of-function macrolactonization of **5**, we also examined the kinetic effect of the S148C mutation on cyclization (Table 1). Steady-state kinetic analysis of Pik $TE_{WT}$ and $TE_{S148C}$ was performed using both methyl-protected Pik hexaketides **4** and **5**, as well as the previously reported NBOM-protected[10] native hexaketide **3**, following a procedure similar to one used previously.[16] Kinetic analysis revealed that the serine to cysteine substitution afforded a superior cyclization catalyst for each substrate tested. Pik $TE_{S148C}$ displayed a 4.3- and 12-fold increase in the $k_{cat}/K_m$ for macrolactonization of both the native (**3**) and C-3 methoxy hexaketide **4** compared to Pik $TE_{WT}$, respectively. In reactions containing methyl-protected epimerized **5**, not only did Pik $TE_{S148C}$ retain the ability to catalyze macrolactonization of the linear substrate, but notably the $k_{cat}/K_m$ was 5.5-fold higher than for Pik $TE_{WT}$-catalyzed hydrolysis. In addition to increased substrate flexibility of Pik $TE_{S148C}$, the gain-of-function mutation provided a 2.6-fold rate enhancement with the native hexaketide **3** over wild type. These kinetic parameters demonstrate that the S148C mutation of Pik TE produces a catalyst with both expanded substrate scope and increased catalytic efficiency.

**Full-Module Processing.** We next investigated whether the engineered Pik $TE_{S148C}$ was able to improve substrate flexibility in the context of full-module catalysis. To accomplish this, we generated PikAIII-$TE_{S148C}$ and incubated it with an analogous series of epimeric Pik pentaketides. LC-HRMS analysis of reaction products revealed two new peaks in the chromatograms not observed in PikAIII-$TE_{WT}$ reactions (Figure S12), with masses and retention times corresponding to 12-membered ring macrolactones. As no authentic standards for the putative novel products were available, a 0.2 mmol scale reaction of **10** with PikAIII-$TE_{S148C}$ was performed and the reaction products were purified and characterized via MS and NMR. Structural determination of the reaction products confirmed that PikAIII-$TE_{S148C}$ was indeed able to generate 11-epi-10-dml **11** from **10**, as well as 3-keto-11-epi-10-dml **12**, due to failed reduction of the $\beta$-keto intermediate by the PikAIII KR domain prior to transfer to the terminal TE domain (Scheme 3).

**Scheme 3. Reaction of PikAIII-$TE_{S148C}$ with C-9-Epimerized Pentaketide 10**[a]




[a]Enzymatic reaction conditions: 1 mM Pik pentaketide, 20 mM (20 equiv) MM-NAC, 8 mM (8 equiv) 2-vinylpyridine, 0.5 mM (50 mol %) $NADP^+$, 2.5 mM (2.5 equiv) glucose-6-phosphate, glucose-6-phosphate dehydrogenase (2 units/mL), 3 $\mu$M (0.3 mol %) PikAIII-$TE_{S148C}$, 8 h, stationary, RT.

Diminished relative KR activity has been previously observed in vitro by the production of both the predicted reduced and 3-keto macrolactones from reactions containing Ery5-TE with DEBS pentaketide.[17] Additionally, PikAIII-PikAIV chimeras[18] lacking a KR domain yielded exclusively 3-keto-10-dml when incubated with Pik pentaketide in vivo, indicating that the WT Pik TE domain is capable of cyclizing the C-3 keto intermediate. We next investigated if the KR-TE domain competition observed in PikAIII-$TE_{S148C}$ reactions with **10** also occurred with the native Pik pentaketide. Accordingly, we performed analytical scale reactions containing PikAIII-$TE_{S148C}$ with its native substrate and analyzed the product distribution using synthetic standards for each product. HPLC quantifica-

**Table 1. Steady-State Kinetic Values for Pik $TE_{WT}$ and $TE_{S148C}$**

| substrate | TE | reaction | $k_{cat}$ ($min^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($mM^{-1}$ $min^{-1}$) |
|---|---|---|---|---|---|
| 3 | WT | cyclization | 101.7 ± 8.2 | 4.06 ± 0.68 | 25.1 ± 0.19 |
| | S148C | cyclization | 261.7 ± 19.6 | 2.40 ± 0.46 | 109 ± 0.21 |
| 4 | WT | cyclization | 11.61 ± 0.24 | 4.46 ± 0.18 | 2.60 ± 0.05 |
| | S148C | cyclization | 44.39 ± 2.01 | 1.42 ± 0.16 | 31.3 ± 0.12 |
| 5 | WT | hydrolysis | 8.58 ± 0.42 | 1.04 ± 0.16 | 8.25 ± 0.16 |
| | S148C | cyclization | 3.37 ± 0.09 | 0.07 ± 0.01 | 50 ± 0.4 |



DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

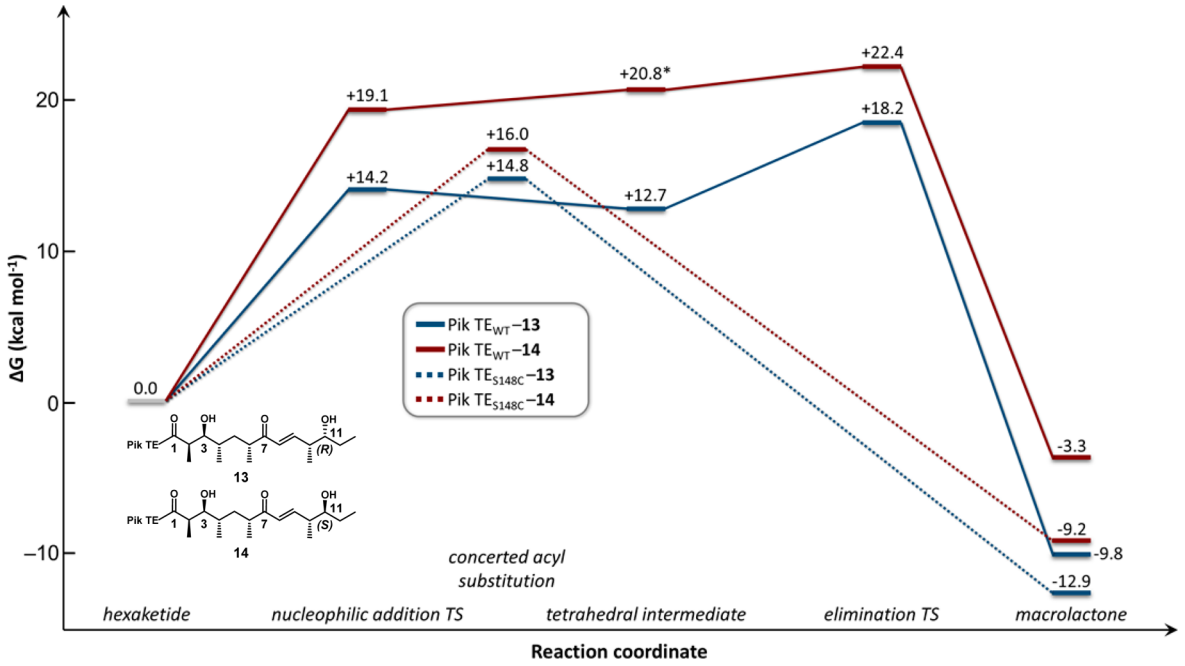



**Figure 6.** Reaction coordinate diagram representing the relative free energies for Pik TE-catalyzed macrolactonization of hexaketides **13** and **14**. Calculations were performed at the PCM/M06-2X/6-31+G(d,p) level using reduced models that define the enzymatic active site (*theozymes*[19]). Relative free energies are in kcal mol$^{-1}$. Only the lowest energy conformers are represented; see Supporting Information for details and all the calculated structures. *This intermediate is higher in energy than its preceding TS due to conformational differences between both stationary points.

tion of the reaction products revealed the conversion of native pentaketide to 3-keto-10-dml and 10-dml (**1**) to be 5.5% and 12.5%, respectively (see Supporting Information). Generation of both the reduced and unreduced macrolactone products indicates competition for the linear hexaketide intermediate between the KR and TE domains even in the context of the native substrate. Isolation of **11** and **12** at 12% yields indicates that the identified thioesterase bottleneck had been alleviated by a single amino acid change (S148C) enabling substrate flux to generate novel epimerized macrolactones.

