Published in final edited form as:
ACS Synth Biol. 2013 November 15; 2(11): 635–642. doi:10.1021/sb4000355.

Sequence, Cloning, and Analysis of the Fluvirucin B_1 Polyketide Syn-thase from *Actinomadura vulgaris*

Tsung-Yi Lin, **Lawrence S. Borketey**, **Gitanjeli Prasad**, **Stephanie A. Waters**, and **Nathan A. Schnarr***
Department of Chemistry, University of Massachusetts, 710 N. Pleasant Street, Amherst, Massachusetts 01003

Abstract

Fluvirucin B_1, produced by *Actinomadura vulgaris*, is a 14-membered macrolactam active against a variety of infectious fungi as well as influenza A. Despite considerable interest from the synthetic community, very little information is available regarding the biosynthetic origins of the fluvirucins. Herein, we report the identification and initial characterization of the fluvirucin B_1 polyketide synthase and related enzymes. The cluster consists of five extender modules flanked by an N-terminal acyl carrier protein and C-terminal thioesterase domain. All but one of the synthase modules contain the full complement of tailoring domains (ketoreductase, dehydratase, and enoyl reductase) as determined by sequence homology with known polyketide synthases. Acitve site analyses of several key components of the cluster are performed to further verify that this gene cluster is associated with production of fluvirucin B_1. This work will both open doors toward a better understanding of macrolactam formation and provide an avenue to genetics-based diversification of fluvirucin structure.

Polyketides remain one of the most clinically important classes of natural products in the fight against infections, pathogens, and cancer.[1] Each metabolite within this class is biosynthesized by dedicated assemblies of proteins termed polyketide synthases (PKSs).[2–12] Mechanistically, PKSs work in analogous fashion to fatty acid synthase (FAS) where sequential decarboxylaive condensations between an acyl carrier protein (ACP)-bound malonate derivative and a ketosynthase (KS)-bound thioester intermediate serves to elongate the polyketide chain (Figure 1). Optional ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains then transform the resulting β-ketothioester to the desired β-functionality (hydroxyl, olefin, methylene).

Over the past several decades, researchers have devised numerous methods for manipulating these processes in an effort to diversify product structure and, ultimately, biological activity.[13–17] Although some success has been realized toward this end, decreased yields of engineered products have limited the scope and efficacy of these methods. At this point it is clear that the two primary factors leading to low yields are substrate selectivity of

Corresponding Authorschnarr@chem.umass.edu.

ASSOCIATED CONTENT

Supporting Information. Results of acyltransferase selectivity analyses, KR sequence analysis, and sequence comparisons between fluvirucin modules are provided This material is available free of charge via the Internet at http://pubs.acs.org

downstream enzymes and disruption of protein-protein interactions when heterologous enzymes are introduced.[18– 23] To circumvent the latter, strategies are needed where genetic alterations can be introduced without dispruting the native three-dimensional PKS architecture.

For modular polyketide synthases this means unearthing assemblies which bear the full complement of tailoring domains (KR, DH, and ER), similar to mammalian FAS, in most, if not all, active modules. Carefully executed mutagenesis of key active site residues should result in all possible β-functionalities while leaving the native protein-protein interactions intact (Figure 2). Therefore, PKSs that produce largely unfunctionalized polyketides (i.e. methylenes at most β-positions) may provide optimal engineering potential.

Fluvirucin B_1 is a 14-membered macrolactam produced by *Actinomadura vulgaris* with moderate to good antifungal and antiviral activities (Figure 3).[24–26] Following assembly of the core macrocycle, a single 3-amino-3,6-dideoxy-L-talopyranose is appended to the sole exocyclic hydroxyl group. The lack of additional ring functionalities peaked our interest as we hypothesized that nearly all active modules should contain the FAS-like domain organization where KR, DH , and ER are all present. If one were to consider PKS systems as evolutionary connected to FAS, the fluvirucin synthase may represent an early link between the two. As a result, we were motivated to explore the biosynthetic origin of fluvirucin B_1 with the ultimate goal of providing a platform for polyketide engineering that circumvented the need for incorportation of heterologous domains to achieve maximal product diversity. Herein, we describe our efforts to unearth and characterize the modular PKS associated with fluvirucin B_1 production in *Actinomadura vulgaris*.

Results

From the core structure of fluvirucin B_1, we hypothesized that the producing PKS would consist of five extender modules assuming that a β-alanine derivative is used as a starter unit in the process. The sole hydroxyl group would therefore arise from a module harboring only a KR domain while all other extender modules would contain the full compliment of KR, DH, and ER domains (Figure 4). Based on the positions of macrolactam ring substituents, we expected that: (1) the first and last modules contained ethylmalonyl-specific AT domains, (2) the second and fourth modules incorporated malonyl groups, and (3) the third module utilized methylmalonate (Figure 4). Finally, ring closure was most likely achieved via a C-terminal thioesterase (TE) domain as is the case with similar macrocyclic polyketides.[27–29] To test these hypotheses and determine the precise arrangement of enzymes within the assembly, we set out to identify and characterize the fluvirucin B_1 biosynthetic gene cluster.

To do so, the producing organism, *Actinomadura vulgaris*, was cultured following published procedures.[24–26] Genomic DNA was isolated and sequenced (Beckman Coulter Genomics) affording 436,311 overlapping sequence fragments. These sequences were partially assembled resulting in 444 consensus sequences ranging in size from 5000 to 170,000 base pairs. The relatively large size of PKS constructs allowed us to quickly identify potential hits by searching each assembled sequence for open reading frames of at least 4000 base pairs.

Our search identified several PKS gene clusters, one of which contained the expected size and module composition of the proposed fluvirucin PKS.

Fluvirucin B_1 PKS Genes

Three modular PKS genes, *flu A-C*, were found to contain an arrangement and composition of domains consistent with the expected fluvirucin PKS assembly (Figure 5). In addition, several PKS-associated genes were uncovered within the cluster including a pair of transciptional regulators (*fluE, fluG*), a glycosyl transferase (*fluF*), a decarboxylase (*fluI*), and a drug transporter (*fluD*) (Figure 5). The *fluJ* sequence shows high similarity to crotonyl-CoA reductases and likely plays a role in ethylmalonate generation as is required for fluvirucin B_1 biosynthesis.

FluA contains modules 1 and 2 of the fluvirucin B_1 assembly. As expected, module 1 has the full compliment of tailoring domains (KR, DH, and ER) while module 2 possesses only a KR domain putatively leading to the sole hydroxyl group on the macrolactam ring. A single loading ACP is found at the N-terminus of FluA similar to the vicenistatin PKS which utilizes an α-methyl-β-alanine starter unit.[30]

FluB contains modules 3 and 4 of the fluvirucin B_1 PKS. Both modules contain KR, DH, and ER domains as would be predicted from the fluvirucin core structure. Finally, FluC consists of module 5 and a C-terminal TE domain. Module 5 again has all three tailoring domains consistent with the lack of functionality at the corresponding macrolactam ring position. As described in detail below, modules 1, 3, and 5 are extremely similar to each other in terms of primary structure as are modules 2 and 4 but similarities drop significantly when comparisons are made between these two groups. It is therefore tempting to speculate that the fluvirucin B1 synthase is composed of modules arising from two separate evolutionary ancestors; one leading to modules 1, 3, and 5 and another leading to modules 2 and 4.

AT Selectivities

Extender unit selectivity for AT domains provides compelling evidence for any link between a given PKS assembly and its associated polyketide product. Using, the SEARCHPKS program developed by Mohanty and coworkers, probable coenzyme A substrates were determined for each of the five putative fluvirucin B_1 synthase AT domains.[31] To our delight, all of the predicted AT domain specificities were consistent with the fluvirucin B_1 core stucture (see supporting info, Figure S1). Specifically, modules 2 and 4 showed high sequence similarity with malonyl-specific AT domains while module 3 was predicted to utilize methylmalonyl-CoA. Modules 1 and 5 returned a single hit for ethylmalonate specificity amidst several methylmalonyl-specific AT domains. To experimentally verify the putative AT selectivities for modules 1, 3, and 5, ketosynthase-acyltransferase (KSAT) didomains were cloned from these modules along with the corresponding ACP domains. Each KSAT didomain was mixed with ACP from the same module followed by introduction of either malonyl-, methylmalonyl-, or ethylmalonyl-*N*-acetylcysteamine (SNAc) thioesters. Following 30min uncubation with each substrate, samples were trypsinized and the extent of AT to ACP transfer for each extender unit was analyzed by LC-MS (Figure 6). Gratifyingly,

experimentally determined AT selectivities for modules 1, 3, and 5 were consistent with those suggested by sequence homology. It is important to note that no direct acylation of ACP was observed with any of the extender units indicating that AT to ACP transfer is the sole mechanism for formation of the acylated ACP (data not shown).

