Title
Engineering site-selective incorporation of fluorine into polyketides.

Journal
Nature Chemical Biology, 18(8)

Authors
Sirirungruang, Sasilada
Ad, Omer
Privalsky, Thomas
et al.

Publication Date
2022-08-01

DOI
10.1038/s41589-022-01070-y

Peer reviewed

HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2023 August 01.

Published in final edited form as:

Nat Chem Biol. 2022 August ; 18(8): 886–893. doi:10.1038/s41589-022-01070-y.

Engineering site-selective incorporation of fluorine into polyketides

Sasilada Sirirungruang[1], **Omer Ad**[2], **Thomas M. Privalsky**[3], **Swetha Ramesh**[1], **Joel L. Sax**[1], **Hongjun Dong**[2], **Edward E. K. Baidoo**[4,5,6], **Bashar Amer**[4,5], **Chaitan Khosla**[3,7], **Michelle C. Y. Chang**[1,2,8,✉]

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA.

[2]Department of Chemistry, University of California, Berkeley, CA, USA.

[3]Department of Chemistry, Stanford University, Stanford, CA, USA.

[4]Joint Bioenergy Institute, Lawrence Berkeley National Laboratory, Emeryville, CA, USA.

[5]Biological Systems and Engineering, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

[6]Department of Energy, Agile BioFoundry, Emeryville, CA, USA.

[7]Department of Chemical Engineering, Stanford University, Stanford, CA, USA.

[8]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA, USA.

Abstract

Although natural products and synthetic small molecules both serve important medicinal functions, their structures and chemical properties are relatively distinct. To expand the molecular diversity available for drug discovery, one strategy is to blend the effective attributes of synthetic and natural molecules. A key feature found in synthetic compounds that is rare in nature is the use of fluorine to tune drug behavior. We now report a method to site-selectively incorporate fluorine into complex structures to produce regioselectively fluorinated full-length polyketides.

✉ **Correspondence and requests for materials** should be addressed to Michelle C. Y. Chang. mcchang@berkeley.edu.

Author contributions

S.S. performed in vivo polyketide production experiments, enzyme characterization experiments, physiological experiments and construction of DNAs with S.R. O.A. synthesized TKL standards and carried out DszAT library generation and screening with J.L.S. T.M.P. and S.S. jointly performed in vitro polyketide production experiments. H.D. assisted with construction of DNAs. E.E.K.B. and B.A. assisted with high-resolution fragmentation analysis of 6dEB and its analogs. M.C.Y.C. and C.K. administered the project. S.S. and M.C.Y.C. wrote the paper with contributions from all authors. All authors designed experiments and analyzed data.

Competing interests

The authors declare no competing interests.

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

Extended data are available for this paper at https://doi.org/10.1038/s41589-022-01070-y.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-022-01070-y.

Peer review information *Nature Chemical Biology* thanks Constance Bailey and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

We engineered a fluorine-selective *trans*-acyltransferase to produce site-selectively fluorinated erythromycin precursors in vitro. We further demonstrated that these analogs could be produced in vivo in *Escherichia coli* on engineering of the fluorinated extender unit pool. By using engineered microbes, elaborate fluorinated compounds can be produced by fermentation, offering the potential for expanding the identification and development of bioactive fluorinated small molecules.

The blending of biological and chemical structure offers enormous potential for accessing new areas of structural space for discovery of function[1–4]. While living organisms are particularly adept at constructing complex bioactive organic scaffolds, their use of the periodic table is limited when compared to human designs[1,5,6]. One such element is fluorine, which is found in only a handful of the $>10^5$ known natural products, yet is highly prevalent as a functional design element among synthetic compounds[7–9]. Fluorine plays a pivotal role in molecular design due to its many unique properties, affecting a broad range of applications in agrochemicals to liquid crystals[10,11]. Indeed, the number of approved drugs containing fluorine has increased over an order of magnitude to reach 20–30% today[7,9]. However, the distinctive properties of fluorine that contribute to widespread usage also create problems in its introduction into target structures[7]. As such, methods to site-selectively install fluorine into complex molecular architectures, especially under mild conditions remain of great interest[12–16].

One especially challenging goal is to introduce fluorine selectively into natural products made by microbes, plants and marine invertebrates[17]. Given their size and complexity, natural products and their variants are particularly difficult to synthesize using purely chemical methods, yet have evolved to contain structural features that can selectively target macromolecules ranging from proteins, nucleic acids and carbohydrates. As such, natural products also make up a large proportion (33.5%) of approved drugs[18]. Thus, combining natural and synthetic features in different ways could allow for increasing the accessible molecular space in which to search for new molecular functions. We seek to expand the synthetic capabilities of living organisms to produce new molecules of interest to human society that combine key functional elements of both natural and synthetic compounds (Fig. 1a). Using this approach, a multistep synthesis can potentially be telescoped to a single scalable growth process from simple universal building blocks such as glucose[19–23].

In particular, we aim to develop a platform to introduce fluorine into complex drug-like scaffolds to allow synthesis of new organofluorines that would otherwise be difficult to access through chemical synthesis. We have focused on modular polyketides as they display diverse molecular structures that can be built with relatively predictable biosynthetic logic[24,25]. Each module of a modular polyketide synthase (PKS) elongates the product backbone by two carbons by a concerted action of its three essential domains: ketosynthase (KS), acyltransferase (AT) and acyl carrier protein (ACP). The AT domain selects for and loads a carboxyacyl coenzyme A (CoA) extender unit to the phosphopantetheinyl modification of the ACP domain, whereas the KS domain then uses the extender unit to elongate the growing polyketide intermediate, before passing it to the following module. Given the hierarchical domain and module organization of the type I modular PKSs that

make these molecules, gene sequence and product structure are directly connected such that changes can be introduced site selectively into the molecule by targeting mutations to the corresponding domain.

Of the more than 10,000 known polyketide natural products, most use methylmalonyl-CoA (R = CH_3, **1**) or malonyl-CoA (R = H, **2**) as an extender unit. More unusual substituents on extender units are rare in native polyketide structures[24,26,27]. Our goal is to expand the diversity of polyketides using a general platform to site-selectively introduce fluorine into full-length targets. Our previous work showed that we could use enzymes and cells to synthesize the fluoromalonyl-CoA (R = F, **3**) extender unit and incorporate it site selectively into simple polyketide fragments[28]. In this system, fluoromalonyl-CoA could be targeted to a particular module for incorporation using a complementation strategy[29,30], where a separate *trans*-AT is used to load the fluorinated extender unit onto a PKS module where the *cis*-AT has been inactivated (Fig. 1b). Unlike *cis*-AT domains that are integral parts of multifunctional assembly-line PKS systems, *trans*-AT domains are free-standing enzymes that interact with ACP domains via intermolecular protein–protein interactions. This simple strategy allows selective incorporation of fluorine in any module containing an inactivated *cis*-AT domain with the evolution of a single fluorine-selective *trans*-AT.

Here, we apply the *trans*-AT strategy to incorporate fluorine substituents site selectively into complex polyketide natural products. By engineering fluorine selectivity into a *trans*-AT enzyme, regioselectively fluorinated polyketide analogs could be produced in vitro. Extending this system to a microbial host, we have further shown that a complete PKS system can be engineered to produce site-selectively fluorinated drug precursors in vitro and in an *Escherichia coli* production host.

Results

Engineering a fluoromalonyl-CoA trans-AT.

Due to the small size of the polyketide fragments previously made[28], we were able to use a wild-type (WT) *trans*-AT from the disorazole biosynthetic pathway (DszAT) that natively charges the PKS with malonyl-CoA extender units[31]. This method could be applied to simple model constructs of the 6-deoxyerythronolide B (6dEB, **8**) synthase (DEBS)[32]. However, to produce fluorinated variants of most complex polyketides, DszAT would need to be engineered to increase its selectivity for fluoromalonyl-CoA while greatly reducing activity on its native substrate malonyl-CoA, which is cellularly abundant[33]. With this goal in mind, we initiated structure-guided protein engineering studies of DszAT to increase its selectivity toward fluoromalonyl-CoA.

The crystal structure of DszAT shows a phenylalanine residue at position 190 near the active site[34] (Supplementary Fig. 1a). F190 is conserved across all malonyl-CoA selective AT domains from *cis*-AT PKS systems[35,36], suggesting that it could be important for selectivity. L87 was found in the second sphere of the active site and hypothesized to be involved in substrate positioning with respect to the oxyanion hole. If the oxyanion hole is weakened, we reasoned that the effect should be less deleterious with fluorinated substrates as they are relatively activated for acyl transfer based on inductive effects (Supplementary Fig. 1b).

Based on these observations, a DszAT saturation mutagenesis library was constructed at position F190 and combined with L87V or L87A mutations to reduce steric bulk. *E. coli* cells expressing the DszAT mutants were lysed and screened by triketide lactone (TKL) formation experiments[37] (Fig. 2a) by adding a purified DEBS module 3 construct with an inactivated AT domain and fused to thioesterase (TE) domain ($Mod3_{DEBS}$ + TE(AT^0)) (Extended Data Fig. 1a). In this assay, a PKS module is used with a diketide starter, *N*-acetylcysteamine thioester of (2*S*,3*R*)-2-methyl-3-hydroxypentanic acid (NDK-SNAC, **4**) to carry out a chain elongation reaction with a carboxyacyl-CoA extender unit containing an R = CH_3, H or F substituent to form a TKL molecule (Supplementary Figs. 2 and 3). Incorporation of methylmalonyl-CoA, malonyl-CoA and fluoromalonyl-CoA extender units in these experiments produces TKL (**5**), desmethyl TKL (H-TKL, **6**) and fluorinated desmethyl TKL (F-TKL, **7**), respectively. Select mutants that showed increased relative production of F-TKL were purified and investigated in vitro with methylmalonyl-CoA, malonyl-CoA and fluoromalonyl-CoA substrates (Extended Data Fig. 1b).

These preliminary screens indicated that F190V DszAT mutant was more selective toward fluoromalonyl-CoA compared to WT enzyme. Steady-state kinetic characterization of the purified mutant further shows that transacylation of an ACP partner is maintained with fluoromalonyl-CoA but substantially decreased with malonyl- and methylmalonyl-CoA, with approximately 6- and 2.5-fold decreases in observed k_{cat}/K values, respectively (Extended Data Fig. 2a). Unproductive hydrolysis of fluoromalonyl-CoA is also reduced in the mutant with around a 2.5-fold decrease in k_{cat}/K values (Extended Data Fig. 2b). Thus, F190V DszAT attains high fluoromalonyl-CoA selectivity by both suppressing malonyl-CoA turnover and fluoromalonyl-CoA hydrolysis.

With this mutant in hand, we then tested TKL formation in vitro with single or mixed carboxyacyl-CoA extender units (R = H, F, Me). When malonyl-CoA was provided as the sole extender unit, the production of H-TKL by DszAT F190V was found to be reduced by an order of magnitude compared to WT (Fig. 2b). Moreover, DszAT F190V was able to moderately increase F-TKL production compared to that produced by the WT enzyme when fluoromalonyl-CoA was provided. Next, the engineered enzyme was subjected to mixed extender unit conditions, which would be required to produce a complex polyketide product using malonyl-CoA, methylmalonyl-CoA or both (Fig. 2c). When presented with an equimolar mixture of all three extender units, WT DszAT produces 7-fold more H-TKL than F-TKL. In comparison, F-TKL production is maintained with DszAT F190V with a 20-fold reduction in H-TKL leading to F-TKL as the dominant product (79 ± 5%). Taken together, these results show that DszAT F190V demonstrates the desired selectivity needed for a fluorine-selective *trans*-AT reagent that can be used in an environment with a complex carboxyacyl-CoA extender unit pool.

Production of regioselectively fluorinated 6dEB analogs.

To test the ability of DszAT F190V for preparation of complex polyketides, we decided to focus on the DEBS system, which has been fully reconstituted in vitro[38]. The polyketide product of the DEBS pathway, 6dEB, contains 21 carbons and 10 stereocenters and is produced through a series of nearly 30 enzymatic steps from one propionyl-CoA starter unit

and six methylmalonyl-CoA extender units[32] (Fig. 3a). The DEBS PKS itself consists of three proteins (DEBS1–3), each about 300 kDa in monomeric mass and together comprise six modules and a loading didomain. Hence, the introduction of single point mutations into one of the DEBS proteins is complicated by their large size as well as homology with other modules encoded on the same protein. Thus, a strategy was developed to selectively inactivate an AT domain by first replacing a section with an antibiotic marker so that the mutated domain can cleanly be reintroduced. Full plasmid sequencing is also needed to ensure that sequences are not lost to recombination.

To produce 2-fluoro-2-desmethyl 6dEB (**9**), the AT domain of module 6 was inactivated (Mod6 AT^0 DEBS) and complemented with either WT or F190V DszAT. When fluoromalonyl-CoA and *trans*-AT were supplied to the system, a mass peak corresponding to a monofluorinated desmethyl 6dEB analog was observed (Fig. 3b, Extended Data Fig. 3a and Supplementary Fig. 4). This product was absent when there was no active site mutation in the AT domain of module 6 of the DEBS system. These results demonstrate the robustness of the complementation strategy to selectively incorporate fluorine using a multimodular PKS in the presence of a mixed extender unit pool.

To confirm regiospecific incorporation of fluoromalonyl-CoA, fragmentation analysis was used to study the substitution pattern of the observed fluorine-containing desmethyl 6dEB analog produced by both F190V and WT DszAT (Supplementary Fig. 5). The fragmentation patterns obtained were consistent with the fluoromalonyl-CoA extender unit being incorporated by module 6, suggesting that the observed fluorine-containing desmethyl 6dEB analog was the expected target, 2-fluoro-2-desmethyl 6dEB (ref. [39]) (Extended Data Fig. 3b and Supplementary Fig. 6). When comparing the system selectivity, the relative amount of 2-fluoro-2-desmethyl 6dEB analog to 6dEB was found higher in the system containing the engineered DszAT than in the system containing WT DszAT (Fig. 3c and Extended Data Fig. 3c).

Next, the AT domain of module 5 was inactivated (Mod5 AT^0 DEBS) with the intention of producing 4-fluoro-4-desmethyl 6dEB (**10**) to show that normal chain elongation could occur after the incorporation of a fluoromalonyl-CoA extender unit. Indeed, we were able to again detect the mass peak corresponding to a monofluorinated desmethyl 6dEB when Mod5 AT^0 DEBS was complemented with WT or F190V DszAT (Fig. 3b and Extended Data Fig. 4a). Similar to the production of 2-fluoro-2-desmethyl 6dEB, the production of this compound was dependent on fluoromalonyl-CoA, *trans*-AT and the *cis*-AT domain active site mutation. Notably, the retention time of this compound differed from the product observed with Mod6 AT^0 DEBS, suggesting that a different regioisomer was made. Fragmentation analysis is consistent with the monofluorinated desmethyl 6dEB analog arising from fluoromalonyl-CoA incorporation by module 5, producing 4-fluoro-4-desmethyl 6dEB (Extended Data Fig. 4b and Supplementary Fig. 7). Again, an increase in system selectivity toward the fluorine-containing analog was observed when F190V DszAT was used (Fig. 3c and Supplementary Fig. 4c).

Production of fluorinated 6-erythronolide B analogs in a microbial host.

With site-selective fluorine incorporation achieved in vitro, we sought to develop an in vivo cellular production system for fluorine-containing polyketide molecules. While in vitro experiments allowed us highly controlled environments for enzymatic production, they can be challenging to scale up for larger scale processes compared to fermentation in engineered microbes. However, the use of an unusual element such as fluorine yields new challenges in the precise control of the intracellular concentration and flux of system components as well as issues in toxicity and cross reactivity.

Previous work showed that feeding fluoromalonate (**11**) to the culture medium of *E. coli* cells expressing $DEBS_{Mod6}$ leads to detectable production of F-TKL[28]. We thus hypothesized that increasing the availability of fluoromalonyl-CoA in the cell, as well as using a complementation strategy, would allow us to increase the yield of F-TKL and produce fluorine-containing polyketides of increasing complexity in living cells. Toward the first goal of increasing the intracellular pool of fluoromalonyl-CoA, both a malonate transporter and a malonyl-CoA synthetase (*Rhodopseudomonas palustris* MatB) were included (Fig. 4a). We then constructed a four-gene, three-plasmid system for in vivo biosynthesis of F-TKL from fluoromalonate as the fluorine source, containing a malonate transporter, MatB, the *trans*-AT and either $Mod3_{DEBS}$ + TE(AT^0) or $Mod6_{DEBS}$ + TE(AT^0) proteins[40] (Fig. 4b).

Comparison of the three malonate transporters MdcF, MadLM and MatC showed that all notably increase the yield of F-TKL produced by both $Mod3_{DEBS}$ + TE(AT^0) and $Mod6_{DEBS}$ + TE(AT^0) (Fig. 4c). While MdcF only demonstrated a moderate effect, expression of either MadLM or MatB increased the F-TKL titer up to 60 μM with MadLM showing less dependence on growth phase[41]. To test that F-TKL formation was dependent on proper acylation of the ACP, the ACP domain was inactivated in the module 3 construct, $Mod3_{DEBS}$ + TE(AT^0 ACP^0). The mutation in the ACP led to a virtual abolishment of F-TKL production, suggesting that production proceeds through the expected covalently tethered intermediate[42] that is competent for multiple chain extension reactions (Fig. 4d). Production appears to be independent of a functional DszAT, suggesting that an endogenous host protein may also be competent at complementation (Fig. 4e). Since *E. coli* does contain a malonyl-CoA-ACP AT (FabD) in the fatty acid synthase pathway, this enzyme could serve as the endogenous *trans*-AT. Despite its low sequence similarity to DszAT and DEBS AT domains, FabD does contain similar structural features (Extended Data Fig. 5) and is competent to acylate the DEBS module 6 ACP domain as shown both by steady-state kinetic analysis and TKL formation (Extended Data Fig. 6). These results show that in vivo fluorinated polyketide yield and selectivity is likely governed by the high intracellular concentration of fluoromalonyl-CoA in the engineered host (Extended Data Fig. 7).

We then sought to produce regioselectively fluorinated 6dEB analogs in vivo with both Mod5 AT^0 and Mod6 AT^0 DEBS. Based on previous findings, we adapted our engineered system to incorporate *E. coli* FabD as the source of *trans*-AT activity. We sought to test the ability of native *E. coli* metabolism to incorporate fluoromalonyl-CoA into complex polyketide products using a high-yielding *E. coli* 6dEB production model system[43] (Fig.

5ab). In *E. coli* 6dEB production system, the host cells express DEBS1–3, along with propionyl-CoA synthetase (PrpE) and propionyl-CoA carboxylase subunits A and B (PccAB). When fed with propionate (**12**), both the propionyl-CoA (**13**) starter unit and methylmalonyl-CoA extender units used to produce 6dEB can be accumulated in vivo.

When provided with propionate precursor, 6dEB production systems expressing Mod5 AT^0 and Mod6 AT^0 DEBS predominantly produce a single species of desmethyl 6dEB with fragmentation patterns consistent with the regioselective production of 4-desmethyl 6dEB and 2-desmethyl 6dEB (ref. [39]), respectively (Extended Data Fig. 8). This result supports that the observed desmethyl 6dEB species were a result of the module with inactivated AT domain incorporating malonyl-CoA instead of methylmalonyl-CoA.

To produce fluorinated 6dEB analogs in vivo, the transport and activation system to accumulate intracellular fluoromalonyl-CoA was added to a *E. coli* 6dEB production system. Although growth inhibition was observed at 5 mM fluoromalonate feeding, it was indistinguishable from the effect seen with malonate feeding, suggesting that this behavior was not specific to fluorine (Extended Data Fig. 9). On expression of MatB and MadLM with feeding of propionate and fluoromalonate, we found that cells expressing either Mod5 AT^0 or Mod6 AT^0 DEBS each produced a distinct monofluorinated desmethyl 6dEB analog (Fig. 5cd). Fragmentation, nuclear magnetic resonance (NMR) analysis and comparison to products produced in vitro supported the assignment to the expected regioisomers for each construct (Mod5 AT^0 DEBS, 4-fluoro-4-desmethyl 6dEB; Mod6 AT^0 DEBS, 2-fluoro-2-desmethyl 6dEB; Extended Data Fig. 10). Taken together, these results show that regiospecific introduction of fluorine substitution can be achieved in complex and full-length polyketide natural products in engineered living cells.

