C-S bond cleavage by a polyketide synthase domain

Ming Ma[a,1], Jeremy R. Lohman[a,1], Tao Liu[b,1], and Ben Shen[a,b,c,d,2]

[a]Department of Chemistry, The Scripps Research Institute, Jupiter, FL 33458; [b]Division of Pharmaceutical Sciences, University of Wisconsin-Madison, Madison, WI 53705; [c]Department of Molecular Therapeutics, The Scripps Research Institute, Jupiter, FL 33458; and [d]Natural Products Library Initiative at The Scripps Research Institute, The Scripps Research Institute, Jupiter, FL 33458

Edited by Jerrold Meinwald, Cornell University, Ithaca, NY, and approved July 15, 2015 (received for review April 29, 2015)

Leinamycin (LNM) is a sulfur-containing antitumor antibiotic featuring an unusual 1,3-dioxo-1,2-dithiolane moiety that is spiro-fused to a thiazole-containing 18-membered lactam ring. The 1,3-dioxo-1,2-dithiolane moiety is essential for LNM's antitumor activity, by virtue of its ability to generate an episulfonium ion intermediate capable of alkylating DNA. We have previously cloned and sequenced the *lnm* gene cluster from *Streptomyces atroolivaceus* S-140. In vivo and in vitro characterizations of the LNM biosynthetic machinery have since established that: (*i*) the 18-membered macrolactam backbone is synthesized by LnmP, LnmQ, LnmJ, LnmI, and LnmG, (*ii*) the alkyl branch at C-3 of LNM is installed by LnmK, LnmL, LnmM, and LnmF, and (*iii*) leinamycin E1 (LNM E1), bearing a thiol moiety at C-3, is the nascent product of the LNM hybrid nonribosomal peptide synthetase (NRPS)-acyltransferase (AT)-less type I polyketide synthase (PKS). Sulfur incorporation at C-3 of LNM E1, however, has not been addressed. Here we report that: (*i*) the bioinformatics analysis reveals a pyridoxal phosphate (PLP)-dependent domain, we termed cysteine lyase (SH) domain (LnmJ-SH), within PKS module-8 of LnmJ; (*ii*) the LnmJ-SH domain catalyzes C-S bond cleavage by using L-cysteine and L-cysteine *S*-modified analogs as substrates through a PLP-dependent β-elimination reaction, establishing L-cysteine as the origin of sulfur at C-3 of LNM; and (*iii*) the LnmJ-SH domain, sharing no sequence homology with any other enzymes catalyzing C-S bond cleavage, represents a new family of PKS domains that expands the chemistry and enzymology of PKSs and might be exploited to incorporate sulfur into polyketide natural products by PKS engineering.

cysteine metabolism | enzyme mechanism | genome mining | pathway engineering | sulfur metabolism

Sulfur is found in many primary metabolites, such as thiamin, biotin, molybdenum cofactors, lipoic acid, iron-sulfur clusters, and nucleosides including 2-thiocytidine, 4-thiouridine, and 2-methylthio-N^6-isopentenyl adenosine (1). There is a wealth of information on how sulfur is incorporated into these primary metabolites, the key step of which is catalyzed by a desulfurase to produce a protein-bound cysteine persulfide by using cysteine as a substrate (2–5). Sulfur is also found in many secondary metabolites (interchangeably referred to as natural products here). The major groups of sulfur-containing natural products are peptides produced by ribosome or nonribosomal peptide synthetases (NRPSs). Sulfur incorporation into these natural products from intact cysteine or methionine is well characterized. However, sulfur incorporation into other natural products remains poorly understood. Only a few cases featured by C-S bond formation or C-S bond cleavage steps were reported to date, and most of the mechanisms require further investigation (6–15).

Leinamycin (LNM) is a sulfur-containing antitumor antibiotic produced by *Streptomyces atroolivaceus* S-140 (16). The structure of LNM is characterized by an unusual 1,3-dioxo-1,2-dithiolane moiety that is spiro-fused to a thiazole-containing 18-membered lactam ring (Fig. 1). The 1,3-dioxo-1,2-dithiolane moiety is essential for LNM's antitumor activity by virtue of its ability to alkylate the N7 position of the deoxyguanosine of DNA through an episulfonium ion intermediate (17–19). We have previously cloned and sequenced the *lnm* gene cluster from *S. atroolivaceus* S-140 (20–22). In vivo and in vitro characterizations of the LNM biosynthetic machinery have since established that: (*i*) the 18-membered hybrid peptide-polyketide macrolactam backbone is synthesized by a hybrid NRPS-acyltransferase (AT)-less type I polyketide synthase (PKS), consisting of LnmQ (adenylation protein), LnmP (peptide carrier protein), LnmI (a hybrid NRPS-AT-less type I PKS), LnmJ (PKS), and LnmG (AT) (21, 23, 24); (*ii*) the alkyl branch at C-3 of LNM is installed by a novel pathway for β-alkylation in polyketide biosynthesis, featuring LnmK (AT/decarboxylase), LnmL (ACP), LnmM (hydroxymethylglutaryl-CoA synthase), and LnmF (enoyl-CoA hydratase) that act on the growing polyketide intermediate tethered to the ACP domain of PKS module-8 of LnmJ (25–27); and (*iii*) leinamycin E1 (LNM E1), bearing a thiol moiety at C-3, is the nascent product of the LNM hybrid NRPS-AT-less type I PKS (28) (Fig. 1). Although sulfur incorporation into the 1,3-dioxo-1,2-dithiolane moiety of LNM has not been addressed, the structures of biosynthetic intermediates and shunt metabolites accumulated in the SB3029 (Δ*lnmK*), SB3030 (Δ*lnmL*), and SB3031 (Δ*lnmM*) mutant strains suggest that the β-alkyl branch at C-3 is installed before the sulfur incorporation, within the PKS module-8 of LnmJ (26), and the structure of LNM E1 further suggests that one sulfur at C-3 is introduced after the β-alkyl branch formation, but before the linear polyketide is released from the PKS module-8 of LnmJ by the thioesterase (TE)-catalyzed cyclization (28). Therefore, the region of approximately 850 aa between the ACP and TE domains in PKS module-8 of LnmJ represent potential candidates responsible for the sulfur incorporation at C-3 of LNM E1 and LNM (Fig. 1).

