

UC Berkeley

UC Berkeley Previously Published Works

Title

Polyketide synthase-based controlled synthesis of polycyclopropanated fuel molecules

Permalink

<https://escholarship.org/uc/item/01102767>

Journal

Nature Communications, 17(1)

ISSN

2041-1723

Authors

Yin, Kevin

Landera, Alexander

Lee, Namil

et al.

Publication Date

2026-05-27

DOI

10.1038/s41467-026-73172-3

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Polyketide synthase-based controlled synthesis of polycyclopropanated fuel molecules

Received: 7 April 2025

Accepted: 30 April 2026

Published online: 27 May 2026

 Check for updates

Kevin Yin^{1,2,3,4}, Alexander Landera^{5,18}, Namil Lee^{1,2,6,7,18}, Anthony T. Iavarone⁶, Suzanne M. Kosina⁸, Thomas D. Young^{1,2,6}, Kai Deng^{1,9}, Justin Baerwald^{1,2,4}, Yan Chen^{1,2}, Jennifer W. Gin^{1,2}, Riley Benedict¹⁰, Yan Chiu^{4,11}, Ezechinyere Ukabiala^{1,2,12,13}, Methun Kamruzzaman^{10,5}, Kunal Poorey^{5,14}, Trent R. Northen^{1,8}, Christopher J. Petzold^{1,2}, Anthe George^{1,5}, Pablo Cruz-Morales^{15,16,17}, Qingyun Dan^{1,2,4,6} ✉ & Jay D. Keasling^{1,2,4,6,11,15} ✉

Reducing carbon emissions from aviation and long-distance transportation sectors requires the development of sustainable biofuels with suitable energy density, freezing point, and other physical properties. We previously demonstrated biological production of high energy polycyclopropanated fatty acids (POP-FAs, class I) using an iterative polyketide synthase (iPKS) pathway in a *Streptomyces* host. Here, we used a computational model of fuel properties to identify chain length and cyclopropanation control as critical steps to engineer this iPKS for biofuel applications. We next explored the natural diversity of POP biosynthesis by investigating homologous pathways. Then, by in vivo gene exchange, we determined cyclopropanase (CP) catalysis to be key for POP-FA engineering. Leveraging both natural and engineered pathway product diversity, we demonstrate targeted production of improved POP-FAs, namely shortened POP-FAs with predicted superior freezing point properties for aviation, as well as fully cyclopropane-saturated POP-FAs which should have superior energy-density. These precise and controllable modifications to POP-FA structure open the door for bioproduction of designer POP fuels.

Decarbonization of all sectors is urgently needed to mitigate human-made climate change. In tandem with the development of low-emission technologies (solar, wind, and lithium-ion batteries), electric vehicles (EVs) offer significant decarbonization potential for short-range land-based transport¹. However, long-range or heavy duty transport requires higher power-to-weight ratios: the energy density of lithium ion batteries is only ~1 MJ/kg² whereas jet fuels must have a heat of combustion of at least 42.8 MJ/kg³. Furthermore, aviation emissions are projected to increase from 1.03 GtCO₂ in 2019 to up to 1.9 GtCO₂ by 2050, and the pathway to emission reduction relies heavily on replacing fossil fuels with biofuels^{4,5}. Thus, the development of energy-

dense biofuels remains important for mitigating CO₂ emissions from aviation and other sectors that are challenging to electrify.

In addition to increased energy density (e.g., via strained bond angles), the next generation of biofuels will need customizable carbon skeleton design to meet the combustion energy, freezing point, and boiling point demands of aviation⁶. Of all hydrocarbons, the strained cyclopropane ring is among the most energy-dense moieties. Energy-dense cyclopropanes were first demonstrably leveraged for fuel purposes with syntin, which was used with oxygen in Soyuz and Proton rockets^{7–10}. However, organic synthesis of syntin was toxic, costly, and derived from petroleum. To address these drawbacks, synthesis of a

A full list of affiliations appears at the end of the paper. ✉ e-mail: qdan@lbl.gov; keasling@berkeley.edu

polycyclopropanated, high performance bio-kerosene¹¹ has recently been demonstrated via the cyclopropanation of myrcene, which can be derived from tree sap¹². Additionally, a recyclable and more efficient catalyst for Simmons-Smith cyclopropanation, HPW/MCM-41, has been recently demonstrated by the same group¹³. Compared to existing synthon synthesis, this new catalytic method is more sustainable; however, both routes still require a stoichiometric amount of diethylzinc. Biological conversion on the other hand can generate fuels from inexpensive, low-energy carbon sources such as lignocellulosic biomass, providing much greater opportunity for economically viable fuels that are carbon neutral or negative^{14,15}. Our approach therefore was to search polycyclopropanated molecules in nature, design plausible polycyclopropanated structures for biofuel applications based on energy property computation, and engineer their biosynthetic pathways for target biofuel synthesis in a microorganism.

A major way of synthesizing structurally complex natural products is through multidomain enzymatic complexes called polyketide synthases (PKSs). Owing to their large design space¹⁶, PKSs are a promising platform for production of designer biomolecules^{17–19} including biofuels^{20–23}, and have the potential to meet the carbon chain length, functional groups, and structural requirements for the energy density, freezing point, and other desired properties of an aviation biofuel. In particular, we recently demonstrated production of polycyclopropanated fatty acids (POP-FAs) by leveraging an iterative PKS (iPKS) pathway which we named fuelimycin (Fig. 1a)²⁴. The polycyclopropanating iPKS (POP-iPKS) contains a multidomain iPKS with ketosynthase (KS), acyltransferase (AT), dehydratase (DH), and acyl-

carrier protein (ACP) domains complexed to three standalone domains: ketoreductase (KR), cyclopropanase (CP), and thioesterase (TE). While the POP-iPKS produced highly cyclopropanated fatty acids, their chain length and the number of cyclopropane rings were not ideal for jet fuels. To approach the properties desired in aviation fuels, control of POP-iPKS pathways must be engineered.

A key factor of POP-iPKS engineering lies in controlled synthesis of cyclopropanes, a canonical topic in organic chemistry. In fact, cyclopropane biosynthesis is widespread in nature and serves a variety of evolutionary purposes. Within the membrane lipids of a broad range of bacteria including *Escherichia coli*, *Salmonella*, and *Streptococcus*, cyclopropane rings improve tolerance to environmental stressors including high osmotic pressure, low pH, and high temperatures^{25,26}. In the mycolic acids of *Mycobacterium tuberculosis*, cyclopropanes are notably associated with increased pathogenesis^{27–29}. Additionally, several bioactive polyketides contain a single cyclopropane ring³⁰, including cyclizidine^{31,32} (Fig. 1b), curacin³³ (Fig. 1c), ambruticin S³⁴, and toblerol A³⁵. Bioproduction of unnatural cyclopropanes has also been demonstrated via carbene transfer³⁶. Recently, a long chain hydrocarbon containing two cyclopropane rings, named sarekensane, has even been reported from the cuticle of *Vertagopus sarekensis*, a springtail (Collembola) arthropod (Fig. 1d)^{37,38}. Finally, formation of more than two cyclopropane rings in a single biomolecule remains unique to POP-iPKS pathways including that of jawsamycin (FR-900848), U-106305, and fuelimycin^{24,39–42} (Fig. 1a). We named these POP-FAs using a “C[x]:CP[y]” nomenclature, whereby x represents the number of carbons in the “main chain” (i.e., excluding the carbons

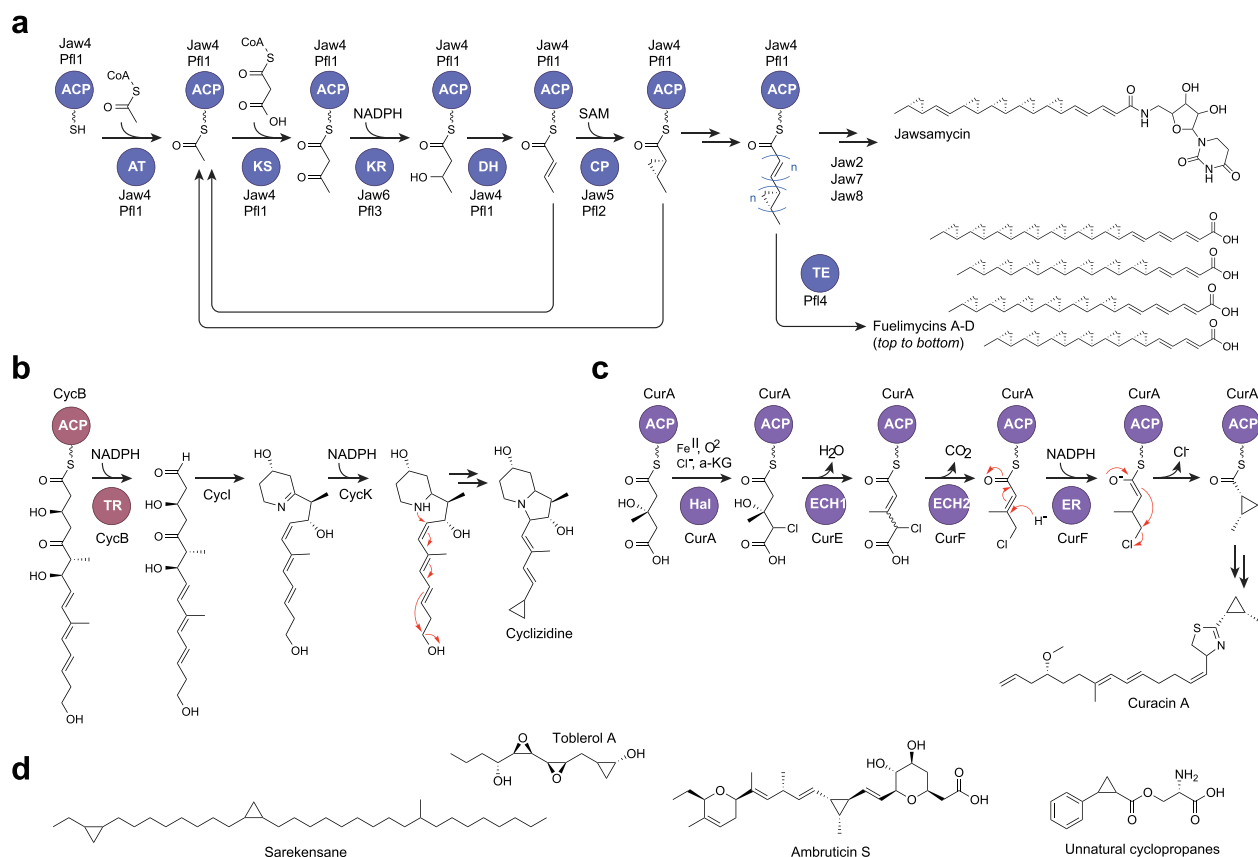


Fig. 1 | Cyclopropane containing natural products and their pathways. a The iterative PKS pathways (jawsamycin, Jaw; fuelimycin, Pfl) for polycyclopropanated (POP) polyketides. **b** The cyclizidine pathway. **c** The curacin pathway. ACP acyl carrier protein, AT acyl transferase, KS ketosynthase, KR ketoreductase, DH dehydratase, CP cyclopropanase, TR terminal reductase, Hal halogenase, ECH1

dehydratase, ECH2 decarboxylase, ER enoyl reductase. **d** Sarekensane, toblerol A, ambruticin S, and unnatural cyclopropanes formed via carbene transfer. Despite the diversity of cyclopropanated structures, none of these molecules have suitable energy density for aviation biofuel applications.

added by the cyclopropanase) and y represents the number of cyclopropane rings. Because POP-FA biosynthesis is iterative, the determining factors for the main-chain length as well as the number and position of cyclopropane rings in the final product(s) are not readily apparent. Without characterized product control mechanisms, design rules cannot be established for engineering POP-iPKS pathways. As such, POP-iPKSs must be further studied to match the engineering potential of assembly-line PKSs with known collinearity between domain architecture and product structure⁴³. Fortuitously, differences in main-chain length and cyclopropanation exist among the fuelimycins (C22:CP7, C20:CP7, C20:CP6, C18:CP6), jawsamycin (C18:CP5), and U-106305 (C18:CP6). Our characterization of the fuelimycin biosynthetic gene cluster (BGC) and its products thus highlights an opportunity to expand genome mining efforts for further discovery of novel POP biochemistry. Subsequently, rational domain exchanges between homologous synthases^{44–48} may reveal the role of each domain in determining the end products and finally open the possibility for programmable biosynthesis of POP molecules.