Discussion

Fluorine is an unusual element that has been widely used to tune the behavior of small molecule synthetic drugs[7,8]. Because of its small size and high electronegativity, the shape of a molecule and its binding to a macromolecular target can be maintained while optimizing both molecular and pharmacokinetic properties. Like other features of synthetic drugs that are essential to their design, fluorine rarely occurs in the natural products made by living systems. Indeed, natural products and synthetic drugs are structurally distinct from each other with regards to characteristics such as stereocenter density, number of rotatable bonds and fraction of sp^3 carbons. One approach to accelerating the pace of drug discovery is to expand the structural space that can be searched and blend the structural complexity and restricted conformational flexibility of natural products with the functional group and elemental diversity of synthetic small molecules. As such, interest in new fluorination methods especially under mild reaction conditions is high. Continuing advances in fluorination methodology have enabled access to a growing range of fluorinated structures[12–16] and our study adds to the growing repertoire of methods to regioselectively incorporate fluorine into complex natural product scaffolds.

Toward this goal, we have developed a method to introduce fluorine site selectively into polyketide natural products that are immediate drug precursors through engineered

biosynthesis. We have engineered a fluoromalonyl-CoA transacylase that can selectively charge a PKS in the presence of the two major extender units, malonyl-CoA and methylmalonyl-CoA. The increased selectivity of the engineered DszAT toward fluoromalonyl-CoA is derived from a reduction in the catalytic efficiency of transacylation activity specifically with malonyl-CoA while maintaining activity with fluoromalonyl-CoA. We also observe an enhancement in yield that may be related to the suppression of fluoromalonyl-CoA hydrolysis. Using this variant, fluorine can be introduced regioselectively into the polyketide backbone to produce two fluorinated analogs of 6dEB, 2-fluoro-2-desmethyl 6dEB and 4-fluoro-4-desmethyl 6dEB, in vitro.

The ability to biosynthesize 4-fluoro-4-desmethyl 6dEB demonstrates that a fluorinated intermediate can be extended downstream by the PKS, suggesting that incorporation of fluorine throughout the polyketide chain is possible. However, the yield of this isomer appears to be lower than 2-fluoro-2-desmethyl 6dEB, where the fluorinated monomer is inserted in the last step, consistent with our previous observations that PKS gatekeeping seems to be greater when extending from a fluorinated monomer compared to chain extension with a fluorinated monomer[42]. Further work to explain these details could help to improve our understanding of the molecular mechanism of gatekeeping in multimodular PKSs for production of natural product analogs. While enzymatic processing of other 6dEB analogs has shown that downstream tailoring is relatively promiscuous, assessment of the impact of fluorine substitution on downstream steps to produce fluorinated analogs of erythromycin would also be valuable.

To produce fluorinated 6dEB analogs in cells, we engineered the intracellular fluoromalonyl-CoA level of the *E. coli* host through the introduction of a malonate transporter and malonyl-CoA synthetase. When fed with fluoromalonate, *E. coli* expressing a malonate transporter and a malonyl-CoA synthetase, along with a single-modular PKS construct, could produce mg l^{-1} levels of fluorinated triketide fragments. However, production yields were lower and more variable for fluorinated 6dEB analogs, possibly due to the low expression of DEBS1–3 compared to single-modular constructs as well as slowed growth under our production conditions. This growth inhibition appears to be related to a general physiological response to malonate as it is not specific to fluorine; however, it seems to be alleviated on expression of a higher flux pathway that consumes the carboxyacyl-CoA extenders, such as a single PKS module or fluorinated bioplastic production[41]. An unintended consequence of DEBS1–3 expression also appeared to be changes in the expression of other pathway enzymes, such as increased malonate transporter expression. In addition to protein expression burden, yield variability among different transformants may also be related to the shared origin of replication for pBP130 and pBP144. Although this was shown to be a stable system[43], it is possible that the copy number could vary between the two plasmids. As such, precise titration and balancing of the different enzymes used for production could assist in increasing pathway flux as well as cell growth, both of which would improve product yield and scalability.

Polyketides have been a valuable source of medicinal compounds and especially desirable targets for exploring new lead compounds. Given the modular nature of their biosynthesis, development of strategies for polyketide modification can be extended to the large number

of other family members. Applications of this platform open opportunities for discovery of new bioactive compounds by expanding the accessible structural space for organofluorines.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-022-01070-y.

Methods

Commercial materials.

Luria-Bertani (LB) medium Miller, LB Agar Miller, Terrific Broth, magnesium sulfate anhydrous and glycerol were purchased from EMD Biosciences. Carbenicillin, glucose, isopropyl-β-D-thiogalactopyranoside (IPTG), sodium chloride, calcium chloride, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, dichloromethane, ethyl acetate, ethylene diamine tetraacetic acid disodium dihydrate (EDTA), sodium phosphate monobasic, potassium phosphate monobasic, ammonium chloride and arabinose were purchased from Fisher Scientific. CoA trilithium sodium salt, malonyl-CoA, methylmalonyl-CoA, malonic acid, methylmalonic acid, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenol pyruvate (PEP), ATP sodium salt, adenylate kinase (myokinase), pyruvate kinase, lactate dehydrogenase, poly(ethyleneimine), β-mercaptoethanol (BME), thiamine pyrophosphate (TPP), α-ketoglutaric acid, sodium phosphate dibasic heptahydrate, cysteamine, 4-hydroxy-6-methyl-2-pyrone acetic anhydride, 1-ethyl-3-(-dimethylaminopropyl)carbodiimide hydrochloride (EDC), 4-dimethylaminopyridine (DMAP), reduced β-nicotinamide adenine dinucleotide 2′-phosphate (NADPH), acetonitrile, dimethyl sulfoxide (DMSO), biotin, thiamine hydrochloride, sodium propionate, ammonium formate, chloramphenicol, β-nicotinamide adenine dinucleotide (NAD^+), reduced β-nicotinamide adenine dinucleotide (NADH), β-nicotinamide adenine dinucleotide 2′-phosphate ($NADP^+$), α-ketoglutarate dehydrogenase and bovine serum albumin were purchased from Sigma-Aldrich. Diethylfluoromalonate was purchased from Sigma-Aldrich or from Matrix Scientific. Spectinomycin (Sp) was purchased from Chem-Impex International. Formic acid, dithiothreitol (DTT) and perchloric acid were purchased from Fluka. 2-methyl-3-oxopentanoic acid ethyl ester was purchased from Fragmenta. Mini-PROTEAN TGX 8–16% and Bio-Rad Protein Assay Dye Reagent concentrate were purchased from Bio-Rad Laboratories. Restriction enzymes, T4 DNA ligase and Phusion DNA polymerase were purchased from New England Biolabs. GoTaqGreen Master Mix was purchased from Promega. Deoxynucleotides (dNTPs) were purchased from Invitrogen. PageRuler Plus prestained protein ladder was purchased from Fermentas (Glen Burnie). Oligonucleotides were purchased from Integrated DNA Technologies, resuspended at a stock concentration of 100 μM in water. DNA purification kits and Ni-NTA agarose were purchased from Qiagen. PD-10 columns Sephadex G-25M was purchased from GE. Complete EDTA-free protease inhibitor was purchased from Roche Applied Science. The

Zymoclean Large Fragment DNA Recovery Kit was purchased from Zymoresearch. Amico Ultra 3,000, 10,000, 30,000 and 100,000 molecular weight cutoff (MWCO) centrifugal concentrators were purchased from EMD Millipore. Chloroform-*d* and D_2O were purchased from Cambridge Isotope Laboratories. ^{19}F-NMR spectra were collected at 25 °C on Bruker AV-600 at the College of Chemistry NMR Facility at the University of California, Berkeley. High-resolution mass spectral analyses were carried out on an Agilent 6530 Quadrupole-Time-of-Flight (QTOF) Accurate Mass spectrometer and an Agilent 6460 Triple Quadrupole (QQQ).

Bacterial strains.

E. coli DH10B-T1^R was used for DNA construction. *E. coli* BL21(DE3)-T1^R, BAP1 (ref. [43]) and BAP1-T1^R (ref. [44]) were used for heterologous protein expression. BAP1 and BAP1-T1^R were used for expression of DEBS modules and DEBS ACP domains as they require modification with phosphopantethiene. BL21(DE3)-T1^R was used for all other proteins.

Gene and plasmid construction.

Standard molecular biology techniques were used to carry out plasmid construction. All PCR amplifications were carried out with Phusion High Fidelity DNA polymerase or with Taq DNA polymerase using GoTaqGreen master mix. For amplification of GC-rich sequences, PCR reactions carried out with Phusion High Fidelity DNA polymerase were performed in the GC buffer supplemented with 10% (v/v) DMSO and 1 M betaine with primer annealing temperatures 4–8 °C below the T_m. DNA assembly was performed using the isothermal Gibson assembly protocol[45]. For analysis and isolation of large DNA fragments from agarose gel, 0.3–0.6% gel and Zymoclean Large Fragment DNA Recovery Kit (Zymoresearch) were used. All constructs were verified by sequencing (UC Berkeley DNA Sequencing Facility and MGH CCIB DNA Core, Center for Computational and Integrative Biology, Massachusetts General Hospital). pFW3 (ref. [40]), pBL12, pBL13, pBL36, pFW98 and pFW100 were gifted by the laboratory of C. Khosla (Stanford University). pBP130 and pBP144 (ref. [43]) were gifts from the laboratory of B. Pfeifer at the University at Buffalo.

DszAT mutant libraries were generated by site-specific mutagenesis at selected residues within the active site of the enzyme. The libraries constructed full saturation at the amino acids at the F190 position (according to DszD numbering), a library at the S86 catalytic residue (S86C, S86D, S86E, S86A) and a double mutant S86; H191 library for the same mutants (S86C; H191A, S86D; H191A and S86E; H191A) was also constructed. Finally, a double mutant library for the F190; L87 residues was generated for select F190 mutants from the single mutant screens a secondary screen for enhanced fluorine substrate selectivity (F190G; L87A, F190G; L87V, F190S; L87A, F190S; L87V, F190T; L87A, F190T; L87V, F190V; L87A, F190V; L87V, F190P; L87A, F190I; L87A and F190I; L87V). The mutant library was constructed by amplification from pFW3.

The F190 mutations were introduced by amplifying pFW3 with two sets of primers. One fragment was amplified with DszAT F190 F1 and various F190 R1 primers (depending on mutation). The second fragment was amplified using various F190 F2 primers (depending

on mutation) and DszAT F190 R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments were inserted into the XbaI-HindIII sites of pFW3 using the Gibson Protocol. The L87 mutations were introduced by amplifying pFW3 with two sets of primers. One fragment was amplified with DszAT L87 F1 and various L87 R1 primers (depending on mutation). The second fragment was amplified using various L87 F2 primers (depending on mutation) and DszAT L87 R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments were inserted into the XbaI-HindIII sites of pFW3 using Gibson assembly.

The H191A mutation was introduced by amplifying pFW3. One fragment was amplified with DszAT H191A F1/R1 and the second fragment was amplified with DszAT H191A F2/R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments were inserted into the NdeI-EcoRI sites of pFW3 using Gibson assembly.

The S86 mutations were introduced by amplifying pET21c-DszAT H191A-His6 or pFW3 with two sets of primers. One fragment was amplified with DszAT S86 F1 and various S86 R1 primers (depending on mutation). The second fragment was amplified using various S86 F2 primers (depending on mutation) and DszAT S86 R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments were inserted into the XbaI-HindIII sits of pFW3 or pET21c-DszAT-H191A-His6 using Gibson assembly.

All double mutants were designed as described above, with the exception of PCR template containing desired mutations from the initial round of screening. For pET16b-His_{10}Pres-$ACP_{DEBSMod6}$, the ACP domain of $Mod6_{DEBS}$ was amplified from pAYC-138 (ref. [40]) using primers PresACP6_Fwd3 and PresACP6_Rev. The PCR product was assembled into the NdeI-BamHI site of a modified pET16b with the Factor Xa cleavage site replaced by the PreScission cleavage site via Gibson assembly.

pFW3_F190V was constructed by amplification from pFW3 with two pairs of primers: DszAT F190 F1/DszAT F190V R1 and DszAT F190V F2/DszAT F190 R2. Primers DszAT F190V R1 and DszAT F190V F2 contain the F190V mutation to be introduced to the plasmid. The two PCR products were inserted into the XbaI-HindIII site of pFW3 via Gibson assembly.

pFW100_DEBS3(Mod5 AT^0) encodes for DEBS3 with S642A mutation. The plasmid was cloned by first replacing the section to be mutated from the parent plasmid with a Cm^R marker and then removing it on insertion of the piece bearing the desired mutation. This strategy allows for identification of mutants using antibiotic selection. Briefly, the Cm^R cassette was first amplified from pACYC184 with the primers pACYC184_CmOperon_pFW100_M5_Fwd_PacI/ pACYC184_CmOperon_pFW100_M5_Rev_SpeI and inserted into the SfiI-BsiWI site of pFW100 via Gibson assembly. *E. coli* cells transformed with the assembly mixture were selected on culture medium containing carbenicillin and chloramphenicol. Plasmids containing both resistance markers were further screened by restriction analysis with SacI. The pFW100 plasmid intermediates with the correct restriction pattern were confirmed by diagnostic sequencing. The target segments were amplified from pFW100 by two primer

pairs to introduce the AT^0 mutation into module 5 of DEBS: pFW100_M5AT0_SfiI_F/pFW100_M5AT0_SfiI_R and pFW100_M5AT0_BsiWI_F/pFW100_M5AT0_BsiWI_R. The PCR products were then inserted via Gibson assembly into the pFW100 intermediate, which was digested with PacI and SpeI to remove the Cm^R cassette. Resulting plasmids were selected for the loss of the Cm^R marker and screened by SacI restriction analysis. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing at MGH CCIB DNA Core, Center for Computational and Integrative Biology, Massachusetts General Hospital.

pFW100_DEBS3(Mod6 AT^0) encodes for DEBS3 with S2107A mutation. The plasmid was constructed using a similar strategy. The Cm^R cassette was amplified from pACYC184 with primers pACYC184_CmOperon_pBP130_M6_Fwd_PacI/pACYC184_CmOperon_pBP130_M6_Rev_SpeI and inserted into the BbvCI-AjuI site of pFW100 via Gibson assembly. *E. coli* transformed with the assembly mixture were selected on culture medium containing carbenicllin and chloramphenicol. Plasmids containing both resistant genes were further screened by restriction analysis with SacI. The pFW100 plasmid intermediates with the correct restriction pattern were confirmed by diagnostic sequencing. Segments of module 6 of DEBS were amplified by two primer pairs to introduce the AT^0 mutation into $DEBS_{Mod6}$. The M6TE-SA-M6-RP–M6TE-SA-M6-FP primer pair was used to amplify a segment of module 6 containing the inactivating mutation to the AT domain from pAYC-138, whereas the M6TE-SA-130-FP–M6TE-SA-130-RP primer pair was used to amplify another contiguous segment from pBP130. These PCR products were inserted via Gibson assembly into the pFW100 intermediate, which was digested with PacI and SpeI to remove the Cm^R cassette. Resulting plasmids were selected for the loss of the Cm^R marker and screened by SacI restriction analysis. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing at MGH CCIB DNA Core, Center for Computational and Integrative Biology, Massachusetts General Hospital.

For p15A-DszAT, the gene encoding DszAT was amplified from pFW3 with the DszATF1–T7TerminatorR1 primer pair. The lac operon was amplified from the same template with the LacCasetteLacI–LacCasetteLacO primer pair.

The PCR products were inserted by Gibson assembly into the ClaI-PacI site of pJA4MCS2, which is pACYC184 (ref. [40]) modified by replacing TcR cassette with a HindIII-XhoI-KpnI-XmaI-PacI-AflII-SFcI-PstI-AscI-AvrII-HincII linker.

For p15A-DszATS86A, the S86A mutation was introduced to p15A-DszAT by amplifying p15A-DszAT with two sets of primers DszATF1–S86AMutationR and S86AMutationF–DszATMutant_CTerm. S86AMutationR and S86AMutationF contain the S86A mutation. The two PCR fragments were inserted by Gibson assembly into the NdeI-HindIII site of p15A-DszAT.

For p15A-DszAT F190V, the F190V mutation was introduced to p15A-DszAT by amplifying p15A-DszAT with two sets of primers: DszATF1–F190VMutationR and F190VMutationF–DszATMutant_CTerm. F190VMutationR and F190VMutationF contain

the S86A mutation. The two PCR fragments were inserted by Gibson assembly into the NdeI-HindIII site of p15A-DszAT.

For pET21c-FabD, the *FabD* gene was amplified from *E. coli* BL21(DE)-T1R using the MalACP_F1–MalACP_R1 primer set. The PCR product was inserted into the NdeI-EcoRI site of pET21c by Gibson assembly.

pBP130_DEBS3(Mod5 AT^0) encodes for DEBS2 and DEBS3 with S642A mutation. The segment of DEBS module 5 containing the desired S642A mutation was amplified with primers M5_BbvCI/M5_NsiI from pFW100_DEBS3(Mod5 AT^0) and gel purified (3.5 kB). pBP130 was digested with BbvCI and NsiI and the resulting 22 kB band was gel purified to provide the backbone. The backbone and insert were then ligated with T4 ligase. The plasmid candidates were screened by restriction analysis with NotI. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing to ensure that no other changes had been introduced.

pBP130_DEBS3(Mod6 AT^0) encodes for DEBS2 and DEBS3 with S642A mutation. pFW100_DEBS3(Mod6 AT^0) was digested with AsiSI and BbvCI and the resulting 6 kB fragment containing the desired S2107A mutation was gel purified. pBP130 was digested with AsiSI and BbvCI and the resulting 20 kB band was gel purified to provide the backbone. The backbone and insert were then ligated with T4 ligase. The plasmid candidates were screened by restriction analysis with XhoI. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing to ensure that no other changes had been introduced.

Expression of His-tagged proteins.

For His_6-MatB, DszAT-His_6, DszAT F190V-His_6, His_{10}-$ACP_{DEBSMod6}$, $Mod3_{DEBS}$ + TE(AT^0)-His_6, $Mod6_{DEBS}$ + TE(AT^0)-His_6, plasmids encoding the proteins of interest were transformed into *E. coli* BL21(DE3)-$T1^R$ (His_6-MatB, DszAT-His_6, DszAT F190V-His_6) or *E. coli* BAP1-$T1^R$ for proteins that require phosphopantetheine modification of ACP domains (His_{10}-$ACP_{DEBSMod6}$, $Mod3_{DEBS}$ + TE(AT^0)-His_6, $Mod6_{DEBS}$ + TE(AT^0)-His_6). Terrific Broth culture medium (1 l) with appropriate antibiotic (carbenicillin, kanamycin, chloramphenicol 50 μg ml^{-1}; spectinomycin 100 μg ml^{-1}) in a 2.5 l ultra-yield flask was inoculated with overnight culture of freshly transformed *E. coli* cells. Cells were grown at 37 °C with shaking at 200 r.p.m. to an optical density (OD_{600}) of 0.6–0.8, at which time they were cold-shocked on ice for 20–40 min. Expression was induced by addition of IPTG (His_6-MatB, His_{10}-$ACP_{DEBSMod6}$, DszAT-His_6, DszAT F190V-His_6 1 mM; $Mod3_{DEBS}$ + TE(AT^0)-His_6, $Mod6_{DEBS}$ + TE(AT^0)-His_6 0.2 mM). Cells were then grown at 16 °C with shaking at 200 r.p.m. overnight and gathered by centrifugation at 8,000*g* for 5 min at 4 °C. Cell pellets were flash frozen in liquid N_2 and stored at −80 °C until purification.

