Title

Module-Based Polyketide Synthase Engineering for de Novo Polyketide Biosynthesis.

Journal

ACS Synthetic Biology, 12(11)

Authors

Nava, Alberto

Roberts, Jacob

Haushalter, Robert

et al.

Publication Date

2023-11-17

DOI

10.1021/acssynbio.3c00282

Peer reviewed

ACS Synthetic Biology

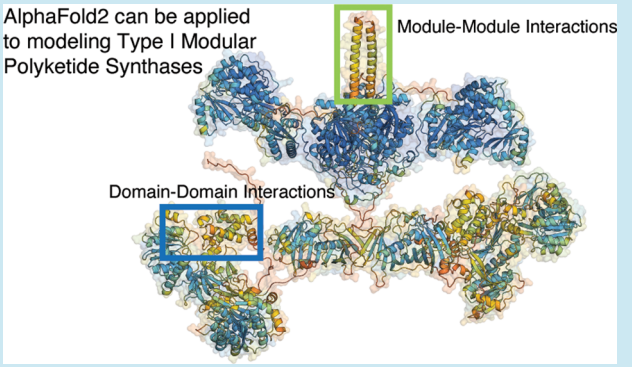
pubs.acs.org/synthbio
Review

Module-Based Polyketide Synthase Engineering for *de Novo* Polyketide Biosynthesis

Alberto A. Nava, Jacob Roberts, Robert W. Haushalter, Zilong Wang, and Jay D. Keasling*

Cite This: *ACS Synth. Biol.* 2023, 12, 3148–3155
Read Online

ABSTRACT: Polyketide retrobiosynthesis, where the biosynthetic pathway of a given polyketide can be reversibly engineered due to the colinearity of the polyketide synthase (PKS) structure and function, has the potential to produce millions of organic molecules. Mixing and matching modules from natural PKSs is one of the routes to produce many of these molecules. Evolutionary analysis of PKSs suggests that traditionally used module boundaries may not lead to the most productive hybrid PKSs and that new boundaries around and within the ketosynthase domain may be more active when constructing hybrid PKSs. As this is still a nascent area of research, the generality of these design principles based on existing engineering efforts remains inconclusive. Recent advances in structural modeling and synthetic biology present an opportunity to accelerate PKS engineering by re-evaluating insights gained from previous engineering efforts with cutting edge tools.



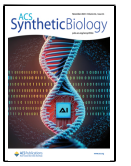
KEYWORDS: *polyketide synthases, structural modeling, retrobiosynthesis, protein engineering*

INTRODUCTION

Retrosynthesis is the concept of designing synthetic organic chemistry routes by working backward from the final product to define a series of achievable reactions from simpler building blocks.[1] Similarly, retrobiosynthesis is the application of a similar concept with the addition of enzyme-catalyzed chemical reactions.[2] Type I polyketide synthases (PKSs) have been heralded as a potential foundation for retrobiosynthesis since their first elucidation as modular enzymatic assembly lines in the early 1990s.[3,4] The polyketides produced by PKSs are a diverse class of natural products with extensive bioactivities including antibacterial (e.g., erythromycin), antifungal (e.g., amphotericin B), and anticancer (e.g., epothilone) properties with agricultural and medicinal applications.[5] Moreover, the stereochemically rich, highly functionalized cores of these compounds pose significant obstacles for synthetic chemists. Thus, the interest and potential in engineering PKSs is based on the valuable activities of these complex compounds, and their colinear biosynthetic logic, meaning that the chemical structure of the compounds produced can be directly inferred by the order and type of the enzymes in the pathway.

PKSs are composed of modules that act as part of an assembly line carrying out stepwise decarboxylative Claisen condensation reactions of acyl-CoA building blocks.[6] The mechanism of polyketide elongation has long been compared to fatty acid synthesis[7,8] in that an acyl-CoA extender unit is first primed by an acyl transferase (AT) onto its cognate acyl carrier protein (ACP) before undergoing a decarboxylative Claisen condensation reaction with the polyketide chain tethered to a β-keto synthase (KS). The polyketide β-carbon is then optionally reduced by a NADPH-dependent ketoreductase (KR) to a hydroxyl group, then dehydrated by a dehydratase (DH) to an enoyl group, and finally reduced to a fully saturated β-carbon by an enoyl reductase (ER). After undergoing elongation and reduction, the polyketide chain is shuffled by the ACP to the KS domain of the downstream module where the assembly line continues with the condensation of a new extender unit. At the end of the assembly line, the most common termination domain is a thioesterase (TE), which either hydrolyzes the linear product off the PKS or catalyzes a cyclization reaction.[9−11] The regio- and stereochemistry of each ketide building block is determined by the combination of conserved catalytic motifs in the KS, KR, DH, and ER.[12] The structure of the final ketide product can be inferred from the order in which these modules interact, which is directed by both protein−protein interactions and substrate specificity of each domain. Therefore,

Received: May 3, 2023
Published: October 23, 2023



https://doi.org/10.1021/acssynbio.3c00282
ACS Synth. Biol. 2023, 12, 3148−3155

there is a direct relationship between the DNA sequence of the biosynthetic cluster and the structure of the molecule produced.

PKSs represent an attractive avenue to the access of truly complex carboxylic molecules, and their unique organization has prompted numerous engineering efforts by mutating or shuffling catalytic parts for the biosynthesis of novel products. They have been successfully engineered to produce numerous interesting bioproducts including potential therapeutics,[13] polymer precursors,[14] specialty chemicals,[15] and biofuels.[16] Nevertheless, many engineered PKSs suffer from reduced catalytic activities.[9] Two primary overarching issues have prevented PKSs from reaching their full potential: neither the protein−protein interactions nor the substrate−protein interactions within PKSs are well characterized.[10,17−19] However, recent structural,[20,21] evolutionary,[22,23] biochemical,[24] and metabolic engineering[16] studies suggest that the traditional boundaries used for module−module connections may not provide optimal interactions between modules, and that there are alternative boundaries that may boost success rates.[24,25] Furthermore, recent progress in the field of machine learning has enabled impressive models in a diverse range of fields including protein structure prediction. The release of Alphafold2[26] and RoseTTAFold[27] for protein structure prediction has prompted a swath of studies investigating complex macromolecular phenomena including protein complex analyses.[28] In this perspective, we discuss how previous module-based PKS engineering efforts can be re-evaluated in the context of modern structural modeling and how the insights gained may initiate a new generation of PKS design principles.

MODULE-BASED POLYKETIDE SYNTHASE ENGINEERING

The natural homology existing between PKS modules from different biosynthetic gene clusters (BGCs) within diverse organisms sparked one of the most ambitious goals of PKS engineering, the lego-ization of PKSs. The lego-ization of PKSs involves the combinatorial expression of PKS parts for the biosynthesis of novel bioproducts (Figure 1A).[30,31] Demonstrating the inherent promiscuity with PKS parts, Menzella et al. successfully biosynthesized over a dozen novel triketide lactones by recombining dozens of PKS modules.[31] However, the dramatically reduced product titers from most of their engineered constructs showed that there are unknown fundamental principles that govern the interactions between PKS modules.

