Chem Soc Rev

Natural and engineered biosynthesis of fluorinated natural products

Journal:	*Chemical Society Reviews*
Manuscript ID:	CS-REV-01-2014-000027.R1
Article Type:	Review Article
Date Submitted by the Author:	21-Apr-2014
Complete List of Authors:	Walker, Mark; University of Illinois, Urbana-Champaign, Chemistry Chang, Michelle; University of California, Berkeley, Chemistry

Natural and engineered biosynthesis of fluorinated natural products

Mark C. Walker[1,2] and Michelle C. Y. Chang[2*]

[1]*Current address: Department of Chemistry, University of Illinois, Urbana-Champaign, Urbana, IL 61801*

[2]*Departments of Chemistry and Molecular & Cell Biology, University of California, Berkeley, Berkeley, CA 94720-1460*

[*]To whom correspondence should be addressed:
Prof. Michelle C. Y. Chang
University of California, Berkeley
Department of Chemistry, 125 Lewis
Berkeley, CA 94720-1460
Email: mcchang@berkeley.edu
Ph.: 510-642-8545
FAX: 510-642-9863

One-sentence summary

Studying the biosynthesis of naturally-occurring and engineered organofluorine natural products may enable the development of new bioactive small molecule targets.

Abstract

Both natural products and synthetic organofluorines play important roles in the discovery and design of pharmaceuticals. The combination of these two classes of molecules has the potential to be useful in the ongoing search for new bioactive compounds but our ability to produce site-selectively fluorinated natural products remains limited by challenges in compatibility between their high structural complexity and current methods for fluorination. Living systems provide an alternative route to chemical fluorination and could enable the production of organofluorine natural products through synthetic biology approaches. While the identification of biogenic organofluorines has been limited, the study of the native organisms and enzymes that utilize these compounds can help to guide efforts to engineer the incorporation of this unusual element into complex pharmacologically active natural products. This review covers recent advances in understanding both natural and engineered production of organofluorine natural products.

1. Introduction

Both biologically-produced and synthetically-derived small molecules serve as important sources for the development of new pharmaceuticals. For example, analysis of new small-molecule drugs approved between 1981 and 2010 reveals that approximately one-third of these compounds are either natural products or their derivatives.[1] While this observation highlights the value of these complex scaffolds, natural products are also often difficult to sufficiently optimize to generate a drug with the appropriate selectivity and pharmacokinetics because of the synthetic

challenges involved in modifying their structures. As such, natural products are also often used to identify macromolecular targets for design of screens for new synthetic lead compounds.[1] Interestingly, the structures of bioactive compounds derived from natural versus synthetic sources are quite divergent.[2] In general, synthetic compounds tend to contain significantly lower structural complexity because of practical concerns in reducing the number of synthetic steps involved in their manufacture. However, they also have the advantage that their structures are not limited to the set of pharmacophores found in nature. One striking example is the recent surge in the use of fluorine in the design of drugs and diagnostics. Indeed, since the introduction of the first fluorinated pharmaceutical – the natural product analog fludrocortisone – in the 1950s, an increasing percentage of pharmaceuticals now contain at least one fluorine substituent,[3] including several of the top-selling drugs on the market.[4]

The unique elemental properties of fluorine have been particularly effective in the design and optimization of synthetic and semisynthetic compounds.[5-8] In summary, the small size of fluorine allows for sterically conservative substitution of hydrogen or oxygen atoms,[9] thus maintaining the overall size and shape of a molecule while tuning many other properties that are important to generating a drug from an initial lead compound. In some cases, the extreme electronegativity of fluorine is used to alter the pK_a of nearby functional groups and change the overall charge of a molecule at the relevant pH.[3] In addition, fluorine can also participate in dipolar interactions that are not available to hydrogen and therefore be introduced to modify binding of fluorinated compounds to their biological targets.[7] Another property that can be altered is *in vivo* drug distribution, which can be affected by changes in hydrophobicity resulting from fluorination.[3] Finally, fluorine introduction is often used to site-specifically block drug modification by

cytochrome P450s, which serve to metabolize and eliminate xenobiotics, allowing the half-life of drugs to be extended *in vivo* and the conversion to toxic compounds to be limited.[8]

Due to the importance of fluorinated small molecules in the drug pipeline, there has been an ongoing focus on new strategies to site-selectively incorporate fluorine into organic compounds, leading to many advances in the development of synthetic reagents and catalysts.[10, 11] An alternative path is to use synthetic biology approaches to engineer the biosynthetic machinery found in living organisms for fluorine utilization, perhaps enabling the merger of the strengths of natural products and their structural complexity with the advantages of fluorine medicinal chemistry in tuning molecular properties. Towards this goal, the study of biological systems that have evolved native organofluorine metabolism can facilitate the design of enzymes, pathways, and organisms that can handle this unusual element.

2. The discovery of biogenic organofluorines

Despite the success of fluorine in the design of synthetic bioactive compounds, nature appears to have evolved only a limited set of biogenic organofluorines. Strikingly, thousands of biogenic organohalogens have been identified to date,[12] but only ~20 of these compound contain fluorine (*Figures 1 and 2*) rather than chlorine or bromine.[13] One possible explanation for the small number of biologically-produced organofluorines is the limited bioavailability of fluoride. While fluorine is the 13th most common element in the earth's crust, the majority resides in insoluble minerals.[14] As a comparison, the concentration of fluoride in sea water is 1.3 ppm whereas chloride is found at 19,000 ppm.[14] Moreover, aqueous fluoride is surrounded by a very tight shell of hydration that must be stripped away to allow for nucleophilic attack.[14] In addition to challenges in nucleophilic C-F bond formation, the electronegativity of fluorine results in a

very high oxidation potential of fluoride that places it out of range of the electrophilic modes of C-F bond formation that are the common route to enzymatic halogenation.[15]

Fluoroacetate (1) was the first fluorinated natural product to be discovered, when it was identified in 1943 as the poison causing the high toxicity of the leaves of the South African plant *Dichapetalum cymosum*.[13] Following that discovery, several more genera of plants from Africa, Australia, and South America were also shown to contain fluoroacetate at tissue concentrations of up to milligrams per gram dry weight.[13] The pathway for fluoroacetate production in these plants has yet to be elucidated and it remains unclear whether the plants or microbial symbionts are responsible for its biosynthesis. Initial studies on calluses from *D. cymosum*, which have not yet been found to contain culturable microorganisms, have shown that fluoroacetate is still produced in the presence of inorganic fluoride but that its production is inconsistent and diminishes over time.[16] However, analysis of soil collected near several species of fluoroacetate-containing plants indicated the presence of organic fluorine,[17] suggesting that plants may be concentrating environmental organofluorines produced by microbes. This question could benefit from further studies to facilitate comparative genomics of the plants themselves or metagenomic analysis of the microbial communities in the surrounding soil.