Here we report diversified natural and chimeric POP-iPKS pathways and demonstrate controllable modifications to the POP-FA structure towards various fuel property targets. First, we identify sequence- and structure-level evidence that POP CP divergence leads to the diversification of POP product chain length and cyclopropane saturation. Then, we demonstrate production of shortened (C12:CP4, C14:CP5) POP-FAs. With shorter chain lengths than fuelimycins and distinct cyclopropane positions, we predict decarboxylated hydrocarbon derivatives of these molecules to be fuels with energy density improvements and comparable freezing points to cyclopropanated myrcene¹¹. Next, we demonstrate production of cyclopropane-saturated POP-FAs (C14:CP6, C16:CP7, C18:CP8, C20:CP9, C22:CP10 and C24:CP11), which are among the most energy-dense molecules that can be produced by a reported biological system. Taken together, our results expand the diversity of POP structures that can be biosynthesized, and lay the groundwork for further rational engineering of POP-iPKS pathways to achieve designer biofuel synthesis and global aviation emissions reduction.

Results

Determining target POP molecules

Upon our first discovery of fuelimycins A–D (Fig. 1a), one critical feature that curtails these molecules from becoming acceptable jet-fuel chemicals is the existence of two or three alkenes in their structures. Being susceptible to oxidative or thermal breakdown^{49,50}, alkenes generally do not meet the stability requirements of sustainable aviation fuels. Furthermore, each double bond represents a missed opportunity for energy-dense cyclopropane ring formation. Based on our previous computational model, the predicted energy density as volumetric net heat of combustion (NHOC) of fuelimycin A-derived fatty acyl methyl ester (FAME), which contains three double bonds (C22:CP7-FAME), is 40.6 MJ/L. In comparison, C12:CP5-FAME's predicted volumetric NHOC is 45.14 MJ/L, despite having ten fewer main-chain carbons and two fewer cyclopropane rings²⁴. Thus, increasing cyclopropane ratio, via replacement of unsaturated hydrocarbons with cyclopropane rings, is a major and necessary fuel upgrade that we sought to pursue biosynthetically.

Furthermore, as aircraft frequently encounter low operating temperatures, aviation fuels must strictly have freezing points below -47°C ³ to be suitable for intercontinental flights and avoid problems associated with crystallization⁵¹. While the freezing point of a structurally similar bio-kerosene (cyclopropanated myrcene) has been characterized¹¹, the estimated titers of fuelimycins have been too low for direct measurement of freezing point and other POP-FA properties (Supplementary Table 1). Instead, we calculated the specific energy, energy density, and freezing point properties of POP molecules theoretically. As our previous model predicted POP-alkanes and -alkenes

to have higher energy densities than POP-FAMES²⁴, we focused on POP-alkanes and -alkenes for our new calculations.

To predict the freezing points of fuel molecules, a machine learning model⁵² trained on a dataset of 19,811 data points⁵³ using 2D PaDel descriptors⁵⁴ demonstrated high accuracy with a mean absolute error (MAE) of 23.69°C after outlier removal refinement. However, the model's performance degraded for cyclopropane-containing oligomers, and particularly for fully cyclopropane-saturated POP molecules. Retraining the model with 37 additional reference freezing points significantly improved its accuracy for these challenging molecules, reducing errors in key examples such as 1,1':2,1''-tercyclohexane and syntin (Supplementary Table 2, Supplementary Data 1). As a result, we observed a positive correlation between the density and freezing point of the POP fuels whose physical properties were calculated. The highest freezing points were found to belong to cyclopropane oligomers without a missing cyclopropane unit. The lowest freezing points were found to occur when a cyclopropane unit is missing and is replaced by a single bond. However, when replaced by a double bond, the freezing point does still decrease considerably relative to the cyclopropane oligomers without a missing cyclopropane unit. The opposite trend occurs when looking at specific energy or energy density. The stereochemistry of a molecule can have profound effects on its melting point. Rotational isomers without bulky chemical groups are able to effectively access all rotational degrees of freedom. These molecules can have lower melting points because it can be harder to adopt a preferred melting point orientation until the temperature drops low enough that molecules are able to align and form crystals. Those with bulky groups can only sample a restricted range of rational angles and thus can more easily conform to a melting point orientation at higher temperature. A simple example of this can be seen from analyzing the trend in melting points for halo-ethanes. The melting points of fluoroethane, chloroethane, bromoethane, and iodoethane are -142.8°C , -138.7°C , -119°C , and -111.1°C , respectively. As the halo group becomes bulkier, the melting point increases because steric hindrance does not allow for free rotation of the $\text{CH}_2\text{-X}$ group, and the preferred orientation for melting is more easily obtained. Of course, there are many other factors that can affect melting point, like atomic charge and dipole moments, and this trend in melting point as a function of bulky chemical groups is only meant to illustrate one factor that affects the melting point of rotational isomers. Our melting point model is trained on a large dataset with a variety of rotational isomers. However, when making predictions it takes in a SMILES notation, which does not include rotational isomer data, and so only takes into account rotational isomerism implicitly.

Our modeled freezing points for syntin and cyclopropanated myrcene were $20\text{--}30^{\circ}\text{C}$ higher than the experimentally determined freezing points of these reference molecules¹¹. To account for this $20\text{--}30^{\circ}\text{C}$ error in our model and align with the freezing point requirements of kerosene-based aviation fuels (e.g., Jet-A at -40°C , Jet-A1 at -47°C), we selected -10°C to -20°C as a conservative freezing point cutoff range (red) for our study³. As expected, our results suggest that the freezing points of POP-FA derivatives decrease as their chain length or total carbon number decreases: alkanes derived from C14-length POP-FAs are acceptable, and C12-length POP-alkanes optimally approach the modeled freezing points of syntin and cyclopropanated myrcene (Fig. 2).

Lastly, to reduce substrate and energy consumption associated with production of sustainable POP fuels while maintaining comparable predicted energy properties to fuelimycins A–D, we reasoned that POP-FAs in the C12–C14 range would lead to more efficient fuel biosynthesis. Based on our previous model, C12:CP3-FAME's predicted volumetric NHOC (38.9 MJ/L) is quite similar to that of fuelimycin B-FAME (C20:CP7-FAME, 39.3 MJ/L)²⁴, despite that fuelimycin B biosynthesis requires input of four more NADPH and four more S-adenosylmethionine (SAM) cofactors. Further, our new model predicted

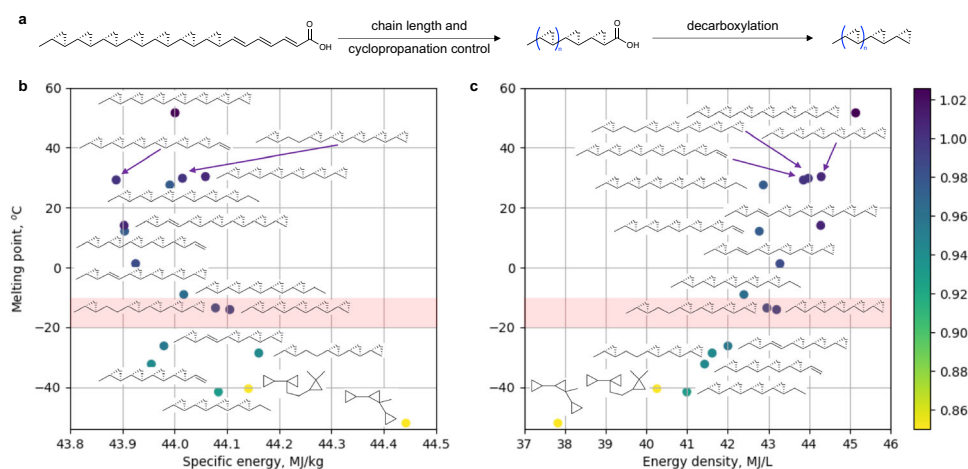


Fig. 2 | Predicted freezing point and energy properties of POP alkanes and other fuels. a Proposed route for POP alkane synthesis. While not the focus of this work, chemical and enzymatic decarboxylation methods may be explored in the future. **b** A plot of freezing point as a function of specific energy (MJ/kg), **c** A plot of

the freezing point as a function of energy density (MJ/L). Scatter points are color-coded based on density at 25 °C (g/cc) based on the colorbar located to the far right of the figure. Source data for this figure are available in Supplementary Data 2.

similarly that C11:CP5 alkane's (derived from C12:CP5 POP-FA) gravimetric NHOC (44.1 MJ/kg) is nearly identical to that of C15:CP7 alkane (44.0 MJ/kg) despite that C16:CP7 POP-FA biosynthesis requires an input of two more NADPH and two more SAM cofactors (Fig. 2). Together with the mentioned benefit of decreasing freezing point, our energy property computation further confirmed our target POP-FAs to be C12–C14 main-chain length.

Bioprospecting POP-iPKS diversity via phylogenetic and structural analysis

As we explored additional POP-iPKS pathway homologs, we renamed the fuelimycin pathway genes in *S. albireticuli* NRRL B-1670 from *pop* to *pfl*—e.g., *pfl1* (iPKS), *pfl2* (CP), *pfl3* (KR), and *pfl4* (TE)—to better distinguish fuelimycin gene/protein homologs from other homologous POP pathways that may also produce POP-FAs. The jawsamycin pathway correspondingly contains core PKS homologs *jaw4* (iPKS), *jaw5* (CP), and *jaw6* (KR), yet with no homologous TE gene.

To diversify POP-FA chemistry and to produce our target molecules, we performed an updated phylogenomic search for pathway homologs. While our previous search covered a broader range of bacterial genomes²⁴, here we used the amino acid sequences of Pfl1 and Pfl2 as queries to search through a more specialized database of 3868 actinobacterial genomes and 2502 MiBiG⁵⁵ entries. From this approach we identified 13 additional candidate BGCs, opening further possibilities for POP biochemical diversity. While we previously identified two major clades including (1) the fuelimycin BGC and its close homologs, and (2) the jawsamycin BGC and its close homologs, our updated search revealed a third clade that includes (3) the *S. longispororuber* BGC and one other homolog (Fig. 3a). We named the core PKS homologs from the *S. longispororuber* BGC *log1* (iPKS), *log2* (CP), *log3* (KR), and *log4* (TE).

To further probe for structural clues indicating POP natural product diversification, we generated AlphaFold 3⁵⁶ models of truncated KS-AT didomains, stand-alone CPs and KRs, and iPKS-CP-KR complexes from the Pfl, Jaw and Log pathways (Supplementary Figs. 1–2). Because of the need to investigate cyclopropanation control, we focused on AlphaFold-based structural analysis of the cyclopropanases. Phylogenetic analysis and AlphaFold 3 prediction revealed that the stand-alone CP enzyme resembles HemN, the *E. coli* oxygen-independent coproporphyrinogen III oxidase⁵⁷ as well as a two-component system involving radical SAM enzyme C10P and methyltransferase C10Q⁵⁸. In agreement with the HemN crystal structure,

AlphaFold 3-predicted POP CP structures reveal a highly conserved N-terminal catalytic subdomain featuring an invariant “C_x-C_xC” motif⁵⁹ (Jaw5 C68, C76 and C79) coordinating the 4Fe-4S cluster, and a less conserved smaller C-terminal subdomain. Interestingly however, these iPKS CPs are generally 20–40 residues longer than HemN at the C-terminus, with varying lengths across different CPs (Supplementary Fig. 3). Moreover, the diversity on both the structure and sequence (Supplementary Fig. 4)⁶⁰ level within the POP CP family suggests that pathway-to-pathway variations in POP product programming may be at least partially encoded in the CP domain, which we sought to test in vivo.

Characterizing natural POP product diversity by heterologous expression

Analyzing Pfl protein abundance effect on fuelimycin titers. In our previous study, we successfully expressed *pfl1-4* genes together with a phosphopantetheinyl transferase (PPTase) gene encoded in ORF.1973 in *Streptomyces coelicolor* M1152, making the fuelimycin-producing POP3.4 strain. However, measuring differences between POP pathways requires sufficiently stable product titers. To improve and stabilize the POP-FA titers, we first probed heterologous expression levels in strain POP3.4. While we observed expression of Pfl1 and Pfl2 at around 1–3% of the total protein abundance, proteomics and RNAseq analysis revealed much lower expression of Pfl3, Pfl4 and ORF.1973 PPTase (Fig. 3b). Natively, the *pfl1-3* genes in *S. albireticuli* NRRL B-1670 strain are likely in a single operon. As the refactoring of the pathway in strain POP3.4 may have altered the native ratio of the Pfl proteins, we decided to re-establish native protein abundance and evaluate the expression of each gene in this cryptic operon, resulting in *S. coelicolor* strain KY2. After measuring similar expression levels of *pfl1-pfl3* in both POP3.4 and KY2 (Fig. 3c), we continued with the POP3.4 refactoring strategy of pairing each gene with its own dedicated promoter, for the ease of further engineering.