**Quantum Mechanical (QM) Modeling of Pik TE Macrolactonization.** Since MD simulations revealed the Pik TE$_{WT}$ active site to be competent in binding the unnatural substrate **5**, we hypothesized the catalytic advantages imparted by the S148C mutation were kinetic in nature as opposed to structural. To further our understanding of the energetics of these processes and how Pik TE$_{S148C}$ is able to overcome the barrier to macrolactone formation with **5**, we turned to QM modeling of the catalytic steps comprising macrolactonization (see Computational Details in the Supporting Information) after formation of the acyl-enzyme intermediate. To accomplish this, we generated abbreviated active site models (*theozymes*[19]) for Pik TE$_{WT}$ and Pik TE$_{S148C}$ containing the native C-3-unprotected hexaketide (Pik TE$_{WT}$-**13** and Pik TE$_{S148C}$-**13**) and the C-11-epi hexaketide (Pik TE$_{WT}$-**14** and Pik TE$_{S148C}$-**14**). Analysis of the resulting free energy landscapes revealed that macrolactonization of the linear hexaketide intermediates is an exergonic process in all four systems after product release (Figure 6). Cyclization of the native hexaketide was more thermodynamically favorable compared to C-11-epimerized hexaketide, particularly with the wild-type enzyme. Moreover, the macrolactonization mechanism with each TE was calculated to change from a stepwise addition−elimination with existence of a tetrahedral intermediate in Pik TE$_{WT}$ to a concerted acyl substitution[20] upon the S148C mutation. This change in mechanism was evidenced by the flat potential energy surface which precludes the formation of a tetrahedral intermediate.

Figure 7 shows the lowest energy transition structures and associated activation barriers ($\Delta G^{\ddagger}$) for the macrolactonization of the acyl intermediates of both the natural **13** and C-11-epimerized **14** substrates with the Pik TE$_{WT}$ and TE$_{S148C}$ protein models. The rate-limiting activation barriers for the TE$_{S148C}$-catalyzed reactions are significantly lower than those for the TE$_{WT}$ reactions, which correspond to the final Ser148-acyl cleavage. Thus, Pik TE$_{S148C}$ performs the macrolactonization of the native and C-11-epimerized hexaketides with $\Delta G^{\ddagger}$ values of 14.8 and 16.0 kcal mol$^{-1}$, respectively, while Pik TE$_{WT}$ displays $\Delta G^{\ddagger}$ values of 18.2 and 22.4 kcal mol$^{-1}$ for the same substrates (Figure 7). These values predict that the rate of reactions containing the native hexaketide after the formation of the acyl-enzyme intermediate is faster in Pik TE$_{S148C}$ compared to TE$_{WT}$.

The calculated increase of 4.2 kcal mol$^{-1}$ in the activation barrier of the macrolactonization of Pik TE$_{WT}$-**13** vs Pik TE$_{WT}$-**14**, agrees well with the experimentally observed lack of epimerized macrolactone formation. This higher macrolactonization energy barrier in the TE$_{WT}$ system increases the difficulty in proceeding from the acyl-enzyme intermediate, which is vulnerable to water hydrolysis.[6] While we did not observe significant water hydrolysis in reactions containing Pik TE$_{WT}$ and the native substrate, incubation with epimerized methyl-protected **5** resulted exclusively in hydrolyzed product, the result of an inability to form the epimerized macrolactone **9**. In contrast, the 6.4 kcal mol$^{-1}$ decrease in activation barrier for Pik TE$_{S148C}$-**14** macrolactonization is consistent with our experimentally observed product formation.