Subsequent studies that investigated the KS domain from downstream modules revealed that the KS domain plays a critical role in the interactions between chimeric modules, and that swapping the KS can, in some cases, resuscitate inactive interactions.[32] Though the failures in KS-catalyzed chain extension are presumably multifaceted and still have not been logically resolved, the gatekeeping behavior of KSs has been supported by in-depth biochemical[33] and evolutionary analyses.[22,23] The evolutionary analyses provide evidence that KS domains more strongly coevolve with upstream domains, suggesting that an alternative definition of a PKS module would have the KS domain as the last domain of a PKS module instead of the first. This alternative definition of a PKS module (starting at AT and ending after KS) is known as the PKS exchange unit (XU),[34] analogous to the definition used in non-ribosomal peptide synthetases (NRPSs).[35,36]

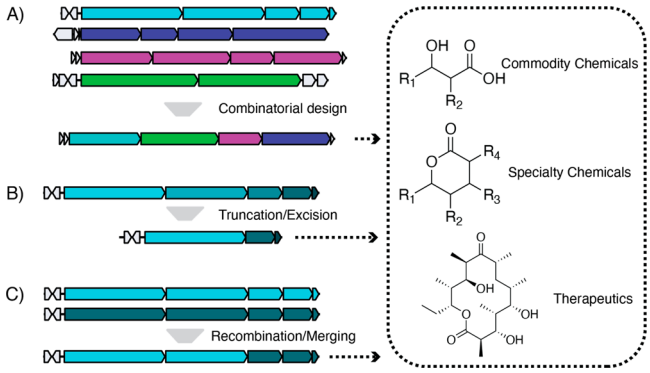


Figure 1. Module-based PKS engineering. (A) Original ideas for a PKS retrobiosynthesis platform involved recombining parts from known PKS biosynthetic gene clusters (BGCs) for the purpose of producing novel molecules. The colors of the genes signify that PKS assembly lines are being composed by a combinatorial library of individual parts (e.g., subunits, modules). (B) Evolutionary guided PKS engineering involves modifying an individual PKS BGC by excision of modules based on recombination principles. The goal being to introduce as few non-native protein−protein interactions as possible. (C) Strategy of recombining two homologous BGCs to produce non-native products. The goal in this strategy is to identify PKS BGCs that allow you to make as few engineering changes as possible in order to achieve a desired final product. On the right side of the figure are a few diverse applications of polyketide molecules: 3-hydroxy acids as commodity chemicals, δ-lactones as specialty chemicals, macrolactones as therapeutics (6-deoxyerythronolide B).

This XU model contrasts with the genetic organization of the domains within open reading frames but is logical given that the structural characteristics of a substrate entering a KS active site are determined by the AT and reducing domains. The biochemical analyses provide evidence that the condensation reaction catalyzed by the KS domain is indeed the rate-limiting step.[37] Notably, this KS gatekeeping activity is much more prevalent in trans-AT PKS assembly lines than in cis-AT PKS systems and has been used to identify generalizable design principles for such systems.[38−40] Trans-AT PKSs are differentiated from cis-AT PKSs in that the AT domain, which is responsible for loading an extender unit onto the ACP, is not located within the same module that performs chain elongation. Instead, the AT domain is either found as a separate, standalone enzyme or occasionally within a different module. Though closely related to cis-AT PKSs, the unique architectural nuances of trans-AT PKSs mean that some but not all engineering strategies can be directly generalized between systems.

Insights about the mechanism of PKS diversity have emerged from the observation that the phylogeny of KS domains aligns closely with the host organism phylogeny and that KS domains are most closely related to other KS domains within the same PKS, and more broadly to other KS domains within the same host organism. This suggests a model that PKSs evolved primarily through horizontal gene/BGC transfer, recombination, gene conversion, and genetic drift, rather than through gene duplication.[41−44] Furthermore, this model implies that many evolutionary means exist to alleviate reduced activity at module boundaries via convergent evolution. It should continue to be a key area of focus to determine what biochemical aspects of PKS evolution can be generalized.

https://doi.org/10.1021/acssynbio.3c00282
ACS Synth. Biol. 2023, 12, 3148−3155

The consideration of these factors has led to some recent evolutionary-inspired PKS engineering strategies (Figure 1B,C). For instance, the Tylosin and Rapamycin PKSs were the target of an accelerated evolution experiment that produced numerous active truncated assembly lines.[13] Notably, the mutant assembly lines were not generated through targeted *in vitro* cloning, but rather through spontaneous *in vivo* homologous recombination. The authors reported that when modules were deleted from the native PKS, the place where recombination generally occurred was not particularly on either end of the KS but rather throughout the KS domain, AT domain, and the interdomain linker region upstream of the ACP, with a notable hot spot for recombination within the middle of the KS that takes advantage of the high sequence homology between KS domains. The Pikromycin PKS is another significant case study system providing evidence that utilizing XU module boundaries can improve activity of chimeric PKSs.[24,45] Moreover, *in vivo* homologous recombination engineering strategies have also been applied to the Pikromycin PKS and similarly demonstrate successful recombination throughout the KS and AT domains.[46] The Aureothin and Neoaureothin PKSs are two homologous BGCs that were successfully engineered to produce non-native products by utilizing XU module boundaries.[47]

The Stambomycin PKS has been subject to one of the most systematic studies of module boundaries to date, where six different approaches to generating a truncated version of the assembly line were applied.[34] The goal of Su et al. was to engineer a successful interaction between module 13 and module 21 of the Stambomycin PKS.[34] They compared six strategies for engineering this interaction including the following: swapping the C-terminal docking domain of module 13 for the C-terminal docking domain of module 20 [nonfunctional]; swapping the C-terminal docking domain and performing a mutation in the KS-ACP interface region of the ACP in module 13 [functional]; swapping the KS domain of module 21 for the KS domain of module 14 utilizing XU boundaries [functional]; swapping the KS domain of module 21 for the KS domain of module 14 utilizing the recombination hot spot boundaries [functional]; swapping the KS domain of module 21 for the KS domain of module 14 utilizing XU boundaries and swapping the KS-ACP interface region of the ACP in module 21 for the KS-ACP interface region of the ACP in module 14 [nonfunctional]; and last swapping the KS domain of module 21 for the KS domain of module 14 utilizing XU boundaries and performing a G to D mutation in the ACP of module 21 [functional].[34] The functional variants were successfully able to produce measurable quantities of truncated final product; notably, the variant that swapped the downstream KS domain using the recombination hot spot junction in the middle of the KS domain achieved the best kinetics.[34]

Recombination at each of the three junctions (Figure 2, Figure S1) have exhibited some success in engineered systems; however, without a systematic study examining module boundaries in a variety of systems, it is difficult to identify generalizable design principles. One promising approach that has recently been applied for the optimization of AT-swap junctions that could potentially be used for more systematically analyzing module−module interactions involves the use of oligonucleotide pools for generating libraries of junctions, a solubility biosensor for performing a preliminary high throughput screen of high expressing variants, and a kinetic analysis of a diverse subset of the library.[48] This approach is amenable to high throughput experimentation and could enable more generalizable junctions by systematically characterizing the space of possible junctions. For now, recent evidence (Table 1) suggests that either the XU module boundary or recombination hot spot are more likely to yield active variants, but it is clear that there is not a one-junction-fits-all type of rule and it remains to be determined whether there are alternative junctions which would yield better activity.