Module 3 Substrate Selectivity

The hydroxyl group generated in module is the only macrolactam ring substituent not introduced by an extender unit. To examine the selectivity in acceptor module (module 3) for the hydroxyl group stereoconfiguration and provide further evidence connecting this PKS with fluvirucin B_1 production, tandem proteolysis/LC-MS was again employed. SNAc thioesters of both enantiomers of 3-hydroxybutyrate were prepared and introduced separately to KSAT3. Following 1h incubation, samples were trypsinized and KS-acylation was observed via LC-MS analysis. As predicted from the fluvirucin B_1 structure, only the (*S*)-3-hydroxybutyryl-SNAc compound, which places the hydroxyl group in the same three-dimensional orientation as the ring hydroxyl of the final product, was accepted by the module 3 KS domain (Figure 7). Very little to no KS-acylation with the (*R*) – isomer was observed.

KR Domains

The Keatinge-Clay lab previously uncovered primary sequence patterns associated with different types of ketoreductase domains commonly found in modular PKS systems.[32] Examination of the fluvirucin KR sequences within this context revealed that all five align with the B1-type KR sequence pattern (see supporting information). This KR family is generally observed when the KR works in concert with other tailoring domains such as DH and ER which is again consistent with the expected enzyme composition for fluvirucin PKS modules 1, 3, 4, and 5. Interestingly, the module 2 KR domain also shows B1-type sequence character despite the absence of DH and ER domains from that module. This observation may further highlight the fascinating evolution of the fluvirucin synthase as discussed below. These results, together with the observed tailoring domain patterns and AT selectivities, provided the necessary evidence to link this polyketide synthase with the biosynthesis of fluvirucin B_1.

Fluvirucin B_1 Synthase Cloning

To confirm the sequences obtained from partial assembly of the *A. vulagaris* genome and with the ultimate goal of reconstituting the entire assembly in *E. coli*, we turned our attention toward cloning each module individually from genomic DNA. Based on alignment with known PKS constructs, we were able to determine effective sequence boundaries for each fluvirucin synthase module. All five modules were cloned separately in pET vectors for expression in *E. coli*. Overexpression of each module was observed in BL21(DE3) cells and gel migration patterns were consistent with calculated protein masses (Figure 8). The fact that *E. coli* seems to respond well to these foreign genes bodes well for our future efforts aimed at generating fluvirucin-derived stuctures in this heterologous host. In the near term, the ability to reliably produce usable quantities of each module will greatly facilitate studies

concerning the substrate specificities and enzyme kinetics that govern fluvirucin B_1 biosynthesis.

Discussion

Fluvirucin B1 is a relatively simple natural product stemming from a rather complex set of biosynthetic transformation. Despite the diminutive size of the PKS responsible for its production in *A vulgaris*, each round of elongation and subsequent β-carbon tailoring requires extensive manipulation of functionality. Four of the five putative extender modules bear the full compliment of tailoring domains meaning that at each of these positions within the assembly, keto-, hydroxyl-, and olefin-containing intermediates are generated en route to the fully saturated product similar to mammalian fatty acid synthase. We have hypothesized that this type of module composition will afford the highest engineering potential as product diversification can be achieved without the need for incorportation of heterlogous domains. In other words, one can potentially access each of the afforementioned functionalities by simple active site mutagenesis of KR, DH, and ER domains leaving the highly evolved protein-protein communication and recognition interfaces in their native states. This is in stark contrast to more popular assemblies like 6-deoxyerythronolide B synthase (DEBS) where nearly all of the extender modules contain, at most, a KR domain, where only ketone functionalities are accessible via similar active site mutagenesis strategies.[2] For this reason, we were eager to uncover the biosynthetic origins of fluvirucin B_1.

As predicted, the fluvirucin B_1 polyketide synthase consists of 5 extender modules flanked by an N-terminal loading ACP and C-terminal TE domain. All but one of the extender modules contains a KR, DH, and ER domain in addition to the required KS, AT, and ACP leading to the relatively unfunctionalized nature of the macrolactam product. Based on this arrangement of composition of modules, β-alanine is expected to serve as the starter unit for fluvirucin B_1 biosynthesis. As strong evidence for this hypothesis, *fluI*, which putatively encodes for a PLP-dependent decarboxylase, displays high homology with both *vinO* from the vicenistatin PKS cluster and *azicN* from the azicemicin PKS cluster.[30,33] The former is responsible for decarboxylation of 3-methylaspartate while the latter decarboxylates aspartic acid itself leading to 3-methyl-β-alanine and β-alanine, respectively. While further studies are needed to confirm the starter unit identity for fluvirucin B_1 biosynthesis, this data strongly suggests a role for β-alanine in the early stages of macrolactam construction.

As alluded to above, thorough analysis of the protein sequences for each module reveals an intriguing trend with implications as to the evolutionary origins of these PKS components. Pairwise alignments between fluvirucin B_1 PKS modules 1, 3, and 5 yield protein sequence identities ranging from 75% −81% (see supporting information). Similarly, modules 2 and 4 show 94% sequence identity. When analogous alignments are executed between these two groups (e.g. module 1 vs. module 2), more typical identities ranging from 60% – 64% are observed. By comparison, sequence identities between modules from the well-characterized DEBS assembly as well as between DEBS modules and fluvirucin PKS modules fall in the more modest 40%–60% range. The similarities between fluvirucin B_1 synthase modules might suggest independent ancestry for modules 1, 3, and 5 versus 2 and 4. It is important to note that the remarkable sequence identites observed within these two groups occurs despite

the fact that each module both accepts and processes appreciably different polyketide intermediates.

Another interesting aspect of the fluvirucin B_1 synthase involves the TE domain. Most macrocycle-forming thioesterases bear a conserved serine residue charged with accepting the fully mature, linear polyketide intermediate from an immediately upstream ACP followed by cyclization and product release. The fluvirucin B_1 TE domain instead uses a cysteine active site for this task. This type of serine to cysteine substitution has been observed in other PKS systems prompting speculation as to possible divergent evolutionary origins between these two active site arrangements.[34–36] Although beyond the scope of this manuscript, this somewhat unique feature of the fluvirucin B_1 synthase should provide additional insights into any kinetic consequences of this switch and thus warrants further study.

Experimental Procedures

Materials

All Biochemicals, chemicals, and media were obtained from Fisher Scientific, all restriction enzymes were obtained from New England Biolabs, and other molecular biological reagents were obtained from Fisher Scientific, New England Biolabs, or Invitrogen. All PCR primers were synthesized by Eurofins MWG Operon.

All the DNA sequences were deposited in the GenBank database under the following accession numbers: **JX308234** (*FluA*), **JX915256** (*FluB*), **JX448408** (*FluC*)

Bacterial strains, culture conditions and DNA purification

Actinomadura vulgaris was purchased through American Type Culture Collection (ATCC) by the accession number ATCC 53715 and used as the source of DNA for shot-gun sequencing service and the cloning of Fluvirucin B_1 polyketide synthase. The strain was cultivated at ambient temperature in the liquid medium of ATCC Medium 172 (N-Z Amine with Soluble Starch and Glucose), which contains 1% glucose, 2% soluble starch, 0.5% yeast extract, 0.5% N-Z amine type A (Sigma C0626), 0.1% $CaCO_3$. The growth of *A vulgaris* at ambient temperature can be observed after 3 days of culture. For genomic DNA extraction purposes, *A. vulgaris* was cultured for 6 days and then genomic DNA was extracted by using the MasterPure gram-positive DNA purification kit (Epicenter, Madison, WI)

All *E. coli* strains used in the study (Top10, BL21(DE3), and BAP1) were propagated at 37 °C in Luria Broth or on Luria agar supplemented with the appropriate antibiotics when needed. All plasmid DNA was prepared using a Qiagen miniprep kit.

Cloning of module 1 of the fluvirucin B_1 polyketide synthase

The PCR reaction for module 1 was performed in a 50μl reaction mixture containing: 1X Phusion GC buffer, 0.2mM dNTP, 0.3mM $MgCl_2$, 5%DMSO, 1μM of each primer (AMod1-NheI-F1: AAAAAAGCTAGCATGAGCCAGTCCGGAAACAGCGAA; Avul-Mod1-R6: CCGCCCAGACATGACCGAACTG), 1U Phusion Hot Start II DNA polymerase

(Thermo scientific, USA) and approximately 450ng of genomic DNA was added as template.

