CHEMBIOCHEM

Supporting Information

Characterization of the Bafilomycin Biosynthetic Gene Cluster from *Streptomyces lohii*

Wei Zhang,[a] Jeffrey L. Fortman,[b] Jacob C. Carlson,[b] Jinyong Yan,[a] Yi Liu,[a] Fali Bai,[a] Wenna Guan,[a] Junyong Jia,[b] Teatulohi Matainaho,[c] David H. Sherman,*[b] and Shengying Li*[a]

cbic_201200743_sm_miscellaneous_information.pdf

Supporting Information

Characterization of the bafilomycin biosynthesis gene cluster from *Streptomyces lohii*

Wei Zhang [a]‡, J. L. Fortman [b]‡, Jacob C. Carlson [b], Jinyong Yan[a], Yi Liu[a], Fali Bai[a], Wenna Guan[a], Junyong Jia [b], Teatulohi Matainaho[c] , David H. Sherman*[b] , and Shengying Li*[a]

[a] *Dr.* W. Zhang, *Dr. J. Yan, Y. Liu, F. Bai, Dr. W. Guan, Professor S. Li*
Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Energy Genetics, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences
189 Songling Road, Qingdao, Shandong 266101 (P. R. China)
Fax: (+86)-532-8066-2778
E-mail: lishengying@qibebt.ac.cn

[b] *Dr. J. L. Fortman, Dr. J.C. Carlson, Dr. J. Jia, Professor D. H. Sherman*
Life Sciences Institute, Departments of Medicinal Chemistry, Chemistry, and Microbiology and Immunology
University of Michigan
210 Washtenaw Avenue, Ann Arbor, MI 48109-2216 (USA)
Fax: (+1)-734-615-3641
E-mail: davidhs@umich.edu

[c] *Professor Teatulohi Matainaho*
Department of Pharmacology
University of Papua New Guinea
Port Morseby (Papua New Guinea)

[‡] *These authors contribute equally to this work*

Genomic Library Construction

Genomic DNA was isolated from a 50 ml culture using a modified protocol from Practical *Streptomyces* genetics (Kieser 2000). Briefly, 4.2g of mycelia were suspended in 20ml of TE buffer containing 2mg/ml lysozyme. The cells were then incubated at 37°C for two hours. Then, 4.8ml of 500mM EDTA and 500µl of proteinase K (20mg/ml) were added. The cells were incubated at 37°C for an additional 5 minutes. 10% SDS was then added to a final concentration of 1% and the cells were incubated for 1 hour at 37°C. This incubation was followed by three equal volume extractions with phenol:chloroform (1:1) and one equal volume extraction with chloroform. The DNA was precipitated from the aqueous fraction by adding 0.7 volumes of isopropanol. Genomic DNA was collected by spooling onto a Pasteur pipette and washed with 70% ethanol. Finally, the DNA was resuspended in 1ml TE. This genomic DNA was utilized for both library construction and initial probe generation.

For library construction, genomic DNA of approximately 40kb was fractionated from the genomic extract by sucrose gradient centrifugation. It was then blunt-end repaired and ligated into a CopyControl fosmid vector (Epicentre) using the manufacturer's protocol. The vectors were packaged into phages using the MaxPlax system protocol and used to transfect *E. coli* EPI300 cells (Epicentre). Based on the average *Streptomyces* genome size of 8Mb and an average insert size of 40kb, it was determined that approximately 1,200 colonies were needed to provide a library of approximately 6x coverage. The library was generated in a 96-well format for ease of replication. Sterile glycerol (50%) was added to the cultures the blocks were sealed and stored at –80°C. This library was used for all subsequent in situ hybridization experiments.