For His_6-PrpE, His_6-Epi, His_6-LDD_{DEBS}, $Mod1_{DEBS}$-His_6, $Mod2_{DEBS}$-His_6, DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6, DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6 and DEBS3(Mod6 AT^0)-His_6, expression plasmids were introduced into *E. coli* BL21(DE3) (His_6-PrpE and His_6-Epi) or *E. coli* BAP1 cells to allow phosphopantetheinyl modification of ACP domains (His_6-LDD_{DEBS}, $Mod1_{DEBS}$-His_6, $Mod2_{DEBS}$-His_6, DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6,

DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6 and DEBS3(Mod6 AT^0)-His_6). Overnight seed cultures were used to inoculate 0.5 l of LB medium with 2% glucose containing the appropriate antibiotic (carbenicillin, kanamycin: 50 μg ml^{-1}) in 2.5 l ultra-yield shake flask. Cells were grown at 37°C with shaking at 200 r.p.m. to an approximate OD_{600} of 0.3, at which point the temperature was slowly lowered. When cells reached OD_{600} of 0.6, at which point cells were at approximately 20 °C, protein expression was induced with IPTG (His_6-PrpE, His_6-Epi, His_6-LDD_{DEBS}, $Mod1_{DEBS}$-His_6, $Mod2_{DEBS}$-His_6, DEBS2: 0.1 mM; DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6, DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6 and DEBS3(Mod6 AT^0)-His_6: 0.5 mM). Cells were then grown at 18 °C with shaking at 200 r.p.m. overnight and gathered by centrifugation at 8,000*g* for 5 min at 4 °C. Cell pellets were flash frozen in liquid N_2 and stored at −80 °C until purification.

Purification of His-tagged proteins.

Cell pellets were resuspended in Lysis Buffer A (200 mM sodium phosphate, 200 mM sodium chloride, 30% (v/v) glycerol, 2.5 mM DTT, pH 7.5) at 5 ml g^{-1} of cell pellet. Lysozyme (0.5 mg ml^{-1}) and complete EDTA-free protease inhibitor cocktail (Roche, 1 eq) were then added before homogenization and sonication. Cleared cell lysates were then obtained after centrifugation at 8,000*g* for 20 min at 4 °C. To precipitate DNA, poly(ethyleneimine) was added dropwise to the soluble fraction at a final concentration of 0.015% (*w/v*). Precipitated DNA was removed by centrifugation at 8,000*g* for 20 min at 4 °C. Protein concentration was determined using A_{280} and protein extinction coefficients calculated from the sequence using the ExPASY ProtParam program.

For His_6-MatB, DszAT-His_6, DszAT F190V-His_6 and FadD-His_6 (ref. [28]), the lysate was diluted threefold with Wash Buffer B (50 mM sodium phosphate, 300 mM sodium chloride, 20% (v/v) glycerol, 20 mM BME, pH 7.5) with 10 mM imidazole and incubated with Ni-NTA agarose resin (2–4 ml) for 45–60 min at 4 °C to bind the His-tagged protein before loading onto the column by gravity flow. The column was washed with Wash Buffer B containing 10 mM imidazole until the eluate was negative for protein content when tested by Bradford protein assay reagent (Bio-Rad). Column was further washed by Wash Buffer B containing 25 mM imidazole until eluate was negative for protein content again. Protein was eluted from the column with Elution Buffer C (50 mM sodium phosphate, 50 mM sodium chloride, 20 mM BME, 20% (v/v) glycerol) containing 250 mM imidazole. Eluted protein was concentrated using a 10 kDa MWCO Amicon Ultra spin concentrator (Millipore) before exchanging into Storage Buffer D (50 mM HEPES, 2.5 mM EDTA, 100 mM sodium chloride, 2.5 mM DTT, 20% (v/v) glycerol) using a PD-10 desalting column containing Sephadex G-25 resin (GE Healthcare). Purified protein was flash frozen in liquid N_2 before storing at −80 °C.

For His_6-$ACP_{DEBSMod6}$, the protein was purified as described above up to protein elution. After elution from the Ni-NTA column, the concentration of the protein solution was determined using Bradford protein assay reagent (Bio-Rad) against BSA as a standard. PreScission protease was then added to the protein solution at a ratio of 1 mg PreScission:25 mg protein substrate. The proteolysis reaction was dialyzed 2× 1:200 overnight at 4 °C into Storage Buffer D without DTT. The mixture was then diluted threefold with Wash

Buffer B without BME containing 10 mM imidazole. The solution was passed three times through Ni-NTA resin (2 ml) to remove PreScission protease and the cleaved His_6 tag before concentrating the eluate using a 3 kDa MWCO Amicon Ultra spin concentrator (Millipore). The purified protein was exchanged into Storage Buffer D without DTT using PD-10 desalting column with Sephadex G-25 resin (GE Healthcare). Thiol-containing reducing agents such as BME and DTT were excluded from buffers starting at the dialysis step as they interfere with transacylation assay. Purified protein was flash frozen in liquid N_2 before storing at −80 °C. Presence of the phosphopantethiene modification was checked by liquid chromatography–TOF (LC–TOF) using positive ionization mode on an Agilent 6224 TOF MS. Samples were buffer exchanged into 10 mM sodium phosphate by 0.5 ml of 7K Zeba Spin desalting column and analyzed on Proswift RP-4H high-performance LC (HPLC) column (1 × 50 mm, room temperature) using a linear gradient from 10 to 100% acetonitrile over 3.5 min with 0.1% (v/v) formic acid as the aqueous mobile phase at flow rate of 0.3 ml min^{-1}. Resulting mass chromatograms were deconvoluted with Agilent MassHunter BioConfirm software.

For $Mod3_{DEBS}$ + TE(AT^0)-His_6, $Mod6_{DEBS}$ + TE(AT^0)-His_6 (ref. [28]), the cleared lysate was incubated with Ni-NTA resin (2–4 ml) for 45–60 min at 4 °C to bind His-tagged protein before loading onto the column by gravity flow. The column was washed with Wash Buffer B containing 10 mM imidazole until eluate was negative for protein content when tested by Bradford protein assay reagent (Bio-Rad). Protein was then eluted from the column with Elution Buffer C containing 100 mM imidazole until eluate was negative for protein content. Eluted protein was concentrated using a 30 kDa MWCO Amicon Ultra spin concentrator and diluted threefold in Anion Exchange Buffer E (50 mM HEPES, 2.5 mM EDTA, 2.5 mM DTT, 20% (v/v) glycerol, pH 7.5) without sodium chloride. The diluted protein solution was loaded onto a 5 ml Hi-Trap Q HP column (GE Healthcare) using an NGC Medium-Pressure Liquid Chromatography System (Bio-Rad). The protein was eluted with a linear gradient from 0 to 1 M sodium chloride in Anion Exchange Buffer E over 30 column volumes (150 ml). Fractions containing the target protein eluted roughly 350 mM sodium chloride were pooled and concentrated with a 30 kD MWCO Amicon Ultra spin concentrator. Purified protein was flash frozen in liquid N_2 before storing at −80 °C.

For DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6, DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6 and DEBS3(Mod6 AT^0)-His_6 (ref. [38]), cell pellets were resuspended in Lysis Buffer F (50 mM sodium phosphate, 10 mM imidazole, 450 mM NaCl, 20% (v/v) glycerol, pH 7.6) at 5 ml g^{-1} of cell pellet. Lysozyme (0.5 mg ml^{-1}) and cOmplete EDTA-free protease inhibitor cocktail (Roche, 1 eq) were then added and incubated at 30 °C for 30 min before sonication. Cleared cell lysates were obtained after centrifugation at 8,000*g* for 20 min at 4 °C and loaded onto column containing Ni-NTA (10 ml per 3 l of cell culture) by gravity flow. The column was washed with Lysis Buffer F until eluate was negative for protein content when tested by Bradford protein assay reagent (Bio-Rad). The column was further washed with Wash Buffer G (50 mM sodium phosphate, 25 mM imidazole, 300 mM sodium chloride, 10% (v/v) glycerol, pH 7.5) before protein was eluted with Elution Buffer H (75 mM sodium phosphate, 500 mM imidazole, 20 mM sodium chloride, 10% (v/v) glycerol, pH 7.5) until eluate was negative for protein content. Eluted protein was loaded onto a 5 ml Hi-Trap Q HP column connected to an AKTA FPLC system. A gradient of 0 to 1 M sodium

chloride was applied over 30 column volumes (150 ml) of Anion Exchange Buffer I (50 mM sodium phosphate, 10% (v/v) glycerol) at flow rate of 5 ml min^{-1}. Fractions (2.5 ml) were collected and analyzed by SDS–PAGE. Fractions containing the target proteins (DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6 eluting at roughly 330–430 mM sodium chloride; DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6 and DEBS3(Mod6 AT^0)-His_6 eluting at roughly 370–500 mM sodium chloride) were pooled and concentrated using a 100 kDa MWCO Amicon Ultra spin concentrator. Purified protein was flash frozen in liquid N_2 before storing at −80 °C.

For His_6-LDD_{DEBS}, $Mod1_{DEBS}$-His_6, $Mod2_{DEBS}$-His_6, His_6-PrpE and His_6-Epi (ref. [38]), cell pellets were resuspended in 35 ml of Lysis Buffer J (50 mM sodium phosphate, 10 mM imidazole, 450 mM NaCl, 10% (v/v) glycerol, pH 7.6) containing one EDTA-Free Protease Inhibitor tablet (Pierce) per liter of cell culture. The cells were lysed by sonication and centrifuged at 20,000*g* for 60 min at 4 °C. The cleared cell lysates were added to Ni-NTA agarose resin (1 ml l^{-1} cell culture) and incubated at 4 °C for 1 h to bind His-tagged protein. The resin was loaded into a Kimble-Kontes Flex column and washed with 20 column volumes of Lysis Buffer J and 10 column volumes of Wash Buffer G. Proteins were eluted with 8 column volumes of Elution Buffer H. The eluate was further purified by anion exchange on a Hi-Trap Q column using a gradient from 0 to 500 mM sodium chloride in Anion Exchange Buffer I on an AKTA FPLC system. Fractions were collected and analyzed by SDS–PAGE. Fractions containing the target protein were pooled and concentrated using a 100 kD MWCO Amicon Ultra spin concentrator. Purified protein was flash frozen in liquid N_2 before storing at −80 °C.

Synthesis of *N*-acetylcysteamine thioester of NDK-SNAC.

The preparation of NDK-SNAC was adapted from literature protocol[46,47]. Ethyl-2-methyl-3-oxopentanoate (1.5 mmol) was dissolved in 10 ml of water and saponified in aqueous sodium hydroxide (10 M, 180 μl; 1.8 mmol). The reaction was stirred overnight at room temperature before pH was adjusted to <1 and reaction was extracted with 2× 35 ml of ethyl acetate. The organic layers were combined, dried over anhydrous magnesium sulfate and evaporated to dryness by rotary evaporation.

The resulting 2-methyl-3-oxopentanoic acid (0.85 mmol) was dissolved in 10 ml of dichloromethane on ice. After 3 min, *N*-acetylcysteamine (86 μl, 0.8 mmol, 0.95 eq) and DMAP (eight pellets) were added to the reaction. After 5 min, ethyl diisopropyl carbodiimide (244 mg, 1.275 mmol; 1.5 eq) was added. The reaction was removed from ice and stirred at room temperature overnight. The reaction was extracted with 2 × 20 ml of saturated aqueous ammonium chloride and 2 × 20 ml and saturated aqueous sodium chloride. The aqueous layers were combined and back extracted with 3 × 20 ml of dichloromethane. All organic fractions were combined, dried over anhydrous magnesium sulfate and evaporated to dryness by rotary evaporation. The resulting 2-methyl-3-oxopentanoyl-SNAC product was confirmed by ^{1}H-NMR and stored in solid form at −80 °C. ^{1}H-NMR (400 MHz, $CDCl_3$ = 7.26 ppm) δ 5.79 (s, N-H), 3.77–3.82 (m, 1H), 3.42–3.49 (m, 2H), 3.03–3.12 (m, 2H), 2.50–2.62 (m, 2H), 1.97 (s, 3H), 1.20 (d, J = 6.8 Hz, 3H), 1.07 (t, J = 7.2 Hz, 3H).

2-Methyl-3-oxopentanoyl-SNAC was enzymatically reduced with the ketoreductase domain from the first module of the pikromycin synthase ($KR_{PIKSMod1}$) and NADPH. This reaction (3 ml) contained 50 mM 2-methyl-3-oxopentanoyl-SNAC (1 M stock in DMSO), 300 mM D-glucose, 100 μM $NADP^+$, 20 U ml^{-1} glucose-1-dehydrogenase and 15 μM $Piks_{Mod1}KR$ in 300 mM HEPES, 100 mM sodium chloride, 10% (v/v) glycerol, pH 7.5. After stirring at room temperature overnight, thin-layer chromatography analysis of crude reaction mixture in 100% ethyl acetate solvent system revealed a product spot at R_f of roughly 0.3. Saturated aqueous sodium chloride (3 ml) was then added to the reaction mixture, which was extracted with 3 × 9 ml of ethyl acetate. The organic layer was dried over anhydrous magnesium sulfate and evaporated to dryness by rotary evaporation. Product was confirmed by ^{1}H-NMR and stored as 1 M solution in 25% (v/v) DMSO at −20 °C. ^{1}H-NMR (400MHz, $CDCl_3$ = 7.26 ppm) δ 3.81–3.84 (m, 1H), 3.43–3.47 (m, 2H), 0.99–3.05 (m, 2H), 2.69–2.75 (m, 1H), 2.62 (s, O-H), 1.97 (s, 3H), 1.43–1.57 (m, 2H), 1.23 (d, J = 7.2 Hz, 3H,), 0.98 (t, J = 7.2 Hz, 3H).

Synthesis of fluoromalonate.

Methanolic sodium hydroxide (2 M, 10.5 ml; 21 mmol) was added to 9:1 (v/v) dichloromethane:methanol (96 ml). Diethylfluoromalonate (1.5 ml, 9.6 mmol) was added and saponified with stirring at room temperature for 1–3 h. Disodium fluoromalonate was isolated by filtration through a Büchner funnel fitted with a filter paper and left to air dry.

Synthesis of fluoromalonyl-CoA.

Fluoromalonyl-CoA was prepared enzymatically from fluoromalonate and CoA using the malonyl-CoA synthetase from *Streptomyces coelicolor* (MatB)[28]. The reaction mixture (5 ml) contained 20 mM fluoromalonate, 4 mM CoA, 5 mM ATP, 20 mM phosphoenol pyruvate, 5 mM TCEP, 10 mM magnesium chloride, 144 U ml^{-1} pyruvate kinase, 80 U ml^{-1} adenylate kinase and 20 μM MatB in 200 mM sodium phosphate, pH 7.5. The reaction was incubated at 37 °C for 1 h. Enzymes were removed by filtration through Nanosep with 10 K Omega membrane spin column before reaction was lyophilized overnight. The residue from lyophilization was resuspended in water (2 ml) and stored frozen at −80 °C until purification. The resuspended crude reaction residue (0.5 ml per injection) was purified on an Eclipse XDB-C18 column (5 μm, 9.4 × 250 mm, room temperature; Agilent) using a linear gradient from 0 to 10% acetonitrile over 30 min with 10 mM ammonium formate as the aqueous mobile phase at flow rate of 4 ml min^{-1}. Fractions (2 ml) were screened using an Agilent 1290 UPLC (ultrahigh-performance liquid chromatograph) connected to a 6460 QQQ mass spectrometer using multiple reaction monitoring (MRM) in positive ionization mode (transition, collision energy, fragmentation voltage: fluoromalonyl-coA (872 → 365, 135, 35)). Fractions were analyzed on a Poroshell HPH-C18 column (2.7 μm, 2.1 × 100 mm, room temperature; Agilent) with a linear gradient from 0 to 20% acetonitrile over 6 min after an initial hold at 0% acetonitrile for 30 s with 10 mM ammonium formate as the aqueous mobile phase at flow rate of 0.6 ml min^{-1}. Fractions containing pure fluoromalonyl-CoA were pooled and lyophilized overnight. The residue was re-dissolved in water and stored at −80 °C. The concentration of fluoromalonyl-CoA was estimated using A_{260} (ε_{260} = 15,400 M^{-1} cm^{-1}).

Lysate assay for screening DszAT mutant library.

DszAT lysate is prepared by inoculating 50 ml of Terrific Broth (containing 50 μg ml^{-1} Carbenicillin) with freshly transformed *E. coli* harboring pFW3 and its derivatives (OD_{600} = 0.05). The cultures were grown to OD_{600} of 0.6–0.8 at 37 °C, cold-shocked on ice for 10–20 min, subsequently induced with 400 μM IPTG and grown overnight at 16 °C. Then 10 ml aliquots of each culture were pelleted by centrifugation at 12,000*g* for 5 min at 4 °C. The cell pellets were then resuspended in 290mM sodium phosphate pH 7.5 at 1 mg ml^{-1} concentration. The cells were lysed with 0.1 mm bacterial glass disruption beads using Mini-BeadBeater-24 (Biospec products) in 2 × 45 s exposures at maximum speed. The resulting homogenate was spun down at 20,000*g* at 4 °C for 20 min and the supernatant was used as the source of DszAT variants in the assay. Lysate from a strain containing empty pET51b+ vector (Invitrogen) was analyzed as negative control for background endogenous transacylation by *E. coli* BL21(DE3)-$T1^R$.

The assay mixture contained equal volume of cell lysate and regeneration system (100 μl of total). Final assay mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenol pyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (15 U ml^{-1}), adenylate kinase (10 U ml^{-1}), methylmalonyl-CoA epimerase (5 μM), CoA (1 mM), MatB (20 μM), methyl- or fluoromalonate (5 mM). The mixture was incubated at 37 °C for 30–45 min and initiated by the addition of NDK-SNAC (5 mM) and DEBS protein (10 μM). The reaction was incubated for 24 h at 37 °C. 50 μl aliquots were removed, quenched with 2.5 μl of 70% (v/v) perchloric acid and centrifuged at 18,000*g* for 10 min to pellet the precipitated protein. The supernatant was removed and flash frozen.

Before analysis, frozen samples were centrifuged at 18,000*g* for 5 min at to remove salts. The supernatant was removed and analyzed by LC–QQQ using MRM in negative ionization mode on an EclipsePlus C18 RRHD column (1.8 μm, 2.1 × 50 mm, room temperature, Agilent) or a Poroshell 120 SB-Aq column (2.7 μm, 2.1 × 50 mm, room temperature, Agilent) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s at 0.6 ml min^{-1}. The transitions are as follows (parent *m/z* → product *m/z*, fragmentation voltage, collision energy): H-TKL: 155 → 97, 100, 5; F-TKL: 173 → 59, 135, 20; H-TTKL: 197 → 95, 135, 5; TKL: 169 → 111, 135, 10. Production of TKLs and tetraketide pyrones (TTKPs) were normalized to the corresponding amounts observed with WT DszAT samples analyzed in conjunction.

In vitro assay for screening selected DszAT mutants.

All assay mixtures (30 μl) contained 5 μM methylmalonyl-CoA epimerase (Epi), 20 μM MatB, 2.5 mM ATP, 15 U ml^{-1} PK, 10 U ml^{-1} adenylate kinase and 50 mM PEP in 400 mM sodium phosphate, 10 mM magnesium chloride and 5 mM TCEP, pH 7.5. Reactions contained 1 mM carboxyacyl-CoA (fluoromalonyl-CoA, malonyl-CoA or methylmalonyl-CoA), 2.5 mM NDK-SNAC, 10 μM $Mod3_{DEBS}$ + TE(AT^0) and 10 μM DszAT variants. Epi, MatB, PK, myokinase, PEP, ATP and carboxyacyl-CoA substrate were preincubated at 37 °C for 30 min. Reactions were then initiated with NDK-SNAC, $Mod3_{DEBS}$ + TE(AT^0) and

trans-AT and incubated for 16 h at 37 °C. The 50 μl aliquots were removed, quenched with 2.5 μl of 70% (v/v) perchloric acid and centrifuged at 18,000*g* for 10 min. The supernatant was removed and stored at −80 °C until analysis. Reactions were analyzed as in lysate assay.

Steady-state kinetic characterization of hydrolysis activity.

The DszAT assay was adapted from a CoA release assay previously described in the literature[48–50]. Assays were performed in 96-well microplate (black, polystyrene, μ-clear, f-bottom, chimney well, med binding; Greiner Bio-One) by monitoring free CoA with 5,5′-dithio-bis-(2-nitrobenzoic acid) (DTNB) at 412 nm using a Synergy Mx microplate reader (Biotek). Reactions were monitored for 4 min using the minimum interval setting between measurements.

