Published in final edited form as:

Mol Biosyst. 2013 March ; 9(3): 478–491. doi:10.1039/c3mb25523a.

Cloning and sequencing of the kedarcidin biosynthetic gene cluster from *Streptoalloteichus* sp. ATCC 53650 revealing new insights into biosynthesis of the enediyne family of antitumor antibiotics[†]

Jeremy R. Lohman^a, Sheng-Xiong Huang^a, Geoffrey P. Horsman^b, Paul E. Dilfer^a, Tingting Huang^a, Yihua Chen^b, Evelyn Wendt-Pienkowski^b, and Ben Shen^{*,a,b,c,d}

^aDepartment of Chemistry, The Scripps Research Institute, Jupiter, Florida 33458, USA

^bDivision of Pharmaceutical Sciences, School of Pharmacy, University of Wisconsin-Madison, Madison, Wisconsin, 53705, USA

^cDepartment of Molecular Therapeutics, The Scripps Research Institute, Jupiter, Florida 33458, USA

^dNatural Products Library Initiative at The Scripps Research Institute, The Scripps Research Institute, Jupiter, Florida 33458, USA

Abstract

Enediyne natural product biosynthesis is characterized by a convergence of multiple pathways, generating unique peripheral moieties that are appended onto the distinctive enediyne core. Kedarcidin (KED) possesses two unique peripheral moieties, a (*R*)-2-aza-3-chloro- β -tyrosine and an iso-propoxy-bearing 2-naphthionate moiety, as well as two deoxysugars. The appendage pattern of these peripheral moieties to the enediyne core in KED differs from the other enediynes studied to date with respect to stereochemical configuration. To investigate the biosynthesis of these moieties and expand our understanding of enediyne core formation, the biosynthetic gene cluster for KED was cloned from *Streptoalloteichus* sp. ATCC 53650 and sequenced. Bioinformatics analysis of the *ked* cluster revealed the presence of the conserved genes encoding for enediyne core biosynthesis, type I and type II polyketide synthase loci likely responsible for 2-aza-L-tyrosine and 3,6,8-trihydroxy-2-naphthionate formation, and enzymes known for deoxysugar biosynthesis. Genes homologous to those responsible for the biosynthesis, activation, and coupling of the L-tyrosine-derived moieties from C-1027 and maduropeptin and of the naphthionate moiety from neocarzinostatin are present in the *ked* cluster, supporting 2-aza-L-tyrosine and 3,6,8-trihydroxy-2-naphthoic acid as precursors, respectively, for the (*R*)-2-aza-3-chloro- β -tyrosine and the 2-naphthionate moieties in KED biosynthesis.

[†]Electronic Supplementary Information (ESI) available: The amino acid sequence of KedA in comparison with other known apoproteins (Fig. S1), the original and revised structures of the KED chromophore (Fig. S2), enediyne natural products whose structures have been determined (Fig. S3), HPLC and MS analysis of the KED chromophore (Fig. S4), SDS-PAGE analysis of the purified KedF (Fig. S5), comparative analysis of the KED, C-1027, and MDP gene cluster supporting the proposed pathway for (*R*)-2-aza-3-chloro- β -tyrosine in KED biosynthesis (Fig. S6), and comparative analysis of the KED, NCS, and MDP gene cluster supporting the proposed pathway for 3-hydroxy-7,8-dimethoxy-6-isopropoxy-2-naphthoic acid in KED biosynthesis. See DOI: 10.1039/b000000x/

^{*}The Scripps Research Institute, 130 Scripps Way, #3A1, Jupiter, Florida 33458, USA; Tel: (561) 228-2456; Fax: (561) 228-2472; shenb@scripps.edu.

Introduction

Kedarcidin (KED) was isolated from *Streptoalloteichus* sp. ATCC 53650 (originally strain L585-6) as a chromoprotein antitumor antibiotic in 1992.^{1–5} The KED apoprotein primary sequence of 114 amino acids was determined by Edman degradation² (Fig. S1), and the solution structure solved by NMR spectroscopy.³ The structure of the KED chromophore was first established on the basis of an extensive spectroscopic analysis in 1992.^{4,5} It has since been revised twice according to total syntheses^{6,7} with the final revised structure shown in Fig. 1 (also see Fig. S2). KED belongs to the enediyne family of antitumor antibiotics, which are of great interest as potent anticancer agents. They possess a reactive enediyne core that is able to abstract hydrogens from the deoxyribose backbone of DNA. Molecular oxygen can then react with the newly formed carbon-centered radicals, leading to site-specific single-stranded or double-stranded breaks, as well as interstrand crosslinks, and ultimately to cell death.^{8–14} The potent anticancer activity of enediynes is offset in clinical applications by their high cytotoxicity. Nevertheless, polymer and antibody conjugates of enediynes have been developed that display reduced general cytotoxicity, thereby allowing for their use in cancer chemotherapies.^{15–20}

