



Published in final edited form as:

J Am Chem Soc. 2022 October 26; 144(42): 19326–19336. doi:10.1021/jacs.2c06631.

Data Science-Driven Analysis of Substrate-Permissive Diketopiperazine Reverse Prenyltransferase NotF: Applications in Protein Engineering and Cascade Biocatalytic Synthesis of (–)-Eurotiumin A

Samantha P. Kelly^{‡,§,†}, Vikram V. Shende^{‡,§,†,||}, Autumn R. Flynn[#], Qingyun Dan^{¥,‡,†}, Ying Ye^{‡,||}, Janet L. Smith^{¥,‡}, Sachiko Tsukamoto^Δ, Matthew S. Sigman[#], David H. Sherman^{‡,†,Θ}

[‡]Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109, USA.

[§]Program in Chemical Biology, University of Michigan, Ann Arbor, MI 48109, USA.

[¥]Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA.

[†]Department of Medicinal Chemistry, University of Michigan, Ann Arbor, MI 48109, USA.

^ΘDepartment of Microbiology and Immunology, University of Michigan, Ann Arbor, MI 48109, USA.

[#]Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA.

^ΔGraduate School of Pharmaceutical Sciences, Kumamoto University, 5-1 Oe-honmachi, Kumamoto 862-0973, Japan.

[‡]These authors contributed equally: Samantha P. Kelly, Vikram V. Shende

Abstract

Prenyltransfer is an early-stage carbon–hydrogen bond (C–H) functionalization prevalent in the biosynthesis of a diverse array of biologically active bacterial, fungal, plant, and metazoan diketopiperazine (DKP) alkaloids. Towards the development of a unified strategy for biocatalytic construction of prenylated DKP indole alkaloids, we sought to identify and characterize a substrate-permissive C2 reverse prenyltransferase (PT). As the first tailoring event within the biosynthesis of cytotoxic notoamide metabolites, PT NotF catalyzes C2 reverse prenyltransfer of brevianamide F. Solving a crystal structure of NotF (in complex with native substrate and prenyl donor mimic dimethylallyl S-thiolodiphosphate (DMSPP)) revealed a large, solvent

Corresponding Authors: Matthew S. Sigman – Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States; matt.sigman@utah.edu; David H. Sherman – Life Sciences Institute, Department of Medicinal Chemistry, Department of Chemistry, and Department of Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109, United States; davidhs@umich.edu.

^{||}Vikram V. Shende – Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, 92037, United States.

[†]Qingyun Dan – Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, California 94608, United States.

^{||}Ying Ye – School of Pharmacy, Huazhong University of Science and Technology, Wuhan 430030, China.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS website.

Full experimental and computational details, NMR spectra, tables, and figures (PDF)

Substrate descriptors and induced fit docking materials (.zip)

The authors declare no competing financial interest.

exposed active site, intimating NotF may possess a significantly broad substrate scope. To assess the substrate selectivity of NotF we synthesized a panel of 30 sterically and electronically differentiated tryptophanyl DKPs, the majority of which were selectively prenylated by NotF in synthetically useful conversions (2 to >99%). Quantitative representation of this substrate library and development of a descriptive statistical model provided insight into the molecular origins of NotF's substrate promiscuity. This approach enabled the identification of key substrate descriptors (electrophilicity, size, and flexibility) that govern the rate of NotF-catalyzed prenyltransfer, and the development of an "induced fit docking (IFD)-guided" engineering strategy improved turnover of our largest substrates. We further demonstrated the utility of PT NotF in tandem with oxidative cyclization using flavin monooxygenase, BvnB. This one-pot, *in vitro* biocatalytic cascade enabled the first chemoenzymatic synthesis of the marine fungal natural product, (–)-eurotiumin A, in three steps and 60% overall yield.

INTRODUCTION

The incredible structural diversity and potent biological activity of C2 reverse prenylated diketopiperazine (DKP) natural products (NPs) has inspired the development of selective chemical methods for reverse prenyltransfer as an entry point for synthesis of prenylated NPs and their analogues.^{1–8} State-of-the-art synthetic methods for C2 reverse prenylation of tryptophan substituents are typically implemented at an early stage, require multiple protecting groups, and fresh preparation of a prenylating reagent, such as prenyl-9-BBN.⁶ The development of a general, biocatalytic method for prenyltransfer, i.e., a site-selective prenyltransferase (PT) with broad substrate scope, would provide an efficient method for installation of prenyl functionalities that circumvents the necessity of protecting groups and enables late-stage diversification of complex substrates with additional tailoring enzymes.

Widespread in primary and secondary metabolism of marine and terrestrial organisms, PTs are responsible for key transformations within biologically active NP biosynthesis, catalyzing the transfer of five carbon building blocks such as dimethylallyl (DMAPP), geranyl (GPP), farnesyl (FPP), and other pyrophosphates onto a variety of substrates.^{9,10} Soluble aromatic PTs that share a conserved fold of five consecutive $\alpha\beta\beta\alpha$ repeats comprise the ABBA-type superfamily of PTs.¹¹ The dimethylallyl tryptophan synthase (DMATS)-type subfamily of ABBA PTs are often encoded in alkaloid biosynthetic gene clusters and catalyze both normal and reverse prenylation of indole and tryptophan derivatives,^{11,12} with varying site-selectivities,^{9,13–15} and often relaxed substrate specificity.^{9,13,15–18} PTs frequently act early in the assembly of biologically active fungal alkaloids, preceding a variety of oxidative tailoring events, including oxygenation (such as (+)-okaramine B **3**),^{19,20} spirocyclization ((+)-brevianamide A **1**),^{21–23} halogenation ((+)-malbrancheamides),²⁴ cycloaddition ((–)-paraherquamides),²⁵ as well as many others (Figure 1A and S1).^{26,27}

We previously identified and characterized DMATS-type PT, NotF, which catalyzes C2 reverse prenylation of native substrate brevianamide F (cyclo(*L*-Trp-*L*-Pro) **5**, Figure 1B) and found it to possess strict substrate selectivity in a limited scope analysis.²⁸ However, recent work using stereoisomeric cyclo(*L/D*-Trp-*L/D*-Pro) and cyclo(*L/D*-Trp-*L/D*-Ala)

substrates suggested that NotF may possess broader flexibility in comparison to homologous DMATS-type PT, BrePT (83% sequence ID, Figure S3).^{16–18,28} To gain molecular insight into how NotF may be able to accommodate a broad range of substrates, and toward our goal of developing a site-selective, yet substrate permissive C2 reverse PT biocatalyst, we sought to further evaluate NotF's substrate specificity using structural, biochemical, and computational approaches.

RESULTS AND DISCUSSION

Structural characterization of NotF.

We solved a ligand-free NotF structure at a 3.2 Å resolution, using molecular replacement with AnaPT (PDB ID: 4LD7) as the search model.²⁹ This NotF structure is consistent with the DMATS-type PTs within the ABBA PT superfamily, and possesses a classic PT fold with 10 inner antiparallel β strands surrounded by outer α helices (Figure 2A).^{9,30–32} A metal ion is absent in the active site due to the lack of a (N/D)DXXD Mg²⁺ coordination motif, also consistent with DMATS-type PTs. In the substrate-free state of NotF, the central active site is open and solvent-accessible. A 3.0 Å NotF structure in complex with its natural substrate brevianamide F (**5**) and a non-hydrolysable DMAPP analog, dimethylallyl S-thiolodiphosphate (DMSPP), was determined by rescreening NotF using co-crystallization. In both the substrate-free and substrate-bound structures, a flexible loop with 21 residues (G328 – Q348) lacks density. However, in the NotF-substrate complex structure, three of the 24 non-crystallographically related copies in the unit cell showed unambiguous densities for both ligands (Figure 2B). This enabled our analysis of substrate binding at a molecular level.

The brevianamide F-bound NotF structure bears only subtle structural differences compared to the ligand-free form, with an RMS value of 0.46 Å for 2565 atoms (Figure 2A). In agreement with other reported fungal indole PT structures,³¹ the diphosphate group of DMSPP is surrounded by four conserved positively charged residues (R122, K212, K282 and K369) forming salt bridges with the negatively charged phosphates (Figure 2C). Furthermore, conserved tyrosines (Y214, Y284, Y371 and Y440) also hydrogen bond with the DMSPP ligand, facilitating diphosphate dissociation, shielding the active site from solvent, and protecting the putative reactive carbocation intermediate.³³ The prenyl acceptor brevianamide F (**5**) resides in a hydrophobic pocket, with the indole moiety tightly packed between a loop region (R97 – S103) and the planar dimethylallyl group (Figure 2D, E). The conserved E108 forms a hydrogen bond with spatially constrained substrate indole N1,³¹ with the C2 acceptor facing the prenyl donor (Figure S8). Although DMSPP C1 is slightly closer to indole C3 (3.5 Å compared to 4.2 Å to indole C2), previous studies have shown that indole C2 is more favorable for Friedel-Crafts alkylation.^{34,35} As multiple brevianamide F PTs have been previously characterized with varying prenylation patterns (normal vs. reverse) and site-selectivities (FtmPT1/FtmB, BrePT/ BvnC, OkaC, etc.), we compared the holo NotF structure with the crystal structure of DMSPP- and **5**-bound FtmPT1 (C2 normal brevianamide F PT PDB: 3O2K) in an attempt to glean structural insight towards site-selectivity and prenylation pattern. While NotF positions the substrate site of prenyltransfer (indole C2) closer in proximity to prenyl C3 (4.2 Å compared to 4.5 Å

to C1), consistent with reverse prenyltransfer, FtmPT1 also binds DMSPP and substrate with indole C2 in closer proximity to prenyl C3 (3.8 Å), rather than prenyl C1 (5.2 Å) (Figures S9 and S10). Thus, we reasoned that rigid structural insight cannot be used to predict reaction outcomes, and perhaps a more dynamic analysis method (molecular dynamics, dynamic docking) might provide improved insight and predictive capabilities.