Assay components were prepared in three solutions in 50 mM sodium phosphate, 1 mM EDTA, 10% (v/v) glycerol, pH 7.5: solution 1, 2.4 mM DTNB; solution 2, 4× fluoromalonyl-CoA and solution 3, 20 μM *trans*-AT enzyme and 0.1 mg ml^{-1} BSA. Solution 1 (25 μl) and solution 2 (25 μl) were added to the wells and preincubated at room temperature for 15–20 min. The reactions were then initiated with solution 3 (50 μl). The final concentrations of fluoromalonyl-CoA were 200, 150, 100, 50, 25, 12.5, 6.25, 3.13, 1.56, 0.78, 0.39 and 0.20 μM. Absorbance values were converted to concentrations using the $\varepsilon_{412\ nm}$ = 14,150 M^{-1} cm^{-1} and a path length of 0.3 cm. Initial velocities (v_0) were defined as the linear portion of the change of product concentration over time curve. The data were fit by nonlinear curve fitting to the standard Michaelis–Menten equation.

Steady-state kinetic characterization of transacylation activity.

The transacylase assay was adapted from a previously described method[29,51,52]. Assays were performed in 96-well microplate (half area, μCLEAR, black polystyrene, medium binding, nonsterile, Greiner Bio-One) with NADH fluorescence (λ_{Ex} = 340 nm, λ_{Em} = 450 nm) monitored using a Synergy Mx microplate reader (Biotek; top optical position at 8 mm read height; gain, 100). Reactions were monitored for 5 min at 30 °C using minimum interval setting between measurements.

Assay components were prepared in three solutions in 50 mM sodium phosphate, 1 mM EDTA, 10% (v/v) glycerol, pH 7.5: solution 1, 2.4 mU μl^{-1} α-ketoglutarate dehydrogenase, 1.6 mM NAD^+, 1.6 mM TPP and 8 mM α-ketoglutarate; solution 2, 4× carboxyacyl-CoA (malonyl-CoA, methylmalonyl-CoA, or fluoromalonyl-CoA) and solution 3, 150 μM $ACP_{DEBSMod6}$, 2× *trans*-AT and 0.1 mg ml^{-1} BSA. Solution 1 (12.5 μl) and solution 2 (12.5 μl) were added to the wells and preincubated at room temperature for 15–20 min. The reactions were then initiated with solution 3 (25 μl).

The final concentrations of malonyl-CoA, methylmalonyl-CoA and fluoromalonyl-CoA were 200, 150, 100, 50, 25, 12.5, 6.25 and 3.13 μM. The final concentrations of DszAT (WT and F190V) were 50 nM for malonyl-CoA and fluoromalonyl-CoA and 10 μM for methylmalonyl-CoA. The final concentrations of FabD were 30 nM for malonyl-CoA, 50 nM for fluoromalonyl-CoA and 100 nM for methylmalonyl-CoA reactions. Controls with no *trans*-AT added were run in parallel for all reactions. Fluorescence was converted to concentration using an NADH standard curve collected externally in the same assay matrix

with carboxyacyl-CoA substrate, ACP domain and DszAT omitted. Initial velocities (v_0) were defined as the linear portion of the change of product concentration over time curve. Transacylation with methylmalonyl-CoA was fit by nonlinear curve fitting to the standard Michaelis–Menten equation. Transacylation of malonyl- and fluoromalonyl-CoA were fit by nonlinear curve fitting to the Hill equation with substrate inhibition based on the exhibited sigmoidal behavior of the kinetic curves at low substrate concentrations. This behavior may have to do with the formation of a ternary complex between the enzyme, carboxyacyl-CoA substrate and ACP cosubstrate, rather than the classical cooperative behavior of multimeric enzymes. The ternary complex for DszAT was previously hypothesized based on the marked increase in the k_{cat}/K_M parameter when ACP is present (transacylation) compared to when it is not (hydrolysis)[29].

Assay for in vitro TKL production.

All assays contained 5 μM methylmalonyl- CoA epimerase (Epi), 20 μM MatB, 2.5 mM ATP, 15 U ml^{-1} PK, 10 U ml^{-1} myokinase and 50 mM phosphoenol pyruvate in 400 mM sodium phosphate, 10 mM magnesium chloride and 5 mM TCEP, pH 7.5.

To compare WT and F190V DszAT under pure and mixed carboxyacyl-CoA extender unit conditions, the assay mixture contained 1 mM fluoromalonyl-CoA, malonyl-CoA and/or methylmalonyl-CoA, 10 mM NDK-SNAC, 10 μM $Mod3_{DEBS}$ + TE(AT^0) and DszAT variants. Epi, MatB, PK, myokinase, PEP, ATP and carboxyacyl-CoA substrate were preincubated at 37 °C for 30 min. Reactions were then initiated with NDK-SNAC, $Mod3_{DEBS}$ + TE(AT^0) and *trans*-AT and incubated for 1 h at 37 °C.

For FabD activity assay, assay mixture contained 1 mM of malonate substrate (fluoromalonate, malonate or methylmalonate), 1 mM of CoA, 1 mM NDK-SNAC, 10 μM $Mod3_{DEBS}$ + TE(AT^0) and 30 μM of *trans*-AT enzyme. Epi, MatB, PK, myokinase, PEP, ATP and acyl-CoA substrate were preincubated at 37 °C for 30 min. The reactions were then initiated with NDK-SNAC, $Mod3_{DEBS}$ + TE(AT^0) and *trans*-AT and incubated overnight at 37 °C.

To quench the reaction, 25 μl of reaction mixture was mixed with 25 μl of 10% (v/v) perchloric acid and centrifuged for 10 min at 4 °C at 20,817*g*. The supernatant was removed and stored at −80 °C until analysis.

Quantification of TKL products by LC–QQQ.

Frozen samples were centrifuged for 10 min at 4 °C at 20,817*g* to remove salts. The supernatant was removed and analyzed by LC–QQQ using MRM in positive ionization mode on an Agilent 6460 QQQ MS. Samples were analyzed on two tandem Ascentis Express RP-amide HPLC columns (2.7 μm, 2.1 mm × 10 cm, 25 °C; Supelco) on an Agilent 1290 UPLC using an isocratic solvent system of 10% acetonitrile in 0.1% (v/v) formic acid as the aqueous mobile phase for 20 min after an initial ramp of 0 to 10% acetonitrile over 0.5 min at flow rate of 0.4 ml min^{-1}. Products were monitored with transitions as follows (parent ion *m/z* → product ion *m/z*, fragmentor, collision energy): TKL (171 → 153, 90,6); H-TKL (157 → 139, 90, 4); F-TKL (175 → 157, 90, 5) and H-TTKP (199 → 111, 100,

9). The identity and the quantity of each compound was verified against synthetic standards using external standard curves. Synthetic standards were obtained as previously described[28].

Assay for in vitro 6dEB analog production.

All assays (Mod6 AT^0 DEBS, 30 μl; Mod3 AT^0 DEBS and Mod5 AT^0 DEBS, 50 μl) contained 2 μM LDD_{DEBS}, 2 μM $Mod1_{DEBS}$, 2 μM $Mod2_{DEBS}$, 2 μM DEBS2 variant (WT or DEBS2(Mod3 AT^0)) and 2 μM DEBS3 variant (WT, DEBS3(Mod5 AT^0) or DEBS3(Mod6 AT^0)), 2 μM MatB, 4 μM Epi, 2 μM PrpE, 1 mM CoA, 4 mM ATP, 2 mM methylmalonate, 0.5 mM NADPH, 0.5 mM propionate in 250 mM sodium phosphate, 10 mM magnesium chloride and 5 mM TCEP, pH 7.5. When used, 10 mM fluoromalonate, 2 mM malonate and 6 μM DszAT variant (WT or F190V) were added. Reaction mixtures were incubated at room temperature for 3 h before quenching and extracting twice with 4× reaction volume of ethyl acetate The combined organic layers were dried using a Speed-Vac SC110 (Savant).

Identification of 6dEB analogs by LC–QTOF and LC–QQQ.

The reaction residue was resuspended in methanol (Mod6 AT^0 DEBS, 30 μl; Mod3 AT^0 DEBS and Mod5 AT^0 DEBS, 25 μl) and centrifuged for 10 min at 4 °C at 20,817*g*. Products were analyzed by LC–QTOF in positive ionization mode on at Agilent 6530B QTOF MS. Samples were analyzed on Ascentis Express RP-amide HPLC columns (2.7 μm, 2.1 mm × 10 cm, 25 °C; Supelco) on an Agilent 1290 UPLC using a linear gradient from 5 to 65% acetonitrile over 16.75 min with 0.1% (v/v) formic acid as the aqueous mobile phase after an initial increase of 0 to 5% acetonitrile over 30 s at flow rate of 0.25 ml min^{-1}. Masses observed correspond to $[M + H^+\text{-}H_2O]^+$ of the target molecules as follows (molecular formula, expected mass, observed mass, Δppm(*exp-obs*)): 6dEB ($C_{21}H_{37}O_5^+$, 369.2636, 369.2639, 0.08 ppm), desmethyl 6dEB ($C_{20}H_{35}O_5^+$, 355.2479, 355.2476, 0.8 ppm) and fluorodesmethyl 6dEB ($C_{20}H_{34}FO_5^+$, 373.2385, 373.2384, 0.3 ppm). Fragmentation analysis was carried out using the targeted tandem MS (MS/MS) mode with fragmentation voltage of 185 and collision energy of 10. From the fragmentation patterns, unique transitions were developed for detection with increased sensitivity with LC–QQQ using MRM in positive ionization mode on an Agilent 6460 QQQ MS with the same LC method. Products were monitored with transitions as follows (parent *m/z* → product *m/z*, fragmentor, collision energy): 6dEB (369.2 → 239.1, 50, 1), 2-desmethyl 6dEB and 4-desmethyl 6dEB (355→239.1, 50, 1), 2-fluoro-2-desmethyl 6dEB and 4-fluoro-4-desmethyl 6dEB (373.2 → 275.1, 105, 0).

In vivo production of TKLs in *E. coli*.

LB medium (50 ml) containing appropriate antibiotics (carbenicillin and chloramphenicol, 50 μg ml^{-1}; spectinomycin, 100 μg ml^{-1}) in a 250-ml baffled flask was inoculated with 1 ml of an overnight LB culture of *E. coli* BAP1 or BAP1-$T1^R$ transformed with relevant plasmids. Cultures were grown at 37 °C with shaking at 200 r.p.m. to OD_{600} = 0.4–0.6. Then, cultures were cooled on ice for 20 min and protein expression was induced with 1 mM IPTG and 0.2% (*w/v*) arabinose. Induced cultures were incubated at 16 °C with shaking for 16–24 h. Cells (40 ml) were gathered by centrifugation at 1,000*g* for 15 min at 4 °C after recording the final OD_{600}. Cells were washed once with 40 ml of M9 medium and

resuspended in M9 medium to OD_{600} of roughly 100. The high-density cell suspension (50 μl) was added to a 1.7-ml microcentrifuge tube along with 5 mM fluoromalonate (1 M stock in water) and 1 mM NDK-SNAC (250 mM stock in 25% DMSO). Cell suspensions were incubated with the two substrates at 16 °C with shaking at 200 r.p.m. for 20–24 h. Then, the cell suspension was centrifuged at 20,817*g* for 1 min at 4 °C. The supernatant (35 μl) was removed and mixed with 10% (v/v) perchloric acid (35 μl) and centrifuged at 20,817*g* for 10 min at 4 °C. The supernatant was removed and stored at −80 °C until analysis.

Malonate/fluoromalonate growth curves.

E. coli BAP1 cultures were transformed with either pFMal(MadLM), which contains both MadLM and MatB, or pFMal(MadLM, no MatB). Cultures were grown to OD_{600} of roughly 0.4 at 37 °C in 6dEB production media[53] with shaking at 200 r.p.m., at which point 200 μl of cultures were collected and supplemented with 1 mM IPTG, 0.2% arabinose and, when used, 20 mM sodium propionate, 5 mM malonate or 5 mM fluoromaloante final concentration. Cultures (100 μl) were then transferred to 96-well microplate (clear, polystyrene, U-bottom, sterile, Greiner Bio-One). Plate was incubated with continuous shaking uncovered at room temperature. Growth was monitored in absorbance mode at 600 nm using a Synergy Mx microplate reader with measurements taken every 2.5 min for 16 h. OD_{600} data were corrected for the path length of 100 μl of liquid (0.3 cm).

In vivo production of 6dEB analogs in *E. coli*.

The 6dEB analogs were produced in *E. coli* BAP1 or BAP1-T1^{R} cultures grown in growth medium previously optimized for 6dEB production[53]. The 6dEB production medium (50 ml) containing appropriate antibiotics in a 250 l baffled flask was inoculated with 1 ml of overnight culture of *E. coli* transformed with relevant plasmids in the same medium. Cultures were grown to OD_{600} = 0.4–0.6 at 37 °C with shaking at 200 r.p.m. Then, cultures were cooled on ice for 20–40 min and protein expression was induced with 1 mM IPTG. Cultures were incubated at 22 °C with shaking at 250 r.p.m. for 16–24 h. Cells were then collected by centrifugation at 10,000*g* for 5 min at 4 °C and resuspended in 1.5 ml of fresh media containing 1 mM IPTG, 0.2% (*w*/*v*) arabinose, 5 mM fluoromalonate, 20 mM propionate and appropriate antibiotics. High-density cell suspensions were grown in 14-ml round-bottom Falcon culture tubes for another 16–24 h at 22 °C with shaking at 250 r.p.m., after which they were collected by centrifugation at 20,817*g* for 1 min at 4 °C. Supernatant (1.5 ml) was removed and extracted with 4 × 0.5 ml ethyl acetate by vortexing for 15–30 s and separating the layers by centrifugation at 13,000*g* for 3 min. The combined organic layers were dried by Speed-Vac SC110 (Savant), resuspended in 30 μl of methanol for analysis. For isolation of product 2-fluoro-2-methyl 6dEB for NMR analysis, production was scaled up to 61 (12 × 500 ml) and followed procedure described above.

Purification of the 2-fluoro-2-desmethyl 6dEB analog from *E. coli* culture medium.

Supernatant (40 ml) was extracted with 4 × 10 ml ethyl acetate by vortexing for 15–30 s and separating the layers by centrifugation at 8,000*g* for 5 min. The combined organic layers were dried by rotary evaporation. The extract was purified on an Eclipse XDB-C18 (5 μm, 9.4 × 250 mm, Agilent) on an Agilent 1200 HPLC using a linear gradient from 15 to 45% acetonitrile over 42 min with water as the aqueous mobile phase after an initial increase

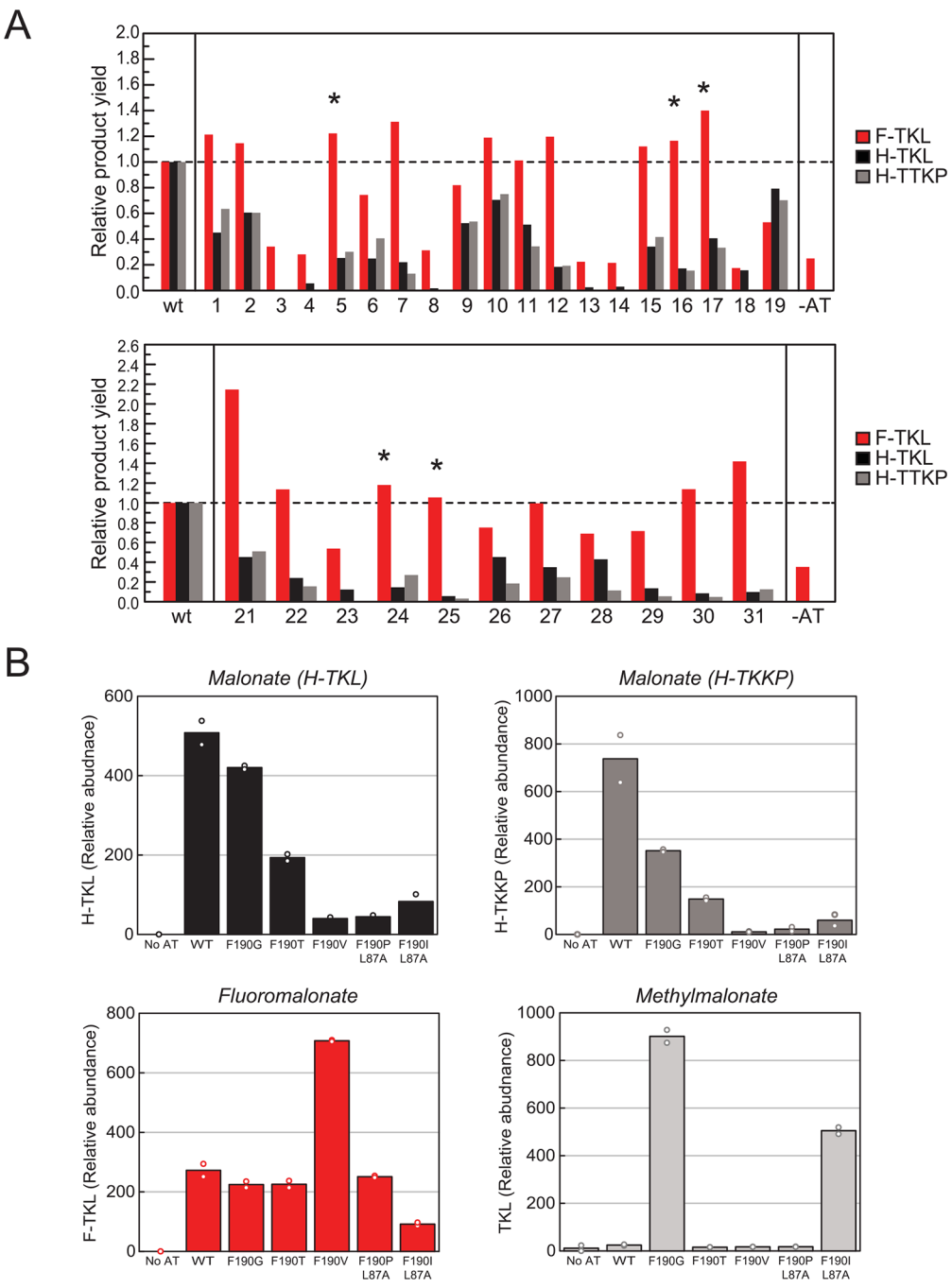
of 0 to 15% acetonitrile over 1.5 min at a flow rate of 4 ml min^{-1}. Fractions (2 ml) were collected and screened using LC–QQQ. Fractions were analyzed on Poroshell 120 EC-C18 (2.7 μm, 2.1 × 50 mm, Agilent) on Agilent 1290 UPLC using a linear gradient from 25 to 60% acetonitrile over 2 min with 0.1% (v/v) formic acid as the aqueous mobile phase after an initial increase of 0 to 25% acetonitrile over 12 s at flow rate of 0.7 ml min^{-1}. Fractions containing 2-fluoro-2-desmethyl 6dEB shown by transition 373.2 → 275.1 were combined and lyophilized to dryness.

^{19}F-NMR of fluorinated 6dEB analog isolated from *E. coli* culture medium.

Lyophilized culture extract containing 2-fluoro-2-desmethyl 6dEB was resuspended in 500 μl of 50:50 MeOH:D_2O mixture for ^{19}F-NMR analysis. Then 5-fluorouracil (1 mM, 100 μl) in D_2O was added to the coaxial insert as internal standard.

^{19}F-NMR spectrum was collected at 298 K on a Bruker AV-600 spectrometer at the College of Chemistry NMR Facility. Tune and shim were optimized. The spectrum was collected using the 19f.cocnmr.av600 pulse program with following parameters: o1p = −200, sw = 200, td = 131,072, d1 = 1 s, ds = 20, ns = 3,200. The resulting spectrum was processed by MestReNova by backward linear prediction from 0 to 128 using Toeplitz method and phased manually. The spectrum was referenced by a set of doublet corresponding to 5-fluorouracil at −171 ppm and ^{19}F-NMR (600 MHz, MeOH:D_2O) δ −195.6 (doublet of doublet, J = 54,12 Hz).