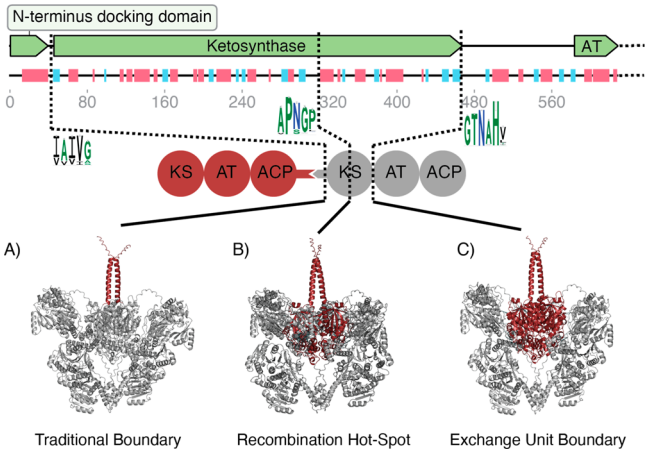


Figure 2. Module junction selected is critical to the solubility and activity of engineered PKSs. There are primarily three junctions that have been explored: (A) the traditional boundary at the start of the KS domain, (B) a recombination hot spot in the middle of the KS domain, and (C) the exchange unit boundary at the end of the KS domain. Shown is Module 21 of the Stambomycin PKS (Sta21) modeled by AlphaFold2 with each module junction (from Su et al.[34]) highlighted in red. Domain annotations within the boundary region are shown at the top of the figure with a secondary structure plot of the KS domain in Sta21 with α-helices as pink-colored boxes and β-sheets as teal-colored boxes. Secondary structure analysis was performed with the PSIPRED server.[29]

Table 1. Investigations into PKS Module Boundaries

Module Boundary	Rationale for Engineering	Product	References
Traditional boundary	High sequence conservation and convenient restriction sites	Triketide lactones	Menzella et al.[31]
Exchange unit boundary	KS gatekeeping hypothesis	Triketide lactones	Chandran et al.[32]
Recombination hot spot	Homology-directed recombination	Rapalogs	Wlodek et al.[13]
Traditional and exchange unit boundary	Co-evolution analysis; after KS was best	Triketide lactones	Miyazawa et al.[24]
All 3 boundaries	Previous literature; middle of KS was best	Mini-stambomycins	Su et al.[34]

POLYKETIDE SYNTHASE DYNAMICS

PKSs are very large and highly dynamic homodimeric proteins, which means that structural characterization of PKS assembly lines was until recently limited to individual domains or didomains. For example, docking domains,[49−51] AT domains,[52,53] KS-AT didomains,[54,55] DH domains,[56,57] KR-ER didomains,[58] TE domains,[59,60] and KR domains[61] have all been individually characterized. However, the activity of PKS

https://doi.org/10.1021/acssynbio.3c00282
ACS Synth. Biol. 2023, 12, 3148−3155

assembly lines is highly dependent on the interactions between domains, which are not able to be completely described from structures of isolated domains. In particular, for the purpose of module-based PKS engineering, there are several structural features that critically determine catalytic activity including the following: (i) the ability of an ACP domain to translocate a polyketide chain to a downstream KS domain, (ii) the ability of a KS domain to catalyze the condensation of a polyketide chain with an extender unit, (iii) and the ability of an ACP domain to translocate a ketide extender unit with an upstream KS domain.

There are several recent reports of larger multidomain structural characterization efforts that reveal insights into how to potentially improve module-based engineering. The first direct description of module−module interactions utilized the ability of small-angle X-ray scattering (SAXS) to capture low-resolution structures of proteins in the Erythromycin PKS without the need for crystallization.[62] This revealed conformations that active PKS modules undergo; however, many key details about intermodule communication could not be elucidated due to the resolution. Cryogenic electron microscopy (cryo-EM) has additionally been used to capture conformational data about a full-length PKS module in the presence of an upstream ACP, resulting in the hypothesis that there are two distinct active sites within the KS domain for intramodule and intermodule ACP interactions.[63,64] However, the arch-shaped structure observed in this system is different from previous extended-shaped models and is being further rebutted by recent structural data.[65] Nonetheless, cryo-EM has been used in tandem with the first high-resolution X-ray crystallographic structure of a full-length PKS module to illuminate many of the dynamic events that must happen for intramodule transacylation and chain extension to occur.[20] Molecular dynamics simulations have been employed to visualize the small-subunit structural changes that occur as an AT domain accepts an extender substrate.[66] High-resolution cryo-EM structures of an Erythromycin PKS module further reveal dynamics of how an AT domains' position relative to its cognate KS domain enable transacylation.[21] In type II fatty acid synthases, several key residues have been identified at the interaction interface between ACP and KS domains that alter the behavior of substrates in the KS active site,[67,68] and detailed mechanisms of KS catalytic activity have been elucidated.[69,70] Though the KS-ACP interface is not identical between type II fatty acid synthases and type I PKSs,[71] the key interplay observed between the ACP domain and its corresponding AT and KS domains sustains the idea that domain−domain interactions incorporating the ACP are crucial to PKS activity. The essentiality of compatible ACP interactions, in particular with the KS domain of a downstream module, has been corroborated in several reports.[17,33] The recent improvement in the resolution of cryo-EM has enabled critical perspectives to the structural dynamics of PKS enzymes and should continue to be employed to further elucidate trajectories of catalytic activity, as has been done in the NRPS field.[72]

PROTEIN MODELING FOR POLYKETIDE SYNTHASES

The release of AlphaFold2[26,73,74] and RoseTTAFold[27] as highly accurate protein structure prediction algorithms based on deep-learning models has enabled numerous impressive protein engineering feats including *de novo* protein design[75] and proteome-scale protein interaction screening.[28] Similar deep-learning models have been used to develop MutCompute,[76] a generalizable protein design algorithm that has successfully been applied to optimize the thermostability, pH tolerance, and kinetics of a PET hydrolase.[77] Furthermore, the advancement of large language models has led to the development of programs including ESMFold,[78] a rapid protein structure prediction algorithm, and ProGen,[79] a generative protein sequence algorithm capable of hallucinating proteins with custom properties. Additional tools have enabled more accessible use to AlphaFold2 including ColabFold[74] and Foldy.[80] Though these modern protein modeling algorithms have enabled incredible achievements, challenges remain in the modeling of large proteins (>3000 amino acids) and in high accuracy protein complex modeling. Nonetheless, these modern protein modeling algorithms may have the potential to augment our understanding and ability to engineer PKSs.

