Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An in vitro platform for the rapid and scalable biocatalysis of complex polyketides

Adrian Keatinge-Clay

adriankc@utexas.edu

UT Austin https://orcid.org/0000-0002-4358-7628

Jie Zhang

The University of Texas at Austin

Article

Keywords:

Posted Date: October 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10935154/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

An In Vitro Platform for the Rapid and Scalable Biocatalysis of Complex Polyketides

Jie Zhang[1] & Adrian T. Keatinge-Clay[1]*

[1]Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX

*Corresponding author. E-mail: adriankc@utexas.edu

ABSTRACT

Modular polyketide synthases (PKSs) are enzymatic assembly lines that naturally produce complex, bioactive compounds challenging to obtain through traditional synthesis or fermentation. Early attempts to harness these molecular factories outside the cell indicated they make poor biocatalysts. However, we observe that a refactored pikromycin synthase optimized for in vitro performance operates as rapidly as iterative fatty acid synthases, two orders of magnitude faster than measured for the equivalently-sized erythromycin synthase. The activities of four engineered variants of the pikromycin synthase were also enhanced through relocating histidine tags, swapping enzymatic/docking domains, and tuning polypeptide stoichiometry. To enable the economic conversion of feedstock chemicals malonate and methylmalonate by these and other synthases into scalable quantities of complex polyketides, we developed the Extender unit to PolyKetide (EPK) platform, which incorporates extender unit and NADPH regeneration systems. In a 10-mL, 2-hour, room-temperature reaction, 1 μM of a refactored pikromycin synthase converts malonate and methylmalonte into 2.8 mM (1 mg/mL) narbonolide with a total turnover number (TTON) of 2800. All five stereocomplex products were readily obtained in pure form at the milligram scale and structurally characterized, demonstrating the advantages of accessing complex polyketide through in vitro biocatalysis as well as the potential to access combinatorial polyketide libraries for medicinal chemistry.

INTRODUCTION

Modular polyketide synthases (PKSs) play the major role in the biosynthesis of many structurally complex, bioactive polyketides, such as the antibacterial erythromycin, the insecticide spinosyn, and the immunosuppressant rapamycin[1, 2]. Because the order of the catalytic domains in these enzymatic assembly lines largely determines the structures of their products, they can seemingly be re-engineered to access designer polyketides and combinatorial libraries for drug development. While synthetic chemists yearn to collaborate with these molecular factories to rapidly forge carbon–carbon bonds, install diverse functional groups, and set stereocenters under mild conditions, harnessing them remains a major challenge.

Most efforts to employ modular PKSs in the biocatalytic production of desired polyketides have been in vivo within *Streptomyces* or *E. coli* hosts (Fig. 1a)[3-6]. In vivo biocatalysis has been historically preferred to in vitro biocatalysis, as cells generate otherwise costly building blocks and coenzymes. However, the activities of assembly lines within cells are affected by many difficult-to-control factors, such as polypeptide stoichiometry, protein stability, building block supply, coenzyme supply, and competing metabolic pathways[7]. Thus, when low titers of desired polyketides are obtained, it can be difficult to determine whether it is the fault of the assembly line, the system housing it, or a combination thereof. While complex polyketides can be accessed through chemical synthesis, routes are usually compound-specific, resource-intensive, and low-yielding[8].

In vitro biocatalysis provides a powerful means of studying, optimizing, and harnessing modular PKSs, as parameters such as polypeptide stoichiometries, substrate concentrations, and reaction conditions can be investigated and optimized for polyketide production. The first measurement of an in vitro reconstituted assembly line was of an engineered version of the 7-module erythromycin PKS (updated modules used throughout)[9, 10] and indicated that PKSs are quite slow, with a k_{app} (system turnover rate) of 1.1 min^{-1} (Fig. 1b)[11]. Our lab subsequently measured the 3-module venemycin PKS to operate with a k_{app} of 37 min^{-1}[12]. Whether this difference is from the smaller size of the venemycin PKS, how the erythromycin PKS was refactored, or reaction conditions was unclear. Encouragingly, milligram quantities of venemycin and similar polyketides were obtained

through in vitro biocatalysis with the help of an extender unit regeneration system[12, 13]. As our lab recently accessed the heptaketide narbonolide through in vivo biocatalysis using a refactored version of the pikromycin PKS in *E. coli*, we sought to perform in vitro biocatalysis with this 7-module synthase and engineered variants to obtain semipreparative quantities of narbonolide and related complex polyketides[4, 14].

Here, we demonstrate the rapid, scalable, and economic in vitro biocatalysis of 5 complex polyketides (**1**–**5**). The activities of their assembly lines were first optimized through tuning reaction conditions, swapping docking/enzymatic domains, repositioning histidine tags, supplying the editing thioesterase PikAV, and adjusting polypeptide stoichiometries. This resulted in a k_{app} of 112 ± 11 min^{-1} (4 μM synthase) for the pikromycin PKS – 2 orders of magnitude faster than observed for the equivalent erythromycin PKS. We then developed the Extender unit to PolyKetide (EPK) platform, which includes extender unit and NADPH regeneration systems, to economically and rapidly obtain milligram quantities of each product and demonstrate the accessibility of diverse complex polyketides.

RESULTS

In vitro reconstitution of a refactored pikromycin synthase and engineered variants

We first aimed to in vitro reconstitute a refactored version of the pikromycin PKS that had been engineered for the in vivo biocatalytic production of narbonolide (**1**) (Figs. 1b–c and Extended Data Figs. 1 and 2a)[4]. While the natural synthase is composed of 4 polypeptides (PikAI–PikAIV), this refactored synthase is split by docking domains from the erythromycin and spinosyn PKSs into 6 polypeptides short enough to be expressed well in *E. coli* K207-3 (**P1P2**^{N}, ^{C}**P2P3**^{N}, ^{C}**P3P4**^{N}, ^{C}**P4P5**^{N}, ^{C}**P5P6**^{N}, and ^{C}**P6P7**; **P1**–**P7** are the pikromycin PKS modules, while ^{N} and ^{C} respectively indicate their AT+(KR/DH/ER)+ACP and KS regions; domains defined in Fig. 1)[15]. Thus, C-terminally His_6-tagged versions these polypeptides were purified from this metabolically-engineered strain of *E. coli* that expresses a phosphopantetheinyl transferase to activate acyl carrier protein (ACP) domains (Fig. S1 and Table S1)[16, 17].

Initial reactions were performed at 2 μM synthase with 1.2 mM (2*RS*)-methylmalonyl-CoA, 0.2 mM malonyl-CoA, and 1.2 mM NADPH (the theoretical ratio for narbonolide synthesis is 6:1:5) in 400 mM potassium phosphate (pH 7.5), using 5 mM tris(2-carboxyethyl)phosphine (TCEP) as a reductant (Figs. 1c, S2, and Extended Data Fig. 2b). An excess of NADPH was provided to both drive polyketide synthesis and compensate for its background oxidation. Time-course analysis showed that **1** increased during the first 8 min and decreased thereafter, suggesting its degradation. When **1** was incubated with individual reaction components, LC/MS analysis revealed its disappearance in the presence of TCEP as well as the concomitant appearance of peaks consistent with narbonolide–TCEP adducts (Extended Data Figs. 2c–f and Fig. S3). The related enone-containing macrolactone 10-deoxymethynolide (10-dml, **2**) is also degraded by TCEP, whereas hexaketide **3**, which does not contain an enone, is not sensitive to TCEP. Narbonolide (**1**) is also slowly degraded in 400 mM potassium phosphate, pH 7.5.

The k_{app} of the refactored pikromycin PKS (**NAR1**) is 21 ± 3 min^{-1} (2 μM synthase) in the absence of TCEP (Extended Data Fig. 2b). After 4 min, nearly all supplied substrates are consumed, and 178 ± 8 μM **1** is generated (89% conversion). Complete conversion is precluded by the hydrolysis of malonyl-CoA by methylmalonyl-specific ATs, the competing synthesis of byproducts such as pentaketide pyrone **6** and triketide lactone **7**, and by the degradation of **1** in phosphate (Fig. S4, *vide infra*). Lower levels of byproducts are generated compared to in vivo biocatalysis[4, 7]. **NAR1** activity was examined across a range of conditions, including different buffer compositions, pH values, glycerol concentrations, and temperatures (Extended Data Fig. 3). The synthase is most active at pH 7.5 and essentially inactive at pH 6.0. The optimal temperature of 31 °C corresponds to the preferred growth temperature of its native host, *Streptomyces venezuelae* ATCC 15439. Since the apparent K_d for the 2 polypeptides of the venemycin PKS had been measured to be 5 μM, how synthase concentration affects k_{app} of the 6-polypeptide **NAR1** was examined[12]. It was observed to increase with synthase concentration (6 ± 1, 11 ± 1, 21 ± 2, and 35 ± 4 min^{-1} at 0.5, 1, 2, and 4 μM synthase, respectively), consistent with the formation of higher-order architecture (Fig. 1c)[18, 19].

Since narbonolide (**1**) could be obtained from **NAR1**, we suspected other complex polyketides may be accessible from related, engineered synthases (Fig. 2). Indeed, anticipated polyketides are

produced by each of the 4 synthases examined: the hexaketide macrolactone 10-dml (**2**) is produced by **DML$_1$**, engineered by inactivating the **P6**KS active site cysteine of **NAR$_1$** [k_{app}: 1.3 ± 0.2 min^{-1} (2 μM synthase)]; the hexaketide δ-lactone **3** is produced by **ΔP3$_1$**, engineered by deleting **P3** of **NAR$_1$** [k_{app} of 0.40 ± 0.04 min^{-1} (2 μM synthase)]; the hexaketide macrolactone **4** is produced by **ΔP6$_1$**, engineered by deleting **P6** of **NAR$_1$** [k_{app} of 1.3 ± 0.3 min^{-1} (2 μM synthase)]; and the heptaketide macrolactone anhydronarbonolide (**5**)[20] is produced by **ANH$_1$**, engineered by swapping **P6** of **NAR$_1$** for **R5** of the rapamycin PKS [k_{app} of 0.40 ± 0.04 min^{-1} (2 μM synthase)]. As these synthases operate at only 2–6% the rate of **NAR$_1$**, competing reactions are more apparent – significant quantities of byproducts are generated by ACP-independent condensations between methylmalonyl-CoA and KS-bound intermediates (Extended Data Fig. 4 and Figs. S4–5), and a large fraction of malonyl-CoA is nonproductively hydrolyzed by methylmalonyl-specific ATs (Extended Data Figs. 4–5). The omission of polypeptides from **NAR$_1$** revealed byproducts **6** and **7** are respectively generated from the offloading of the **P2**KS-bound diketide and **P4**KS-bound tetraketide by methylmalonyl-CoA, similar to in vitro studies of polypeptides from the erythromycin synthase in which KS-bound intermediates were offloaded by fluoromalonyl-CoA[21]. While hydrolysis of malonyl-CoA by methylmalonyl-specific ATs had been previously observed, the extent of this side-reaction was surprising[22]. Malonyl-CoA added to **ΔP3$_1$**, a synthase comprised of only methylmalonyl-specifice ATs, is consumed 5 times more rapidly than methylmalonyl-CoA.