Extended Data

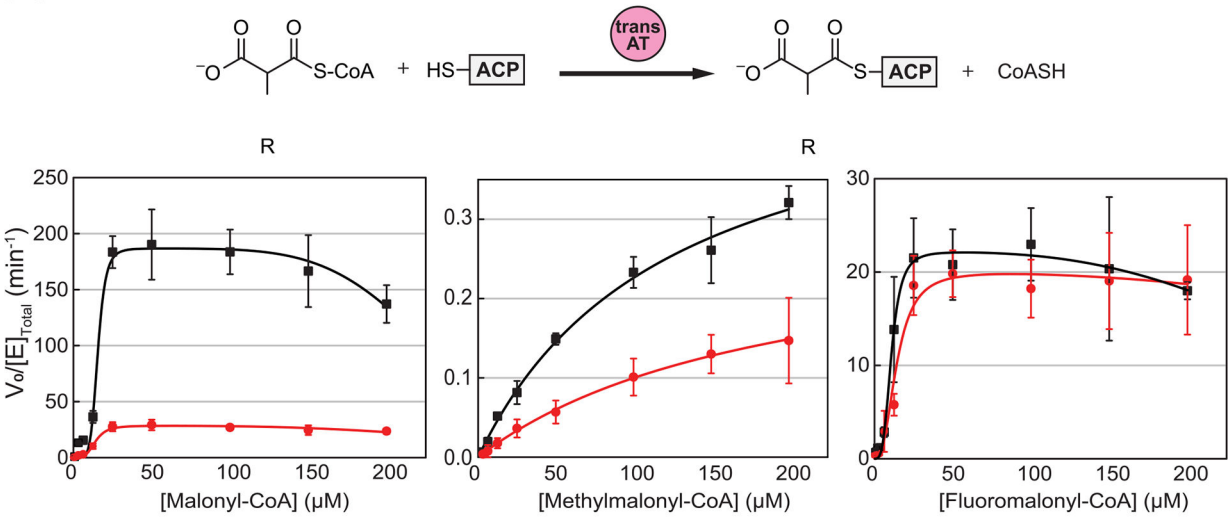


Extended Data Fig. 1 |. trans-AT library screening.

(A) Lysate of *E. coli* expressing following DszAT mutants were screened through the triketide lactone production assay with purified $Mod3_{DEBS}$ + TE(AT^0) protein, malonate or fluoromalonate extender unit, malonate-coeznzyme A ligase MatB and NDK-SNAC: **1**, F190A; **2**, F190C; **3**, F190D; **4**, F190E; **5**, F190G; **6**, F190H; **7**, F190I; **8**, F190K; **9**, F190L; **10**, F190M; **11**, F190N; **12**, F190P; **13**, F190Q; **14**, F190R; **15**, F190S; **16**,

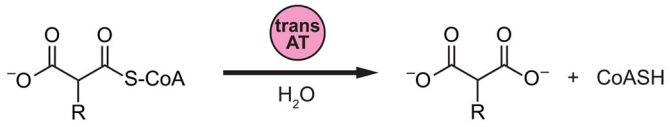
F190T; **17**, F190V; **18**, F190W; **19**, F190Y; **21**, F190G L87V; **22**, F190G L87A; **23**, F190I L87V; **24**, F190I L87A; **25**, F190P L87A; **26**, F190S L87V; **27**, F190S L87A; **28**, F190T L87V; **29**, F190T L87A; **30**, F190V L87V; **31**, F190V L87A. Product formation was monitored by LC-QQQ using negative ionization mode (transition): F-TKL (173→59), H-TKL (155→97), H-TTKP (197→95), TKL (169→111). The amounts of F-TKL (red), H-TKL (black), and H-TTKP (gray) products were determined by integrating extracted ion counts for the relevant species and are shown normalized by the amount of corresponding product produced by wild-type enzyme. H-TTKP arises from two chain extensions with malonyl-CoA. Mutants that showed potentially increased selectivity towards fluoromalonyl-CoA (underlined) compared to wild-type enzyme were selected for in vitro screening. **(B)** Selected DszAT mutants from lysate screen (F190G, F190T, F190V, F190P L87A and F190I L87A) were overexpressed, purified and assessed through in vitro triketide lactone experiment with purified $Mod3_{DEBS}$ + TE(AT^0) protein, MatB and NDK-SNAC for the production of H-TKL and H-TTKP from malonyl-CoA (top panels), F-TKL from fluoromalonyl-CoA (bottom left) and TKL from methylmalonyl-CoA (bottom right). The amounts of products were determined as in (A). Data are mean of two technical replicates.

A

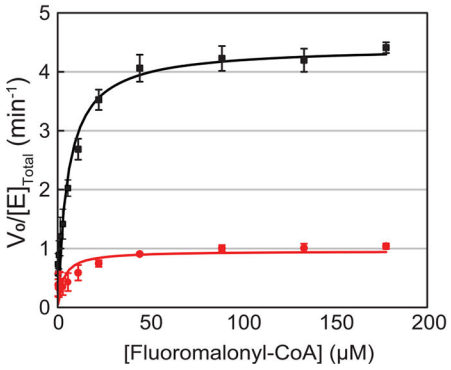


Substrate	R	DszAT	k_{cat} (min^{-1})	K (μM)	k_{cat}/K (min^{-1} $μM^{-1}$)	K_i (μM)	n
Malonyl-CoA	H	WT	187 ± 8.5	16 ± 1.5	12 ± 1.7	233 ± 16	6.2 ± 2.3
		F190V	29 ± 1.5	14 ± 1.1	2.0 ± 0.3	282 ± 37	4.0 ± 1.3
Methylmalonyl-CoA	CH_3	WT	0.51 ± 0.04	123 ± 21	0.004 ± 0.001	na	na
		F190V	0.29 ± 0.08	193 ± 95	0.0015 ± 0.0012	na	na
Fluoromalonyl-CoA	F	WT	22 ± 1.6	11 ± 1.3	2.1 ± 0.4	304 ± 73	3.5 ± 1.4
		F190V	20 ± 2.1	14 ± 2.6	1.4 ± 0.4	542 ± 446	2.6 ± 1.0

B



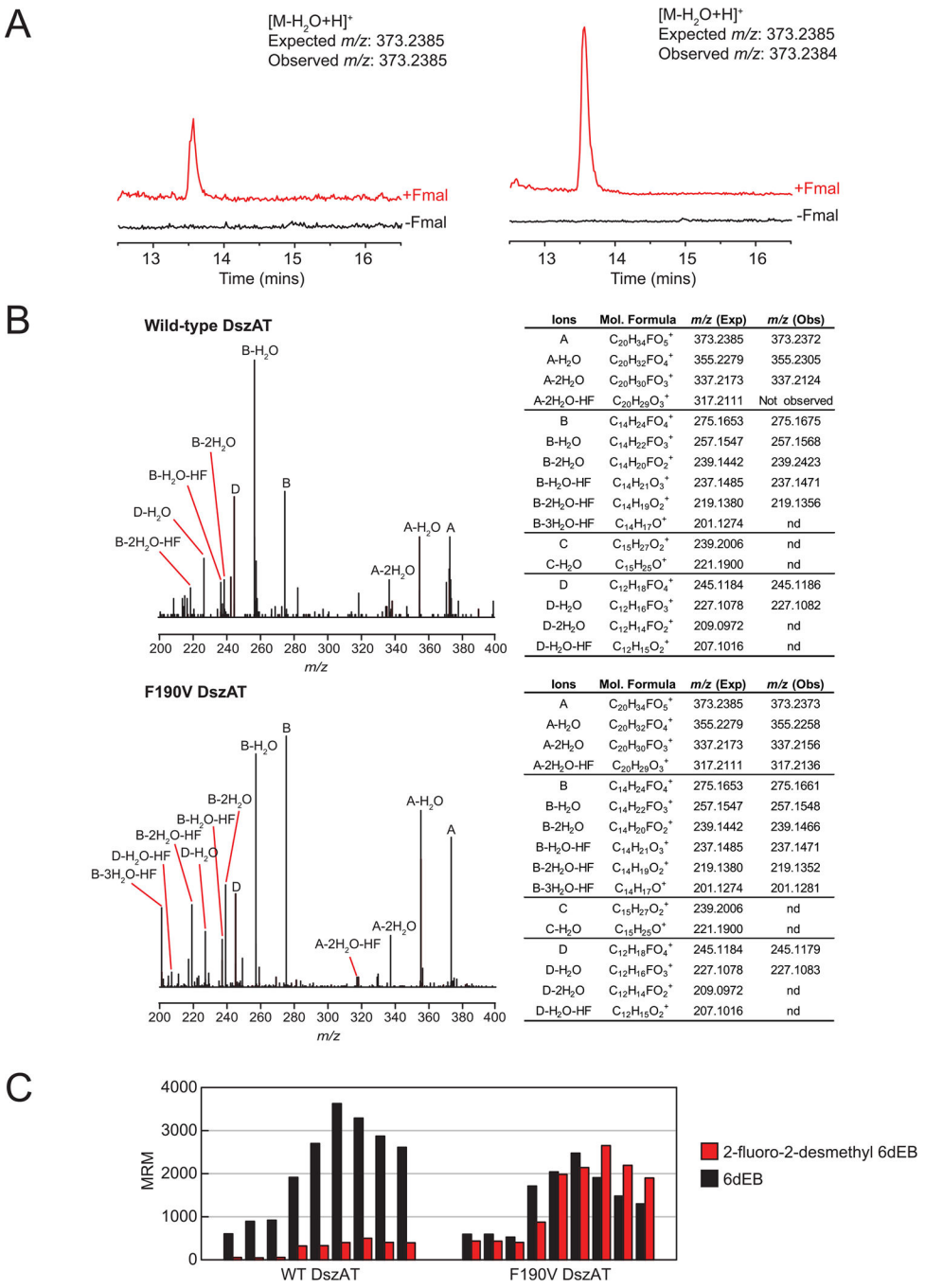
DszAT	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($min^{-1}μM^{-1}$)
Wild-type	4.4 ± 0.1	6.1 ± 0.7	0.72 ± 0.1
F190V	0.96 ± 0.06	3.5 ± 1.1	0.28 ± 0.11



Extended Data Fig. 2 |. Characterization of DszAT F190V.

(A) Steady-state kinetic analysis of transacylation reaction of malonyl-CoA (left), methylmalonyl-CoA (middle) and fluoromalonyl-CoA (right) by wild-type (black) and F190V (red) DszAT with 75 uM of $ACP_{DEBSMod6}$ as acyl accepter. Transacylation activity was measured by α-ketoglutarate dehydrogenase coupled assay monitoring CoA release. The dose-response curves for transacylation of methylmalonyl-CoA by WT and F190V DszAT were fit with Michaelis-Menten equation. The dose-response curves for WT and F190V DszAT transacylation with malonyl- and fluoromalonyl-CoA exhibited sigmoidal

behavior with evidence of inhibition at high substrate concentrations and were therefore fit with the Hill equation modified for substrate inhibition. Data shown are mean ± s.d. of three technical replicates. **(B)** Steady-state kinetic analysis of hydrolysis reaction of fluoromalonyl-CoA by wild-type and F190V DszAT. Hydrolytic activity was measured by DTNB assay monitoring CoA release. The dose-response curves were obtained from non-linear fitting of data to Michaelis-Menten equation. Data points are mean ± s.d. of three technical replicates.

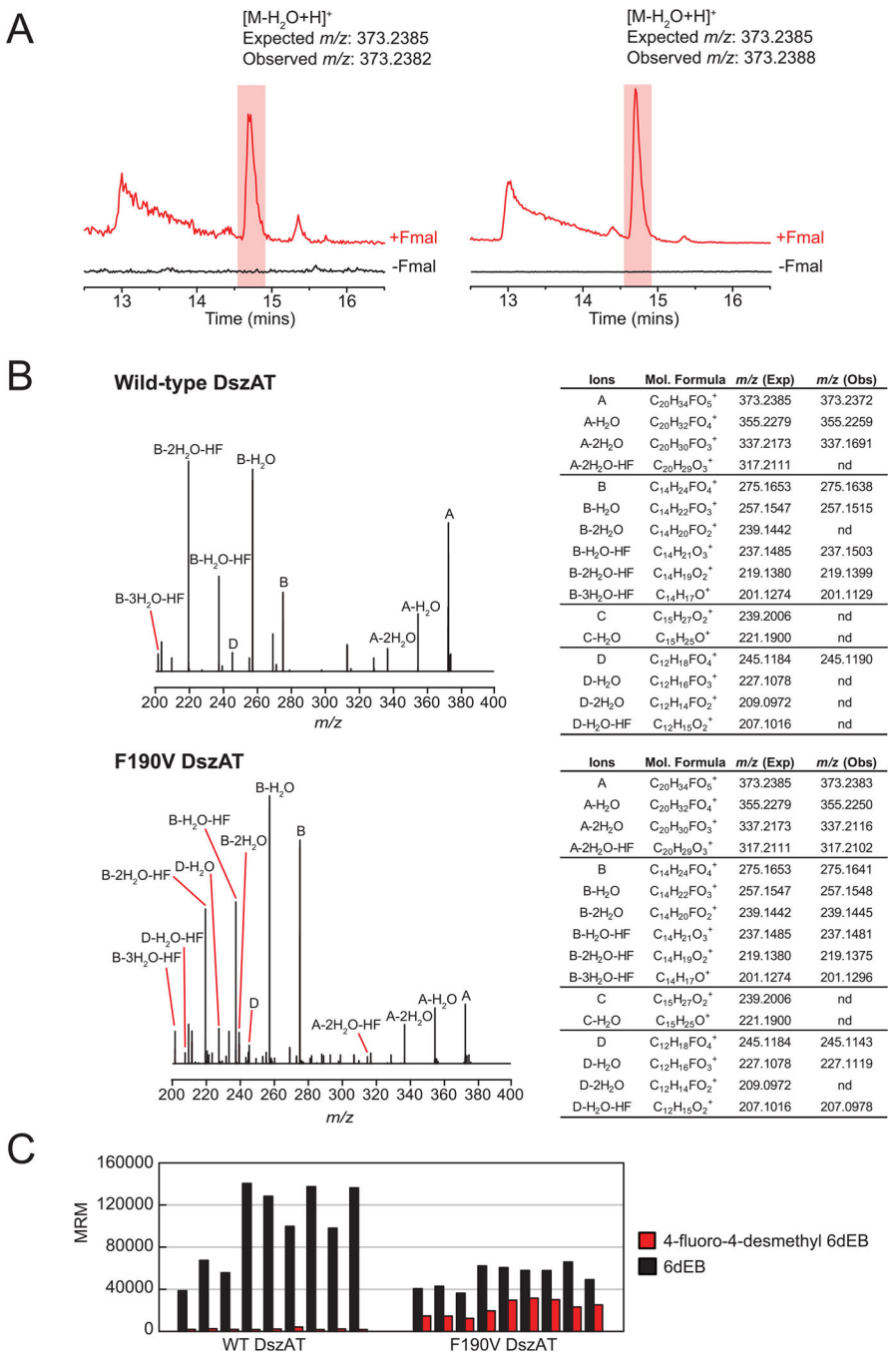


Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2372
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2305
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2124
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	Not observed
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1675
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1568
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.2423
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1471
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.1380	219.1356
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	nd
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1186
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1082
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	nd
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	nd

Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2373
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2258
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2156
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	317.2136
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1661
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1548
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.1466
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1471
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.1380	219.1352
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1281
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1179
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1083
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	nd
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	nd

Extended Data Fig. 3 |. Generation of 2-fluoro-2-desmethyl-6dEB analog through in vitro reconstitution of Mod6 AT0 DEBS.

(A) Extracted ion chromatograms showing exact mass of in vitro production of monofluorinated desmethyl 6dEB by Mod6 AT^0 DEBS complemented with WT DszAT (left) or F190V DszAT (right). Reactions containing 2 μM each of LDD_{DEBS}, $Mod1_{DEBS}$, $Mod2_{DEBS}$, DEBS2 and DEBS3(Mod6 AT^0), 6 μM WT or F190V DszAT, 2 mM methylmalonate and when used 10 mM fluoromalonate (+Fmal, red) were analyzed by LC-QTOF. Chromatograms are representative of at least three technical replicates. **(B)** Fragmentation pattern of the monofluorinated desmethyl 6dEB analog produced by Mod6 AT^0 DEBS complemented by WT DszAT or F190V DszAT is consistent with 2-fluoro-2-desmethyl 6dEB, which is expected when fluoromalonyl-CoA is incorporated by module 6. Fragmentation pattern shows presence of ions in A, B and D families (Supplementary Fig. 5) with mass shifts according to a substitution of $-CH_3$ with -F group on carbon 2, 4, 6 or 8 of 6dEB. Observed loss of HF from daughter ions expected to contain fluorine substitution further supports the presence of fluorine in the molecule. Notably, the most prominent fragment in 6dEB MS/MS spectrum, the C ion, was entirely missing in the MS/MS spectrum of the monofluorinated desmethyl 6dEB analog[39]. This lack of detectable fragment prevents narrowing down the definite location of the fluorine substitution. However, the suppression of ion C was similarly observed when malonate was provided to the system in place of fluoromalonate (Supplementary Fig. 6) and is consistent with the reactivity of 6dEB analog with a modification near the C3 position. The lack of methyl substituent at C2 as in 2-fluoro-2-methyl 6dEB would reduce the reactivity of A ion with respect to the hydride shift from C3 to C9 necessary to form C ion. Spectra are representative of at least three technical replicates (nd, not detected). **(C)** Relative amounts of 2-fluoro-2-desmethyl 6dEB and 6dEB by complementation of Mod6 AT^0 DEBS by WT or F190V DszAT. The amounts of products were determined by integrating extracted ion counts for the relevant species monitored by LC-QQQ (transition): 6dEB (369.2→239.1) and 2-fluoro-2-desmethyl 6dEB (373.2 → 275.1). Although the absolute amounts of the two observed products vary among replicates; F190V DszAT reactions consistently produce higher ratio of fluorinated to non-fluorinated product than WT DszAT.



Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2372
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2259
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.1691
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	nd
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1638
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1515
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	nd
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1503
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.1380	219.1399
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1129
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1190
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	nd
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	nd
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	nd

Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2383
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2250
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2116
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	317.2102
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1641
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1548
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.1445
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1481
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.1380	219.1375
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1296
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1143
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1119
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	nd
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	207.0978

Extended Data Fig. 4 |. Generation of 4-fluoro-4-desmethyl-6dEB analog through in vitro reconstitution of Mod5 AT0 DEBS.

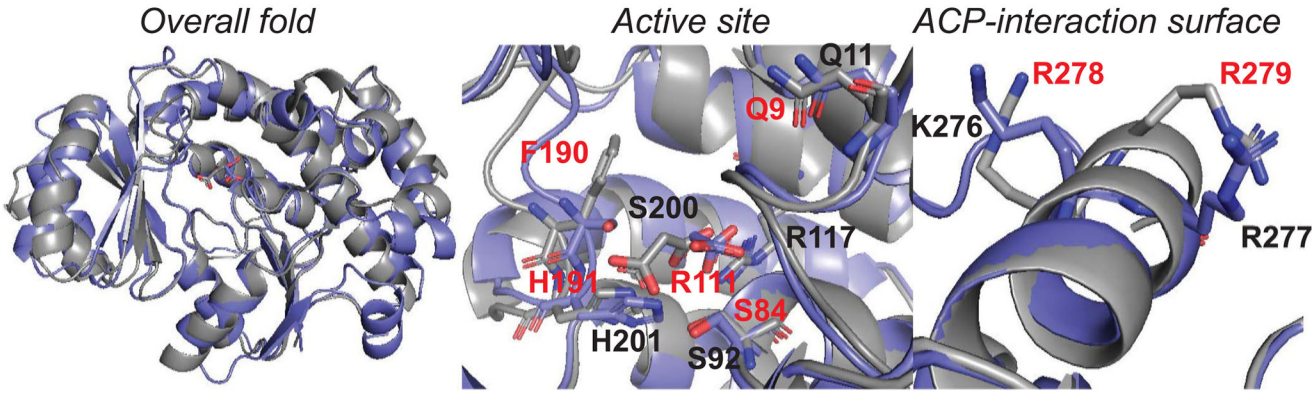
(A) Extracted ion chromatograms showing exact mass of in vitro production of monofluorinated desmethyl 6dEB by Mod5 AT^0 DEBS complemented with WT DszAT (left) or F190V DszAT (right). Reactions containing 2 μM each of LDD_{DEBS}, $Mod1_{DEBS}$, $Mod2_{DEBS}$, DEBS2 and DEBS3(Mod5 AT^0), 6 μM WT or F190V DszAT, 2 mM methylmalonate and when used 10 mM fluoromalonate (+Fmal, red) were analyzed by LC-QTOF. Chromatograms are representative of at least three technical replicates. **(B)**

Fragmentation pattern of the monofluorinated desmethyl 6dEB analog produced by Mod5 AT^0 DEBS complemented by WT DszAT or F190V DszAT is consistent with 4-fluoro-4-desmethyl 6dEB, which is expected when fluoromalonyl-CoA is incorporated by Mod5. The observed daughter ion masses are necessarily similar to those of 2-fluoro-2-desmethyl 6dEB observed in Mod6 AT^0 DEBS system, due to fluorine substitution in similar positions. In short, mass shifts according to a substitution of $-CH_3$ with -F group on carbon 2, 4, 6 or 8 of 6dEB, loss of HF from daughter ions and absence of C ion are observed. Spectra are representative of at least three technical replicates (nd, not detected). **(C)** Relative amounts of 4-fluoro-4-desmethyl 6dEB and 6dEB by complementation of Mod5 AT^0 DEBS by WT or F190V DszAT. The amounts of products were determined by integrating extracted ion counts for the relevant species monitored by LC-QQQ (transition): 6dEB (369.2 → 239.1) and 2-fluoro-2-desmethyl 6dEB (373.2 → 275.1).