Given the structural similarity of fluoroacetate to the common metabolite acetate (2), it can often utilize the same metabolic pathways. For example, the seeds of the plant *Dichapetalum toxicarium* have been observed to contain ω-fluoro-fatty acids (5), which include ω-fluorooleic acid and ω-fluoropalmitic acid and account for the majority of known organofluorines. Further examination of the ω-fluoro-fatty acid pools showed the distribution of their lengths and levels of desaturation were similar to those of non-fluorinated fatty acids (6), suggesting that

fluoroacetate is being used as the starter unit by the normal fatty acid biosynthetic machinery.[18] Interestingly, fluorine has only been observed at the ω-position, indicating that fluorine is not natively incorporated as an extender unit by fatty acid synthases. Endogenous host metabolic pathways are also responsible for fluoroacetate toxicity (*Figure 1*). As with ω-fluoro-fatty acid biosynthesis, fluoroacetate is activated to fluoroacetyl-coenzyme A (CoA) (3). Fluoroacetyl-CoA can be then be condensed with oxaloacetate by citrate synthase to produce (*2R,3R*)-fluorocitrate (7).[19] Upon reaction with the aconitase, fluoride is eliminated and instead of producing fluoroisocitrate, results in 4-hydroxy-transaconitate (12), which effectively inhibits the aconitase irreversibly based on its extremely slow off-rate.[20] In addition to this "lethal synthesis",[21] fluorocitrate binds tightly to citrate transporters, blocking citrate trafficking into and out of mitochondria.[22]

Besides fluoroacetate and its metabolites, only one other biogenic organofluorine has been identified. Nucleocidin, a 4´-fluorodeoxyadenosine derivative, was first isolated from *Streptomyces calvus* in the 1950s. However, for a period of time production of this metabolite could not be repeated and the biosynthetic route for the production of this organofluorine has not yet been elucidated. Recently, production may have been achieved again by the generation of rifamycin-resistant mutants of *S. calvus*.[23] Determination of the biosynthetic pathway for this organofluorine would be informative as it likely diverges from fluoroacetate pathways while maintaining some potential similarities.

3. Organofluorine biosynthesis in *Streptomyces cattleya*

The only host of a characterized carbon-fluorine bond-forming enzyme is the soil-dwelling bacterium, *Streptomyces cattleya* (*Figure 2*). During studies to enhance the production of

thienamycin by *S. cattleya*, fluoride was introduced inadvertently into the media as a contaminant and resulted in the observation of new antibiotics, which were determined to be fluoroacetate and fluorothreonine (18).[24] When cultured with sodium fluoride, *S. cattleya* produces fluoroacetate and fluorothreonine upon entry into stationary phase and continues to do so until most of the available fluoride is consumed.[25] The ratio of fluoroacetate to fluorothreonine has been observed to begin at 1:1 but then reaches approximately 3:1 over time.[25] Given its defined ability to produce organofluorines in cell culture, *S. cattleya* has become the model organism for studying organofluorine metabolism.

3.1. Fluorinase. The key step in organofluorine biosynthesis is formation of the C-F bond. While several possibilities were proposed for this reaction,[14] experiments where *S. cattleya* cell lysates were incubated with various cofactors and inorganic fluoride revealed that addition of (*S*)-adenosyl-L-methionine (SAM) (13) leads to the production of 5′-fluoro-5′-deoxyadenosine (FDA) (14), 5′-fluoro-5′-deoxyinosine (FDI), and fluoroacetate.[26] To further characterize the initial fluorination product, FDA and FDI were incubated individually with *S. cattleya* cell lysates. Only FDA was further processed to fluoroacetate, suggesting that FDA is an intermediate in the organofluorine biosynthetic pathway and that FDI is likely a shunt product arising from adventitious deamination.[26] Identification of the substrates for the C-F bond forming reaction enabled the native purification of a 5′-fluoro-5′-deoxyadenosine synthase, also called the fluorinase, from *S. cattleya*.[27] Edman degradation of the fluorinase yielded its N-terminal sequence and allowed subsequent identification of the gene (*flA*) through PCR amplification from *S. cattleya* genomic DNA with degenerate primers.[28] The fluorinase was then expressed heterologously in *Escherichia coli* for biochemical and structural studies. Co-crystal structures of the fluorinase with either SAM or FDA and methionine were obtained and provide initial insight

into fluorinase-catalyzed C-F bond formation.[28] In the FDA-methionine product structure, there appear to be several short hydrogen bonds between the enzyme and the fluorine atom of FDA. These hydrogen bonds were thus proposed to participate in driving the desolvation of the fluoride ion,[29] allowing it to initiate an S_N2 reaction with the 5′-carbon of SAM[30, 31] to form the C-F bond. Morever, the SAM appears to be bound in a strained conformation in the substrate complex, activating the 5′-carbon for nucleophilic attack. Despite these modes of activation, the measured rate of FDA formation by fluorinase purified from heterologous sources is slow (k_{cat} ~ 0.07 min^{-1}) and the K_M for fluoride is high (~2 mM), indicating the challenges involved in binding and activating fluoride in aqueous media.[28]

3.2. 5′-fluoro-5′-deoxyadenosine phosphorylase. Using the fluorinase gene sequence to design probes, an organofluorine biosynthetic gene cluster was identified from a *S. cattleya* genomic cosmid library.[32] A clustered gene (*flB*) encoding an enzyme homologous to methylthioadenosine phosphorylase from the methionine salvage pathway was found to lie directly upstream of the fluorinase. Biochemical characterization of this enzyme showed it was competent to catalyze the formation of 5-fluoro-5-deoxyribose-1-phosphate (FDR1P) (15) from FDA and inorganic phosphate.[32] Substrate selectivity studies with adenosine and FDA also showed that the phosphorylase demonstrates a preference for the fluorinated substrate over its non-fluorinated congener. A knockout strain of *S. cattleya* lacking this gene was generated and found to be deficient in the production of fluoroacetate and fluorothreonine, further suggesting this enzyme is on the pathway to organofluorine biosynthetic.[33]

3.3. 5-fluoro-5-deoxyribose-1-phosphate isomerase. Although *flB* was found within the biosynthetic gene cluster, it is interesting to note that genes encoding other downstream enzymes

in the organofluorine biosynthetic pathway could not be located in the fluorinase gene cluster. Thus, the next step was proposed to be analogous to a methylthioribose-1-phosphate isomerase based on the common features between the fluoroacetate and methionine salvage pathways. Studies with *S. cattleya* cell lysates were consistent with this hypothesis as 5-fluoro-5-deoxyribulose-1-phosphate (FDRib1P) (16) was found as an intermediate.[34] A conserved methylthioribose-1-phosphate isomerase was then identified from the *S. cattleya* genome by PCR amplification using degenerate primers. Subsequent biochemical characterization of this enzyme showed it could produce FDRib1P from FDR1P.[35] Disruption of the corresponding gene leads to loss of organofluorine production when *S. cattleya* is cultured in defined media,[33] but the knockout strain could still produce fluoroacetate and fluorothreonine when grown in rich media.[36] Examination of the *S. cattleya* genome revealed the presence of an unusual methylthioribose-1-phosphate isomerase fused to a methylthioribulose-1-phosphate dehydratase, which is the next enzyme in the methionine salvage pathway. When a double knockout strain lacking both methylthioribose-1-phosphate isomerases was grown in rich media, no fluoroacetate or fluorothreonine was observed.[36] These results suggest the isomerase step is part of the organofluorine biosynthetic pathway and that either copy of the methylthioribose-1-phosphate isomerase is capable of providing this activity.