Remarkably, no hydrogen bond or charged-charged interaction involving the DKP core or proline side chain are present in the complex structure. The NotF active site pocket leaves substantial unoccupied space surrounding the fused DKP-proline bicycle, and the binding site remains mostly solvent-exposed in the ligand-free structure. Based on these structures, we reasoned that NotF may be able to accommodate substrates with amino acid substitutions at the proline side chain without affecting the indole interaction. Thus, we set out to evaluate the scope of NotF's substrate accommodation for potential biocatalytic production of NPs and their derivatives.

Substrate scope of NotF.

Using a series of sterically and electronically differentiated tryptophanyl-DKP substrates (**5**, **8** – **36**), we interrogated the substrate promiscuity of NotF. This PT was able to convert all 30 DKP substrates, often displaying complete consumption of substrate (Figure 3 Tables S2–S60), to generate C2 reverse prenylated DKP products including NPs talathermophilin E (**S1** from **8**),³⁶ preechinulin (**37** from **9**),³⁷ fellutanine B (**S2** from **35**),³⁸ and deoxybrevianamide E (**6** from **5**) (Figure S1).³⁹ DKP substrates containing small, nonpolar side chains (**5**, **9** – **12**, **15**, **18**, **23**) were converted with high efficiency, reflected by $\geq 70\%$ conversion in all cases. NotF also tolerates bulky (**21** – **36**) and polar (**14**, **15**, **20**, **21**, **28**) side chains with conversions ranging from 2 to 99%. The high substrate promiscuity of NotF offers an efficient biocatalytic method for selective chemoenzymatic synthesis of diverse C2 reverse prenylated DKPs. Furthermore, high conversion of alkyne-containing substrate **18** (76%) enables potential late-stage diversification via annulation reactions such as Larock indole synthesis⁴⁰ or biorthogonal modification through one-pot diversification via click chemistry.^{41,42}

Previous reports investigating PT substrate scope revealed that use of non-native substrates can induce a change of both chemo- and site-selectivity with respect to both the indole core and DMAPP.^{14,15,43} To confirm NotF's site-selectivity, we performed preparative scale biocatalytic prenyltransfer on select substrates varying in steric bulk including cyclo(*L*-Trp-*L*-Ala) (**9**), cyclo(*L*-Trp-*L*-Val) (**13**), cyclo(*L*-Trp-*L*-Pip) (**23**), and tetrahydrocarbazole-containing DKP **36**. Isolation and NMR structural characterization of all four reactions confirmed C2 reverse prenyltransfer with isolated yields (54–98%) consistent with those from analytical scale reactions (Figures 3, S335 – S356).

Overall, the results demonstrate NotF's utility as a substrate-permissive catalyst through selective C2 reverse prenylation of 30 tryptophanyl-DKPs representative of fungal indole alkaloid NP structures, suggesting NotF's exceptional potential as a biocatalyst in the assembly of this class of NPs. Aside from the native NotF product deoxybrevianamide E (**6**) serving as a direct precursor to various fungal bicyclo[2.2.2]diazaoctane indole

alkaloids, preechinulin (**37**) is a likely intermediate in the biosynthesis of variegolortide A (**4**), aspechinulin B (**2**), and (–)-uncarilin B (**S8**), among many others (Figures 1 and S1).^{44–46} Similarly, C2 reverse prenylation of cyclo(*L*-Trp-*L*-Val) (**13**) results in a likely precursor to cristatumin F (**S5**),⁴⁷ and cyclo(*L*-Trp-*L*-Pip) (**23**) has been utilized as a synthetic intermediate toward the marcfortines and related indole alkaloids (**S10**, Figure S1).^{2,48,49} We hypothesized the mono C2 reverse prenylation of substrate cyclo(*L*-Trp-*L*-Trp) (**35**) may give rise to an on pathway intermediate to okaramine B (**3**, Figure 1),¹⁹ however, we observed NotF efficiently catalyzes the formation of a di-prenylated product by LCMS. Full characterization of the major product by NMR revealed NotF iteratively reverse prenylates the C2 position of both indole heterocycles, resulting in highly selective formation of symmetrical product **S98** (Figures S362-S366), a known synthetic intermediate towards gypsetin (**44**).⁵⁰ The ability to selectively C2 reverse prenylate these substrates therefore represents a solution to a key transformation within these biosynthetic systems.

Multivariate Linear Regression Modeling.

While NotF converted all 30 of these substrates, we were unable to identify clear trends between steric or electronic factors of substrates and the observed magnitude of catalytic turnover. In the absence of distinct structural factors (from both enzyme and substrates) that may govern substrate specificity of NotF, our access to a wide variety of structurally differentiated substrates and ability to efficiently acquire kinetic data provided a unique opportunity to gain insight into the molecular basis for substrate prenylation. We hypothesized that substrate structure and conformation may be related to kinetics of binding and dissociation and reasoned that a substrate-profiling approach, aided by computational tools, might clarify the basis for substrate turnover. Thus, we opted to gain insight into NotF's substrate permissiveness by employing statistical modeling techniques, an approach traditionally used to parameterize small molecule catalysts.^{51,52} Three-dimensional unrestrained conformational ensembles of each substrate structure were collected, and quantitative chemical descriptors were acquired from density functional theory (DFT) calculations of the substrates at the B3LYP level of theory. These features included electronic (e.g., natural bond orbital (NBO) charges) and steric descriptors (e.g., Sterimol values).⁵³ We calculated a wide range of descriptors to provide an unbiased set of parameters from which a forward-stepwise linear regression algorithm could determine the best-fit descriptors (details of all descriptors calculated can be found in the Supporting Information). Regression of this library of descriptors to the log of empirically determined pseudo-first order initial rates (Figure 3), enabled the building of a statistical model (Figure 4A).

Our best fit model is comprised of three terms, each with a negative coefficient, indicating that as magnitude of the descriptor values rise, the relative rate of reaction is reduced (Figure 4A). This model presents robust values for statistical evaluation (where the Training R^2 is high, and the Test R^2 and leave-one-out cross-validation Q^2 agree with both the Training and Test R^2 values, Figure 4A). Substrates containing unsaturated moieties, many of which are aromatic, possess high $\omega_{\min}$ values (defined as the substrate electrophilicity index) and are prenylated at slower rates than those that are saturated (substrates with the 10 largest $\omega_{\min}$ values are **18**, **21**, **24**, **29 – 31**, and **33 – 36**). The active site of NotF consists of several

aromatic residues (F216, Y266, F268, W424) that surround the binding pocket of native substrate **5** (Figure 4B). We reasoned substrates with an aromatic or unsaturated group (and high $\omega_{\min}$ values) in the variable side chain likely form noncovalent π -interactions with the aromatic residues, which may restrict product dissociation, resulting in decreased turnover (Figures S126 – S128). The natural bond orbital term (relating to substrate flexibility, Figure 4C) also possesses an inverse relationship with the rate, indicating substrates with highly dynamic NBO charges at this position also convert more slowly. This is exemplified in substrate pairs **22** and **23**, **29** and **30**, and **35** and **36**, which illustrate how restricting degrees of rotational freedom through tethering of the DKP side chain (as in **23**, **30**, and **36**) is reflected in increased rate of these substrates relative to their unconstrained pair. The third term, Sterimol B5_{max}, indicates that a large maximum width of each substrate diminishes the rate of the reaction (Figure 4C). Analysis of the NotF binding site bound to the native substrate **5** shows this is likely due to repulsive steric interactions within the active site in the case of bulkier substrates (such as **32**, **35** and **36**, Figure 3). These interactions may demand major active site repositioning for substrate accommodation, leading to slower substrate binding and ultimately decreased turnover. While the model describes the differences observed in initial rates, we found that in many cases we cannot correlate this explanation to the total conversion achieved in these reactions. We hypothesize that additional factors, such as product inhibition, may slow the rate once the substrates no longer achieve saturating concentration, effectively decreasing overall conversion.

By utilizing statistical modeling to correlate substrate structures to initial reaction rates, we efficiently uncovered key substrate features that lead to differing rates of prenylation. Notably, this approach bypasses computationally intensive techniques that do not inherently reveal these features, such as molecular dynamics (MD) simulations. However, we sought to corroborate our model explicitly, and turned to induced-fit docking (IFD),^{54–56} a molecular docking protocol where both substrate and enzyme are flexible, to further investigate our assertions. In the case of substrates with an aromatic DKP moiety (such as **29**), IFD shows π -interactions between the phenyl ring and Y266 (Figure S129). The binding pose of the non-aromatic counterpart, **32**, is similar to that of **29**, however interaction with Y266 is lacking. This finding substantiates the $\omega_{\min}$ term in our statistical model, indicating π -interactions between arene-containing DKPs and the binding pocket of NotF.

When **5** is bound to NotF, the residues in the pocket closely surround the substrate. But when **36**, a substrate with a high B5_{max} value, is docked via IFD we observe major repositioning of nearby residues Y266, L193, and P352 to accommodate the steric bulk of **36** (Figure 5A, B). The RMSD of active-site residues (within 5 Å of ligand) when **36** is docked is 1.56 Å, relative to the docking pose of **5**. When **12**, a substrate with small B5_{max} value, is docked this repositioning does not occur, and the RMSD for active site residues is 1.06 Å (Table S63). These IFD results support the hypothesis that large B5_{max} values lead to major NotF residue repositioning upon substrate binding.

IFD-guided engineering NotF for C2 reverse prenylation of large, aromatic substrates.