## DISCUSSION

**TE Substrate Tolerance.** In vivo engineering of the DEBS biosynthetic pathway has shown that DEBS TE possesses some


DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

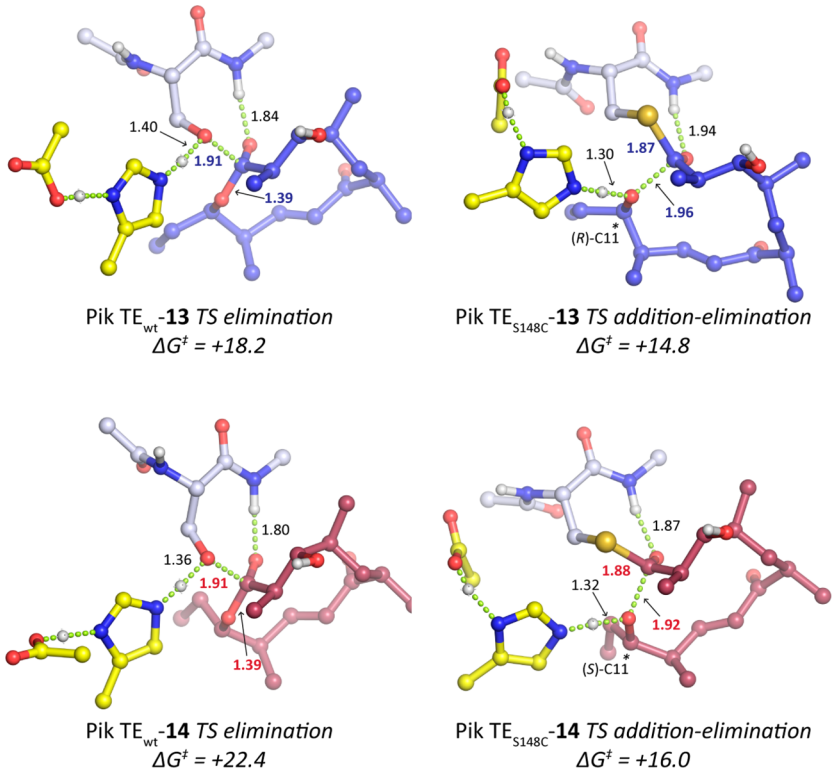



**Figure 7.** Lowest energy rate-limiting transition structures calculated with PCM/M06-2*X*/6-31+G(d,p) for the abbreviated active site models (*theozymes*[19]) of the Pik TE$_{WT}$ (left) and Pik TE$_{S148C}$ (right) catalyzed macrolactonization of native hexaketide **13** (top, in blue) and C-11-epi hexaketide **14** (bottom, in red). Activation free energies ($\Delta G^{\ddagger}$) calculated from the corresponding Pik TE hexaketides are given in kcal mol$^{-1}$ and distances in angstroms. Relevant breaking/forming C−O and C−S bonds are shown in boldface. Nonpolar hydrogens have been removed for clarity. See Supporting Information for all calculated structures along the reaction pathway.

tolerance to modifications in both the length and the functionality of the linear polyketide substrate. Modifications in the DEBS pathway have yielded a large number of 6-dEB analogs from differential extender unit incorporation and reductive processing[21−28] as well as macrolactones ranging in size from 6- to 16-membered rings.[29−32] However, the titers of these unnatural products are greatly diminished compared to wild type production levels,[6] and the inherent complexity of in vivo biosynthesis has prevented identification of pathway bottlenecks, such as TE domains.

Initial in vitro biochemical characterization of the DEBS TE[33,34] provided further evidence for the relatively high substrate tolerance of PKS TEs for acylation and hydrolysis, as terminally (omega) hydroxylated fatty acids and substrates resembling simplified DEBS heptaketides were all hydrolyzed by DEBS TE. However, the ability of PKS TEs to cyclize substrates other than their native linear intermediates has proved to be much more limited. In fact, TE mediated macrolactonization has only been observed in a select few studies.[10,16,35−42]

In addition to their native substrates, Pik TE has been shown to catalyze macrolactonization of C-3 methyl **4** and NBOM-protected derivatives,[10] though not C-7 reduced analogs.[16,36] DEBS TE has been shown to catalyze macrolactonization of unnatural mimics of the DEBS heptaketide.[39,42] However, there are no reports of either the Pik or DEBS TEs catalyzing macrolactonization of a substrate containing a nucleophilic hydroxyl group with an unnatural, epimerized (*S*)-configuration. When probed for the ability to form an epimeric heptaketide mimic of 6-dEB, the DEBS TE displayed a high level of stereoselectivity for the natural (*R*)-configuration, and exclusively hydrolyzed the unnatural (*S*)-stereoisomer,[42] adding to the observations in this study for strict stereoselectivity in TE-catalyzed macrolactonization.

**PKS Catalysis with Engineered TE Domains.** In an preceding article, we describe the identification of the TE domain as a catalytic bottleneck in the processing of unnatural substrates. Accordingly, we have engineered the TE domain for increased substrate flexibility and introduced it into PikAIII-TE to yield the engineered variant PikAIII-TE$_{S148C}$.