The enediynes represent a steadily growing family of natural products with remarkable molecular architectures. Since the structural elucidation of neocarzinostatin (NCS)²¹ and calicheamicin (CAL),²² the first two members of the family, in the 1980s, 14 enediynes have now been structurally confirmed, which include three probable enediynes isolated as aromatized products^{17,19,20,23} (Fig. S3). Structurally, the enediynes are characterized by an unsaturated 9- or 10-membered carbocyclic ring featuring a diyne conjugated to a central double bond or an incipient double bond. The 9-membered enediyne chromophores are typically isolated noncovalently bound to an apoprotein (Fig. S1), and the resulting complex is termed a chromoprotein; examples include C-1027, NCS, maduropeptin (MDP), and KED. There are exceptions where 9-membered enediynes lack an apoprotein, including N1999A2, the enediyne precursors of sporolides (SPO) and possibly the cyanosporasides^{19,20} (Fig. S3). All 10-membered enediynes known to date are discrete small molecules that do not require sequestration by an apoprotein; examples include CAL, esperamicin (ESP), dynemicin (DYN), namenamycin, shishijimicin, and uncialamycin (Fig. S3). Upon release from the apoprotein the 9-membered enediyne chromophore undergoes a Bergman or Myers-Saito rearrangement, yielding a benzenoid diradical that initiates oxidative DNA damage, thereby triggering cell death. The 10-membered enediynes typically need a base or reducing agent to initiate a similar rearrangement that subsequently damages DNA leading to cell death.^{8–10,15–17,19,20}

While the enediyne core defines the enediyne family of natural products, they are always decorated with various peripheral moieties that modulate the biological activity and specificity of the individual enediyne natural products. The biosynthetic gene clusters for four 9-membered enediynes (C-1027,²⁴ NCS,²⁵ MDP,²⁶ and SPO²⁷) and three 10-membered enediynes (CAL,²⁸ ESP,^{29,30} and DYN³¹) have been cloned and partially characterized. Comparative studies of these biosynthetic machineries have revealed: (i) enediyne core biosynthesis is initiated by the enediyne polyketide synthase (PKS), but it is the enediyne PKS-associated enzymes that channel a nascent common polyene intermediate into 9- or 10-membered enediyne cores,^{32,33} (ii) biosynthesis of the peripheral moieties varies widely in the nature of precursors from primary metabolism, featuring much novel chemistry and enzymology,^{34–57} and (iii) a convergent biosynthetic strategy between the enediyne core and the varying peripheral moieties finally furnishes the myriad of functionalities found in the enediyne family of natural products.^{19,20}

Inspired by the findings from comparative studies of the enediyne biosynthetic machineries, we decided to clone and characterize the KED biosynthetic machinery to shed new insights into biosynthesis of the enediyne family of antitumor antibiotics. We are particularly intrigued by the following observations: (i) amino acid sequencing revealed three variants of the KED apoproteins with varying N-termini (Fig. S1), (ii) the 2-aza-3-chloro- β -tyrosine moiety that has not been seen in any other natural product, (iii) a deceptively simple isopropoxy group at the 2-naphthionate moiety, the biosynthesis of which has little literature precedence, and (iv) the peripheral moieties are appended to the core with an unusual stereochemistry that differs from the other enediynes characterized to date (Fig. 1) (also see Fig. S3 for comparison).

Here we present: (i) the cloning and annotation of the *ked* biosynthetic gene cluster from *Streptoalloteichus* sp. ATCC 53650, (ii) a convergent biosynthetic pathway for the KED chromophore on the basis of sequence analysis and comparisons to the other cloned 9- and 10-membered enediyne gene clusters,^{24–31} and (iii) in vivo characterization of *kedE* and *kedE10* and in vitro characterization of KedF further supporting the proposed pathway for KED biosynthesis.

Results

Confirmation of KED Chromoprotein Production

The KED chromoprotein was purified to homogeneity guided by a bioassay against *M. luteus*.³⁴ The KED chromophore was released from the purified KED chromoprotein by EtOAc extraction. The KED chromophore, purified under these conditions, was shown by HPLC analysis to be a mixture of the enediyne and aromatized forms, and the enediyne form was completely converted into the aromatized form at room temperature overnight; the identities of both enediyne and aromatized forms of the KED chromophore were confirmed by high resolution mass spectrometry (Fig. S4). For the enediyne form of the KED chromophore, HRESIMS yielded an $[M + H]^+$ ion at m/z 1030.37338 for the enediyne form of the KED chromophore, consistent with its predicted molecular formula of $C_{53}H_{60}N_3O_{16}Cl$ (calculated $[M + H]^+$ ion at m/z 1030.37349).^{2,5} For the aromatized form of the KED chromophore, HRESIMS revealed an $[M + H]^+$ ion at m/z 1032.39109, consistent with the molecular formula of $C_{53}H_{62}N_3O_{16}Cl$ (calculated $[M + H]^+$ ion at m/z 1032.38914), differing from the enediyne form by the presence of two additional protons as would be predicted for the aromatized KED chromophore (Fig. 1 and Fig. S4).^{58,59} These results re-confirm that *Streptoalloteichus* sp. ATCC 53650 in our possession harbors functional KED biosynthetic machinery.^{1–5} Under the conditions described, the isolated yield of the KED chromoprotein complex is estimated to be 50 mg/L.^{2,5}

Cloning, Sequencing, and Annotation of the *ked* Gene Cluster

The enediyne PKS gene is the hallmark of enediyne biosynthetic clusters,^{32,33} and as such, degenerate primers previously utilized to clone enediyne genes and clusters were used to localize the *ked* cluster.²⁹ A PCR-amplified internal fragment of *kedE* (probe-1) was first used as a probe to screen the *Streptoalloteichus* sp. ATCC 53650 cosmid library, resulting in the isolation of cosmid pBS16002. Iterations of chromosomal walking from pBS16002 using probe-2, -3, -4, and -5 afforded the four additional overlapping cosmids pBS16003, pBS16004, pBS16005, and pBS16006. Together, the five overlapping cosmids cover 135 kb of contiguous DNA (Fig. 2A), complete DNA sequence of which led to the identification of 117 orfs (Fig. 2B). The overall GC content of the sequenced region is 73.2%, characteristic for the *Actinomycetels*.⁶⁰