Guided by our observations from IFD, we sought to engineer a NotF variant with enhanced rate with larger substrates. IFD of substrate **36** revealed active site residues (Y266, L193,

P352) adopted alternative conformations to accommodate increased substrate volume, and aromatic residues (F216, Y266), participating in π - π stacking (Figure 5A, B, S125). To evaluate the impact of steric and electronic interactions between these residues and substrates, they were all selected for substitution with alanine via site-directed mutagenesis. As Y266 participates in both π - π stacking and steric interactions, we generated a series of mutations of this residue to valine, threonine, and proline to assess the effect of various residue sizes at this position. Mutagenesis of L193 to alanine failed to generate a soluble NotF variant and as a surrogate, we generated a L193S mutant based on homology to structurally characterized PT, AtaPT (PDB: 5KCY).⁵⁷ While NotF variants evaluated did not improve the turnover rate for medium sized aromatic substrates **24** nor **34**, we observed improved rates for the larger indolic substrates, carbazole **36** and cyclo(*L*-Trp-*L*-Trp) **35** (Figure 5C). Variants NotF_{Y266A}, NotF_{Y266V}, and NotF_{P352A} catalyzed prenyltransfer to **36** at rates 1.1, 1.3 and 1.03 times faster than WT respectively and NotF_{Y266A}, NotF_{Y266V}, NotF_{Y266P}, and NotF_{L193S} prenylated **35** at rates 1.5, 1.7, 1.3, and 1.6 times greater than WT, respectively. However, while all variants evaluated retained their site- and chemoselectivity in the prenylation of **36**, close inspection of product profiles for **35** revealed that the observed increased reaction rate was accompanied by a change in selectivity (Figures S187, S189, S191, S194, S196, S198, S200, S203, and S205).

These results demonstrate that decreasing steric bulk at positions L193, Y266, and P352 improves the reaction rate of NotF with large, aromatic substrates **35** and **36**. We attribute this result to the size and electrophilicity of Y266, as it is the only residue found to participate in π - π stacking (Figure S129) as well as major active site repositioning upon IFD with **36** (Figure 5). While reducing the size of this residue to alanine showed substantial improvement of the rate of reaction with these substrates, more modest reduction of size to valine resulted in a greatly enhanced reaction rate, indicating this residue's participation in substrate binding. The increased rate of NotF_{L193S} and NotF_{P352A} offers experimental evidence to the substantial repositioning of these residues observed in IFD **36** in the NotF crystal structure. Through analysis of reaction profiles, we observed differences in product distribution between substrate **35** and NotF variants. These mutagenesis experiments corroborate our model and interpretation as we observe significant rate improvement with previously low turnover substrates. However, the unexpected erosion of selectivity for substrate **35** demonstrates the multifaceted challenge in engineering improved prenyltransferase variants with broad substrate scope while conserving site selectivity for both substrates in this bimolecular reaction.

One-pot biocatalytic cascade to (–)-eurotiumin A.

The fused pyrroloindoline scaffold is prevalent in a variety of prenylated DKP NPs with potent biological activity including anticancer ardeemin (**43**),^{58,59} antimicrobials roquefortine C (**45**)⁶⁰ and drimentine A (**46**),⁶¹ and the structurally analogous diketomorpholine NP shornephine A (**42**, Figure 5).⁶² Given NotF's broad substrate specificity and good to high turnover of their DKP precursors, we aimed to develop a biocatalytic cascade approach towards prenylated pyrroloindoline DKPs.

We recently identified BvnB, a stereoselective flavin monooxygenase (FMO) that catalyzes indole 2,3 epoxidation and subsequent cyclization to the 3-hydroxypyrroloindoline product brevianamide E (**47**) (Figure 6, Figure S230).²³ Based on these properties, its use in tandem with NotF for facile assembly of prenylated pyrroloindoline NPs in a one-pot reaction was explored.

After validating BvnB's ability to catalyze *in vitro* epoxidation and cyclization of deoxybrevianamide E (**6**) to brevianamide E (**47**, Figure S231), we assessed potential compatibility between NotF and BvnB for the development of a biocatalytic cascade to generate prenylated hydroxypyrroloindoline NPs and corresponding analogues.⁶³ Gratifyingly, in a one-pot cascade with both biocatalysts, we observed high conversion of native substrate **5** to **47** (Figure S232). In control reactions lacking NotF, BvnB failed to convert non-prenylated substrate **5**, suggesting a requirement for molecular recognition of the C2 reverse prenyl substituent by BvnB for productive catalysis.

Given the selective substrate recognition observed with native substrates **5** and **6**, we envisioned employing the NotF-BvnB cascade in a stereoselective, biocatalytic synthesis of fungal 3-hydroxypyrroloindoline NP, eurotiumin A (**41**),⁶⁴ whose synthesis has not been previously reported. As in reactions with **5**, we observed selective substrate recognition by each biocatalyst with no conversion of cyclo(*L*-Trp-*L*-Ala) (**9**) by BvnB in the absence of NotF (Figure 7). Scale up of the NotF-BvnB cascade with **9** and structural characterization of final product confirmed the first biocatalytic, stereoselective synthesis of eurotiumin A (**41**, 60% isolated yield over three steps, with a single chromatographic purification), and no detectable stereoisomers or shunt products. This process enabled determination of the optical rotation of **41** as $[\alpha]_D^{24} -335.08$ hereon referred to as (–)-eurotiumin A. Thus, the “one-pot” use of NotF and BvnB provides an efficient route towards the hydroxypyrroloindoline scaffold present in NPs with a range of biological activities,⁶⁵ which are amenable to further synthetic diversification.⁶⁶

CONCLUSION

Herein, we describe the biochemical, structural, and statistical characterization of the C2 reverse prenyltransferase (PT), NotF, which revealed the molecular basis for broad flexibility toward tryptophanyl-DKP substrates. We previously reported NotF's high selectivity for brevianamide F (**5**) with no turnover of unnatural substrates cyclo(*L*-Trp-*L*-Trp) (**35**) and cyclo(*L*-Trp-*L*-Tyr) (**33**),²⁸ and we reason that improved methods for NotF expression and purification resulted in more stable and active protein sample. We now demonstrate effective C2 reverse prenylation of 30 DKP substrates, and provide a straightforward, protecting group free, biocatalytic strategy to overcome a synthetically challenging C–H functionalization in indole alkaloid biosynthesis. NotF is capable of directly generating several prenylated NPs from simple DKP precursors (accessible in two synthetic steps), including talathermophilin E (**S1**), preechinulin (**37**), deoxybrevianamide E (**6**), and fellutanine B (**S2**), and precursors to additional NP classes including the bicyclo[2.2.2]diazaoctane indole alkaloids **1** and **7**, polyprenylated cristatumin F (**S5**), oxidatively cyclized okaramine B (**3**), as well as the cycloechinulins and variecolortide A (**4**, Figures 1 and S1).

Data science-driven approaches can provide powerful methods for enzyme characterization that supplement our current molecular tools for better understanding of biocatalyst substrate scope. Statistical modeling of substrate features can be used to rationalize performance in biocatalytic reactions and enable engineering of improved biocatalysts with enhanced activity and broad applicability. In our study, the implementation of statistical modeling through multivariate linear regression (MLR) of steric and electronic molecular descriptors enabled our identification of substrate features that influence the rate of prenyltransfer by NotF.

Our model found that substrates with high electrophilicity, flexibility, and steric bulk are turned over at slower rates, a result unlikely to be gleaned from MD simulations alone. Using computationally inexpensive IFD, we found that larger and more flexible substrates require rearrangement of active site residues that likely attenuate turnover to product. While previous studies have used statistical analysis to understand enzyme catalysis and guide engineering efforts,^{67–70} this work showcases the integration of substrate profiling using MLR with IFD as a computationally cost-effective and simple interrogation of substrate scope. To accompany random, site-saturation,⁷¹ structure-guided,²⁴ and MD-guided^{72,73,74} mutagenesis strategies that our group has previously relied upon, we have harnessed the predictive power of machine learning and statistical modeling to further expand our protein engineering toolbox via statistical modeling and “IFD-guided” engineering, and aim to further develop NotF variants and other NP biosynthetic enzymes with expanded substrate scope (Figure 8). We specifically envision both continuing to engineer NotF_{Y266V} for turnover of highly flexible, bulky substrates, in addition to engineering NotF to accommodate more bulk surrounding the indole substrate to access further functionalized biosynthetic intermediates and probe downstream tailoring events of numerous indole alkaloid biosynthetic pathways.^{1,26,36,47,62,75,76}

Finally, we developed a one-pot methodology using PT NotF and FMO BvnB that provides an efficient route towards the 3-hydroxypyrroloindoline NP scaffold from simple and readily accessible diketopiperazine building blocks. This biocatalytic cascade was applied towards the first synthesis of antioxidant NP (–)-eurotiumin A as a single stereoisomer, in 60% yield over three steps (two-step synthesis of cyclo(*L*-Trp-*L*-Ala), one step biocatalytic cascade) involving one final chromatographic purification. With demonstrated crosstalk between notoamide and brevianamide biosynthetic enzymes, we envision further diversification of prenylated DKP-containing NPs. In summary, we have employed structural and statistical modeling approaches to demonstrate the molecular basis for NotF’s broad substrate scope, and its application in the biocatalytic assembly of the 3-hydroxypyrroloindoline scaffold.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors would like to acknowledge Rajani Arora, UM LSI Multimedia and Social Media specialist, for her contribution through the submitted cover art and TOC figure, Dr. Sean Newmister for his contribution to crystallographic data, as well as Robert Hohlman for assistance in substrate synthesis. The authors also thank the

NSF CCI Center for Selective C–H Functionalization (CHE-1700982), NIH grant R35 GM118101, and the Hans W. Vahlteich Professorship (to D.H.S.), NIH grant R01 DK042303, and the Margaret J. Hunter Professorship (to J.L.S.). M.S.S. and A.R.F. thank NIH grant R35 GM136271 for partial support. Computational resources were provided by the Center for High Performance Computing (CHPC) at the University of Utah. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. GM/CA@APS (beamlines 23ID-B and 23ID-D) has been funded by the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006, P30GM138396). The Eiger 16M detector at GM/CA-XSD was funded by NIH grant S10 OD012289.