The thermal cycler (Mastercycler ep gradient, Eppendorf) was programmed according to the following "2-step" amplification profile: 3min denaturation at 98°C, then 10 initial cycles of 10sec denaturation at 98°C, 5min annealing and elongation at 72°C, followed by 27 cycles of 10sec denaturation at 98°C, 5min + (5sec/cycle) elongation at 72°C, and a final extension step at 72°C for 5min. The amplified DNA fragments (7221bp) were then subjected to 0.8% agarose gel and single bands were excised and purified using a QIAquick Gel Extraction Kit (QIAGEN, Germany) according to the instructions from the manufacturer. Subsequently, it was subjected to restriction enzyme digestion with NheI and NotI and the digested products were ligated to pre-digested pET-21b to obtain pTL-A01.

Cloning of module 2 of the fluvirucin B_1 polyketide synthase

The PCR reaction for pre-module 2 was performed in a 50μl reaction mixture containing: 1X Phusion GC buffer, 0.2mM dNTP, 0.3mM $MgCl_2$, 15% glyerol, 0.5M sulfolane[37], 1μM of each primer (Avul-Mod2-pF6: CCCATCAACACCCACACCCT; Avul-Mod2-R7: GCCATCCACAGGTAGCGGTTG), 1U Phusion Hot Start II DNA polymerase (Thermo scientific, USA) and approximately 450ng of genomic DNA was added as template.

The thermal cycler was programmed according to the following "stepdown" amplification profile: 3min denaturation at 98°C, then 10 initial cycles of 10sec denaturation at 98°C, 30sec annealing at 72–68°C, 6min elongation at 72°C where the annealing temperature was decreased by 0.4°C per cycle, followed by 27 cycles of 10sec denaturation at 98°C, 30sec annealing at 68°C, 6min + (5sec/cycle) elongation at 72°C, and a final extension step at 72°C for 5min.

The amplified DNA fragments (6760bp) were then purified, and directly inserted into the plasmid vector PCR-Blunt II Topo (Zero blunt TOPO PCR cloning kit, Invitrogen) to obtain pM2-44-4-3 (pTL-preM2). After obtaining pTL-preM2, the same PCR protocol was performed as stated above, except for using the following primers: AMod2-NdeI-F8: AAAAAACATATGACGCTGGTGTTCGACCAC; AMod2- HindIII-R1: TTTTTTAAGCTTGGACGCGCCGAGCTGGTC. The DNA amplicons (5265bp) were digested by NdeI and HindIII, and then were ligated to pre-digested pET-21b to obtain pTL-A02.

Cloning of module 3 of the fluvirucin B_1 polyketide synthase

The PCR reaction for module 3 was performed in a 50μl reaction mixture containing: 1X Phusion GC buffer, 0.2mM dNTP, 0.3mM $MgCl_2$, 7%DMSO, 1μM of each primer (AMod3-EcoRI-F2: AAAAAAGAATTCGATGGCCACTGACGACAAGTTCCGG; AMod3-R2: TTTTTTGTGGACGTGGACGCGGCTCGGAC), 1U Phusion Hot Start II DNA polymerase (Thermo scientific, USA) and approximately 450ng of genomic DNA was added as template.

The thermal cycler was programmed as for module 1, and the amplified DNA fragments (6338bp) were purified, and directly digested by EcoRI and NotI. The digested products were ligated to pre-digested pET-21b to obtain pTL-A03.

Cloning of module4 of Fluvirucin B_1 polyketide synthase

The PCR protocol for pre-module 4 is the same as for pre-module 2, except for using the following primers: AMod4-HindIII-F8: AAAAAA AAGCTT CGGCAAGATCATCCTGACCATGC and AMod4h-R10: CGGGTACATGCCCAAGGAGTTGA are for M4-1f fragments (5863bp); AMod4h-F7: CAACGCACAAGACATCCAACA and AMod4-XhoI-R7: TTTTTTCTCGAGCAGGCCCTGGTCGATCAGCGAGAAGAG C are for M4-2f fragments (4339bp). M4-1f fragments and M4-2f fragments are separately inserted into the plasmid vector PCR-Blunt II Topo to obtain pTL-M4-1f and pTL-M4-2f Later, the pTL-M4-1f plasmids were digested by HindIII and MluI to obtain M4-1f fragments again to clone into pTL-M4-2f to harvest pTL-preM4.

Finally, the pTL-preM4 was digested by NotI and XhoI and ligated to pre-digested pET-21b to obtain pTL-A04.

Cloning of module 5 + TE of the fluvirucin B_1 polyketide synthase

The PCR reaction for module 5 was performed in a 50μl reaction mixture containing: 1X Phusion GC buffer, 0.2mM dNTP, 0.3mM $MgCl_2$, 10%DMSO, 1μM of each primer (AMod5-NheI-F3: AAAAAAGCTAGCATGGCTGACGAAGAGAAGCTCCTC; AMod5-HindIII-R2: TTTTTTAAGCTTCGCGCCGTTCGA), 1U Phusion Hot Start II DNA polymerase (Thermo scientific, USA) and approximately 450ng of genomic DNA was added as template.

The thermal cycler was programmed as for module 1, and the amplified DNA fragments (7196bp) were then extracted and subjected to NheI and HindIII double digestion. Finally, the digested products were ligated to pre-digested pET-21b to obtain pTL-A05.

Cloning of Flu KSAT1, KSAT3, KSAT5 and Flu ACP1, ACP3, ACP5 domains of the fluvirucin B_1 polyketide synthase

The DNA sequence encoding Flu-KSAT1 was amplified from pTL-A01 by the PCR protocol described for pre-module 2. Flu KSAT1 was constructed as an NheI-EcoRI fragment by using following primers, pTL-KSAT1-F: TTTTTTGCTAGCGAGCCCATCGCGATCGTC and pTL-KSAT1-R: AAAAAAGAATTCTGGTCCACGGCGGCCTGG. This NheI-EcoRI fragment was cloned into the pET21b expression vector to yield plasmid pTL-KSAT1.

The DNA sequence encoding Flu-KSAT3, Flu-KSAT5 and Flu-ACP1, Flu-ACP3, Flu-ACP5 were cloned similarly by the corresonding templates the corresponding primers as follows, pTL-KSAT3-F1: TTTTTTGCTAGCATGGCCACTGACGACAAG, pTL-KSAT3-R1: AAAAAAGAATTCGGATCCACCCGGGTCAGG; pTL-KSAT5-F1: TTTTTTGCTAGCATGGCTGACGAAGAGAAG, pTL-KSAT5-R2:

AAAAAAGAATTCTGATCCACCCGAGCCTG; pTL-ACP1-F: TTTTTTGCTAGCCTGACCGGACTACCGGCG, pTL-ACP1-R: AAAAAAGAATTCGTGACACGCTGGAGCAG; pTL-ACP3-F: TTTTTTGCTAGCCTGACCGGCCTGCCCGCG, pTL-ACP3-R: AAAAAAGAATTCGTGACACGCTGGAGCAG; pTL-ACP5-F: TTTTTTGCTAGCCTGGCCGGGCTGTCG, pTL-ACP5-R: AAAAAAGAATTCGCGATCTCCTCCGCCAG

The resulting plamids for Flu-KSATs (pTL-KSAT1, pTL-KSAT3, pTL-KSAT5) were transformed into BL21(DE3). The plamids constructed for Flu-ACPs (pTL-ACP1, pTL-ACP3, pTL-ACP5) were introduced into BAP1.

General Procedure for Protein Expression and Isolation

E. coli (BL-21) bearing the appropriate plasmid were grown in 1L shake cultures of LB-ampicillin media at 37°C until the OD_{600} was between 0.6 and 0.8. Overexpression was induced by adding 200μL of 1M IPTG at appropriate induction temperature (see below) for 16 hours. After this point, all work was carried out at 4°C. Cells were pelleted by spinning at 6000 RPM for 10 minutes and resuspended in 10mL of lysis buffer (20 mM Tris-HCl, 150mM NaCl, 1mM Na_2EDTA, 1mM EGTA, 1% Triton, 2.5mM sodium pyrophosphate, 1mM β-glycerophosphate, 1mM Na_3VO_4, 1μg/mL leupeptin, pH 8). Cells were lysed for five 30 second intervals with a 60 second cool down period between each. Lysed cells were spun at 14,000 rpm for 60 minutes. The lysate supernatant was equilibrated with 3mL of Ni-NTA bead slurry for 60 minutes. The mixture was then poured into a 15mL column and the supernatant eluted. The column was then washed with two 15mL portions of wash buffer (50mM phosphate, 300mM NaCl, 50mM imidazole, pH 8.0), and eluted with 3mL of elution buffer (50mM phosphate, 300mM NaCl, 250mM imidazole, pH 8.0). The purified protein was loaded onto a 100kDa cutoff centrifugal concentrator and diluted to 15mL with storage buffer (100mM Tris, 1mM EDTA, 1mM dithioerythritol, 10% glycerol, pH 8) and spun at 3000 rpm. Dilution and filtration was repeated a total of three times. Protein concentration were determined by Bradford assay with an average concentration of approximately 500 μM. Proteins were flash frozen and stored at −80°C until use.