Probe Generation

The first round of library screening was intended to isolate any type I polyketide synthase (PKS) genes present in this organism. To accomplish this, degenerate PCR primers, 4UU (5' – MGI GAR GCI YTI CAR ATG GAY CCI CAR CAR MG) and 5LL (5' – GGR TCN CCI ARI TGI GTI CCI GTI CCR TGI GC), designed to a conserved portion of the type I polyketide ketosynthase (KS) domain were utilized in a PCR reaction using genomic DNA as the template. The reactions were carried out using taq polymerase and a Bio-Rad iCycler thermocycler using the following parameters: 95°C for 3:00 followed by thirty cycles of melting at 95°C (0:30), an annealing gradient from 48°C-71°C (1:00) and extension at 72°C (3:00). A final 7:00 extension cycle at 72°C completed the run. A small amount of product was amplified at adjacent temperature points 52.4°C and 49.7°C. This product was purified by gel electrophoresis and used as template in a subsequent reaction, using the following parameters: 95°C for 3:00 followed by thirty cycles of melting at 95°C (0:30), annealing 52.4°C (1:00) and extension at 72°C (3:00) and a final extension cycle at 72°C (7:00). The amplified product was purified by gel electrophoresis and extracted from the gel using Wizard® SV Gel and PCR Clean-up System (Promega) following the manufacturers instructions. To establish the identity and desired heterogeneity of this probe, the PCR product was ligated into the cloning vector pGEM T-Easy (Lucigen). Sequencing confirmed our PCR probe to be homologous to KS gene fragments. The PCR product was then radiolabeled with P^{32} dCTP using a Radprime Labeling Kit (Invitrogen) following the manufacturer's instructions. The initial round of probing revealed a portion of the target gene cluster. Therefore, all subsequent probing experiments were performed to find cosmids containing overlapping DNA sequence. These experiments were carried out using the same general procedure. However, the primers for these experiments were designed based on DNA sequence identified by the previous round of probing and sequencing. The size range of subsequent probes was ~400-800bp. The PCR parameters for these reactions were as follows: 95°C for 3:00 followed by 30 cycles of melting at 95°C (1:00), annealing 60°C (1:00) and extension at 72°C (1:00), followed by a final extension cycle at 72°C (5:00).

Library Screening

Initial candidates were selected from a library of 1152 clones by probing with a radiolabelled ketosynthase gene fragment.

Fosmid library clones were grown in 96-well blocks and replica plated onto sterile nylon or nitrocellulose membranes placed on LB agar containing chloramphenicol (12.5 µg/ml) and grown at 37°C overnight for subsequent hybridization. Cell lysis, DNA denaturing, cross-linking and in situ hybridization were carried out using standard protocols (Sambrook, 2001). The hybridized membranes were exposed to x-ray film overnight. Clones correlating to images on the film were cultured from the parent library and subjected to further analyses.

Restriction Digestion and Southern Hybridization

The hybridization analysis revealed 12 clones positive for type I KS genes. These were analyzed by restriction mapping with BamHI and BglII, and the digested DNA was then transferred to a nitrocellulose membrane. A Southern hybridization analysis was carried out using standard protocols (Sambrook, 2001) and the mixed KS probe used for the initial colony blot hybridization.

Sub-Cloning and Sequencing

Candidate fosmids were shotgun sequenced. Fosmids were randomly sheared by nebulization (90 seconds at 10psi) in an AeroMist Treatment reservoir (Inhalation Plastics Inc.). The DNA was blunt-end repaired utilizing an End-It™ DNA end repair kit (Epicentre) and size selected by gel electrophoresis. Fragments of 3-5kb were then ligated into the plasmid pSMART (Lucigen) and electroporated into E. cloni cells (Lucigen) following the manufacturer's instructions. Subclone plasmid DNA was isolated from 2ml cultures using standard alkaline lysis protocols (Sambrook, 2001). DNA sequencing was conducted by the sequencing core facility at the University of Michigan and Agencourt® Bioscience Corporation. Primers SL1 and SR2 (Lucigen) were utilized by both facilities. Fosmid end-sequencing was conducted exclusively at the University of Michigan core facility, using pCC1™/pEpiFOS™ forward and reverse sequencing primers (Epicentre).

Sequence was assembled using Seqman software (DNAStar). Translation analysis was carried out using the MacVector open reading frame analysis function.

Knock outs

Gene disruption in *Streptomyces lohii* was performed using a modified version of the REDIRECT system (Gust, 2003). A suicide knockout vector was created using fosmid 8D3 from the cluster sequencing effort by replacing the gene of interest with a selectable disruption cassette. This fosmid was transferred by interspecies conjugation from *E. coli* into our host strain *Streptomyces lohii*. Due to initially low efficiences of conjugation, we modified our disruption cassette to include the entire origin of transfer region of the RK2 plasmid, instead of the minimized region used in the published REDIRECT system. To generate this new disruption cassette, the full origin of transfer from the pKC1139 plasmid was PCR amplified, treated with T4 PNK and blunt-end ligated into pIJ773 vector that had been digested with BsiWI and treated with Klenow fragment and antartic phosphatase. The orientation of insertion and sequence fidelity of the resulting cassette were verified by DNA sequencing.