We here report that: (*i*) the bioinformatics analysis reveals a PLP-dependent domain, we termed cysteine lyase (SH) domain (LnmJ-SH), together with a domain of unknown function (DUF), located between the ACP and TE domains in PKS module-8 of LnmJ; (*ii*) the LnmJ-SH domain catalyzes C-S bond cleavage by

Significance

Sulfur incorporation into natural products remains poorly understood except for those derived from intact cysteine or methionine. Leinamycin (LNM) is a sulfur-containing antitumor antibiotic featuring an unusual 1,3-dioxo-1,2-dithiolane moiety. A pyridoxal phosphate-dependent domain, termed cysteine lyase (SH) domain, is identified within the LNM polyketide synthase (PKS) module-8 of LnmJ. The LnmJ-SH domain catalyzes C-S bond cleavage by using L-cysteine and L-cysteine *S*-modified analogs as substrates, shares no sequence homology with any other enzymes catalyzing C-S bond cleavage, and represents a new family of PKS domains. This study establishes L-cysteine as the origin of the C-3 sulfur of LNM, expands the chemistry and enzymology of PKS, and sets the stage to incorporate sulfur into polyketide natural products by PKS engineering.

Author contributions: B.S. designed research; M.M., J.R.L., and T.L. performed research; M.M., J.R.L., T.L., and B.S. analyzed data; and M.M., J.R.L., and B.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The amino acid sequences reported in this paper have been deposited in the NCBI database. For a list of accession numbers, see *SI Appendix*.

[1]M.M., J.R.L., and T.L. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: shenb@scripps.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508437112/-/DCSupplemental.

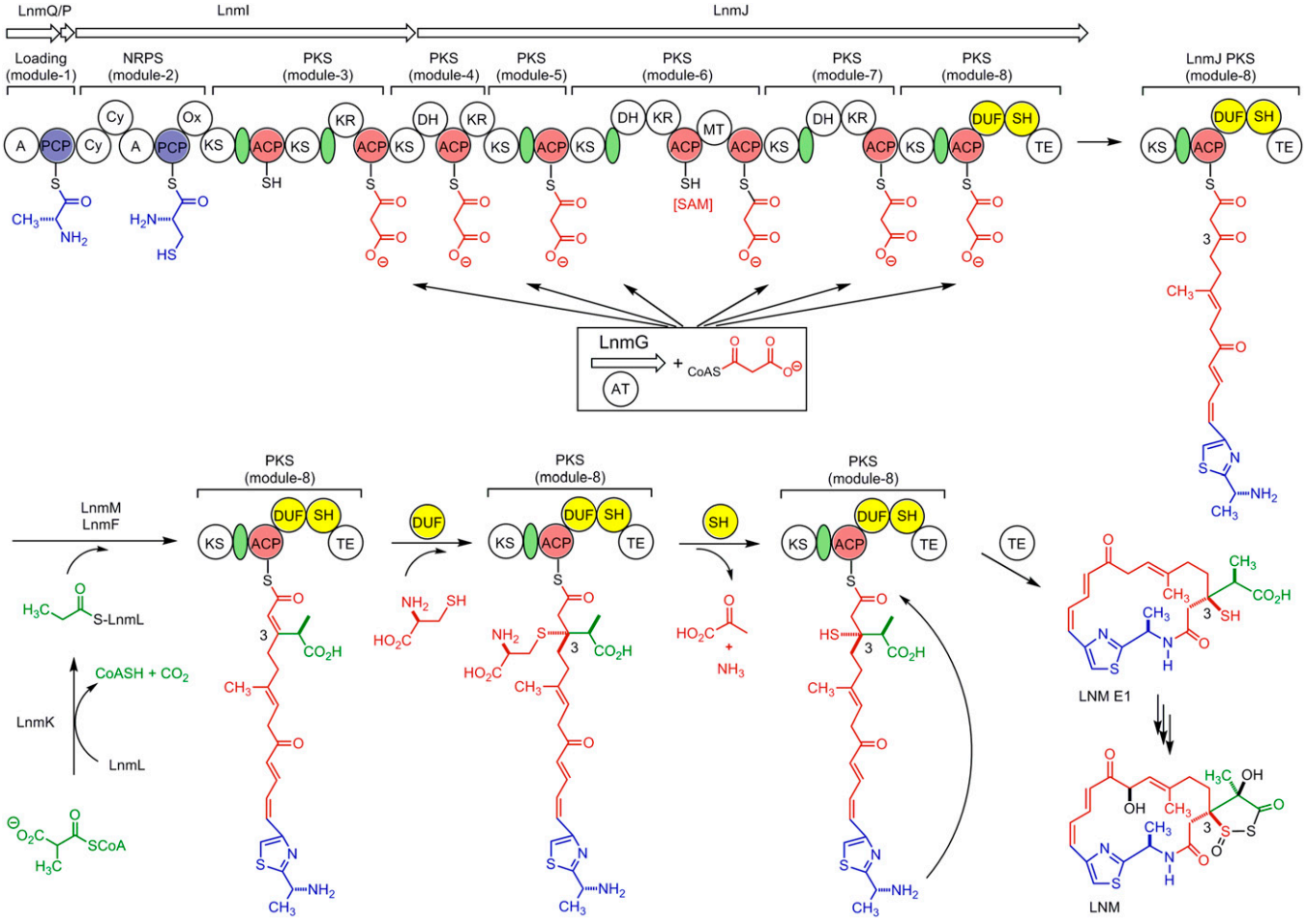

Fig. 1. Proposed biosynthetic pathway for LNM featuring (*i*) the LnmQPIJ hybrid NRPS-AT–less type I PKS with the discrete LnmG AT loading the malonyl CoA extender units to all six PKS modules (21, 23, 24), (*ii*) the LnmKLMF enzymes catalyzing introduction of the β-alkyl branch at C-3 (25-27), (*iii*) the DUF and SH domains of PKS module-8 of LnmJ catalyzing sulfur incorporation into C-3 of LNM from L-cysteine characterized in this study, and (*iv*) LNM E1 as the nascent product of the LNM hybrid NRPS-AT-lee type I PKS (28), setting the stage to investigate the tailoring steps that convert LNM E1 to LNM. Color coding indicates the moieties installed by NRPS (blue), PKS (red), β-alkyl branch (green), and other tailing enzyme (black), and the green ovals denote AT docking domains. A, adenylation domain; AT, acyltransferase; Cy, condensation/cyclization; DH, dehydratase; DUF, domain of unknown function; KR, ketoreductase; KS, ketosynthase; MT, methyltransferase; Ox, oxidation; PCP, peptidyl carrier protein; SAM, S-adenosylmethionine; SH, PLP-dependent cysteine lyase domain; TE, thioesterase.

using L-cysteine and L-cysteine *S*-modified analogs as substrates through a PLP-dependent β-elimination reaction, establishing L-cysteine as the origin of sulfur at C-3 of LNM and LNM E1; and (*iii*) the LnmJ-SH domain, sharing no sequence homology with any other enzymes catalyzing C-S bond cleavage in natural product biosynthesis, represents a new family of PKS domains that expands the chemistry and enzymology of PKSs.

