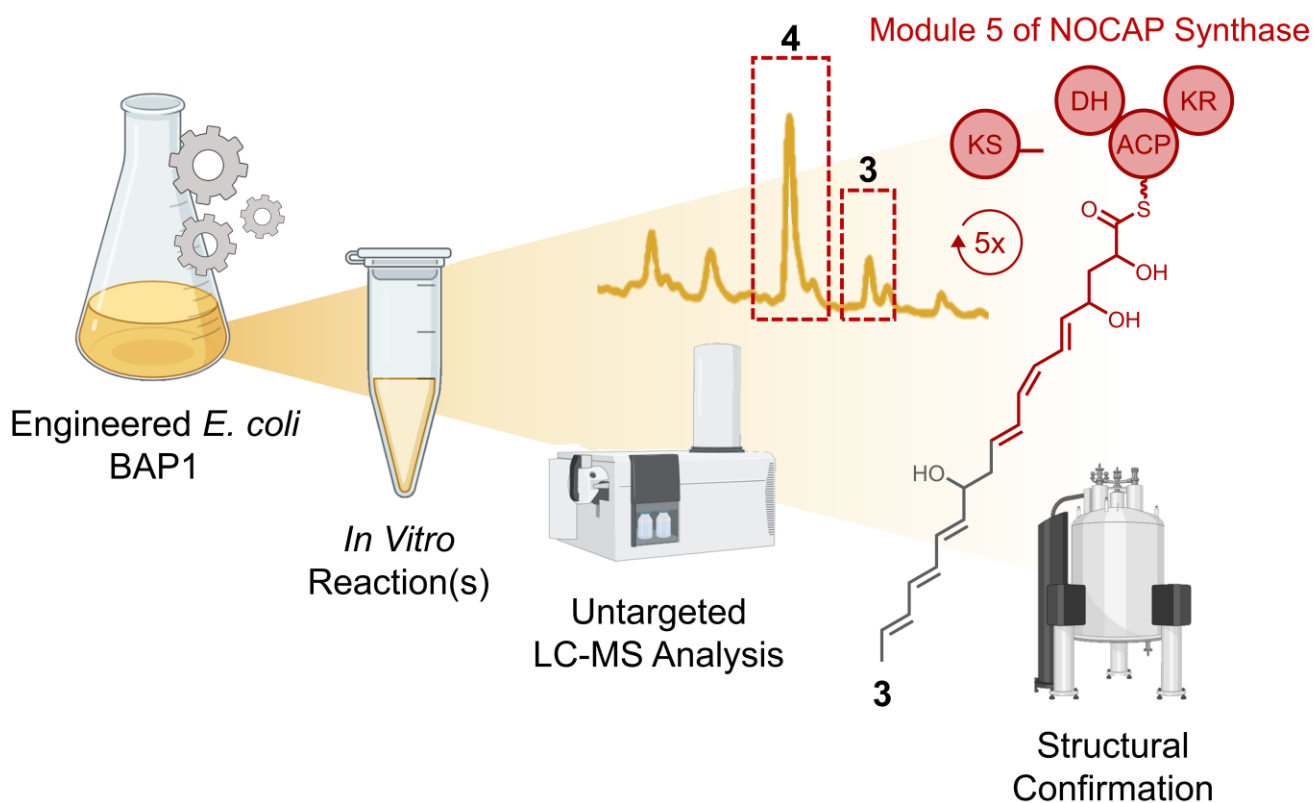


Analysis of a “Split-and-Stuttering” Module of an Assembly Line Polyketide Synthase

Katarina M. Guzman[†], Kai P. Yuet[‡], Stephen R. Lynch^{‡,⊥}, Corey W. Liu^{§,^,⊥}, and Chaitan Khosla^{†,‡,^,*}

[†]Department of Chemical Engineering, [‡]Department of Chemistry, [§]Department of Structural Biology, [^]Stanford ChEM-H, Stanford University, Stanford, CA 94305, *corresponding author