REFERENCES

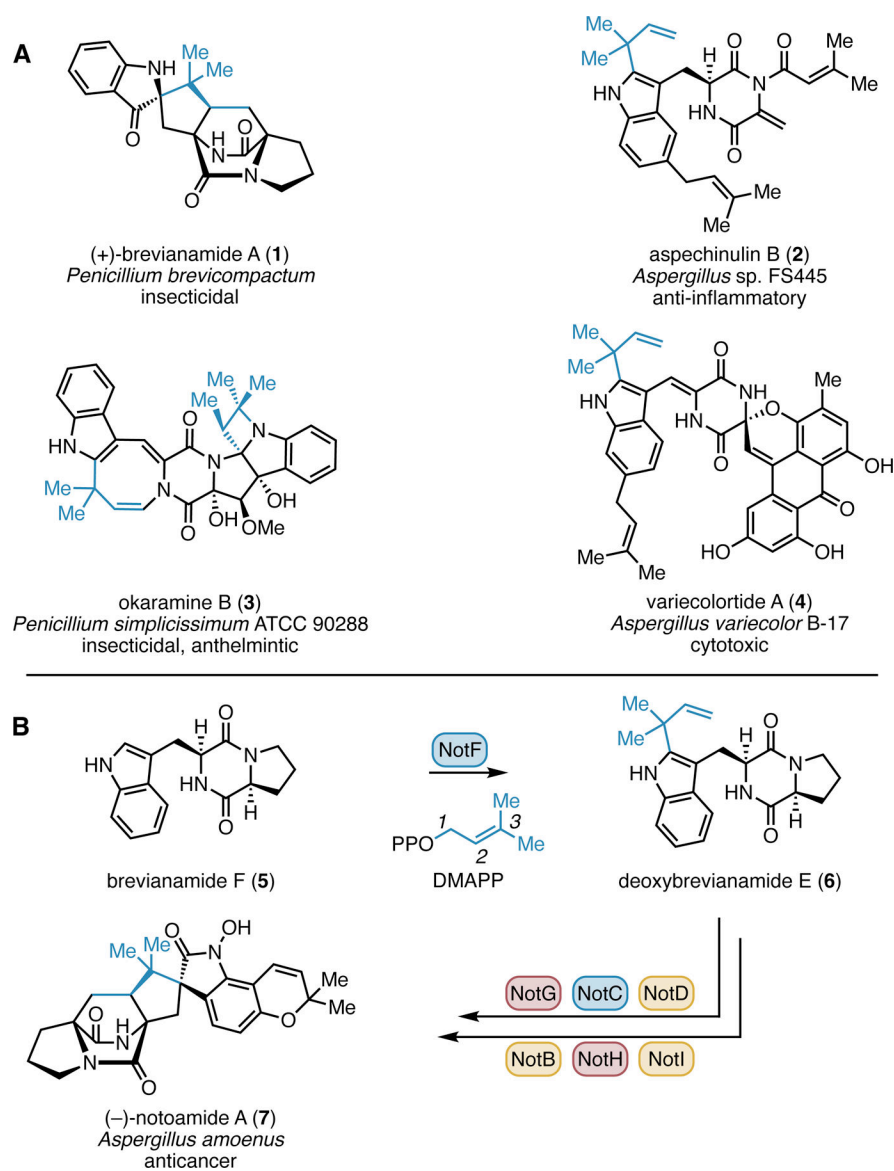
- (1). Grubbs AW; Artman GD; Williams RM Concise Syntheses of the 1,7-Dihydropyrano[2,3-g]Indole Ring System of the Stephacidins, Aspergamides and Norgeamides. *Tetrahedron Lett* 2005, 46 (52), 9013–9016. 10.1016/j.tetlet.2005.10.112.
- (2). Greshock TJ; Grubbs AW; Williams RM Concise, Biomimetic Total Synthesis of d,l-Marcfortine C. *Tetrahedron* 2007, 63 (27), 6124–6130. 10.1016/j.tet.2007.03.016. [PubMed: 18596842]
- (3). Miller KA; Welch TR; Greshock TJ; Ding Y; Sherman DH; Williams RM Biomimetic Total Synthesis of Malbrancheamide and Malbrancheamide B. *J. Org. Chem* 2008, 73 (8), 3116–3119. 10.1021/jo800116y. [PubMed: 18345688]
- (4). Li S-M Applications of Dimethylallyltryptophan Synthases and Other Indole Prenyltransferases for Structural Modification of Natural Products. *Appl. Microbiol. Biotechnol* 2009, 84 (4), 631–639. 10.1007/s00253-009-2128-z. [PubMed: 19633837]
- (5). Li S-M Prenylated Indole Derivatives from Fungi: Structure Diversity, Biological Activities, Biosynthesis and Chemoenzymatic Synthesis. *Nat Prod Rep* 2010, 27 (1), 57–78. 10.1039/B909987P. [PubMed: 20024094]
- (6). Zhao L; May JP; Huang J; Perrin DM Stereoselective Synthesis of Brevianamide E. *Org. Lett* 2012, 14 (1), 90–93. 10.1021/ol202880y. [PubMed: 22126228]
- (7). Lund S; Hall R; Williams GJ An Artificial Pathway for Isoprenoid Biosynthesis Decoupled from Native Hemiterpene Metabolism. *ACS Synth. Biol* 2019, 8 (2), 232–238. 10.1021/acssynbio.8b00383. [PubMed: 30648856]
- (8). Yu H; Zong Y; Xu T Total Synthesis of (–)-Penicimutanin A and Related Congeners. *Chem. Sci* 2020, 11 (3), 656–660. 10.1039/C9SC05252F.
- (9). Winkelblech J; Fan A; Li S-M Prenyltransferases as Key Enzymes in Primary and Secondary Metabolism. *Appl. Microbiol. Biotechnol* 2015, 99 (18), 7379–7397. 10.1007/s00253-015-6811-y. [PubMed: 26216239]
- (10). Alhassan A; Abdullahi M; Uba A; Umar A Prenylation of Aromatic Secondary Metabolites: A New Frontier for Development of Novel Drugs. *Trop. J. Pharm. Res* 2014, 13 (2), 307. 10.4314/tjpr.v13i2.22.
- (11). Bonitz T; Alva V; Saleh O; Lupas AN; Heide L Evolutionary Relationships of Microbial Aromatic Prenyltransferases. *PLoS ONE* 2011, 6 (11), e27336. 10.1371/journal.pone.0027336. [PubMed: 22140437]
- (12). Tello M; Kuzuyama T; Heide L; Noel JP; Richard SB The ABBA Family of Aromatic Prenyltransferases: Broadening Natural Product Diversity. *Cell. Mol. Life Sci* 2008, 65 (10), 1459–1463. 10.1007/s00018-008-7579-3. [PubMed: 18322648]
- (13). Li S-M Evolution of Aromatic Prenyltransferases in the Biosynthesis of Indole Derivatives. *Phytochemistry* 2009, 70 (15–16), 1746–1757. 10.1016/j.phytochem.2009.03.019. [PubMed: 19398116]
- (14). Rudolf JD; Wang H; Poulter CD Multisite Prenylation of 4-Substituted Tryptophans by Dimethylallyltryptophan Synthase. *J. Am. Chem. Soc* 2013, 135 (5), 1895–1902. 10.1021/ja310734n. [PubMed: 23301871]
- (15). Mori T Enzymatic Studies on Aromatic Prenyltransferases. *J. Nat. Med* 2020, 74 (3), 501–512. 10.1007/s11418-020-01393-x. [PubMed: 32180104]
- (16). Yin S; Yu X; Wang Q; Liu X-Q; Li S-M Identification of a Brevianamide F Reverse Prenyltransferase BrePT from *Aspergillus Versicolor* with a Broad Substrate Specificity towards

