Biosynthetic Studies of Aziridine Formation in Azicemicins

Yasushi Ogasawara and Hung-wen Liu*

Division of Medicinal Chemistry, College of Pharmacy, and Department of Chemistry and Biochemistry, University of Texas, Austin, Texas 78712

Supporting Information

Materials. Isotopically-labeled precursors used in the feeding experiments include sodium [1-^{13}C]acetate (99 atom % ^{13}C, Isotec (Champaign, IL)), sodium [1,2-$^{13}C_2$]acetate (99 atom % ^{13}C, Isotec), L-[3,3,3-$^{2}H_3$]alanine (99.8 atom % D, CDN Isotopes (Quebec, Canada)), D,L-[2,3,3-$^{2}H_3$]serine (98 atom % D, Cambridge Isotope Laboratories (Andover, MA)), and D,L-[2,3,3-$^{2}H_3$]aspartic acid (99 atom % D, CDN Isotopes). Enzymes and molecular weight standards used for the cloning experiments were purchased from Invitrogen (Carlsbad, CA) or New England Biolabs (Ipswich, MA). Kits for DNA gel extraction and minipreps are products of Qiagen (Valencia, CA). Antibiotics and chemicals were acquired from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA). Deep vent (exo-) DNA polymerase and KOD hot start DNA polymerase were obtained from New England Biolabs and Novagen (Madison, WI), respectively. Growth medium components were purchased from Becton Dickinson (Sparks, MD). Oligonucleotide primers were ordered from Integrated DNA Technologies (Coralville, IA) or Invitrogen.

General. NMR spectra were recorded on a Varian Unity 500 MHz spectrometer at the NMR facility of the Department of Chemistry and Biochemistry, University of Texas at Austin, and chemical shifts (δ in parts per million) are reported relative to that of the solvent peak ($\delta = 7.26$ for $CDCl_3$ in ^{1}H NMR spectra). Genetic manipulations of *E. coli* were performed according to standard protocols. DNA sequencing was performed by the Core Facilities of the Institute of Cellular and Molecular Biology, University of Texas at Austin.

Culture Conditions and Isolation of Azicemicin A. The azicemicin-producing bacteria, *Kibdelosporangium* sp. MJ126-NF4, was kindly provided by Dr. Yuzuru Akamatsu (Microbial Chemistry Research Center, Tokyo, Japan), and was maintained on yeast-starch agar plates (0.2% yeast

extract, 1% soluble starch, 1.5% agar (all w/v)). To isolate azicemicins, *Kibdelosporangium* sp. MJ126-NF4 cells were first grown in seed medium (2% galactose, 2% dextrin, 1% Bacto-Soytone, 0.5% corn steep liquor, 1% glycerol, 0.2% $(NH_4)_2SO_4$, and 0.2% $CaCO_3$, pH 7.4). An aliquot (20 mL) of the 4-day preculture was transferred into 1 L of the production medium (2% dextrin, 1% Bacto-Soytone, 2% glycerol, 0.3% yeast extract, 0.2% $(NH_4)_2SO_4$, and 0.2% $CaCO_3$, pH 7.4) and the culture was incubated with shaking at 30 °C for 5 days. Supplementation with isotopically-labeled compounds was carried out as follows. Equal portions of sodium [1-^{13}C]acetate (a total of 1 g) was added to a growth culture (1 L) by pulse feeding after 42, 64, 86, and 108 h of inoculation. Similarly, portions of sodium [1,2-$^{13}C_2$]acetate (total 1 g) and D,L-[2,3,3-$^2H_3$]serine (total 0.5 g) were separately added to cultures by pulse feeding 48, 64, 84, and 102 h after inoculation. For L-[3,3,3-$^2H_3$]alanine (total 0.5 g) and D,L-[2,3,3-$^2H_3$]aspartic acid (total 0.5 g), the feedings were carried out 46, 70, and 92 h after inoculation. After five days, the culture broths were centrifuged at 9000 × *g* for 30 min. The supernatants were extracted three times with ethyl acetate and the organic extracts were collected. After removal of the organic solvent by evaporation, azicemicin A was purified through several silica gel chromatography steps using a combination of two different solvent systems ($CHCl_3$:MeOH = 10:1, or ethyl acetate). The structure of the purified product was verified by NMR and MS analysis. (Table S1, Figure S1)

Table S1: ^{13}C NMR results from the incorporation of ^{13}C-labeled acetate.