Here we demonstrate the potential applicability of AlphaFold2 to the modeling of PKSs. We have shown that AlphaFold2 is capable of reasonably modeling an entire PKS module (Figure S2) and in some cases demonstrates asymmetric reaction chambers (Figure S3). Each PKS model generated by AlphaFold2 has an extended-shape conformation, and we have not observed any models that have an arch-shaped conformation. This is likely due to the majority of experimentally derived structures of KS-AT domains, which form part of the data set on which AlphaFold2 is trained, being in an extended-shape conformation. An additional attractive feature would be the ability to model multiple modules at the same time. Unfortunately, we have been unable to obtain an AlphaFold2 model of two full-length modules at the same time due to two main reasons: (i) the GPU memory required scales quadratically with the length of the protein leading to out-of-memory errors for protein complexes that are too large and (ii) reducing the memory usage by reducing the size of the multiple sequence alignments input into AlphaFold2 leads to artifacts in the resultant structure such as overlapping chains (Figure S4). Alternatively, we have been able to demonstrate that it is possible to truncate the upstream module to facilitate modeling with a downstream module, thus making it possible to visualize an ACP interacting with both a downstream KS-AT (Figure 3A, Figure S5) and an upstream KS-AT (Figure 3B, Figure S5). In a useful demonstration of how KS-ACP modeling can inform and boost engineering success, Buyachuihan et al. utilized AlphaFold2 models of the Venemycin synthase to improve product yields by optimizing docking domains that take into account the flexibility required at the KS-ACP interface.[81]

These capabilities beg the question of whether structural modeling could be used to predict the likelihood that a given ACP could successfully interact with a given KS domain. We used Alphafold to model the 144 intersubunit ACP and KS-AT interactions experimentally characterized by Menzella et al.[31] and applied the state-of-the-art protein complex prediction algorithm described by Humphreys et al.[28] to calculate the probability that a given upstream ACP would successfully interact with a given downstream KS, the hypothesis being that the predicted interaction probability between natively interacting modules would be higher than non-natively interacting modules and that it would correlate with overall production titers. Overall, we observed limited correlation between the predicted interaction probability and the observed production data from Menzella et al.[31] (Figure S6). This result is not

https://doi.org/10.1021/acssynbio.3c00282
ACS Synth. Biol. 2023, 12, 3148−3155

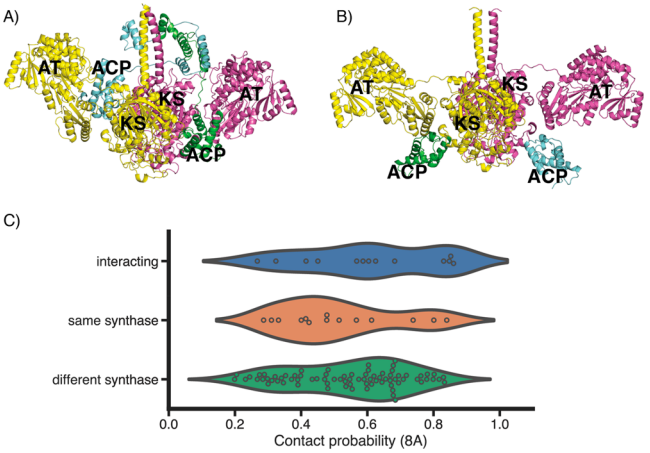


Figure 3. Modern deep-learning protein models applied to PKSs. (A) An AlphaFold2 model of an upstream ACP interacting with a downstream KS-AT with corresponding docking domains. Shown is the Lipomycin PKS Module 1 ACP (green and blue) and Module 2 KS-AT (yellow and pink). (B) An AlphaFold2 model of an ACP interacting with an upstream KS-AT. Shown is the Epothilone PKS Module 7 KS-AT (yellow and pink) and ACP (green and blue). (C) A violin heat map of AlphaFold2 predicted protein−protein contact probabilities (defined as the highest predicted contact probability over all pairs of residues in two protein chains at a threshold of 8 Å[28]) for each of the ACP and KS-AT interactions experimentally characterized in Menzella et al.[31] Combinations are categorized based on whether they natively interact (blue), whether they are natively in the same PKS BGC (orange), or otherwise from different PKS BGCs (green).

surprising as it is clear that the product titer depends on many factors including non-native substrate tolerance and chimeric protein stability. Also, despite our initial hypothesis that the predicted interaction probability between natively interacting domains would be higher than that between non-natively interacting domains, we see a limited correlation between the predicted interaction probability and whether the interaction between an ACP and KS is natively occurring (Figure 3C). It is possible that the docking domains included in the models, which are non-native to each combination (except for eryM2 ACP and eryM3 KS-AT), lead to discordance in the predictions, or it could also be possible that AlphaFold2 does not have the precision and resolution necessary to accurately differentiate between the highly homologous domains between PKS modules. Ultimately, it is likely that the successful interaction between KS and ACP domains is necessary but not sufficient for productivity and that a more holistic analysis of protein−protein and substrate−protein interactions is necessary for inference. Advancements in the way that the co-evolutionary signal is calculated for colinear multidomain and multisubunit enzyme assembly lines like PKSs should boost the accuracy of these algorithms, by for example using antiSMASH[82] annotations to update multiple sequence alignments. Additionally, the advancement and incorporation of deep-learning tools capable of handling substrate−protein interactions, such as DiffDock,[83] should enable more comprehensive modeling of PKS behavior. Lastly, the ability of AlphaFold2 to generate unique stable conformations of PKSs should enable detailed visualizations of PKS dynamics. Ultimately, these tools are at the stage where unique and novel insights can already be obtained, and further improvements will only accelerate what is possible in PKS engineering.

OUTLOOK

PKSs have shown promise as a retrobiosynthesis platform for production of a wide variety of small molecules. PKS intermodule communication has naturally been a topic of significant interest as it is frequently the bottleneck in engineered systems. PKS collinearity presents an interesting opportunity for machine-learning scientists to test modern protein design strategies on a system with a reduced design space. Modern protein design programs present a complementary opportunity for PKS engineers to obtain optimized PKS designs that take into account evolutionary signals, structural elements, and empirical evidence. The collaboration between these groups should prove beneficial.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.3c00282.