Results

Bioinformatics Analysis of the PKS Module-8 of LnmJ Revealing Novel Domains. Inspired by the proposal that sulfur incorporation at C-3 takes place after β-alkyl branch formation and before release of the mature hybrid peptide-polyketide intermediate from the ACP domain by TE-catalyzed cyclization to afford LNM E1, we reexamined the PKS module-8 of LnmJ. We previously noted the atypical architecture of PKS module-8 of LnmJ but were unable to assign function to the unknown domains (21, 22, 28). Close examination of PKS module-8 of LnmJ revealed two additional domains between ACP and TE, previously not known to be associated with PKSs (Fig. 1). C-terminal to the ACP is a DUF [DUF2156 as annotated by PFAM (29), residues 6299–6563, E value 2.9×10^{-18}], with weak homology to the acyl-CoA acyltransferase superfamily [InterProScan (30); E value 2.2×10^{-6}]. The region between the DUF and TE, we now termed the SH domain, shares similarity with the family of aromatic amino acid β-eliminating lyases (ABLs) (residues 6689–6964, PFAM E value 2.3×10^{-48}). BLAST searches of the LnmJ-SH domain revealed homology to tyrosine phenol lyases (TPL) and tryptophan indole lyases (TIL), canonical members of the ABL family (Fig. 2). TPL and TIL are PLP-dependent enzymes that catalyze C_β-C_γ bond cleavage of tyrosine and tryptophan to phenol and indole, respectively (31) (*SI Appendix*, Fig. S1). Sequence alignments and molecular modeling of the selected ABLs and LnmJ-SH domain indicate conservation of all residues involved in PLP and amino acid substrate carboxylate binding (*SI Appendix*, Fig. S2).

BLAST searches also revealed homologs of LnmJ-SH (E value between 7.0×10^{-66} and 2.0×10^{-107}) within PKSs from various organisms, none of which, however, had previously been functionally annotated. The SH domain is always located immediately C-terminal to a DUF (Figs. 2 and 3*A* and *SI Appendix*, Fig. S3). Organisms other than *S. atroolivaceus* S-140 that possess PKS modules with DUF-SH domain architecture include *Actinokineospora* sp. EG49, *Aquimarina muelleri*, *Burkholderia pseudomallei* (various subspecies), *Catenulispora acidiphilia* DSM 44928, *Fulvivirga*

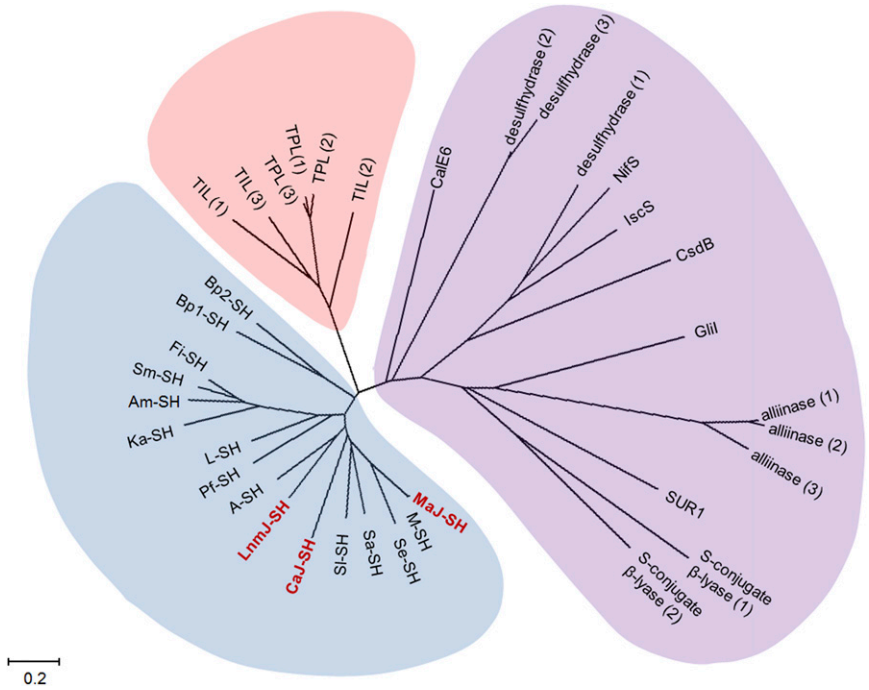

Fig. 2. The phylogenetic analysis for LnmJ-SH in comparison with selected members of ABLs and selected enzymes catalyzing C-S bond cleavage from natural product biosynthetic pathways. LnmJ-SH forms a distinct clade with functionally uncharacterized homologs, as exemplified by CaJ-SH and MaJ-SH, from various organisms (cyan). Selected enzymes catalyzing C-S bond cleavage from both primary metabolism, as exemplified by the cysteine desulfurases, cysteine desulfhydrases, and cysteine *S*-conjugate β-lyases, and natural product biosynthetic pathways, as exemplified by cysteine sulfoxide lyases, methionine γ-lyases, and SUR1, fall into another distinct clade (purple). Members of the ABLs form the third clade (orange) lying between the two clades mentioned above. See *SI Appendix* for accession numbers of the selected enzymes or domains used in this phylogenetic analysis.

imtechensis, *Kordia algicida* OT-1, *Lysobacter* sp. ATCC 53042, *Micromonospora aurantiaca* ATCC 27029, *Micromonospora sp.* L5, *Pseudoqulbenkiania ferrooxidans*, *Saccharothrix espanaensis*, *Salinispora arenicola*, *Sporocytophaga myxococcoides*, and *Streptomyces leeuwenhoekii* (Fig. 2). None of these organisms, however, is known to produce LNM or analogs.

The PLP-Dependent C-S Bond Cleavage Activity of LnmJ-SH Domain. We successfully cloned and expressed the *lnmJ-SH* fragment in *Escherichia coli* and purified the LnmJ-SH domain to homogeneity (Fig. 3*B*). The purified LnmJ-SH was yellow in color; displayed UV absorption peak at 420 nm, indicative of bound PLP; and eluted as a homotetramer upon size-exclusion chromatography (*SI Appendix*). Given the sequence homology of LnmJ-SH to TPL and TIL (*SI Appendix*, Fig. S2), we tested the cleavage activities of LnmJ-SH with L-tyrosine, L-tryptophan, D- and L-cysteine, and L-cysteine *S*-modified analogs (Fig. 3 *A* and *C* and *SI Appendix*, Fig. S1). LnmJ-SH was active with L-cysteine and all of the L-cysteine *S*-modified analogs tested, but not with L-tyrosine, L-tryptophan, and D-cysteine, as monitored by production of pyruvate [as methylquinoxalinol after derivatization by *o*-phenylenediamine (OPD); *SI Appendix*, Fig. S4]. The reactions with L-cysteine (**1**), L-cystine (**2**), and *S*-phenyl-L-cysteine (**4**) were further verified by detection of hydrogen sulfide (*SI Appendix*, Fig. S5), thiocysteine (*SI Appendix*, Fig. S6), and thiophenol (*SI Appendix*, Fig. S7), respectively.