Modular PKSs can operate as fast as FASs

We sought to engineer versions of each assembly line that outpace competing reactions, beginning with the refactored pikromycin synthase (Extended Data Fig. 6). Since in vivo biocatalysis of narbonolide is improved by employing the native pikromycin docking domain motifs[4], we reintroduced **P3**DD, **P5**DD, and **P6**DD. The k_{app} of the resulting assembly line, **NAR$_2$**, is 29 ± 4 min^{-1} (1 μM synthase), 2.4-fold greater than **NAR$_1$**. Removing the docking domain motifs that had been introduced into **P4**, which does not natively contain them, had no effect on assembly line activity [k_{app} (**NAR$_3$**): 29 ± 2 min^{-1} (1 μM synthase)]. Since both termini participate in docking interactions, the locations of histidine tags were investigated with **NAR$_4$**–**NAR$_7$**. These designs revealed that tag relocation from the C-terminus to the loop between the KS and AT domains

slightly improves activity [k_{app} (**NAR$_7$**): 37 ± 3 min^{-1} (1 μM synthase)]. Splitting the first polypeptide to facilitate protein purification did not significantly affect synthase activity [k_{app} (**NAR$_8$**): 36 ± 4 min^{-1} (1 μM synthase)] (Figs. 3a–b). As with **NAR$_1$**, the rate of narbonolide production by **NAR$_8$** increases with synthase concentration, plateauing at k_{app} of 85 ± 11 min^{-1} with 4 μM synthase at 31 °C (Fig. 3c). The natural, 3-module venemycin PKS was re-examined and observed to plateau at a k_{app} of 43 ± 2 min^{-1} with 2 μM synthase (Extended Data Fig. 7).

Premature decarboxylation of ACP-bound extender units results in acetyl and propionyl intermediates that stall assembly lines. As the addition of the editing thioesterase PikAV that removes these groups improved in vivo biocatalysis by the refactored pikromycin PKS, it was investigated[4, 23]. The k_{app} of **NAR$_1$** (2 μM synthase) increases 1.5-fold in the presence of 1.75–7 μM PikAV (Figs. 3d and S6). Supplying 3.5 μM PikAV to **NAR$_8$** increases its activity 1.3-fold, enabling the highest k_{app} observed in this study: 112 ± 11 min^{-1} (4 μM synthase). This is equivalent to several well-studied, iterative type I fatty acid synthases (FASs; *vide infra*)[24-27]. Under these conditions, 0.9 mM narbonolide is formed from 1 mM malonyl-CoA in 3 min (90% conversion).

Although relocating histidine tags to the loop between KS and AT yielded the most active synthases, C-terminally-tagged polypeptides are more readily purifiable. Thus, the purification-friendly, 7-polypeptide **NAR$_9$** was engineered [k_{app} (**NAR$_9$**): 27 ± 5 min^{-1} (1 μM synthase)] (Fig. 3a). This design was selected for the scaled-up in vitro biocatalysis of narbonolide and as the basis for redesigning the **DML**, **ΔP3**, **ΔP6**, and **ANH** synthases (Extended Data Fig. 8).

Optimizing domain compositions

DML$_1$, **ΔP3$_1$**, **ΔP6$_1$**, and **ANH$_1$** are 94–98% less active than **NAR$_1$**, the synthase from which they are derived. This is likely a result of suboptimal domain-domain and polyketide-domain interactions. We sought to learn how to optimize domain-domain interactions using the in vitro platform by studying model synthases in which polyketide-domain interactions are removed from the equation. The 6th modules of the pikromycin and erythromycin PKSs, **P6** and **E6**, perform identical chemistry on hexaketide intermediates possessing identical α–ζ substituents (Fig. 4a). Thus, assembly lines in which domains in **P6** and the surrounding region were swapped with equivalent domains from the erythromycin PKS were combinatorially constructed. KSs and the

processing enzymes upstream of them are now known to closely evolutionarily co-migrate and are grouped together within the updated module[9, 10]; KSs physically associate with the acyl-ACPs upstream of them to enable the transacylation of correctly processed acyl chains[28, 29]. The synthase in which **P5**KS, **P6**AT, **P6**KR, and **P6**ACP (the traditional module, $\mathbf{x_1y_1}$) are swapped contains 2 disrupted ACP/KS interfaces and is one of the least active. Although the synthase in which **P6**AT, **P6**KR, **P6**ACP, and **P6**KS (the updated module, $\mathbf{x_2y_2}$) are swapped was quite active, the most active synthase was constructed by swapping only **P6**KR, **P6**ACP, and **P6**KS ($\mathbf{x_3y_2}$). The synthase in which **P6**KR, **P6**ACP, **P6**KS, and **P7**AT (equivalent to the "XUA" set of domains that yielded the best results in recent engineering of the mediomycin PKS, $\mathbf{x_3y_3}$) are swapped is much less active[5]. The domain-swapping strategy was also applied to optimize $\mathbf{ANH_1}$ and yielded equivalent results – swapping **P6(**KR+ACP+KS) for **R5(**DH+KR+ACP+KS) ($\mathbf{x_2y_2}$) boosts productivity 2.5-fold (Fig. 4b).

The Sherman lab observed 10-dml (**2**) production from incubating a pentaketide substrate with PikAIII and a version of PikAIV in which the **P6**KS reactive cysteine is mutated to alanine. The production rate increased when the **P7**ACP serine that is phosphopantetheinylated is also mutated to alanine. We observe this mutation to enhance activity 9.3-fold [k_{app} ($\mathbf{DML_3}$): 6.5 ± 0.9 min^{-1}] (Extended Data Figure 8). However, the deletion of **P7**ACP enhances activity 11-fold [k_{app} ($\mathbf{DML_4}$): (7.6 ± 0.9 min^{-1})]. With $\mathbf{DML_4}$ at 2 μM and PikAV at 3.5 μM, k_{app} is 13.6 ± 1.4 min^{-1}.

Increased flux through stoichiometric tuning

A general practice in metabolic engineering is to increase the concentration of enzymes suspected to be at the bottleneck of a biochemical pathway. Thus, we examined the effect of increasing the concentrations of the assembly line polypeptides most affected through the engineering performed to generate the **DML**, **ΔP3**, **ΔP6**, and **ANH** synthases. The concentrations of the 3rd and 4th polypeptides of the module-deleted synthase $\mathbf{\Delta P3_3}$, $\mathbf{^{C}P2P4^{N}}$ and $\mathbf{^{C}P4P5^{N}}$, were varied while maintaining the others at 1 μM (Fig. 5a and Extended Data Fig. 8). Activity is not enhanced by independently increasing the concentration of the upstream polypeptide, $\mathbf{^{C}P2P4^{N}}$, but is enhanced by independently increasing the concentration of the downstream polypeptide, $\mathbf{^{C}P4P5^{N}}$. Interestingly, when the concentration of the downstream polypeptide, $\mathbf{^{C}P4P5^{N}}$, is increased, activity can be further enhanced by increasing the concentration of the upstream polypeptide,

^{C}**P2P4**^{N}, but only until it equals the concentration of the downstream polypeptide, ^{C}**P4P5**^{N}. With 4 μM ^{C}**P2P4**^{N} and 8 μM ^{C}**P4P5**^{N}, the k_{app} for **ΔP3$_3$** is enhanced 4-fold to 12.5 ± 1.2 min^{-1}. The same trend was observed when titrating the 6th and 7th polypeptides of **ANH$_3$**: its k_{app} is enhanced 12-fold with 8 μM ^{C}**P5R5**^{N} and 16 μM ^{C}**R5P7** to 16.9 ± 1.1 min^{-1} (Fig. 5b). The activities of **DML$_4$** and **ΔP6$_3$** were not enhanced by increasing the concentrations of their 2 downstream polypeptides.

In vitro biocatalysis yields semi-preparative quantities of diverse complex polyketides

To enable scaled-up biocatalytic reactions with **NAR$_9$**, **DML$_4$**, **ΔP3$_3$**, **ΔP6$_3$**, and **ANH$_3$**, the Extender unit to PolyKetide (EPK) platform was developed (Fig. 6). Its reaction mixture includes regeneration systems for complex polyketide synthesis – adenylate kinase, pyruvate kinase, MatB, and methylmalonyl-CoA epimerase for malonyl- and (2*S*)-methylmalonyl-CoA regeneration as well as glucose dehydrogenase for NADPH regeneration. Under platform conditions, malonyl-CoA can be produced at 31 μM min^{-1}, and a 1:1 mixture of (2*R*)- and (2*S*)-methylmalonyl-CoA can be produced at 88 μM min^{-1} (ATP is supplied 4 min before the PKS polypeptides to generate the starting pool of extender units). These rates support narbonolide formation in a 100-μL reaction containing 1 μM **NAR$_9$** and 3.5 μM PikAV at room temperature (23 ºC).

Next, a 10-mL EPK reaction was performed to test the semi-preparative production of narbonolide (Fig. 6a). Its initial rate was equivalent to that of a reaction containing malonyl-CoA, (2*RS*)-methylmalonyl-CoA and NADPH, but no regeneration enzymes (43 versus 47 μM min^{-1}), indicating that extender unit and coenzyme regeneration keep pace with the assembly line. The narbonolide (**1**) concentration reached 1 mg mL^{-1} (2.8 mM) after 2 h. Following ethyl acetate extraction and HPLC purification, 7.2 mg of pure narbonolide was obtained, enabling its characterization by NMR and x-ray diffraction (Fig. S10 and Extended Data Table 1). This is the first time the crystal structure of this iconic polyketide has been reported.

During 8-h reactions with slower synthases, GDH was observed to precipitate. However, after increasing its concentration to 10 μM and supplying 5% (v/v) glycerol, each of **DML$_4$**, **ΔP3$_3$**, **ΔP6$_3$** and **ANH$_3$** yielded their anticipated polyketides (**2**–**5**) in quantities sufficient for characterization by NMR (and x-ray crystallography for **4**; Figs. 6b and S11–21, Extended Data Table 1). The clean

background of the in vitro reactions enabled the detection of unexpected products that could easily have been missed working up in vivo biocatalytic reactions. A compound observed while purifying **5** from the **ANH$_3$** reaction was analyzed by LC/MS and NMR and determined to be the linear product **8** (Figures S27–31). Similarly, a compound observed while purifying macrolactone **4** from the **ΔP6$_3$** reaction was determined to be its C2-epimer, **9** (Figures S32–36).

DISCUSSION

Complex polyketides are currently accessed through chemical synthesis or fermentation rather than through in vitro biocatalysis. Early in vitro studies of modular PKSs led to the notion that harnessing them outside the cell is impractical: a reconstituted erythromycin synthase was measured to produce 6-deoxyerythronolide B at a sluggish 1.1 min^{-1}[11], and the highest observed TTON measured for PKS polypeptides was 300 for the conversion of a pentaketide substrate to narbonolide by PikAIII and PikAIV[30]. In addition these unimpressive biocatalytic parameters, potential deterrents to in vitro biocatalysis include the challenges of pathway engineering and polypeptide purification as well as the high costs of required substrates and coenzymes.

Here, we demonstrate that complex polyketides can be obtained through in vitro biocatalysis and that this route offers several advantages. Design and reaction optimization for the refactored pikromycin PKS led to **NAR$_8$**, which can synthesize narbonolide with a k_{app} of 112 ± 11 min^{-1}, two orders of magnitude faster than measured for the equivalently-sized erythromycin PKS. Thus, each module in **NAR$_8$** operates at least at 1.9 cycles per second, similar to rates measured for iterative, type I fatty acid synthases (FASs) from *Saccharomyces cerevisiae*, *Corynebacterium ammoniagenes*, and *Gallus gallus* at 18, 2.0, and 3.4 cycles per second, respectively[24-27]. While the reaction chambers of the structurally-characterized hexameric *S. cerevisiae* FAS and dimeric *Dictyostelium discoideum* are quite different, both contain relatively mobile ACP domains that transiently dock rigidly-oriented AT, KS, KR, DH, and ER domains[31, 32]. Thus, pikromycin synthase is likely comprised of highly organized reaction chambers, which overlap such that ACPs convey intermediates from one to the next. Here, we show PKSs are not only faster than expected but also more biocatalytically robust, possessing a half-life of ~1 day and capable of high turnover. For the 2-hour, 10-mL, room temperature EPK reaction in which 1 μM **NAR$_9$** generates 2.8 mM

narbonolide, the TTON is a solid 2800. EPK reactions are relatively inexpensive, as they only utilize catalytic quantities of ATP, CoA, and NADPH. Thus, the complex polyketide narbonolide can be accessed through in vitro biocatalysis more rapidly and economically than through chemical synthesis. Likewise, narbonolide can be accessed more rapidly through in vitro biocatalysis than in vivo biocatalysis; follow-up studies will determine whether it is also more economical[4].