Functional Assignments of Genes within the *ked* Cluster

Functional assignments of individual orfs were made by comparison of the deduced gene products with proteins of known or predicted functions in the database as summarized in Table 1. Sequence analysis by BLAST comparison and InterProScan of putative orfs suggested that the *ked* gene cluster minimally spans ~105 kb. Starting from *kedE11* and concluding at *kedS1*, the *ked* cluster contains 81 orfs in 21 operons that encode KED biosynthesis, regulation, and resistance (Figure 2B and Table 1). The orfs flanking the *ked* cluster encode proteins of unknown functions or with similarities to enzymes involved in aromatic amino acid metabolism.

Among the orfs identified within the *ked* cluster include: (i) one (*kedE*) encodes the enediyne PKS and 17 (*kedE1* to *kedE11*, *kedM*, *kedS*, *kedL*, *kedJ*, *kedD2*, *kedF*) encodes accessory enzymes for enediyne core biosynthesis, (ii) six (*kedY* and *kedY1* to *kedY5*) encode enzymes for tailoring the (*R*)-2-aza-3-chloro- β -tyrosine moiety and its coupling to the enediyne core, (iii) five (*kedN1* to *kedN5*) encode enzymes for modifying the 2-naphthionate moiety and its coupling, via (*R*)-2-aza-3-chloro- β -tyrosine, to the enediyne core, (iv) ten (*kedS1* to *kedS10*) encode enzymes for biosynthesis of the two sugar moieties and their couplings to the enediyne core, (v) three (*kedR1*, *kedR2*, *kedR3*) encode proteins for pathway regulation, (vi) three (*kedA*, *kedX*, and *kedX2*) encode elements for resistance, and (vii) four (*kedU1* to *kedU4*) encode proteins of unknown function. In addition, there are two loci, termed type II PKS locus consisting of 18 genes (*kedU11* to *kedU28*) and type I PKS locus consisting of 15 gene (*kedU31* to *kedU45*), inserted within the *ked* cluster, that are unprecedented among all enediyne clusters known to date. They serve as candidates encoding biosynthesis of the nascent precursors for the (*R*)-2-aza-3-chloro- β -tyrosine and 2-naphthionate moieties (Fig. 3).

The KED Apoprotein Keda

Bioinformatics analysis revealed a single gene, *kedA*, within the *ked* cluster, for which the deduced gene product matched the isolated Keda apoprotein.^{2,3} The *kedA* gene is translated as a 145-amino acid protein, and SignalP analysis predicted a leader peptide that is cleaved between A31 and A32, resulting in a 114- acid protein with a predicted N-terminus of Ala-Ser-Ala-Ala-Val (Fig. S1). This is consistent with the amino acid sequence of the isolated mature Keda apoprotein.^{2,3} The two additional known variants of Keda, with a Ser-Ala-Ala-Val and an Ala-Ala-Val terminus, respectively, could be accounted for by either the promiscuity of the leader peptide cleavage site or partial proteolysis of the N-terminus of the mature Keda during isolation. The *kedA* gene has also been independently cloned recently from *Streptoalloteichus* sp. ATCC 53650 and sequenced, overexpression of which in *Streptoalloteichus* sp. ATCC 53650 resulted in a 2-fold enhancement of KED titer.⁶¹

In Vivo Characterization of Kede-Kede10 as Enediyne PKS-Thioesterase for Heptaene Production

We have previously shown the production of heptaene as a hallmark for enediyne biosynthesis, which has been detected from all enediyne producers examined to date and can be produced upon co-expression of the enediyne *pksE* and associated thioesterase (TE) in either *E. coli* or *Streptomyces lividans*.^{32,33} Co-expression of *kedE-kedE10* in *E. coli*, with co-expressions of both *sgcE-sgcE10* as a positive control³³ and *sgcE(C211A)*- as a negative control,³³ indeed resulted in the production of heptaene, the identity of which was confirmed by HPLC analysis in comparison with an authentic standard (Fig. 4).

In Vitro Characterization of KedF as an Epoxide Hydrolase

The *kedF* gene was predicted to encode an epoxide hydrolase, and epoxide hydrolases, such as SgcF⁴⁸ and NcsF2,⁴⁹ have been shown to play a critical role in enediyne biosynthesis, setting up the stereochemistry of the enediyne core for appending the peripheral moieties. KedF was overproduced in *E. coli*, purified to homogeneity (Fig. S5), and directly assayed for epoxide hydrolase activity using racemic styrene epoxide as a substrate mimic as described previously for SgcF⁴⁸ and NcsF2.⁴⁹ HPLC analysis of the reaction mixture showed a single product, the identity of which was confirmed to be the expected 1-phenyl-1,2-ethanediol upon comparison with an authentic standard. To investigate the substrate specificity, KedF was incubated with either (*R*)- or (*S*)-styrene oxide as a substrate. Chiral HPLC analysis of resultant products showed (*S*)-1-phenyl-1,2-ethanediol as the major product, with 85% enantiomeric excess (ee), from (*S*)-styrene oxide and (*R*)-1-phenyl-1,2-ethanediol, with 54% ee, from (*R*)-styrene oxide (Fig. 5A).