Protein	Plasmid	Induction Temperature	Yield (mg/L)
module 1	pTL-A01	25°C	50
module 2	pTL-A02	18°C	40
module 3	pTL-A03	25°C	60
module 4	pTL-A04	18°C	50
module 5	pTL-A05	25°C	4
KSAT1	pTL-KSAT1	25°C	80
KSAT3	pTL-KSAT3	25°C	80
KSAT5	pTL-KSAT5	25°C	80
ACP1	pTL-ACP1	25°C	100
ACP3	pTL-ACP3	25°C	100

Protein	Plasmid	Induction Temperature	Yield (mg/L)
ACP5	pTL-ACP5	25°C	100

Synthesis of (*R*)- and (*S*)-3-hydroxybutyryl-SNAc

To a round bottom flask with sir bar was added DCM (10mL), EDCI.HCl (1.2eq), R-3-hydroxybutyric acid (or (S)-3-hydroxybutyric acid)(1.1eq), DMAP (0.02eq) and N-acetyl cysteamine (SNAc) (1eq). Reaction was stirred overnight at room temperature. The reaction was diluted with 10mL DCM and 20mL H_2O. The organic phase was washed with saturated NH_4Cl (aq), $NaHCO_3$ (aq) and brine. The reaction was dried with anhydrous sodium sulfate and concentrated to yield pure titled product as a clear liquid (91% yield).

(*S*)- 3-hydroxybutyryl-SNAc

^{1}H NMR (400 MHz, CH_3Cl-*d*) δ ppm 1.33 (3 H, d, *J*=6.82 Hz), 1.98 (3 H, s), 2.65 - 2.76 (2 H, m), 2.78 - 2.82 (1 H, d), 3.04 - 3.10 (2 H, t), 3.27 - 3.31 (1 H, m), 3.42-3.46(3H, q), 6.36 (1 H, br. s.)

^{13}C NMR (101 MHz, CH_3Cl -*d*) δ ppm 19.76, 21.35, 26.95, 28.50, 34.57, 37.48, 49.19, 168.60, 195.73

LRMS [M+H] for $C_8H_{15}NO_3S$ calcd 206.1 found 206.1

(*R*)- 3-hydroxybutyryl-SNAc

^{1}H NMR (400 MHz, CH_3Cl-*d*) δ ppm 1.33 (3 H, d, *J*=6.82 Hz), 1.98 (3 H, s), 2.65 - 2.76 (2 H, m), 2.78 - 2.82 (1 H, d), 3.04 - 3.10 (2 H, t), 3.27 - 3.31 (1 H, m), 3.42-3.46(3H, q), 6.36 (1 H, br. s.)

^{13}C NMR (400 MHz, CH_3Cl -*d*) δ ppm 19.76, 21.35, 26.95, 28.50, 34.57, 37.48, 49.19, 168.60, 195.73

LRMS [M+H] for $C_8H_{15}NO_3S$ calcd 206.1 found 206.1

Synthesis of malonyl and substituted malonyl SNAc thoesters

These syntheses we carried out using established procedures[38]. A general outline follows:

To a solution of appropriate malonic or substituted malonic acid (1eq) in dry THF (5mL), was added pyridine (2.2eq) and t- butanol (1.8eq). The solution was cooled to 0°C. Methanesulfonyl chloride (1.05eq) was then added dropwise over a 10 minute period. The reaction mixture was warmed to room temperature and stirred for 3 hours. The mixture was filtered and the resulting filtrate diluted with water. The pH was adjusted to ~12 and washed 3X with dichloromethane. The aqueous layer was acidified (pH ~2), extracted 3X with dichloromethane and dried with sodium sulfate. The product was then coupled to SNAc via EDC coupling. In an RB flask, the acid (1.1eq) was dissolved in dichloromethane. EDCI (1.2eq), DMAP (0.02eq) and SNAc (1eq) were then added and reaction stirred at room temperature overnight. The mixture was diluted with water and dichloromethane. The

organic layer was washed with NH_4Cl, $NaHCO_3$, and brine, and then dried with Na_2SO_4. Concentration *in vacuo* yielded the expected product. The product was dissolved in TFA at 0°C. After stirring at 0°C for 24 hours, TFA was evaporated. The product was diluted with diethyl ether and concentrated. This was repeated three times. The crude product was purified by chromatography to yield the titled compound.

Malonyl SNAc thioester (pale white solid) (70% yield)

^{1}H NMR (400 MHz, *DMSO-d*$_6$) δ ppm 1.79 (br. s., 3 H), 2.95 (br. s., 2 H), 3.18 (br. s., 2 H), 3.65 (br. s., 2 H), 8.06 (br. s., 1 H)

^{13}C NMR (400 MHz, *DMSO-d*$_6$) δ ppm 21.04, 27.06, 36.57, 48.21, 165.97, 167.91, 190.24

LRMS [M+H] for $C_7H_{11}NO_4S$ calcd 206.0 found 206.1

Methyl Malonyl SNAc thioester (pale white solid) (63% yield)

^{1}H NMR (400 MHz, CH_3Cl -*d*) δ ppm 1.44 (3 H, d, *J*=7.07 Hz), 1.99 (3 H, s), 3.06-3.17 (2 H, m), 3.40 - 3.52 (2 H, m), 3.63 - 3.77 (1 H, m), 6.71 (1 H, t, *J*=5.56 Hz), 10.35 (1 H, br. s.)

^{13}C NMR (400 MHz, CHLOROFORM-*d*) ppm 12.13, 20.74, 26.59, 37.64, 52.07, 170.24, 170.59, 194.96

LRMS [M+H] for $C_8H_{13}NO_4S$ calcd 220.1 found 220.1

Ethyl malonyl SNAc thioester (pale yellow solid) (51% yield)

^{1}H NMR (400 MHz, CH_3Cl -*d*) δ ppm 1.01 - 1.06 (2 H, m), 2.01 (3 H, s), 3.05 (1 H, dd, *J*=13.26, 6.69 Hz), 3.11 - 3.22 (1 H, m), 3.37 (1 H, t, *J*=7.20 Hz), 3.48 (2H, q, *J*=5.89 Hz), 3.56 (1 H, t, *J*=7.45 Hz), 6.59 (1H, br. s.), 11.53 (1H, br. s.)

^{13}C NMR (101 MHz, CH_3Cl -*d*) δ ppm 9.83, 20.63, 20.98, 26.62, 37.73, 51.02, 59.44, 170.96, 172.31, 194.21

LRMS [M+H] for $C_9H_{15}NO_4S$ calcd 234.1 found 234.1

Conclusions

In summary, we have identified and characterized the putative PKS genes associated with fluvirucin B_1 aglycone biosynthesis in *A. vulgaris*. The number and composition of modules as well as predicted AT specificities are consistent with the fluvirucin B_1 structure. The abundance of tailoring domains within the assembly is expected to provide increased engineering potential, through straightforward active site mutagenesis. Reconstitution of fluvirucin B_1 aglycone biosynthesis in a more workable host will greatly facilitate these studies and efforts to do so are currently underway in our laboratory.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by start-up funding from the University of Massachusetts, Amherst (to N.A.S). The authors would like to thank James Chambers and Min Chen for helpful discussions.