Synthases are quite sensitive to changes imparted through PKS engineering. The module deletions that yielded **ΔP3$_1$** and **ΔP6$_1$**, the module-swapping that yielded **ANH$_1$**, and the point mutation that yielded **DML$_1$** each result in suboptimal domain-domain and polyketide-domain interactions, leading to 94–98% decreases in activity relative to **NAR$_1$**. Even the swapping of docking domains led to a 59% difference in activity between **NAR$_2$**, which harbors the pikromycin docking domains, and **NAR$_1$**, which harbors docking domains from the erythromycin and spinosyn PKSs. We describe several methods to boost the activities of engineered synthases, including determining the optimal set of domains for swapping and increasing the stoichiometries of the polypeptides most affected by the engineered change. While the best results from EPK reactions are obtained from synthases that operate with a k_{app} of at least 10 min^{-1}, even **ΔP6$_3$**, which operates at 3.8 ± 0.5 min^{-1}, yields 2.7 mg of pure hexaketide macrolactone **4** from a 20-mL EPK reaction.

The EPK platform reaches far beyond providing scalable access to a desired complex polyketide. As demonstrated, it can quantitatively measure assembly line activities independent of cellular metabolism to quickly provide feedback to design-build-test-learn PKS engineering cycles. The much smaller volumes and complexity enabled by the platform compared to in vivo biocatalysis greatly facilitate product purification and characterization by NMR and crystallography. It enables structural biology studies of PKS higher-order architecture through advanced techniques such as cryo-electron tomography and atomic force microscopy. The EPK platform could also aid in the synthesis of large, combinatorial libraries of complex polyketides through automated liquid handling of PKS polypeptides in multi-well plates to accelerate the medicinal chemistry needed to optimize complex polyketide drug leads. Thus, this biotechnological innovation has opened diverse collaborations with some of Nature's most powerful synthetic factories.

Reporting Summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this Article.

Data availability

All data are available through the Supplementary Information document and the Source Data file. The X-ray crystallographic coordinates for the structures reported in this study, **1** and **4**, have been deposited at the Cambridge Crystallographic Data Centre (CCDC), under deposition numbers 2513770 and 2578074, respectively. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

Biological materials. E. coli K207-3 cells are available from the authors upon request.

Acknowledgements. Serhii Vasylevskyi, Director of the X-ray Facility at The University of Texas at Austin, solved the crystal structures of narbonolide (**1**) and **4**.

Funding. A.T.K. discloses support for the research of this work from the NIH [grant number GM106112] and the Welch Foundation [grant number F-2245].

Author contributions statement. J.Z. and A.T.K. designed the project. J.Z. performed all the experiments. J.Z. and A.T.K. wrote the manuscript.

Competing interests statement. The authors declare no competing interests.

Inclusion & ethics statement. All participants in this research were properly credited.

Additional information. Supplementary methods, figures, and tables are available in the Supplementary Information document. Correspondence and requests for materials should be addressed to A.T.K.

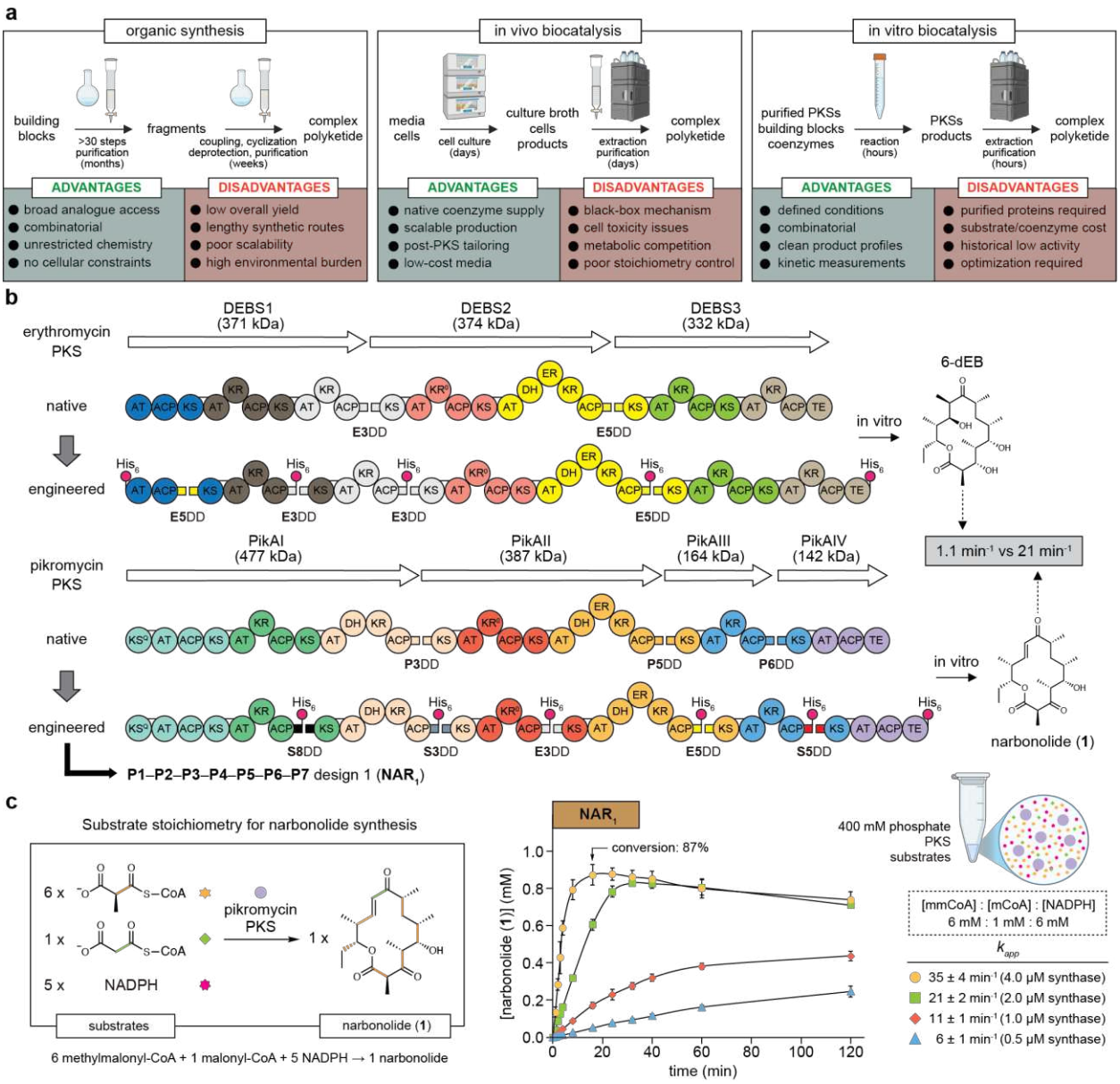


Figure 1. In vitro biocatalysis with modular PKSs. a) The advantages and disadvantages of 3 routes to accessing desired polyketides – chemical synthesis, in vivo biocatalysis, and in vitro biocatalysis – are compared. b) The first report of an in vitro reconstituted PKS was of an engineered version of the erythromycin PKS, which was measured to operate with a k_{app} of 1.1 min^{-1} (2 μM synthase). The first version of the pikromycin PKS tested here (**NAR$_1$**) operates with a k_{app} of 21 min^{-1} (2 μM synthase). His_6-tags and docking domains (**E**, erythromycin; **P**, pikromycin; **S**, spinosyn) are indicated. AT, acyltransferase; DH, dehydratase; ER, enoylreductase; ACP, acyl carrier protein; KS, ketosynthase; TE, thioesterase. c) (Methyl)malonyl-CoA substrates

and NADPH coenzyme, supplied in slight excess, are rapidly converted to narbonolide by the pikromycin PKS (87% in 16 min). The dependence of k_{app} on enzyme concentration is indicative of multimeric architecture.

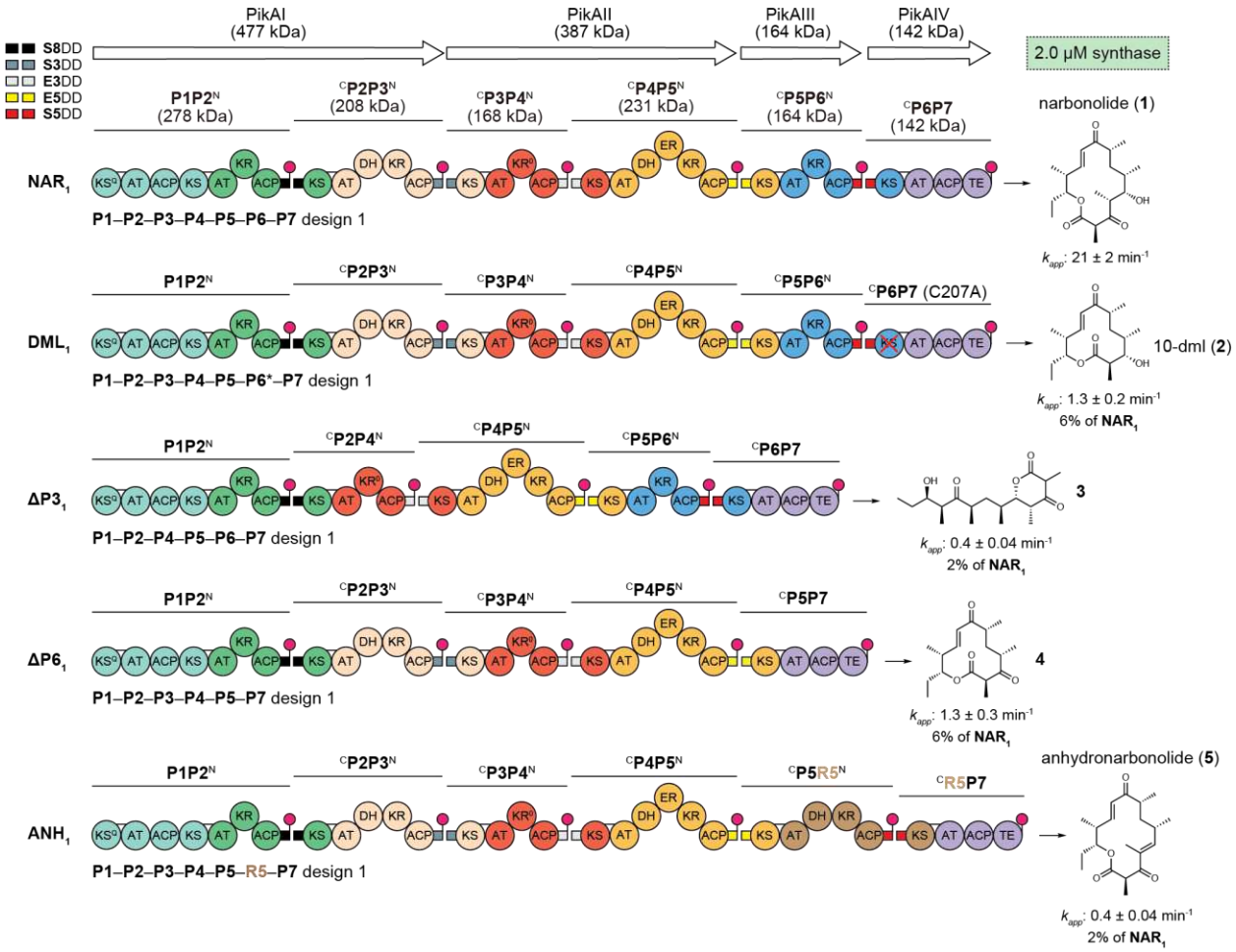


Figure 2. Initial designs for the pikromycin PKS and 4 engineered variants. These assembly lines produce heptaketides narbonolide (**1**) and anhydronarbonolide (**5**) as well as hexaketides 10-dml (**2**), **3**, and **4**. The k_{app} values, measured for each initial design at 2 μM synthase, show the engineered variants are at least an order of magnitude slower than the refactored pikromycin PKS **NAR1**.