To investigate the enantioselectivity of KedF, the steady state kinetic parameters of KedF towards (*R*)- and (*S*)-styrene oxides were determined by adopting the previously developed continuous spectrophotometric assay.^{48,49,62} A plot of initial velocity versus the concentration of (*S*)-styrene oxide displayed Michaelis-Menten kinetics, yielding a k_{cat} of $36.6 \pm 1.1 \text{ min}^{-1}$, a K_{M} of $0.91 \pm 0.10 \text{ mM}$, and a $k_{\text{cat}}/K_{\text{M}}$ value of $40.2 \pm 4.6 \text{ mM}^{-1} \text{ min}^{-1}$, while assays with (*R*)-styrene oxide afforded a k_{cat} of $35.1 \pm 2.4 \text{ min}^{-1}$, a K_{M} of $3.50 \pm 0.64 \text{ mM}$, and a $k_{\text{cat}}/K_{\text{M}}$ value of $10.0 \pm 1.9 \text{ mM}^{-1} \text{ min}^{-1}$ (Fig. 5B). Thus, KedF preferentially hydrolyzes (*S*)-styrene epoxide with a 4.0-fold greater specificity constant. The preference of KedF for the an (*S*)-epoxide is consistent with its proposed role in generating a 13-(*S*)-vicinal diol intermediate in KED biosynthesis (Fig. 3).

Discussion

Cloning of the *ked* Cluster from *Streptoalloteichus* sp. ATCC 53650

We set out to clone and sequence the *ked* gene cluster to further our understanding of enediyne core biosynthesis and to explore the novel chemistry governing the biosynthesis of the peripheral moieties, as exemplified by the (*R*)-2-aza-3-chloro- β -tyrosine and the isopropoxy-bearing 2-naphthoic acid (Fig. 1). The general method we developed previously to access the enediyne PKS and associated genes by PCR³⁰ and the knowledge we have gained by characterizing the C-1027,²⁴ NCS,²⁵ and MDP²⁶ biosynthetic machinery greatly expedited the cloning and sequencing of the *ked* cluster from *Streptoalloteichus* sp. ATCC 53650. The *ked* cluster was localized to a 105-kb contiguous DNA region, consisting of 81 orfs that encode KED biosynthesis, resistance, and regulation (Fig. 2 and Table 1). The cluster boundaries were assigned on the basis of bioinformatics analysis, pending future experimental confirmation. The difficulty in developing a genetic system for *Streptoalloteichus* sp. ATCC 53650, in spite of exhaustive effort, has also prevented us from verifying the *ked* cluster directly by in vivo experiments. Nevertheless, the identity of the cloned gene cluster to encode KED biosynthesis is supported by: (i) the finding of *kedA* within the cloned *ked* cluster that encodes the previously isolated KED apoprotein,^{2,5} (ii) production of the signature heptaene product for enediyne biosynthesis upon co-expression of *kedE-kedE10* in *E. coli*^{32,33,35} and (iii) in vitro characterization of KedF as an epoxide hydrolase using a substrate mimic that affords a vicinal diol product with the regio- and absolute stereochemistry as would be expected for the KED chromophore.^{6,7,48,49}

Biosynthesis of the Two Deoxysugars and Their Incorporation

Identification of the ten sugar biosynthesis genes within the *ked* cluster and their deduced functions supported a divergent pathway for biosynthesis of the two sugars from the common precursor D-glucose-1-phosphate (Fig. 2B and Table 1).^{35,63,64} Thus, as depicted

in Fig. 3A, D-glucose-1-phosphate is first converted into the common intermediate NDP-2,6-dideoxy-4-keto-D-glucose, and three of the five enzymes needed are encoded within the *ked* cluster (KedS1, KedS2, and KedS3). The enzymes responsible for the first two steps, a D-glucopyranosyl-1-nucleotidyltransferase and a NDP-glucose-4,6-dehydratase, are most likely provided by other biosynthetic pathways in *Streptoalloteichus* sp. ATCC 53650, and biosynthetic crosstalk between sugar biosynthetic pathways has been noted previously.⁶⁵ NDP-2,6-dideoxy-4-keto-D-glucose is then diverged by KedS4 and KedS5, affording NDP-L-mycarose, and by KedS7, KedS8, and kedS9, affording NDP-kedarosamine, respectively, both of which are finally coupled to the enediyne core by the two glycosyltransferases, KedS6 and KedS10.^{63,64}

Biosynthesis of the (*R*)-2-Aza-3-chloro- β -tyrosine Moiety and Its Incorporation