3.4. Fluoroacetaldehyde dehydrogenase. To gain further insight into the fluoroacetate pathway, *S. cattleya* was grown on ^{13}C-labeled carbon sources and it was shown that labels incorporated at the C-1 or C-2 positions of fluoroacetate were also incorporated into the analogous C-3 or C-4 positions of fluorothreonine. These results suggest that fluoroacetate and fluorothreonine share a common precursor, which was then found to be fluoroacetaldehyde (17).[37] An enzyme capable of oxidizing fluoroacetaldehyde to fluoroacetate was purified from *S.*

cattleya and subjected to Edman degradation sequencing to determine its N-terminal sequence. [38] Notably, this enzyme is related to chloroacetaldehyde dehydrogenases from dichloroethane degradation pathways.[36]

3.4. Fluorothreonine transaldolase. Labeling studies also were used to show that incubation of *S. cattleya* whole cells with [1-^{2}H]-fluoroacetaldehyde results in the production of fluorothreonine with a significant level of deuterium incorporation at the C-3 position, implying that fluoroacetaldehyde is used directly in fluorothreonine production.[37] A PLP-dependent transaldolase capable of producing fluorothreonine from threonine and fluoroacetaldehyde was isolated from *S. cattleya*[39] and found to be critical for fluorothreonine production in genetic knockout studies.[33] This enzyme is different from typical threonine aldolases, which interconvert glycine and acetaldehyde with threonine, since it carries out a transaldolase reaction instead with threonine itself. Furthermore, the fluorothreonine transaldolase is fused to a second zinc-dependent aldolase domain of unknown function.

3.5. 5-fluoro-5-deoxyribulose-1-phosphate aldolase. The remaining steps in the fluoroacetate and fluorothreonine biosynthetic pathway that link FRib1P to fluoroacetaldehyde have yet to be characterized. The most direct route would be an aldolase capable of producing fluoroacetaldehyde and dihydroxyacetone phosphate (DHAP) directly from FDRib1P. Indeed, a fuculose-1-phosphate aldolase from *Streptomyces coelicolor* has been shown to be able to carry out this reaction.[35] Attempts to purify an enzyme from *S. cattleya* that could carry out the reverse reaction, the condensation of fluoroacetaldehyde and DHAP, yielded an aldolase that produced 5-fluoro-5-deoxyxylulose-1-phosphate,[35] which suggests that an epimerase could act on 5Rib1P prior to cleavage. Another possibility is that the cryptic aldolase domain fused to the fluorothreonine could be involved, but this hypothesis has yet to be experimentally validated.[39]

3.6. Other host organisms. Given that fluoroacetate and its derivatives account for the majority of known biogenic organofluorines, it is possible that a similar pathway is responsible for organofluorine biosynthesis in other organisms. The fluorinase is the most unique enzyme in the organofluorine biosynthetic pathway and a member of the SAM hydroxide adenosyltransferase family of proteins, which also contains SAM hydrolases[40] and a chlorinase involved in salinosporamide A biosynthesis.[41] These enzymes have been identified in bacteria and archaea but have yet to be plants, which may indicate that microbes participate in the production of plant-based fluoroacetate. Recent genome sequencing efforts have expanded the number of members in this family and proteins with >80% sequence identity to the fluorinase have now been identified and organisms harboring these genes have been shown to produce organofluorines.[42]

4. The fluorine physiology of *S. cattleya*

Global analysis of the transcriptional response of *S. cattleya* to fluoride addition was explored using RNA-seq in order to begin probing aspects of fluorine physiology beyond fluoroacetate and fluorothreonine biosynthesis.[36] By examining the relative levels of transcription of the pathway genes at 0.5, 2, 6, and 48 h after the addition of fluoride, it was shown that the fluorinase, the 5′-fluoro-5′-deoxyadenosine phosphorylase, the fluoroacetaldehyde dehydrogenase, and the fluorothreonine transaldolase all share a similar transcriptional pattern despite the lack of clustering of the latter two genes. These genes are upregulated both by presence of fluoride as well as by entry into stationary phase. In fact at 48 h after the addition of fluoride, the fluorinase is the sixth most highly-transcribed gene in the entire genome, which is consistent with an important role in *S. cattleya* physiology as well as its low turnover rate. In contrast, while the genes for the methylthioribose-1-phosphate isomerases and

for two putative fuculose-1-phosphate aldolases are transcribed, they do not share this pattern of fluoride or growth stage regulation.

Similar to other antibiotic hosts, the fluorinase biosynthetic cluster also contains a putative resistance gene, the fluoroacetyl-CoA thioesterase (*flK*).[32] Although *flK* is not essential under these growth conditions,[36] it reverses the lethal synthesis of fluorocitrate by hydrolyzing fluoroacetyl-CoA back to fluoroacetate with a ~10^6-fold selectivity for the fluorinated substrate over the intracellularly-abundant congener, acetyl-CoA.[43, 44] It is interesting to note that this mechanism is quite different from the dehalogenase enzymes that have been used to protect ruminants from fluoroacetate poisoning[45] and preserves the C-F bond. Biochemical analysis of the enzymes involved in host fluoroacetate toxicity (*Figure 1*) indicated that the enzymes from *S. cattleya* do not demonstrate any particular fluorine selectivity above orthologs from organisms that do not biosynthesize organofluorines. Taken together, the biochemical, genetic, and transcriptional profiling data suggested that coordinated regulation of organofluorine biosynthesis and the tricarboxylic acid (TCA) cycle serves as the major mechanism of organofluorine resistance in *S. cattleya*, with fluoroacetate and fluorothreonine biosynthesis initiated as cells approach stationary phase and shut down the TCA cycle.[36]

The initial source of fluoride-based gene regulation remains to be determined. One possibility is that fluoride is recognized by either a transcription factor or riboswitch.[46] On the other hand, the fluorinase is constitutively expressed and basal levels of organofluorines could also be involved in regulation. There are four predicted transcriptional regulators clustered with the fluorinase gene, with one in particular (*flG*) being upregulated in the presence of fluoride. Other changes in the transcriptional landscape in the presence of fluoride are also observed.[36] For example, 5-carbon carbohydrate metabolism is altered when fluoride is added to the growth

media. As the organofluorine biosynthetic pathway includes 5-carbon sugar intermediates, these changes could be involved in controlling the pools of organofluorines. Also notable is the upregulation of enzymes involved in 3-carbon carbohydrate metabolism, which would be a product of the organofluorine biosynthetic pathway if DHAP is produced. Further exploration of the molecular basis for this transcriptional response could provide additional insight into how organisms sense fluoride or organofluorines.

5. Engineering complex organofluorine biosynthesis

While much remains to be learned about natural organofluorine biosynthesis, work towards engineering the production of more complex fluorinated natural products using the biosynthetic machinery of microbes and plants is underway (*Figure 3*). Semisynthetic efforts to chemically modify natural product scaffolds to introduce fluorine have shown that these compounds can exihibit improved pharmacological behavior upon fluorination.[47-50] Thus, the development of orthogonal methods for site-selective fluorine incorporation may allow us to expand the scope of accessible structures for discovery of new bioactive compounds. A major challenge in using biological systems to introduce fluorine is that the organofluorine building block must be able to compete with native building blocks that are already present in the host organism. One successful strategy to overcoming this problem is to disrupt the production of an unusual building block used in the biosynthesis of the target natural product and then to rescue the disruption by providing the host with a fluorinated analog (*Figure 3B*). It is also possible to intercept the biosynthetic pathway using synthetic intermediates with modified enzymes (*Figure 3C*). However, these approaches limit the scope of the structures that can be approached as well as the sites of fluorine incorporation. Thus, another approach is to look to building blocks with universal biosynthetic potential that are broadly used for biosynthesis of natural products. In this

case, the production of these building blocks cannot be completely disrupted, as they are often required for normal cellular function as well as the synthesis of the rest of the target compound. As such, the biosynthetic machinery itself needs to be modified in order to allow for fluorine-selective building block incorporation (*Figure 3D*). Progress towards addressing all these issues has been made in several different classes of natural products, including polyketides, ribosomally- and nonribosomally-synthesized peptides, as well as alkaloids.