Based on typical PKS domain stoichiometry⁴³, we hypothesize that Pfl1 (iPKS), Pfl2 (CP), Pfl3 (KR) and Pfl4 (TE) together should form a protein complex in a 1:1:1:1 ratio. Thus, we reasoned that increasing *pfl3* and *pfl4* expression levels may lead to more functional Pfl complexes and ultimately higher POP-FA titers. Moreover, as a trans-acting TE, higher protein abundance of Pfl4 may lead to early termination of polyketide biosynthesis and shorter-chain POP-FAs. To test these hypotheses, we increased *pfl3* and *pfl4* gene copy numbers by utilizing additional integration sites and replaced underperforming promoters

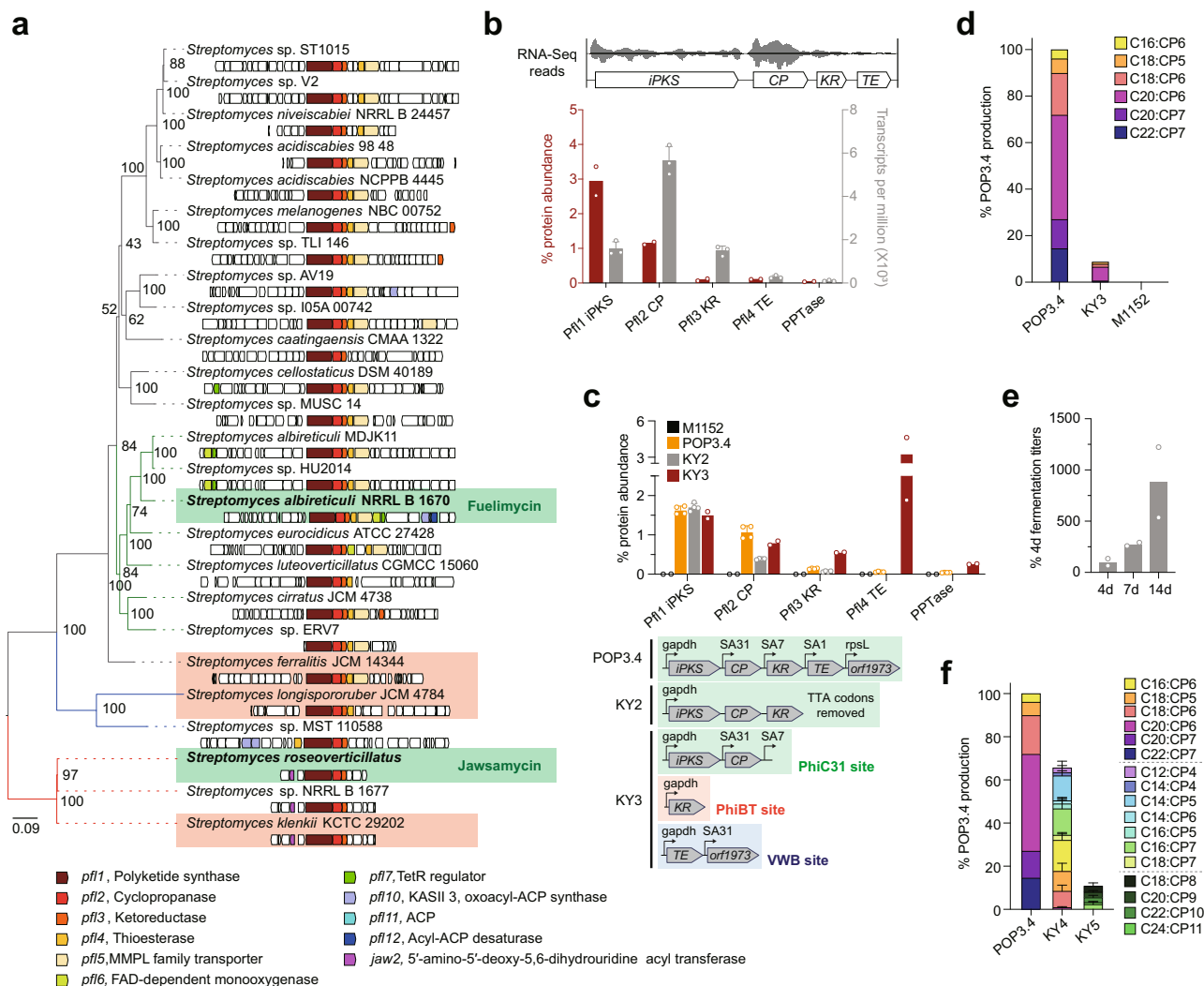


Fig. 3 | Strategies to improve POP-FA production. a Diversity and conservation of POP-*iPKS* BGCs. Pathways highlighted in green have characterized products; genes from pathways highlighted in orange were selected for characterization in this study. **b** Transcriptomic (gray, $n = 3$) and proteomic (red, $n = 2$) analysis of *pfl1-pfl4* and *orf1973* expression in POP3.4 strain. **c** Genetic reorganization of Pfl pathway and their effect on protein expression levels (M1152 and KY3, $n = 2$; POP3.4 and KY2, $n = 4$), and on **d** POP-FA production ($n = 2$). **e** Effect of fermentation time on POP-FA

production ($n = 2$). **f** Production titers and diversity of POP-FAs from fuelimycin (POP3.4, $n = 2$), jawsamycin (KY4, $n = 5$) and longisporuber (KY5, $n = 4$) POP pathways. All data are presented as mean values; error bars indicate the s.d. where n = biological replicates and two experimental repeats, with the exception of transcriptomics (no repeats). All strain information can be found in Supplementary Data 4. Source data for this figure is available in the Source Data file; the full transcriptomics data is available in GSA under accession code CRA039551.

(Fig. 3c). Many *pfl* expression ratios were achieved with the aforementioned strategies, making *pfl3/pfl4*-overexpressing KY3 and KY6, *pfl3*-overexpressing KY7 and *pfl4*-overexpressing KY8. We then checked the POP-FA product profiles of these engineered strains.

To our surprise, we observed a universal decrease of POP-FA titers compared to POP3.4 strain (Fig. 3d; Supplementary Fig. 5). This observation suggested that matching the native *pfl* expression ratio may be important for optimal heterologous POP-FA production. Furthermore, *pfl4* overexpression did not significantly alter the main-chain length distribution of POP-FAs. Alongside its distinctive feature of being a Hotdog-fold TE²⁴, this result reinforces that *Pfl4* behaves very differently from common stand-alone thioesterases⁶¹, motivating us to test alternative *pop* pathways for shortened C12-C14 POP-FA production.

Testing alternative hosts and cultivation media optimization. Before moving on to homologous *pop* pathway investigation, we decided to investigate common metabolic engineering strategies for POP-FA titer and production stability improvement. We recently reported that

exogenous addition of pyrroloquinoline quinone (PQQ) improves production of secondary metabolites in a broad range of *Streptomyces*, as well as activates previously cryptic BGCs⁶². However, exogenous addition of PQQ⁶² did not activate *pfl1-4* in the native *S. alberticuli* NRRL B-1670 strain (Supplementary Fig. 6). Furthermore, media supplementation of PQQ, glucose, and malonate⁶³⁻⁶⁵ each did not improve fuelimycin production in POP3.4 (Supplementary Fig. 7).

We also tested other fast-growth *Streptomyces* chassis strains, including *S. venezuelae* ATCC 15439 and *S. albus* J1074. We observed expression of all heterologous proteins in *S. venezuelae* ATCC 15439^{66,67} (Supplementary Fig. 8) yet only ~10% of fuelimycin production compared to POP3.4 (Supplementary Fig. 9). Additionally, *pfl2*, *pfl3* and ORF.1973 expressed poorly in *S. albus* J1074 (Supplementary Fig. 8), confirming *S. coelicolor* to be the preferred host for our study.

Finally, synthesis of polyketides and other natural products primarily occurs during a late phase of growth, after *Streptomyces* has switched to secondary metabolism⁶⁸⁻⁷⁰. We thus extended the cultivation time from previously reported 4 days to 7 days, and observed 1.9- to 4.4-fold increased and more consistent product titers. Further

extending cultivation time to 14 days reduced consistency, but presented an additional increase in titers by 1.8- to 4.8-fold, to a total increase of 4- to 18-fold as compared to 4 day cultivations (Fig. 3e). We determined that together with optimized POP-FA extraction methods⁷¹, a 7-day cultivation time enabled a balance of consistent, increased titers and rapid Design-Build-Test cycles to explore other *pop* gene clusters in *S. coelicolor*, detect a wide range of POP-FA products with diverse chain length, and systematically compare product profiles despite a potentially limited titer for minor POP-FA products. Given the increased yet inconsistent titers, we reserved scaled-up 14-day cultivations for compound isolation and more detailed analysis of products from selected pathways.

Identifying product profiles of homologous *pop* gene clusters.

Besides fuelimycins, POP-containing jawsamycins were also reported in *Streptomyces roseovorticillatus*⁴². Compared to major fuelimycin products, which range from C18 to C22 in main-chain length and contain 6 to 7 cyclopropane rings, the jawsamycin pathway has a maximum reported main-chain length of C18 and maximum reported cyclopropane number of 5⁴². However, the relative amounts of the jawsamycin major and minor products have not been described. Furthermore, lacking a native TE, the jawsamycin pathway does not produce any POP-FA. Therefore, to directly compare the profile of POP-FA products between jawsamycin and fuelimycin pathways, we introduced the core jawsamycin PKS genes *jaw4* (iPKS), *jaw5* (CP), and *jaw6* (KR) together with *pfl4* (TE) and ORF.1973 (PPTase) from *S. albireticuli* into *S. coelicolor* M1152, creating strain *S. coelicolor* KY4. We refer to the polycyclopropane PKS in KY4 as JJJJ with the first letter designating the iPKS, the second letter the CP, the third letter the KR, and the fourth letter the TE. J refers to *jaw* subunits, and P refers to *pfl* subunits. We then compared the POP-FA products from each *pop* pathway by LC-MS analysis.

The POP3.4 strain (PPPP) produced fuelimycins ranging from 16 to 22 carbons in main-chain length, with at least one double bond. We named them class I POP-FAs. In comparison, the jawsamycin core pathway in strain KY4 produced a shorter POP-FA profile, ranging from 12 to 18 carbons long, achieving our major engineering goal of shortening POP-FA main-chain length. The chain length differences between POP3.4 and KY4 also opened the possibility of POP-FA chain length control via a hybrid *pop* system. Moreover, unlike fuelimycins, jawsamycin has a carbon-carbon double bond between two cyclopropane rings as a result of a cyclopropanation skip in its second iterative cycle^{39,42,72} (Fig. 1a). We name the products of this modified jawsamycin pathway (JJJP), which we expect would share the skipped-cyclopropane pattern, class II POP-FAs.

We noticed that, out of all 13 POP-FAs produced by POP3.4 and KY4, only 2 from KY4 were fully cyclopropanated; these cyclopropane-saturated POP-FAs comprised less than 20% of total KY4 products, driving us to search for PKSs that would produce a higher cyclopropane saturation ratio. From our phylogenetic analysis, we identified a highly divergent POP-iPKS pathway from *Streptomyces longispororuber* (Fig. 3a). We therefore introduced the core *S. longispororuber* POP-iPKS pathway—*log1* (iPKS), *log2* (CP), *log3* (KR), and *log4* (TE)—into *S. coelicolor* M1152 creating strain KY5, harboring the LLLL PKS (L refers to *log* subunits). As a result, KY5 (LLLL) produced class III POP-FAs between 16 and 24 carbons long, all of which were fully saturated with cyclopropane rings (Fig. 3f; Supplementary Fig 10), presenting a distinct class of molecules compared to the larger number of double bonds in both fuelimycin (POP3.4) and jawsamycin (KY4) pathway products.

The cyclopropane position changes proposed in these different classes of POP-FAs were previously unreported to the best of our knowledge, and may offer a novel strategy for cyclopropane synthesis and engineering in a biological system. Our theoretical calculations suggest that alkanes derived from class II and III POP-FAs have increased specific energy compared to alkanes derived from class I

POP-FAs. Additionally, class III POP alkanes achieve a comparable predicted freezing point as class II POP alkanes in one fewer round of CO₂-emitting chain extension (Fig. 2; Supplementary Data 2). Control over the placement of alkene group(s) may also enable fine-tuning of properties, like freezing point, with downstream chemical modifications such as methylation or hydrogenation. However, class I POP-FAs in the POP3.4 strain remained the most productive: KY4 (JJJP) and KY5 (LLLL) achieved only 66% and 10% of the total product titers of POP3.4, respectively. With these diverse pathways available and with three classes of POP-FA products characterized, we next sought to investigate as well as combine the product titer, chain length, and/or cyclopropane saturation advantages of each pathway.

Investigating POP-iPKS programming with chimeric pathways

The recombining of domains or modules from different type I multi-modular PKS pathways into a chimeric PKS is a proven approach for producing novel or desired polyketides^{20,22,23}. For iterative syntheses, in vivo gene exchange serves the additional purpose of elucidating product programming mechanisms due to the nature of the pathway architecture^{44–46,48}.