TOC/Abstract Graphic

Abstract

Notwithstanding the “one-module-one-elongation-cycle” paradigm of assembly line polyketide synthases (PKSs), some PKSs harbor modules that iteratively elongate their substrates through a defined number of cycles. While some insights into module iteration, also referred to as “stuttering”, have been derived through *in vivo* and *in vitro* analysis of a few PKS modules, a general understanding of the mechanistic principles underlying module iteration remains elusive. This report serves as the first interrogation of a stuttering module from a *trans*-AT subfamily PKS that is also naturally split across two polypeptides. Previous work has shown that Module 5 of

the NOCAP (**n**ocardiosis **a**ssociated **p**olyketide) synthase iterates precisely three times in the biosynthesis of its polyketide product, resulting in an all *trans*-configured triene moiety in the polyketide product. Here we describe the intrinsic catalytic properties of this NOCAP synthase module. Through complementary experiments *in vitro* and in *E. coli*, the “split-and-stuttering” module was shown to catalyze up to five elongation cycles, although its dehydratase domain ceased to function after three cycles. Unexpectedly, the central olefinic group of this truncated product had a *cis* configuration. Our findings set the stage for further in-depth analysis of a structurally and functionally unusual PKS module with contextual biosynthetic plasticity.

Main

Polyketide synthases (PKSs) are multifunctional enzymes responsible for the biosynthesis of structurally diverse polyketide products, many of which are clinically used as medicines (1). Therefore, understanding their catalytic diversity and mechanisms has been of high interest to researchers in the hope of engineering novel pharmacologically useful molecules. One class of PKSs, called assembly line PKSs, consists of multiple catalytic modules each of which is responsible for a well-defined set of chemical transformations on the growing polyketide chain. Minimally a PKS module harbors a ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP) domain (2). Some PKS modules (hereafter designated *cis*-AT PKSs) include the AT as an integral domain of a multifunctional polypeptide, whereas other modules (designated *trans*-AT PKSs) engage stand-alone AT proteins (3).

Canonically, each module of an assembly line PKS catalyzes a single elongation and modification cycle, where after the growing polyketide chain is either translocated onto the KS domain of the next module or it is off-loaded from the assembly line (4, 5). Such a collinear architecture has facilitated prediction of structural features of the resulting polyketide products (6). However, not all PKSs operate according to a collinear one-module-one-elongation-cycle model (3). For example, the NOCAP synthase (7, 8) harbors a *trans*-AT module (Module 5) that catalyzes three successive elongation and modification cycles on the growing polyketide chain yielding products **1** and **2** (**Figure 1**) (9,10). After each cycle, the elongated and modified polyketide product is passed from the ACP domain of Module 5 back to its own KS domain (11). This type of programmed module iteration, also known as “stuttering” (12, 13), has been identified in a few other assembly line PKSs (14) such as the stigmatellin (15), lankacidin (16), borrelidin (17), neo-aureothin and aureothin (18) PKS pathways.

While the governing principles that control programmed iteration are not well understood, some insights have been gained through *in vitro* and *in vivo* analysis of *cis*-AT PKSs. Module 1 of the aureothin synthase (*cis*-AT PKS), which catalyzes two elongation and modification cycles in the context of the complete assembly line PKS, was expressed in *Streptomyces lividans* (19). The resulting strain produced a spectrum of compounds that had undergone up to 4 rounds of elongation, although products resulting from 3 and 4 rounds of iteration were orders of magnitude less abundant. Additionally, an *in vitro* analysis of Module 5 of the borrelidin PKS revealed that the KS domain of this stuttering module accepted non-native substrates corresponding to the

chain lengths of its natural intermediates (20). Collectively, these and other findings have led to a mechanistic model in which the KS specificity at least partly dictates the number of iterations observed. However, the downstream module also has been suggested to exhibit a gatekeeping role (21). Given the lack of prior studies on stuttering modules from *trans*-AT PKSs, we interrogated Module 5 of the NOCAP synthase. Of particular interest to us was its additionally unusual “split-and-stuttering” feature, with its KS and ACP domains distributed across separate proteins (**Figure 1**).

In order to investigate the potential gatekeeping role of downstream modules in the NOCAP synthase we sought to develop a truncated assembly line excluding Modules 6-8-TR. Fortunately, a system for functionally expressing the entire NOCAP synthase in *E. coli* BAP1 (22) was recently developed which was adapted for the present study (10). Three plasmids with unique antibiotic resistance markers were introduced into *E. coli* BAP1 (**Figure 2, Figure S1**). Plasmid pCK-KPY178 encodes the malonyl-CoA synthetase MatB from *Streptomyces coelicolor* (23), the *Rhizobium leguminosarum* malonate carrier protein MatC (24), the loading module (Module L) and *trans*-AT-TEII (*tAT*-TEII) of the NOCAP synthase. Plasmid pCK-KPY222 encodes Modules 1-3 with engineered docking domains to facilitate protein-protein interactions (25). Finally, pKMG14 or alternatively pKMG15 encodes Module 4 and the DH-ACP-KR or the DH-ACP domains of Module 5, respectively. In the absence of a specific product release mechanism in these truncated assembly lines, we relied on spontaneous hydrolysis of polyketide products from ACP₅.