- Tryptophan-Containing Cyclic Dipeptides. *Appl. Microbiol. Biotechnol* 2013, 97 (4), 1649–1660. 10.1007/s00253-012-4130-0. [PubMed: 22660767]
- (17). Schuller JM; Zocher G; Liebhold M; Xie X; Stahl M; Li S-M; Stehle T Structure and Catalytic Mechanism of a Cyclic Dipeptide Prenyltransferase with Broad Substrate Promiscuity. *J. Mol. Biol* 2012, 422 (1), 87–99. 10.1016/j.jmb.2012.05.033. [PubMed: 22683356]
- (18). Yang K; Li S-M; Liu X; Fan A Reinvestigation of the Substrate Specificity of a Reverse Prenyltransferase NotF from *Aspergillus* sp. MF297–2. *Arch. Microbiol* 2020, 202 (6), 1419–1424. 10.1007/s00203-020-01854-7. [PubMed: 32185409]
- (19). Lai C-Y; Lo I-W; Hewage RT; Chen Y-C; Chen C-T; Lee C-F; Lin S; Tang M-C; Lin H-C Biosynthesis of Complex Indole Alkaloids: Elucidation of the Concise Pathway of Okaramines. *Angew. Chem. Int. Ed* 2017, 56 (32), 9478–9482. 10.1002/anie.201705501.
- (20). Kato N; Furutani S; Otake J; Noguchi A; Kinugasa K; Kai K; Hayashi H; Ihara M; Takahashi S; Matsuda K; Osada H Biosynthesis and Structure–Activity Relationship Studies of Okaramines That Target Insect Glutamate-Gated Chloride Channels. *ACS Chem. Biol* 2018, 13 (3), 561–566. 10.1021/acscchembio.7b00878. [PubMed: 29384650]
- (21). Fraley AE; Caddell Haatveit K; Ye Y; Kelly SP; Newmister SA; Yu F; Williams RM; Smith JL; Houk KN; Sherman DH Molecular Basis for Spirocycle Formation in the Paraherquamide Biosynthetic Pathway. *J. Am. Chem. Soc* 2020, 142 (5), 2244–2252. 10.1021/jacs.9b09070. [PubMed: 31904957]
- (22). Fraley AE; Tran HT; Kelly SP; Newmister SA; Tripathi A; Kato H; Tsukamoto S; Du L; Li S; Williams RM; Sherman DH Flavin-Dependent Monooxygenases NotI and NotI' Mediate Spiro-Oxindole Formation in Biosynthesis of the Notoamides. *ChemBioChem* 2020, 21 (17), 2449–2454. 10.1002/cbic.202000004. [PubMed: 32246875]
- (23). Ye Y; Du L; Zhang X; Newmister SA; McCauley M; Alegre-Requena JV; Zhang W; Mu S; Minami A; Fraley AE; Adrover-Castellano ML; Carney NA; Shende VV; Qi F; Oikawa H; Kato H; Tsukamoto S; Paton RS; Williams RM; Sherman DH; Li S Fungal-Derived Brevianamide Assembly by a Stereoselective Semipinacolase. *Nat. Catal* 2020, 3 (6), 497–506. 10.1038/s41929-020-0454-9. [PubMed: 32923978]
- (24). Fraley AE; Garcia-Borràs M; Tripathi A; Khare D; Mercado-Marin EV; Tran H; Dan Q; Webb GP; Watts KR; Crews P; Sarpong R; Williams RM; Smith JL; Houk KN; Sherman DH Function and Structure of Mala/MalA', Iterative Halogenases for Late-Stage C–H Functionalization of Indole Alkaloids. *J. Am. Chem. Soc* 2017, 139 (34), 12060–12068. 10.1021/jacs.7b06773. [PubMed: 28777910]
- (25). Dan Q; Newmister SA; Klas KR; Fraley AE; McAfoos TJ; Somoza AD; Sunderhaus JD; Ye Y; Shende VV; Yu F; Sanders JN; Brown WC; Zhao L; Paton RS; Houk KN; Smith JL; Sherman DH; Williams RM Fungal Indole Alkaloid Biogenesis through Evolution of a Bifunctional Reductase/Diels–Alderase. *Nat. Chem* 2019, 11 (11), 972–980. 10.1038/s41557-019-0326-6. [PubMed: 31548667]
- (26). Li S; Srinivasan K; Tran H; Yu F; Finefield JM; Sunderhaus JD; McAfoos TJ; Tsukamoto S; Williams RM; Sherman DH Comparative Analysis of the Biosynthetic Systems for Fungal Bicyclo[2.2.2]Diazaoctane Indole Alkaloids: The (+)/(–)-Notoamide, Paraherquamide and Malbrancheamide Pathways. *MedChemComm* 2012, 3 (8), 987. 10.1039/c2md20029e. [PubMed: 23213353]
- (27). Fraley AE; Sherman DH Enzyme Evolution in Fungal Indole Alkaloid Biosynthesis. *FEBS J* 2020, 287 (7), 1381–1402. 10.1111/febs.15270. [PubMed: 32118354]
- (28). Ding Y; Wet J. R. de; Cavalcoli J; Li S; Greshock TJ; Miller KA; Finefield JM; Sunderhaus JD; McAfoos TJ; Tsukamoto S; Williams RM; Sherman DH Genome-Based Characterization of Two Prenylation Steps in the Assembly of the Stephacidin and Notoamide Anticancer Agents in a Marine-Derived *Aspergillus* sp. *J. Am. Chem. Soc* 2010, 132 (36), 12733–12740. 10.1021/ja1049302. [PubMed: 20722388]
- (29). Yu X; Zocher G; Xie X; Liebhold M; Schütz S; Stehle T; Li S-M Catalytic Mechanism of Stereospecific Formation of Cis-Configured Prenylated Pyrroloindoline Diketopiperazines by Indole Prenyltransferases. *Chem. Biol* 2013, 20 (12), 1492–1501. 10.1016/j.chembiol.2013.10.007. [PubMed: 24239009]

- (30). Kuzuyama T; Noel JP; Richard SB Structural Basis for the Promiscuous Biosynthetic Prenylation of Aromatic Natural Products. *Nature* 2005, 435 (7044), 983–987. 10.1038/nature03668. [PubMed: 15959519]
- (31). Metzger U; Schall C; Zocher G; Unsold I; Stec E; Li S-M; Heide L; Stehle T The Structure of Dimethylallyl Tryptophan Synthase Reveals a Common Architecture of Aromatic Prenyltransferases in Fungi and Bacteria. *Proc. Natl. Acad. Sci* 2009, 106 (34), 14309–14314. 10.1073/pnas.0904897106. [PubMed: 19706516]
- (32). Jost M; Zocher G; Tarcz S; Matuschek M; Xie X; Li S-M; Stehle T Structure–Function Analysis of an Enzymatic Prenyl Transfer Reaction Identifies a Reaction Chamber with Modifiable Specificity. *J. Am. Chem. Soc* 2010, 132 (50), 17849–17858. 10.1021/ja106817c. [PubMed: 21105662]
- (33). Luk LYP; Tanner ME Mechanism of Dimethylallyltryptophan Synthase: Evidence for a Dimethylallyl Cation Intermediate in an Aromatic Prenyltransferase Reaction. *J. Am. Chem. Soc* 2009, 131 (39), 13932–13933. 10.1021/ja906485u. [PubMed: 19743851]
- (34). Lee S-L; Floss HG; Heinsteins P Purification and Properties of Dimethylallylpyrophosphate: Tryptophan Dimethylallyl Transferase, the First Enzyme of Ergot Alkaloid Biosynthesis in *Claviceps*. *Sp. SD 58. Arch. Biochem. Biophys* 1976, 177 (1), 84–94. 10.1016/0003-9861(76)90418-5. [PubMed: 999297]
- (35). Tanaka S; Shiomi S; Ishikawa H Bioinspired Indole Prenylation Reactions in Water. *J. Nat. Prod* 2017, 80 (8), 2371–2378. 10.1021/acs.jnatprod.7b00464. [PubMed: 28803474]
- (36). Guo J-P; Tan J-L; Wang Y-L; Wu H-Y; Zhang C-P; Niu X-M; Pan W-Z; Huang X-W; Zhang K-Q Isolation of Talatrophilins from the Thermophilic Fungus *Talaromyces thermophilus* YM3–4. *J. Nat. Prod* 2011, 74 (10), 2278–2281. 10.1021/np200365z. [PubMed: 21967034]
- (37). Allen CM Monoisoprenylated Cyclo-L-Alanyl-L-Tryptophanyl. Biosynthetic Precursor of Echinulin. *J. Am. Chem. Soc* 1973, 95 (7), 2386–2387. 10.1021/ja00788a059.
- (38). Kozlovsky AG; Vinokurova NG; Adanin VM; Burkhardt G; Dahse H-M; Gräfe U New Diketopiperazine Alkaloids from *Penicillium fellutanum*. *J. Nat. Prod* 2000, 63 (5), 698–700. 10.1021/np9903853. [PubMed: 10843594]
- (39). Steyn PS Austamide, a New Toxic Metabolite from *Aspergillus ustus*. *Tetrahedron Lett* 1971, 12 (36), 3331–3334. [https://doi.org/10.1016/S0040-4039\(00\)80000-0](https://doi.org/10.1016/S0040-4039(00)80000-0) [PubMed: 5411111]
- (40). Chuang KV; Kieffer ME; Reisman SE A Mild and General Larock Indolization Protocol for the Preparation of Unnatural Tryptophans. *Org. Lett* 2016, 18 (18), 4750–4753. 10.1021/acs.orglett.6b02477. [PubMed: 27598827]
- (41). Rostovtsev VV; Green LG; Fokin VV; Sharpless KB A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective “Ligation” of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed* 2002, 41 (14), 2596–2599.
- (42). Baskin JM; Bertozzi CR Bioorthogonal Click Chemistry: Covalent Labeling in Living Systems. *QSAR Comb. Sci* 2007, 26 (11–12), 1211–1219. 10.1002/qsar.200740086.
- (43). Rudolf JD; Poulter CD Tyrosine O-Prenyltransferase SirD Catalyzes S-, C-, and N-Prenylations on Tyrosine and Tryptophan Derivatives. *ACS Chem. Biol* 2013, 8 (12), 2707–2714. 10.1021/cb400691z. [PubMed: 24083562]
- (44). Wang W-L; Zhu T-J; Tao H-W; Lu Z-Y; Fang Y-C; Gu Q-Q; Zhu W-M Three Novel, Structurally Unique Spirocyclic Alkaloids from the Halotolerant B-17 Fungal Strain of *Aspergillus varicolor*. *Chem. Biodivers* 2007, 4 (12), 2913–2919. 10.1002/cbdv.200790240. [PubMed: 18081101]
- (45). Liu Z; Chen Y; Li S; Hu C; Liu H; Zhang W Indole Diketopiperazine Alkaloids from the Deep-Sea-Derived Fungus *Aspergillus* sp. FS445. *Nat. Prod. Res* 2021, 1–9. 10.1080/14786419.2021.1925271.
- (46). Geng C-A; Huang X-Y; Ma Y-B; Hou B; Li T-Z; Zhang X-M; Chen J-J (±)-Uncarilins A and B, Dimeric Isoechinulin-Type Alkaloids from *Uncaria rhynchophylla*. *J. Nat. Prod* 2017, 80 (4), 959–964. 10.1021/acs.jnatprod.6b00938. [PubMed: 28225280]
- (47). Zou X; Li Y; Zhang X; Li Q; Liu X; Huang Y; Tang T; Zheng S; Wang W; Tang J A New Prenylated Indole Diketopiperazine Alkaloid from *Eurotium cristatum*. *Molecules* 2014, 19 (11), 17839–17847. 10.3390/molecules191117839. [PubMed: 25372398]