Engineering control of chain length. The fuelimycin products have on average 3–6 more total carbons than the jawsamycin products, thus opening an interesting opportunity to investigate PKS main-chain length control. To investigate the effect of each PKS domain on the chain length profile of its products, we constructed chimeric pathways in which 1–2 domains from the fuelimycin PKS (PPPP) were substituted with the corresponding jawsamycin domain (J)(Fig. 4a, b).

We first observed that KR substitution from Pfl3 to Jaw6 (strain KY9 harboring PJJP) did not significantly impact the chain length distribution of POP-FA products; similar to the native fuelimycin pathway (POP3.4), 60–70% of total products had a main-chain length of 20–22 carbons, while 20–35% of products had a main-chain length of 18 carbons.

Next, we tested replacement of the fuelimycin iPKS with the jawsamycin iPKS (strain KY10 harboring JPPP) and observed a decrease in the proportion of POP-FAs with 20 carbons (main-chain length) from 50–65% to 35–40%, an increase in abundance of C18 main chain products from 20–25% to 40–55%, and virtually no detection of C22 main chain products.

Finally, substitution of the fuelimycin CP domain with that from jawsamycin (strain KY12 harboring PJPP) resulted in a more significant decrease in main-chain length profile: approximately 55% of total POP-FAs were C18 (main chain), and the remaining 45% were C16. By comparison, the core jawsamycin pathway in KY4 (JJJP) produced the shortest profile: only 30% of total products were C18 (main chain), 45% were C16, and 25% were C14 and C12 (Fig. 4). These results support that the CP, and to a lesser extent the iPKS, are primary determinants, and thus engineering targets, for POP-FA chain length.

Engineering control of cyclopropane ratio. Having found modifications that alter POP-FA main-chain length, we next sought to engineer cyclopropane ratio. An increase in the cyclopropane ratio is desirable, as each cyclopropane ring has a strain energy of 27 kcal/mol⁷³, and thus additional cyclopropane rings increase the energy density of the POP-FA molecule while replacing energetically unfavorable, oxidatively unstable double bonds. Given that the CP domain performs the cyclopropanation reaction, we reasoned that it would play a major role in determining cyclopropane ratio, alongside its demonstrated role in determining main-chain length. To test this hypothesis, from our earlier phylogenetic analysis, we selected three CP genes from *S. ferralitis* (*ferCP*), *S. klenkii* (*kleCP*), and *S. longispororuber* (*log2*) (Fig. 5a) based on evolutionary distance, substituted them into POP3.4, and confirmed their expression with proteomics analysis (Supplementary Fig. 11).

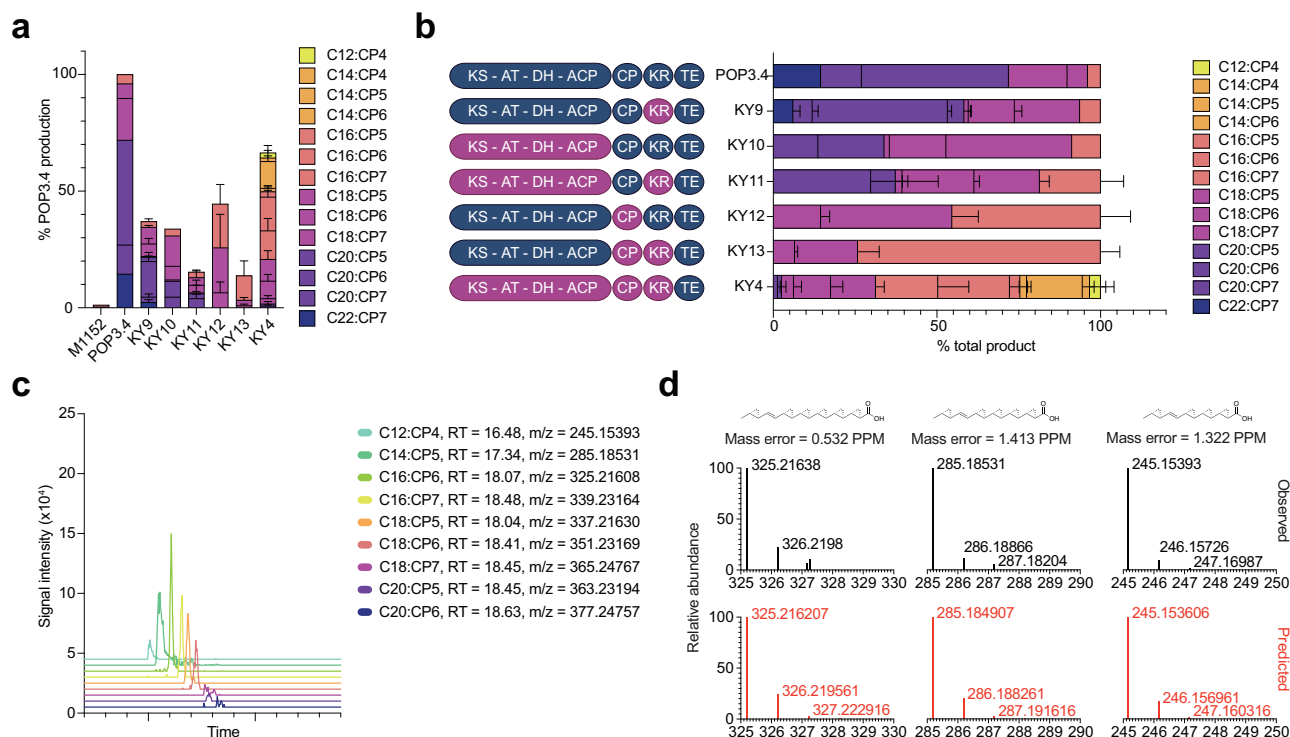


Fig. 4 | Diversification of POP-FA chain lengths. Comparisons of **a** total production and **b** chain length profile of POP-FAs from chimeric strains. Data are presented as mean values; for $n > 2$, error bars indicate the s.d. of biological replicates ($n = 1$ for M1152; $n = 2$ for POP3.4 and KY10; $n = 3$ for KY11, KY12, and KY13; $n = 4$ for KY9; $n = 5$ for KY4); two experimental repeats. **c** Extracted ion

chromatograms for strain KY4 (*jaw4-jaw6*, *pfl4*; JJJP); three experimental repeats. **d** Observed (black) and predicted (red) ions of C16:CP6, C14:CP5, and C12:CP4, two experimental repeats. All strain information can be found in Supplementary Data 4. Source data for this figure is available in the Source Data file.

Notably in line with their evolutionary proximity, the *ferCP* substitution in strain KY14 (PFPP) resulted in a product profile resembling that of the full fuelimycin pathway (POP3.4 with PPPP), while the *kleCP* substitution in strain KY15 (PKPP) resulted in a product profile resembling that of the jawsamycin CP substitution in KY12 (PJPP). While approximately 65% of total POP-FA products from strains POP3.4 (PPPP) and KY14 (PFPP) contained 3 double bonds and 25–30% of products contained 2 double bonds, KY15 (PKPP) and KY12 (PJPP) produced a higher cyclopropane ratio, with 40% of total products containing 3 double bonds and 40–45% of total products containing 1 double bond. Finally, the *log2* CP substitution in strain KY16 (PLPP) resulted in a further increase to the cyclopropane ratio of POP-FA products, with 40% of total products containing 1 double bond and 45% being fully cyclopropanated POP-FAs (C16:CP7, C14:CP6) (Fig. 5).

Notably, KY16 (PLPP) matched the higher product titers of POP3.4, and compared to KY5 (LLLL), produced a shorter POP-FA profile with 60–70% of products having 16 carbons in the main chain. KY16 (PLPP) had a $> 400\%$ increased class III POP-FA titer (and 900% increased total titer) compared to KY5 (LLLL), further illustrating in vivo gene exchange as an efficient strategy for PKS-based POP-FA engineering. In addition to its role in chain length determination, these results demonstrate the CP domain to be a key engineering target for cyclopropanation control.

Given its combined titer, cyclopropanation, and chain length advantages, we selected strain KY16 (PLPP) for 14-day cultivations (Fig. 3f), scaled up to 3 L, for the purpose of HPLC purification and structural characterization of products. In bioreactor format, we again observed a late-stage production increase corresponding to glucose depletion⁷⁴ (Supplementary Fig. 12); however, relative production titers remained higher when strains were cultivated in flasks. Based on our standard curves (Supplementary Table 1; Supplementary Figs. 13–14), strain KY16 produced an estimated 5.40–6.24 mg/L total

POP-FAs in 50 mL flask cultures and 1.37–1.58 mg/L total POP-FAs in 1 L bioreactor cultures. Subsequent LC-MS/MS analysis resulted in fragmentation spectra that confirm the cyclopropane-saturated structures of class III POP-FAs (Supplementary Figs. 15–18); however, our MS/MS analysis confirmed only the number and not the position of carbon-carbon double bonds in unsaturated products from KY16 (PLPP) (Supplementary Figs. 19–22). In total, our engineered strains present an increase to POP-FA diversity from 6 to at least 18 distinct molecules (Table 1; Supplementary Data 3), without including potential variations in carbon-carbon double bond position.

POP-iPKS thioesterase does not play a major role in product programming. Next, to investigate the role of POP-TE on POP product profile, we built and tested a chimeric POP complex with fuelimycin *pfl1-pfl3* (iPKS, CP, and KR genes) and *log4* (TE), resulting in strain KY17 (PPPL). Although we previously observed that KY5 (LLLL) strictly produced fully cyclopropane-saturated POP-FAs, the KY17 (PPPL) product profile instead resembled that of POP3.4, suggesting that POP-iPKS TE domains are promiscuous and do not significantly control POP-FA product profiles (Supplementary Fig. 23). This further supports that the addition of a POP-TE into the jawsamycin pathway, e.g., strain KY4 (JJJP), should produce class II POP-FAs and not modify the distinct cyclopropanation pattern of jawsamycins.

N-terminal subdomain of Log2 cyclopropanase exhibits product programming activity. As described above, comparison of predicted cyclopropanase structures revealed a conserved N-terminal subdomain and a variable C-terminal helix (Supplementary Fig. 3). To investigate the role of each subdomain on product profile, we constructed chimeric *pfl2-log2* and *log2-pfl2* cyclopropanases using two different subdomain boundaries (Supplementary Fig. 24), which we named L-L-P, L-P-P, P-L-L, and P-P-L, and substituted them for the

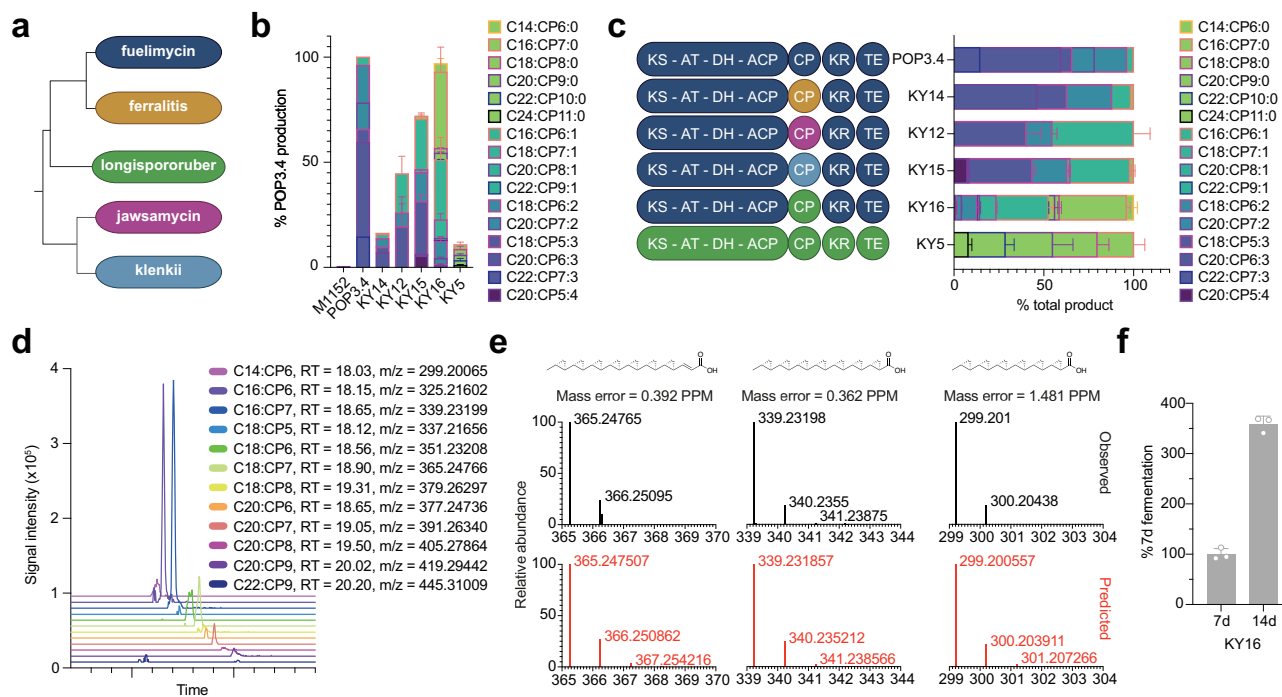


Fig. 5 | Diversification of POP-FA cyclopropanation levels. a Diversity of CP homologs selected for chimeric pathways. Comparisons of **b** total production and **c** cyclopropanation profile of POP-FAs from chimeric strains. To highlight cyclopropane ratios, an alternative “C[x]:CP[y]:[z]” nomenclature is used for POP-FAs, where x represents main chain carbon length, y represents the number of cyclopropane rings, and z represents the number of C = C double bonds. The data are presented as mean values; for $n > 2$, error bars indicate the s.d. of biological replicates ($n = 1$ for M1152; $n = 2$ for POP3.4 and KY14; $n = 3$ for KY12, KY15 and KY16; $n = 4$

for KY5); two experimental repeats **d** Extracted ion chromatograms for strain KY16 (*pfl1*, *log2*, *pfl3*, *pfl4*; PLPP); three experimental repeats. **e** Observed (black) and predicted (red) ions of C18:CP7, C16:CP7, and C14:CP6, two experimental repeats. The position of CP-rings for C18:CP7 is not determined. **f** Effect of fermentation time on POP-FA production in strain KY16 (PLPP). The data are presented as mean values; error bars indicate the s.d. of biological replicates ($n = 3$) with two experimental repeats. All strain information can be found in Supplementary Data 4. Source data for this figure is available in the Source Data file.