LnmJ-SH Catalyzing C-S Bond Cleavage of Cysteine-Adduct Polyketide Substrate Mimics. To determine whether LnmJ-SH acts on the ACP-tethered polyketide intermediate (Fig. 1), we next synthesized the *N*-acetylcysteamine and pantetheine thioesters of the L-cysteine-polyketide adduct, **9** and **10,** as substrate mimics (Fig. 3*D* and *SI Appendix*, Figs. S8–S20). Upon incubation with LnmJ-SH, **9** and **10** afforded products **16** and **17**, respectively, whereas no product formation was observed from the negative controls using boiled LnmJ-SH (*SI Appendix*, Fig. S21). Compound **16** was purified from repeated reactions and analyzed by NMR and high-resolution electrospray mass spectrometry (HR-ESI-MS) (*SI Appendix*). The HR-ESI-MS analysis of **16** afforded the $[M + H]^+$ ion at *m/z* 236.0776 (calculated $[M + H]^+$ ion for $C_9H_{17}NO_2S_2$ at *m/z* 236.0778) and $[M + Na]^+$ ion at *m/z* 258.0596 (calculated $[M + Na]^+$ ion for $C_9H_{17}NO_2S_2$ at *m/z* 258.0598) (*SI Appendix*, Fig. S22). The ^1H NMR spectrum of **16** gave the resonances attributed to two coupled methylene groups at δ_H 3.36 (t, J = 6.7 Hz, H-3) and δ_H 3.04 (t, J = 6.7 Hz, H-4), one methylene group at δ_H 2.94 (s, H-6), one methyl group at δ_H 1.94 (s, H-1) and two overlapped methyl groups at δ_H 1.49 (s, H-8 and H-9). The two coupled methylene groups at δ_H 3.36 (t, J = 6.7 Hz, H-3) and δ_H 3.04 (t, J = 6.7 Hz, H-4), and the methyl group at δ_H 1.94 (s, H-1), together with the resonances at δ_C 21.1 (C-1), δ_C 172.0 (C-2), δ_C 38.7 (C-3), and δ_C 28.0 (C-4) in the ^{13}C NMR spectrum, suggested there was a S-substituted *N*-acetylcysteamine moiety in **16**. The N-acetylcysteamine moiety was confirmed by the correlations of H-1/C-2, H-3/C-2, H-3/C-4, and H-4/C-3 in the HMBC spectrum. The HMBC spectrum also showed the correlations of H-4/C-5, H-6/C-5, H-8 (9)/C-6, H-8 (9)/C-7, and H-6/C-8 (9), establishing the thioester linkage between an *N*-acetylcysteamine and a 3-substituted isovaleric acid. The substitution at C-3 of the isovaleric acid was established as a thiol group by the molecular formula of $C_9H_{17}NO_2S_2$. Thus, **16** was unambiguously identified as the C-S bond cleavage product of **9** (Fig. 3*D* and *SI Appendix*, Figs. S23–S27). The identity of **17** as the C-S bond cleavage product of **10** was similarly established (Fig. 3*D*), based primarily on the HR-ESI-MS analysis that afforded the $[M + H]^+$ ion at *m/z* 395.1670 (calculated $[M + H]^+$ ion for $C_{16}H_{30}N_2O_5S_2$ at *m/z* 395.1673) and $[M + Na]^+$ ion at *m/z* 417.1490 (calculated $[M + Na]^+$ ion for $C_{16}H_{30}N_2O_5S_2$ at *m/z* 417.1493) (*SI Appendix*, Fig. S28).

Kinetic Analysis of LnmJ-SH with L-Cysteine and L-Cysteine *S*-Modified Analogs as Substrates. Optimization of the LnmJ-SH catalyzed C-S bond cleavage reaction revealed maximal activity in 0.1 mM PLP, 20 mM KCl, and a pH of 8.5 when buffered with Tris·HCl (*SI Appendix*, Fig. S29). The pseudo-first–order kinetic parameters of LnmJ-SH toward L-cysteine (**1**) and L-cysteine *S*-modified analogs, including L-cystine (**2**), *S*-*t*-butylmercapto-L-cysteine (**3**), *S*-phenyl-L-cysteine (**4**), *S*-benzyl-L-cysteine (**5**), *S*-*t*-butyl-L-cysteine (**6**), *S*-ethyl-L-cysteine (**7**), *S*-methyl-L-cysteine (**8**), and the polyketide intermediate mimics **9** and **10** (Fig. 3), were determined by monitoring pyruvate production (Table 1 and *SI Appendix*, Fig. S4). The K_m or k_{cat} values for all substrates, except **1**, **9**, and **10,** could not be determined because of their low solubility at higher concentrations in the reaction buffer. However, their k_{cat}/K_m values can be obtained based on the data obtained in the low concentration range. The K_m and k_{cat} values for **9** and **10** were determined at the pH of 7.2 because of significant nonenzymatic reverse β-addition of **9** and **10**, resulting in the loss of L-cysteine, at pH 8.5 (*SI Appendix*, Fig. S21). The K_m and k_{cat} values for **1** and **4** at pH of 7.2 were also determined for comparison (Table 1). Among all of the substrates tested, **3** and **4** exhibited the highest activities, with k_{cat}/K_m values of 1.6 ± 0.3 and 1.5 ± 0.3 $mM^{-1}min^{-1}$, respectively. The catalytic efficiencies of LnmJ-SH toward **1** and **4** decreased as the pH was lowered from 8.5 to 7.2 (Table 1). At pH of 7.2, the k_{cat}/K_m values for **1**, **4**, **9**, and **10** ranged from (23 ± 5) × 10^{-3} to 0.38 ± 0.01 $mM^{-1}·min^{-1}$, indicating substrate tolerance of LnmJ-SH (*SI Appendix*, Fig. S30).