A

```
FabD   1    MTQFAFVFPGQGSQTVGMLADMAASYPIVEETFAEASAALGYDLWALTQQGPAEELNKTW  60
DszAT  1    --MKAYMFPGQGSQAKGMGRALFDAFPALT---ARADGVLGYSIRALCQDDPDQRLSQTQ  55
                *::********: **   :  ::* :    * *...***.:.** *:.* : *.:*.

FabD   61   QTQPALLTASVALYRVWQQQGGKAPAMMAGHSLGEYSALVCAGVIDFADAVRLVEMRGKF  120
DszAT  56   FTQPALYVVNALSY-LKRREEEAPPDFLAGHSLGEFSALFAAGVFDFETGLALVKKRGEL  114
             ***** ....  * :...:   .* ::********:***..***:**  .: **: **::

FabD   121  MQEAVPEGTGAMAAIIGLDDASIAKACEEAAEGQVVSPVNFNSPGQVVIAGHKEAVERAG  180
DszAT  115  MGDA---RGGGMAAVIGLDEERVRELLDQNG-ATAVDIANLNSPSQVVISGAKDEIARLQ  170
            * :*     *.***:****:  : :  :: . . .*. .*:***.****:* *: : *

FabD   181  AACKAAGAKRALPLPVSVPSHCALMKPAADKLAVELAKITFNAPTVPVVNNVDVK-CETN  239
DszAT  171  VPFEAAGAKKYTVLRVSAAFHSRFMRPAMVEFGRFLEGYDFAPPKIPVISNVTARPCKAD  230
            .. :*****.   * **.. *. :*.**  ::.  *    * .*.:**:.** .. *:::

FabD   240  GDAIRDALVRQLYNPVQWTKSVEYMAAQGVEHLYEVGPGKVLTGLTKRIVDTLTASALNE  299
DszAT  231  G--IRAALSEQIASPVRWCESIRYLMGRGVEEFVECGHGIVLTGLYAQI--------RRD  280
            *  ** **  *: .**.* :*: *: ..*** : * * * *****  .*        .:

FabD   300  PSAMAAALEL                                                  309
DszAT  281  AQPLVVA---                                                  287
            ...:..*
```

B

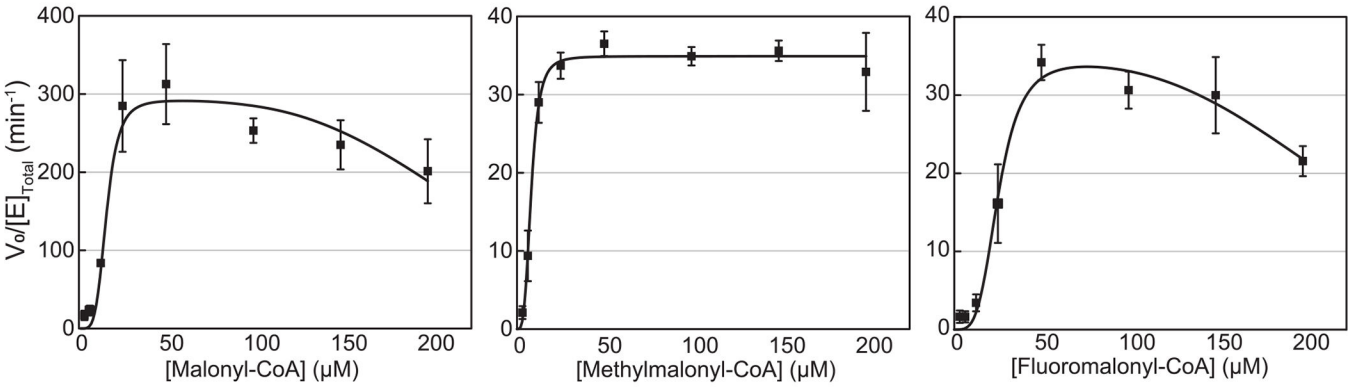


Extended Data Fig. 5 |. Sequence and structural analysis of E. coli FabD.

(A) Sequence comparison between FabD and DszAT. Conserved functional residues are shown in red; specificity-determining residue are shown in blue; and ACP-interacting residues are shown in green. **(B)** Structural comparison between Fab (blue) and DszAT (grey). The two proteins have the same overall structure (left, rmsd = 1.7 Å), active site architecture (middle) and ACP-interacting surface (right). Both proteins contain two subdomains with the active site in between. The substrate-binding pocket is largely polar. In both structures, malonate is bound by salt bridge interactions between its β-carboxy

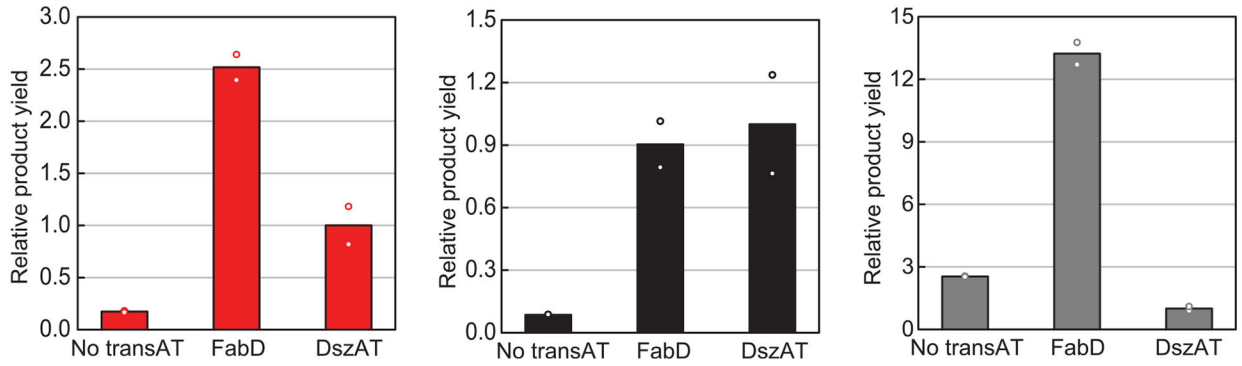
group and the side chain amine groups of R117/R111, positioning the carbonyl carbon for nucleophilic attack near the catalytic serine S84/S92, which is found in nearly identical position in both proteins. Notably, in DszAT, F190 is found positioned near the Cα of malonate that may prevent binding of substrate with bulkier substituent such as a methyl group at the α-position. In FabD, this residue is replaced with S200, which may provide more flexibility in terms of substrate binding. Interactions between AT and ACP domains has previously been shown to rely on a small number of charged amino acid residues located near the active sites of the two protein domains[40,54]. R278 and R279 of DszAT interact with E70 of its native protein partner, DSZS ACP1. Similarly, K180 of DszAT interacts with D45 of DSZS ACP1. Structural alignment shows that these positively charged interface residues of DszAT, R278, R279 and K180, correspond to K276, R277 and R190 of FabD, maintaining the charge at the interface for interaction with an ACP partner. Active site, specificity-determining, and interface residues are shown in sticks.

A



Substrate	k_{cat} (min^{-1})	$K_{0.5}$ (μM)	$k_{cat}/K_{0.5}$ (min^{-1} $μM^{-1}$)	K_i (μM)	n
Malonyl-CoA	293 ± 20	15 ± 1.5	20 ± 3.3	230 ± 19	4.3 ± 1.5
Methylmalonyl-CoA	36 ± 0.9	8.3 ± 0.4	4.2 ± 0.3	429 ± 92	3.4 ± 0.5
Fluoromalonyl-CoA	35 ± 3	25 ± 2	1.4 ± 0.2	230 ± 15	3.8 ± 1.2

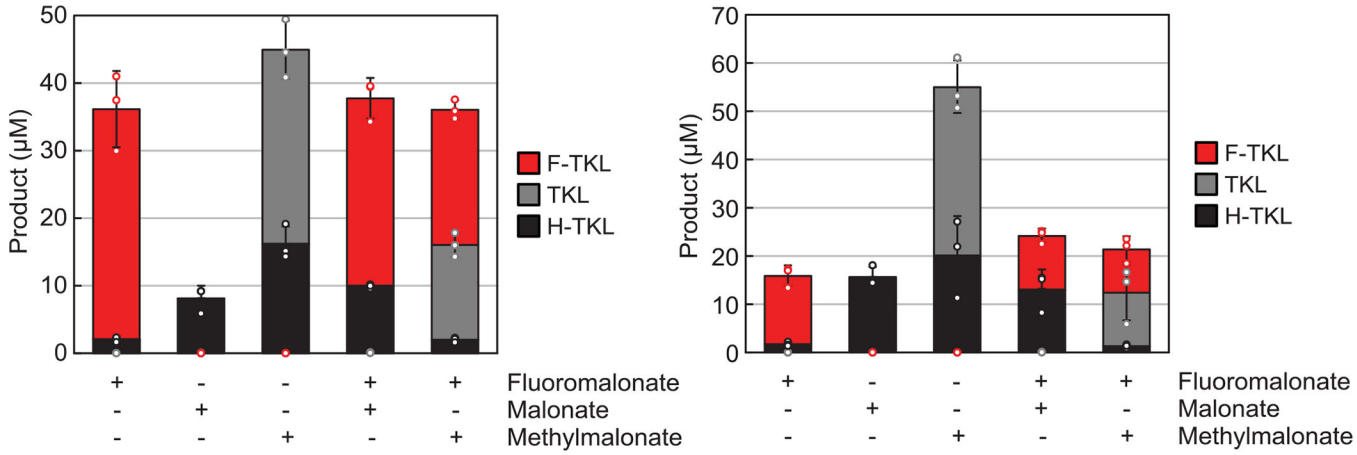
B



Extended Data Fig. 6 |. Characterization of E. coli FabD transacylation activity with the DEBS PKS.

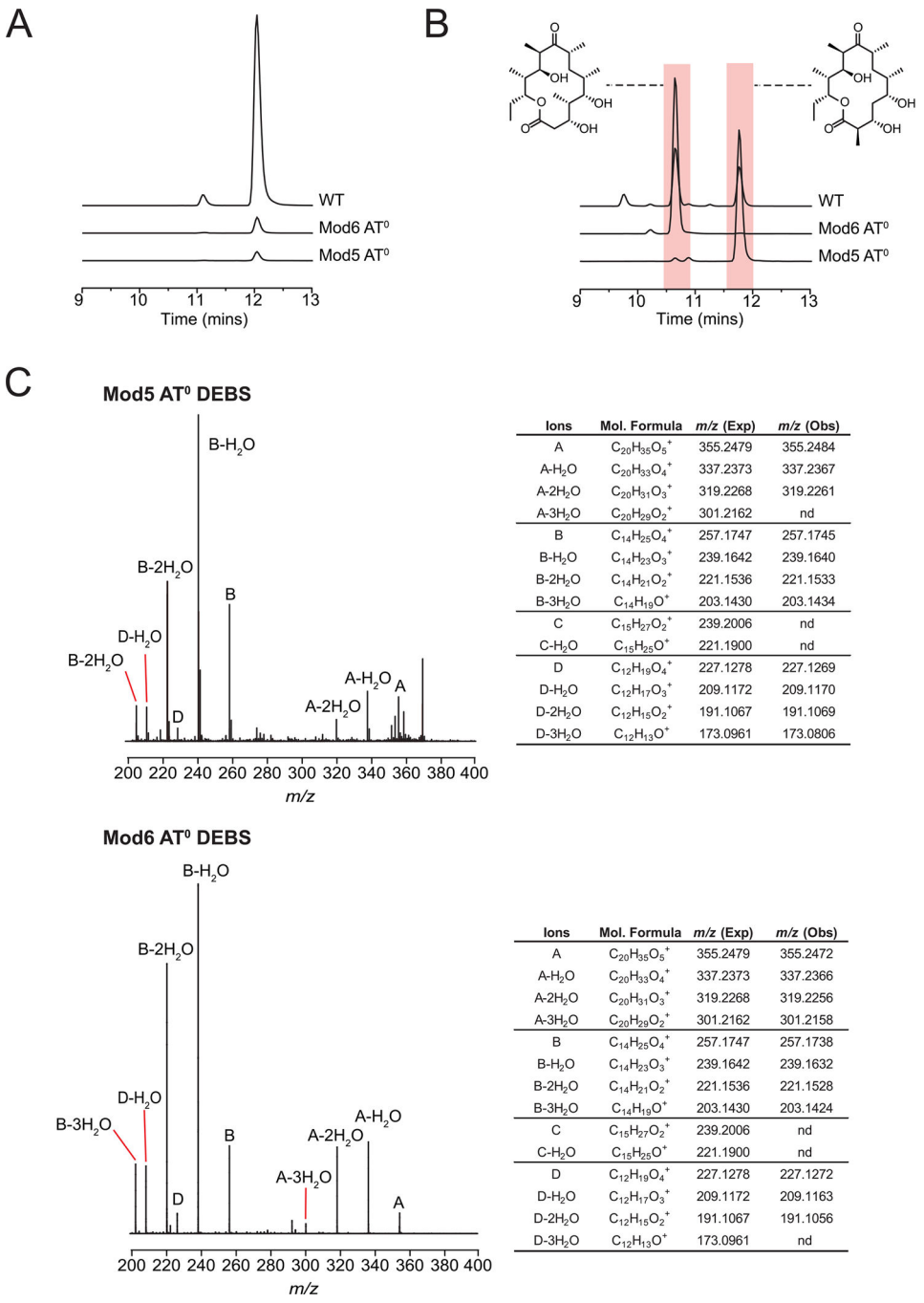
(A) Steady-state kinetic characterization of transacylation reaction of malonyl-CoA (left), methylmalonyl-CoA (middle) and fluoromalonyl-CoA (right) catalyzed by FabD with 75 μM of $ACP_{DEBSMod6}$ as acyl accepter. Activity was measured by α-ketoglutarate dehydrogenase coupled assay, monitoring coenzyme A release by NADH fluorescence. Curves were obtained from non-linear curve fitting of data to Hill equation modified for substrate inhibition model. Table contains k_{cat}, $K_{0.5}$, $k_{cat}/K_{0.5}$, K_i and n values. Data shown are mean ± s.d of three technical replicates. Error in k_{cat}/K_M is obtained by propagation from individual kinetic terms**. (B)** In vitro triketide lactone assay for FabD complementation of $Mod3_{DEBS}$ + TE(AT^0). Normalized representation of the amount of F-TKL (left), H-TKL (middle) and TKL (right) products are shown for reactions containing 10 μM $Mod3_{DEBS}$

+ TE(AT^0)), 30 μM FabD or WT DszAT, 1 mM NDK-SNAC, 1 mM fluoromalonate (left), malonate (middle) or methylmalonate (right), and 1 mM coenzyme A under regenerative condition. Reactions were quenched and analyzed by LC-QQQ after 24 h. The amounts of products were determined by integrating extracted ion counts for the relevant species (transition): F-TKL (175 → 157), H-TKL (157 → 139) and TKL (171 → 153). Data are mean of two technical replicates.



Extended Data Fig. 7 |. Influence of extender unit availability on the in vivo selectivity of chain elongation by single-modular DEBS constructs in engineered E. coli.

Concentrated cell suspension of *E. coli* BAP1 expressing $Mod3_{DEBS}$ + TE(AT^0) (left) or $Mod6_{DEBS}$ + TE(AT^0) (right), MatB and MadLM were provided with 1 mM NDK-SNAC and 5 mM fluoromalonate, malonate and/or methylmalonate and analyzed by LC-QQQ after 24 h. The concentrations of products were determined by integrating extracted ion counts for the relevant species (transition): F-TKL (175 → 157), H-TKL (157 → 139) and TKL (171 → 153), and comparing to external standard curves generated using synthetic standards of the molecules. Data revealed that the product profile follows the corresponding precursor profile, suggesting that the outcome is mainly governed by the provided precursors. Data are mean ± s.d. of three biological replicates.

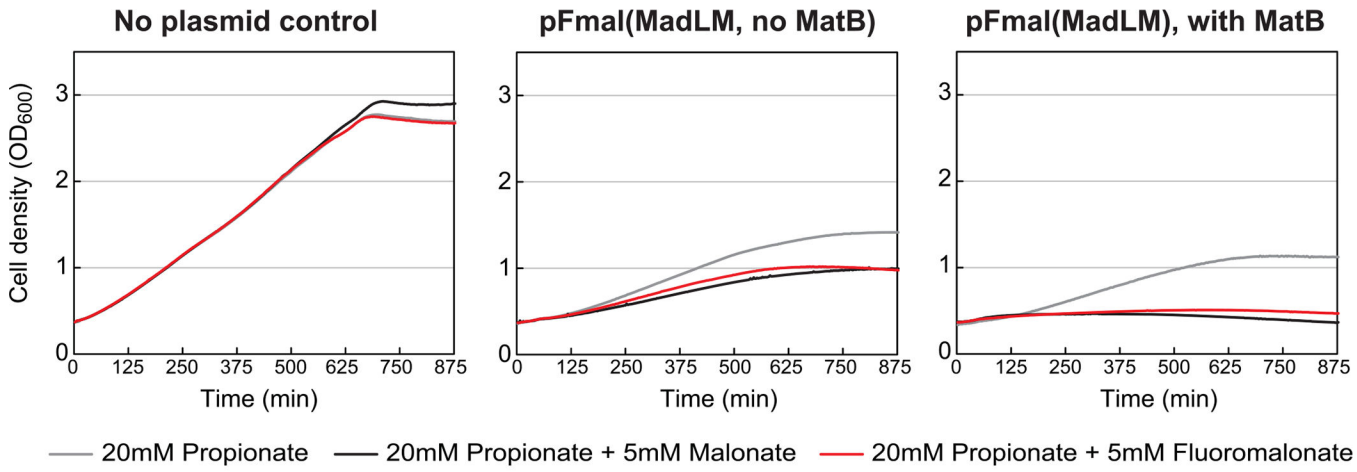


Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{35}O_5^+$	355.2479	355.2484
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.2367
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2261
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	nd
B	$C_{14}H_{25}O_4^+$	257.1747	257.1745
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.1640
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1533
B-3H_2O	$C_{14}H_{19}O^+$	203.1430	203.1434
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{19}O_4^+$	227.1278	227.1269
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.1170
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	191.1069
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	173.0806