5.1. Polyketide natural products. Polyketides comprise a large class of natural products that are made from the iterative condensation and tailoring of acyl-coenzyme A building blocks by large multimodular enzymes or enzyme complexes called polyketide synthases (PKSs).[51] Many members of this family are currently used clinically and continue to be the focus of efforts to optimize their pharmacological properties. One example is rapamycin, which is an immunosuppressant from *Streptomyces hygroscopicus* produced by a type I PKS system. The three PKS enzymes involved in its biosynthesis carry out the assembly-line incorporation and tailoring of 14 extender units beginning with an unusual 4,5-dihydroxycyclohex-1-enecarboxylic acid starter unit.[52] If the biosynthetic pathway to this precursor from shikimic acid is disrupted, *S. hygroscopicus* will only produce rapamycin when provided with exogenous starter units that are accommodated by the system (*Figure 3B*).[53] Thus, the production of rapamycin can be rescued with the addition of fluorohydroxycyclohexane carboxylic acids to generate fluorinated rapamycin analogs (*Figure 4A*).[54]

In addition to the rescue with the starter unit, synthetic intermediates can also be accommodated by PKS systems, which can accept non-native thioesters as substrates. In this case, the biosynthetic machinery itself is modified and the PKS modules responsible for the initial steps in biosynthesis are removed. Production of the final product is then complemented

by addition of a synthetic thioester intermediate to yield the corresponding polyketide (*Figure 3C*). This concept has been illustrated using the 6-deoxyerythronolide B (DEB) synthase that is responsible for the synthesis of DEB, the macrolide precursor to the antibiotic erythromycin.[55] The removal of the loading and first modules of the DEB synthase, which respectively carry out incorporation of the starter unit and the first extender unit, results in no detectable polyketide product with the modified DEBS when expressed heterologously in *Streptomyces coelicolor*. Culturing this host with a synthetic diketide thioester, (*2R**, *3S**)-5-fluoro-3-hydroxy-2-methylpentanoate *N*-propionylcysteamine thioester (fluoro-NDK-SNPC), then allows the activity of the first two modules to be complemented and 6-deoxy-15-fluoroerythronolide to be produced. Furthermore when the truncated DEBS is expressed in the native erythromycin producer, *Saccharopolyspora erythraea*, production of 15-fluoroerythromycin is observed (*Figure 4B*). These results demonstrate that subsequent tailoring enzymes are competent to process the fluorinated intermediates and that the presence of a fluorine substituent in one part of the chain does not prevent downstream elongation or processing steps.

Although PKS systems are known to be promiscuous with regard to their use of starting substrates, it is quite difficult to carry out chain extension reactions with non-native extender units. However, this approach is important for the modification of positions that would be difficult to access through synthetic intermediates. Work with FK506 provides an example a system where the presence of an unusual extender unit enables the incorporation of fluorine into a polyketide natural product (*Figure 3B*).[49] FK506 is a macrolide with immunosuppressant activity similar to rapamycin and its biosynthesis requires an allylmalonyl-CoA extender unit. In contrast to modules that utilize the universal malonyl-CoA and methylmalonyl-CoA extenders, ones involved in the incorporation of more exotic extenders seem to demonstrate relaxed

substrate specificity. Thus, a variety of 4- to 6-carbon carboxylic acids can be fed to the host organism, *Streptomyces* sp. KCTC 11604BP, when allylmalonyl-CoA biosynthesis is disrupted and be used to replace allylmalonyl-CoA. Among these substrates is 4-fluorocrotonate, which is converted to fluoroethylmalonyl-CoA by the host metabolism and utilized by the PKS to produce fluoro-FK506 (*Figure 4C*).[49]

Another interesting example that highlights the plasticity of the PKS modules encoding unusual extenders is the polyketide/nonribosomal peptide, salinosporamide A, produced by the marine actinomycete *Salinospora tropica*. This natural product is synthesized using a unique chloroethylmalonyl-CoA extender unit. Interestingly, the enzyme that is responsible for the chlorination is the closest biochemically-characterized relative of the fluorinase from *S. cattleya*.[41] The chlorinase produces 5′-chloro-5′-deoxyadenosine from SAM and chloride, which yields 4-chlorocrotonyl-CoA as the product of the pathway and chloroethylmalonyl-CoA upon reductive carboxylation. When the gene encoding the chlorinase in *S. tropica* was replaced in the chromosome with the gene for the fluorinase, culturing this strain of *S. tropica* with fluoride leads to the production of fluorosalinosporamide (*Figure 4D*).[56] The production of fluorosalinosporamide is the only example to date of a complex fluorinated natural product being biosynthesized from fluoride but shows that such a transformation is possible.

Perhaps the most general route to site-selective fluorination of polyketide natural products would be to engineer the biosynthetic machinery to accept fluorinated versions of malonyl-CoA and methylmalonyl-CoA, as the vast majority of polyketides characterized to date use only one or both of these two extenders units (*Figure 3D*).[51] Initial work towards demonstrating the potential of this approach has recently been reported using fluoromalonyl-CoA as an example (*Figure 4E*).[57] The production of fluoromalonyl-CoA can be engineered from either fluoroacetate

or fluoromalonate and used for *in vitro* chain extension reactions using truncated DEBS model systems to produce fluorinated polyketide products. Regioselective fluorine incorporation could then be achieved by inactivating the acyltransferase domain through an active site mutation in the module responsible for chain extension at the desired site for fluorination and then rescuing this activity with a trans-acting acyltransferase (*Figure 4E*), although the specific mechanism through which this rescue is achieved remains to de determined. In addition, *in vivo* production of fluorinated polyketides was also demonstrated in *Escherichia coli* by taking advantage of the fact that it does not have a large endogenous pool of competing methylmalonyl-CoA.[57] This overall strategy could allow for the replacement of an individual malonyl-CoA or methylmalonyl-CoA extender unit with a fluoromalony-CoA unit, greatly expanding the scope of target structures accessible from biological systems.

5.2. Peptide natural products. In addition to fluorinated polyketides, the production of nonribosomally- and ribosomally-synthesized peptides can also engineered by manipulating precursor supply (*Figure 3B*). The nonribosomal peptide antibiotic balhimycin is produced by *Amycolatopsis mediterranei* and contains two β-hydroxytyrosine extender units. Since it is non-essential, the gene responsible for hydroxylation of tyrosine at the β-position can be disrupted to inhibit production of balhimycin. Upon feeding the organism with 2-fluoro-β-hydroxytyrosine or 3,5-difluoro-β-hydroxytyrosine, the corresponding fluorinated balhimycins are produced (*Figure 5A*).[58]

Since peptide natural products can also be derived from ribosomally-synthesized precursors, techniques to incorporate noncanonical amino acids into proteins can be applied to their production. The biosynthetic machinery involved in the production of the lantibiotic lichenicidin

was heterologously expressed in an *Escherichia coli* strain deficient in tryptophan biosynthesis. When the culture was supplemented with 4-fluorotryptophan, incorporation of the fluorinated amino acid is observed in place of tryptophan (*Figure 5B*).[59] In this regard, site-selective methods for non-canonical amino acid insertion can also be applied to these systems.[60]

5.3. Alkaloids. The production of novel fluorinated alkaloids has been successfully engineered in plants by modifying the biosynthetic pathways involved in precursor supply (*Figure 3B*). In this regard, tryptamine is the starter unit for a large number of monoterpene alkaloids produced by the plant *Catharanthus roseus* and is produced via the decarboxylation of tryptophan. The production of tryptamine and associated alkaloids are dramatically reduced by knocking down tryptophan decarboxylase expression using RNA silencing.[61] These plants deficient in tryptophan decarboxylase activity can then be fed 5-fluorotryptamine, resulting in the production of fluorinated analogs of downstream alkaloids (*Figure 5C*). The fluorinated precursors are accommodated at different levels downstream, but some of the fluorinated analogs, such as fluoroserpentine and fluorotabersonine, can be produced at similar levels to the non-fluorinated congeners in the wild-type plant. These results show that although the majority of natural product engineering has taken place in bacteria, it can also be achieved in plants, which are a key source of bioactive natural products.