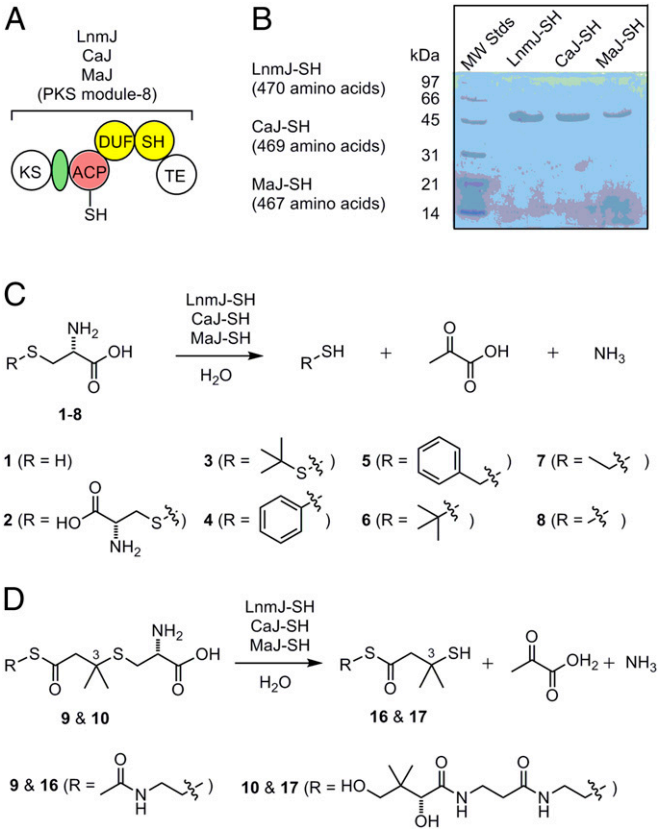

Fig. 3. LnmJ-SH, CaJ-SH, and MaJ-SH as a new family of PKS domains that catalyzes C-S bond cleavage. (*A*) Domain organization and architectural conservation of the PKS modules featuring the DUF and SH domains reported in this study. (*B*) SDS/PAGE (12%) of the purified SH domains (calculated molecular mass, including the extra amino acids from expression vector, in daltons): LnmJ-SH (49,609.6), CaJ-SH (49,699.9), and MaJ-SH (48,539.8). (*C*) LnmJ-SH, CaJ-SH, and MaJ-SH catalyze C-S bond cleavage using L-cysteine (**1**) and L-cysteine *S*-modified analogs (**2–8**) as substrates. (*D*) LnmJ-SH, CaJ-SH, and MaJ-SH catalyze C-S bond cleavage using *N*-acetylcysteamine and pantetheine thioester of the L-cysteine-polyketide adducts **9** and **10** as substrate mimics of the ACP-tethered growing polyketide intermediate in LNM biosynthesis, affording the corresponding products **16** and **17** featuring a thiol moiety at C-3 (Fig. 1).

PLP-Dependent C-S Bond Cleavage Activities of the LnmJ-SH Domain Homologs. To examine whether other LnmJ-SH homologs could catalyze the similar C-S bond cleavage, we overproduced in *E. coli* and purified two selected LnmJ-SH homologs, CaJ-SH from *C. acidiphilia* DSM 44928 and MaJ-SH from *M. aurantiaca* ATCC 27029, to homogeneity (Fig. 3). Although not known as LNM or analog producers, bioinformatics analysis revealed that both strains contain a biosynthetic gene cluster homologous to the *lnm* gene clusters (*SI Appendix*, Fig. S3). The purified CaJ-SH and MaJ-SH were both yellow in color, indicative of bound PLP, and catalyzed C-S bond cleavage by using all of the substrates tested for LnmJ-SH (Fig. 3 and Table 1). Both enzymes also exhibited the highest catalytic efficiency toward **4**. Compared with the same substrate, the catalytic efficiencies for the three enzymes in general increase noticeably (~40-fold) in the order of LnmJ-SH < CaJ-SH < MaJ-SH (Table 1). However, for **9** and **10**, the three SH domains exhibited similar k_{cat}/K_m values ranging from (23 ± 5) × 10^{-3} to (51 ± 8) × 10^{-3} $mM^{-1}·min^{-1}$ (i.e., ~twofold). Among the three SH domains, MaJ-SH appeared to be most sensitive to pH, displaying a sharp decrease in k_{cat}, as exemplified for **1** from 106 ± 6 min^{-1} at pH 8.5 to 6.7 ± 0.2 min^{-1} at pH 7.2 (~16-fold). The K_m and k_{cat} values of MaJ-SH for **4** could be determined at pH 7.2, benefited from the reduced k_{cat} value at the lower pH (Table 1 and *SI Appendix*, Fig. S30).

Discussion

LnmJ-SH and Homologs Are Distinct from All Sulfur-Incorporating Enzymes Known to Date. Sulfur-containing metabolites are well known from both primary and secondary metabolism (1–15). The best-characterized sulfur-incorporating enzymes from primary metabolism include the following: (*i*) the group I and II cysteine desulfurases that use L-cysteine as the substrate and catalyze the production of a protein-bound persulfide and alanine, as exemplified by NifS, IscC, and CsdB (1–3); (*ii*) the D-cysteine desulfhydrases that use D-cysteine as the substrate and catalyze the β-elimination reaction to produce hydrogen sulfide, ammonia, and pyruvate (32); and (*iii*) the cysteine *S*-conjugate β-lyases that use L-cysteine-*S*-conjugates as the substrates and catalyze the β-elimination reaction to produce a thiol, ammonia, and pyruvate (33). They are all PLP-dependent enzymes that share significant sequence homology and fall into a large clade upon phylogenetic analysis (Fig. 2).

Sulfur incorporation into secondary metabolites is less well characterized except for those peptide natural products derived from intact incorporation of cysteine and methionine by either ribosomal or NRPS mechanism. Recent literature has witnessed a flurry of reports on natural product biosynthesis featuring C-S bond formation or C-S bond cleavage steps (6–15). Thus, as summarized in *SI Appendix*, Fig. S31, in the biosynthesis of lincomycin A, LmbT catalyzes the C-S bond formation between GDP-D-α-D-lincosamide and ergothioneine (EGT), and LmbV then catalyzes the sulfur exchange between the resultant EGT *S*-conjugate and mycothiol (MSH) to generate the new C-S bond (6). In the biosynthesis of BE-7585A, the C-S bond formation in the 2-thioglucose moiety is catalyzed by BexX with the sulfur provided by sulfur carrier protein and activating protein involved in primary metabolite biosynthesis (7). The C-S bond formation in ergothioneine and ovothiol biosynthesis is catalyzed by EgtB using γ-glutamylcysteine as the sulfur donor and by OvoA using cysteine as the sulfur donor, respectively (8). The C-S bond formation in gliotoxin biosynthesis is catalyzed by GliG using glutathione as the sulfur donor, and the