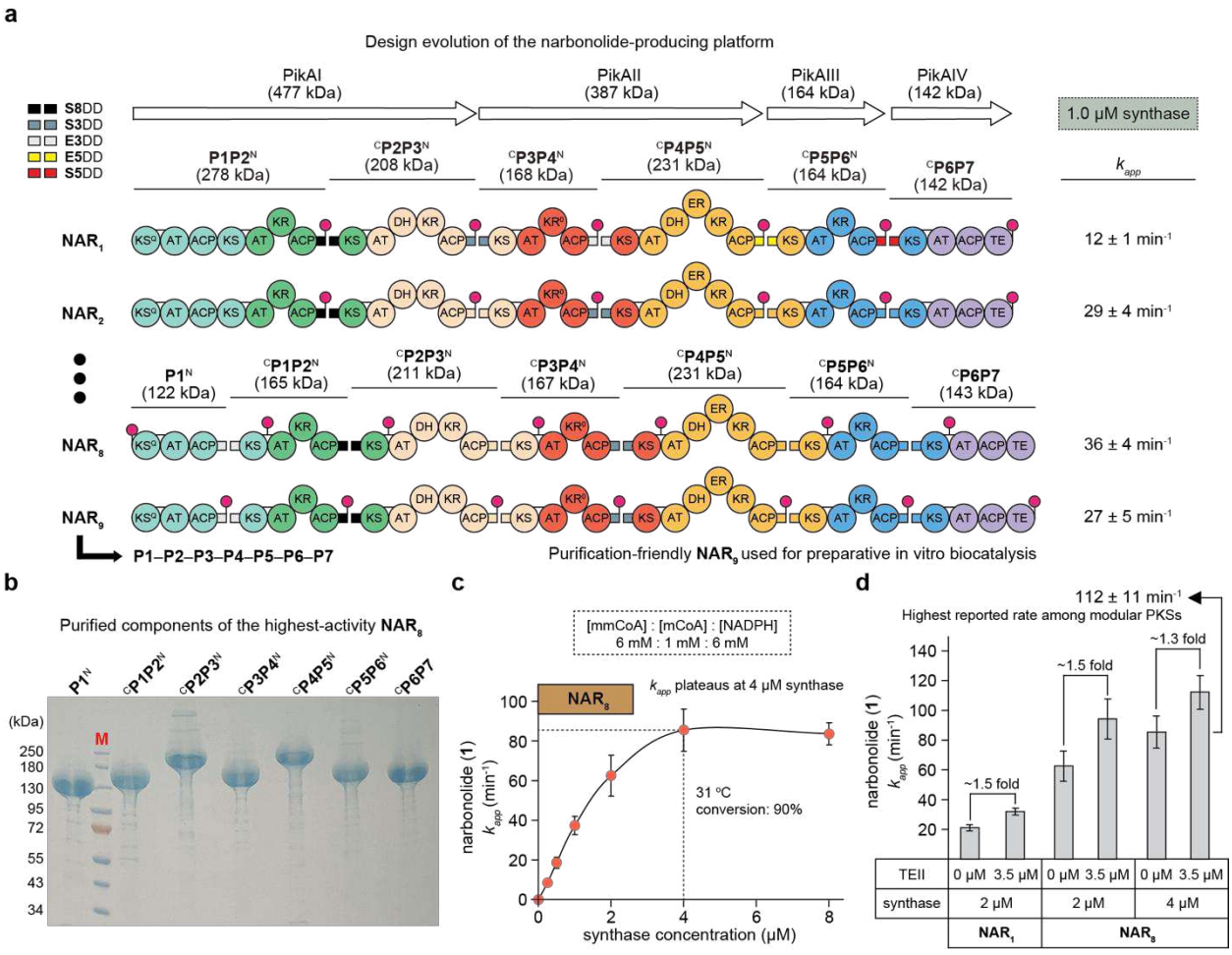


Figure 3. Optimizing pikromycin PKS for narbonolide production. a) Several versions of the pikromycin PKS, differing in polypeptide boundaries, docking domains, and histidine tag locations were engineered. Replacing heterologous docking domains with those native to the pikromycin PKS increased assembly line activity 140%. Relocating histidine tags to the loop between KS and AT increased activity 24%. While **NAR8** showed the highest turnover rate, the C-terminally tagged polypeptides of **NAR9** were more readily purified. Thus, **NAR9** was chosen for preparative polyketide production. b) SDS-PAGE analysis of the 7 purified polypeptides of **NAR8**. c) The k_{app} of **NAR8** increases with [synthase] until 4 μM. The optimal temperature is 31 °C (see Extended Data Fig. 3f). d) In the presence of 3.5 μM PikAV, **NAR8** operates with a k_{app} of 112 ± 11 min^{-1} (4 μM synthase), the highest rate measured for a modular PKS.

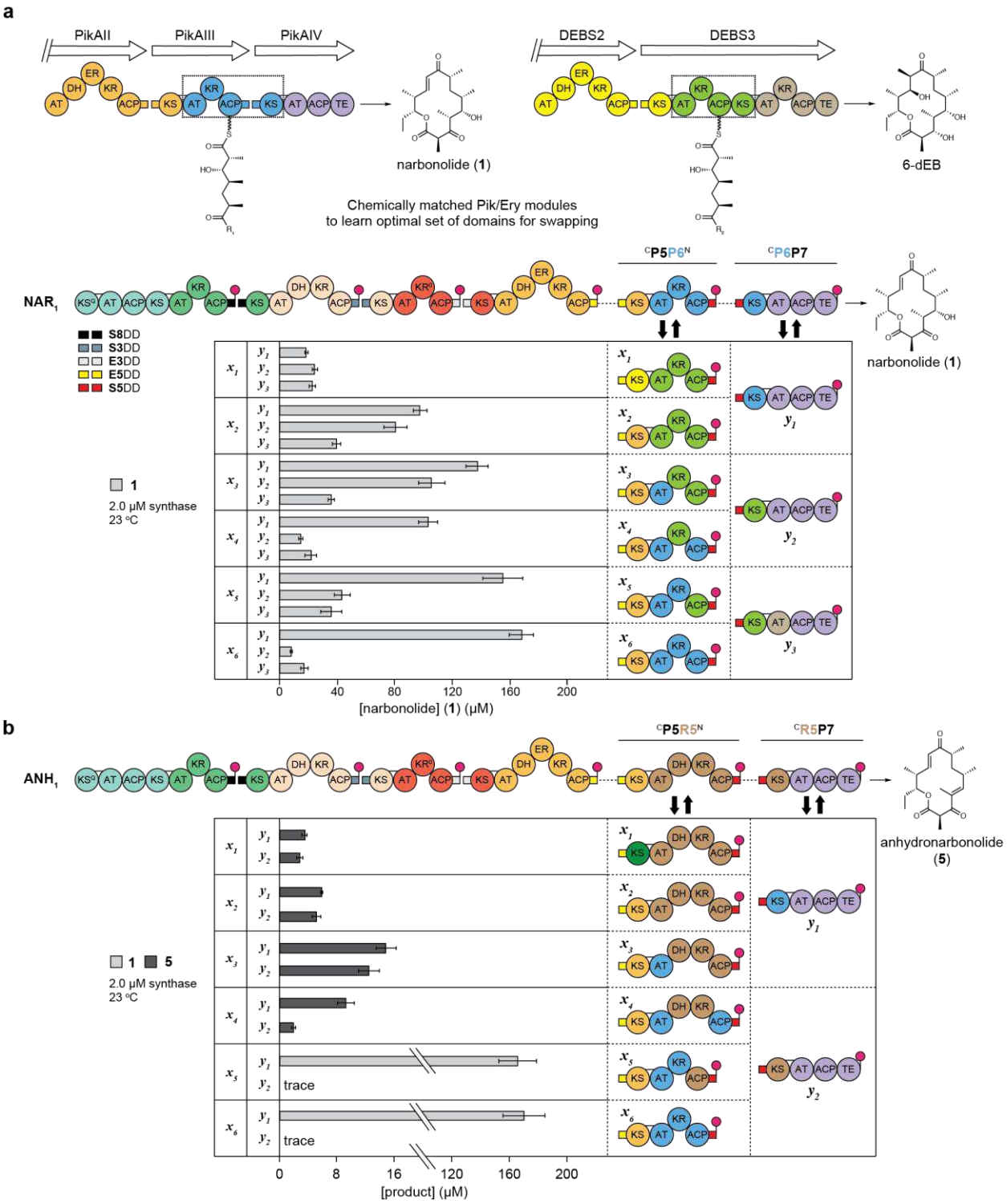


Figure 4. Optimizing domain compositions. a) Replacing **P6** of **P1-P2-P3-P4-P5-P6-P7** with **E6** entails swapping the AT, KR, ACP, and KS domains (**x2**, **y2**). The resulting synthase, **P1-P2-P3-P4-P5-E6-P7**, produces narbonolide at half the rate. However, when only KR, ACP, and KS are swapped (**x3**, **y2**), the resulting synthase produces narbonolide 25% faster than **P1-P2-P3-P4-P5-**

E6-**P7**. b) Similarly, when only the KR, ACP, and KS domains of **P6** are replaced with the equivalent domains of **R5** (**x_2**,**y_2**), the resulting synthase is 2.5-fold more active than **P1**-**P2**-**P3**-**P4**-**P5**-**R5**-**P7** (**x_3**,**y_2**).