- (48). Polonsky J; Merrien M-A; Prangé T; Pascard C; Moreau S Isolation and Structure (X-Ray Analysis) of Marcfortine A, a New Alkaloid from *Penicillium roqueforti*. J Chem Soc Chem Commun 1980, No. 13, 601–602. 10.1039/C39800000601.
- (49). Chen M; Liu C-T; Tang Y Discovery and Biocatalytic Application of a PLP-Dependent Amino Acid γ -Substitution Enzyme That Catalyzes C–C Bond Formation. J. Am. Chem. Soc 2020, 142 (23), 10506–10515. 10.1021/jacs.0c03535. [PubMed: 32434326]
- (50). Schkeryantz JM; Woo JCG; Siliphaivanh P; Depew KM; Danishefsky SJ Total Synthesis of Gypsetin, Deoxybrevianamide E, Brevianamide E, and Tryprostatin B: Novel Constructions of 2,3-Disubstituted Indoles. J. Am. Chem. Soc 1999, 121 (51), 11964–11975. 10.1021/ja9925249.
- (51). Santiago CB; Guo J-Y; Sigman MS Predictive and Mechanistic Multivariate Linear Regression Models for Reaction Development. Chem. Sci 2018, 9 (9), 2398–2412. 10.1039/C7SC04679K. [PubMed: 29719711]
- (52). Toste FD; Sigman MS; Miller SJ Pursuit of Noncovalent Interactions for Strategic Site-Selective Catalysis. Acc. Chem. Res 2017, 50 (3), 609–615. 10.1021/acs.accounts.6b00613. [PubMed: 28945415]
- (53). Harper KC; Bess EN; Sigman MS Multidimensional Steric Parameters in the Analysis of Asymmetric Catalytic Reactions. Nat. Chem 2012, 4 (5), 366–374. 10.1038/nchem.1297. [PubMed: 22522256]
- (54). Induced Fit Docking Protocol 2015–2, Glide Version 6.4, Prime Version 3.7, Schrödinger, LLC, New York, NY, 2015.
- (55). Sherman W; Day T; Jacobson MP; Friesner RA; Farid R Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. J. Med. Chem 2006, 49 (2), 534–553. 10.1021/jm050540c. [PubMed: 16420040]
- (56). Sherman W; Beard HS; Farid R Use of an Induced Fit Receptor Structure in Virtual Screening. Chem. Biol. Htmleint Glyphamp Asciiamp Drug Des 2006, 67 (1), 83–84. 10.1111/j.1747-0285.2005.00327.x.
- (57). Chen R; Gao B; Liu X; Ruan F; Zhang Y; Lou J; Feng K; Wunsch C; Li S-M; Dai J; Sun F Molecular Insights into the Enzyme Promiscuity of an Aromatic Prenyltransferase. Nat. Chem. Biol 2017, 13 (2), 226–234. 10.1038/nchembio.2263. [PubMed: 27992881]
- (58). Hochlowski E; Mullally MM; Spanton SG; Whittern DN; Hill P; McAlpine JB II. Isolation and Elucidation of the Structure of 5-N-acetylardeemin and Two Congeners. J. Antibiot. (Tokyo) 1993, 7, 380–386.
- (59). Chou T-C; Depew KM; Zheng Y-H; Safer ML; Chan D; Helfrich B; Zatorska D; Zatorski A; Bornmann W; Danishefsky SJ Reversal of Anticancer Multidrug Resistance by the Ardeemins. Proc. Natl. Acad. Sci 1998, 95 (14), 8369–8374. 10.1073/pnas.95.14.8369. [PubMed: 9653193]
- (60). Clark B; Capon RJ; Lacey E; Tennant S; Gill JH Roquefortine E, a Diketopiperazine from an Australian Isolate of *Gymnoascus reessii*. J. Nat. Prod 2005, 68 (11), 1661–1664. 10.1021/np0503101. [PubMed: 16309319]
- (61). Lacey E; Power M; Wu Z; Richards RW Terprenylated Diketopiperazines, (Drimentines) WO1998009968A1, 1998.
- (62). Khalil ZG; Huang X; Raju R; Piggott AM; Capon RJ Shornephine A: Structure, Chemical Stability, and P-Glycoprotein Inhibitory Properties of a Rare Diketomorpholine from an Australian Marine-Derived *Aspergillus* sp. J. Org. Chem 2014, 79 (18), 8700–8705. 10.1021/jo501501z. [PubMed: 25158286]
- (63). Tsunematsu Y; Ishikawa N; Wakana D; Goda Y; Noguchi H; Moriya H; Hotta K; Watanabe K Distinct Mechanisms for Spiro-Carbon Formation Reveal Biosynthetic Pathway Crosstalk. Nat. Chem. Biol 2013, 9, 818–825. 10.1038/nchembio.1366. [PubMed: 24121553]
- (64). Zhong W-M; Wang J-F; Shi X-F; Wei X-Y; Chen Y-C; Zeng Q; Xiang Y; Chen X-Y; Tian X-P; Xiao Z-H; Zhang W-M; Wang F-Z; Zhang S Eurotiumins A–E, Five New Alkaloids from the Marine-Derived Fungus *Eurotium* sp. SCSIO F452. Mar. Drugs 2018, 16 (4), 136. 10.3390/md16040136. [PubMed: 29690501]
- (65). Ruiz-Sanchis P; Savina SA; Albericio F; Álvarez M Structure, Bioactivity and Synthesis of Natural Products with Hexahydropyrrolo[2,3-b]Indole. Chem. - Eur. J 2011, 17 (5), 1388–1408. 10.1002/chem.201001451. [PubMed: 21268138]

- (66). Adhikari AA; Chisholm JD Lewis Acid Catalyzed Displacement of Trichloroacetimidates in the Synthesis of Functionalized Pyrroloindolines. *Org. Lett* 2016, 18 (16), 4100–4103. 10.1021/acs.orglett.6b02024. [PubMed: 27486831]
- (67). Kmuníček J; Hynková K; Jedlicka T; Nagata Y; Negri A; Gago F; Wade RC; Damborský J Quantitative Analysis of Substrate Specificity of Haloalkane Dehalogenase LinB from *Sphingomonas paucimobilis* UT26[†]. *Biochemistry* 2005, 44 (9), 3390–3401. 10.1021/bi047912o. [PubMed: 15736949]
- (68). Mazurenko S; Prokop Z; Damborsky J Machine Learning in Enzyme Engineering. *ACS Catal* 2020, 10 (2), 1210–1223. 10.1021/acscatal.9b04321.
- (69). Neubauer PR; Pienkny S; Wessjohann L; Brandt W; Sewald N Predicting the Substrate Scope of the Flavin-Dependent Halogenase BrvH. *ChemBioChem* 2020, 21 (22), 3282–3288. 10.1002/cbic.202000444. [PubMed: 32645255]
- (70). Siedhoff NE; Schwaneberg U; Davari MD Machine Learning-Assisted Enzyme Engineering. In *Methods in Enzymology*; Elsevier, 2020; Vol. 643, pp 281–315. 10.1016/bs.mie.2020.05.005. [PubMed: 32896285]
- (71). Koryakina I; Kasey C; McArthur JB; Lowell AN; Chemler JA; Li S; Hansen DA; Sherman DH; Williams GJ Inversion of Extender Unit Selectivity in the Erythromycin Polyketide Synthase by Acyltransferase Domain Engineering. *ACS Chem. Biol* 2017, 12 (1), 114–123. 10.1021/acscchembio.6b00732. [PubMed: 28103677]
- (72). Shende VV; Khatri Y; Newmister SA; Sanders JN; Lindovska P; Yu F; Doyon TJ; Kim J; Houk KN; Movassaghi M; Sherman DH Structure and Function of NzeB, a Versatile C–C and C–N Bond-Forming Diketopiperazine Dimerase. *J. Am. Chem. Soc* 2020, 142 (41), 17413–17424. 10.1021/jacs.0c06312. [PubMed: 32786740]
- (73). Yang S; DeMars MD; Grandner JM; Olson NM; Anzai Y; Sherman DH; Houk KN Computational-Based Mechanistic Study and Engineering of Cytochrome P450 MycG for Selective Oxidation of 16-Membered Macrolide Antibiotics. *J. Am. Chem. Soc* 2020, 142 (42), 17981–17988. 10.1021/jacs.0c04388. [PubMed: 33030347]
- (74). Newmister SA; Srivastava KR; Espinoza RV; Caddell Haatveit K; Khatri Y; Martini RM; Garcia-Borrás M; Podust LM; Houk KN; Sherman, David. H. Molecular Basis of Iterative C–H Oxidation by TamI, a Multifunctional P450 Monooxygenase from the Tirandamycin Biosynthetic Pathway. *ACS Catal* 2020, 10 (22), 13445–13454. 10.1021/acscatal.0c03248. [PubMed: 33569241]
- (75). Kato H; Yoshida T; Tokue T; Nojiri Y; Hirota H; Ohta T; Williams RM; Tsukamoto S Notoamides A–D: Prenylated Indole Alkaloids Isolated from a Marine-Derived Fungus, *Aspergillus* sp. *Angew. Chem. Int. Ed* 2007, 46 (13), 2254–2256. 10.1002/anie.200604381.
- (76). Imhoff JF; Labes A; Wiese J Bio-Mining the Microbial Treasures of the Ocean: New Natural Products. *Biotechnol. Adv* 2011, 29 (5), 468–482. 10.1016/j.biotechadv.2011.03.001. [PubMed: 21419836]

**Figure 1.**

A) Prenylated DKP-containing NPs. Reverse prenyl groups are colored in blue; B) C2 reverse prenyltransfer catalyzed by NotF en route to (-)-notoamide A with P450s colored in red and flavin-dependent enzymes in yellow.