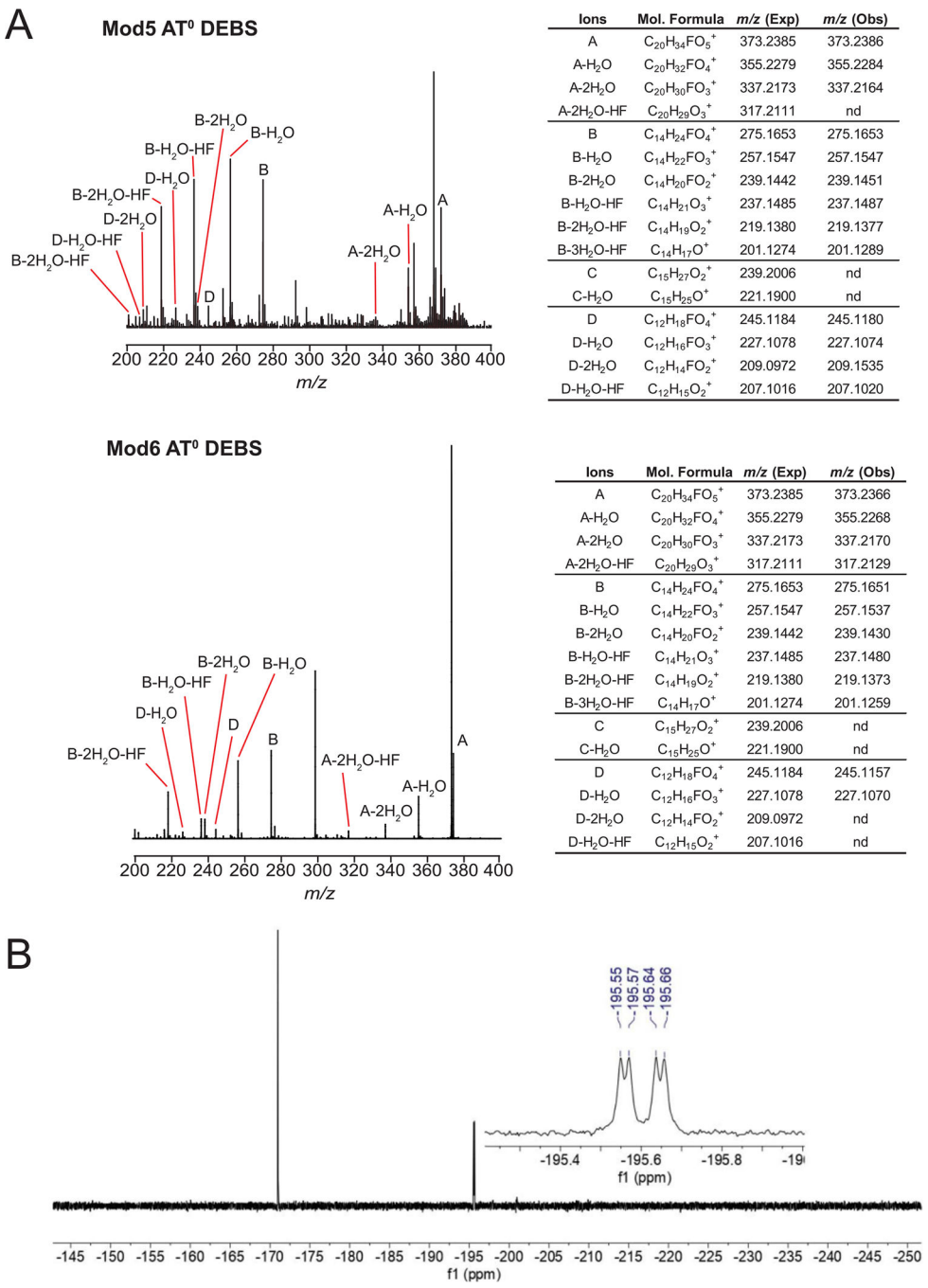
Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{35}O_5^+$	355.2479	355.2472
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.2366
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2256
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	301.2158
B	$C_{14}H_{25}O_4^+$	257.1747	257.1738
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.1632
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1528
B-3H_2O	$C_{14}H_{19}O^+$	203.1430	203.1424
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{19}O_4^+$	227.1278	227.1272
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.1163
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	191.1056
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	nd

Extended Data Fig. 8 |. In vivo production of desmethyl 6dEB analogs by engineered E. coli.
(A) Extracted ion chromatograms showing the production of 6dEB by *E. coli* expressing DEBS, Mod5 AT^0 DEBS, or Mod6 AT^0 DEBS. *E. coli* BAP1 harboring variants of pBP130 and pBP144 were grown to OD_{600} = 0.4–0.6 at 37 °C with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose and cultures provided with 20 mM sodium propionate. Cultures were then incubated with substrates at 22 °C with shaking at 250 rpm for 1 day, after which, culture media were collected, extracted and analyzed. 6dEB product was monitored by LC-QQQ using transition 369.1→239.1.

Chromatograms are representative of at least three biological replicates. **(B)** Extracted ion chromatograms showing the production of desmethyl-6dEB analogs by *E. coli* expressing DEBS, Mod5 AT^0 DEBS, or Mod6 AT^0 DEBS. Samples were prepared as in (A). Desmethyl-6dEB products were monitored by LC-QQQ using transition: 355.2→225.1. Chromatograms are representative of at least three biological replicates. **(C)** Fragmentation patterns of the desmethyl 6dEB analogs produced by *E. coli* harboring Mod5 AT^0 DEBS (top) or Mod6 AT^0 DEBS (bottom). The observed patterns are consistent with those of the desmethyl-6dEB analogs produced by in vitro reconstitution of corresponding enzyme systems. Spectra are representative of at least three biological replicates (nd, not detected).



Extended Data Fig. 9 |. Growth curves of E. coli BAP1 expressing pFmal plasmids.
Various *E. coli* BAP1 strains were grown to OD_{600} ~ 0.4 at 37 °C with shaking at 200 rpm, at which point 100 μL of culture was transferred to 96-well plate and induced with 1 mM IPTG/0.2% (w/v) arabinose. The no plasmid control strain (left), strain expressing only the MadLM transporter (middle) and strain expressing the MadLM transporter and MatB (right) were supplemented with 20 mM sodium propionate, 5 mM malonate or 5 mM fluoromalonate as noted in the legend. Cell growth was monitored at OD_{600} using a microplate reader. Data are representative of 3 biological replicates.



Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2386
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2284
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2164
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	nd
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1653
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1547
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.1451
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1487
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.1380	219.1377
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1289
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1180
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1074
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	209.1535
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	207.1020

Ions	Mol. Formula	*m/z* (Exp)	*m/z* (Obs)
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2366
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2268
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2170
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	317.2129
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1651
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1537
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.1430
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1480
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.1380	219.1373
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1259
C	$C_{15}H_{27}O_2^+$	239.2006	nd
C-H_2O	$C_{15}H_{25}O^+$	221.1900	nd
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1157
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1070
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	nd
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	nd

Extended Data Fig. 10 |. In vivo production of monofluorinated desmethyl 6dEB analogs by engineered E. coli.

(A) Fragmentation pattern and tabulated masses of daughter ions of monofluorinated desmethyl 6dEB analog produced by *E. coli* expressing Mod5 AT^0 DEBS (top)or Mod6 AT^0 DEBS (bottom). The observed patterns are consistent with those of the monofluorinated desmethyl-6dEB analogs produced by in vitro reconstitution of corresponding enzyme systems. Spectra are representative of at least three biological replicates. **(B)** ^{19}F-NMR spectrum of 2-fluoro-2-desmethyl 6dEB isolated from culture media extract of *E. coli*

expressing Mod6 AT^0 DEBS, MatB and MadLM. Concentrated cell suspension was provided with 5 mM fluoromalonate and 20 mM propionate and incubated at 22 °C with shaking at 250 rpm. After 24 h incubation, culture medium was extracted with ethyl acetate, dried via rotary evaporation. 2-fluoro-2-desmethyl 6dEB was then purified from ethyl acetate extract of culture medium through HPLC and fractions were screened using LC-QQQ. Fractions showing presence of 2-fluoro-2-desmethyl 6dEB with transition 373.2 → 275.1 were combined, lyophilized and resuspended in 500 μL of 50:50 MeOH:D_2O mixture for ^{19}F-NMR analysis. Spectrum was collected on Bruker AV600 with following parameters: o1p = −200, sw = 200, d1 = 1 s, d0 −20, n = 3200. 5-fluorouracil (1 mM) was used as internal standard. Spectrum reveals a set of signals between −195 and −196 ppm displaying a doublet of doublet splitting pattern (J = 54, 12 Hz) consistent with the β-hydroxy-α-fluoro-carbonyl motif expected of 2-fluoro-2-desmethyl 6dEB. The observed chemical shift value is similar to that observed of other compounds with similar α-fluoro-β-hydroxy ester motif[55]. Based on vicinal coupling constant ($^3J_{F\text{-}H}$) and known (S)-orientation of hydroxy group on carbon 3, the observed molecule is assigned as 2-(*R*)-fluoro-2-desmethyl 6dEB.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the Pfeifer laboratory at University at Buffalo, The State University of New York for pBP130 and pBP144 plasmids. We thank J. Fang for helpful discussions as well as the C. Chang, Zhang and Francis laboratories at UCB for use of equipment. This work was supported by grant nos. NIH 1 R01 GM123181–01 (M.C.Y.C.) and NIH 1 R35 GM141799 (C.K.). S.S. and O.A. acknowledge the support of a National Institutes of Health NRSA Training grant 1 T32 GM066698. H.D. acknowledges the UC Berkeley Tang Distinguished Scholars Program for a postdoctoral fellowship. Instruments in the UC Berkeley College of Chemistry NMR Facility were supported in part by grant no. NIH S10OD024998.

Data availability

Source data are provided with this paper. All data associated with this study are contained in the published article or are available from the corresponding author upon reasonable request.

References

1. Clardy J & Walsh C Lessons from natural molecules. Nature 432, 829–837 (2004). [PubMed: 15602548]
2. Karageorgis G, Foley DJ, Laraia L & Waldmann H Principle and design of pseudo-natural products. Nat. Chem. 12, 227–235 (2020). [PubMed: 32015480]
3. Rodrigues T, Reker D, Schneider P & Schneider G Counting on natural products for drug design. Nat. Chem. 8, 531–541 (2016). [PubMed: 27219696]
4. Luo X et al. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. Nature 567, 123–126 (2019). [PubMed: 30814733]
5. Stratton CF, Newman DJ & Tan DS Cheminformatic comparison of approved drugs from natural product versus synthetic origins. Bioorg. Med. Chem. Lett. 25, 4802–4807 (2015). [PubMed: 26254944]
6. Ertl P & Schuhmann T A systematic cheminformatics analysis of functional groups occurring in natural products. J. Nat. Prod. 82, 1258–1263 (2019). [PubMed: 30933507]

7. Müller K, Faeh C & Diederich F Fluorine in pharmaceuticals: looking beyond intuition. Science 317, 1881–1886 (2007). [PubMed: 17901324]
8. Wang J et al. Fluorine in pharmaceutical industry: fluorine-containing drugs introduced to the market in the last decade (2001–2011). Chem. Rev. 114, 2432–2506 (2014). [PubMed: 24299176]
9. O'Hagan D & Deng H Enzymatic fluorination and biotechnological developments of the fluorinase. Chem. Rev. 115, 634–649 (2015). [PubMed: 25253234]
10. Purser S, Moore PR, Swallow S & Gouverneur V Fluorine in medicinal chemistry. Chem. Soc. Rev. 37, 320–330 (2008). [PubMed: 18197348]
11. Cheng Q & Ritter T New directions in C–H fluorination. TRECHEM 1, 461–470 (2019).
12. Hull KL, Anani WQ & Sanford MS Palladium-catalyzed fluorination of carbon-hydrogen bonds. J. Am. Chem. Soc. 128, 7134–7135 (2006). [PubMed: 16734446]
13. Cho EJ et al. The palladium-catalyzed trifluoromethylation of aryl chlorides. Science 328, 1679–1681 (2010). [PubMed: 20576888]
14. Scattolin T, Bouayad-Gervais S & Schoenebeck F Straightforward access to *N*-trifluoromethyl amides, carbamates, thiocarbamates and ureas. Nature 573, 102–107 (2019). [PubMed: 31485055]
15. Roque JB, Kuroda Y, Göttemann LT & Sarpong R Deconstructive fluorination of cyclic amines by carbon-carbon cleavage. Science 361, 171–174 (2018). [PubMed: 30002251]
16. Planas O, Wang F, Leutzsch M & Cornella J Fluorination of arylboronic esters enabled by bismuth redox catalysis. Science 367, 313–317 (2020). [PubMed: 31949081]
17. Eustáquio AS, O'Hagan D & Moore BS Engineering fluorometabolite production: fluorinase expression in *Salinispora tropica* yields fluorosalinosporamide. J. Nat. Prod. 73, 378–382 (2010). [PubMed: 20085308]
18. Newman DJ & Cragg GM Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J. Nat. Prod. 83, 770–803 (2020). [PubMed: 32162523]
19. Peralta-Yahya PP, Zhang F, del Cardayre SB & Keasling JD Microbial engineering for the production of advanced biofuels. Nature 488, 320–328 (2012). [PubMed: 22895337]
20. Choi YJ & Lee SY Microbial production of short-chain alkanes. Nature 502, 571–574 (2013). [PubMed: 24077097]
21. Hug JJ, Krug D & Müller R Bacteria as genetically programmable producers of bioactive natural products. Nat. Rev. Chem. 4, 172–193 (2020).
22. Ro DK et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440, 940–943 (2006). [PubMed: 16612385]
23. Niquille DL et al. Nonribosomal biosynthesis of backbone-modified peptides. Nat. Chem. 10, 282–287 (2018). [PubMed: 29461527]
24. Hertweck C The biosynthetic logic of polyketide diversity. Angew. Chem. Int. Ed. Engl. 48, 4688–4716 (2009). [PubMed: 19514004]
25. Staunton J & Weissman KJ Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. 18, 380–416 (2001). [PubMed: 11548049]
26. Wilson MC & Moore BS Beyond ethylmalonyl-CoA: the functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity. Nat. Prod. Rep. 29, 72–86 (2012). [PubMed: 22124767]
27. Ray L & Moore BS Recent advances in the biosynthesis of unusual polyketide synthase substrates. Nat. Prod. Rep. 33, 150–161 (2016). [PubMed: 26571143]
28. Walker MC et al. Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. Science 341, 1089–1094 (2013). [PubMed: 24009388]
29. Dunn BJ, Watts KR, Robbins T, Cane DE & Khosla C Comparative analysis of the substrate specificity of *trans*- versus *cis*-acyltransferases of assembly line polyketide synthases. Biochemistry 53, 3796–3806 (2014). [PubMed: 24871074]
30. Musiol-Kroll EM et al. Polyketide bioderivatization using the promiscuous acyltransferase KirCII. ACS Synth. Biol. 6, 421–427 (2017). [PubMed: 28206741]
31. Carvalho R, Reid R, Viswanathan N, Gramajo H & Julien B The biosynthetic genes for disorazoles, potent cytotoxic compounds that disrupt microtubule formation. Gene 359, 91–98 (2005). [PubMed: 16084035]

32. Khosla C, Tang Y, Chen AY, Schnarr NA & Cane DE Structure and mechanism of the 6-deoxyerythronolide B synthase. Ann. Rev. Biochem. 76, 195–221 (2007). [PubMed: 17328673]
33. Zha W, Rubin-Pitel SB, Shao Z & Zhao H Improving cellular malonyl-CoA level in *Escherichia coli* via metabolic engineering. Metab. Eng. 11, 192–198 (2009). [PubMed: 19558964]
34. Mathews II et al. The conformational flexibility of the acyltransferase from the disorazole polyketide synthase is revealed by an X-ray free-electron laser using a room-temperature sample delivery method for serial crystallography. Biochemistry 56, 4751–4756 (2017). [PubMed: 28832129]
35. Yadav G, Gokhale RS & Mohanty D Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. J. Mol. Biol. 328, 335–363 (2003). [PubMed: 12691745]
36. Starcevic A et al. ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. Nucleic Acids Res. 36, 6882–6892 (2008). [PubMed: 18978015]
37. Gokhale RS, Tsuji SY, Cane DE & Khosla C Dissecting and exploiting intermodular communication in polyketide synthases. Science 284, 482–485 (1999). [PubMed: 10205055]
38. Lowry B et al. In vitro reconstituÿtion and analysis of the 6-deoxyerythronolide B synthase. J. Am. Chem. Soc. 135, 16809–16812 (2013). [PubMed: 24161212]
39. Ashley GW & Carney JR API-mass spectrometry of polyketides. II. Fragmentation analysis of 6-deoxyerythronolide B analogs. J. Antibiot. 57, 579–589 (2004).
40. Wong FT, Chen AY, Cane DE & Khosla C Protein–protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. Biochemistry 49, 95–102 (2010). [PubMed: 19921859]
41. Thuronyi BW, Privalsky TM & Chang MCY Engineered fluorine metabolism and fluoropolymer production in living cells. Angew. Chem. Int. Ed. Engl. 56, 13637–13640 (2017). [PubMed: 28861937]
42. Ad O, Thuronyi BW & Chang MCY Elucidating the mechanism of fluorinated extender unit loading for improved production of fluorine-containing polyketides. Proc. Natl Acad. Sci. USA 114, E660–E668 (2017). [PubMed: 28096394]
43. Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE & Khosla C Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. Science 291, 1790–1792 (2001). [PubMed: 11230695]
44. Dong H, Liffland S, Hillmyer MA & Chang MCY Engineering in vivo production of à-branched polyesters. J. Am. Chem. Soc. 141, 16877–16883 (2019). [PubMed: 31547647]
45. Gibson DG et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345 (2009). [PubMed: 19363495]
46. Bonnett S et al. Acyl-CoA subunit selectivity in the pikromycin polyketide synthase PikAIV: steady-state kinetics and active-site occupancy analysis by FTICR-MS. Chem. Biol. 18, 1075–1081 (2011). [PubMed: 21944746]
47. Piasecki SK et al. Employing modular polyketide synthase ketoreductases as biocatalysts in the preparative chemoenzymatic syntheses of diketide chiral building blocks. Chem. Biol. 18, 1331–1340 (2011). [PubMed: 22035802]
48. Ushimaru K, Sangiambut S, Thomson N, Sivaniah E & Tsuge T New insights into activation and substrate recognition of polyhydroxyalkanoate synthase from *Ralstonia eutropha*. Appl. Microbiol. Biotechnol. 97, 1175–1182 (2013). [PubMed: 22543354]
49. Jia K, Cao R, Hua DH & Li P Study of Class I and Class III polyhydroxyalkanoate (PHA) synthases with substrates containing a modified side chain. Biomacromolecules 17, 1477–1485 (2016). [PubMed: 26974339]
50. Wodzinska J et al. Polyhydroxybutyrate synthase: evidence for covalent catalysis. J. Am. Chem. Soc. 118, 6319–6320 (1996).
51. Molnos J, Gardiner R, Dale GE & Lange R A continuous coupled enzyme assay for bacterial malonyl-CoA:acyl carrier protein transacylase (FabD). Anal. Biochem. 319, 171–176 (2003). [PubMed: 12842120]

52. Dunn BJ, Cane DE & Khosla C Mechanism and specificity of an acyltransferase domain from a modular polyketide synthase. Biochemistry 52, 1839–1841 (2013). [PubMed: 23452124]

53. Pistorino M & Pfeifer BA Efficient experimental design and micro-scale medium enhancement of 6-deoxyerythronolide B production through *Escherichia coli*. Biotechnol. Prog. 25, 1364–1371 (2009). [PubMed: 19634176]

54. Kosol S, Jenner M, Lewandowski JR & Challis GL Protein–protein interactions in trans-AT polyketide synthases. Nat. Prod. Rep. 35, 1097–1109 (2018). [PubMed: 30280735]

55. Ocampo R, Dolbier WR, Abboud KA & Zuluaga F Catalyzed Reformatsky reactions with ethyl bromofluoroacetate for the synthesis of α-fluoro-β-hydroxy acids. J. Org. Chem. 67, 72–78 (2002). [PubMed: 11777441]