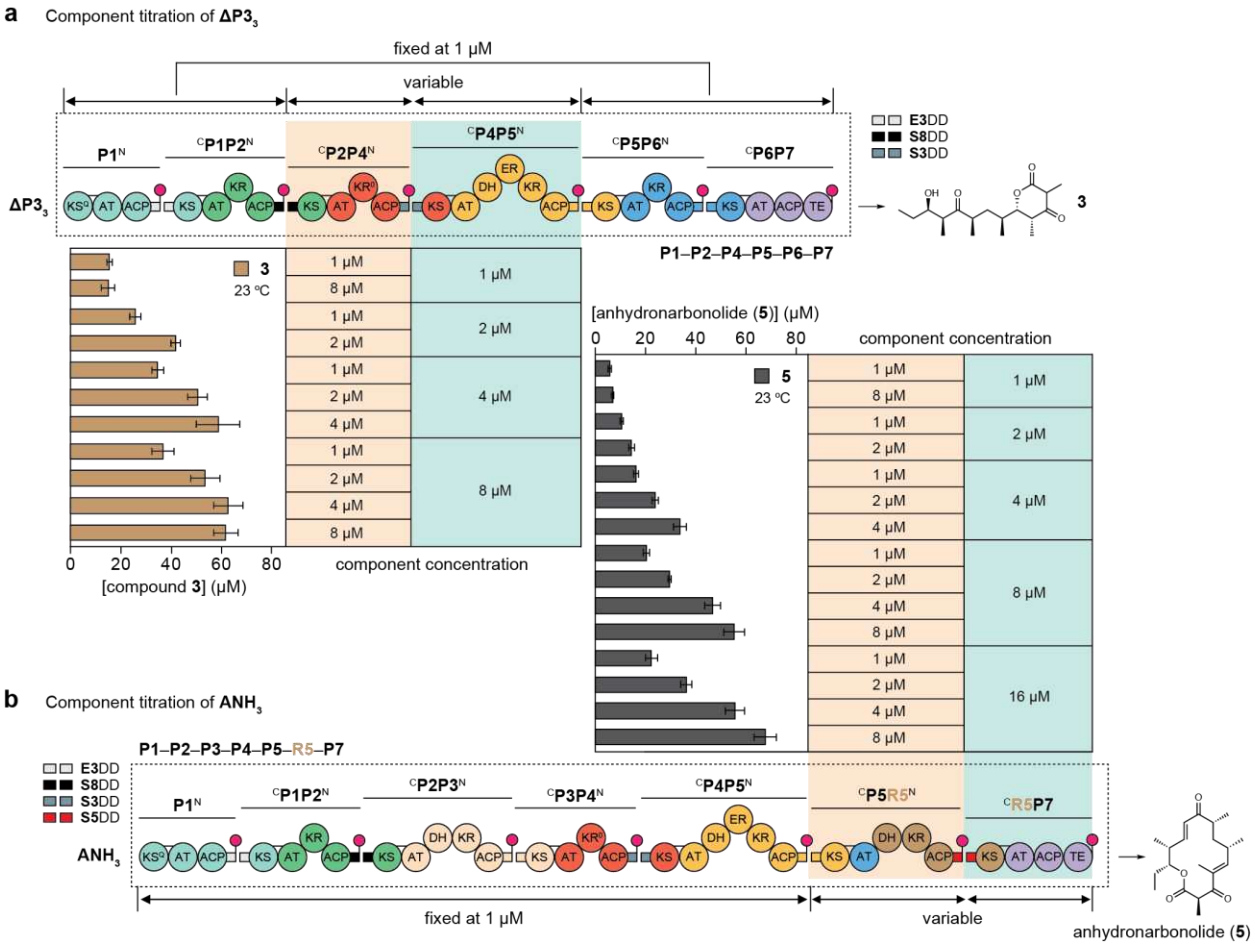


Figure 5. Optimizing polypeptide stoichiometry. a) Stoichiometry-tuning of the module-deleted synthase **ΔP3** shows that increasing the concentration of the polypeptide downstream of the deletion, **^{C}P4P5^N**, increases production of hexaketide **3** more than increasing the concentration of the polypeptide from which the module was deleted, **^{C}P2P4^N**. b) Stoichiometry-tuning of the module-swapped synthase **AHN** shows that increasing the concentration of both polypeptides containing the module swap, **^{C}P5R5^N** and **^{C}R5P7**, improves the production rate for heptaketide **5** (supplying 8 μM **^{C}P5R5^N** and 16 μM **^{C}R5P7** boosted productivity 12-fold).

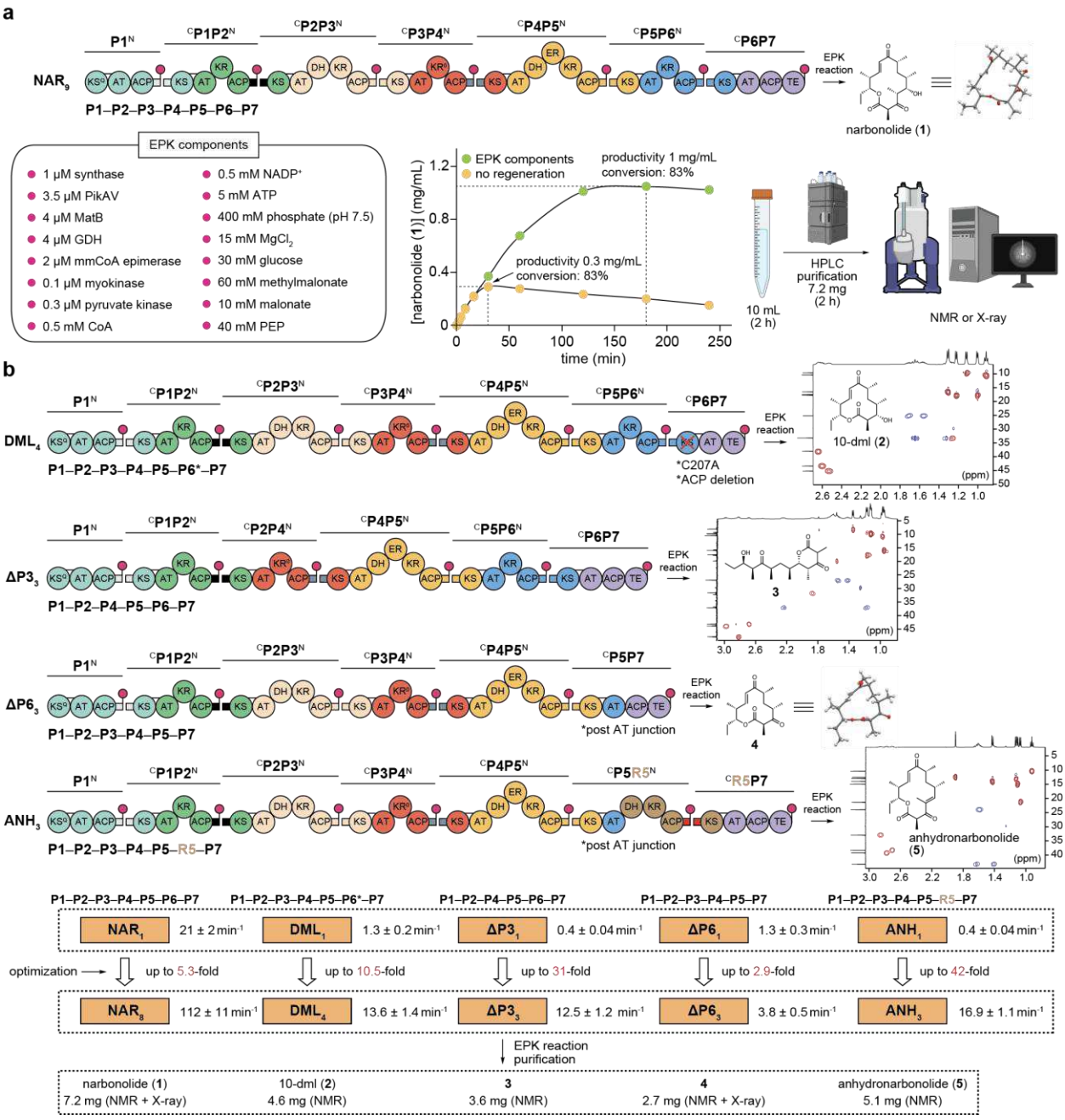
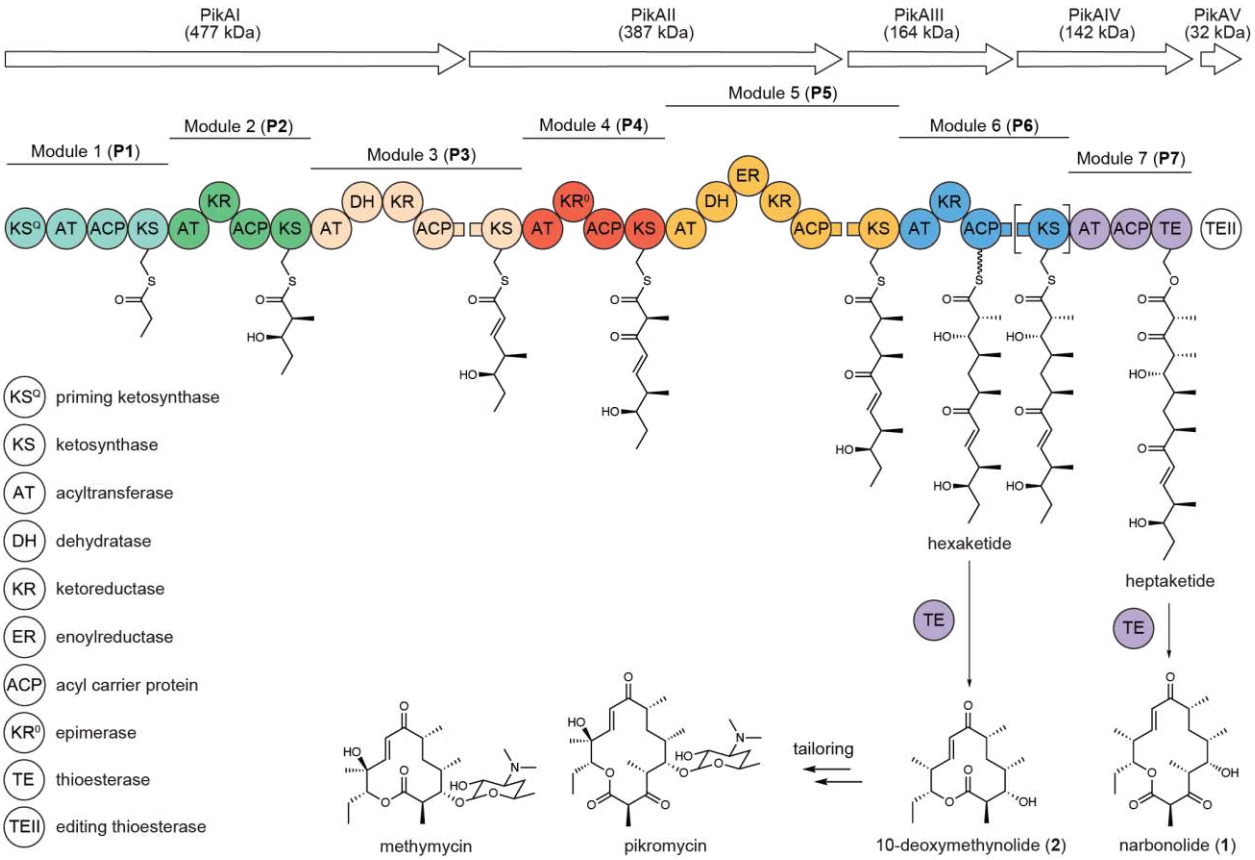
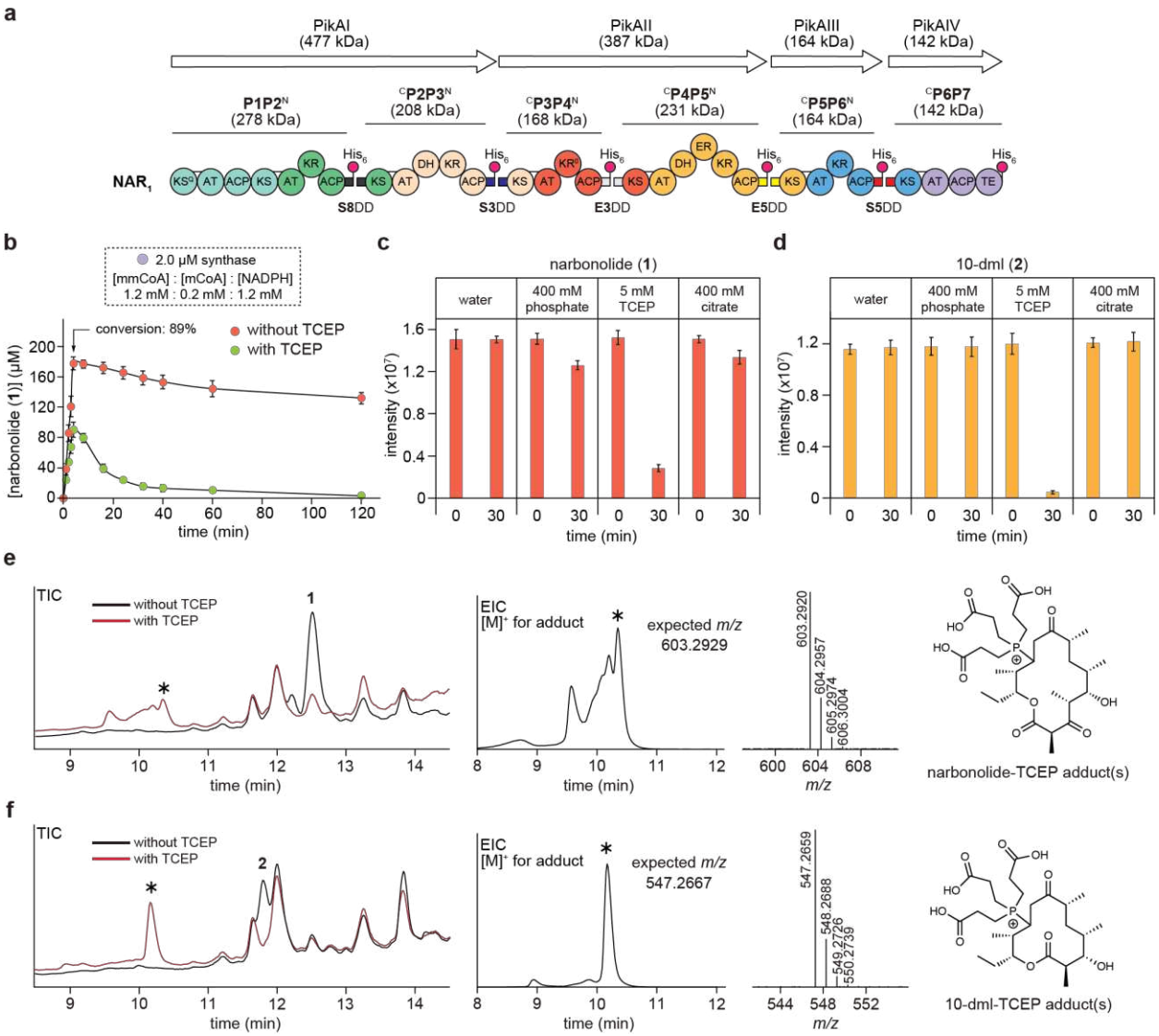