2-Aza- β -tyrosine is not known as a natural product, nor has it been found as a part in any other natural product. 2-Aza-L-tyrosine has been isolated from *Streptomyces chibaensis* SF-1346,⁶⁶ but nothing is known about its biosynthesis. Therefore, we did not know a priori what candidate genes to look for that would encode for (*R*)-2-aza-3-chloro- β -tyrosine biosynthesis within the *ked* cluster. Remarkably, comparative analysis of the *ked* cluster with the C-1027 and MDP clusters unveiled a subset of six genes, *kedY*, *kedY1* to *kedY5*, as well as *kedE6*, that are absolutely conserved among the three gene clusters (Fig. S6).^{24,25,56,57} It is these findings that inspired us to propose a pathway for (*R*)-2-aza-3-chloro- β -tyrosine biosynthesis starting from 2-aza-L-tyrosine, in a mechanistic analogy to the biosynthesis, activation, and incorporation of the L-tyrosine-derived moieties in C-1027 and MDP (Fig. S6). Thus, as depicted in Fig. 3B, 2-aza-L-tyrosine is first converted to (*R*)-2-aza- β -tyrosine, catalyzed by KedY4, a 4-methylideneimidazole-5-one (MIO) containing aminomutase.^{36–43,56} Loading of (*R*)-2-aza- β -tyrosine to the free standing peptidyl carrier protein KedY2 by the discrete adenylation enzyme KedY1 activates (*R*)-2-aza- β -as the (*R*)-2-aza- β -tyrosyl-S-KedY2 intermediate. The latter is chlorinated by KedY3, a FAD-dependent halogenase requiring the KedE6 flavin reductase, and finally coupled to the enediyne core via an ester linkage catalyzed by the discrete condensation enzyme KedY5.^{45–49,54} The high sequence homology between KedY1 to KedY5, as well as KedE6, and their counterparts in the C-1027 and MDP biosynthetic machinery supports the proposed pathway for (*R*)-2-aza-3-chloro- β -tyrosine in KED biosynthesis.^{56,57} The distinct substrate specificity, as exemplified by KedY1 for 2-aza-L-tyrosine vs SgcC1 for L-tyrosine, regiospecificity, as exemplified by KedY3 for C-6 chlorination of (*R*)-2-aza- β -tyrosyl-S-KedY2 vs SgcC3 for C-3 chlorination of (*S*)- β -tyrosyl-S-SgcC2, and enantiospecificity, as exemplified by KedY4 affording (*R*)-2-aza- β -tyrosine vs SgcC4 affording (*S*)- β -tyrosine, provide outstanding opportunities to investigate structure-and-activity relationship of this set of fascinating enzymes.

Bioinformatics analysis, however, failed to yield clues for the biosynthetic origin of 2-aza-L-tyrosine. In the absence of any other apparent candidates, we now propose, based more on necessity rather than on bioinformatics data, that the 18-gene type II PKS locus may play a role in 2-aza-L-tyrosine biosynthesis (Fig. 2B and Table 1). This locus has an identical genetic organization and shares high sequence homology with a locus from *Salinispora tropica* (Table 1), which resides near the SPO enediyne cluster but its functions are unknown.²⁷ There are two sets of ketosynthase α and β (KS α and KS β) within this locus. The KS α of both sets lacks the canonical C-H-H/N active site motifs but retain the active site residue cysteines (C-E-A for KedU20 and C-E-S for KedU22), while the KS β of both sets lacks the active site residue cysteine (P-D-A for KedU21 and, AD-G for KedU23). KSs with noncanonical active site motifs are rare but known, and they represent an emerging family of enzymes catalyzing a broad range of chemistry.^{67–69} On the assumption that this locus does play a role in 2-aza-L-tyrosine biosynthesis, one could envisage 2-aza-L-

phenylalanine, either free or tethered to a carrier protein, as a penultimate intermediate of the pathway. Hydroxylation of 2-aza-L-phenylalanine, catalyzed by KedY, a FAD-dependent monooxygenase requiring the KedE6 flavin reductase, finally affords 2-aza-L-tyrosine.^{45,47} Although our attempt to express this type II PKS locus, with or without *kedY*, in selected heterologous hosts failed to produce detectable amount of 2-aza-L-phenylalanine or 2-aza-L-tyrosine, this proposal now sets the stage to investigate 2-aza-L-tyrosine biosynthesis in *S. chibaensis* SF-1346.⁶⁶

Biosynthesis of the iso-Propoxy Bearing 2-Naphthonate Moiety and Its Incorporation

The 2-naphthonate moiety is most likely of polyketide origin, but the exact nature of the nascent linear polyketide intermediate and its subsequent folding pattern to afford the 2-naphthonate backbone cannot be predicted in the absence of isotope labeling experiments. Similar aromatic polyketide moieties have been found in other enediyne natural products, as exemplified by the benzoic acid moiety in MDP and the 1-naphthoic acid moiety in NCS, and the biosynthesis of both moieties are catalyzed by the iterative type I PKSs, MdpB^{26,52} and NcsB,^{25,49–51} respectively (Fig. S7). Inspired by this biosynthetic precedence, we took a close examination of the orfs within the *ked* cluster and identified, in addition to *kedE* that encodes the enediyne PKS, *kedU38* that resides in the middle of the 15-gene type I PKS locus, encodes a type I PKS with a similar domain organization as MdpB and NcsB (Fig. S7).^{25,26} On the basis of these findings, we now propose that the type I PKS locus may play a role in the biosynthesis of the 2-naphthonate moiety. It could be imagined that KedU38 catalyzes the formation of a nascent intermediate, which is further modified by the other activities within the type I PKS locus to yield 3,6,8-trihydroxy-2-naphthoic acid as a key intermediate (Fig. S7). However, all attempts to express the type I PKS locus in selected heterologous failed to produce detectable amount of the proposed 2-naphthoic acid intermediates, therefore this proposal awaits experimental verification.