Table 1. Pseudo-first–order kinetic parameters of LnmJ-SH, CaJ-SH, and MaJ-SH toward selected substrates

	LnmJ-SH			Ca-SH			Ma-SH		
Substrates*	K_m, mM	k_{cat}, min^{-1}	k_{cat}/K_m, $mM^{-1}{\cdot}min^{-1}$	K_m, mM	k_{cat}, min^{-1}	k_{cat}/K_m, $mM^{-1}{\cdot}min^{-1}$	K_m, mM	k_{cat}, min^{-1}	k_{cat}/K_m, $mM^{-1}{\cdot}min^{-1}$
1	35 ± 5	2.4 ± 0.1	0.07 ± 0.01	40 ± 5	6.0 ± 0.3	0.15 ± 0.02	42 ± 5	106 ± 6	2.5 ± 0.3
2			0.22 ± 0.01			1.12 ± 0.07			13 ± 1
3			1.6 ± 0.3			2.15 ± 0.02			27 ± 3
4			1.5 ± 0.3			4.53 ± 0.07			68 ± 3
5			0.20 ± 0.01			0.26 ± 0.02			0.63 ± 0.07
6			$(23 \pm 1) \times 10^{-3}$			$(69 \pm 3) \times 10^{-3}$			0.381 ± 0.001
7			$(14 \pm 1) \times 10^{-3}$			0.416 ± 0.002			0.137 ± 0.002
8			$(6.1 \pm 0.2) \times 10^{-3}$			0.22 ± 0.01			$(48 \pm 1) \times 10^{-3}$
1†	13 ± 2	0.82 ± 0.04	$(63 \pm 1) \times 10^{-3}$	18 ± 4	5.2 ± 0.4	0.29 ± 0.01	6.3 ± 0.8	6.7 ± 0.2	1.1 ± 0.1
4†			0.38 ± 0.01			0.93 ± 0.02	1.9 ± 0.2	8.7 ± 0.3	4.6 ± 0.5
9†	32 ± 6	0.75 ± 0.06	$(23 \pm 5) \times 10^{-3}$	42 ± 6	2.2 ± 0.1	$(51 \pm 8) \times 10^{-3}$	33 ± 5	1.24 ± 0.07	$(38 \pm 6) \times 10^{-3}$
10†	30 ± 3	0.72 ± 0.03	$(24 \pm 3) \times 10^{-3}$	23 ± 2	0.66 ± 0.02	$(29 \pm 3) \times 10^{-3}$	23 ± 4	1.07 ± 0.06	$(47 \pm 9) \times 10^{-3}$

*Assays were carried out in 0.2 mM PLP, 20 mM KCl, 100 mM Tris·HCl, pH 8.5 (*SI Appendix*).
†Assays were carried out in 0.2 mM PLP, 20 mM KCl, 100 mM Tris·HCl, pH 7.2 (*SI Appendix*).

subsequent C-S bond cleavage is catalyzed by a pyridoxal phosphate (PLP)-dependent enzyme, GliI (9). The C-S bond formation in sulfoquinovosyl diacylglycerol biosynthesis is catalyzed by SQD1 using sulfite as the sulfur donor (10). The C-S bond formation in grixazone biosynthesis is proposed by a nonenzymatic reaction between the quinone imine intermediate and *N*-acetylcysteine (11). The C-S bond cleavage in glucosinolate biosynthesis in plants is catalyzed by SUR1, which belongs to the PLP-dependent cysteine *S*-conjugate β-lyases (12). The C-S bond cleavage in the biosynthesis of volatile sulfur containing natural products in plants is catalyzed by cysteine sulfoxide lyases, as exemplified by alliinases in the genus *Allium* (13, 14). Finally, CalE6 from the calicheamicin biosynthetic gene cluster in *Micromonospora echinospora* was found to catalyze C-S bond cleavage in vitro, using methionine as the substrate, to produce methanethiol, although no evidence was provided to correlate this reaction to sulfur incorporation in calicheamicin biosynthesis (15). While the emerging sulfur-incorporating enzymes from natural product biosynthetic pathways are very diverse, those that catalyze C-S bond cleavage are all PLP-dependent enzymes, including SUR1, cysteine sulfoxide lyases, CalE6, and GliI, and fall into the same clade as the cysteine desulfurases, cysteine desulfhydrases, and cysteine *S*-conjugate β-lyases from primary metabolism (Fig. 2).

Remarkably, LnmJ-SH shares no sequence homology with any of the sulfur-incorporating enzymes known to date from either primary or secondary metabolisms. Instead, LnmJ-SH forms a distinct clade, including CaJ-SH, MaJ-SH, and many other homologs from various organisms, none of which was functionally annotated previously (Fig. 2). Characterization of LnmJ-SH and selected homologs CaJ-SH and MaJ-SH in this study therefore unveils a new family of PKS domains that catalyze C-S bond cleavage, using L-cysteine and L-cysteine-*S*-conjugates as substrates (Fig. 3).

LnmJ-SH Representing a New Family of PKS Domains That Catalyze C-S Bond Cleavage. We have unambiguously established that the LnmJ-SH family of PKS domains catalyzes C-S bond cleavage, using L-cysteine and L-cysteine-*S*-conjugates as substrates (Fig. 3). These findings expand the chemistry and enzymology of PKSs, and the LnmJ-SH family of PKS domains might be exploited to incorporate sulfur into polyketide natural products by PKS engineering.

It should be noted that the LnmJ-SH family of PKS domains shares significant sequence homology with the ABL family of enzymes. All of the residues required for PLP and amino acid substrate binding are conserved (*SI Appendix*, Fig. S3), but the two families fall into two distinct clades phylogenetically (Fig. 2). Members of the ABL family enzymes, such as TPL and TIL, are known to use cysteine as an alternative substrate, catalyzing C-S bond cleavage, albeit at significantly reduced catalytic efficiency (V_{max}/K_m) of 4% (33) and 1% (34), respectively, relative to their natural substrates tyrosine and tryptophan. In contrast, LnmJ-SH, CaJ-SH, and MaJ-SH showed no activity toward tyrosine and tryptophan under all conditions examined. The LnmJ-SH family of enzymes therefore apparently has evolved to be L-cysteine and L-cysteine-*S*-conjugates specific. Interestingly, substrate promiscuity toward L-cysteine *S*-modified analogs is also known for ABLs (35, 36). For example, when *S*-(*o*-nitrophenyl)-L-cysteine (SOPC) was tested as an alternative substrate, both TPL and TIL exhibited higher catalytic efficiency toward SOPC than their native substrates, presumably due to *o*-nitrothiophenol being a better leaving group (35, 36). These results are reminiscent of the findings from the current study for the SH domains, which exhibited the highest catalytic efficiency toward *S*-phenyl-L-cysteine (**4**) among all of the substrates tested (Table 1). Regardless of the observed substrate promiscuity, the LnmJ-SH family of new PKS domains catalyzes C-S bond cleavage and prefers amino acid substrates bearing minimally a thio-side chain and L-amino group (Figs. 1 and 3).