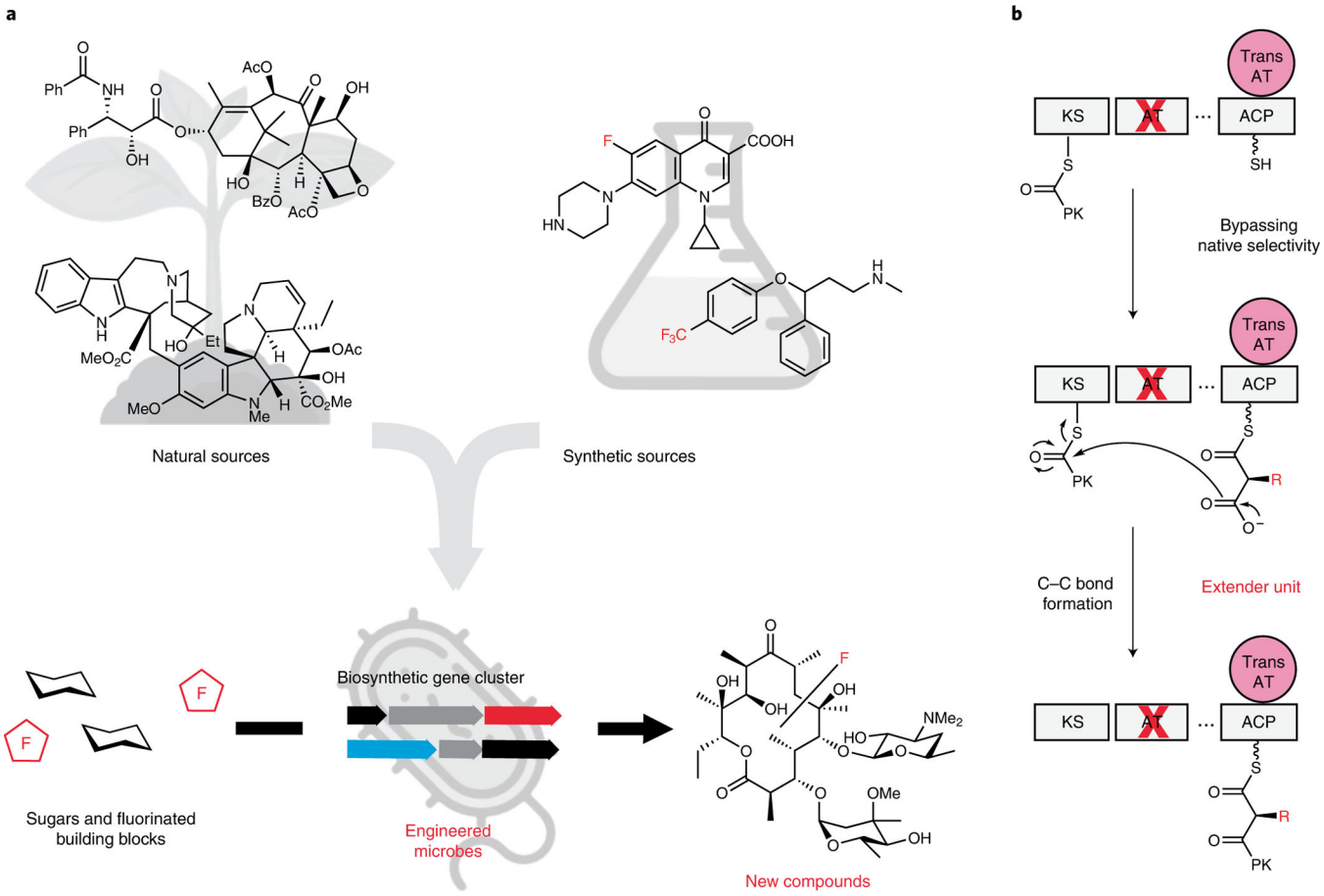


Fig. 1 |. Merging biological and chemical strategies to expand the scope of fluorinated small molecule production.

a, By merging the ability of living systems to produce complex molecular architecture with design elements from synthetic chemistry we can expand the accessible scope of molecular space for discovery of new function. **b**, Insertion of a nonnative extender unit can be achieved by inactivating the *cis*-AT within a modular PKS and complementation in with a standalone *trans*-AT. This allows the native selectivity of the *cis*-AT, which is the dominant domain involved in selecting the identity of the extender unit, to be bypassed.

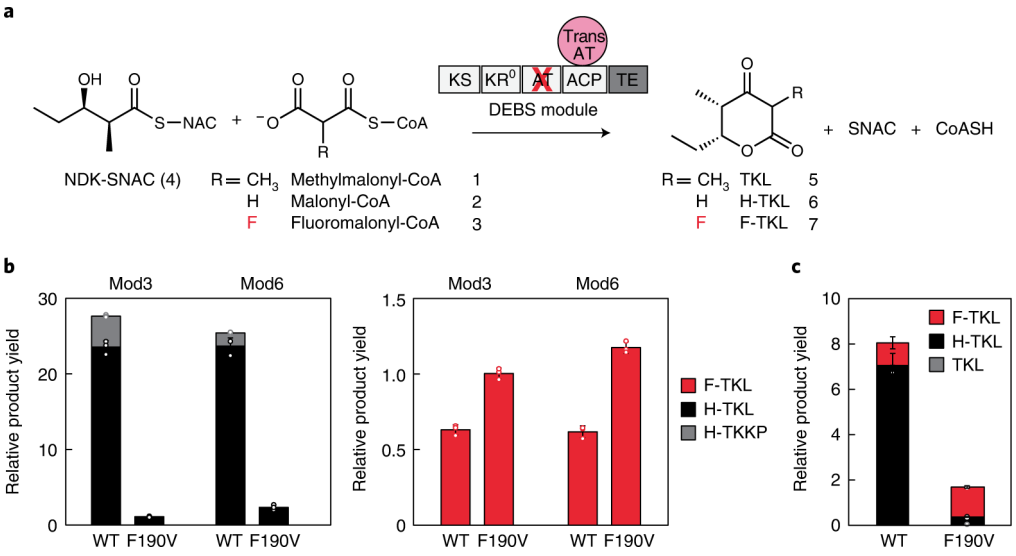


Fig. 2 |. Engineering of fluoromalonyl-CoA-specific trans-AT.

a, Formation of TKLs from *N*-acetylcysteamine thioester of NDK-SNAC and carboxyacyl-CoA extender unit catalyzed by single-modular PKS construct. **b**, In vitro TKL assay comparing WT and F190V DszAT under pure malonyl-CoA (left) and fluoromalonyl-CoA (right) extender unit conditions. Normalized representation of the amount of the products arising from malonyl-CoA, H-TKL (single extension, black) and H-TTKP (double extension, gray) and from fluoromalonyl-CoA, F-TKL (red), are shown. Data were normalized to the amount of product (left H-TKL; right F-TKL) produced by Mod3 DEBS + TE(AT^0) and F190V DszAT. Data are mean ± s.d. of three technical replicates. **c**, In vitro TKL assay comparing WT and F190V DszAT under mixed carboxyacyl-CoA extender unit condition. Normalized representation of the amount of H-TKL (black, malonyl-CoA product), TKL (gray, methylmalonyl-CoA product) and F-TKL (red, fluoromalonyl-CoA product) are shown. Data were normalized to the amount of F-TKL produced by WT DszAT. Data are mean ± s.d. of three replicates.

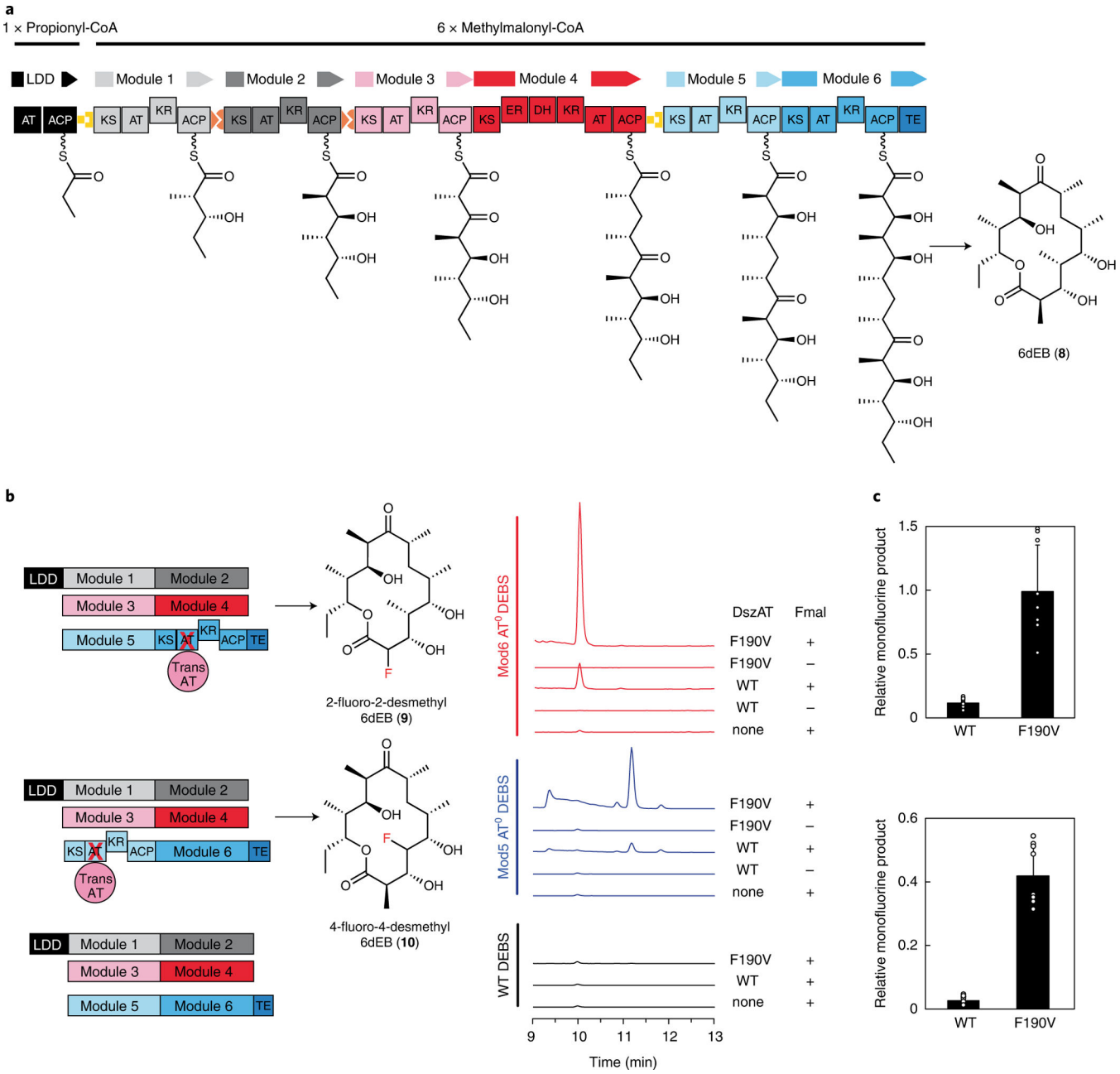


Fig. 3 |. In vitro generation of regioselectively fluorinated 6dEB analog.

a, The DEBS produces the aglycone precursor of the antibiotic erythromycin and consists of a loading didomain and six modules, assembling one propionyl-CoA and six methylmalonyl-CoA molecules into the 6dEB product consisting of 21 carbons and 10 stereocenters. **b**, The fluorine substituent was incorporated regioselectively to produce 2-fluoro-2-desmethyl 6dEB and 4-fluoro-4-desmethyl 6dEB analogs by in vitro reconstitution of Mod6 AT^0 DEBS and Mod5 AT^0 DEBS, respectively. Reactions contained 2 mM methylmalonate and, when indicated, 10 mM fluoromalonate. Extracted ion chromatograms shown are representative of at least three technical replicates. **c**, System selectivity of fluoromalonyl-CoA incorporation by Mod6 AT^0 DEBS (top) and Mod5 AT^0 DEBS (bottom)

by WT and F190V DszAT. Relative monofluorine product was calculated from the ratio between integrated extracted ion counts for monofluorinated desmethyl 6dEB analog to 6dEB of each replicate. Data are mean ± s.d. of nine replicates.

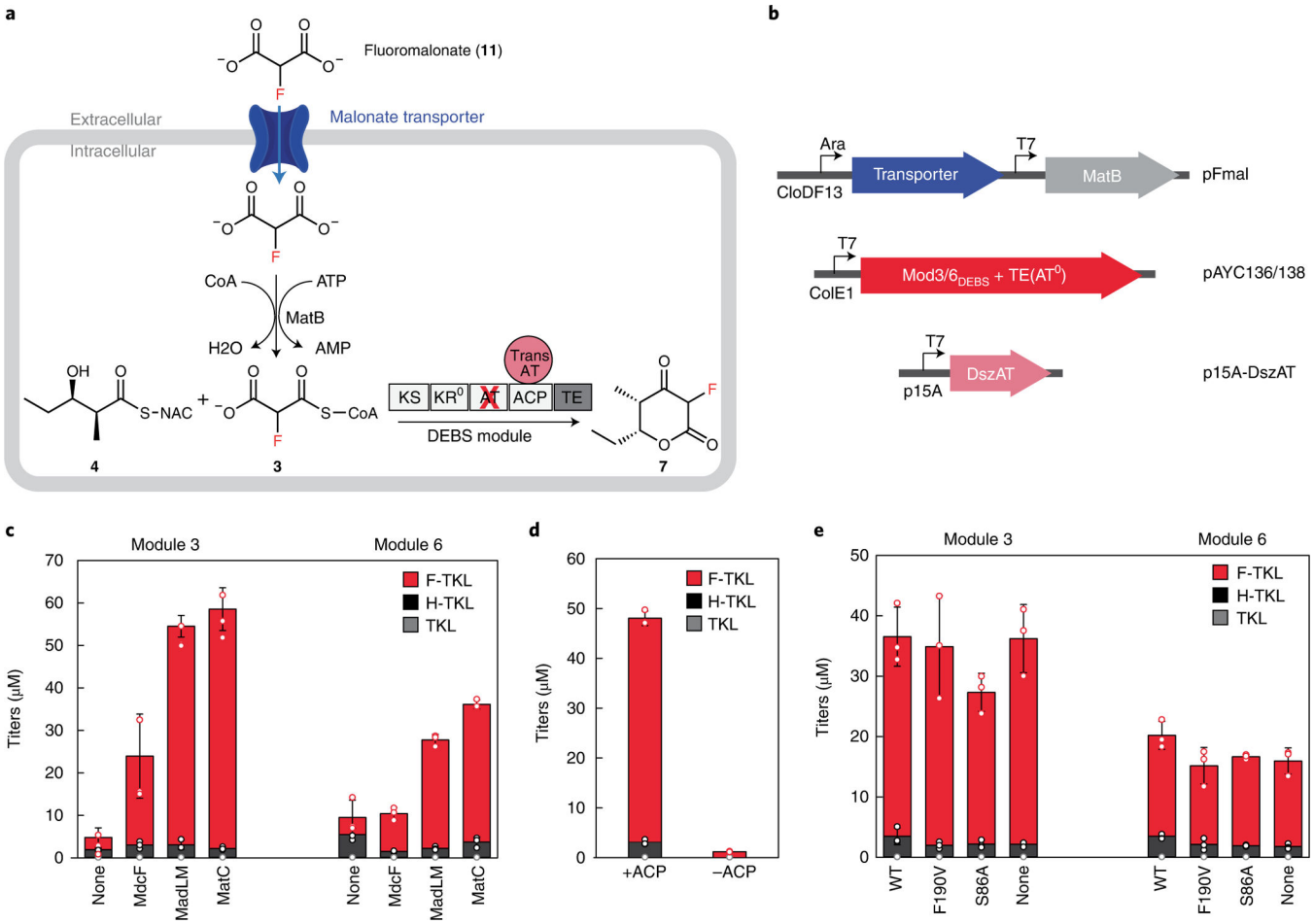


Fig. 4 |. Cellular production of fluorinated polyketide model compounds by engineered E. coli.

a, Complementation strategy for fluorinated polyketide fragments production in *E. coli*. In addition to the PKS and the *trans*-AT, the cellular production system requires a specialized transport system to import fluoromalonate into the intracellular space and activate to the fluoromalonyl-CoA extender unit. Cells natively maintain a high level of ATP, which provides the necessary ATP regeneration. **b**, Engineered pathway for fluorinated TKL (F-TKL) production in *E. coli*. The system includes three plasmids with compatible origins and antibiotic markers, which together encode for four pathway proteins. pFmal (cloDF13 origin, Sp^R) series encode for a malonate transporter and malonyl-CoA synthetase MatB. pAYC136 and 138 (ColE1 origin, Cb^R) encode for $Mod3_{DEBS}$ + TE(AT^0) and $Mod6_{DEBS}$ + TE(AT^0), respectively. p15A-DszAT (p15A origin, Cm^R) encodes for the DszAT variant. **c**, Production titers of TKLs showing the contribution of malonate transporters (*Klebsiella pneumoniae* MdcF, *Pseudomonas fluorescens* MadLM and *Streptomyces coelicolor* MatC) when NDK-SNAC and fluoromalonate were provided to production cultures. Data are mean ± s.d. of three biological replicates. **d**, Production titers of TKLs by *E. coli* expressing pathway with or without an active ACP domain showing the requirement for the domain in F-TKL production. Data are mean ± s.d. of three biological replicates. **e**, Production titers of TKLs by *E. coli* expressing pathway with an active (WT and F190V) or an inactive (S86A and none) *trans*-AT. Data are mean ± s.d. of three biological replicates.

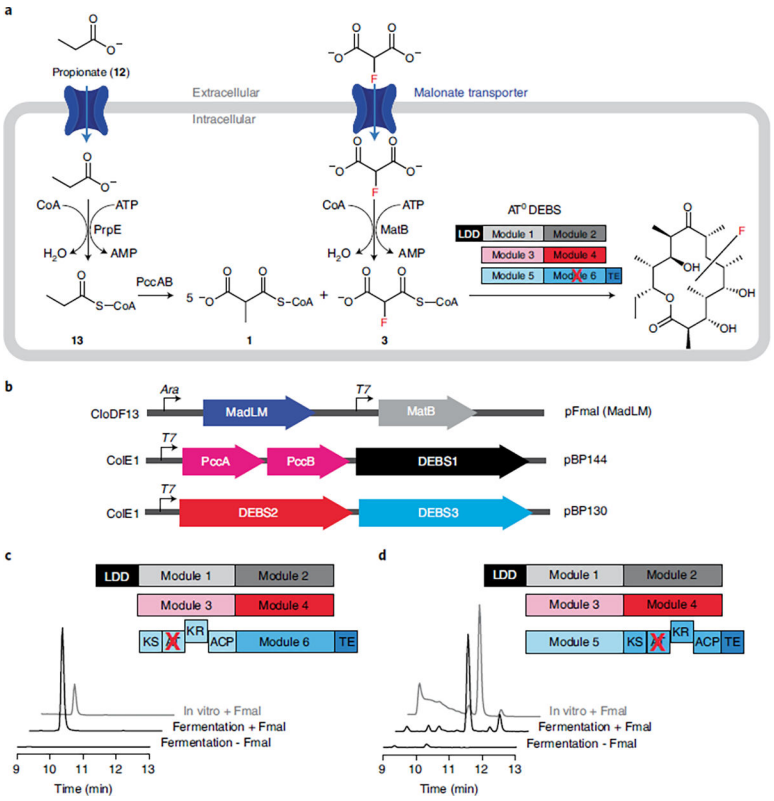


Fig. 5 |. Regiospecific incorporation of fluorine into complex polyketides by E. coli production host.

a, Cellular system for production of fluorinated 6dEB analogs in *E. coli*. In addition to DEBS, the cellular production system requires a transport and activation system to accumulate propionyl-CoA starter unit and methylmalonyl-CoA and fluoromalonyl-CoA extender units intracellularly. Propionate is activated to propionyl-CoA by PrpE incorporated in the genome of BAP1 strain. Propionyl-CoA is then converted to methylmalonyl-CoA by propionyl-CoA carboxylase PccAB, encoded by pBP144. Fluoromalonate is transported into the intracellular space by malonate transporter MadLM and activated to fluoromalonyl-CoA by MatB. Cells natively maintain a high level of ATP, which provides the necessary ATP regeneration. **b**, Engineered pathway for production of fluorinated 6dEB analogs in *E. coli*. The system includes three plasmids with compatible origins and antibiotic markers, which together encode for six pathway proteins. pFmal(MadLM) plasmid (cloDF13 origin, Sp^R) encodes for malonate transporter MadLM, and malonyl-CoA synthetase MatB. pBP130 (ColE1 origin, Cb^R) and pBP144 (ColE1 origin, Km^R) encode for DEBS1, DEBS2 and DEBS3 proteins. pBP144 also encodes for a propionyl-CoA carboxylase (PccAB). **c**, Extracted ion chromatograms showing the generation of the 4-fluoro-4-desmethyl 6dEB analog (*m/z* 373.2 → 275.1) in culture medium of *E. coli* expressing Mod5 AT^0 DEBS. Chromatograms shown are representative of at least three biological replicates. **d**, Extracted ion chromatograms showing the generation of 2-fluoro-2-desmethyl 6dEB analog (*m/z* 373.2 → 275.1) in culture medium of *E. coli* expressing Mod6 AT^0 DEBS. Chromatograms shown are representative of at least three biological replicates.