One to three liters of *E. coli* BAP1 [pCK-KPY178/pCK-KPY222/pKMG14] supernatant was subjected to solid phase extraction (SPE) followed by high performance liquid chromatography (HPLC) (**Figure S2**). Encouragingly, two compounds **3** and **4** with distinct UV signatures were identified (**Figure 2, Figures S3-S4**). Control strains harboring pKMG15 in lieu of pKMG14 or lacking either plasmid failed to produce either **3** or **4** (**Figure S5**). High-resolution mass spectrometry (HRMS) analysis revealed that **3** had a molecular formula of C₂₀H₂₈O₅ (*m/z*: [M-H]⁻ calcd. 347.1864, obsd. 347.1845, 5.5 ppm), while **4** had a molecular formula of C₁₈H₂₆O₅ (*m/z*: [M-H]⁻ calcd. 321.1707, obsd. 321.1704, 0.9 ppm) (**Figure S6**). Similar to the relationship between **1** and **2**, these predicted formulas differed by a mass shift consistent with an unsaturated two-carbon (C₂H₂) moiety. To verify the anticipated biosynthetic relationship between **3** and **4**, an *E. coli* BAP1 strain lacking Module 3 [pCK-KPY178/pCK-KPY102/pKMG14] was engineered (**SI Methods**). The resulting strain yielded **4** but not **3** (**Figure S5**), confirming a requirement of Module 3 for the biosynthesis of the latter compound. However, because these HRMS values were inconsistent with our expectation that Module 5 catalyzed three elongation and modification cycles, we turned to NMR spectroscopy to elucidate the structures of **3** and **4**.

The structures of **3** and **4** were uncovered by extensive 1-D and 2-D NMR analysis, including ¹H 1D, ¹³C 1D, COSY, HSQC, HMBC, and ROESY experiments (**Figure 3, Figures S7-S31, Table 1**). The structures deduced from NMR analysis suggested that Module 5 catalyzed five elongation cycles in the biosynthesis of these compounds, but that the final two cycles were not accompanied by dehydration of the β-hydroxyacyl intermediate (**Figure 4**,

Figure S32). Furthermore, the presence of a cross-peak in the ROESY spectrum implied the existence of a *cis*-olefinic bond between C₈-C₉ in contrast to the all-*trans* configuration of **1** and **2** (**Figures S28 and S31**). To our knowledge this is the first reported instance of “imperfect stuttering”, where the programmed sequence of chain modification reactions is not faithfully catalyzed in every iteration. The unexpected chain length and configuration of **3** (and **4**) raised the possibility that endogenous *E. coli* enzymes played a role in its biosynthesis. For instance, FabA from the fatty acid synthase of *E. coli* iteratively catalyzes the dehydration of ACP-bound β -hydroxyacyl intermediates generated during fatty acid biosynthesis. FabA can also reversibly isomerize *trans*-2-decenoyl-ACP to *cis*-3-decenoyl-ACP (26). Therefore, to rule out the possibility of artifactual involvement of host enzymes, we sought to functionally reconstitute the truncated hexamodular derivative of the NOCAP synthase *in vitro* using purified proteins.

The following proteins were expressed and purified *via* immobilized metal ion chromatography followed by anion exchange chromatography: DH₅-ACP₅-KR₅, DH₅-ACP₅ and *t*AT-TEII from the NOCAP synthase, and *S. coelicolor* MatB (**Figure S33, SI Methods**). Additionally, Module L, Modules 1-2, Module 3 and Module 4-KS₅ were purified on a StrepII-Tag affinity column. Modules L, 1-2, 3, 4-KS₅ and DH₅-ACP₅-KR₅ were incubated in the presence of *t*AT-TEII, MatB, NADPH, S-adenosyl methionine (SAM), malonate, ATP, and CoASH. Parallel reactions were performed without malonate or with [2-¹³C]-, [1,3-¹³C₂]-, or [¹³C₃]-malonate to ensure the products detected were of polyketide origin. Product formation relied on spontaneous hydrolysis from ACP₅. The predominant product of the truncated hexamodular PKS *in vitro* was identical to **3**, as judged by HRMS (**Figure 5**). The observation of +10, +10, and +20 mass shifts in reactions containing [2-¹³C]-, [1,3-¹³C₂]-, and [¹³C₃]-malonate, respectively, supported our hypothesis that *in vitro*-derived **3** a) originates from the truncated assembly line and b) had undergone five rounds of elongation at Module 5 (**Figure S34**). **4** was not synthesized in detectable amounts *in vitro*. These results largely support our proposed biosynthetic pathway. However, despite considerable effort, these *in vitro* reactions did not yield sufficient material for ROESY verification of **3**'s geometry.