Direct comparison of PikAIII-TE$_{WT}$ to PikAIII-TE$_{S148C}$ with a panel of stereoisomer Pik pentaketides demonstrated the success of this approach as a 0.2 mmol scale reaction with **10** resulted in the production of two new macrolactone products, 11-epi-10-dml **11** and 3-keto-11-epi-10-dml **12**. The 1:1 isolation of **11** and **12** from reactions of **10** with PikAIII-TE$_{S148C}$ indicates that the TE was indeed responsible for the failed catalysis using PikAIII-TE$_{WT}$, as the KR domain retained a sufficient level of activity toward the unnatural intermediate. This demonstration of TE engineering offers a potential means for increasing the substrate flexibility of biosynthetic pathways for the production of natural product analogs.

In a previous report we assayed a series of different thio- and oxoester-protected hexaketides against the terminal Pik module (PikAIV) as well as the excised TE domain.[10] We found that thio- or oxoester choice determined the catalytic route of the hexaketide in the presence of PikAIV in vitro, with thiophenol and *N*-acetylcysteamine thioesters having a 10:1 preference for either full-module processing or direct cyclization, respectively. Moreover, previous studies have demonstrated that TE


DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

domains can function as hydrolases with a large degree of substrate flexibility.[16,17] Thus, we reasoned that the low yields from the PikAIII-$TE_{S148C}$ reactions with **10** are likely due to direct loading of the $TE_{S148C}$ domain and subsequent hydrolysis. Indeed, control reactions (see Supporting Information) confirmed (i) the S148C mutation overrides the KS preference previously enjoyed by thiophenol thioesters and (ii) enzymatic reactions performed without extender unit or reductive cofactors showed a much higher hydrolysis of Pik pentaketide with PikAIII-$TE_{S148C}$ relative to wild type.

Recently, Schaffer et al. reported a TE domain with a native cysteine nucleophile critical to the cyclization of a strained $\beta$-lactone ring in the obafluorin (Obi) nonribosomal peptide pathway.[43] Of note, an analogous cysteine to serine mutation in the Obi TE resulted solely in the hydrolysis product, further implicating the catalytic advantage of a cysteine TE in the cyclization of strained ring systems. Examples of modular type I polyketide TE domains containing cysteine active site nucleophiles are notably rare.[6] Further studies will focus on delineating the biosynthetic parameters that select for cysteine or serine active site residues.

**Computational Investigation of the TE Domain Catalysis.** Our investigation revealed mutual recognition, TE-substrate shape complementarity, and intrinsic substrate structural preferences to be critical for the hexaketide to reach a conformation productive toward macrolactonization. MD simulations revealed hydrophobic interactions between the substrate and TE residues lining the active site as being vital for guiding the hexaketide to a catalytically competent *pro*-cyclic conformation. These findings are consistent with mutational analysis of DEBS TE where exchange of potential hydrogen-bonding residues did not substantively affect the specificity constant for hydrolysis of four unnatural thioester substrates. In this case, Wang and Boddy suggested that hydrophobic interactions between the active site and substrate are the main driving force of substrate specificity.[37] Additionally, our results are in agreement with the available crystal structures of Pik TE bound with phosphonate substrate mimics, which displayed a lack of specific TE-substrate polar contacts.[7,8]

Recently, Chen et al.[44] reported theoretical investigations of the thioesterase domain from the erythromycin biosynthetic pathway (DEBS TE) to describe the mechanistic parameters that determine the catalytic partitioning of substrates to either macrocyclic or linear hydrolysis products. MD simulations coupled with QM calculations were performed on systems of DEBS TE modeled with the native DEBS heptaketide and Pik hexaketide, which both lead to macrolactonization, as well as two diastereomers of a reduced C-7 hydroxyl analog of Pik hexaketide that result exclusively in substrate hydrolysis. Analysis of the resulting MD simulations provided findings consistent with those in this report, particularly highlighting the importance of the formation of a substrate pre-reaction state through induced-fit mutual recognition between the enzyme and the substrate for macrolactonization to occur. Consistent with the present study, the authors found a hydrogen bond between the lactonizing hydroxyl group of the substrate and the catalytic histidine as well as hydrophobic interactions to be critical for formation of a catalytically competent pre-reaction state.