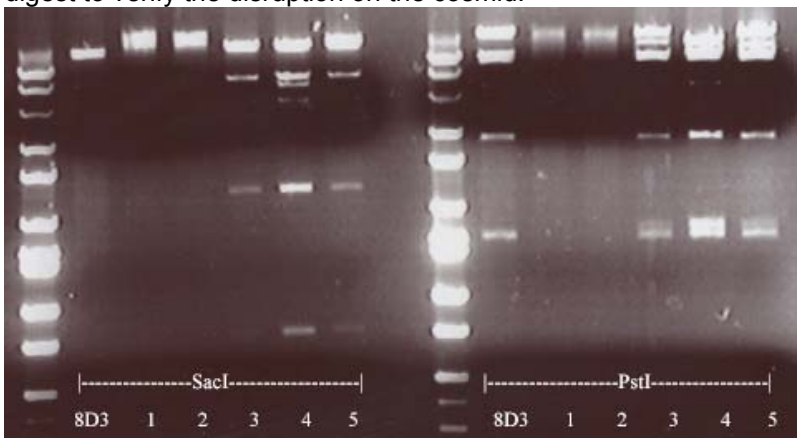
Primers used to amplify oriT for insertion into pIJ773 (the ends contain restriction sites that were relevant for a different cloning project but were not employed in this work):
JC57: AAAAAACACCTGCCGGTCTTGCCTTGCTCGTCG
JC58: AAAAAAAGGCCTTTTCCGCTGCATAACCCTGC

The BafAIII ORF contains two PKS modules, and the region targeted for replacement was the entirety of the second module, using the BafAIII stop codon as the 3' boundary and the predicted beginning of Module 2 as the 5' boundary. When this region of fosmid 8D3 is replaced by the REDIRECT cassette, there remains approximately 17 kb of native sequence on each side of the insert for homologous recombination into the chromosome.

These primers were:

JC213: CGGCCCCGGAAGCGGGCACGACAACGGCCGCCACACCCATTCCGGGGATCCGTCGACC
JC214: GCGGACGGGGGTGAGGTGCGCGGACCGGTGCCCCCGTCATGTAGGCTGGAGCTGCTTC

To create the cosmid suicide vector, the disruption cassette was amplified from the (pIJ773 + full oriT) template using primers JC213/214, which contain 40 bp of homology to the targeted region on either side. This PCR product was gel purified and transformed into freshly prepared electrocompetent, arabinose induced cells of Epi300 bearing pKD119 (expresses the lambda red recombinase proteins) and cosmid 8D3. The recovered outgrowth of this transformation was plated to LB with apramycin (50 ug/mL) and chloramphenicol (12.5 ug/mL). Apramycin resistant colonies were picked and grown in 2 mL liquid cultures, miniprepped, and analyzed by *SacI* and *PstI* digest to verify the disruption on the cosmid.

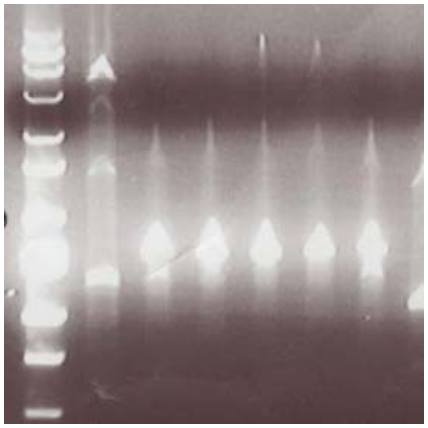


Cosmid 8D3	BafAIII	Cosmid 8D3	BafAIII
SacI Digest		PstI Digest	
14178	13284	17770	11889
13284	7790	9401	9401
7790	6875	8823	8823
5772	5772	6272	6272
5612	5612	3638	3638
410	2294	1541	1624
391	752		1541
8	410		
	391		
	8		

Figure S1. Restriction analysis of the knockout construct made from fosmid 8D3. Five disruption clones were evaluated in comparison to the wild-type cosmid. The anticipated digest patterns for the wild-type and KO are shown above. The digest pattern of clone 4 was identified as correct, and this clone was sequenced using primers JC57 and JC213 to verify the integrity of the disruption cassette.