Notes and references

1. O'Hagan, D. The Polyketide Metabolites. New York: E. Horwood; 1996.
2. Khosla C, Tang Y, Chen AY, Schnarr NA, Cane DE. Structure and mechanism of the 6-deoxyerythronolide B synthase. Annu. Rev. Biochem. 2007; 76:195–221. [PubMed: 17328673]
3. Staunton J, Weissman KJ. Polyketide biosynthesis: A millennium review. Nat. Prod. Rep. 2001; 18:380–416. [PubMed: 11548049]
4. Khosla C. Natural product biosynthesis: A new interface between enzymology and medicine. J. Org. Chem. 2000; 65:8127–8133. [PubMed: 11101363]
5. Donadio S, Katz L. Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea* . Gene. 1992; 111:51–60. [PubMed: 1547954]
6. Donadio S, Staver MJ, Mcalpine JB, Swanson SJ, Katz L. Modular organization of genes required for complex polyketide biosynthesis. Science. 1991; 252:675–679. [PubMed: 2024119]
7. Smith S, Tsai S-C. The type I fatty acid and polyketide synthases: a tale of two megasynthases. Nat. Prod. Rep. 2007; 24:1041–1072. [PubMed: 17898897]
8. Smith S. The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. FASEB J. 1994; 8:1248–1259. [PubMed: 8001737]
9. Wakil SJ. Fatty acid synthase, a proficient multifunctional enzyme. Biochemistry. 1989; 28:4523–4530. [PubMed: 2669958]
10. Asturias FJ, Chadick JZ, Cheung IK, Stark H, Witkowski A, Joshi AK, Smith S. Structure and molecular organization of mammalian fatty acid synthase. Nat. Struct. Mol. Biol. 2005; 12:225–232. [PubMed: 15711565]
11. Smith S. Architectural options for a fatty acid synthase. Science. 2006; 311:1251–1252. [PubMed: 16513973]
12. Maier T, Jenni S, Ban N. Architecture of mammalian fatty acid synthase at 4.5 Å resolution. Science. 2006; 311:1258–1262. [PubMed: 16513975]
13. Schnarr, NA.; Khosla, C. Chemical Biology: From Small Molecules to Systems Biology and Drug Design. Wiley-VCH; Weinheim: 2006.
14. Donadio S, Sosio M. Strategies for combinatorial biosynthesis with modular polyketide synthases. Combinatorial Chem. High Throughput Screening. 2003; 6:489–500.
15. Walsh CT. Combinatorial biosynthesis of antibiotics: Challenges and opportunities. ChemBioChem. 2002; 3:124–134.
16. Staunton J, Wilkinson B. Combinatorial biosynthesis of polyketides and nonribosomal peptides. Curr. Opin. Chem. Biol. 2001; 5:159–164. [PubMed: 11282342]
17. Cane DE, Walsh CT, Khosla C. Harnessing the biosynthetic code: combinations, permutations, and mutations. Science. 1998; 282:63–68. [PubMed: 9756477]
18. McDaniel R, Thamchaipenet A, Gustafsson C, Fu H, Betlach M, Ashley G. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. Proc. Natl. Acad. Sci. USA. 1996; 96:1846–1851. [PubMed: 10051557]
19. Xue Q, Ashley G, Hutchinson C, Santi D. A multiplasmid approach to preparing large libraries of polyketides. Proc. Natl. Acad. Sci. USA. 1999; 96:11740–11745. [PubMed: 10518520]
20. Ruan X, Pereda A, Stassi DL, et al. Acyltransferase domain substitutions in erythromycin polyketide synthase yield novel erythromycin derivatives. J. Bacteriology. 1997; 179:6416–25.
21. Kapur S, Chen AY, Cane DE, Khosla C. Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. Proc. Natl. Acad. Sci. USA. 2010; 107:22066–22071. [PubMed: 21127271]

22. Tran L, Broadhurst WR, Tosin M, Cavalli A, Weissman KJ. Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases. Chem. Biol. 2010; 17:705–716. [PubMed: 20659683]
23. Chandran SS, Menzella HG, Carney JR, Santi DV. Activating Hybrid Modular Interfaces in Synthetic Polyketide Synthases by Cassette Replacement of Ketosynthase Domains. Chem. Biol. 2006; 13:469–474. [PubMed: 16720267]
24. Naruse N, Tenmkyo O, Kawano K, Tomita K, Ohgusa N, Miyaki T, Konishi M, Oki T. Fluvirucins A1, A2, B1, B2, B3, B4 and B5, new antibiotics active against influenza A virus. I. Production, isolation, chemical properties and biological activities. J. Antibiot. 1991; 44:733–740. [PubMed: 1880063]
25. Naruse N, Tsuno T, Sawada Y, Konishi M, Oki T. Fluvirucins A1, A2, B1, B2, B3, B4 and B5, new antibiotics active against influenza A virus. II. Structure determination. J. Antibiot. 1991; 44:741–755. [PubMed: 1880064]
26. Naruse N, Konishi M, Oki T. Fluvirucins A1, A2, B1, B2, B3, B4 and B5, new antibiotics active against influenza A virus. III. The stereochemistry and absolute configuration of fluvirucin A1. J. Antibiot. 1991; 44:756–761. [PubMed: 1880065]
27. Kohli RM, Walsh CT. Enzymology of acyl chain macrocyclization in natural product biosynthesis. Chem. Commun. 2003; 3:297–307.
28. Parenty A, Moreau X, Campagne JM. Macrolactonizations in the total synthesis of natural products. Chem. Rev. 2006; 106:911–939. [PubMed: 16522013]
29. Moutevelis-Minakakis P, Neokosmidi A, Filippakou M, Stephens D, Dennis EA, Kokotos G. Synthesis of lipophilic 2-oxoamides based on γ-aminobutyric and δ-aminovaleric analogues and their activity against phospholipase A2. J. Pept. Sci. 2007; 13:634–641. [PubMed: 17631670]
30. Ogasawara Y, Katayama K, Minami A, Otsuka M, Eguchi T, Kakinuma K. Cloning, Sequencing, and Functional Analysis of the Biosynthetic Gene Cluster of Macrolactam Antibiotic Vicenistatin in *Streptomyces halstedii* . Chem. Biol. 2004; 11:79–86. [PubMed: 15112997]
31. Yadav G, Gokhale RS, Mohanty D. SEARCHPKS: a program for detection and analysis of polyketide synthase domains. Nucleic Acids Res. 2003; 31:3654–3658. [PubMed: 12824387]
32. Keatinge-Clay AT. A Tylosin Ketoreductase Reveals How Chirality Is Determined in Polyketides. Chemistry & Biology. 2007; 14:898–908. [PubMed: 17719489]
33. Ogasawara Y, Liu H-W. Biosynthetic studies of aziridine formation in azicemicins. J. Am.Chem.Soc. 2009; 131:18066–18068. [PubMed: 19928906]
34. Irwin DM. Evolution of an active-site codon in serine proteases. Nature. 1998; 336:429–430. [PubMed: 3194031]
35. Witzowski A, Naggert J, Witkowska HE, Randhawa ZI, Smith S. Utilization of an active serine 101- cysteine mutant to demonstrate the proximity of the catalytic serine 101 and histidine 237 residues in thioesterase II. J. Biol. Chem. 1992; 267:18488–18492. [PubMed: 1526985]
36. Witzowski A, Witkowska HE, Smith S. Reengineering the specificity of a serine active-site enzyme. Two active-site mutations convert a hydrolase to a transferase. J. Biol. Chem. 1994; 269:379–383. [PubMed: 8276823]
37. Chakrabarti R, Schutt CE. The enhancement of PCR amplification by low molecular-weight sulfones. Gene. 2001; 274:293–298. [PubMed: 11675022]
38. SangJoon, Mo, et al. Biosynthesis of the Allylmalonyl-CoA Extender Unit for the FK506 Polyketide Synthase Proceeds through a Dedicated Polyketide Synthase and Facilitates the Mutasynthesis of Analogues. J. Am. Chem. Soc. 2011; 133:976–985. [PubMed: 21175203]