To understand the energetic consequences of the S148C mutation during the macrolactonization process, we performed DFT analysis of Pik $TE_{WT}$ and $TE_{S148C}$ modeled as acyl-enzyme intermediates with the native and C-11-epimerized hexaketides. The results of our QM calculations revealed a significant kinetic advantage in the reactions catalyzed by Pik $TE_{S148C}$. The S148C mutation provides a mechanistic change during the macrolactonization step from a two-step transesterification in the $TE_{WT}$ reaction (i.e., addition−elimination) to a lower energy single concerted step in the $TE_{S148C}$ pathway.

Overall, our combined computational method for investigating Pik TE catalysis using MD simulations in concert with QM calculations provides a plausible explanation for improved substrate flexibility and catalytic efficiency of Pik $TE_{S148C}$. According to the results from our MD simulations, epimerization of the hexaketide C-11 stereocenter generates a substrate with a reduced propensity for acquiring a catalytically competent conformation within the TE active site. However, the ability of the hexaketides to reach a conformation viable for catalysis is not the only factor affecting macrocylicization, especially if this step is not rate-limiting.[44] Hence, even if a substrate has a poor propensity to arrange in a productive conformation and/or lactonization is structurally hindered (as with the C-11-epimerized hexaketide) a significant acceleration through a key single mutation can overcome these structural limitations, boosting reactivity even with unnatural substrates.

## CONCLUSION

Based on the TE catalytic stringency observed in the processing of unnatural substrates,[9] we focused the current study on investigating full-length substrate analogs as probes for the Pik TE domain and describe the identification of a single active site mutation that generates a more effective macrolactonization catalyst. Remarkably, this single S148C mutation provided a Pik TE variant with increases in both substrate scope and catalytic efficiency. Notably, Pik $TE_{S148C}$ was able to catalyze the cyclization of a hexaketide with an epimerized nucleophilic hydroxyl to generate a novel epimerized macrolactone, while $TE_{WT}$ catalyzed complete hydrolysis of the same substrate due to inefficient macrocyclization. Furthermore, application of the engineered TE through reactions containing PikAIII-$TE_{S148C}$ allowed for the full-module processing of a C-9-epimerized Pik pentaketide **10** to the corresponding epimeric macrolactones **11** and **12**. Computational investigation of both variants revealed that a single mutation greatly lowers the activation barriers to macrolactonization; as a consequence, the catalytic process occurs faster and the substrate specificity previously dictated by shape complementarity is kinetically overcome, in line with the observed increase in $k_{cat}$.

Notably, in the context of our system, the KR domain of PikAIII-$TE_{S148C}$ displayed diminished relative activity for both the native Pik pentaketide as well as **10** with an epimerized distal hydroxyl stereocenter. The product distribution indicates competition between the KR mediated $\beta$-keto reduction and TE cyclization of the linear hexaketide intermediate. This competition is remarkable since the KR domain is fully competent toward processing the native substrate and no unreduced products are observed in reactions containing PikAIII-$TE_{WT}$, *suggesting that the engineered TE domain is able to outcompete the native catalytic sequence.*[45,46] While this domain competition for the hexaketide intermediate diminished the product yields for 11-epi-10-dml **11**, it provides insight into the effects of engineered domains on the sequence of catalytic events in PKS catalysis.

While additional biochemical studies of excised domains and full-modules with full-length analogs of native substrates are necessary to support and expand our understanding of PKS



DOI: 10.1021/jacs.7b06436
J. Am. Chem. Soc. 2017, 139, 13456−13465

function, it is becoming apparent that efficient production of specific, designer macrolide analogs will require significant pathway engineering. However, production of a desired natural product analog should be obtainable through the following workflow: (i) a targeted domain is engineered to perform an unnatural function; (ii) downstream modules are biochemically characterized in vitro with the resulting unnatural polyketide in order to identify bottlenecks; and (iii) catalytically inefficient domain(s) are engineered with the goal of restoring effective processing to generate unnatural products. Indeed, as noted in a recent review by Weissman,[47] it is perhaps more realistic to envision the future application of PKS engineering as a synthetic biology tool for producing specific, high-value natural product derivatives through targeted reprogramming of modular type I polyketide pathways rather than generation of natural product libraries.

The results described herein lay the groundwork for future engineering of PKS TE domains in order to generate more flexible catalysts for the production of novel natural product analogs.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b06436.