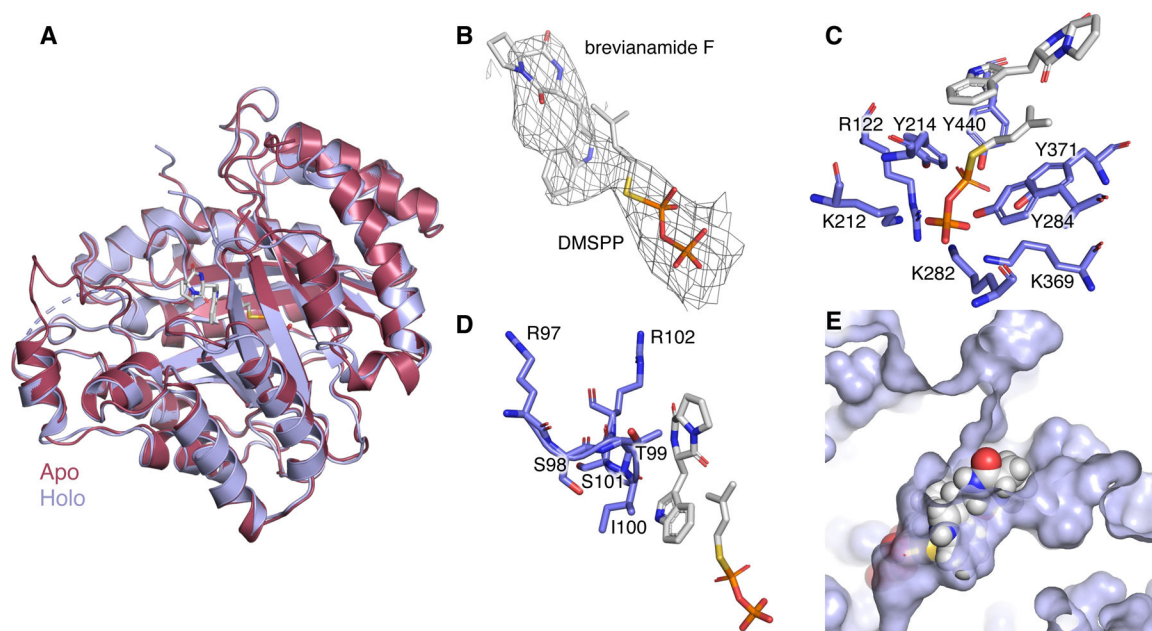
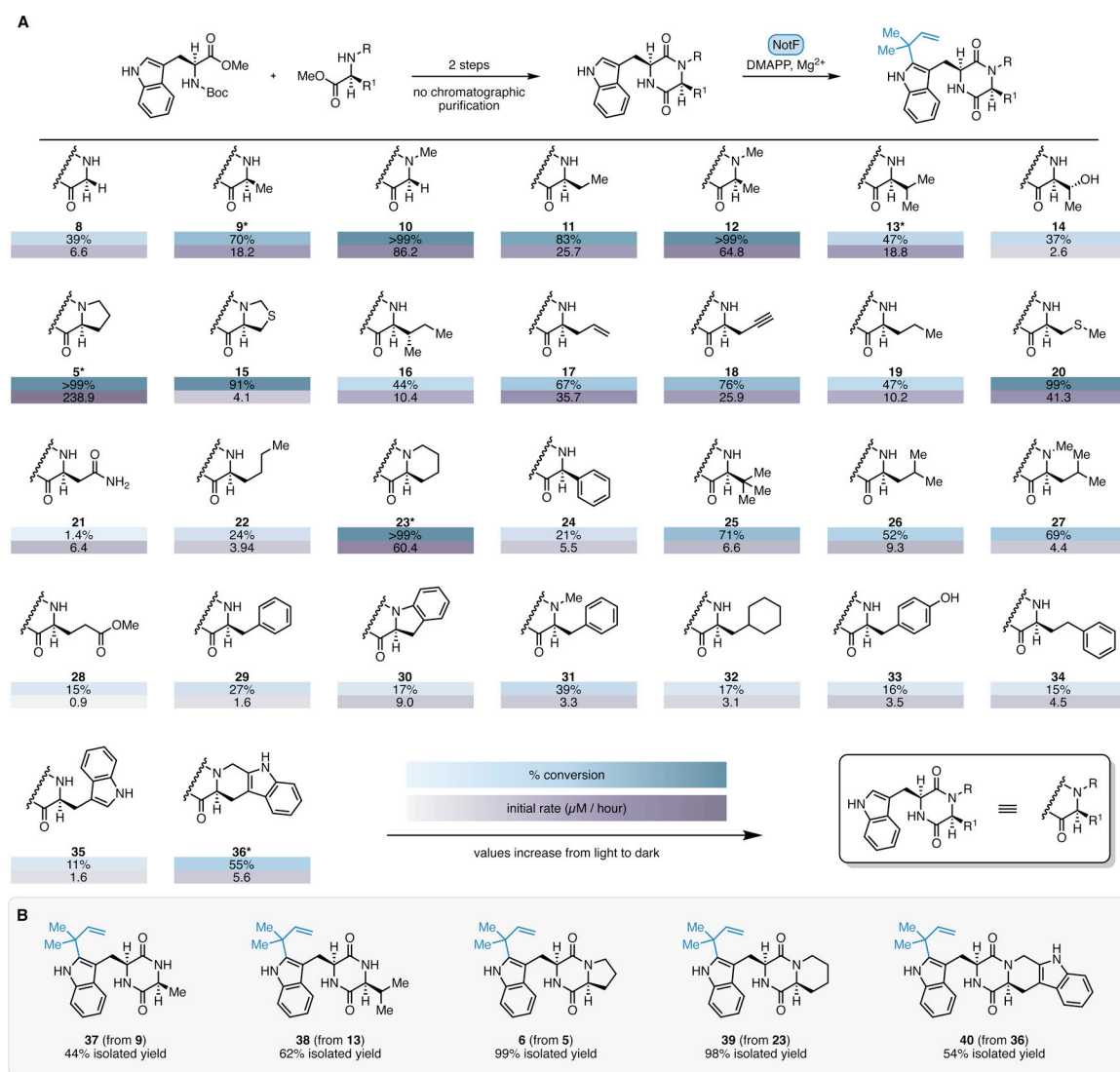
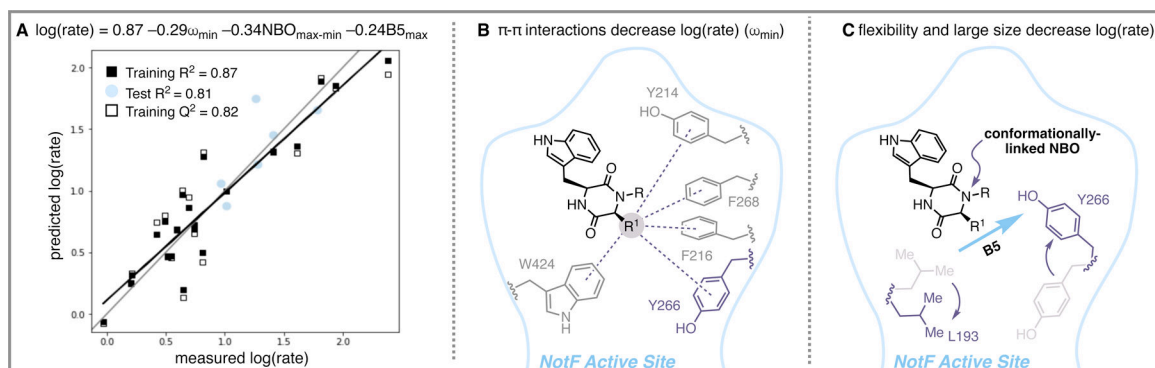


Figure 2.

Crystal structure of NotF in complex with brevianamide F (**5**) and DMSPP. A) structural alignment of NotF apo (maroon, PDB ID 6VY9) and holo (purple, PDB ID 6VYA) structures; B) F_0-F_c map with ligands omitted from refinement at contour level 1.0 MATH presents unambiguous density for both ligands; C) conserved residues in the active site interact with DMSPP; D-E) brevianamide F is tightly packed in the reaction chamber. The chamber surface is colored in purple.

**Figure 3.**

A) General reaction scheme and substrate scope of NotF using synthesized panel of tryptophanyl DKPs with % conversion in blue and initial rate (μM per hour) in purple; B) Reverse C2 prenylated products generated on preparatory scale and characterized by NMR.

**Figure 4.**

Linear model and mechanistic interpretation of molecular feature relation to initial rate ($\mu\text{M}/\text{hour}$). A) The linear regression model shown indicates three terms which lead to decreased rate. The Training and Test R^2 values agree with each other, indicating a successfully descriptive model; B) The DKP substrate framework shown in a cartoon of the active site of NotF. Illustrated are key aromatic residues present within 5 Å of native substrate; C) The DKP substrate framework shown in a cartoon of the active site of NotF. In the crystal structure of native substrate **5** bound to NotF, active site residues tightly surround the binding pocket, indicating expansion in width or length coupled with increased flexibility would demand major residue repositioning.