Using the sequence alignment and structures of TPL and TIL (PDB ID codes 2VFL, 2TPL, 2EZ1, 2EZ2, and 1AX4), we generated a structural model of LnmJ-SH (*SI Appendix*, Fig. S32). TPL and TIL are catalytic as dimers, which are stabilized by the binding of two potassium cations, as such, the active site of each monomer is closed off from solvent by dimerization (31, 37–39). These dimers form higher-order tetramers in solution and crystal structures. The residues responsible for potassium binding and dimerization are mutated in LnmJ-SH (S6874[K], W6677[G], A6694[E], A6696[Y], corresponding TPL residue in brackets), although LnmJ-SH retains potassium dependence and a tetrameric quaternary structure. A few regions of LnmJ-SH, likely responsible for alteration of function, could be predicted. The loop between H6994 and P7000 in LnmJ-SH is missing 11 residues, and the C-terminal β-α-β-motif, responsible for covering the active site and substrate specificity, is likely truncated in LnmJ-SH starting at residue L7038. In addition, the region between R6764 and R6774 in LnmJ-SH is lacking three residues. Together, the truncations and differences in oligomeric interfaces likely allow LnmJ-SH to accommodate a large substrate such as the growing peptide-polyketide intermediate proposed for LNM biosynthesis (Fig. 1).

A Proposal for C-3 Sulfur Incorporation in LNM Biosynthesis. In the proposed biosynthetic pathway for LNM, after introduction of the β-alkyl branch at C-3, there is a requirement for β-addition of a sulfur nucleophile at the C-3 carbon of the α,β-unsaturated linear polyketide intermediate (Fig. 1). While the characterization of LnmJ-SH in this study establishes that the sulfur at C-3 of

LNM is ultimately derived from L-cysteine, we have not accounted for how the C-S bond is formed by the β-addition. Recently, β-additions, with carbon nucleophiles (i.e., Michael addition), to generate an alkyl branch in the growing polyketide chain have been observed (40–42). On the basis of the absolutely conserved KS-ACP-DUF-SH-TE domain architecture of the PKS module-8 of LnmJ and its homologs (Figs. 1 and 3 and *SI Appendix*, Fig. S3), we now propose that a β-addition of L-cysteine, with a sulfur nucleophile, to the α,β-unsaturated polyketide intermediate to afford the cysteine adduct, and LnmJ-DUF serves as the candidate responsible for the β-addition (Fig. 1). This proposal is supported by the ability of LnmJ-SH, CaJ-SH, and MaJ-SH to catalyze C-S bond cleavage of the L-cysteine *S*-modified analogs, especially **9** and **10**, mimics of the proposed biosynthetic intermediate for LNM E1 (Figs. 1 and 3) with similar k_{cat}/K_m values (Table 1). The observed K_m values of LnmJ-SH, CaJ-SH, and MaJ-SH with all tested substrates (**1–10**) are between 1.9 ± 0.2 and 42 ± 6 mM (Table 1), which are too high to be relevant under the physiological reactions. These high K_m values may result from the structural difference between the native substrates of the SH domains and the substrate mimics tested, the fact that the SH domains were studied in isolation (i.e., in truncation lacking the context of other domains within the native PKS modules), or a combination of both (Fig. 1). However, in this proposed pathway, the LnmJ-SH domain would not need tight substrate binding as the L-cysteine-adduct intermediate would be tethered to the cognate ACP, overcoming a need for low K_m values. These hypotheses would also be consistent with the prediction that the active site of LnmJ-SH is likely more open to solvent than the other ABL family members in the structural modeling (*SI Appendix*, Fig. S32). Further studies including structural elucidation of LnmJ-SH and functional characterization of LnmJ-DUF are clearly warranted to define the precise course of events leading to sulfur incorporation at C-3 in LNM biosynthesis.

Materials and Methods

Materials, methods, and detailed experimental procedures are provided in *SI Appendix*. Included in *SI Appendix* are cloning and expression of LnmJ-SH, CaJ-SH, and MaJ-SH; detection of pyruvate, hydrogen sulfide, thiocysteine and thiophenol; chemical synthesis of compound **9** and **10**; detection of compound **16** and **17**; optimization of LnmJ-SH reaction; and ^{1}H and ^{13}C NMR spectra and HR-ESI-MS spectra for compound **9–11** and **15–17**.

ACKNOWLEDGMENTS. We thank Kyowa Hakko Kogyo Co. Ltd for the wild-type *S. atroolivaceus* S-140 strain and the NMR Core facility at the Scripps Research Institute in obtaining ^{1}H and ^{13}C NMR data. This work was supported in part by NIH Grant CA106150.