X-ray crystallographic data for **S2** (CIF)
Full experimental details and spectroscopic data (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Authors

*houk@chem.ucla.edu
*gonzalo.jimenez@unirioja.es
*davidhs@umich.edu

### ORCID

K. N. Houk: 0000-0002-8387-5261
Gonzalo Jiménez-Osés: 0000-0003-0105-4337
David H. Sherman: 0000-0001-8334-3647


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This research was supported by Rackham Merit Predoctoral Fellowships (D.A.H. and A.A.K.), American Foundation for Pharmaceutical Education Predoctoral Fellowship (D.A.H.), and NIH T-32-CA009676 (A.A.K.). Support for L.F. was provided through a grant from the National Science Foundation REU Program (NSF DBI #1263079). NIH Grants R35 GM118101, GM076477, the Hans W. Vahlteich Professorship (to D.H.S.) and MINECO/FEDER (grants CTQ2015-70524-R and RYC-2013-14706 to G.J.O.) are gratefully acknowledged. Calculations were performed using the Extreme Science and Engineering Discovery Environment (XSEDE), which is funded by the NSF (OCI-1053575), UR (Beronia cluster), and the UCLA Institute of Digital Research and Education (IDRE). We are grateful to Dr. J. W. Kampf for small molecule X-ray crystallography.

## ■ REFERENCES


(1) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug Discovery* **2008**, *7*, 608.
(2) Gaynor, M.; Mankin, A. *Curr. Top. Med. Chem.* **2003**, *3*, 949.
(3) Mankin, A. S. *Curr. Opin. Microbiol.* **2008**, *11*, 414.
(4) Bulkley, D.; Innis, C. A.; Blaha, G.; Steitz, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 17158.
(5) Woodward, R. B. *Angew. Chem.* **1957**, *69*, 50.
(6) Horsman, M. E.; Hari, T. P. A.; Boddy, C. N. *Nat. Prod. Rep.* **2016**, *33*, 183.
(7) Akey, D. L.; Kittendorf, J. D.; Giraldes, J. W.; Fecik, R. A.; Sherman, D. H.; Smith, J. L. *Nat. Chem. Biol.* **2006**, *2*, 537.
(8) Giraldes, J. W.; Akey, D. L.; Kittendorf, J. D.; Sherman, D. H.; Smith, J. L.; Fecik, R. A. *Nat. Chem. Biol.* **2006**, *2*, 531.
(9) Hansen, D. A.; Koch, A. A.; Sherman, D. H. *J. Am. Chem. Soc.* **2017**, DOI: 10.1021/jacs.7b06434, (preceding paper in this issue).
(10) Hansen, D. A.; Koch, A. A.; Sherman, D. H. *J. Am. Chem. Soc.* **2015**, *137*, 3735.
(11) Stang, E. M.; White, M. C. *Angew. Chem., Int. Ed.* **2011**, *50*, 2094.
(12) Inanaga, J.; Hirata, K.; Saeki, H.; Katsuki, T.; Yamaguchi, M. *Bull. Chem. Soc. Jpn.* **1979**, *52*, 1989.
(13) Tietze, L. F.; Schützenmeister, N.; Grube, A.; Scheffer, T.; Baag, M. M.; Granitzka, M.; Stalke, D. *Eur. J. Org. Chem.* **2012**, *2012*, 5748.
(14) Witkowski, A.; Witkowska, H. E.; Smith, S. *J. Biol. Chem.* **1994**, *269*, 379.
(15) Abrahmsen, L.; Tom, J.; Burnier, J.; Butcher, K. A.; Kossiakoff, A.; Wells, J. A. *Biochemistry* **1991**, *30*, 4151.
(16) He, W.; Wu, J.; Khosla, C.; Cane, D. E. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 391.
(17) Mortison, J. D.; Kittendorf, J. D.; Sherman, D. H. *J. Am. Chem. Soc.* **2009**, *131*, 15784.
(18) Chemler, J. A.; Tripathi, A.; Hansen, D. A.; O'Neil-Johnson, M.; Williams, R. B.; Starks, C.; Park, S. R.; Sherman, D. H. *J. Am. Chem. Soc.* **2015**, *137*, 10603.
(19) Tantillo, D. J.; Jiangang, C.; Houk, K. N. *Curr. Opin. Chem. Biol.* **1998**, *2*, 743.
(20) Yang, W.; Drueckhammer, D. G. *J. Am. Chem. Soc.* **2001**, *123*, 11004.
(21) McDaniel, R.; Thamchaipenet, A.; Gustafsson, C.; Fu, H.; Betlach, M.; Betlach, M.; Ashley, G. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 1846.
(22) Tang, L.; Fu, H.; McDaniel, R. *Chem. Biol.* **2000**, *7*, 77.
(23) Reeves, C. D.; Murli, S.; Ashley, G. W.; Piagentini, M.; Hutchinson, C. R.; McDaniel, R. *Biochemistry* **2001**, *40*, 15464.
(24) Ruan, X.; Pereda, A.; Stassi, D. L.; Zeidner, D.; Summers, R. G.; Jackson, M.; Shivakumar, A.; Kakavas, S.; Staver, M. J.; Donadio, S.; Katz, L. *J. Bacteriol.* **1997**, *179*, 6416.
(25) Liu, L.; Thamchaipenet, A.; Fu, H.; Betlach, M.; Ashley, G. *J. Am. Chem. Soc.* **1997**, *119*, 10553.
(26) Stassi, D. L.; Kakavas, S. J.; Reynolds, K. A.; Gunawardana, G.; Swanson, S.; Zeidner, D.; Jackson, M.; Liu, H.; Buko, A.; Katz, L. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 7305.
(27) Xue, Q.; Ashley, G.; Hutchinson, C. R.; Santi, D. V. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 11740.
(28) Marsden, A. F. A.; Wilkinson, B.; Cortés, J.; Dunster, N. J.; Staunton, J.; Leadlay, P. F. *Science* **1998**, *279*, 199.
(29) McDaniel, R.; Kao, C. M.; Fu, H.; Hevezi, P.; Gustafsson, C.; Betlach, M.; Ashley, G.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1997**, *119*, 4309.
(30) Kao, C. M.; McPherson, M.; McDaniel, R. N.; Fu, H.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1997**, *119*, 11339.
(31) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1995**, *117*, 9105.
(32) Jacobsen, J. R.; Hutchinson, C. R.; Cane, D. E.; Khosla, C. *Science* **1997**, *277*, 367.
(33) Gokhale, R. S.; Hunziker, D.; Cane, D. E.; Khosla, C. *Chem. Biol.* **1999**, *6*, 117.
(34) Aggarwal, R.; Caffrey, P.; Leadlay, P. F.; Smith, C. J.; Staunton, J. *J. Chem. Soc., Chem. Commun.* **1995**, 1519.
(35) Boddy, C. N.; Schneider, T. L.; Hotta, K.; Walsh, C. T.; Khosla, C. *J. Am. Chem. Soc.* **2003**, *125*, 3428.
(36) Aldrich, C. C.; Venkatraman, L.; Sherman, D. H.; Fecik, R. A. *J. Am. Chem. Soc.* **2005**, *127*, 8910.