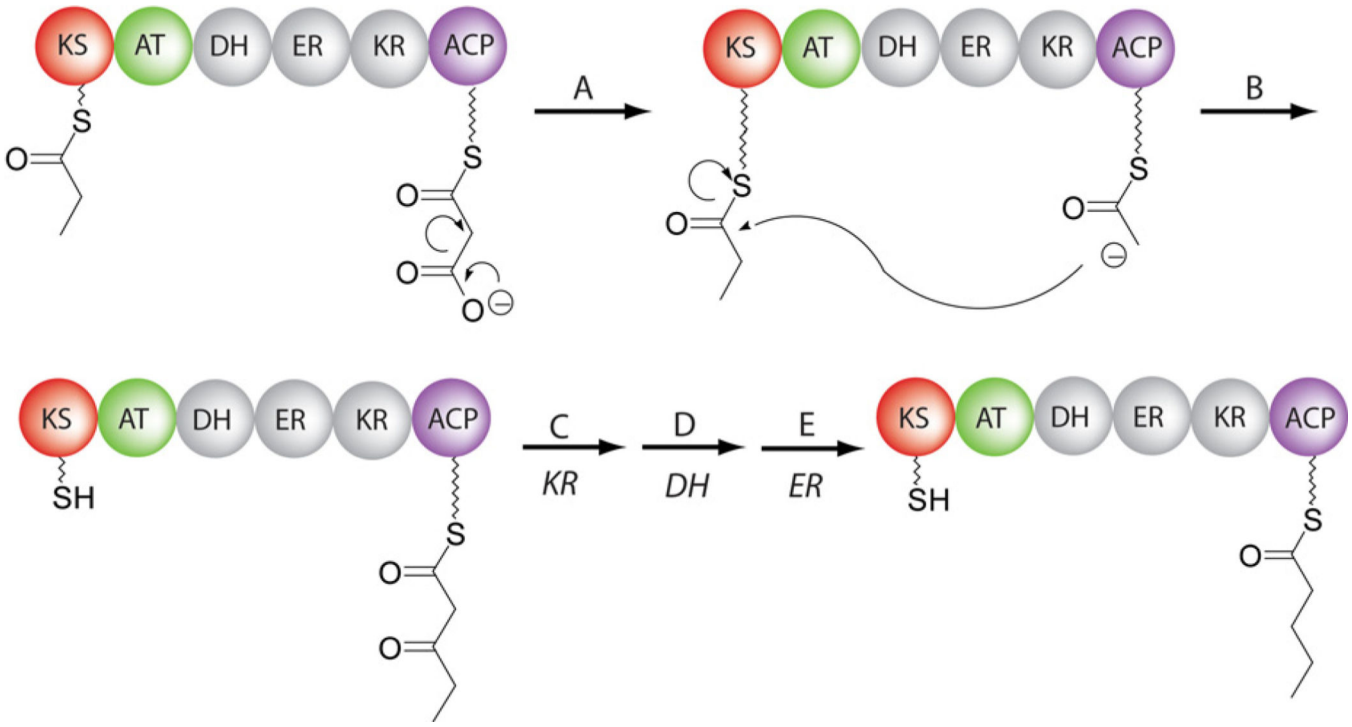


Figure 1.

Proposed mechanism for polyketide and fatty acid formation in a modular PKS and FAS, respectively (A) KS-mediated decarboxylation of ACP-bound malonate forms an ACP-bound enolate (B) Claisen-like condensation between the KS-bound chain and extender unit produces an ACP-bound β-ketothioester. (C) KR domain reduces the β-ketothioester to a β-hydroxythioester. (D) DH domain dehydrates β-hydroxythioester to an enoyl thioester. (E) ER domain reduces the enoyl group to form a saturated acyl-chain. See text for abbreviations.

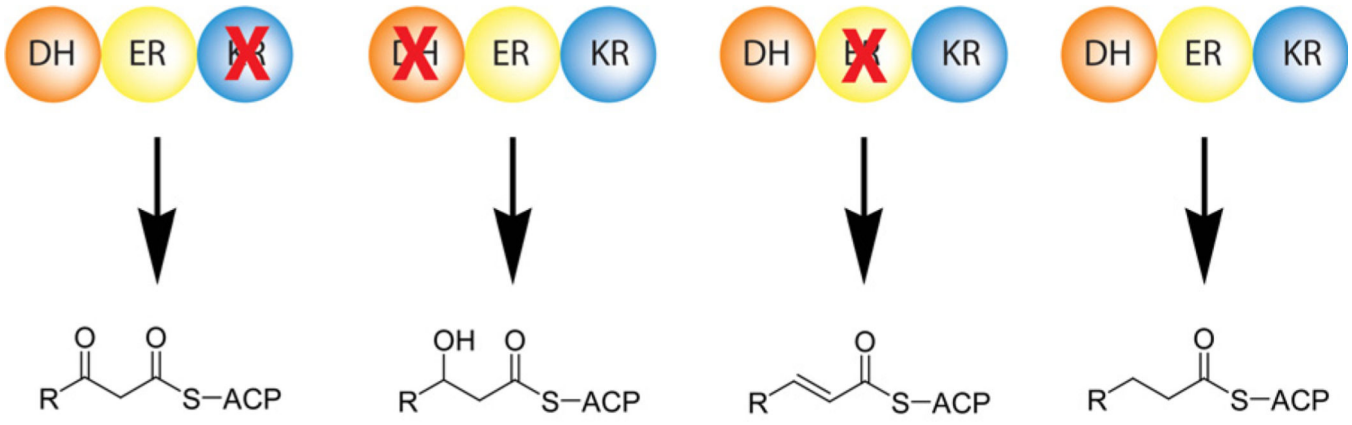


Figure 2.

Schematic diagram of a mutational inactivation strategy for production of novel compounds form fatty acid-like PKS modules. Alterations in chemical structure are designed at the genetic level.

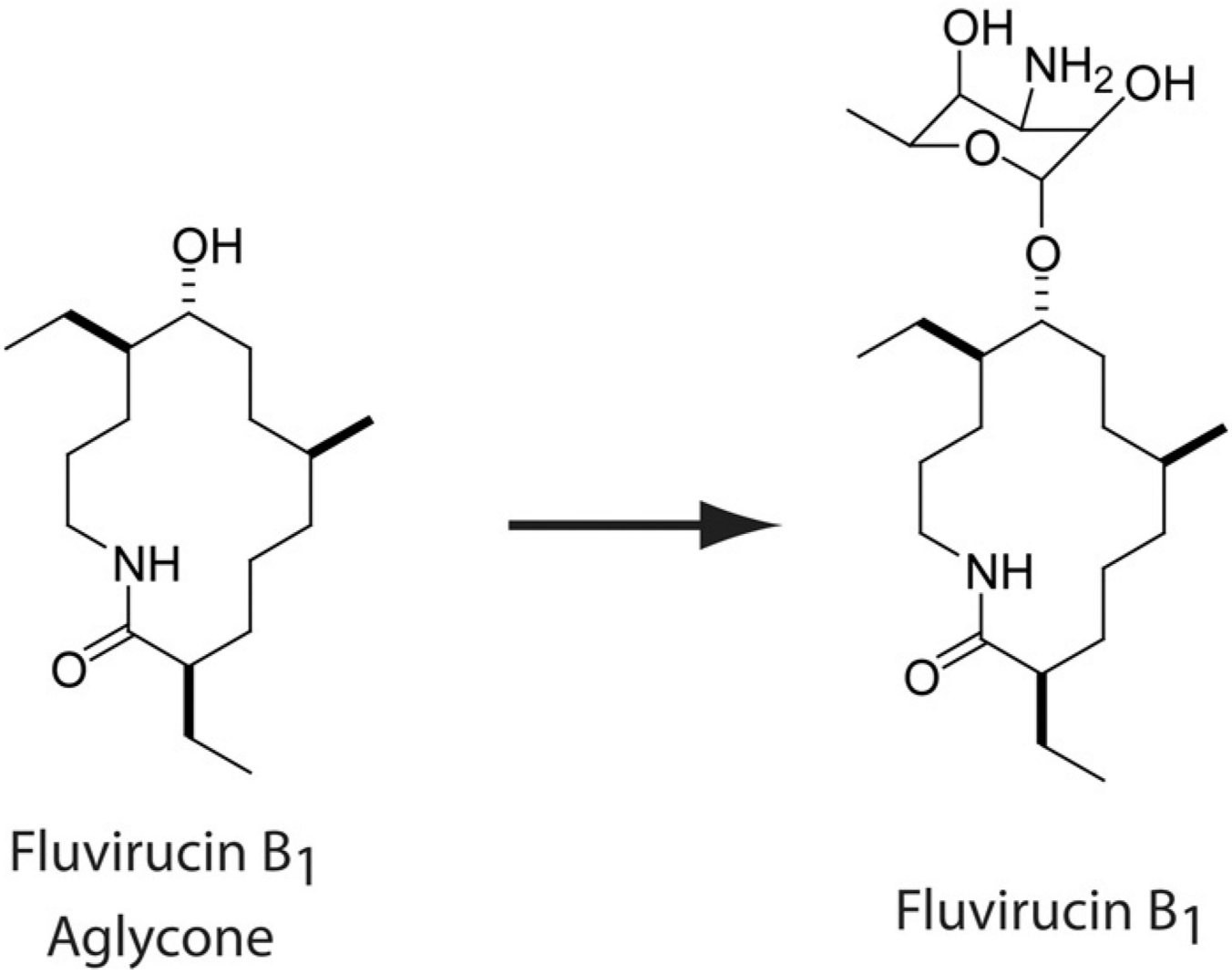
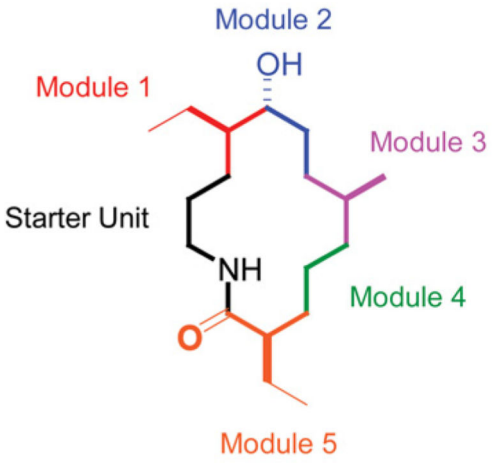


Figure 3.
Structure of fluvirucin B_1 aglycone and fluvirucin B_1.