6. Conclusions

Since first being identified in the 1940s, a fairly small number of biogenic organofluorines have been identified. The rarity of these compounds is somewhat surprising given the ever-increasing number of synthetic bioactive organofluorines. Studies of the biosynthesis of these compounds were initially hampered by limited amounts of available biomass and inconsistent production by the plants in which they were identified. However, discovery of the

organofluorine-producing bacterium, *S. cattleya*, has allowed for the elucidation of a biosynthetic pathway for fluoroacetate and fluorothreonine and initial progress into understanding its fluorine physiology.

Along with the study of natural systems, several routes for engineered incorporation of fluorine into pharmacologically important classes of natural products have been developed. While the many of these engineered fluorinated natural products have been found to date to have similar or lower bioactivity than their non-fluorinated analogs, the development of these strategies will allow for the application of medicinal chemistry type lead optimization on natural products with promising biological activities in years to come.

Acknowledgements

Research in our laboratory in this area is funded by the generous support of the University of California, Berkeley and the National Institutes of Health (1 DP2 OD008696).

References

1. D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2012, **75**, 311-335.
2. T. Henkel, R. M. Brunne, H. Müller and F. Reichel, *Angew. Chem. Int. Ed.*, 1999, **38**, 643-647.
3. W. K. Hagmann, *J. Med. Chem.*, 2008, **51**, 4359-4369.
4. D. O'Hagan, *J. Fluorine Chem.*, 2010, **131**, 1071-1081.
5. B. K. Park, N. R. Kitteringham and P. M. O'Neill, *Annu. Rev. Pharmacol. Toxicol.*, 2001, **41**, 443-470.
6. H.-J. Böhm, D. Banner, S. Bendels, M. Kansy, B. Kuhn, K. Müller, U. Obst-Sander and M. Stahl, *Chembiochem*, 2004, **5**, 637-643.
7. K. Müller, C. Faeh and F. Diederich, *Science*, 2007, **317**, 1881-1886.
8. S. Purser, P. R. Moore, S. Swallow and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320-330.
9. A. Bondi, *J. Phys. Chem.*, 1964, **68**, 441-451.
10. J.-A. Ma and D. Cahard, *Chem. Rev.*, 2008, **108**, PR1-PR43.

11. T. Liang, C. N. Neumann and T. Ritter, *Angew. Chem. Int. Ed.*, 2013, **52**, 8214-8264.
12. G. W. Gribble, *Naturally occurring organohalogen compounds - A comprehensive update*, Springer, 2009.
13. D. O'Hagan and D. B. Harper, *J. Fluorine Chem.*, 1999, **100**, 127-133.
14. D. B. Harper and D. O'Hagan, *Nat. Prod. Rep.*, 1994, **11**, 123-133.
15. C. S. Neumann, D. G. Fujimori and C. T. Walsh, *Chem. Biol.*, 2008, **15**, 99-109.
16. N. Grobbelaar and J. J. M. Meyer, *J. Plant Physiol.*, 1990, **135**, 550-553.
17. R. J. Hall and R. B. Cain, *New Phytol.*, 1972, **71**, 839-853.
18. J. T. G. Hamilton and D. B. Harper, *Phytochemistry*, 1997, **44**, 1129-1132.
19. D. D. Clarke, *Neurochem. Res.*, 1991, **16**, 1055-1058.
20. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, **93**, 13699-13703.
21. R. A. Peters, *Proc. R. Soc. London, B*, 1952.
22. N. V. Goncharov, R. O. Jenkins and A. S. Radilov, *J. Appl. Toxicol.*, 2006, **26**, 148-161.
23. K. Fukuda, T. Tamura, Y. Segawa, Y. Mutaguchi and K. Inagaki, *Actinomycetologica*, 2009, **23**, 51-55.
24. M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch and E. Inamine, *J. Antibiot.*, 1986, **39**, 259-265.
25. K. A. Reid, J. T. Hamilton, R. D. Bowden, D. O'Hagan, L. Dasaradhi, M. R. Amin and D. B. Harper, *Microbiology*, 1995, **141 (Pt 6)**, 1385-1393.
26. C. Schaffrath, S. L. Cobb and D. O'Hagan, *Angew. Chem. Int. Ed.*, 2002, **41**, 3913-3915.
27. C. Schaffrath, H. Deng and D. O'Hagan, *FEBS Lett.*, 2003, **547**, 111-114.
28. C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan and J. H. Naismith, *Nature*, 2004, **427**, 561-565.
29. X. Zhu, D. A. Robinson, A. R. Mcewan, D. O'Hagan and J. H. Naismith, *J. Am. Chem. Soc.*, 2007, **129**, 14597-14604.
30. D. O'Hagan, R. J. M. Goss, A. Meddour and J. Courtieu, *J. Am. Chem. Soc.*, 2003, **125**, 379-387.
31. H. M. Senn, D. O'Hagan and W. Thiel, *J. Am. Chem. Soc.*, 2005, **127**, 13643-13655.
32. F. Huang, S. F. Haydock, D. Spiteller, T. Mironenko, T.-L. Li, D. O'Hagan, P. F. Leadlay and J. B. Spencer, *Chem. Biol.*, 2006, **13**, 475-484.
33. C. Zhao, P. Li, Z. Deng, H.-Y. Ou, R. P. Mcglinchey and D. O'Hagan, *Bioorg. Chem.*, 2012, **44**, 1-7.
34. M. Onega, R. P. Mcglinchey, H. Deng, J. T. G. Hamilton and D. O'Hagan, *Bioorg. Chem.*, 2007, **35**, 375-385.

35. H. Deng, S. M. Cross, R. P. Mcglinchey, J. T. G. Hamilton and D. O'Hagan, *Chem. Biol.*, 2008, **15**, 1268-1276.

36. M. C. Walker, M. Wen, A. M. Weeks and M. C. Y. Chang, *ACS Chem. Biol.*, 2012, **7**, 1576-1585.

37. S. J. Moss, C. D. Murphy, J. T. Hamilton, W. C. McRoberts, D. O'Hagan, C. Schaffrath and D. B. Harper, *Chem. Commun.*, 2000, 2281-2282.