Quality measures of AlphaFold2 model of Stambomycin PKS Module 21 (Figure S1); AlphaFold2 model of Lasalocid PKS Module 14 (Figure S2); AlphaFold2 model of Lipomycin PKS Module 1 demonstrating asymmetric reaction chamber (Figure S3); example of steric artifacts caused by modeling two consecutive PKS modules with AlphaFold2 (Figure S4); quality measures of AlphaFold2 models of KS-AT and ACP interactions used in Figure 3 (Figure S5); graphs of interaction probability between KS and ACP domains (Figure S6) (PDF)

PDB files of AlphaFold2 models (ZIP)

AUTHOR INFORMATION

Corresponding Author

Jay D. Keasling − *Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, California 94608, United States; Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Chemical and Biomolecular Engineering and Department of Bioengineering, University of California, Berkeley, Berkeley, California 94720, United States; Center for Synthetic Biochemistry, Shenzhen Institutes for Advanced Technologies, Shenzhen 518055, P.R. China; The Novo Nordisk Foundation Center for Biosustainability, Technical University Denmark, Kongens Lyngby 2800, Denmark;* orcid.org/0000-0003-4170-6088; Email: jdkeasling@lbl.gov

Authors

Alberto A. Nava − *Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, California 94608, United States; Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, California 94720, United States*

Jacob Roberts − *Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, California 94608, United States; Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Bioengineering, University of California, Berkeley, Berkeley, California 94720, United States*

https://doi.org/10.1021/acssynbio.3c00282
ACS Synth. Biol. 2023, 12, 3148−3155

Robert W. Haushalter − *Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, California 94608, United States; Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

Zilong Wang − *Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, California 94608, United States; Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.3c00282

Author Contributions

A.A.N. and J.R. performed structural modeling and analysis. All authors contributed to the writing and editing of the manuscript.

Notes

The authors declare the following competing financial interest(s): J.D.K. has financial interests in Amyris, Ansa Biotechnologies, Apertor Pharma, Berkeley Yeast, Demetrix, Lygos, Napigen, ResVita Bio, and Zero Acre Farms.

■ ACKNOWLEDGMENTS

This work was part of the DOE Joint BioEnergy Institute (https://www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research through Contract [DE-AC02-05CH11231] between Lawrence Berkeley National Laboratory and the U.S. Department of Energy, the DOE Distinguished Scientist Fellow Program to J.D.K., and a grant (EE0008926) from the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office. J.B.R. was supported in part by a fellowship award under Contract [FA9550-21-F-0003] through the National Defense Science and Engineering Graduate (NDSEG) Fellowship Program, sponsored by the Air Force Research Laboratory (AFRL), the Office of Naval Research (ONR), and the Army Research Office (ARO). A.A.N. was supported by a National Science Foundation Graduate Research Fellowship, fellow ID [2018253421]. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

■ REFERENCES

(1) Corey, E. J. Robert Robinson Lecture. Retrosynthetic Thinking—essentials and Examples. *Chem. Soc. Rev.* **1988**, *17* (0), 111−133.

(2) Bachmann, B. O. Biosynthesis: Is It Time to Go Retro? *Nat. Chem. Biol.* **2010**, *6* (6), 390−393.

(3) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. An Unusually Large Multifunctional Polypeptide in the Erythromycin-Producing Polyketide Synthase of Saccharopolyspora Erythraea. *Nature* **1990**, *348* (6297), 176−178.

(4) Donadio, S.; Staver, M.; McAlpine, J.; Swanson, S.; Katz, L. Modular Organization of Genes Required for Complex Polyketide Biosynthesis. *Science* **1991**, *252* (5006), 675−679.

(5) Risdian, C.; Mozef, T.; Wink, J. Biosynthesis of Polyketides in Streptomyces. *Microorganisms* **2019**, *7*, 124.

(6) Grininger, M. Enzymology of Assembly Line Synthesis by Modular Polyketide Synthases. *Nat. Chem. Biol.* **2023**, *19* (4), 401−415.

(7) Maier, T.; Leibundgut, M.; Ban, N. The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* **2008**, *321* (5894), 1315−1322.

(8) Sherman, D. H.; Smith, J. L. Clearing the Skies over Modular Polyketide Synthases. *ACS Chem. Biol.* **2006**, *1* (8), 505−509.

(9) Weissman, K. J. Genetic Engineering of Modular PKSs: From Combinatorial Biosynthesis to Synthetic Biology. *Nat. Prod. Rep.* **2016**, *33* (2), 203−230.

(10) Barajas, J. F.; Blake-Hedges, J. M.; Bailey, C. B.; Curran, S.; Keasling, J. D. Engineered Polyketides: Synergy between Protein and Host Level Engineering. *Synthetic and Systems Biotechnology* **2017**, *2* (3), 147−166.

(11) Liu, Y.; Cruz-Morales, P.; Zargar, A.; Belcher, M. S.; Pang, B.; Englund, E.; Dan, Q.; Yin, K.; Keasling, J. D. Biofuels for a Sustainable Future. *Cell* **2021**, *184* (6), 1636−1647.

(12) Keatinge-Clay, A. T. Stereocontrol within Polyketide Assembly Lines. *Nat. Prod. Rep.* **2016**, *33* (2), 141−149.

(13) Wlodek, A.; Kendrew, S. G.; Coates, N. J.; Hold, A.; Pogwizd, J.; Rudder, S.; Sheehan, L. S.; Higginbotham, S. J.; Stanley-Smith, A. E.; Warneck, T.; Nur-E-Alam, M.; Radzom, M.; Martin, C. J.; Overvoorde, L.; Samborskyy, M.; Alt, S.; Heine, D.; Carter, G. T.; Graziani, E. I.; Koehn, F. E.; McDonald, L.; Alanine, A.; Rodriguez Sarmiento, R. M.; Chao, S. K.; Ratni, H.; Steward, L.; Norville, I. H.; Sarkar-Tyson, M.; Moss, S. J.; Leadlay, P. F.; Wilkinson, B.; Gregory, M. A. Diversity Oriented Biosynthesis via Accelerated Evolution of Modular Gene Clusters. *Nat. Commun.* **2017**, *8* (1), 1206.

(14) Hagen, A.; Poust, S.; Rond, T. de; Fortman, J. L.; Katz, L.; Petzold, C. J.; Keasling, J. D. Engineering a Polyketide Synthase for In Vitro Production of Adipic Acid. *ACS Synth. Biol.* **2016**, *5* (1), 21−27.

(15) Menzella, H. G.; Carney, J. R.; Li, Y.; Santi, D. V. Using Chemobiosynthesis and Synthetic Mini-Polyketide Synthases To Produce Pharmaceutical Intermediates in Escherichia Coli. *Appl. Environ. Microbiol.* **2010**, *76* (15), 5221−5227.