(37) Wang, M.; Boddy, C. N. *Biochemistry* **2008**, *47*, 11793.
(38) Heberlig, G. W.; Wirz, M.; Wang, M.; Boddy, C. N. *Org. Lett.* **2014**, *16*, 5858.
(39) Hari, T. P.; Labana, P.; Boileau, M.; Boddy, C. N. *ChemBioChem* **2014**, *15*, 2656.
(40) Kudo, F.; Kitayama, T.; Kakinuma, K.; Eguchi, T. *Tetrahedron Lett.* **2006**, *47*, 1529.
(41) Wang, M.; Zhou, H.; Wirz, M.; Tang, Y.; Boddy, C. N. *Biochemistry* **2009**, *48*, 6288.
(42) Pinto, A.; Wang, M.; Horsman, M.; Boddy, C. N. *Org. Lett.* **2012**, *14*, 2278.
(43) Schaffer, J. E.; Reck, M. R.; Prasad, N. K.; Wencewicz, T. A. *Nat. Chem. Biol.* **2017**, *13*, 737.
(44) Chen, X.-P.; Shi, T.; Wang, X.-L.; Wang, J.; Chen, Q.; Bai, L.; Zhao, Y.-L. *ACS Catal.* **2016**, *6*, 4369.
(45) Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R.; Hakansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. *Nature* **2014**, *510*, 512.
(46) Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R.; Hakansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. *Nature* **2014**, *510*, 560.
(47) Weissman, K. J. *Nat. Prod. Rep.* **2016**, *33*, 203.



DOI: 10.1021/jacs.7b06436
*J. Am. Chem. Soc.* 2017, 139, 13456−13465