Regardless the exact biosynthetic origin of the 2-naphthonate moiety, comparative analysis of *ked* cluster to the MDP and NCS clusters further unveiled a subset of five genes, KedN1 to KedN5, with high sequence homology to the tailoring enzymes for the 1-naphthonate moiety in NCS biosynthesis (Fig. S7).^{25,26,49–51} These findings lend additional support to the intermediacy of 3,6,8-trihydroxy-2-naphthoic acid in KED biosynthesis. Thus, as depicted in Fig. 3C, 3,6,8-trihydroxy-2-naphthoic acid could be C-7 hydroxylated by the KedN3 P-450 monooxygenase, triple O-methylated by the KedN1 O-methyltransferase, and tandem C-methylated to furnish the isopropoxy group by the KedN5 radical SAM methyltransferase. The fully modified 2-naphthoic acid is finally activated by KedN2 as a naphthonyl CoA and coupled to the enediyne core via an amide linkage to the 2-aza-3-chloro- β -tyrosine moiety by the KedN4 acyltransferase. The KedN5-catalyzed tandem C-methylation of an O-CH₃ group is unusual for isopropoxy group biosynthesis. A similar mechanism has been proposed for CndI, which was identified to C-methylate an O-CH₃ group to afford an ethoxy group for chondrochloren biosynthesis in *Chondromyces crocatus* Cm c5.⁷⁰ The fact that KedN5 shows significant sequence homology to CndI (24% identity/37% similarity) supports the proposed role of KedN5 in KED biosynthesis.

The Enediyne Core Biosynthesis and Convergent Biosynthesis for the KED Chromophore

By comparing and contrasting the seven enediyne gene clusters known to date [i.e., the four 9-membered enediynes of C-1027,²⁴ NCS,²⁵ MDP,²⁶ and SPO²⁷ and the three 10-membered enediynes of CAL,²⁸ DYN (partial),³¹ and ESP (partial)^{29,30}], we have previously shown that (i) both 9- and 10-membered enediyne clusters share an absolutely conserved five-gene cassette, known as the enediyne PKS cassette, consisting of *E*, *E3*, *E4*, *E5* and *E10*, (ii) PKS chemistry (i.e., E-E10) does not direct biosynthetic divergence between 9- and 10-membered enediynes,^{32,33} (iii) it is the 9- or 10-membered pathway specific enediyne PKS

accessory enzymes that most likely morph a common nascent polyketide intermediate into the distinct enediyne core structures,^{32,33,53} and (iv) the final assembly of the enediyne chromophores features a convergent biosynthetic logic that employs varying coupling chemistry^{44,46,53,54} and often exploits epoxide-forming and epoxide-opening enzymes in activating the endiyne cores^{48,49} and setting up the stereochemistry for the attachment of the peripheral moieties (Fig. 6).^{19,20,53,57}

The *ked* cluster now joins the growing list of enediyne biosynthetic machinery, supporting the emerging paradigm for enediyne core biosynthesis^{19,20,53,56,57} but also revealing new insights. Thus, the *ked* cluster also harbors the absolutely conserved five-gene enediyne PKS cassette (Fig. 6), whose PKS chemistry is demonstrated by the production of the hallmark heptaene product for enediyne biosynthesis upon co-expression of *kedE-kedE10* in *E. coli* (Fig. 4).^{32,33} Flanking the *ked* enediyne PKS cassette are the highly conserved 13 genes, *kedE1*, *kedE2*, *kedE6* to *kedE9*, *kedE11*, *kedD2*, *kedF*, *kedJ*, *kedL*, *kedM*, and *kedS* (Fig. 2B and Table 1), that are highly conserved among the five 9-membered enediyne gene clusters known to date (Fig. 6).^{33,53} They encode the 9-membered enediyne pathway specific PKS accessory enzymes for endiyne core biosynthesis, including *KedF* whose epoxide hydrolase activity was demonstrated in vitro to afford a vicinal diol with the same regio- and absolute stereochemistry as would be for the KED enediyne core (Fig. 5). It should be noted that while the five genes consisting of the enediyne PKS cassette are typically clustered, the organization of genes encoding the accessory enzymes is less conserved, scattering on either side of the enediyne PKS cassette within the gene cluster (Fig. 6). They nonetheless show significant sequence homology, ensuring their identification upon careful bioinformatics analysis (Table 1). These observations should now be taken into consideration in future effort to identify and annotate new enediyne biosynthetic gene clusters. Finally, the fully modified and activated KED enediyne core intermediate is coupled with the two deoxysugars, the (*R*)-2-aza-3-chloro- β -tyrosine, and the 3-hydroxy-7,8-dimethoxy-6-isopropoxy-2-naphthionate moiety, and *KedS6*, *KedS10*, *KedN4*, and *KedY5* are proposed to catalyze these coupling steps, respectively, the timing of which is pending future determination (Fig. 3D). It has long been speculated that the convergent molecular logic for enediyne biosynthesis presents outstanding opportunities to engineer new enediyne natural products by combinatorial strategies.^{12–14,19,20} The availability of the *ked* cluster and the novel chemistry associated with the KED biosynthetic machinery surely will enrich the enediyne genetic toolbox and facilitate such engineering effort.

KED Biosynthesis and Structural Revisions

The structure of KED chromophore has been revised twice since it was first published^{4–7} (Fig. S2). The original structure had the (*R*)-2-aza-3-chloro- α -tyrosine moiety with the 2-naphthamide linked at the α -amino position.^{4,5} This structure was subsequently revised to (*R*)-2-aza-3-chloro- β -tyrosine with the 2-naphthamide linked at the β -amino position.⁶ Cloning and sequencing of the *ked* cluster in the current study supports this revision, as *KedY4* is similar to *SgcC4* and *MdpC4*, two MIO-containing aminomutases that have been characterized in vitro to catalyze the conversion of α -tyrosine to β -tyrosine⁵⁶ (Fig. S6), supporting the intermediacy of (*R*)-2-aza-3-chloro- β -tyrosine in KED biosynthesis (Fig. 3).