Position	Chemical Shift	[1-^{13}C]acetate relative ^{13}C intensities	[1,2-$^{13}C_2$]acetate coupling constant (Hz)
1	206.5	3.2	38.3
2	47.2	1.0	*
3	70.8	2.0	48.0
4	41.5	1.0	31.5
4a	41.1	3.1	31.5
5	37.5	1.0	41.2
6	200.7	2.6	41.2
6a	105.7	0.9	64.5
7	164.6	2.3	64.5
7a	110.2	1.1	63.6
8	150.7	1.8	63.6
9	133.5	0.9	70.8
10	154.2	1.4	70.8
11	98.3	1.0	61.2
11a	131.5	1.7	61.2
12	143.9	1.0	78.0
12a	123.8	1.8	78.0
12b	75.5	0.9	38.3
1'	44.0	1.0	48.0
2'	31.7	1.1	*
9-OMe	61.0	0.9	-
12-OMe	62.3	0.8	-
N-Me	46.9	0.8	-

* Signal intensities increased

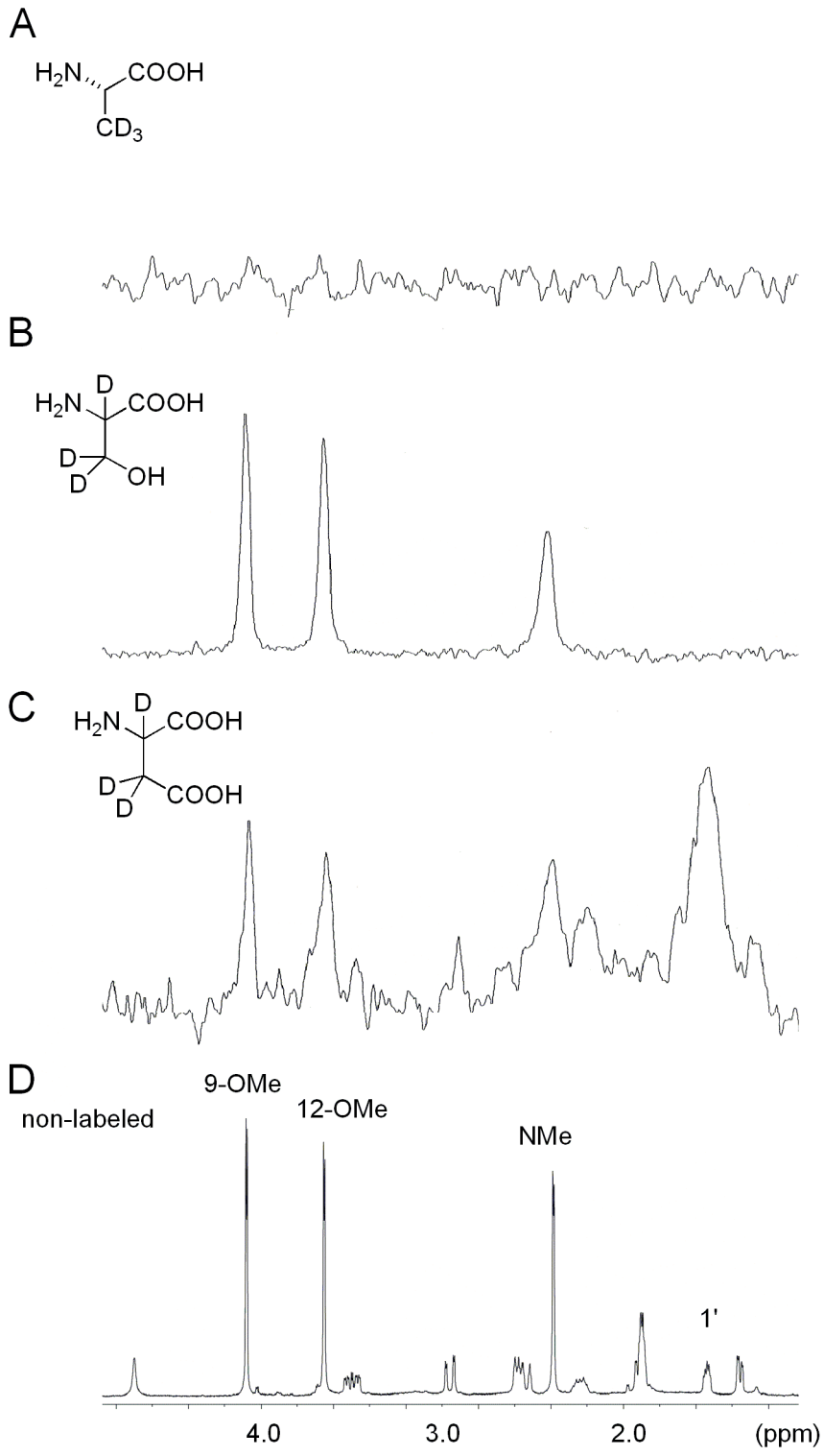

Figure S1: Feeding experiments with deuterium labeled amino acids. The ^{2}H-NMR spectrum (77 MHz, $CHCl_3$) of labeled azicemicin A with deuterated (A) alanine,(B) serine, and (C) aspartate. (D) ^{1}H NMR spectrum (500 MHz, $CDCl_3$) of non-labeled azicemicin A.

Identification and Cloning of the Azicemicin Biosynthetic Gene Cluster. The genomic DNA of *Kibdelosporangium* sp. was isolated according to a published procedure.[1] For construction of the genomic library, the chromosomal DNA was partially digested with *Sau*3AI, isolated by ethanol precipitation, and ligated into the pOJ446 cosmid that had been subjected to linearization by *Hpa*I, calf