In summary, this is the first description of a *trans*-AT “split-and-stuttering” module that preferentially iterates for additional cycles beyond those observed in the context of the full PKS assembly line. We therefore propose that Module 6 of the NOCAP synthase regulates the overall polyketide chain length by preferentially recognizing the ACP₅-bound intermediate after it has undergone three elongation and modification cycles on Module 5. The inability of DH₅ to catalyze dehydration after three elongation cycles revealed an unprecedented feature of this PKS domain. Future efforts will aim to elucidate the underlying mechanisms that govern module 5's particularly unusual iterative behavior. Overall, our findings demonstrate the remarkable contextual biosynthetic plasticity of Module 5 of the NOCAP synthase while highlighting the opportunity for future polyketide engineering through a deeper understanding of iterative PKS modules.

Experimental Section

For detailed descriptions, see the Supporting Information.

Cloning. Genes were PCR amplified from *Nocardia* genomic DNA or codon optimized for *E. coli* and cloned into pET and Duet vectors using restriction enzyme-based methods, Gibson Assembly, or commercial assembly master mixes. Plasmids were cloned in *E. coli* DH5 α , Stellar, or *E. coli* TOP10.

***E. coli* cultures.** *E. coli* BAP1 cultures housing the appropriate plasmids and antibiotics were grown at 30°C until an OD₆₀₀ of 0.2 at which point they were induced with isopropyl- β -D-galactopyranoside (IPTG) and moved to 16°C for an additional 72 hrs.

Protein Expression and Purification. Proteins were expressed in *E. coli* BAP1 and purified by either Ni-NTA- or Strep-Tactin-based affinity chromatography followed by anion exchange chromatography.

***In Vitro* Assays and LC-MS.** Purified proteins (4-8 μ M) were incubated with appropriate cofactors for 24 hrs at 25°C. Reactions (50 μ L) were quenched with 50 μ L methanol, spun down, and the top-layer (~50 μ L) was analyzed by LC-MS (Agilent 6545 Q-TOF LC-MS system).

NMR. Spectra were acquired on a 600 MHz Varian Inova spectrometer (Stanford University) and a 900 MHz Bruker NMR spectrometer at the Central California 900 MHz NMR Facility (University of California, Berkeley). Samples were prepared in 250-750 μ L of CD₃OD.

Associated Content

The Supporting Information is available free of charge at #

Detailed methods, figures, and tables ([PDF](#))

Author Information

Corresponding Author

Chaitan Khosla – *Departments of Chemical Engineering, Chemistry, and Stanford Chem-H, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0001-6529-495X;*
Email: khosla@stanford.edu

Authors

Katarina M. Guzman – *Department of Chemical Engineering, Stanford University, Stanford, California 94305, United States*

Kai P. Yuet – *Department of Chemistry, Stanford University, Stanford, California 94305, United States*

Stephen R. Lynch – *Department of Chemistry, Stanford University, Stanford, California 94305, United States*

Corey W. Liu – *Department of Structural Biology and Stanford ChEM-H, Stanford University, Stanford, California 94305, United States*

Complete contact information is available at: <#>

Author Contributions

S.R.L.[‡] and C.W.L.[‡] contributed equally.

Notes

The authors declare no competing financial interest.

Acknowledgements

We thank members of the Khosla Lab for thoughtful discussions. We also thank Theresa McLaughlin (Stanford University Mass Spectrometry facility) and Jeffrey G. Pelton (University of California, Berkeley) for technical assistance performing LC-MS and NMR experiments respectively. This work was supported by National Institutes of Health (NIH) Grant 5R01 GM087934 (to C.K.) (with extension -26S1 (to C.K and K.M.G.)), NIH Grant F32 GM123637 (to K.P.Y.) and a Stanford Graduate Fellowship (SGF) Award (to K.M.G). This work utilized the Stanford Cancer Institute Proteomics/Mass Spectrometry Shared Resource, which is supported by NIH Grant P30 CA124435, and the 900 MHz Bruker NMR spectrometer (funded by NIH Grant P41 GM068933) at the Central California 900 MHz NMR Facility.