Figure 6. The Extender unit to PolyKetide (EPK) platform provides rapid, economic access to complex polyketides. a) The platform includes extender unit and NADPH regeneration systems that help convert malonate and methylmalonate to polyketide products. It was optimized by monitoring narbonolide (**1**) production from **NAR9**. The platform maintained high activity after 30 min, in contrast to reactions lacking regeneration systems. From a 2-hour, 10-mL reaction conducted at 23 ºC, 7.2 mg pure narbonolide was obtained, enabling its crystal structure to be reported for the first time. b) EPK reactions were also performed with the 4 engineered variants of

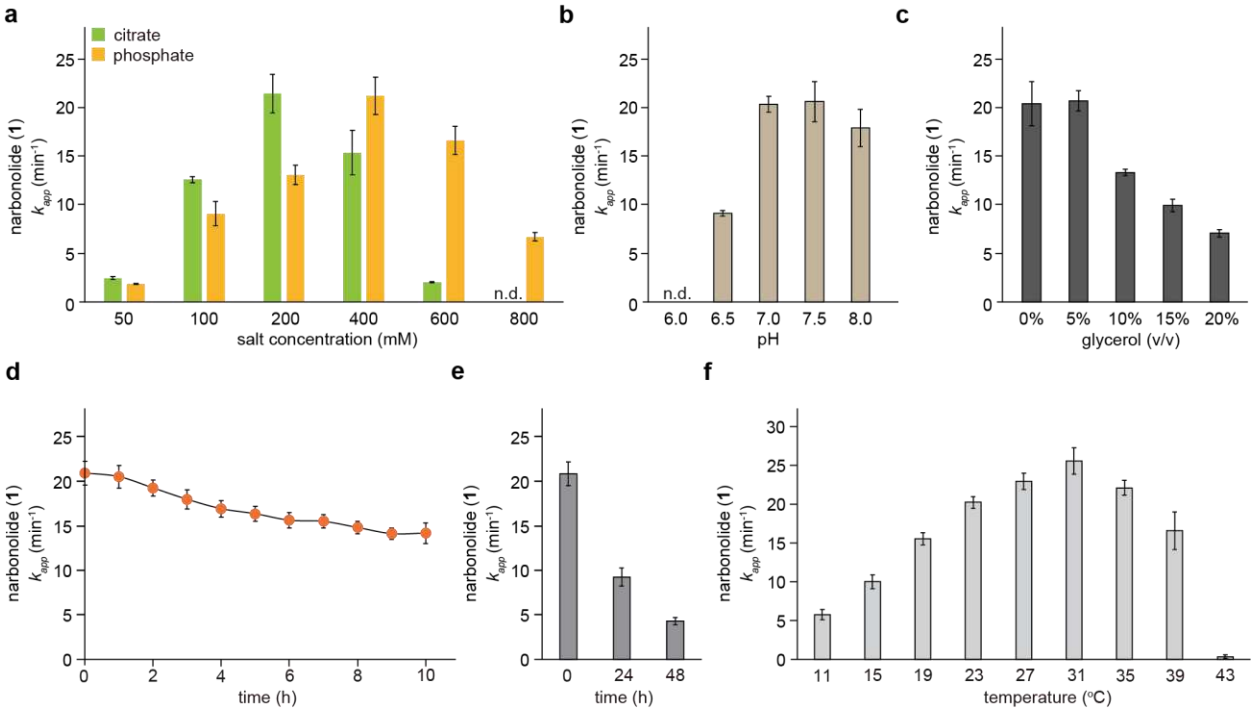
the pikromycin synthase that were optimized (up to 42-fold) for in vitro biocatalysis. They also yielded semi-preparative quantities of complex polyketides, which were structurally characterized by NMR spectroscopy and x-ray crystallography (Figs. S9–S35).



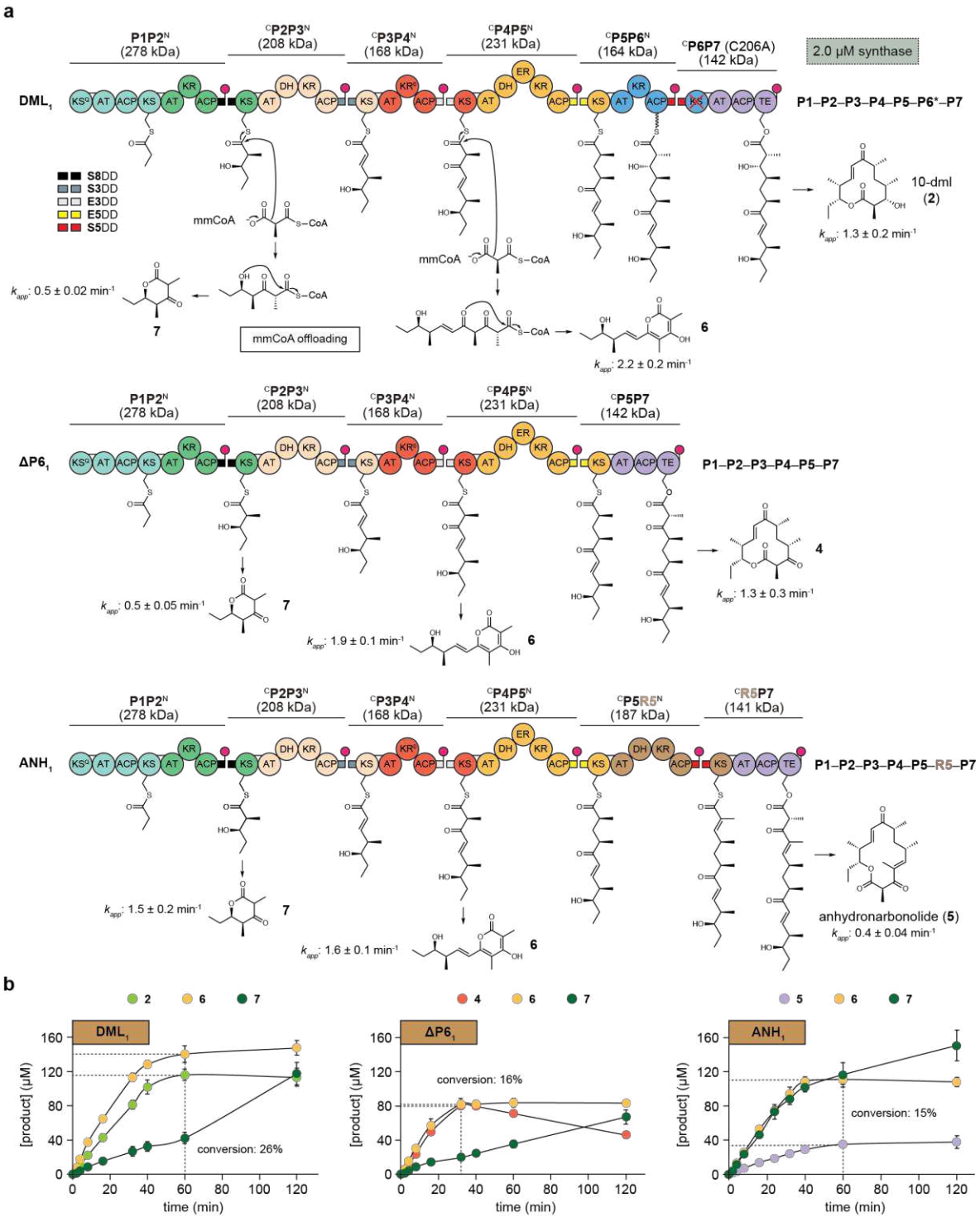
Extended Data Figure 1. The pikromycin polyketide synthase (PKS) and its products. Four polypeptides house 7 modules (updated definition), **P1**–**P7**, that synthesize the macrolactones narbonolide (**1**) and 10-deoxymethynolide (10-dml, **2**), precursors of the macrolide antibiotics pikromycin and methymycin. Inactivation of the most downstream KS results in the production of 10-dml. Rectangles represent the docking domain motifs, ^{C}DD and ^{N}DD.