cyclopropanase in the fuelimycin pathway, creating strains KY20–KY23, respectively. However, all four chimeric cyclopropanases exhibited poor protein expression, with only KY20 [P(L-L-P)PP] expressing a cyclopropanase to at least 0.35–0.5% total protein abundance (Supplementary Fig 12). Correspondingly, only KY20 produced detectable titers of POP-FAs, 55–80% of which contained a single double bond, while the remaining 20–45% were fully cyclopropanated, and all detected POP-FAs were C18 or shorter (Supplementary Fig 25). This product profile, compared to that of POP3.4, suggests that the *log2* N-terminal subdomain alone was sufficient to increase cyclopropane-saturation over the *pfl* pathway. Thus, variable residues within the otherwise conserved catalytic N-terminal subdomain of the cyclopropanase may influence the programming of this iterative pathway.

Discussion

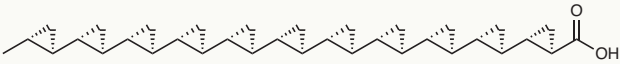
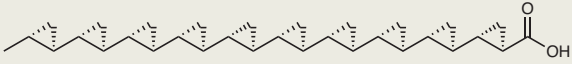
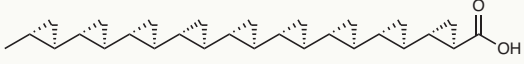
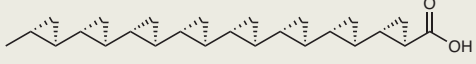
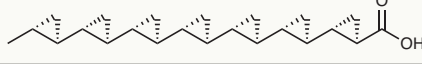
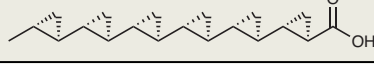
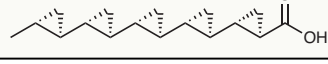
We previously demonstrated bioproduction of energy-dense class I POP fuels. However, the chain length and cyclopropane ratio of these POP-FAs were not ideal for aviation or rocketry applications²⁴. To develop competitive biojet fuels, the iterative PKS mechanisms of POP-FA biosynthesis would need to be understood and engineered. Generally, type I iPKSs have been shown to exhibit a diversity of catalytic programming mechanisms involving the substrate or intermediate specificity of AT, KS, CP, KR, ER, TE, or other domains as well as gatekeeping activity or domain-domain interactions⁷⁵. Our research represents a pioneering investigation into the product programming function of an iPKS CP: we discovered sequence- and structure-level diversity in the CP, and subsequently demonstrated the potential for this domain to be engineered to control the profile of released POP polyketides. We report production of the shortest and most aviation-compatible POP-FAs to date, as well as production of more energy-

dense class III POP-FAs for the first time, thus diversifying POP-FAs to be more customizable towards being energy efficient, thermally stable, energy dense, and/or having a lower freezing temperature.

After comparing POP-FA product profiles between chimeric strains with different CP homologs, we discovered that similarities and differences in chain length and cyclopropanation profiles aligns with CP phylogeny. We can thus predict that uncharacterized CP homologs with closer evolutionary distance to *pfl2* and *ferCP* will likely favor longer POP-FAs with a higher degree of unsaturation. By contrast, CP homologs with closer evolutionary distance to *jaw5* or *kleCP* may favor shorter POP-FAs while those closer to *log2* CP may favor high degrees of cyclopropane saturation. Finally, as bioprospection efforts expand to mine more non-*Streptomyces* genomes for POP-iPKS BGCs, particularly divergent CPs can be prioritized for characterization, and our POP3.4-derived host microbes can serve as a suitable engineering platform.

It was previously hypothesized⁴² that product programming involves a balance between condensation (KS) and cyclopropanation (CP) reaction rates (Fig. 6). The interplay between KS and CP may explain the shorter product profile of chimeric KY16 (PLPP) strain as compared to both of its native LLLL (KY5) and fuelimycin (POP3.4, PPPP) counterparts. After the 6th or 7th iterative polyketide extension, fuelimycin Pfl PKS begins to favor (i.e., exhibit faster reaction rates on) unsaturated intermediates while the *S. longispororuber* Log PKS continues to favor fully cyclopropane-saturated intermediates. The Pfl1 KS outcompetes the Pfl2 CP at those steps where unsaturations are favored. Conversely, the Log2 CP outcompetes the Log1 KS at the corresponding 6th and 7th extensions by the Log PKS, which favors cyclopropanation. In the chimeric PKS, both the Pfl1 KS and Log2 CP would act relatively fast on the 6th and 7th extensions and would no longer consistently form the same intermediates. As the intermediate

Table 1 | Cyclopropane-saturated (class III) POP-FAs produced in this study

Chemical Structure	Name	KY4 (JJJP)	KY5 (LLLL)	KY16 (PLPP)
	C24:CP11	–	+	–
	C22:CP10	–	+	–
	C20:CP9	–	+	+
	C18:CP8	–	+	+
	C16:CP7	+	+	+
	C14:CP6	+	–	+
	C12:CP5	–	–	+

preferences of a chimeric pathway diverge and conflict, condensation by the KS may stall when cyclopropane-saturated extensions occur and cyclopropanation reactions may stall when unsaturated extensions occur. These stalled intermediates could then be released as shorter POP-FAs. Intriguingly however, when increasing cultivation time to 14 days, the product profile of KY16 (PLPP) shifted with a notable increase in the longer C20:CP9 product (Supplementary Fig 26). Here, late-stage mobilization of cellular triacylglycerol (TAG)⁷⁴ increases carbon flux towards polyketide biosynthesis; this may improve the kinetics of the Pfl1 KS reaction to outcompete the TE and perform additional extensions. Finally, the specificity of Pfl1 KS may strictly disfavor substrates longer than C20-length intermediates, as a similar increase in cultivation time to 14 days did not significantly shift the product profile of the native fuelimycin (POP3.4, PPPP) pathway (Supplementary Fig 27). If purification of these enzymatic subunits is achieved and an *in vitro* POP-FA production platform is established, this complex interplay of competing enzyme kinetics and substrate flux can be more deeply explored. Additionally, as more efficient protein engineering methods^{76,77} become applicable to megasynthases, the redesign of KS and CP specificities toward shorter substrates may provide the framework for stricter programming of POP-iPKS products.

Extensive pathway and host engineering must be pursued to achieve the titers, rates, and yields necessary for economically viable fuels. In particular, at least a 1000-fold increase in titers is necessary to reach g/L scale production. Although this is a significant metabolic engineering challenge, such 1000-fold titer increases^{78–80} have been previously achieved. Furthermore, heterologous production on the g/L scale has been demonstrated in *Streptomyces*⁸¹ as well as in other production hosts⁸². These improvements to POP production titers may also enable fuels property testing, NMR structural elucidation, and *in vitro* investigation and engineering of downstream enzymatic activity on purified POP-FAs. Next, strategies for enzymatic or chemical

decarboxylation of POP-FAs into POP-alkenes and the following reduction to eventually give POP-alkanes, which we predict to have superior fuel properties to the current aviation fuels²⁴, must be investigated.

As the light-duty transportation sector has rapidly electrified⁸³, challenges remain for the development of sustainable yet energy-dense aviation, rocketry and long-haul transportation fuels. The continued development of cyclopropanated fuel molecules demonstrates that cyclopropane rings provide sought-after energy properties. Compared to organic synthesis, bioproduction of cyclopropanated fuels will minimize expenses, hazards, and greenhouse gas emissions. Additionally, organic synthesis approaches to polycyclopropane formation either depend entirely on the structure of existing olefin substrates^{11,13}, or involve extensive considerations for precise structure manipulation, including the addition and removal of protecting groups, conversion or removal of *cis* vs *trans* isomers or diastereomers, and identification of chain elongation routes that do not result in cyclopropane loss^{84–86}. By contrast, we have systematically shown that a POP-iPKS pathway can be engineered towards targeted chain lengths, cyclopropane saturation levels, and potentially cyclopropane positions. With the ability to precisely design cyclopropane chemistry, our work lays the foundation for the development of the highest performance jet fuels that can be biosynthesized from a sustainable carbon source.

Methods

To identify POP BGCs, 3868 actinobacterial genomes and 2502 MiBiG⁵⁵ entries were analyzed with a modified CORASON pipeline^{24,71,87} using the amino acid sequences of Pfl1 and Pfl2 as queries. As previously described²⁴, these modifications to CORASON optimized the processing of large genome datasets, eliminated deprecated dependencies, and eliminated the trimming of sequence alignments and

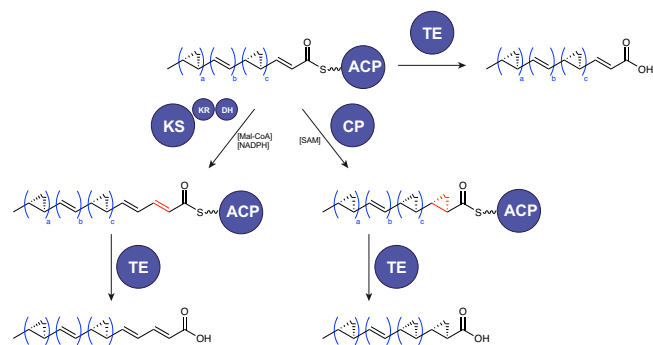


Fig. 6 | Proposed competition-driven product programming mechanism. During each extension cycle, three reactions compete for the ACP-bound intermediate. Depending on the specificity and kinetics of competing enzymes, as well as on substrate flux and likely other factors, each intermediate may undergo either a cyclopropanation-skipping extension (in which the KS outcompetes the CP and TE), a cyclopropanation (in which the CP outcompetes the KS and TE), or chain release (in which the TE outcompetes the KS and CP).

reformatting. The code for this modified pipeline is freely available at [<https://github.com/pablo-genomes-to-vials-cruz>]. *Streptomyces coelicolor* M1152 was obtained from the John Innes Centre and *S. albireticuli* NRRL B-1670 from the NRRL collection. *S. venezuelae* ATCC 15439 was obtained from ATCC. The pfl1, 2, 3, 4 and ORF.1973 genes were cloned from previously described plasmids²⁴ or from genomic DNA. The *jaw4*, *5*, *6*, *ferCP*, *kleCP*, and *log2* genes were codon optimized for *Streptomyces coelicolor* using GenSmart™ Optimization and/or BaseBuddy^{88,89}, and synthesized (GenScript, USA). Promoter elements were cloned from previously described plasmids²⁴ and sequences are provided in Supplementary Data 6.

Cloning and integration of plasmids

All primers and genes used in this study can be found in Supplementary Data 5–6. All plasmids were assembled using NEBuilder® HiFi DNA Assembly. The assembled plasmids were transformed into XL1-Blue (Agilent) or 5- α (NEB) competent *E. coli*, purified and sequenced (Plasmidsaurus, Azenta) to verify their integrity. For integration into *Streptomyces* hosts, a standard conjugation host *E. coli* ET12567 carrying a helper plasmid pUZ8002 was transformed with the verified plasmid and incubated with *Streptomyces*. Ex-conjugants were selected with the appropriate antibiotics, and propagated in SFM media. Successful plasmid integrations were confirmed by polymerase chain reaction (PCR). Additional method details can be found in previously described protocols^{24,71}.