intestinal alkaline phosphatase (CIAP) treatment, and digestion with *Bam*HI. Gigapack III Gold packaging extract (Stratagene, La Jolla, CA) was used to perform the *in vitro* packaging based on the manufacturer's protocol. Degenerate PCR primers were designed based on multiple sequence alignments of several known KSα's from actinomycetes: KSaF2, 5′-GGSTGCACSTCVGGSMTSGAC-3′; KSaR1, 5′-CCGATSGCRCCSAGSGAGTG-3′. (S = G + C, V = A + C + G, M = A + C, R = A + G) The PCR based screening of the genomic library was carried out with the following conditions: 1 cycle at 95 °C for 2 min, 30 cycles of 95 °C for 20 sec, 51 °C for 10 sec, and 70 °C for 30 sec, using deep vent (exo-) DNA polymerase.

Sequencing and Homology Analysis. Sequencing was performed using a capillary-based AB 3700 DNA analyzer, and the DNA sequencing data were assembled and analyzed using the Vector NTI Suite program (Version 10, Invitrogen). Open reading frame (ORF) assignments were made with the assistance of FramePlot[2] and the Vector NTI software. The resulting ORFs were analyzed by BLAST[3]. These genes and their proposed functions are shown in Figure 1 and Table S2.

Table S2: Proposed function of each ORF in the azicemicin biosynthetic gene cluster.