	Predicted AT Specificity	Predicted Tailoring Domains
Module 1	ethylmalony-CoA	KR, DH, ER
Module 2	malonyl-CoA	KR
Module 3	methylmalonyl-CoA	KR, DH, ER
Module 4	malonyl-CoA	KR, DH, ER
Module 5	ethylmalonyl-CoA	KR, DH, ER

Figure 4.
Predicted tailoring domains and AT selectivities for each fluvirucin B1 synthase module based on the fluvirucin B1 aglycone structure.

A

Summary of the *flu* Genes and Comparison with Database

Protein	Amino Acids	Proposed function	Sequence similarity (Protein, Origin)	Identity/Similarity	Accession No.
FluD	506	MFS-type efflux pump	Efflux pump/drug resistance transporter, *Streptomyces sp. MP39-85*	49%/67%	ACO94494.1
FluE	299	TetR family transcriptional regulator	transcriptional regulator of TetR family, *Stackebrandtia nassauensis*	37%/50%	YP_003510265.1
FluF	425	glycosyltransferase	Glycosyltransferase, *Streptomyces sp. Tu21*	46%/67%	ABQ28821.1
FluG	351	LysR family transcriptional regulator	putative transcriptional regulator of LysR family, *Streptomyces himastatinicus* ATCC 53653	37%/51%	ZP_07292079.1
FluH	317	proline iminopeptidase	prolyl aminopeptidase, *Streptomyces halstedii*	67%/82%	ZP_07299775.1
FluI	420	decarboxylase	Decarboxylase, *Streptomyces halstedii*	62%/75%	BAD08372.1
FluJ	448	crotonyl-CoA reductase	crotonyl-CoA reductase *Streptomyces violaceusniger Tu 4113*	74%/85%	ZP_07607278.1
FluA	3969	PKS modules 1 and 2	lasalocid modular polyketide synthase , *Streptomyces lasaliensis*	53%/65%	CAQ64687.1
FluB	4375	PKS modules 3 and 4	lasalocid modular polyketide synthase , *Streptomyces lasaliensis*	55%/67%	CAQ64687.1
FluC	2396	PKS module 5	polyketide synthase, *Streptomyces graminofaciens*	53%/66%	BAJ16468.1

B

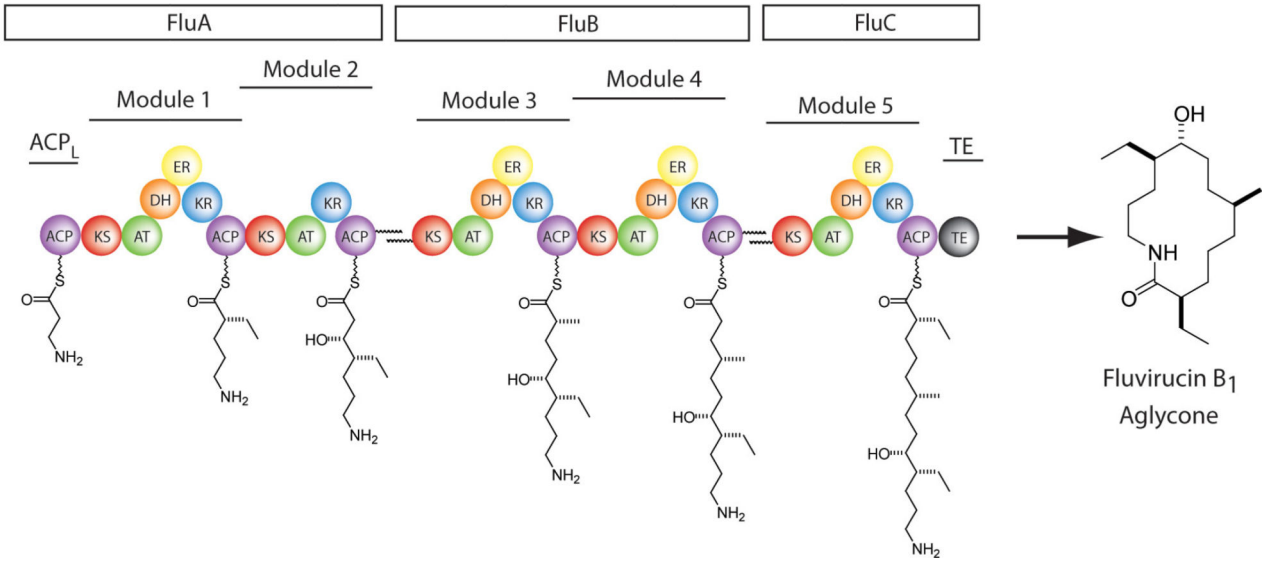


Figure 5.

Gene organization for the fluvirucin B_1 PKS cluster. **A**: Gene cluster organization and gene products identified along with comparison to known homologs. **B**: Schematic diagram for the putative fluvirucin B_1 PKS. The assembly consists of five extender modules flanked by an N-terminal loading ACP and C-terminal thioesterase (TE) domain. FluA consists of the loading ACP, module 1 and module 2. FluB consists of modules 3 and 4. FluC consists of module 5 and the TE domain. See text for abbreviations.

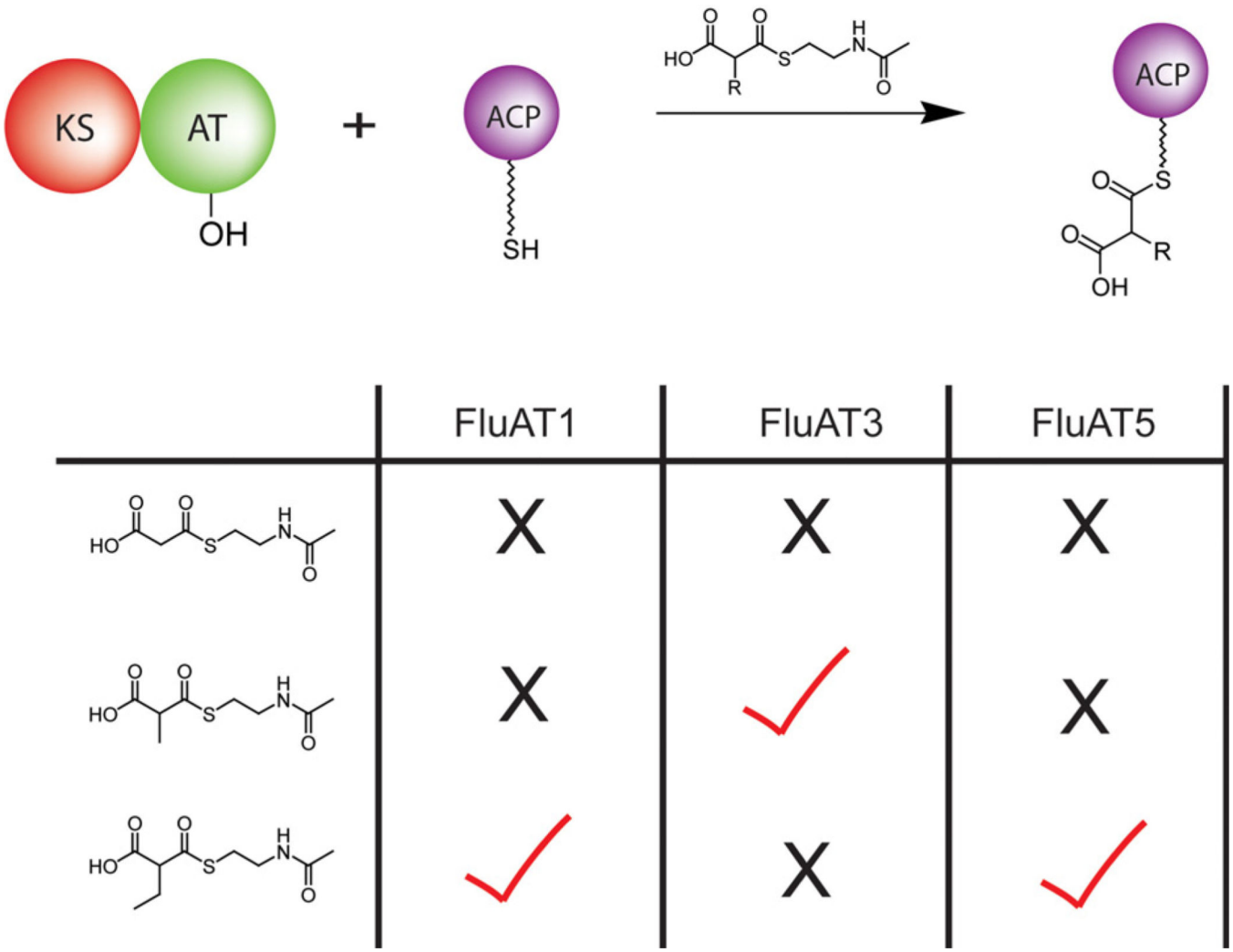


Figure 6.