(16) Yuzawa, S.; Mirsiaghi, M.; Jocic, R.; Fujii, T.; Masson, F.; Benites, V. T.; Baidoo, E. E. K.; Sundstrom, E.; Tanjore, D.; Pray, T. R.; George, A.; Davis, R. W.; Gladden, J. M.; Simmons, B. A.; Katz, L.; Keasling, J. D. Short-Chain Ketone Production by Engineered Polyketide Synthases in Streptomyces Albus. *Nat. Commun.* **2018**, *9* (1), 4569.

(17) Klaus, M.; Ostrowski, M. P.; Austerjost, J.; Robbins, T.; Lowry, B.; Cane, D. E.; Khosla, C. Protein-Protein Interactions, Not Substrate Recognition, Dominate the Turnover of Chimeric Assembly Line Polyketide Synthases. *J. Biol. Chem.* **2016**, *291* (31), 16404−16415.

(18) Klaus, M.; Grininger, M. Engineering Strategies for Rational Polyketide Synthase Design. *Nat. Prod. Rep.* **2018**, *35* (10), 1070−1081.

(19) Annaval, T.; Paris, C.; Leadlay, P. F.; Jacob, C.; Weissman, K. J. Evaluating Ketoreductase Exchanges as a Means of Rationally Altering Polyketide Stereochemistry. *Chembiochem* **2015**, *16* (9), 1357−1364.

(20) Bagde, S. R.; Mathews, I. I.; Fromme, J. C.; Kim, C.-Y. Modular polyketide synthase contains two reaction chambers that operate asynchronously. *Science* **2021**, *374* (6568), 723−729.

(21) Cogan, D. P.; Zhang, K.; Li, X.; Li, S.; Pintilie, G. D.; Roh, S.-H.; Craik, C. S.; Chiu, W.; Khosla, C. Mapping the catalytic conformations of an assembly-line polyketide synthase module. *Science* **2021**, *374* (6568), 729−734.

(22) Zhang, L.; Hashimoto, T.; Qin, B.; Hashimoto, J.; Kozone, I.; Kawahara, T.; Okada, M.; Awakawa, T.; Ito, T.; Asakawa, Y.; Ueki, M.; Takahashi, S.; Osada, H.; Wakimoto, T.; Ikeda, H.; Shin-ya, K.; Abe, I. Characterization of Giant Modular PKSs Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides. *Angew. Chem., Int. Ed.* **2017**, *56* (7), 1740−1745.