Transcriptomics and proteomics analyses

To investigate transcription levels of the genes in the fuelimycin pathway, we lysed *Streptomyces coelicolor* M1152 POP3.4 with lysozyme at 37 °C for 45 min and extracted RNA using QIAGEN RNeasy Kit. The extracted RNA was then used for whole transcriptome sequencing with rRNA depletion (Azenta, USA). The RNA sequencing reads were mapped to the *Streptomyces coelicolor* M1152 pPOP3.4 reference genome sequence (BioProject PRJNA774980) using CLC Genomics Workbench software (CLC Bio, Aarhus, Denmark). The mapped read data were exported in BAM file format, then converted to a GFF file format to include read counts for each genomic position. The GFF file was subsequently visualized using SignalMap (v2.0.0.5; Roche NimbleGen, Basel, Switzerland).

For proteomics analysis, *Streptomyces* strains were cultured for 4 days in R5 medium with nalidixic acid (30 μ g/ml) and spectinomycin (100 μ g/ml) selection with 4.5 mm glass beads for mycelial dispersion. Cells were harvested and stored at –80 °C until further processing. Protein was extracted from cell pellets and tryptic peptides were

prepared by following established proteomic sample preparation protocol⁹⁰. Briefly, cell pellets were resuspended in Qiagen P2 Lysis Buffer (Qiagen, Germany) to promote cell lysis. Proteins were precipitated with addition of 1 mM NaCl and 4 x vol acetone, followed by two additional washes with 80% acetone in water. The recovered protein pellet was homogenized by pipette mixing with 100 mM ammonium bicarbonate in 20% methanol. Protein concentration was determined by the detergent compatible (DC) protein assay (BioRad, USA). Protein reduction was accomplished using 5 mM tris 2-(carboxyethyl)phosphine (TCEP) for 30 min at room temperature, and alkylation was performed with 10 mM iodoacetamide (IAM; final concentration) for 30 min at room temperature in the dark. Overnight digestion with trypsin was accomplished with a 1:50 trypsin:total protein ratio. The resulting peptide samples were analyzed on an Agilent 1290 UHPLC system coupled to a Thermo Scientific Orbitrap Exploris 480 mass spectrometer for discovery proteomics⁹¹. Briefly, peptide samples were loaded onto an Ascentis® ES-C18 Column (Sigma-Aldrich, USA) and were eluted from the column by using a 10 minute gradient from 98% solvent A (0.1% formic acid in water) and 2% solvent B (0.1% formic acid in acetonitrile) to 65% solvent A and 35% solvent B. Eluting peptides were introduced to the mass spectrometer operating in positive-ion mode and were measured in data-independent acquisition (DIA) mode with a duty cycle of 3 survey scans from *m/z* 380 to *m/z* 985 and 45 MS2 scans with precursor isolation width of 13.5 *m/z* to cover the mass range. DIA raw data files were analyzed by an integrated software suite DIA-NN⁹². The databases used in the DIA-NN search (library-free mode) was *Streptomyces coelicolor* latest Uniprot proteome FASTA sequences plus the protein sequences of the heterologous proteins and common proteomic contaminants. DIA-NN determines mass tolerances automatically based on first pass analysis of the samples with automated determination of optimal mass accuracies. The retention time extraction window was determined individually for all MS runs analyzed via the automated optimization procedure implemented in DIA-NN. Protein inference was enabled, and the quantification strategy was set to Robust LC = High Accuracy. Output main DIA-NN reports were filtered with a global FDR = 0.01 on both the precursor level and protein group level. The Top3 method, which is the average MS signal response of the three most intense tryptic peptides of each identified protein, was used to plot the quantity of the targeted proteins in the samples^{93,94}. The generated mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD057441⁹⁵. DIA-NN is freely available for download from <https://github.com/vdemichev/DiaNN>.

Streptomyces cultivation, POP-FA extraction, LC-MS and LC-MS/MS analysis

Routine cultivation, spore collection, conjugation, and/or protoplast transformation of *Streptomyces* were performed using standard methods⁹⁶. A comprehensive protocol for POP-FA production, extraction, purification, LC-MS analysis has been previously reported^{24,71}; protocol modifications are highlighted in the following sections.

For POP-FA production, *Streptomyces* strains were first cultivated in 3 mL TSB medium⁷¹ with nalidixic acid (30 μ g/ml) and spectinomycin (100 μ g/ml) selection in 24-deep-well plates with 4.5 mm glass beads for mycelial dispersion for 3–4 days. 1 mL of each seed culture was then inoculated into 50 mL R5 medium with nalidixic acid (30 μ g/ml) and spectinomycin (100 μ g/ml) selection in 250 mL Erlenmeyer baffled flasks with 4.5 mm glass beads and cultivated for 4–7 days. For bioreactor experiments, three 2 L (1.8 L max working volume) Eppendorf DASGIP bioreactors were used, each at a starting volume of 1 L. A 50% Dissolved Oxygen cascade setpoint was used, utilizing agitation stir rate changes from 300 to 1300 rpm with a Rushton Impeller. An airflow rate of 30 SL/h was maintained. The DASGIP MX4×4 module

controlled the airflow of sterile dry air. Sampling involved discarding the dead-space of the dip-tube sample port, followed by drawing a 30 mL sample. The setpoint for pH was 7 with a deadband range of 0.2; both acid (H₂SO₄ 15%) and base (NH₄OH 15%) were controlled using the DASGIP MP8 modules.

The POP-FA products were obtained from the *Streptomyces* mycelium, which was lysed by sonication. The sonication pulse settings were modified to 15 s on/15 s off. The pH of the lysate was adjusted to 4, and methanol and chloroform were added to the lysate in a 1:1:1 ratio. The methanol:chloroform:lysate mixture was then vortexed for 10 min, followed by 5 min of centrifugation at $<1000 \times g$ to ensure complete separation of fractions. The chloroform fraction was recovered, and the solvent evaporated. The remaining dense liquid product was resuspended in methanol, vortexed for 10 min and incubated at 20 °C for 8–12 h before solids were separated by centrifugation and removed. These crude extracts were analyzed using the previously reported LC-MS method²⁴; in brief, a 1200 series LC system (Agilent) with a Viva C4 column (Restek, catalog number 9512511) was connected to an LTQ-Orbitrap-XL mass spectrometer with an electrospray ionization source (Thermo Fisher Scientific) and operated in negative ion mode. The resulting data was analyzed using Thermo Xcalibur 2.0.6 Qual Browser software. Using culture extracts of *S. coelicolor* M1152 (no POP-FA production) as a negative control, the abundance of each POP-FA product was calculated by their extracted-ion chromatogram (EIC) peak area; because minor products do not impact total titer nor product profile estimates as significantly as major products, we accepted a rough estimate of minor product abundances when their EICs were more complex or contained split peaks.

For LC-MS/MS analysis of KY16 (PLPP) products, POP-FA were separated using reverse phase chromatography and detected using a Thermo Orbitrap Exploris 120 mass spectrometer; all optimized parameters are available in Supplementary Data 7. Raw datafiles were converted to mzML format using ThermoRawFileParser (version 1.4.2)⁹⁷. MSI and MS2 data were extracted in R using the RaMS package⁹⁸.

Physical property prediction methods

The physical properties predicted in this paper are density, enthalpy of vaporization, specific energy, energy density, and freezing point. Density and enthalpy of vaporization were calculated using the thermo python module⁹⁹, which contains an ordered list of methods for enthalpy of vaporization and density, respectively (ordered based on accuracy). Many of the methods available require thermodynamic parameters, such as T_c, P_c, V_c, and ω . A script was written to calculate these parameters using the Joback method¹⁰⁰. The values of these parameters were then made available to the methods which require them. The ordered list of methods were called iteratively using a loop, until a method was found that had available all the necessary parameters to calculate the target thermodynamic property (either enthalpy of vaporization or density).

In order to calculate the specific energy or the energy density, the enthalpy of combustion must first be calculated. This was done by first optimizing the geometry of each molecule using the B3LYP/6-311 G** method^{101,102}. Subsequent frequency calculations were then carried out at the same level and theory so as to ensure that the optimized geometry is a minimum (a requirement fulfilled by ensuring that all frequencies are real and positive). B3LYP was developed in 1993 and has enjoyed wide success in predicting accurate molecular geometries, even for strained ring systems^{103–106}. While the basis set chosen can have an effect on molecular geometry, research indicates that, in most cases, the choice of basis set has only a small impact on molecular geometry^{107,108}. The 6-311 G** basis set is a split valence triple ζ basis set, and was chosen as a compromise between efficiency and accuracy.

A single point calculation was then carried out using the composite method CBS-QB3¹⁰⁹, which provides enthalpies that are accurate to within a few kilojoules/mole¹¹⁰. Similar calculations were performed for

O₂(Σ_g^-), CO₂, and H₂O. With all enthalpies calculated, and knowledge of the coefficients of the balanced combustion equation, the enthalpy of combustion was obtained. A small correction is required to account for the vaporization of each target molecule, which can be included by calculating the enthalpy of vaporization for each target molecule. Aside from the enthalpy of vaporization, all of the calculations necessary to obtain the enthalpy of combustion were performed using the Gaussian 09 quantum chemistry software¹¹¹. Using this method, the enthalpy of combustion is obtained in units of kJ/mol. In order to obtain the specific energy in units of MJ/kg, the enthalpy of combustion must be divided by the molecular weight of the target molecule in units of kg/mol, and the result must be divided by 1000 (the moles cancel out, and kJ are converted to MJ). From the specific energy (units in MJ/kg), the energy density can be obtained by multiplying the specific energy by the density in units of kg/L (which is equivalent to units of g/cc). For synten, the literature density of 0.851 g/cc was used¹¹, although the calculated density of 0.905 g/cc represents an error of only 6.34%. For cyclopropanated myrcene, there is only a literature density at 20 °C, so a calculated density at 25 °C was used.

Freezing point model

A freezing point model was previously obtained by training a machine learning (ML) model⁵² using a database of 19,811 freezing points⁵³. Here, a brief description of how the model was developed is given. Chemical descriptors were obtained using 2D PaDel descriptors⁵⁴. 2D descriptors provide topological information, and in the training of the freezing point model, this resulted in 1263 features. Several methods were used for training, including linear regression, ensemble-based regression, gradient boosted regression, extreme gradient boosted regression, and deep learning. The best initial performance was obtained when no feature engineering was performed, and when the extreme gradient boosted regression method was used to train the model. Performance was evaluated by calculating the mean absolute error using fivefold cross validation. The best performing model was refined by removing outliers from the training set. Outliers were identified by using the z-score method, which removed 844 freezing points from our training set. The best performing model was one with a mean absolute error of 23.69 °C.

Evaluating the performance of the initial freezing point model against cyclopropane oligomers

Although the initial freezing point model has an excellent mean absolute error, it is important to evaluate its performance against the type of molecules produced in this work. Supplementary Table 1 shows some notable comparisons of the initial freezing point model against data taken from the available literature. The model performs well when a cycloalkane unit is skipped in the polymerization step. This is shown in entry one of the table, where the measured freezing point is 11.5 °C, and the initial freezing point model yields an estimate of 11.7 °C. However, the model's performance is severely degraded when there are no cycloalkane units skipped. This is shown in entry 2 of the table, where the measured freezing point is 18.85 °C, and the initial freezing point model yields an estimate of 52.08 °C. Furthermore, entry 3 and 4 of the table, which are examples of cyclopropane trimers, clearly shows a degraded ability for the initial freezing point model to provide reasonable estimates of the freezing point for the molecules being produced in this work.

In order to improve the performance of our freezing point model, we retrained the model, inserting 37 additional freezing points into the training set. The structures and freezing points of the additional freezing points are available in Supplementary Data 1. The retrained model performs much better than the original freezing point model, as seen from the last column of Supplementary Table 2. The retrained model predicts a freezing point for 1,1':2',1''-tercyclohexane (entry 2 in Supplementary Table 2) of 5.5 °C vs. 52.08 °C using the initial freezing

point model. The retrained model also does better at estimating the freezing point of syntin (entry 3 in Supplementary Table 2), as well as the cyclopropanated myrcene molecule (entry 4 in Supplementary Table 2).

Statistics & reproducibility

No experimental repeats were performed for transcriptomics analysis or bioreactor experiments. Two or more experimental repeats were performed for proteomics analysis and for comparisons of POP-FA production titers and profiles. The quantity of the targeted proteins in the cell culture samples was plotted using the average mass spectrometry signal response of the three most intense tryptic peptides of each identified protein (Top3 method); less intense peptides were not included in protein quantitation analysis. No other data were excluded from the analyses. No statistical method was used to predetermine sample size.