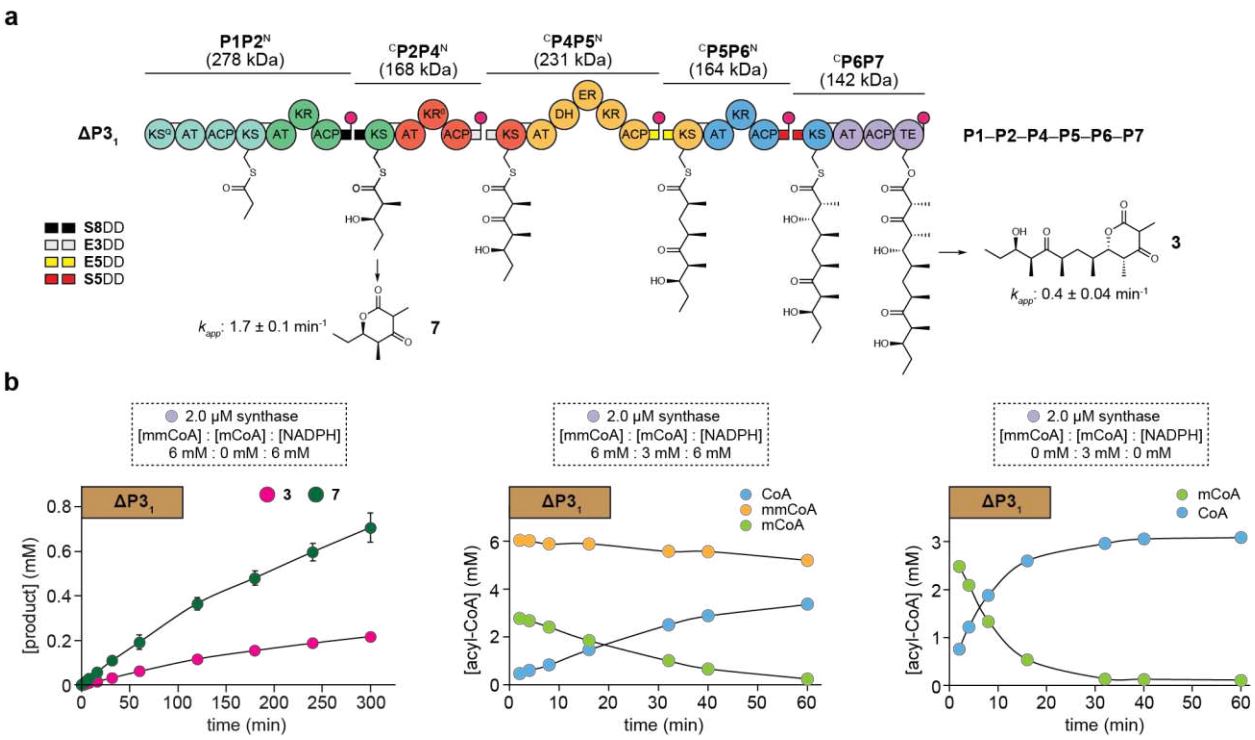
Extended Data Figure 2. In vitro biocatalysis of narbonolide (**1**). a) His_6-tags (pink circles) were engineered at the C-termini of the 6 polypeptides of a refactored pikromycin PKS that generates narbonolide in *E. coli* K207-3 to facilitate their purification. The polypeptides noncovalently assemble through docking domain (DD) motifs imported from modules of the spinosyn (**S**) and erythromycin (**E**) PKSs. b) Narbonolide produced by **NAR1** (2 μM synthase) disappears in reactions containing TCEP. c) Pure narbonolide is degraded rapidly by TCEP and slowly by phosphate and citrate. d) Pure 10-dml (**2**) is degraded rapidly by TCEP. New compounds are observed by LC/MS when (e) narbonolide (**1**) or (f) 10-dml (**2**) is incubated with TCEP for 30 minutes. TIC (total ion chromatogram), EIC (extracted ion chromatogram).



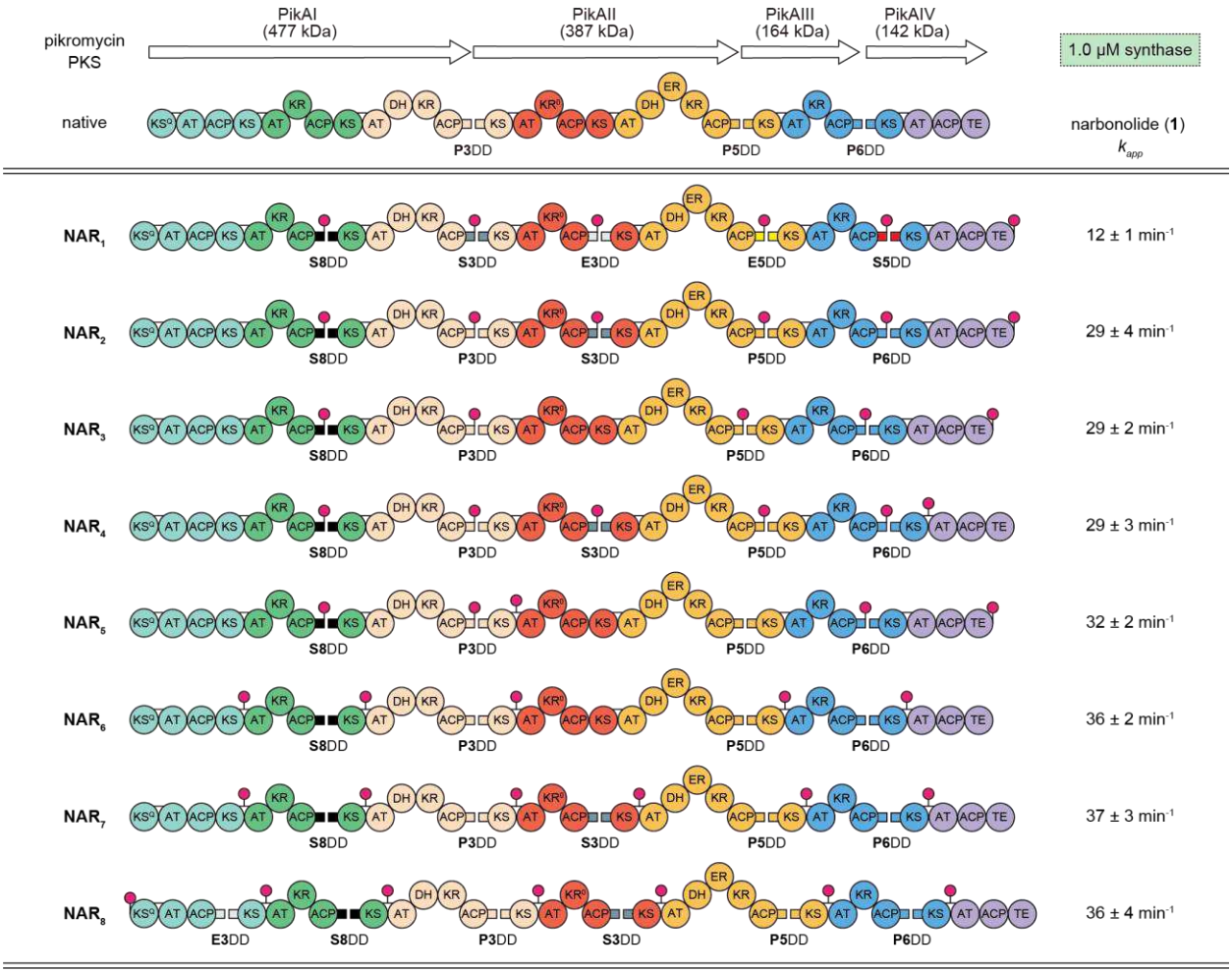
Extended Data Figure 3. Determining the optimal conditions for narbonolide (**1**) synthesis. Production is optimal in a) 200 mM citrate or 400 mM phosphate, b) at pH 7.5, c) in 5% (v/v) glycerol, d) when the synthase is left at room temperature (23 °C) for hours or e) days, and f) at 31 °C.



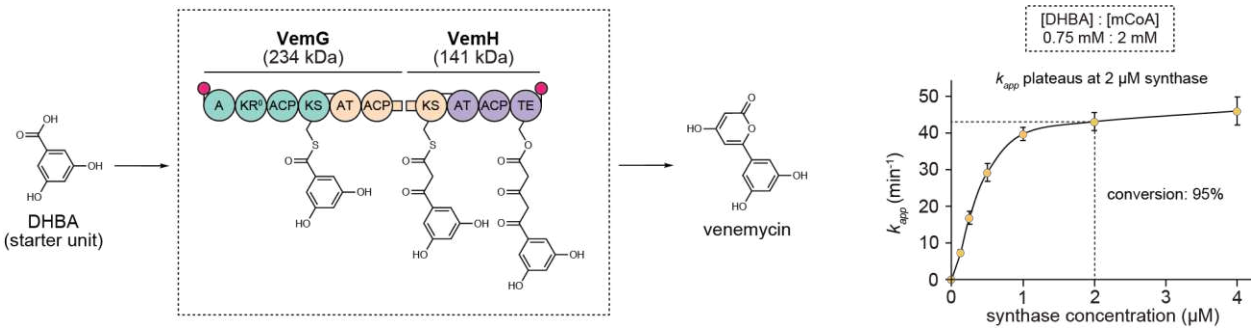
Extended Data Figure 4. Product profiles of **DML**$_1$, **ΔP6**$_1$, and **ANH**$_1$. The diketide intermediate bound to **P2**KS and the tetraketide intermediate bound to **P4**KS are susceptible to offloading by (2*S*)-methylmalonyl-CoA. The resulting generation of triketide **6** and pentaketide **7** competes with the synthesis of the desired polyketides, 10-dml (**2**), **4**, and anhydronarbonolide (**5**).



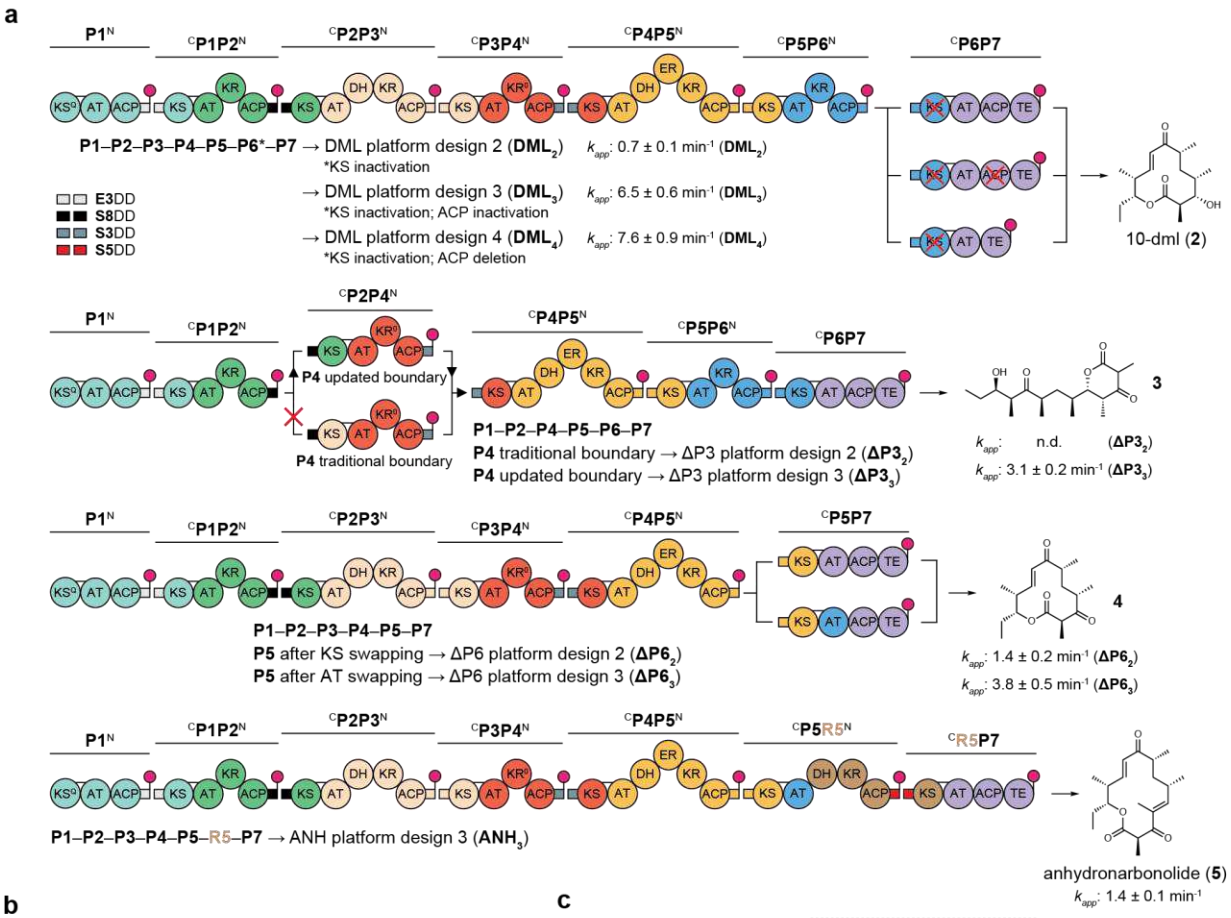
Extended Data Figure 5. Reactions competing with the production of desired polyketide from a module-deleted PKS. a) **P1-P2-P4-P5-P6-P7** produces hexaketide **3** as well as byproduct **7** through methylmalonyl-CoA offloading. b) While this synthase does not contain any malonyl-specific ATs, malonyl-CoA supplied to the reaction is consumed 5 times faster than methylmalonyl-CoA. Malonyl-CoA hydrolysis by methylmalonyl-selective AT domains is even faster (k_{app}: 100 min^{-1}) in the absence of methylmalonyl-CoA.