The second revision was of the stereochemistry of the KED enediyne core, initially inverting the entire enediyne core into its enantiomer and subsequently revising the C-10/C-11 disubstitution pattern from trans- to cis-configuration (Fig. S2).^{4–7} The C-10/C-11 cis-disubstitution provides a α -glycosidic and an ether linkage to the L-mycaroside and the (*R*)-2-aza-3-chloro- β -tyrosine moiety, respectively, in the KED chromophore (7). Intriguingly, similar glycosidic and ether/ester linkages to deoxysugar, tyrosine-derived moieties (C-1027 and MDP) and the 1-naphthionate moiety (NCS) are also present in other enediyne natural products, but the relative stereochemistry of these disubstitutions are in the

trans-configuration, as exemplified by the C-1027, MDP, and NCS chromophores (Fig. S3). Comparative studies of KED biosynthesis to those of C-1027, MDP, and NCS now provide opportunities to decipher the mechanism, thereby controlling and exploiting the regio- and stereochemistry in appending the peripheral moieties to each of the endiayne cores for enediayne biosynthesis and structural diversity.^{12–14,19}

Conclusion

Kedarcidin, a member of the enediayne family of antitumor antibiotics, features a novel molecular architecture. The kedarcidin biosynthetic gene cluster is cloned from *Streptoalloteichus* sp. ATCC 53650 and sequenced and annotated. The identity of the cloned gene cluster to encode KED biosynthesis is supported by: (i) finding the *kedA* gene within the cloned *ked* cluster that encodes the previously isolated KED apoprotein, (ii) production of the signature heptaene product for enediayne biosynthesis upon co-expression of *kedE-kedE10*, encoding the enediayne PKS and the associated type II TE, in *E. coli*, and (iii) in vitro characterization of KedF as an epoxide hydrolase using a substrate mimic that affords a vicinal diol product with the regio- and absolute stereochemistry as would be expected for the KED chromophore. Comparative analysis between *ked* and the other cloned 9- and 10-membered enediayne gene clusters supports a convergent biosynthetic pathway for the KED chromophore, an emerging paradigm for the enediayne family of natural products, but the KED biosynthetic machinery is also predicted to feature much novel chemistry.

Experimental

Bacterial Strains, Plasmids, and Sequence Analysis

Streptoalloteichus sp. ATCC 53650, the KED producer, and *Micrococcus luteus* ATCC 9431, the test organism for assay of the antibacterial activity of KED, were from American Type Culture Collection (Rockville, MD). SuperCos1, Gigapack III XL and *E. coli* XL1-Blue MR cells (Stratagene, La Jolla, CA), pGEM-T Easy and pSP72 (Promega, Madison, WI), and pETDuet-1, pRSFDuet-1, and *E. coli* BL21(DE3) cells (Novagen, Madison, WI) were from commercial sources. pANT841,⁷¹ pBS1050,³² pBS1051,³² and pBS1065³² were described previously. DIG-labeling kit and calf intestinal phosphatase (Roche, Indianapolis, IN), T4 DNA ligase (Promega), and restriction enzymes (New England Biolabs Ipswich, MA or Invitrogen, Carlsbad, CA) were from commercial sources. DNA sequencing was carried out at the University of Wisconsin-Madison Biotechnology Center (Madison, WI). Sequence analysis was carried out using BLASTN available from NCBI, and contiguous DNA was compiled using Lasergene (DNASTAR Inc., Madison, WI). Open reading frames (orfs) were predicted using ORFfinder from NCBI and Genemark,⁷² and protein sequences were analyzed using PSI-BLAST and InterProScan.⁷³ All recombinant DNA manipulations were performed by following standard procedures^{60,74} or the manufacturers' instructions.

Production, Isolation, and Analysis of KED

KED production, isolation, and analysis were carried out essentially by following the literature procedures.^{2,5} Thus, *Streptoalloteichus* sp. ATCC 5360 was grown on TSB agar plate⁶⁰ for single colonies. Seed inoculum was prepared by introducing the colony periphery of petri dish cultures into 250-mL flasks containing 50 mL of TSB medium,⁶⁰ followed by shaking at 250 rpm and 28 °C for two days. Production fermentation was carried out by adding 3 mL of seed inoculum into each of the ten 250-mL flasks containing 50 mL of production medium (3% glycerol, 1% pharmedia, 1.5% distiller's solubles extract, 1% fish meal, 0.05% KH₂PO₄, and 0.6% CaCO₃, pH 7.0), and shaking at 250 rpm and 28 °C for five days. The fermentation culture was centrifuged (8000 rpm, 4 °C, 35 min) and filtered to remove mycelia. The supernatant was slowly adjusted to pH 5.0 with 2 N HCl while stirring,

followed by centrifugation (12000 rpm, 4 °C, 35 min) to remove precipitates. The resulting supernatant was mixed with DEAE-cellulose resin equilibrated in buffer (0.05 M Tris-HCl, pH 5.6). The resulting DEAE-cellulose resin was washed with the same buffer twice and eluted with the same buffer containing 1 M NaCl. The eluate was dialyzed against Milli-Q H₂O at 4 °C overnight using a 10 kDa molecular weight cutoff membrane. The dialyzed solution was lyophilized, dissolved in 4 mL cold H₂O, and applied to a DEAE-cellulose column equilibrated in 0.05 M Tris-HCl, pH 5.6. The column was washed with cold H₂O and eluted stepwise with 0.1 M, 0.2 M, and 0.3 M NaCl. Fractions were assayed against *M. luteus*,³⁴ and the active fractions were combined and lyophilized to afford a yellow powder. Further purification was achieved using Sephadex G-75 chromatography eluting with cold H₂O at 4 °C. Again, fractions were followed by assay against *M. luteus*, and active fractions were combined and lyophilized to give pure KED chromoprotein.