Reporting summary

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

Data availability

The DNA sequences of the fuelimycin BGC from *Streptomyces albir-eticuli* NRRL-B1670 are available at MiBig as [BGC0002138](#) with Protein IDs PAU45552.1 (*pfl1*), PAU45553.1 (*pfl2*), PAU45554.1 (*pfl3*), and PAU45555.1 (*pfl4*). Strains and plasmids used in this study, and listed with JPub IDs in Supplementary Data 4, are available upon request through the Joint BioEnergy Institute Strain Registry [<https://public-registry.jbei.org/folders/946>]. The RNAseq data generated in this study have been deposited in the Genome Sequence Archive¹¹² in National Genomics Data Center¹¹³, China National Center for Bioinformatics / Beijing Institute of Genomics, Chinese Academy of Sciences, under the GSA accession code [CRA039551](#). The protein mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier [PXD057441](#). The processed LC-MS data generated in this study are provided in the Source Data file and Supplementary Data 8. Protein structures used in this work include PDB IDs [4MZO](#), [SBP1](#), [7CPX](#), [1OLT](#). Source data are provided with this paper.

Code availability

The ML model⁵² for freezing point predictions will be detailed in the final published version of the referenced manuscript, which is currently a preprint at the time of writing [<https://doi.org/10.26434/chemrxiv-2024-0jfmh>], while data used to re-train this model is provided in Supplementary Data 1. The code for this model is accessible at DOE CODE repository¹¹⁴: [<https://doi.org/10.11578/dc.20260318.3>].

References

- Calvin, K. et al. Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland. <https://doi.org/10.59327/ipcc/ar6-9789291691647> (2023).
- Liao, J. C., Mi, L., Pontrelli, S. & Luo, S. Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nat. Rev. Microbiol.* **14**, 288–304 (2016).
- Zhang, C., Hui, X., Lin, Y. & Sung, C.-J. Recent development in studies of alternative jet fuel combustion: Progress, challenges, and opportunities. *Renew. Sustain. Energy Rev.* **54**, 120–138 (2016).
- Gössling, S., Humpe, A., Fichert, F. & Creutzig, F. COVID-19 and pathways to low-carbon air transport until 2050. *Environ. Res. Lett.* **16**, 034063 (2021).
- Bergero, C. et al. Pathways to net-zero emissions from aviation. *Nat. Sustain.* **6**, 404–414 (2023).
- Liu, Y. et al. Biofuels for a sustainable future. *Cell* **184**, 1636–1647 (2021).
- Sutton, G. P. History of Liquid Propellant Rocket Engines. (AIAA, 2006).
- Sutton, G. P. & Biblarz, O. Rocket Propulsion Elements. (John Wiley & Sons, 2011).
- Zou, J.-J., Zhang, X. & Pan, L. High-Energy-Density Fuels for Advanced Propulsion: Design and Synthesis. (John Wiley & Sons, 2020).
- Edwards, T. Liquid fuels and propellants for aerospace propulsion: 1903–2003. *J. Propul. Power* **19**, 1089–1107 (2003).
- Liu, Y. et al. Synthesis of strained high-energy rocket bio-kerosene via cyclopropanation of myrcene. *Fuel Process. Technol.* **201**, 106339 (2020).
- Kolichski, M. B., Cocco, L. C., Mitchell, D. A. & Kaminski, M. Synthesis of myrcene by pyrolysis of β -pinene: Analysis of decomposition reactions. *J. Anal. Appl. Pyrolysis* **80**, 92–100 (2007).
- Liu, Y. et al. HPW/MCM-41 catalytic Simmons-Smith cyclopropanation of olefins for synthesis of high-energy-density fuel. *Chem. Eng. Sci.* **283**, 119366 (2024).
- Field, J. L. et al. Robust paths to net greenhouse gas mitigation and negative emissions via advanced biofuels. *Proc. Natl. Acad. Sci. Usa.* **117**, 21968–21977 (2020).
- Tilman, D., Hill, J. & Lehman, C. Carbon-negative biofuels from low-input high-diversity grassland biomass. *Science* **314**, 1598–1600 (2006).
- Pang, B. et al. Technical advances to accelerate modular type I polyketide synthase engineering towards a retro-biosynthetic platform. *Biotechnol. Bioprocess Eng.* **24**, 413–423 (2019).
- Kao, C. M., Katz, L. & Khosla, C. Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science* **265**, 509–512 (1994).
- Regentin, R. et al. Precursor-directed biosynthesis of novel triketide lactones. *Biotechnol. Prog.* **20**, 122–127 (2004).
- Menzella, H. G. et al. Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **23**, 1171–1176 (2005).
- Yuzawa, S. et al. Short-chain ketone production by engineered polyketide synthases in *Streptomyces albus*. *Nat. Commun.* **9**, 4569 (2018).
- Curran, S. C. et al. Probing the flexibility of an iterative modular polyketide synthase with non-native substrates in vitro. *ACS Chem. Biol.* **13**, 2261–2268 (2018).
- Zargar, A. et al. A bimodal PKS platform that expands the biological design space. *Metab. Eng.* **61**, 389–396 (2020).
- Zargar, A. et al. Chemoinformatic-guided engineering of polyketide synthases. *J. Am. Chem. Soc.* **142**, 9896–9901 (2020).
- Cruz-Morales, P. et al. Biosynthesis of polycyclopropanated high energy biofuels. *Joule* **6**, 1590–1605 (2022).
- Grogan, D. W. & Cronan, J. E. Jr Cyclopropane ring formation in membrane lipids of bacteria. *Microbiol. Mol. Biol. Rev.* **61**, 429–441 (1997).
- Poger, D. & Mark, A. E. A ring to rule them all: The effect of cyclopropane fatty acids on the fluidity of lipid bilayers. *J. Phys. Chem. B* **119**, 5487–5495 (2015).
- Glickman, M. S., Cox, J. S. & Jacobs, W. R. Jr A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. *Mol. Cell* **5**, 717–727 (2000).
- Rao, V., Fujiwara, N., Porcelli, S. A. & Glickman, M. S. Mycobacterium tuberculosis controls host innate immune activation through cyclopropane modification of a glycolipid effector molecule. *J. Exp. Med.* **201**, 535–543 (2005).
- Rao, V., Gao, F., Chen, B., Jacobs, W. & Glickman, M. Trans-cyclopropanation of mycolic acids on trehalose dimycolate

- suppresses Mycobacterium tuberculosis -induced inflammation and virulence. *J. Clin. Invest.* **116**, 1660–1667 (2006).
30. Ma, S., Mandalapu, D., Wang, S. & Zhang, Q. Biosynthesis of cyclopropane in natural products. *Nat. Prod. Rep.* **39**, 926–945 (2022).
 31. Huang, W., Kim, S. J., Liu, J. & Zhang, W. Identification of the Polyketide Biosynthetic Machinery for the Indolizidine Alkaloid Cyclizidine. *Org. Lett.* **17**, 5344–5347 (2015).
 32. Peng, H. et al. Deciphering piperidine formation in polyketide-derived indolizidines reveals a thioester reduction, transamination, and unusual imine reduction process. *ACS Chem. Biol.* **11**, 3278–3283 (2016).
 33. Khare, D. et al. Conformational switch triggered by α -ketoglutarate in a halogenase of curacin A biosynthesis. *Proc. Natl. Acad. Sci.* **107**, 14099–14104 (2010).
 34. Julien, B., Tian, Z.-Q., Reid, R. & Reeves, C. D. Analysis of the ambruticin and jerangolid gene clusters of *Sorangium cellulosum* reveals unusual mechanisms of polyketide biosynthesis. *Chem. Biol.* **13**, 1277–1286 (2006).
 35. Ueoka, R., Bortfeld-Miller, M., Morinaka, B. I., Vorholt, J. A. & Piel, J. Toblerols: Cyclopropanol-Containing Polyketide Modulators of Antibiosis in *Methylobacteria*. *Angew. Chem. Int. Ed. Engl.* **57**, 977–981 (2018).
 36. Huang, J. et al. Complete integration of carbene-transfer chemistry into biosynthesis. *Nature* **617**, 403–408 (2023).
 37. Möllerke, A., Bello, J., Leinaas, H. P. & Schulz, S. Cyclopropane hydrocarbons from the springtail *Vertagopus sarekensis*—A new class of cuticular lipids from arthropods. *J. Nat. Prod.* **87**, 85–97 (2024).
 38. Möllerke, A. & Schulz, S. Small animals with unique chemistry – the natural product chemistry of Collembola. *Nat. Prod. Rep.* <https://doi.org/10.1039/d4np00049h> (2024).
 39. Yoshida, M. et al. A novel antifungal antibiotic, FR-900848. I. Production, isolation, physico-chemical and biological properties. *J. Antibiot.* **43**, 748–754 (1990).
 40. Kuo, M. S. et al. Discovery, isolation, structure elucidation, and biosynthesis of U-106305, a cholesteryl ester transfer protein inhibitor from UC 11136. *J. Am. Chem. Soc.* **117**, 10629–10634 (1995).
 41. Hiratsuka, T. et al. Biosynthesis of the structurally unique polycyclopropanated polyketide-nucleoside hybrid jawsamycin (FR-900848). *Angew. Chem. Int. Ed. Engl.* **53**, 5423–5426 (2014).
 42. Hiratsuka, T., Suzuki, H., Minami, A. & Oikawa, H. Stepwise cyclopropanation on the polycyclopropanated polyketide formation in jawsamycin biosynthesis. *Org. Biomol. Chem.* **15**, 1076–1079 (2017).
 43. Soohoo, A. M., Cogan, D. P., Brodsky, K. L. & Khosla, C. Structure and Mechanisms of Assembly-Line Polyketide Syntheses. *Annu. Rev. Biochem.* <https://doi.org/10.1146/annurev-biochem-080923-043654> (2024).
 44. Fisch, K. M. et al. Rational domain swaps decipher programming in fungal highly reducing polyketide synthases and resurrect an extinct metabolite. *J. Am. Chem. Soc.* **133**, 16635–16641 (2011).
 45. Kakule, T. B., Lin, Z. & Schmidt, E. W. Combinatorialization of fungal polyketide synthase-peptide synthetase hybrid proteins. *J. Am. Chem. Soc.* **136**, 17882–17890 (2014).
 46. Yang, X.-L. et al. Molecular basis of methylation and chain-length programming in a fungal iterative highly reducing polyketide synthase. *Chem. Sci.* **10**, 8478–8489 (2019).
 47. Hayashi, S. et al. Control Mechanism for Carbon-Chain Length in Polyunsaturated Fatty-Acid Synthases. *Angew. Chem. Int. Ed. Engl.* **58**, 6605–6610 (2019).
 48. Cox, R. J. Curiouser and curiouser: Progress in understanding the programming of iterative highly-reducing polyketide synthases. *Nat. Prod. Rep.* **40**, 9–27 (2023).
 49. Kopinke, F. D., Zimmermann, G., Reyniers, G. C. & Froment, G. F. Relative rates of coke formation from hydrocarbons in steam cracking of naphtha. 2. Paraffins, naphthenes, mono-, di-, and cycloolefins, and acetylenes. *Ind. Eng. Chem. Res.* **32**, 56–61 (1993).
 50. Pereira, R. C. C. & Pasa, V. Effect of mono-olefins and diolefins on the stability of automotive gasoline. *Fuel* **85**, 1860–1865 (2006).
 51. Affens, W. A., Hall, J. M., Holt, S. & Hazlett, R. N. Effect of composition on freezing points of model hydrocarbon fuels. *Fuel (Lond.)* **63**, 543–547 (1984).
 52. Kamruzzaman, M., Landera, A., Menon, N. C., Allendorf, M. D. & Poorey, K. P2MAT: A machine learning (ML) driven software for Property Prediction of MATERIAL. Preprint at <https://doi.org/10.26434/chemrxiv-2024-0jfmh> (2024).
 53. Bradley, J.-C., Lang, A. S. & Williams, A. J. Jean-Claude Bradley Open Melting Point Dataset. (2014).
 54. Yap, C. W. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **32**, 1466–1474 (2011).
 55. Terlouw, B. R. et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res.* **51**, D603–D610 (2023).
 56. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
 57. Layer, G., Moser, J., Heinz, D. W., Jahn, D. & Schubert, W.-D. Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J.* **22**, 6214–6224 (2003).
 58. Jin, W.-B., Wu, S., Jian, X.-H., Yuan, H. & Tang, G.-L. A radical S-adenosyl-L-methionine enzyme and a methyltransferase catalyze cyclopropane formation in natural product biosynthesis. *Nat. Commun.* **9**, 2771 (2018).
 59. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* **29**, 1097–1106 (2001).
 60. Yariv, B. et al. Using evolutionary data to make sense of macromolecules with a ‘face-lifted’ ConSurf. *Protein Sci.* **32**, e4582 (2023).
 61. Kotowska, M. & Pawlik, K. Roles of type II thioesterases and their application for secondary metabolite yield improvement. *Appl. Microbiol. Biotechnol.* **98**, 7735–7746 (2014).
 62. Wang, X. et al. Elucidation of genes enhancing natural product biosynthesis through co-evolution analysis. *Nat. Metab.* **6**, 933–946 (2024).
 63. Lombó, F. et al. Enhancing the atom economy of polyketide biosynthetic processes through metabolic engineering. *Biotechnol. Prog.* **17**, 612–617 (2001).
 64. Park, S. R., Ahn, M. S., Han, A. R., Park, J. W. & Yoon, Y. J. Enhanced flavonoid production in *Streptomyces venezuelae* via metabolic engineering. *J. Microbiol. Biotechnol.* **21**, 1143–1146 (2011).
 65. Ye, S., Magadán-Corpas, P., Pérez-Valero, Á., Villar, C. J. & Lombó, F. Metabolic engineering strategies for naringenin production enhancement in *Streptomyces albidoflavus* J1074. *Microb. Cell Fact.* **22**, 167 (2023).
 66. Xue, Y., Zhao, L., Liu, H. W. & Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. *Proc. Natl. Acad. Sci. USA* **95**, 12111–12116 (1998).
 67. Lee, N. et al. Re-classification of *Streptomyces venezuelae* strains and mining secondary metabolite biosynthetic gene clusters. *iScience* **24**, 103410 (2021).
 68. Nieselt, K. et al. The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* **11**, 10 (2010).