38. C. D. Murphy, S. J. Moss and D. O'Hagan, *Appl. Environ. Microbiol.*, 2001, **67**, 4919-4921.

39. C. D. Murphy, D. O'Hagan and C. Schaffrath, *Angew. Chem. Int. Ed.*, 2001, **40**, 4479-4481.

40. A. S. Eustáquio, J. Härle, J. P. Noel and B. S. Moore, *Chembiochem*, 2008, **9**, 2215-2219.

41. A. S. Eustáquio, F. Pojer, J. P. Noel and B. S. Moore, *Nat. Chem. Biol.*, 2008, **4**, 69-74.

42. H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith and D. O'Hagan, *Chembiochem*, 2014, **15**, 364-368.

43. A. M. Weeks, S. M. Coyle, M. Jinek, J. A. Doudna and M. C. Y. Chang, *Biochemistry*, 2010, **49**, 9269-9279.

44. A. M. Weeks and M. C. Y. Chang, *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, 19667-19672.

45. K. Gregg, B. Hamdorf, K. Henderson, J. Kopecny and C. Wong, *Appl. Environ. Microbiol.*, 1998, **64**, 3496-3498.

46. J. L. Baker, N. Sudarsan, Z. Weinberg, A. Roth, R. B. Stockbridge and R. R. Breaker, *Science*, 2012, **335**, 233-235.

47. C.-H. Liang, S. Yao, Y.-H. Chiu, P. Y. Leung, N. Robert, J. Seddon, P. Sears, C.-K. Hwang, Y. Ichikawa and A. Romero, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 1307-1310.

48. A. Rivkin, T.-C. Chou and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2005, **44**, 2838-2850.

49. S. Mo, D. H. Kim, J. H. Lee, J. W. Park, D. B. Basnet, Y. H. Ban, Y. J. Yoo, S.-w. Chen, S. R. Park, E. A. Choi, E. Kim, Y.-Y. Jin, S.-K. Lee, J. Y. Park, Y. Liu, M. O. Lee, K. S. Lee, S. J. Kim, D. Kim, B. C. Park, S.-g. Lee, H. J. Kwon, J.-W. Suh, B. S. Moore, S.-K. Lim and Y. J. Yoon, *J. Am. Chem. Soc.*, 2011, **133**, 976-985.

50. Y. Kobayashi, H. Wada, C. Rossios, D. Takagi, M. Higaki, S. i. Mikura, H. Goto, P. J. Barnes and K. Ito, *J. Pharmacol. Exp. Ther.*, 2013, **345**, 76-84.

51. J. Staunton and K. J. Weissman, *Nat. Prod. Rep.*, 2001, **18**, 380-416.

52. P. A. S. Lowden, B. Wilkinson, G. A. Böhm, S. Handa, H. G. Floss, P. F. Leadlay and J. Staunton, *Angew. Chem.*, 2001, **113**, 799-801.

53. M. A. Gregory, H. Petkovic, R. E. Lill, S. J. Moss, B. Wilkinson, S. Gaisser, P. F. Leadlay and R. M. Sheridan, *Angew. Chem.*, 2005, **117**, 4835-4838.

54. R. J. M. Goss, S. Lanceron, A. Deb Roy, S. Sprague, M. Nur-e-Alam, D. L. Hughes, B. Wilkinson and S. J. Moss, *Chembiochem*, 2010, **11**, 698-702.

55. S. L. Ward, R. P. Desai, Z. Hu, H. Gramajo and L. Katz, *J. Ind. Microbiol. Biotechnol.*, 2007, **34**, 9-15.

56. A. S. Eustáquio, D. O'Hagan and B. S. Moore, *J. Nat. Prod.*, 2010, **73**, 378-382.

57. M. C. Walker, B. W. Thuronyi, L. K. Charkoudian, B. Lowry, C. Khosla and M. C. Y. Chang, *Science*, 2013, **341**, 1089-1094.

58. S. Weist, B. Bister, O. Puk, D. Bischoff, S. Pelzer, G. J. Nicholson, W. Wohlleben, G. Jung and R. D. Süssmuth, *Angew. Chem. Int. Ed.*, 2002, **41**, 3383-3385.

59. F. Oldach, R. Al Toma, A. Kuthning, T. Caetano, S. Mendo, N. Budisa and R. D. Süssmuth, *Angew. Chem. Int. Ed.*, 2012, **51**, 415-418.

60. C. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413-444.

61. W. Runguphan, J. J. Maresh and S. E. O'Connor, *Proc. Natl. Acad. Sci. U.S.A.*, 2009, **106**, 13673-13678.

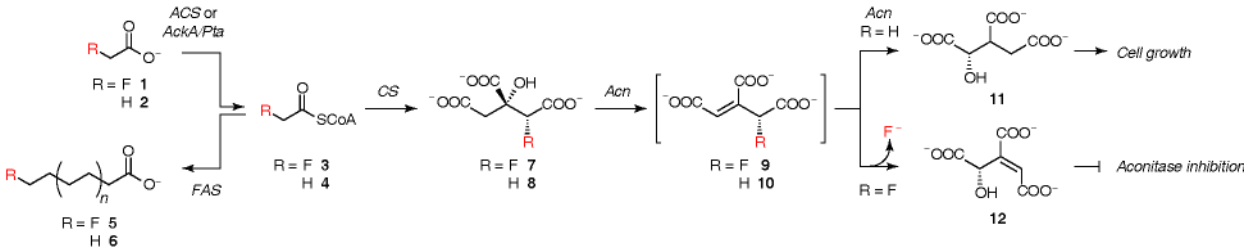


Figure 1. Metabolic pathways of characterized biogenic organofluorines. Fluoroacetate (1) is activated through the same routes as acetate (2), which produces fluoroacetyl-CoA (3) and acetyl-CoA (4), respectively. Fluoroacetyl-CoA can be used by fatty acid synthases (FAS) as a starter unit to initiate formation of ω-fluoro-fatty acids, which account for the majority of characterized organofluorines (5), in an analogous manner to canonical fatty acids (6). It can also be converted to fluorocitrate (7) while the normal cellular metabolite, acetyl-CoA, yields citrate (8) as an intermediate of the tricarboxylic acid (TCA) cycle. When the aconitase (Acn) acts on fluorocitrate and citrate, it produces fluoroaconitate (9) and aconitate (10), respectively. With the physiological substrate, the aconitate is rehydrated to generate isocitrate (11), which is an intermediate of the TCA cycle and used for cell growth. However, elimination of fluoride occurs from fluoroaconitate to yield 4-hydroxy-transaconitate (12), a potent inhibitor of the aconitase. Since the inhibitor is already bound to the active site and has an extremely slow off-rate, fluorocitrate is an effectively irreversible inhibitor that leads to disruption of the TCA cycle. (ACS, acetyl-CoA synthetase; AckA, acetate kinase; Pta; CS, citrate synthase; Acn, aconitase)