References

- [1] Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **2004**, 303, 1805–1810.
- [2] Katz, L. The DEBS Paradigm for Type I Modular Polyketide Synthases and Beyond, *Methods in Enzymol.* **2009**, 459, 113–142.
- [3] Gay, D. C., Gay, G., Axelrod, A. J., Jenner, M., Kohlhaas, C., Kampa, A., Oldham, N. J., Piel, J., and Keatinge-Clay, A. T. A Close Look at a Ketosynthase from a Trans-Acyltransferase Modular Polyketide Synthase. *Structure* **2014**, 22, 444–451.
- [4] Donadio, S., Staver, M., Mcalpine, J., Swanson, S., and Katz, L. Modular organization of genes required for complex polyketide biosynthesis. *Science* **1991**, 252, 675–679.

- [5] Khosla, C., Herschlag, D., Cane, D. E., and Walsh, C. T. Assembly Line Polyketide Synthases: Mechanistic Insights and Unsolved Problems. *Biochemistry* **2014**, 53, 2875–2883.
- [6] Helfrich, E. J. N., Ueoka, R., Dolev, A., Rust, M., Meoded, R. A., Bhushan, A., Califano, G., Costa, R., Gugger, M., Steinbeck, C., Moreno, P., and Piel, J. Automated structure prediction of trans-acyltransferase polyketide synthase products. *Nat. Chem. Bio.* **2019**, 15, 813–821.
- [7] Komaki, H., Ichikawa, N., Hosoyama, A., Takahashi-Nakaguchi, A., Matsuzawa, T., Suzuki, K., Fujita, N., and Gono, T. Genome based analysis of type-I polyketide synthase and nonribosomal peptide synthetase gene clusters in seven strains of five representative *Nocardia* species. *BMC Genomics* **2014**, 15, 323.
- [8] Kageyama, A., Yazawa, K., Mukai, A., Kohara, T., Nishimura, K., Kroppenstedt, R. M., and Mikami, Y. *Nocardia araoensis* sp. nov. and *Nocardia pneumoniae* sp. nov., isolated from patients in Japan. *Int. J. Syst. Evol. Microbiol.* **2004**, 54, 2025–2029.
- [9] Kuo, J., Lynch, S. R., Liu, C. W., Xiao, X., and Khosla, C. Partial In Vitro Reconstitution of an Orphan Polyketide Synthase Associated with Clinical Cases of Nocardiosis. *ACS Chem. Bio.* **2016**, 11, 2636–2641.
- [10] Yuet, K. P., Liu, C. W., Lynch, S. R., Kuo, J., Michaels, W., Lee, R. B., Mcshane, A. E., Zhong, B. L., Fischer, C. R., and Khosla, C. Complete reconstitution and deorphanization of the 3 MDa nocardiosis-associated polyketide synthase. *J. Am. Chem. Soc.* **2020**, 142 (13), 5952–5957.
- [11] Busch, B., Ueberschaar, N., Sugimoto, Y., and Hertweck, C. Interchain Retrotransfer of Aureothin Intermediates in an Iterative Polyketide Synthase Module. *J. Am. Chem. Soc.* **2012**, 134, 12382–12385.
- [12] He, J., and Hertweck, C. Iteration as Programmed Event during Polyketide Assembly; Molecular Analysis of the Aureothin Biosynthesis Gene Cluster. *Chemistry & Biology* **2003**, 10, 1225–1232.
- [13] Moss, S. J., Martin, C. J., and Wilkinson, B. Loss of collinearity by modular polyketide synthases: a mechanism for the evolution of chemical diversity. *Nat. Prod. Rep.* **2004**, 21, 575–593.
- [14] Nivina, A., Yuet, K. P., Hsu, J., and Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chemical Reviews* **2019**, 119, 12524–12547.
- [15] Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R. The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J. Biol. Chem.* **2002**, 277, 13082–13090.