ORF	Proposed Function	Typical Homology	Identity/Similarity (%)
azicR2	regulator	two-component system response regulator, *Streptomyces coelicolor*	64 / 75
azicR1	regulator	two-component system sensor kinase, *Streptomyces coelicolor*	46 / 56
azicO1	dehydrogenase	dehydrogenase, *Streptomyces steffisburgensis*	57 / 68
azicZ	aspartate racemase	aspartate racemase, *Polaromonas sp.*	43 / 61
azicU1	unknown	secreted protein, *Streptomyces coelicolor*	32 / 46
azicU2	unknown	hypothetical protein, *Streptomyces coelicolor*	59 / 69
azicO2	P450	P450 hydroxylase, *Streptomyces atroolivaceus*	43 / 62
azicO3	oxygenase	FAD-dependent oxygenase, SimA7, *Streptomyces antibioticus*	64 / 77
azicY	drug-resistance transporter	transmembrane efflux protein, *Streptomyces hygroscopicus*	40 / 53
azicX	PMP oxidase	pyridoxamine 5'-phosphate oxidase, ActVA2, *Streptomyces coelicolor*	49 / 64
azicU3	unknown	hypothetical protein, *Nocardia farcinica*	37 / 57
azicR3	regulator	repressor-response regulator, med-ORF30, *Streptomyces sp.* AM-7161	57 / 74
azicO4	oxygenase	FAD-dependent monooxygenase, UrdE, *Streptomyces fradiae*	65 / 75
azicA	ketosynthase α subunit	ketosynthase α subunit, PgaA, *Streptomyces sp.* PGA64	74 / 87
azicB	ketosynthase β subunit	ketosynthase β subunit, SimA2, *Streptomyces antibioticus*	71 / 83
azicC	acyl carrier protein	minimal PKS acyl carrier protein, UrdC, *Streptomyces fradiae*	69 / 81
azicD	polyketide ketoreductase	polyketide ketoreductase, PgaD, *Streptomyces sp.* PGA64	82 / 87
azicE	cyclase	bifunctional cyclase/aromatase, PgaL, *Streptomyces sp.* PGA64	71 / 77
azicO5	oxygenase	monooxygenase, UrdM, *Streptomyces fradiae*	53 / 65
azicF	reductase	reductase, LanV, *Streptomyces cyanogenus*	59 / 76
azicG	decarboxylase	decarboxylase, PgaI, *Streptomyces sp*. PGA64	79 / 89
azicH	FMN-reductase	NADPH-dependent FMN reductase, UrdO, *Streptomyces fradiae*	57 / 68
azicO6	oxygenase	$FMNH_2$-utilizing oxygenase, *Streptomyces avermitilis*	48 / 61
azicO7	oxygenase	polyketide synthesis hydroxylase, TcmG, *Bacillus thuringiensis*	36 / 50
azicU4	unknown	hypothetical protein, *Mycobacterium vanbaalenii*	22 / 36
azicU5	unknown	hypothetical protein, *Alkaliphilus metalliredigenes*	27 / 46
azicU6	unknown	hypothetical protein, *Mycobacterium sp.*	36 / 51
azicI	fatty acid desaturase	delta fatty acid desaturase, *Streptomyces coelicolor*	53 / 66
azicU7	unknown	hypothetical protein, *Thermobifida fusca*	35 / 49
azicR5	regulator	two-component system response regulator, *Streptomyces hygroscopicus*	59 / 72
azicR4	regulator	two component sensor kinase, *Streptomyces hygroscopicus*	37 / 48
azicK	acyltransferase	acyltransferase, AknF, *Streptomyces galilaeus*	49 / 63
azicJ	cyclase	cyclase, JadI, *Streptomyces venezuelae*	57 / 67
azicU8	unknown	hypothetical protein, *Frankia alni*	50 / 63
azicL	methyltransferase	O-methyltransferase, *Streptomyces kanamyceticus*	45 / 64
azicU9	unknown	hypothetical protein, *Salinispora tropica*	39 / 50
azicM	adenylyltransferase	amino acid adenylation, *Salinispora tropica*	55 / 68
azicN	decarboxylase	PLP-dependent decarboxylase, VinO, *Streptomyces halstedii*	41 / 54
azicP	acyl carrier protein	PKS/NRPS acyl-carrier-protein, VinL, *Streptomyces halstedii*	38 / 63
azicQ	acyl-CoA dehydrogenase	acyl-CoA dehydrogenase, *Salinispora tropica*	57 / 69
azicS	acyl-CoA dehydrogenase	acyl-CoA dehydrogenase, *Salinispora tropica*	57 / 69
azicT	phosphopantetheinyltransferase	phosphopantetheinyl transferase, *Salinispora tropica*	60 / 68
azicU10	unknown	hypothetical protein, *Salinispora tropica*	36 / 48
azicV	adenylyltransferase	amino acid adenylation, *Salinispora arenicola*	56 / 68
azicW	drug-resistance transporter	drug-registance transporter, *Nocardia farcinica*	69 / 81