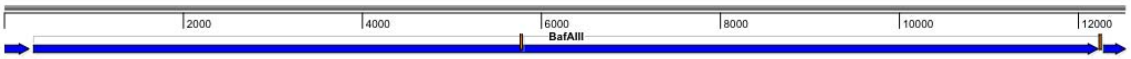
The validated KO construct was electroporated into ET12567 + pUZ8002 and the resulting strain was used in interspecies conjugation with spores of *S. lohii* as previously described (Kieser, 2000). Following conjugation the mating mixture was spread onto MS agar (90 mm, 25 mL plate) and after 20 hours growth at 30°C was flooded with 500 ug apr and 1250 ug nalidixic acid in 1 mL sterile water. After 3 days of continued incubation at 30°C, colonies appeared and were picked to 2xYT with the same antibiotics. 12 resulting colonies were then restreaked to fresh 2xYT with the same antibiotics, and from the resulting streaks, 5 single colonies (one from each of 5 streaks), were used for downstream genetic analysis.

Strains were inoculated from agar plates into 2xYT broth with 30 ug/mL apramycin and grown for 3 days at 30°C with shaking. 2 mL of culture were used for isolation of genomic DNA as previously described (Kieser, 2000).

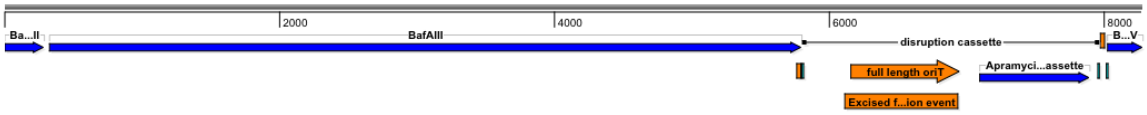


Expected product for wild-type: 6560 bp
Expected product for disruption: 2303 bp
Expected product for disruption and recombinant excision of oriT fragment: 1473 bp
JC215: CGCTCGCGGACTTCCTGGGC
JC216: CGCGTACGCACACGGCGTGG

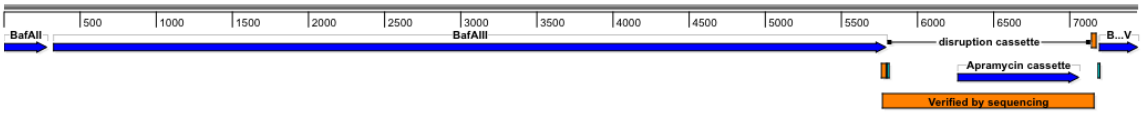
Figure S2. Genetic confirmation of the disruption was accomplished by PCR amplification using genomic DNA from mutants as template. Diagnostic primers were chosen to border the region targeted for deletion (JC215/JC216). Products corresponding to the wild-type locus were lost in the mutants as expected, but the new product resulting from the presence of the disruption cassette was of a size smaller than anticipated. Sequencing of this product revealed that the disruption cassette was present, but a recombination event between the original minimal oriT and the introduced full length oriT excised a 830 bp segment from the cassette. This excision was completely within the oriT and did not affect the apramycin resistance marker of the cassette, or any unintended flanking sequences from the bafilomycin cluster.



Wild type bafilomycin cluster BafAIII region above. The small orange boxes represent the boundaries of the region to be disrupted.



Anticipated disruption of BafAIII with full length oriT



Observed disruption of BafAIII with oriT partially excised

Figure S3. Illustrations of the knockouts described herein. The sequence context of the disruption was verified by sequencing of the PCR products. All 5 clones were verified to contain this disruption.

To analyze metabolite profiles, the five mutant strains and one wild-type strain were inoculated from agar plates into 2xYT with 30 ug/mL apramycin and grown for 3 days at 30°C with shaking. 300 uL of seed culture was transferred to 12 mL of Md2 media supplemented with 30 ug/mL apramycin and grown for 3 days at 30°C with shaking. The cultures were centrifuged to remove cells and the supernatant was extracted twice with an equal volume of dichloromethane. The organic extracts from each culture were pooled and dried in vacuo, and redissolved in DMSO. Extracts were analyzed by LC-MS on a C-18 column eluted with a gradient of acetonitrile in water with 0.1% formic acid.

Analytical methods

LC-MS analysis was performed on a Shimadzu 2010 EV ESI spectrometer by using an XBridgeTM C18 3.5 um, 2.1 x 50 mm column with a MeCN and H_2O solvent system supplemented with formic acid (0.1% v/v), and elution was with a linear gradient of increasing MeCN at a flow rate of 0.2 mL/min. Bafilomycin metabolites were detected by UV absorbance at 288 nm and their identity was confirmed by mass spectrometry.