- [16] Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K., and Kinashi, H. The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. *Mol. Microbiol.* **2003**, 48, 1501–1510.
- [17] Olano, C., Wilkinson, B., Sánchez, C., Moss, S. J., Sheridan, R., Math, V., Weston, A. J., Braña, A. F., Martín, C. J., Oliynyk, M., Méndez, C., Leadlay, P. F., and Salas, J. A. Biosynthesis of the angiogenesis inhibitor borrelidin by *Streptomyces parvulus* Tü4055: cluster analysis and assignment of functions. *Chem. Biol.* **2004**, 11, 87–97.
- [18] Traitcheva, N., Jenke-Kodama, H., He, J., Dittmann, E., and Hertweck, C. Non-colinear polyketide biosynthesis in the aureothin and neo-aureothin pathways: an evolutionary perspective. *ChemBioChem* **2007**, 8, 1841–1849.
- [19] Busch, B., Ueberschaar, N., Behnken, S., Sugimoto, Y., Werneburg, M., Traitcheva, N., He, J., and Hertweck, C. Multifactorial Control of Iteration Events in a Modular Polyketide Assembly Line. *Angew. Chem. Int. Ed.* **2013**, 52, 5285–5289.
- [20] Curran, S. C., Hagen, A., Poust, S., Chan, L. J. G., Garabedian, B. M., Rond, T. D., Baluyot, M.-J., Vu, J. T., Lau, A. K., Yuzawa, S., Petzold, C. J., Katz, L., and Keasling, J. D. Probing the Flexibility of an Iterative Modular Polyketide Synthase with Non-Native Substrates in Vitro. *ACS Chem. Biol.* **2018**, 13, 2261–2268.
- [21] Sugimoto, Y., Ishida, K., Traitcheva, N., Busch, B., Dahse, H.-M., and Hertweck, C. Freedom and Constraint in Engineered Noncolinear Polyketide Assembly Lines. *Chemistry & Biology* **2015**, 22, 229–240.
- [22] Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **2001**, 291, 1790–1792
- [23] Hughes, A. J.; Keatinge-Clay, A. Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* **2011**, 18, 165–176.
- [24] Lombo, F.; Pfeifer, B.; Leaf, T.; Ou, S.; Kim, Y. S.; Cane, D. E.; Licari, P.; Khosla, C. Enhancing the atom economy of polyketide biosynthetic processes through metabolic engineering. *Biotechnol. Prog.* **2001**, 17, 612–617.
- [25] 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, 16404–16415.
- [26] Leesong M, Henderson BS, Gillig JR, Schwab JM, Smith JL. Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure* **1996**, 4, 253–256.

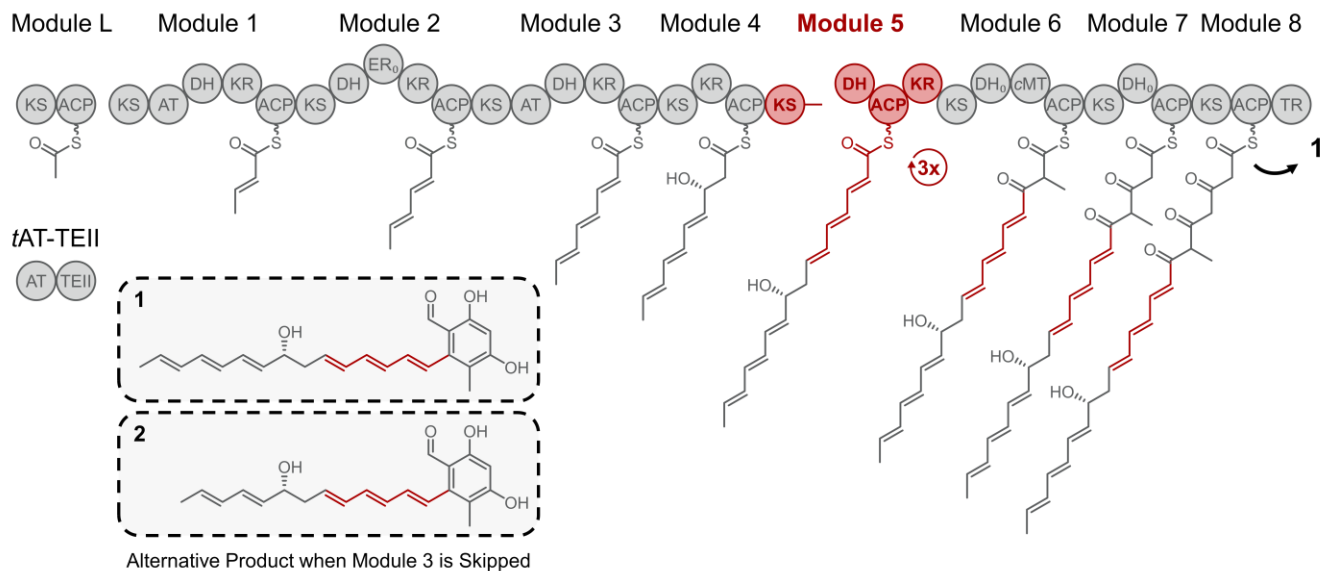


Figure 1. Biosynthetic pathway of **1** via Modules L-8-TR + tAT-TEII of NOCAP synthase. **2** is an alternative product generated when Module 3 is skipped. Key: KS, ketosynthase; AT, acyltransferase; DH, dehydratase; KR, ketoreductase; ER, enoylreductase; cMT, C-methyltransferase; ACP, acyl carrier protein; TR, thioester reductase; and TEII, thioesterase II. ₀; subscript implies inactive domain.