Extended Data Figure 6. Synthase architecture affects narbonolide (**1**) production. **NAR1**– **NAR8** test how docking domains, His_6-tags, and polypeptide boundaries influence PKS activity. Using the native pikromycin docking domains and relocating His_6-tags from the C-terminus to the flexible loop between KS and AT improved PKS activity. Introducing docking domains to split polypeptides did not change PKS activity.



Extended Data Figure 7. In vitro reconstitution of venemycin biosynthesis. Purified VemG and VemH were supplied with 3,5-dihydroxybenzoic acid (DHBA) and malonyl-CoA to generate venemycin. k_{app} increased with synthase concentration and plateaued at ~2 μM synthase, reaching 43 ± 2 min^{-1} with ~95% conversion.



Extended Data Figure 8. Optimizing assembly line designs. a) Rate measurements for versions of the engineered **DML**, **ΔP3**, **ΔP6**, and **ANH** synthases. Subsequent increases in k_{app} were obtained through tuning polypeptide stoichiometries. b) Rates of product/byproduct formation for

versions of **NAR** and **DML** at 1.0 μM synthase. While the rate of byproduct formation is similar between **DML$_2$**–**DML$_4$**, the rate of 10-dml (**2**) production is significantly greater for **DML$_3$** and **DML$_4$**, possibly because holo-**P7**ACP inhibited acyl-**P6**ACP docking to TE. c) The k_{app} of **DML$_4$** plateaus at 2 μM synthase, a slightly lower concentration than observed for **NAR$_8$** (4 μM). The addition of thioesterase II (PikAV) increased the **DML$_4$** k_{app} 1.4-fold to 13.6 ± 1.4 min^{-1}.

Compound	Platform	PKS concentration	EPK	PikAV	Isolated yield	Characterization
1	**NAR$_9$**	1 μM all	EPK (10 mL, 2 h)	Yes	7.2 mg	NMR, X-ray
2	**DML$_4$**	1 μM all	optimized EPK 5% (v/v) glycerol 10 μM GDH 20 mL 8 h	Yes	4.6 mg	NMR
3	**ΔP3$_3$**	^{C}**P2P4**^{N}: 2 μM; ^{C}**P4P5**^{N}: 2 μM all others: 1 μM		No	3.6 mg	NMR
4	**ΔP6$_3$**	1 μM all		No	2.7 mg	NMR, X-ray
5	**ANH$_3$**	^{C}**P5**R5^{N}: 4 μM; ^{C}R5**P7**: 16 μM all others: 1 μM		No	5.1 mg	NMR

Extended Data Table 1. EPK reaction conditions, yields, and structural characterizations.

REFERENCES

[1] Keatinge-Clay, A. T. (2017) The Uncommon Enzymology of Cis-Acyltransferase Assembly Lines, *Chem Rev 117*, 5334-5366.

[2] Nivina, A., Yuet, K. P., Hsu, J., and Khosla, C. (2019) Evolution and Diversity of Assembly-Line Polyketide Synthases, *Chem Rev 119*, 12524-12547.

[3] Peng, H., Ishida, K., Sugimoto, Y., Jenke-Kodama, H., and Hertweck, C. (2019) Emulating evolutionary processes to morph aureothin-type modular polyketide synthases and associated oxygenases, *Nat Commun 10*, 3918.

[4] Miyazawa, T., Bista, R., and Keatinge-Clay, A. T. (2026) Refactoring pikromycin synthase for the modular biosynthesis of macrolide antibiotics in E. coli, *Nature Synthesis*.

[5] Kudo, K., Hashimoto, T., Awakawa, T., Zhang, L., Nishimura, T., Hashimoto, J., Kozone, I., Kagaya, N., Suenaga, H., Keatinge-Clay, A. T., Abe, I., and Shin-Ya, K. (2026) Skeletal editing via multi-step engineering of a modular polyketide synthase, *Nat Commun 17*.

[6] Miyazawa, T., Fitzgerald, B. J., and Keatinge-Clay, A. T. (2021) Preparative production of an enantiomeric pair by engineered polyketide synthases, *Chem Commun (Camb) 57*, 8762-8765.

[7] Zhang, J., Bista, R., Miyazawa, T., and Keatinge-Clay, A. T. (2023) Boosting titers of engineered triketide and tetraketide synthases to record levels through T7 promoter tuning, *Metab Eng 78*, 93-98.

[8] Gao, X., Woo, S. K., and Krische, M. J. (2013) Total synthesis of 6-deoxyerythronolide B via C-C bond-forming transfer hydrogenation, *J Am Chem Soc 135*, 4223-4226.

[9] 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., and Abe, I. (2017) Characterization of Giant Modular PKSs Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides, *Angew Chem Int Ed Engl 56*, 1740-1745.

[10] Keatinge-Clay, A. T. (2017) Polyketide Synthase Modules Redefined, *Angew Chem Int Ed Engl 56*, 4658-4660.

[11] Lowry, B., Robbins, T., Weng, C. H., O'Brien, R. V., Cane, D. E., and Khosla, C. (2013) In vitro reconstitution and analysis of the 6-deoxyerythronolide B synthase, *J Am Chem Soc 135*, 16809-16812.

[12] Miyazawa, T., Hirsch, M., Zhang, Z., and Keatinge-Clay, A. T. (2020) An in vitro platform for engineering and harnessing modular polyketide synthases, *Nat Commun 11*, 80.

[13] Hughes, A. J., and Keatinge-Clay, A. (2011) Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of Streptomyces coelicolor MatB, *Chem Biol 18*, 165-176.

[14] Xue, Y., Zhao, L., Liu, H. W., and Sherman, D. H. (1998) A gene cluster for macrolide antibiotic biosynthesis in Streptomyces venezuelae: architecture of metabolic diversity, *Proc Natl Acad Sci U S A 95*, 12111-12116.

[15] Collin, S., and Weissman, K. J. (2025) Docking domains from modular polyketide synthases and their use in engineering, *Nat Commun 16*, 6690.

[16] Murli, S., Kennedy, J., Dayem, L. C., Carney, J. R., and Kealey, J. T. (2003) Metabolic engineering of Escherichia coli for improved 6-deoxyerythronolide B production, *J Ind Microbiol Biotechnol 30*, 500-509.

[17] Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases, *Biochemistry 37*, 1585-1595.
[18] Dang, T., EswarKumar, N., Tripathi, S. K., Yan, C., Wang, C. H., Cao, M., Paul, T. K., Agboluaje, E. O., Xiong, M. P., Ivanov, I., Ho, M. C., and Zheng, Y. G. (2024) Oligomerization of protein arginine methyltransferase 1 and its functional impact on substrate arginine methylation, *J Biol Chem 300*, 107947.
[19] Keatinge-Clay, A. T., Miyazawa, T., Zhang, J., Ray, K. A., Lutgens, J. D., Bista, R., and Lin, S. N. (2023) Crystal structures reveal the framework of cis -acyltransferase modular polyketide synthases, *bioRxiv*.
[20] Ogura, H., Furuhata, K., Kuwano, H., and Harada, N. (1975) Stereochemistry of macrolides. I. 1a-Conformation of aglycones of pikromycin and narbomycin and their derivatives, *J Am Chem Soc 97*, 1930-1934.
[21] Ad, O., Thuronyi, B. W., and Chang, M. C. (2017) Elucidating the mechanism of fluorinated extender unit loading for improved production of fluorine-containing polyketides, *Proc Natl Acad Sci U S A 114*, E660-E668.
[22] Dunn, B. J., Cane, D. E., and Khosla, C. (2013) Mechanism and specificity of an acyltransferase domain from a modular polyketide synthase, *Biochemistry 52*, 1839-1841.
[23] Kim, B. S., Cropp, T. A., Beck, B. J., Sherman, D. H., and Reynolds, K. A. (2002) Biochemical evidence for an editing role of thioesterase II in the biosynthesis of the polyketide pikromycin, *J Biol Chem 277*, 48028-48034.
[24] Gajewski, J., Pavlovic, R., Fischer, M., Boles, E., and Grininger, M. (2017) Engineering fungal de novo fatty acid synthesis for short chain fatty acid production, *Nat Commun 8*, 14650.
[25] Jayakumar, A., Tai, M. H., Huang, W. Y., al-Feel, W., Hsu, M., Abu-Elheiga, L., Chirala, S. S., and Wakil, S. J. (1995) Human fatty acid synthase: properties and molecular cloning, *Proc Natl Acad Sci U S A 92*, 8695-8699.
[26] Stuible, H. P., Meurer, G., and Schweizer, E. (1997) Heterologous expression and biochemical characterization of two functionally different type I fatty acid synthases from Brevibacterium ammoniagenes, *Eur J Biochem 247*, 268-273.
[27] Fichtlscherer, F., Wellein, C., Mittag, M., and Schweizer, E. (2000) A novel function of yeast fatty acid synthase. Subunit alpha is capable of self-pantetheinylation, *Eur J Biochem 267*, 2666-2671.
[28] Hirsch, M., Fitzgerald, B. J., and Keatinge-Clay, A. T. (2021) How cis-Acyltransferase Assembly-Line Ketosynthases Gatekeep for Processed Polyketide Intermediates, *ACS Chem Biol 16*, 2515-2526.
[29] Hirsch, M., Desai, R. R., Annaswamy, S., and Keatinge-Clay, A. T. (2024) Mutagenesis Supports AlphaFold Prediction of How Modular Polyketide Synthase Acyl Carrier Proteins Dock With Downstream Ketosynthases, *Proteins 92*, 1375-1384.
[30] Aldrich, C. C., Beck, B. J., Fecik, R. A., and Sherman, D. H. (2005) Biochemical investigation of pikromycin biosynthesis employing native penta- and hexaketide chain elongation intermediates, *J Am Chem Soc 127*, 8441-8452.
[31] Singh, K., Bunzel, G., Graf, B., Yip, K. M., Neumann-Schaal, M., Stark, H., and Chari, A. (2023) Reconstruction of a fatty acid synthesis cycle from acyl carrier protein and cofactor structural snapshots, *Cell 186*, 5054-5067 e5016.

[32] Hoogerbrugge, G., Keatinge-Clay, A. T., and Marcotte, E. M. (2026) Serendipity and the Slime Mold: A Visual Survey of High-Molecular-Weight Protein Assemblies Reveals the Structure of the Polyketide Synthase Pks16, *Mol Cell Proteomics 25*, 101484.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Compound4.cif
- Compound1.cif
- Compound4CheckCIF.pdf
- SIsubmission.pdf
- Compound1CheckCIF.pdf