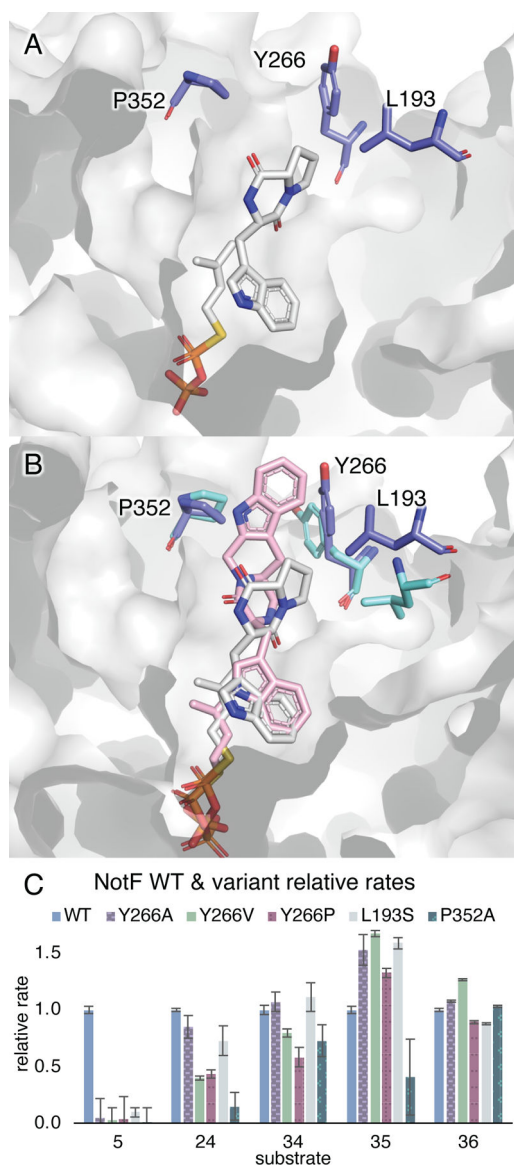


Figure 5. IFD-guided engineering of a NotF variant with enhanced rate with bulky, aromatic substrate, **36**. A) Active site residues L193, F216, and Y266 with in substrate-bound NotF crystal structure; B) Overlay of substrate (brevianamide F and DMSPP in grey, active site residues in purple) bound crystal structure with IFD pose of substrate **36** bound (pink, DMSPP in pink), with repositioned active site residues in cyan; C) Relative rate comparison between NotF WT and variants with substrates **5**, **24**, **34**, **35**, and **36**.

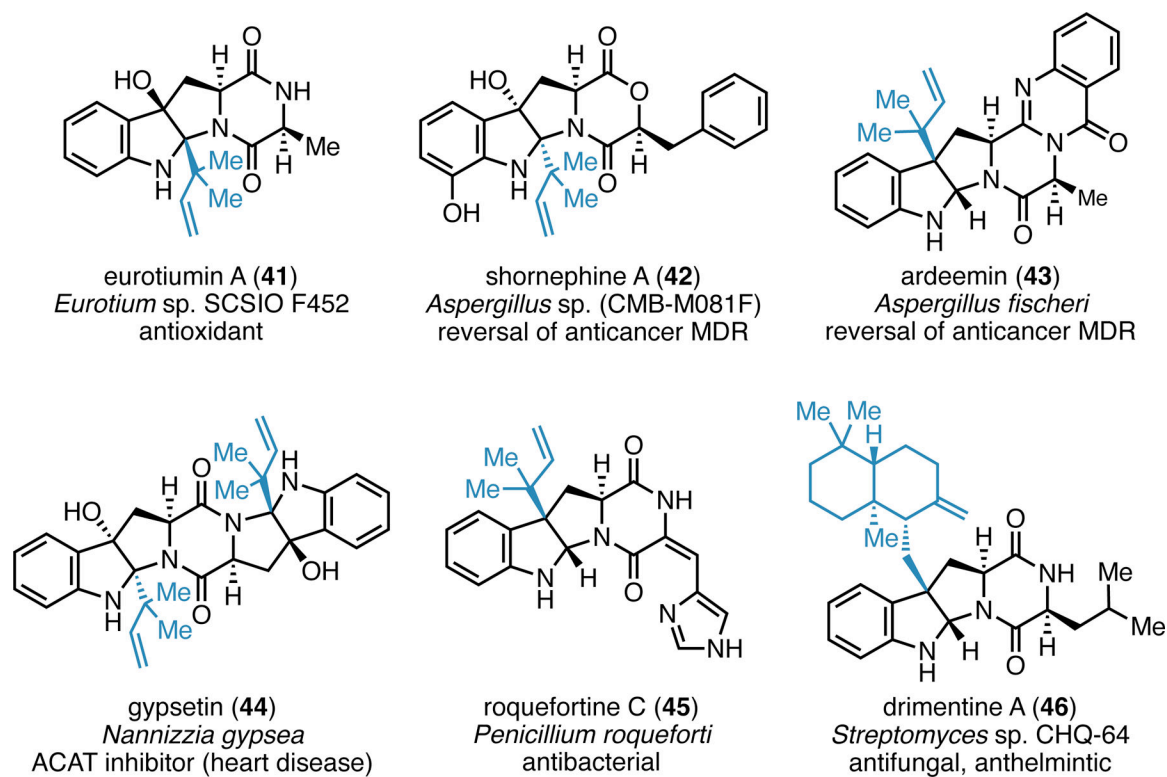


Figure 6.
Biologically active prenylated pyrroloindoline DKP-containing NPs.

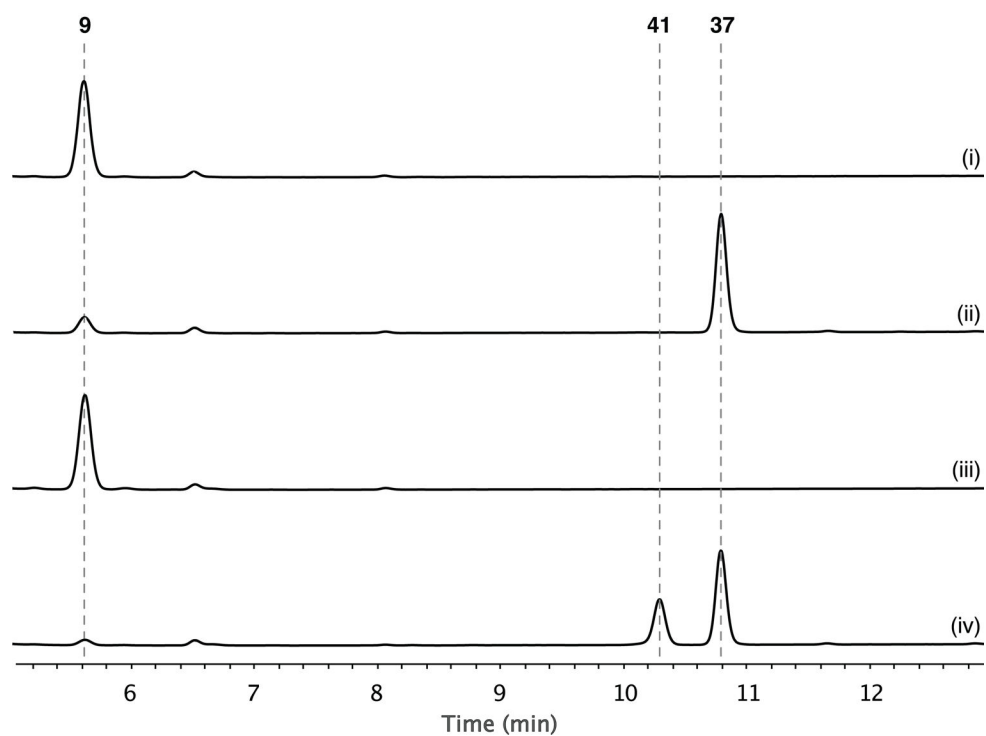


Figure 7. One-pot, diastereoselective biocatalytic cascade with PT NotF and FMO BvnB towards 3-hydroxypyrroloindoline NPs brevianamide E (47) and eurtiumin A (41): HPLC traces of (i) no enzyme control; (ii) 9 + NotF; (iii) 9 + BvnB; (iv) 9 + NotF + BvnB.

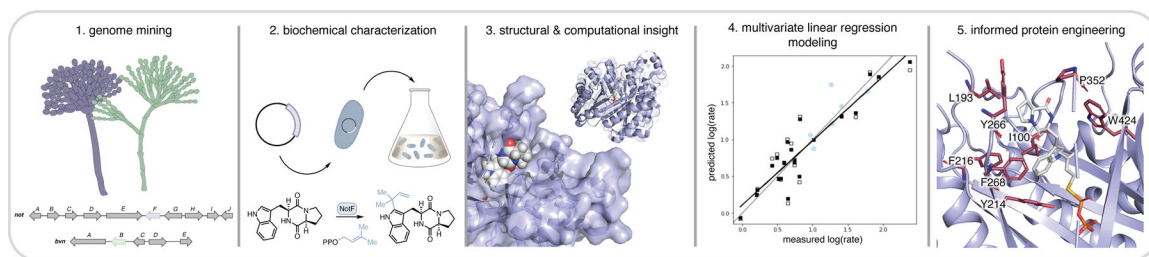


Figure 8.

Enzyme discovery and characterization pipeline used in this work. Our typical workflow includes more computationally intensive MD simulations and/or QM/MM calculations, but here we introduce a new statistical-driven module via multivariate linear regression modeling (MLR) that will enable future engineering of NotF and other biocatalysts.