(23) Hirsch, M.; Fitzgerald, B. J.; Keatinge-Clay, A. T. How Cis-Acyltransferase Assembly-Line Ketosynthases Gatekeep for Processed Polyketide Intermediates. *ACS Chem. Biol.* **2021**, *16* (11), 2515−2526.
(24) Miyazawa, T.; Hirsch, M.; Zhang, Z.; Keatinge-Clay, A. T. An in vitro platform for engineering and harnessing modular polyketide synthases. *Nat. Commun.* **2020**, *11* (1), 80.
(25) Keatinge-Clay, A. T. Polyketide Synthase Modules Redefined. *Angew. Chem., Int. Ed.* **2017**, *56* (17), 4658−4660.
(26) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.
(27) Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; Millán, C.; Park, H.; Adams, C.; Glassman, C. R.; DeGiovanni, A.; Pereira, J. H.; Rodrigues, A. V.; van Dijk, A. A.; Ebrecht, A. C.; Opperman, D. J.; Sagmeister, T.; Buhlheller, C.; Pavkov-Keller, T.; Rathinaswamy, M. K.; Dalwadi, U.; Yip, C. K.; Burke, J. E.; Garcia, K. C.; Grishin, N. V.; Adams, P. D.; Read, R. J.; Baker, D. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* **2021**, *373*, 871−876.
(28) Humphreys, I. R.; Pei, J.; Baek, M.; Krishnakumar, A.; Anishchenko, I.; Ovchinnikov, S.; Zhang, J.; Ness, T. J.; Banjade, S.; Bagde, S. R.; Stancheva, V. G.; Li, X.-H.; Liu, K.; Zheng, Z.; Barrero, D. J.; Roy, U.; Kuper, J.; Fernández, I. S.; Szakal, B.; Branzei, D.; Rizo, J.; Kisker, C.; Greene, E. C.; Biggins, S.; Keeney, S.; Miller, E. A.; Fromme, J. C.; Hendrickson, T. L.; Cong, Q.; Baker, D. Computed Structures of Core Eukaryotic Protein Complexes. *Science* **2021**, *374* (6573), No. eabm4805.
(29) Buchan, D. W. A.; Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 Years on. *Nucleic Acids Res.* **2019**, *47* (W1), W402−W407.
(30) Sherman, D. H. The Lego-ization of polyketide biosynthesis. *Nat. Biotechnol.* **2005**, *23* (9), 1083−1084.
(31) Menzella, H. G.; Reid, R.; Carney, J. R.; Chandran, S. S.; Reisinger, S. J.; Patel, K. G.; Hopwood, D. A.; Santi, D. V. Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **2005**, *23* (9), 1171−1176.
(32) Chandran, S. S.; Menzella, H. G.; Carney, J. R.; Santi, D. V. Activating Hybrid Modular Interfaces in Synthetic Polyketide Synthases by Cassette Replacement of Ketosynthase Domains. *Chem. Biol.* **2006**, *13* (5), 469−474.
(33) Klaus, M.; Buyachuihan, L.; Grininger, M. Ketosynthase Domain Constrains the Design of Polyketide Synthases. *ACS Chem. Biol.* **2020**, *15* (9), 2422−2432.
(34) Su, L.; Hôtel, L.; Paris, C.; Chepkirui, C.; Brachmann, A. O.; Piel, J.; Jacob, C.; Aigle, B.; Weissman, K. J. Engineering the stambomycin modular polyketide synthase yields 37-membered mini-stambomycins. *Nat. Commun.* **2022**, *13* (1), 515.
(35) Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C.-P.; Bode, H. B. De novo design and engineering of non-ribosomal peptide synthetases. *Nat. Chem.* **2018**, *10* (3), 275−281.
(36) Bozhüyük, K. A. J.; Linck, A.; Tietze, A.; Kranz, J.; Wesche, F.; Nowak, S.; Fleischhacker, F.; Shi, Y.-N.; Grün, P.; Bode, H. B. Modification and de novo design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nat. Chem.* **2019**, *11* (7), 653−661.
(37) Khosla, C.; Tang, Y.; Chen, A. Y.; Schnarr, N. A.; Cane, D. E. Structure and Mechanism of the 6-Deoxyerythronolide B Synthase. *Annu. Rev. Biochem.* **2007**, *76*, 195−221.
(38) Vander Wood, D. A.; Keatinge-Clay, A. T. The Modules of Trans-Acyltransferase Assembly Lines Redefined with a Central Acyl Carrier Protein. *Proteins* **2018**, *86* (6), 664−675.
(39) Nguyen, T.; Ishida, K.; Jenke-Kodama, H.; Dittmann, E.; Gurgui, C.; Hochmuth, T.; Taudien, S.; Platzer, M.; Hertweck, C.; Piel, J. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat. Biotechnol.* **2008**, *26* (2), 225−233.
(40) Helfrich, E. J. N.; Piel, J. Biosynthesis of polyketides by trans -AT polyketide synthases. *Nat. Prod. Rep.* **2016**, *33* (2), 231−316.
(41) Nivina, A.; Yuet, K. P.; Hsu, J.; Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chem. Rev.* **2019**, *119* (24), 12524−12547.
(42) Nivina, A.; Herrera Paredes, S.; Fraser, H. B.; Khosla, C. GRINS: Genetic Elements That Recode Assembly-Line Polyketide Synthases and Accelerate Their Diversification. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (26), e2100751118.
(43) Anzai, Y.; Saito, N.; Tanaka, M.; Kinoshita, K.; Koyama, Y.; Kato, F. Organization of the biosynthetic gene cluster for the polyketide macrolide mycinamicin in Micromonospora griseorubida. *FEMS Microbiol. Lett.* **2003**, *218* (1), 135−141.
(44) Stinear, T. P.; Mve-Obiang, A.; Small, P. L. C.; Frigui, W.; Pryor, M. J.; Brosch, R.; Jenkin, G. A.; Johnson, P. D. R.; Davies, J. K.; Lee, R. E.; Adusumilli, S.; Garnier, T.; Haydock, S. F.; Leadlay, P. F.; Cole, S. T. Giant plasmid-encoded polyketide synthases produce the macrolide toxin of Mycobacterium ulcerans. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (5), 1345−1349.
(45) Miyazawa, T.; Fitzgerald, B. J.; Keatinge-Clay, A. T. Preparative production of an enantiomeric pair by engineered polyketide synthases. *Chem. Commun.* **2021**, *57* (70), 8762−8765.
(46) Chemler, J. A.; Tripathi, A.; Hansen, D. A.; O'Neil-Johnson, M.; Williams, R. B.; Starks, C.; Park, S. R.; Sherman, D. H. Evolution of Efficient Modular Polyketide Synthases by Homologous Recombination. *J. Am. Chem. Soc.* **2015**, *137* (33), 10603−10609.
(47) Peng, H.; Ishida, K.; Sugimoto, Y.; Jenke-Kodama, H.; Hertweck, C. Emulating evolutionary processes to morph aureothin-type modular polyketide synthases and associated oxygenases. *Nat. Commun.* **2019**, *10* (1), 3918.
(48) Englund, E.; Schmidt, M.; Nava, A. A.; Klass, S.; Keiser, L.; Dan, Q.; Katz, L.; Yuzawa, S.; Keasling, J. D. Biosensor Guided Polyketide Synthases Engineering for Optimization of Domain Exchange Boundaries. *Nat. Commun.* **2023**, *14* (1), 4871.
(49) Whicher, J. R.; Smaga, S. S.; Hansen, D. A.; Brown, W. C.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. Cyanobacterial Polyketide Synthase Docking Domains: A Tool for Engineering Natural Product Biosynthesis. *Chem. Biol.* **2013**, *20* (11), 1340−1351.
(50) Broadhurst, R. W.; Nietlispach, D.; Wheatcroft, M. P.; Leadlay, P. F.; Weissman, K. J. The Structure of Docking Domains in Modular Polyketide Synthases. *Chem. Biol.* **2003**, *10* (8), 723−731.
(51) Buchholz, T. J.; Geders, T. W.; Bartley, F. E.; Reynolds, K. A.; Smith, J. L.; Sherman, D. H. Structural Basis for Binding Specificity between Subclasses of Modular Polyketide Synthase Docking Domains. *ACS Chem. Biol.* **2009**, *4* (1), 41−52.
(52) Park, H.; Kevany, B. M.; Dyer, D. H.; Thomas, M. G.; Forest, K. T. A Polyketide Synthase Acyltransferase Domain Structure Suggests a Recognition Mechanism for Its Hydroxymalonyl-Acyl Carrier Protein Substrate. *PLoS One* **2014**, *9* (10), No. e110965.
(53) Liew, C. W.; Nilsson, M.; Chen, M. W.; Sun, H.; Cornvik, T.; Liang, Z.-X.; Lescar, J. Crystal Structure of the Acyltransferase Domain of the Iterative Polyketide Synthase in Enediyne Biosynthesis. *J. Biol. Chem.* **2012**, *287* (27), 23203−23215.
(54) Tang, Y.; Chen, A. Y.; Kim, C.-Y.; Cane, D. E.; Khosla, C. Structural and Mechanistic Analysis of Protein Interactions in Module 3 of the 6-Deoxyerythronolide B Synthase. *Chem. Biol.* **2007**, *14* (8), 931−943.
(55) Tang, Y.; Kim, C.-Y.; Mathews, I. I.; Cane, D. E.; Khosla, C. The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (30), 11124−11129.