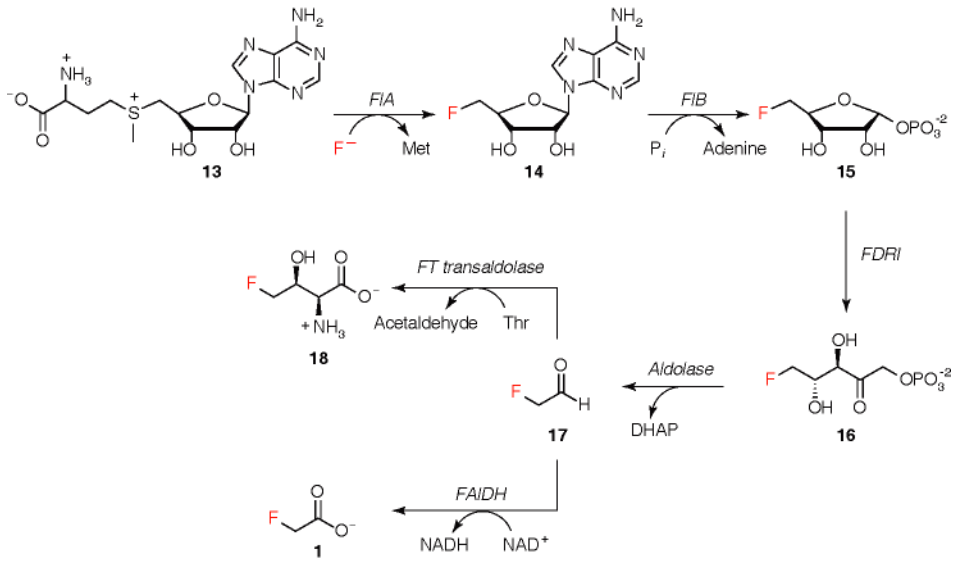


Figure 2. Organofluorine biosynthetic pathway in *S. cattleya*. SAM (13) and inorganic fluoride are converted to FDA (14) by the fluorinase (FlA). FDA is then converted to FDR1P (15) by the 5′-fluoro-5′-deoxyadenosine phosphorylase (FlB). Next, a 5-fluoro-5-deoxyribose-1-phosphate isomerase (FDRI) opens the ring and isomerizes the sugar to form FDRib1P (16). A yet to be identified aldolase is hypothesized to cleave FDRib1P to produce fluoroacetaldehyde (17), which is then oxidized to fluoroacetate (1) by the fluoroacetaldehyde dehydrogenase (FAlDH) or converted to fluorothreonine (18) by the fluorothreonine transaldolase (FT transaldolase).

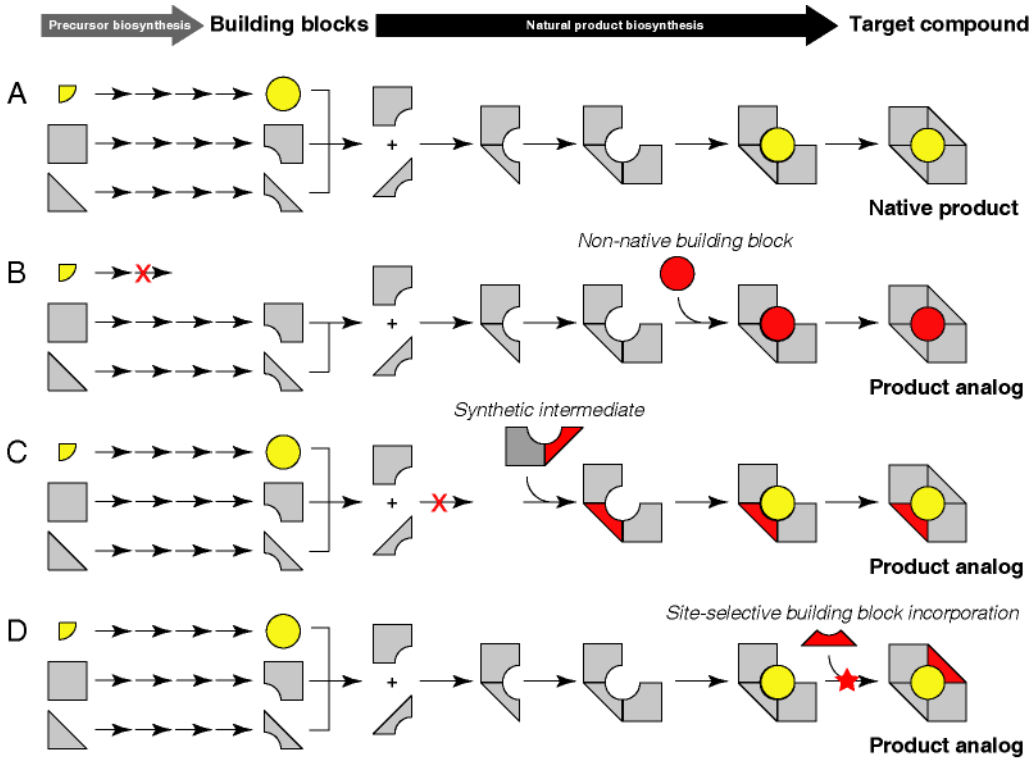


Figure 3. Strategies for the production of fluorinated natural products by engineering precursor and natural product biosynthesis. (A) Wild-type natural product biosynthesis with a native precursor supply. (B) Engineered natural product production through the disruption of the biosynthesis of a unique extender unit (red X) and feeding of non-native fluorinated building block. (C) Engineered natural product production through the disruption of initial steps in natural product biosynthesis (red X) and feeding a fluorinated synthetic intermediate. (D) Engineered natural product biosynthesis through the design of a fluorine-specific module (red star) of the biosynthetic machinery.

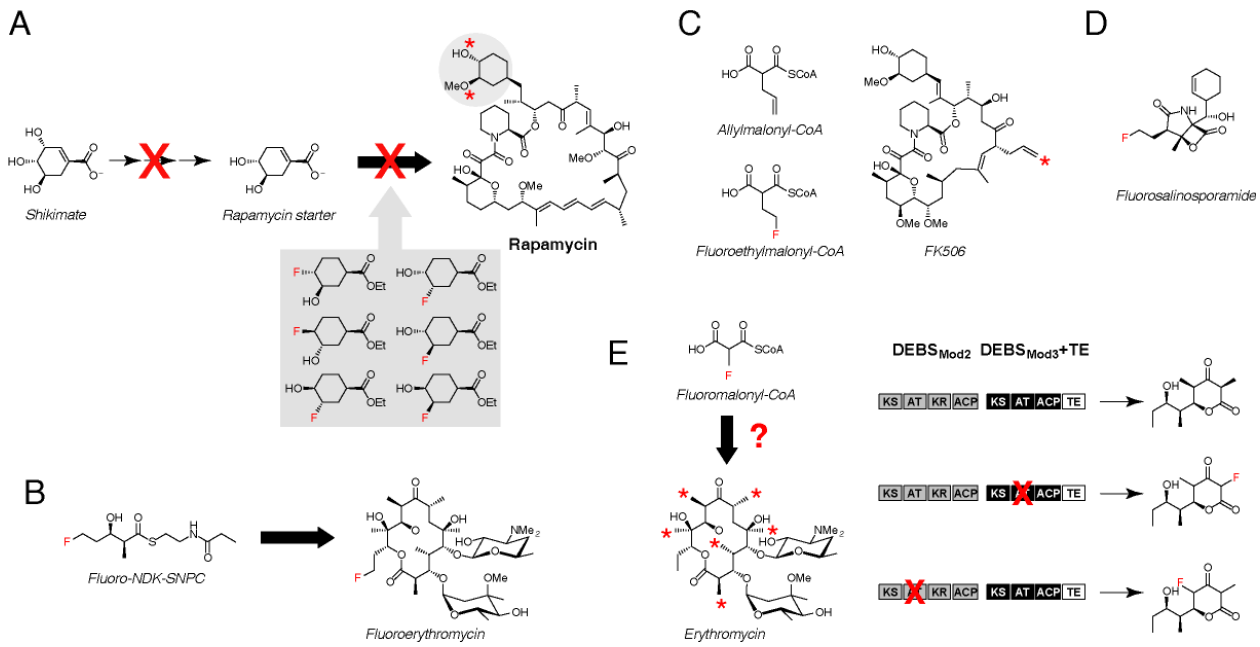


Figure 4. Engineering the production of fluorinated polyketides. (A) Disruption of the biosynthetic pathway to produce the starter unit for rapamycin can be rescued by the addition of synthetic analogs. The sites of fluorine incorporation are indicated with red stars. (B) Fluoro-NDK-SNPC is a synthetic intermediate can be fed to a truncated PKS system and be incorporated into the final product. (C) The allylmalony-CoA extender unit used in FK506 biosynthesis can be replaced with fluoroethylmalonyl-CoA when the production of the native extender is disrupted. The site of fluorine incorporation is indicated with a red star. (D) The chlorine substituent of salinosporamide can be replaced with fluorine when the gene encoding the fluorinase from *S. cattleya* is inserted in place of the chlorinase found in the native host. (E) Strategies for utilization of fluoromalonyl-CoA as an extender unit increases the breadth of fluorinated structures that can be approached (red stars indicate potential sites of fluorination) (left). Site-selective fluorination can be achieved using DEBS model systems (right).