Comparison of Bafilomycin and Concanamycin Biosynthesis

During the isolation of the bafilomycin biosynthesis gene cluster, the genetic architecture for the biosynthesis of the 18-membered plecomacrolide concanamycin was published (Haydock, 2005). While these two natural products have obvious differences, their similarities are striking (Figure S4). The isobutyrate-specific AT-ACP loading domain is absent from the concanamycin cluster. Instead, this cluster contains an additional two module ORF and incorporates the more common acetate starter unit. The only other difference in the genes associated with the assembly of the macrolactone cores is the addition of an entire module to the penultimate ORF (*conE*) in the PKS portion of the pathway. This additional module incorporates the extremely unusual ethylmalonate extender unit. The result is the 18-membered ring, an ethyl side chain and an additional hydroxyl group. Outside of these differences, the PKS portions of these two clusters are remarkably similar. Three of the five megasynthases from the bafilomycin cluster have homologs in the concanamycin cluster that are identical in their domain organization. Indeed, even the other two ORFs share their domain arrangement outside of the differences listed above.

Both clusters contain genes involved in methoxymalonate biosynthesis. The fact that each of the ORFs from this pathway is roughly 70% identical at the amino acid level suggests that these two pathways are closely related. However, while the bafilomycin contains these genes in what appears to be a discreet operon, they are discontinuous in the concanamycin cluster. Genes encoding sugar biosynthesis and attachment enzymes are in an operon that lies between the homologs of *bafE* and *bafF*.

Concanamycin is tailored by addition of a carbamoylated sugar on the C21 hydroxy, while bafilomycin is tailored with a fumarate or fumarate-cyclopentenone moiety. Both the sugar biosynthesis and carbamoyl transferase genes are found within the concanamycin biosynthesis cluster. However, the 18-membered plecomacrolide virsutomycin A is composed of the same macrolactone core as concanamycin, but is tailored with a fumarate and 2-amino-3-hydroxycyclopent-2-enone like bafilomycin (Omura, 1983). This suggests that the evolution between bafilomycin

and concanamycin may have occurred in a number of steps. Other 16-membered plecomacrolides show variability in their starter units. The common polyketide starters of acetate and propionate are implicated by the existence of bafilomycin analogs with methyl or ethyl side chains at C23. The pentadiene side chain of micromonosporolide A suggests the addition of an acetate-loading module followed by two full extension modules. Taken together, the loss of the cyclopentenone and changes in the loading domains suggest a high level of plasticity near the 5' end of the cluster. The structure of the 16-membered plecomacrolide formamicin along with the differences that mark the generation of the 18-membered ring, suggests there is also plasticity between *bafAIII* and *bafAIV*. The placement of the sugar operon also hints at a possible integration site. However, there are no other sequenced plecomacrolide clusters to support this observation.

Taken as a whole the differences between these two biosynthesis gene clusters suggest that they are highly related, both to each other and most likely to the group as a whole. The addition of the bafilomycin biosynthesis cluster to the PKS body of knowledge provides an important piece of data for the eventual elucidation of the origin of the plecomacrolides.