1. Mueller EG (2006) Trafficking in persulfides: Delivering sulfur in biosynthetic pathways. *Nat Chem Biol* 2(4):185–194.
2. Kessler D (2006) Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes. *FEMS Microbiol Rev* 30(6):825–840.
3. Hidese R, Mihara H, Esaki N (2011) Bacterial cysteine desulfurases: Versatile key players in biosynthetic pathways of sulfur-containing biofactors. *Appl Microbiol Biotechnol* 91(1):47–61.
4. Fontecave M, Ollagnier-de-Choudens S, Mulliez E (2003) Biological radical sulfur insertion reactions. *Chem Rev* 103(6):2149–2166.
5. Marquet A (2001) Enzymology of carbon-sulfur bond formation. *Curr Opin Chem Biol* 5(5):541–549.
6. Zhao Q, Wang M, Xu D, Zhang Q, Liu W (2015) Metabolic coupling of two small-molecule thiols programs the biosynthesis of lincomycin A. *Nature* 518(7537):115–119.
7. Sasaki E, et al. (2014) Co-opting sulphur-carrier proteins from primary metabolic pathways for 2-thiosugar biosynthesis. *Nature* 510(7505):427–431.
8. Song H, Leninger M, Lee N, Liu P (2013) Regioselectivity of the oxidative C-S bond formation in ergothioneine and ovothiol biosyntheses. *Org Lett* 15(18):4854–4857.
9. Scharf DH, et al. (2012) Epidithiol formation by an unprecedented twin carbon-sulfur lyase in the gliotoxin pathway. *Angew Chem Int Ed Engl* 51(40):10064–10068.
10. Lin CI, McCarty RM, Liu HW (2013) The biosynthesis of nitrogen-, sulfur-, and high-carbon chain-containing sugars. *Chem Soc Rev* 42(10):4377–4407.
11. Suzuki H, Furusho Y, Higashi T, Ohnishi Y, Horinouchi S (2006) A novel o-aminophenol oxidase responsible for formation of the phenoxazinone chromophore of grixazone. *J Biol Chem* 281(2):824–833.
12. Mikkelsen MD, Naur P, Halkier BA (2004) Arabidopsis mutants in the C-S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. *Plant J* 37(5):770–777.
13. Iranshahi M (2012) A review of volatile sulfur-containing compounds from terrestrial plants: Biosynthesis, distribution and analytical methods. *J Essent Oil Res* 24(4):393–434.
14. Kubec R, et al. (2011) Precursors and formation of pyrithione and other pyridyl-containing sulfur compounds in drumstick onion, Allium stipitatum. *J Agric Food Chem* 59(10):5763–5770.
15. Song H, Xu R, Guo Z (2015) Identification and characterization of a methionine γ-lyase in the calicheamicin biosynthetic cluster of Micromonospora echinospora. *ChemBioChem* 16(1):100–109.
16. Hara M, et al. (1989) Leinamycin, a new antitumor antibiotic from Streptomyces: Producing organism, fermentation and isolation. *J Antibiot (Tokyo)* 42(12):1768–1774.
17. Asai A, et al. (1996) Thiol-meidated DNA alkylation by the novel antitumor antibiotic leinamycin. *J Am Chem Soc* 118(28):6802–6803.
18. Gates KS (2000) Mechanisms of DNA damage by leinamycin. *Chem Res Toxicol* 13(10):953–956.
19. Fekry MI, et al. (2011) Noncovalent DNA binding drives DNA alkylation by leinamycin: Evidence that the Z,E-5-(thiazol-4-yl)-penta-2,4-dienone moiety of the natural product serves as an atypical DNA intercalator. *J Am Chem Soc* 133(44):17641–17651.
20. Cheng YQ, Tang GL, Shen B (2002) Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in Streptomyces atroolivaceus S-140. *J Bacteriol* 184(24):7013–7024.
21. Cheng YQ, Tang GL, Shen B (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc Natl Acad Sci USA* 100(6):3149–3154.
22. Tang GL, Cheng YQ, Shen B (2004) Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. *Chem Biol* 11(1):33–45.
23. Tang GL, Cheng YQ, Shen B (2006) Polyketide chain skipping mechanism in the biosynthesis of the hybrid nonribosomal peptide-polyketide antitumor antibiotic leinamycin in Streptomyces atroolivaceus S-140. *J Nat Prod* 69(3):387–393.
24. Tang GL, Cheng YQ, Shen B (2007) Chain initiation in the leinamycin-producing hybrid nonribosomal peptide/polyketide synthetase from Streptomyces atroolivaceus S-140. Discrete, monofunctional adenylation enzyme and peptidyl carrier protein that directly load D-alanine. *J Biol Chem* 282(28):20273–20282.
25. Liu T, Huang Y, Shen B (2009) Bifunctional acyltransferase/decarboxylase LnmK as the missing link for beta-alkylation in polyketide biosynthesis. *J Am Chem Soc* 131(20):6900–6901.
26. Huang Y, et al. (2011) Characterization of the lnmKLM genes unveiling key intermediates for β-alkylation in leinamycin biosynthesis. *Org Lett* 13(3):498–501.
27. Lohman JR, Bingman CA, Phillips GN, Jr, Shen B (2013) Structure of the bifunctional acyltransferase/decarboxylase LnmK from the leinamycin biosynthetic pathway revealing novel activity for a double-hot-dog fold. *Biochemistry* 52(5):902–911.
28. Huang SX, et al. (2015) Leinamycin E1 acting as an anticancer prodrug activated by reactive oxygen species. *Proc Natl Acad Sci USA* 112(27):8278–8283.
29. Finn RD, et al. (2010) The Pfam protein families database. *Nucleic Acids Res* 38(Database issue):D211–D222.
30. Zdobnov EM, Apweiler R (2001) InterProScan–an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17(9):847–848.
31. Phillips RS, Demidkina TV, Faleev NG (2003) Structure and mechanism of tryptophan indole-lyase and tyrosine phenol-lyase. *Biochim Biophys Acta* 1647(1-2):167–172.
32. Nagasawa T, Ishii T, Kumagai H, Yamada H (1985) D-Cysteine desulfhydrase of Escherichia coli. Purification and characterization. *Eur J Biochem* 153(3):541–551.
33. Kumagai H, et al. (1972) Purification, crystallization and properties of tyrosine phenol lyase from Erwinia herbicola. *Agric Biol Chem* 36(3):472–482.
34. Newton WA, Morino Y, Snell EE (1965) Properties of crystalline tryptophanase. *J Biol Chem* 240(3):1211–1218.
35. Chen H, Gollnick P, Phillips RS (1995) Site-directed mutagenesis of His343–>Ala in Citrobacter freundii tyrosine phenol-lyase. Effects on the kinetic mechanism and rate-determining step. *Eur J Biochem* 229(2):540–549.
36. Zakomirdina LN, et al. (2002) Tryptophan indole-lyase from Proteus vulgaris: Kinetic and spectral properties. *Biochemistry (Mosc)* 67(10):1189–1196.
37. Sundararaju B, et al. (1997) The crystal structure of Citrobacter freundii tyrosine phenol-lyase complexed with 3-(4′-hydroxyphenyl)propionic acid, together with site-directed mutagenesis and kinetic analysis, demonstrates that arginine 381 is required for substrate specificity. *Biochemistry* 36(21):6502–6510.
38. Milić D, et al. (2006) Structures of apo- and holo-tyrosine phenol-lyase reveal a catalytically critical closed conformation and suggest a mechanism for activation by K^+ ions. *Biochemistry* 45(24):7544–7552.
39. Antson AA, et al. (1993) Three-dimensional structure of tyrosine phenol-lyase. *Biochemistry* 32(16):4195–4206.
40. Lim S-K, et al. (2009) iso-Migrastatin, migrastatin, and dorrigocin production in Streptomyces platensis NRRL 18993 is governed by a single biosynthetic machinery featuring an acyltransferase-less type I polyketide synthase. *J Biol Chem* 284(43):29746–29756.
41. Bretschneider T, et al. (2013) Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature* 502(7469):124–128.
42. Seo J-W, et al. (2014) Comparative characterization of the lactimidomycin and iso-migrastatin biosynthetic machineries revealing unusual features for acyltransferase-less type I polyketide synthases and providing an opportunity to engineer new analogues. *Biochemistry* 53(49):7854–7865.