69. Esnault, C. et al. Strong antibiotic production is correlated with highly active oxidative metabolism in *Streptomyces coelicolor* M145. *Sci. Rep.* **7**, 200 (2017).
70. Zong, G. et al. Use of elicitors to enhance or activate the antibiotic production in streptomyces. *Crit. Rev. Biotechnol.* **42**, 1260–1283 (2022).
71. Yin, K., Cruz-Morales, P., Whitford, C. M. & Keasling, J. D. Heterologous production of polycyclopropanated fatty acids and their methyl esters in *Streptomyces*. *STAR Protoc.* **4**, 102190 (2023).
72. Barrett, A. G. M., Kasdorf, K., Tustin, G. J. & Williams, D. J. Determination of the full structure and absolute stereochemistry of the antifungal agent FR-900848: an X-ray crystallographic study of (1 R, 3 S, 4 R, 6 S, 7 S, 9 R, 10 S, 12 R)-quatercyclopropyl-1,12-dimethanediyl di-4-bromobenzoate. *J. Chem. Soc., Chem. Commun.* **0**, 1143–1144 (1995).
73. Cremer, D. & Gauss, J. Theoretical determination of molecular structure and conformation. 20. Reevaluation of the strain energies of cyclopropane and cyclobutane carbon-carbon and carbon-hydrogen bond energies, 1,3 interactions, and sigma-aromaticity. *J. Am. Chem. Soc.* **108**, 7467–7477 (1986).
74. Wang, W. et al. Harnessing the intracellular triacylglycerols for titer improvement of polyketides in *Streptomyces*. *Nat. Biotechnol.* **38**, 76–83 (2019).
75. Wang, J., Deng, Z., Liang, J. & Wang, Z. Structural enzymology of iterative type I polyketide synthases: various routes to catalytic programming. *Nat. Prod. Rep.* **40**, 1498–1520 (2023).
76. Jiang, K. et al. Rapid in silico directed evolution by a protein language model with EVOLVEpro. *Science* **387**, eadr6006 (2025).
77. Singh, N. et al. A generalized platform for artificial intelligence-powered autonomous enzyme engineering. *Nat. Commun.* **16**, 5648 (2025).
78. Tan, G.-Y. et al. Heterologous biosynthesis of spinosad: An omics-guided large polyketide synthase gene cluster reconstitution in *Streptomyces*. *ACS Synth. Biol.* **6**, 995–1005 (2017).
79. Du, G. et al. Influence of cluster-situated regulator PteF in filipin biosynthetic cluster on avermectin biosynthesis in *Streptomyces avermitilis*. *Biol. (Basel)* **13**, 344 (2024).
80. Zhang, J. et al. A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature* **609**, 341–347 (2022).
81. Wang, X. et al. Heterologous production of chlortetracycline in an industrial grade *Streptomyces rimosus* host. *Appl. Microbiol. Biotechnol.* **103**, 6645–6655 (2019).
82. Banerjee, D. et al. Genome-scale and pathway engineering for the sustainable aviation fuel precursor isoprenol production in *Pseudomonas putida*. *Metab. Eng.* **82**, 157–170 (2024).
83. Yuan, M., Thellufsen, J. Z., Lund, H. & Liang, Y. The electrification of transportation in energy transition. *Energy* **236**, 121564 (2021).
84. Falck, J. R., Mekonnen, B., Yu, J. & Lai, J.-Y. Synthesis of the polycyclopropane antibiotic FR-900848 via the horeau gambit. *J. Am. Chem. Soc.* **118**, 6096–6097 (1996).
85. Barrett, A. G. M. & Kasdorf, K. Total synthesis of the pentacyclopropane antifungal agent FR-900848. *J. Am. Chem. Soc.* **118**, 11030–11037 (1996).
86. Verbicky, C. A. & Zercher, C. K. Olefin cross-metathesis in the preparation of polycyclopropanes: formal synthesis of FR-900848. *Tetrahedron Lett.* **41**, 8723–8727 (2000).
87. Navarro-Muñoz, J. C. et al. A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* **16**, 60–68 (2020).
88. Schmidt, M. et al. Maximizing heterologous expression of engineered type I polyketide synthases: Investigating codon optimization strategies. *ACS Synth. Biol.* **12**, 3366–3380 (2023).
89. Zulkower, V. & Rosser, S. DNA Chisel, a versatile sequence optimizer. *Bioinformatics* **36**, 4508–4509 (2020).
90. Chen, Y. et al. Alkaline-SDS cell lysis of microbes with acetone protein precipitation for proteomic sample preparation in 96-well plate format. *PLoS One* **18**, e0288102 (2023).
91. Chen, Y., Gin, J. & J. Petzold, C. Discovery proteomic (DIA) LC-MS/MS data acquisition and analysis. *J. Exp. Bot.* <https://doi.org/10.17504/protocols.io.e6nvwk1z7vmk/v2> (2022).
92. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
93. Silva, J. C., Gorenstein, M. V., Li, G.-Z., Vissers, J. P. C. & Gromann, S. J. Absolute Quantification of Proteins by LCMSE. *Mol. Cell. Proteom.* **5**, 144–156 (2006).
94. Ahrné, E., Molzahn, L., Glatter, T. & Schmidt, A. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics* **13**, 2567–2578 (2013).
95. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
96. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. *Practical Streptomyces Genetics* (The John Innes Foundation, 2000).
97. Hulstaert, N. et al. ThermoRawFileParser: Modular, scalable, and cross-platform RAW file conversion. *J. Proteome Res.* **19**, 537–542 (2020).
98. William, K. & Ingalls, A. E. Tidy data neatly resolves mass spectrometry's ragged arrays. *R. J.* **14**, 193–202 (2022).
99. Caleb Bell and Contributors. <https://github.com/CalebBell/thermo>. Thermo: Chemical properties component of Chemical Engineering Design Library (ChEDL) <https://github.com/CalebBell/thermo> (2016–2024).
100. Joback, K. G. & Reid, R. C. Estimation of pure-component properties from group-contributions. *Chem. Eng. Commun.* **57**, 233–243 (1987).
101. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
102. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B Condens. Matter* **37**, 785–789 (1988).
103. Hou, X.-J. & Huang, M.-B. Internal rotations of vinylcyclopropane, cyclopropanecarboxaldehyde, and cyclopropanecarboxylic acid: a DFT B3LYP study. *Theochem* **585**, 93–104 (2002).
104. Bernardi, F., Bottoni, A. & Miscione, G. P. DFT study of the palladium-catalyzed cyclopropanation reaction. *Organometallics* **20**, 2751–2758 (2001).
105. Bachrach, S. M. DFT study of [2,2]-, [3,3]-, and [4,4]paracyclophanes: strain energy, conformations, and rotational barriers. *J. Phys. Chem. A* **115**, 2396–2401 (2011).
106. Novak, I. Ring strain in [n]ladderanes. *J. Phys. Chem. A* **112**, 10059–10063 (2008).
107. El-Azhary, A. A. & Suter, H. U. Comparison between optimized geometries and vibrational frequencies calculated by the DFT methods. *J. Phys. Chem.* **100**, 15056–15063 (1996).
108. Wang, M. et al. Molecular geometries and vibrational contributions to reaction thermochemistry are surprisingly insensitive to the choice of basis sets. *J. Chem. Theory Comput.* **19**, 5036–5046 (2023).
109. Montgomery, J. A. Jr, Frisch, M. J., Ochterski, J. W. & Petersson, G. A. A complete basis set model chemistry. VII. Use of the minimum population localization method. *J. Chem. Phys.* **112**, 6532–6542 (2000).
110. Pokon, E. K., Liptak, M. D., Feldgus, S. & Shields, G. C. Comparison of CBS-QB3, CBS-APNO, and G3 predictions of gas phase deprotonation data. *J. Phys. Chem. A* **105**, 10483–10487 (2001).

111. Frisch, M. J. et al. Uranyl extraction by N, N-dialkylamide ligands studied by static and dynamic DFT simulations. *Gaussian* 09 (2016).
112. Zhang, S. et al. The GSA family in 2025: A broadened sharing platform for multi-omics and multimodal data. *Genomics Proteom. Bioinforma.* **23**, qzaf072 (2025).
113. CNCB-NGDC Members and Partners. Database resources of the National Genomics Data Center, China National Center for Bioinformatics in 2025. *Nucleic Acids Res.* **53**, D30–D44 (2025).
114. Kamruzzaman, M. & Poorey, K. sandialabs/P2MAT. DOE CODE repository. <https://doi.org/10.11578/dc.20260318.3> (2026).

Author contributions

K.Y., P.C.-M., Q.D., and J.D.K. contributed to the experimental design. K.Y., R.B., Y.Chui, E.U. and Q.D. performed molecular cloning and strain construction. K.Y., Q.D., K.D., A.T.I., S.K. and T.R.N. performed cultivation, POP-FA extractions, preparative HPLC, LC-MS and LC-MS/MS analysis. K.Y., J.B., and Q.D. performed bioreactor experiments. N.L. and K.Y. performed transcriptomics analysis. Y.Chen, J.W.G., C.J.P., and K.Y. performed proteomic analyses. A.L. M.K., K.P. and A.G. developed physical property calculations. P.C.-M. and Q.D. performed bioinformatic and structural analyses. K.Y., N.L., P.C.-M., T.D.Y., Q.D. and J.D.K. evaluated the data and prepared the manuscript with the assistance of all other authors.

Funding

This work was funded by the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under contract DEAC02-05CH11231 between DOE and Lawrence Berkeley National Laboratory, and by the DOE Distinguished Scientist Fellow Program and Bakar Fellow Fund to J.D.K. Part of this research was supported by the Laboratory Directed Research and Development program (Project 231332) at Sandia National Laboratories, a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525 to A.L., M.K., K.P. and A.G. Work by P.C.M. was supported by the Novo Nordisk Foundation with grants NNF20CC0035580 and NNF24SA0100980.

Competing interests

J.D.K. has financial interests in Ansa Biotechnologies, Apertor Pharma, Berkeley Yeast, BioMia, Cyklos Materials, Demetrix, Lygos, Napigen, ResVita Bio, and Zero Acre Farms. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-73172-3>.

Correspondence and requests for materials should be addressed to Qingyun Dan or Jay D. Keasling.

Peer review information *Nature Communications* thanks Fong Tian Wong who co-reviewed with Yifeng Wei, Valery Polyakov and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

¹Joint Bioenergy Institute, Emeryville, CA, USA. ²Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ³Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. ⁴Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA, USA. ⁵Biomass Science and Conversion Technology Department, Sandia National Laboratories, Livermore, CA, USA. ⁶QB3 Institute, University of California, Berkeley, CA, USA. ⁷Graduate School of Engineering Biology, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea. ⁸Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ⁹Department of Biomaterials and Biomanufacturing, Sandia National Laboratories, Livermore, CA, USA. ¹⁰Department of Chemistry, University of California, Berkeley, CA, USA. ¹¹Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. ¹²Department of Bioengineering, University of California, Berkeley, CA, USA. ¹³Department of Bioengineering, University of California San Francisco, San Francisco, CA, USA. ¹⁴Department of Material Lifecycle Management, Sandia National Laboratories, Livermore, CA, USA. ¹⁵The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet 220, Kongens Lyngby, Denmark. ¹⁶The Novo Nordisk Foundation Biotechnology Research Institute for the Green Transition, Technical University of Denmark, Kongens Lyngby, Denmark. ¹⁷Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. ¹⁸These authors contributed equally: Alexander Landera, Namil Lee. ✉ e-mail: qdan@lbl.gov; keasling@berkeley.edu