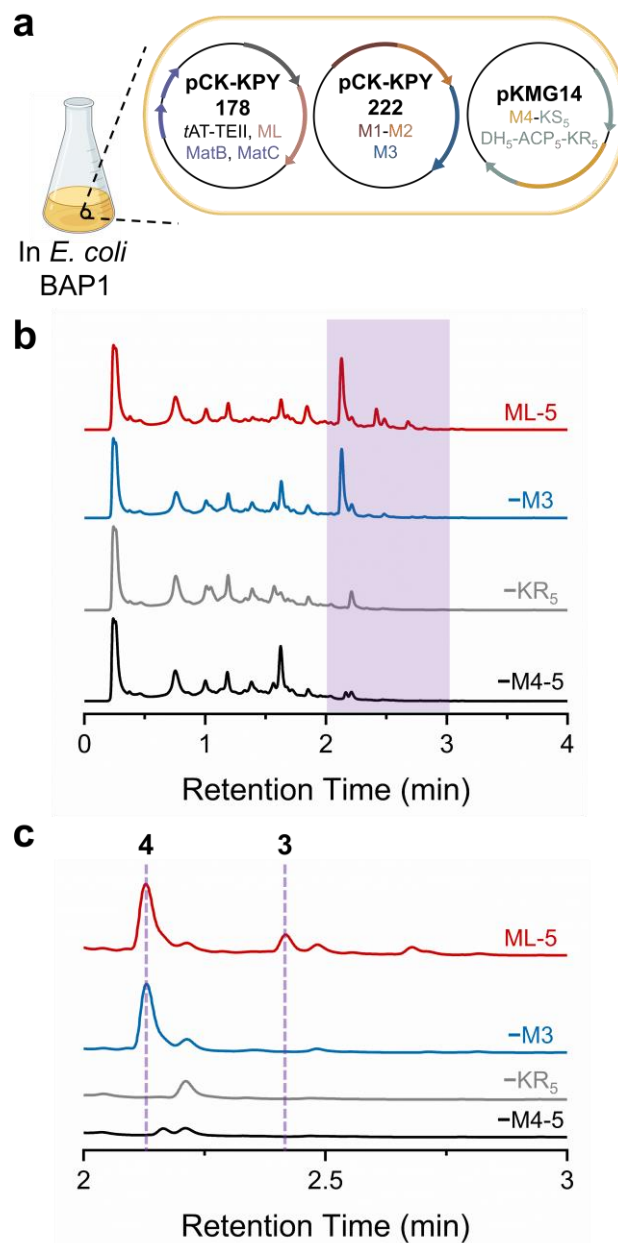


Figure 2. Compounds **3** and **4** identified in (a) an *E. coli* BAP1 strain housing NOCAP synthase Modules L-5 [pCK-KPY178/pCK-KPY222/pKMG14]. (b) HPLC UV chromatogram (at 280 ± 10 nm) comparing crude extracts from *E. coli* BAP1 supernatants. *E. coli* BAP1 strains: [pCK-KPY178/pCK-KPY222] (black, omit Module 4-KS₅ and DH₅-ACP₅-KR₅); [pCK-KPY178/pCK-KPY222/pKMG15] (grey, omit KR₅); [pCK-KPY178/pCK-KPY102/pKMG14] (blue, omit Module 3); and [pCK-KPY178/pCK-KPY222/pKMG14] (red, Modules L through 5). (c) Enlarged region of HPLC UV trace for clarity. Data acquired on an Agilent 6545 Q-TOF LC-MS system.