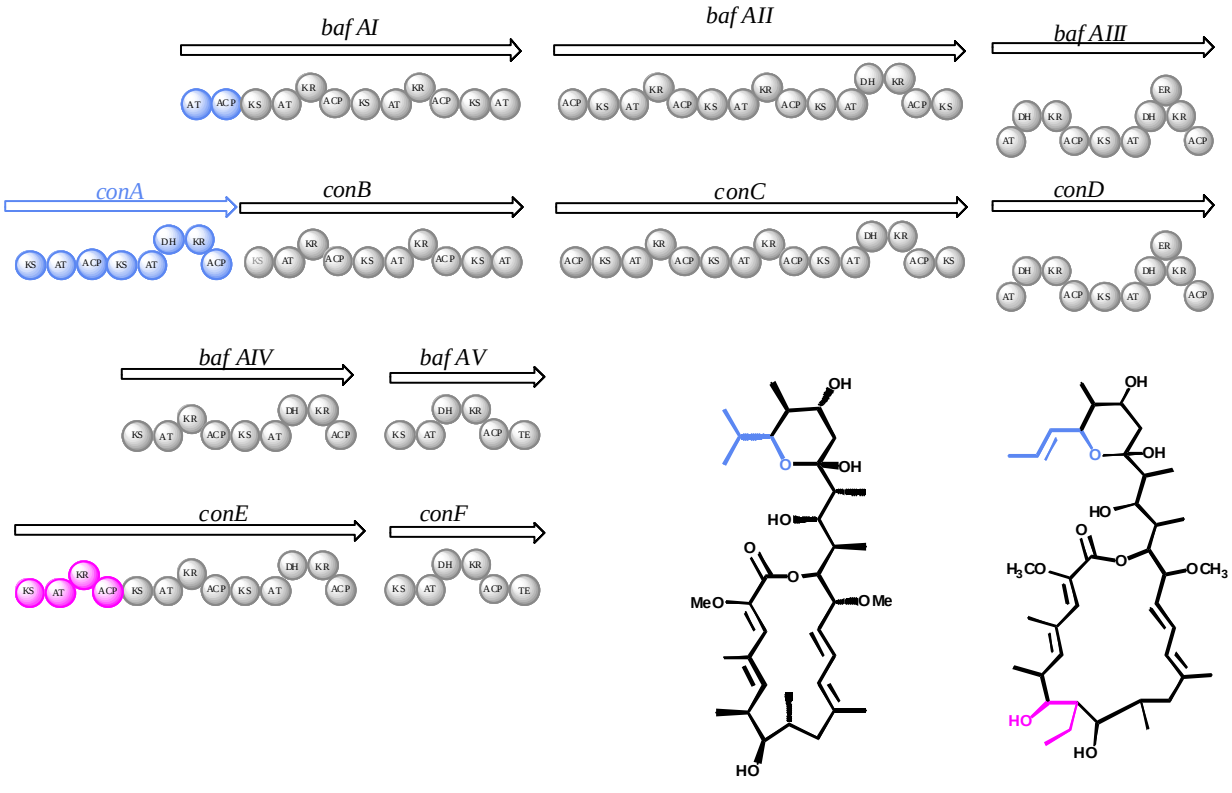
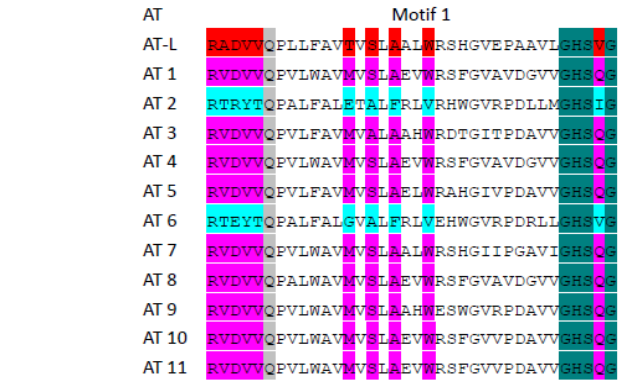


Figure S4. Differences in the genetic architecture and chemical structure of bafilomycin and concanamycin. Differences are highlighted.

Figure S5. Protein sequence alignments of Bafilomycin ATs (the two conserved sequence motifs used to predict substrate specificity are boxed)



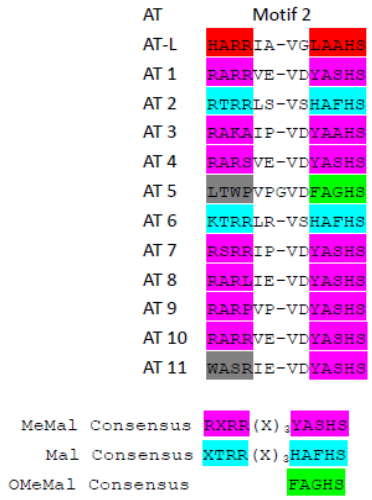


Figure S6. AT Substrate Specificity Consensus

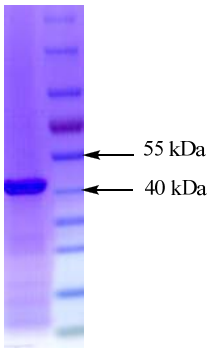


Figure S7. SDS-PAGE analysis for BafZ purified from *E. coli* BL21(DE3)