Schematic diagram of and observed results for the AT substrate selectivity studies of module 1, 3, and 5. Red checks indicate that the substrate shown on the left is transferred from the indicated AT to the ACP domain while a black X indicated that no substrate transfer was observed. See supporting information for raw LC-MS data. FluATX = fluvirucin AT domain from module X.

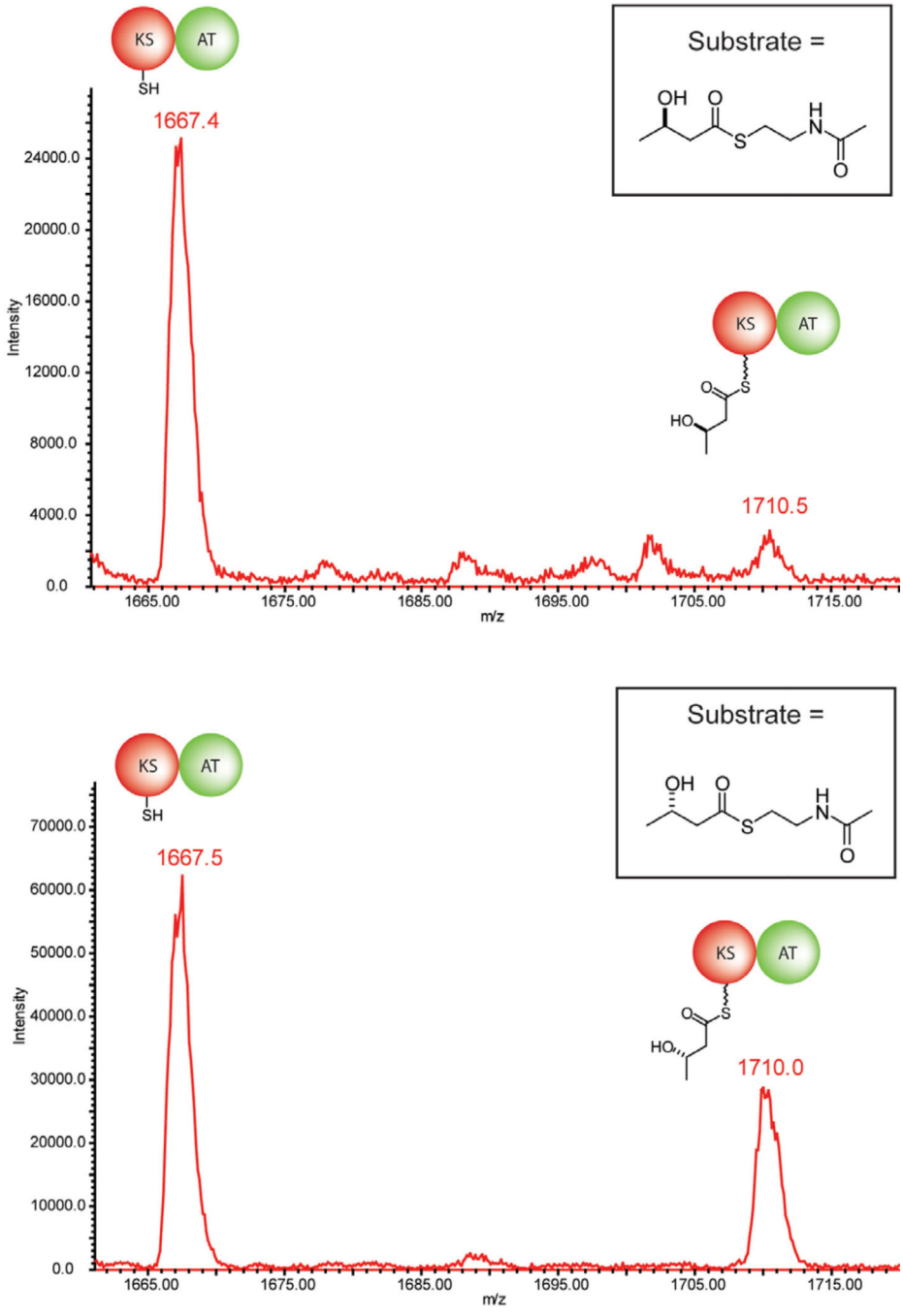


Figure 7.

LC-MS data for module 3 KS-acylation with (*R*)-3-hydroxybutyryl-SNAc (Top) and (*S*)-3-hydroxybutyryl-SNAc (Bottom). Only the (*S*)-isomer is accepted by module 3 KS as is expected from the fluvirucin B_1 structure. Peaks are labeled with the corresponding acylated or unacylated KSAT3 didomain. All peaks are M/Z = +2

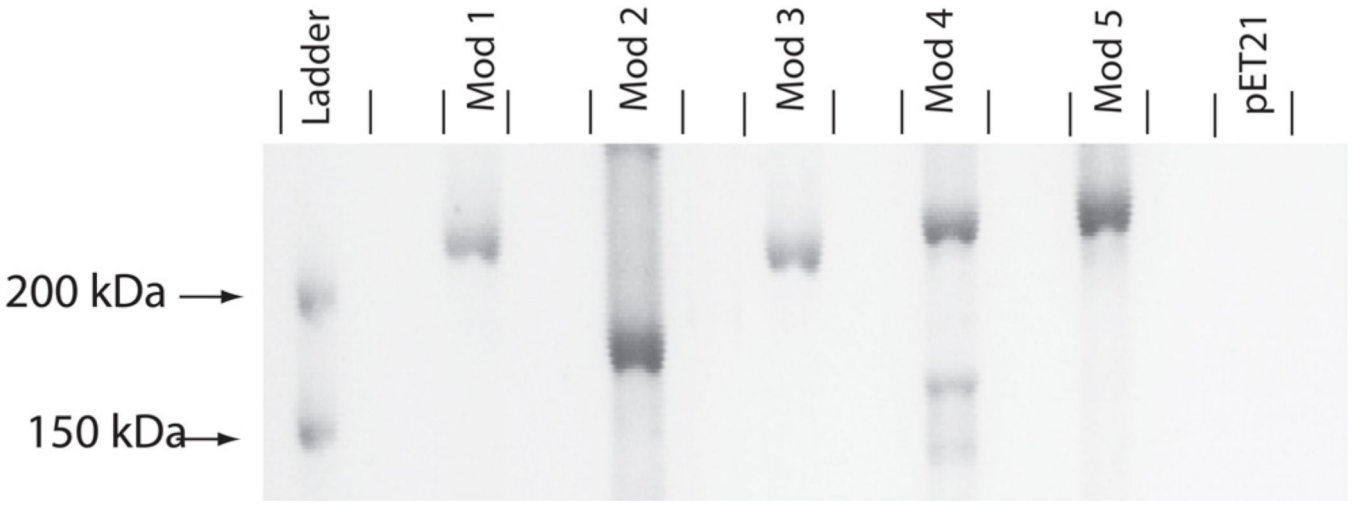


Figure 8.

PAGE analysis of fluvirucin modules overexpressed in *E. coli* following Ni-NTA affinity purification. Lanes are marked with the corresponding protein or blank pET21 vector. Mod = Module. Approximate protein molecular weights: Module 1 = 230 kDa, Module 2 = 185 kDa, Module 3 = 220 kDa, Module 4 = 240 kDa, Module 5 = 254 kDa. % Acrylamide = 7.0