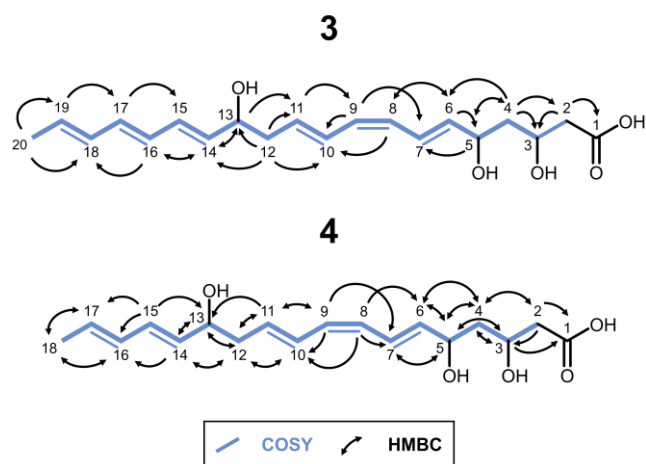


Figure 3. Structures of **3** and **4** assembled from 2D NMR data.

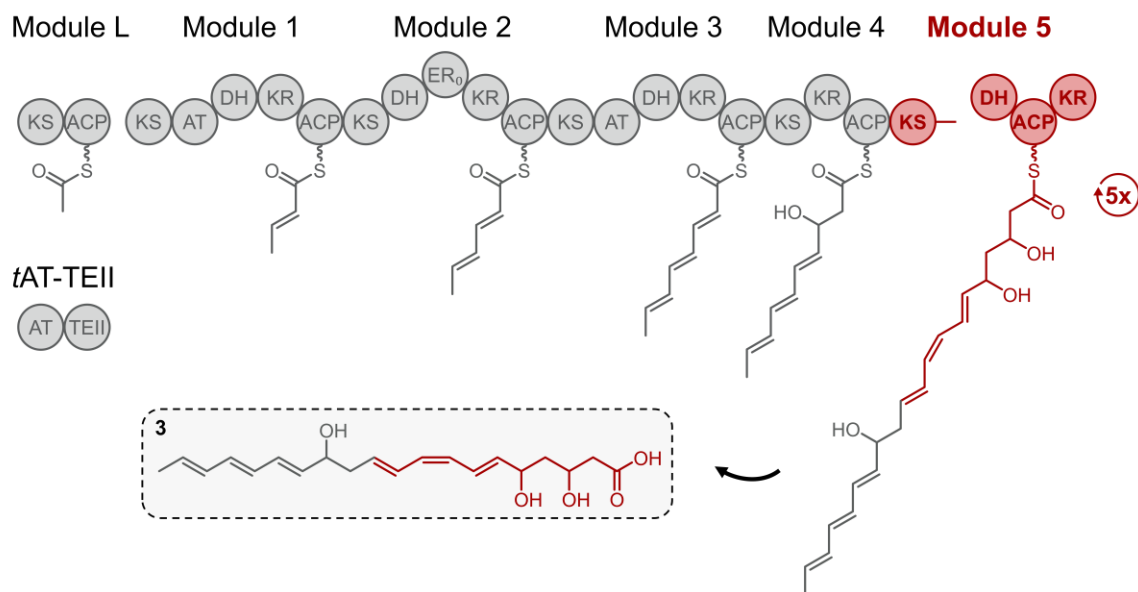


Figure 4. Proposed biosynthesis of **3** from NOCAP synthase Modules L-5 + *tAT-TEII*.

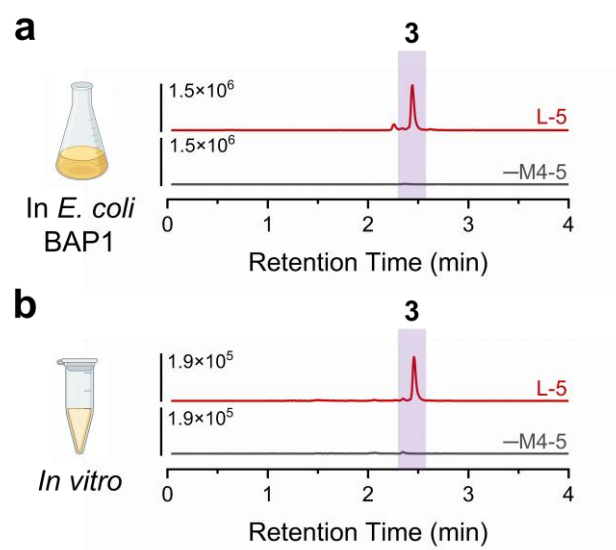


Figure 5. (a) EICs of **3** synthesized in *E. coli* BAP1 and (b) *in vitro* (upper red traces). Negative controls (lower grey traces) omitted Module 4-KS₅ and DH₅-ACP₅-KR₅.