Expression of AzicM and AzicV in *E. coli*. The *azicM* and *azicV* genes were each PCR-amplified from appropriate cosmids using primers with engineered *Nde*I and *Eco*RI restriction sites at the 5′ and 3′ termini, respectively. The gene for *azicM* was amplified using PCR primers AzicM-*N*-(*Nde*I): 5′-CGCATATGACCCTGTTACACCACATCG-3′ and AzicM-*C*-(*EcoR*I): 5′-CGGCGAATTCTCATCCA-CCGTCGGCC-3′, with the following PCR conditions: 1 cycle at 95°C for 2 min, 30 cycles of 95 °C for 20 sec, 51 °C for 10 sec, and 70 °C for 30 sec, using KOD hot start DNA polymerase. The primers for *azicV* were AzicV-N-(NdeI): 5′-CATATGCAGACCATGTACGACTGG-3′ and AzicV-C-(EcoRI): 5′-TAAGTGACCTCGCTCGAATTCAGGC-3′, and the PCR conditions were 1 cycle at 95 °C for 2 min, 30 cycles of 95 °C for 20 sec, 50 °C for 10 sec, 70 °C for 30 sec, with KOD hot start DNA polymerase. The introduced *Nde*I and *Eco*RI restriction sites are underlined. The PCR-amplified genes were purified, digested with the appropriate restriction enzymes, and ligated into the pET28b(+) vector digested with the same enzymes. The sequences of each construct were confirmed by DNA sequencing. The resulting plasmids, pET28_azicM and pET28_azicV, were used to transform *E. coli* BL21 star (DE3) strains (Invitrogen) for protein over-expression. An overnight culture of each transformant was grown in 2 L of LB medium supplemented with 30 μg/mL kanamycin. The culture was grown at 37 °C until an OD_{600} of 0.6 was reached, after which 0.3 mM isopropyl β-D-thiogalactoside (IPTG) was added, and the culture was grown at 18 °C for 12 additional hours. The harvested wet-cells were suspended in 60 mL of buffer (50 mM Tris·HCl, 300 mM NaCl, pH 8.0) and disrupted by sonication. The cell lysate was clarified by centrifugation (15000 rpm, 30 min), and AzicM and AzicV were purified from the supernatant using Ni-NTA column chromatography.

Assays for AzicM and AzicV. Enzyme reactions were carried out at 30 °C in the presence of 1 mM ATP, 2 mM of the appropriate amino acid, 0.5 mM $MgCl_2$, and 4 μM of either AzicM or AzicV in a total volume of 500 μL of 50 mM Tris·HCl buffer, pH 8.0. Amino acids used for the assay were purchased from Sigma-Aldrich (Milwaukee, WI) or TCI America (Portland, OR), with the exception of (*S*)- and (*R*)-aziridine-2-carboxylic acid, which were chemically synthesized using reported method.[4] After a 30 min incubation period, the reactions were quenched by boiling for 2 min. The protein precipitates were removed by centrifugation and the supernatants were transferred to a new tube and subjected to a pyrophosphate (PP_i) quantitation assay. Each 200 μL sample was diluted with an equal volume of water and then mixed with 25 μL of the molybdate reagent (2.5% $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$ in 5 M H_2SO_4), 50 μL of bisulfate solution A (10% $NaHSO_3$, 0.5% Na_2SO_3) and 25 μL of a 10% thioglycerol

solution. After 10 min, 0.4 mL of isoamyl alcohol was added, and the solutions were mixed and centrifuged for 1 min at 500 × *g*. The top layers were discarded and the bottom layers (400 μL), which contained the PP_i-molybdate complex, was transferred to a new tube. 50 μL of bisulfite solution B (1:15 dilution of bisulfate solution A), 25 μL of the thioglycerol solution, and 50 μL of ethanol were then added to this solution. After mixing, the absorbance at 575 nm of the resulting solution was recorded. The amount of PP_i generated in each reaction was calculated using a standard curve obtained with a tetrapotassium pyrophosphate standard. The relative rates of each enzyme for the various amino acid substrates were obtained by normalization with respect to the rate of the substrate with the largest activity of each enzyme (AzicM: D-aspartate, 0.028 min^{-1}, and AzicV: L-cysteine, 0.078 min^{-1},)

(1) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, England, 2000.
(2) http://watson.nih.go.jp/~jun/cgi-bin/frameplot.pl
(3) http://blast.ncbi.nlm.nih.gov/Blast.cgi
(4) Nakajima, K.; Takai, F.; Tanaka, T.; Okawa, K. *Bulletin of the Chemical Society of Japan* **1978**, *51*, 1577-1578.