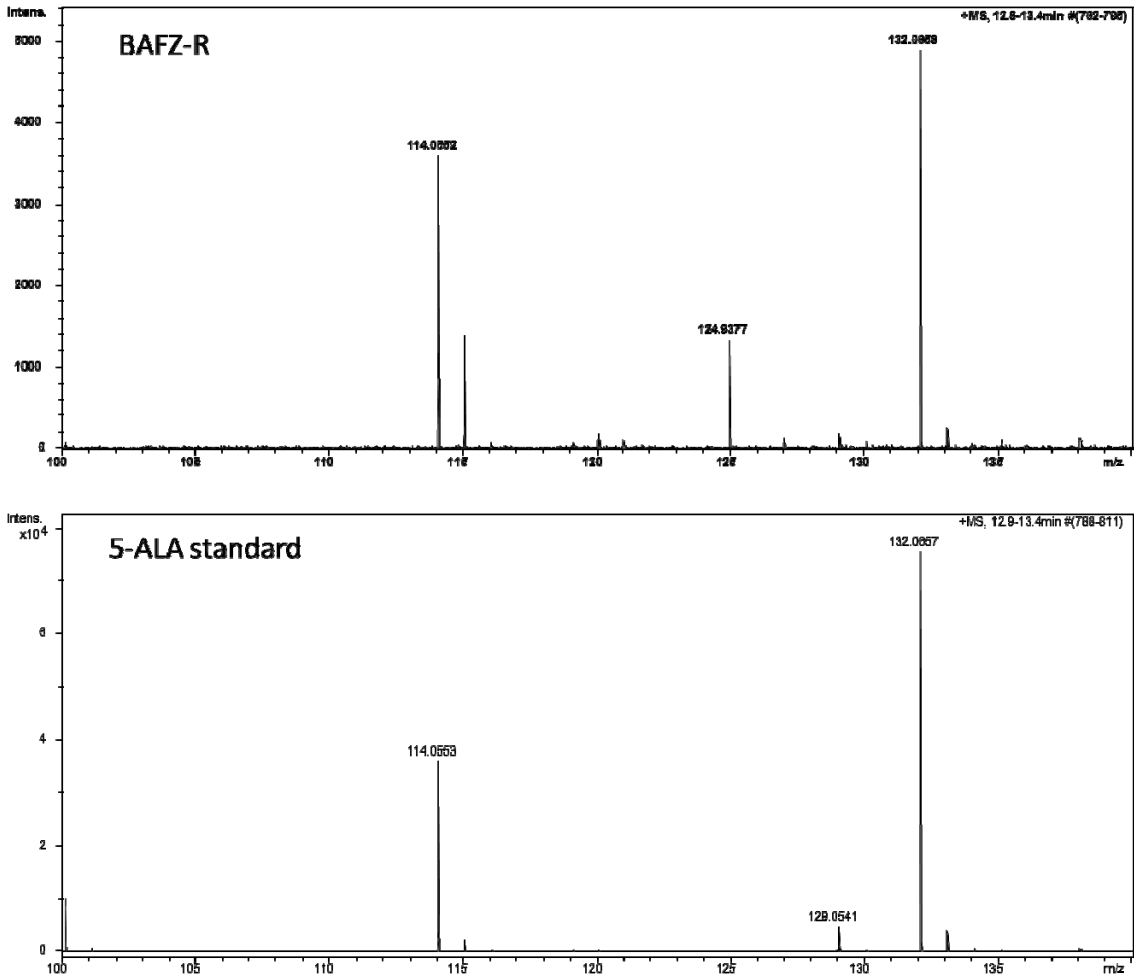


Figure S8. High Resolution Mass spectra measured during HR-LCMS. (the upper panel shows the mass spectrum of the 5-ALA product from BafZ catalyzed reaction; the lower panel shows the mass spectrum of the 5-ALA authentic standard.)

References

T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, 2000. Practical *Streptomyces* Genetics. The John Innes Foundation, Norwich, UK.

Sambrook, J., and D. W. Russell. 2001. Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

B. Gust, G. L. Challis, K. Fowler, T. Kieser, K. F. Chater, *Proc. Natl. Acad. Sci. U S A.* **2003**, *100*, 1541

S. F. Haydock, A. N. Appleyard, T. Mironenko, J. Lester, N. Scott, P. F. Leadlay, *Microbiol.*, **2005**, *151*, 3161

S. Omura, N. Imamura, K. Hinotozawa, K. Otoguro, G. Lukas, R. Faghih, R. Tolmann, B. H. Arison, J. L. Smith, *J. Antibiot.*, **1983**, *36*, 1783