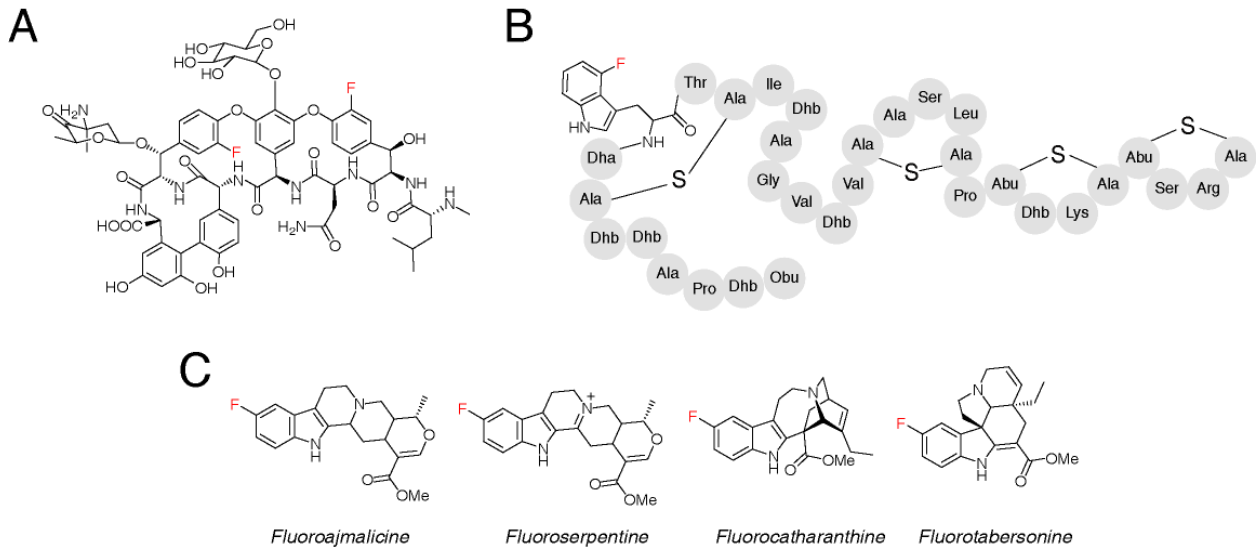


Figure 5. Other classes of engineered fluorinated natural products. (A) Incorporation of 2-fluoro-β-hydroxytyrosine into the nonribosomal peptide balhimycin. (B) Replacement of trypophan with 4-fluorotryptophan to produce a fluorinated analog of the ribosomally-synthesized peptide product lichenicidin. (C) Production of fluorinated terpene indole alkaloids in engineered plant cells.

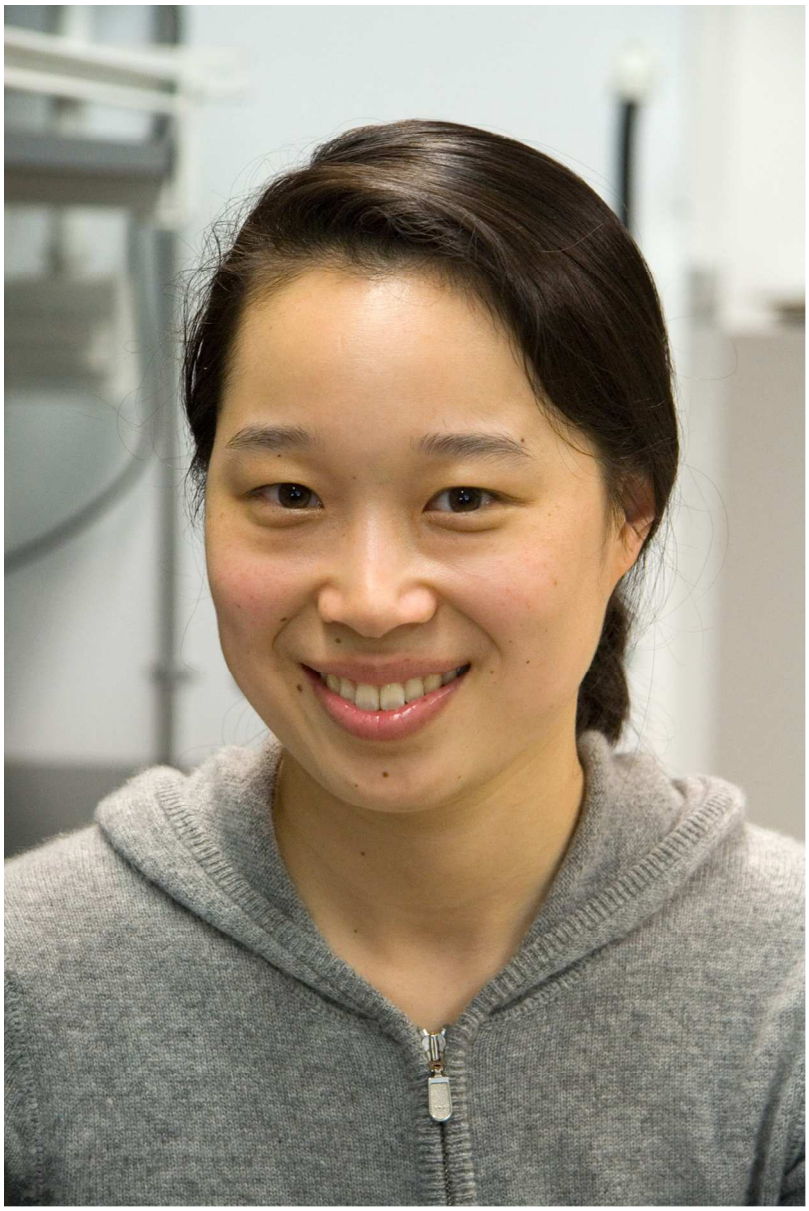
Mark Walker

Mark graduated from Oberlin College in 2004 with a B.A. in Biochemistry. He then obtained his Ph.D. in Molecular and Cell Biology in 2012 at the University of California, Berkeley in the laboratory of Michelle Chang studying organofluorine biosynthesis in natural and engineered systems. He is currently a postdoctoral fellow in the group of Wilfred van der Donk at the University of Illinois, Urbana-Champaign.

Michelle Chang

Michelle graduated from the University of California, San Diego in 1997 with a B.S. in Biochemistry and B.A. in French Literature. She then obtained her Ph.D. in Chemistry in 2004 at the Massachusetts Institute of Technology as a joint student with Daniel Nocera and JoAnne Stubbe. After completing her postdoctoral studies with Jay Keasling at the University of California, Berkeley, she started her independent career in 2007 with a joint appointment in the Departments of Chemistry and Molecular and Cell Biology at UC Berkeley.

564x534mm (72 x 72 DPI)



195x292mm (150 x 150 DPI)