https://doi.org/10.1021/acssynbio.3c00282
ACS Synth. Biol. 2023, 12, 3148−3155

(56) Keatinge-Clay, A. Crystal Structure of the Erythromycin Polyketide Synthase Dehydratase. *J. Mol. Biol.* **2008**, *384* (4), 941−953.
(57) Barajas, J. F.; McAndrew, R. P.; Thompson, M. G.; Backman, T. W. H.; Pang, B.; de Rond, T.; Pereira, J. H.; Benites, V. T.; Martín, H. G.; Baidoo, E. E. K.; Hillson, N. J.; Adams, P. D.; Keasling, J. D. Structural Insights into Dehydratase Substrate Selection for the Borrelidin and Fluvirucin Polyketide Synthases. *J. Ind. Microbiol. Biotechnol.* **2019**, *46* (8), 1225−1235.
(58) Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Clay, A. T. Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nat. Chem. Biol.* **2012**, *8* (7), 615−621.
(59) Curran, S. C.; Pereira, J. H.; Baluyot, M.-J.; Lake, J.; Puetz, H.; Rosenburg, D. J.; Adams, P.; Keasling, J. D. Structure and Function of BorB, the Type II Thioesterase from the Borrelidin Biosynthetic Gene Cluster. *Biochemistry* **2020**, *59* (16), 1630−1639.
(60) Tsai, S. C.; Miercke, L. J.; Krucinski, J.; Gokhale, R.; Chen, J. C.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. Crystal Structure of the Macrocycle-Forming Thioesterase Domain of the Erythromycin Polyketide Synthase: Versatility from a Unique Substrate Channel. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (26), 14808−14813.
(61) Liu, C.; Yuan, M.; Xu, X.; Wang, L.; Keatinge-Clay, A. T.; Deng, Z.; Lin, S.; Zheng, J. Substrate-Bound Structures of a Ketoreductase from Amphotericin Modular Polyketide Synthase. *J. Struct. Biol.* **2018**, *203* (2), 135−141.
(62) Edwards, A. L.; Matsui, T.; Weiss, T. M.; Khosla, C. Architectures of Whole-Module and Bimodular Proteins from the 6-Deoxyerythronolide B Synthase. *J. Mol. Biol.* **2014**, *426*, 2229−2245.
(63) Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structure of a modular polyketide synthase. *Nature* **2014**, *510* (7506), 512−517.
(64) Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **2014**, *510* (7506), 560−564.
(65) Klaus, M.; Rossini, E.; Linden, A.; Paithankar, K. S.; Zeug, M.; Ignatova, Z.; Urlaub, H.; Khosla, C.; Köfinger, J.; Hummer, G.; Grininger, M. Solution Structure and Conformational Flexibility of a Polyketide Synthase Module. *JACS Au* **2021**, *1* (12), 2162−2171.
(66) Kalkreuter, E.; Bingham, K. S.; Keeler, A. M.; Lowell, A. N.; Schmidt, J. J.; Sherman, D. H.; Williams, G. J. Computationally-Guided Exchange of Substrate Selectivity Motifs in a Modular Polyketide Synthase Acyltransferase. *Nat. Commun.* **2021**, *12* (1), 2193.
(67) Milligan, J. C.; Lee, D. J.; Jackson, D. R.; Schaub, A. J.; Beld, J.; Barajas, J. F.; Hale, J. J.; Luo, R.; Burkart, M. D.; Tsai, S.-C. Molecular Basis for Interactions between an Acyl Carrier Protein and a Ketosynthase. *Nat. Chem. Biol.* **2019**, *15* (7), 669−671.
(68) Beld, J.; Cang, H.; Burkart, M. D. Visualizing the Chain-Flipping Mechanism in Fatty-Acid Biosynthesis. *Angew. Chem. Weinheim Bergstr. Ger.* **2014**, *126* (52), 14684−14689.
(69) Chisuga, T.; Nagai, A.; Miyanaga, A.; Goto, E.; Kishikawa, K.; Kudo, F.; Eguchi, T. Structural Insight into the Reaction Mechanism of Ketosynthase-Like Decarboxylase in a Loading Module of Modular Polyketide Synthases. *ACS Chem. Biol.* **2022**, *17* (1), 198−206.
(70) Dickinson, M. S.; Miyazawa, T.; McCool, R. S.; Keatinge-Clay, A. T. Priming Enzymes from the Pikromycin Synthase Reveal How Assembly-Line Ketosynthases Catalyze Carbon-Carbon Chemistry. *Structure* **2022**, *30* (9), 1331−1339.e3.
(71) Kapur, S.; Chen, A. Y.; Cane, D. E.; Khosla, C. Molecular Recognition between Ketosynthase and Acyl Carrier Protein Domains of the 6-Deoxyerythronolide B Synthase. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (51), 22066−22071.
(72) Patel, K. D.; MacDonald, M. R.; Ahmed, S. F.; Singh, J.; Gulick, A. M. Structural Advances toward Understanding the Catalytic Activity and Conformational Dynamics of Modular Nonribosomal Peptide Synthetases. *Nat. Prod. Rep.* **2023**, *40*, 1550.
(73) Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.; Blackwell, S.; Yim, J.; Ronneberger, O.; Bodenstein, S.; Zielinski, M.; Bridgland, A.; Potapenko, A.; Cowie, A.; Tunyasuvunakool, K.; Jain, R.; Clancy, E.; Kohli, P.; Jumper, J.; Hassabis, D. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2021, 2021.10.04.463034 (accessed Sept 14, 2023). DOI: 10.1101/2021.10.04.463034
(74) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: making protein folding accessible to all. *Nat. Methods* **2022**, *19* (6), 679−682.
(75) Anishchenko, I.; Pellock, S. J.; Chidyausiku, T. M.; Ramelot, T. A.; Ovchinnikov, S.; Hao, J.; Bafna, K.; Norn, C.; Kang, A.; Bera, A. K.; DiMaio, F.; Carter, L.; Chow, C. M.; Montelione, G. T.; Baker, D. De Novo Protein Design by Deep Network Hallucination. *Nature* **2021**, *600* (7889), 547−552.
(76) Shroff, R.; Cole, A. W.; Diaz, D. J.; Morrow, B. R.; Donnell, I.; Annapareddy, A.; Gollihar, J.; Ellington, A. D.; Thyer, R. Discovery of Novel Gain-of-Function Mutations Guided by Structure-Based Deep Learning. *ACS Synth. Biol.* **2020**, *9* (11), 2927−2935.
(77) Lu, H.; Diaz, D. J.; Czarnecki, N. J.; Zhu, C.; Kim, W.; Shroff, R.; Acosta, D. J.; Alexander, B. R.; Cole, H. O.; Zhang, Y.; Lynd, N. A.; Ellington, A. D.; Alper, H. S. Machine Learning-Aided Engineering of Hydrolases for PET Depolymerization. *Nature* **2022**, *604* (7907), 662−667.
(78) Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. Evolutionary-Scale Prediction of Atomic Level Protein Structure with a Language Model. *Science* **2023**, *379*, 1123.
(79) Madani, A.; Krause, B.; Greene, E. R.; Subramanian, S.; Mohr, B. P.; Holton, J. M.; Olmos, J. L.; Xiong, C.; Sun, Z. Z.; Socher, R.; Fraser, J. S.; Naik, N. Large Language Models Generate Functional Protein Sequences across Diverse Families. *Nat. Biotechnol.* **2023**, *41*, 1099.
(80) Roberts, J. B.; Nava, A. A.; Pearson, A. N.; Incha, M. R.; Valencia, L. E.; Ma, M.; Rao, A.; Keasling, J. D. Foldy: A Web Application for Interactive Protein Structure Analysis. *bioRxiv*, 2023, 2023.05.11.540333 (accessed Sept 14, 2023). DOI: 10.1101/2023.05.11.540333.
(81) Buyachuihan, L.; Zhao, Y.; Schelhas, C.; Grininger, M. Docking Domain Engineering in a Modular Polyketide Synthase and Its Impact on Structure and Function. *ACS Chem. Biol.* **2023**, *18* (7), 1500−1509.
(82) Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; van Wezel, G. P.; Medema, M. H.; Weber, T. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **2021**, *49*, W29−W35.
(83) Corso, G.; Stärk, H.; Jing, B.; Barzilay, R.; Jaakkola, T. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. *arXiv [q-bio.BM]*, 2022, 2210.01776, DOI: 10.48550/arXiv.2210.01776