To dissociate the KED chromophore from the apoprotein, 5 mg of purified KED chromoprotein was dissolved in 0.2 mL of 0.1 M potassium phosphate buffer, pH 4.3, and extracted twice with 0.3 mL of EtOAc each at 4 °C. The combined EtOAc extract was evaporated in vacuum, and the residue was subjected to high resolution electrospray ionization mass spectrometry (HRESIMS) analysis on an IonSpec HiResMALDI FT mass spectrometer with a 7 tesla superconducting magnet. A portion of the EtOAc extract was also left at room temperature overnight and then similarly evaporated to dryness and analyzed by HRESIMS. The freshly prepared and the overnight EtOAc extracts were also subjected to HPLC analysis. HPLC was carried out on a Varian HPLC system equipped with Prostar 210 pumps, a photodiode array detector, and an Atima-C18 column (5 µm, 4.6 mm x 250 mm, Grace Davison Discovery Sciences, Deerfield, IL). The column was developed at flow rate of 1 mL/min with a linear gradient from 100% buffer A (0.01 M potassium phosphate, pH 6.8) to 20% buffer A/80% buffer B (80% CH₃CN in 0.01 M potassium phosphate buffer, pH 6.8) in 35 min, monitored at 320 nm.

Cosmid DNA Library Construction, Screening, and Sequencing

A SuperCos1 cosmid library was constructed using partially digested (*Sau3AI*) *Streptoalloteichus* sp. ATCC 53650 chromosomal DNA followed by dephosphorylation with calf intestinal phosphatase according to standard procedures.^{60,74} After an overnight ligation at 16 °C, the mixture was packaged using Gigapack III XL and used to transfect *E. coli* XL1 Blue MR cells following the manufacturer's instructions (Stratagene). A 3.5-kb internal fragment of *kedE* was PCR amplified from total genomic DNA using Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA) and the following pair of primers (forward 5'-GGCGGCGGVTACACSGTSGACGGMGCCTGC-3'/reverse, 5'-CCCATSCCGACSCCGGACCASACSGACCAYTCCA-3', where M = A or C; S = C or G; V = A, C, or G; Y = C or T) as described previously.³⁰ The PCR product was cloned into pGEM-T Easy to afford pBS16001, confirmed to encode an internal fragment of an enediynes PKS gene by sequencing,^{29,39} and used to prepare the DIG-labeled probe (probe-1). Probe-1 was then used to screen the cosmid library by colony hybridization, yielding three positive clones. One of the positive clones, pBS16002, was end-sequenced using the following pair of primers (forward 5'-GGGAATAAGGGCGACACGGG-3'/reverse 5'-GCTTATCGATGATAAGCGGTC-3') and confirmed to encode a part of the *ked* cluster. Four additional rounds of chromosomal walking from pBS16002 were subsequently carried out using probe-2, -3, -4, and -5, respectively to isolate overlapping cosmids that cover the entire *ked* cluster (Fig. 2A). Thus, probe-2 and probe-3 were prepared by PCR from pBS16002 using the following pairs of primers (probe-2, forward 5'-GGTACTACCTGCTGTGC-3'/reverse 5'-GGTCTTGGTGAAGCTGC-3') and (probe-3, forward 5'-CGATCAAGTCGATCCTGACC-3'/reverse 5'-GGTCGCTGGTGATGTCGTCG-3'), respectively. Screening the cosmid library by colony

hybridization with probe-2 and probe-3, respectively, resulted in the isolation of pBS16003 and pBS16004. Similarly, probe-4 was prepared by PCR from pBS16004 using the following pairs of primers (forward 5'-GGAGGTCGAGGTGCGTGC-3'/reverse 5'-GGTTCCACGTGATCAGC-3') and used to screen the cosmid library to isolate pBS16005. Probe-5 was prepared by PCR from pBS16005 using the following pairs of primers (forward 5'-GCTGTGCCTGGTGGACCTGACC-3'/reverse 5'-GCAGCAGGTCGAGGTGC-3') and used to screen the cosmid library to isolate pBS16006. Finally, the five overlapping cosmids (i.e., pBS16002, pBS16003, pBS16004, pBS16005, and pBS16006) were similarly end-sequenced to confirm their candidacy for complete sequencing (Fig. 2A).

The five overlapping cosmids were used to generate subclone libraries for complete DNA sequence determination. The resultant DNA sequences were compiled and assembled into contigs, and gaps were filled in by primer walking or by subcloning fragments covering the gaps and subsequently sequencing the cloned fragments (Fig. 2B and Table 1).

The *kedE-kedE10* Co-expression Construct for *E. coli* Expression

To construct the *kedE-kedE10* co-expression plasmid, a 2400-bp SstI-MluI fragment, containing the 1.8-kb of the 3' region of *kedE* together with *kedE10* was first cloned from pBS16002 and ligated into the same sites of pANT841 to afford pBS16007. A 4354-bp XmnI-SstI fragment containing the 5' region of *kedE* together with ~400 bp of upstream sequence was next cloned from pBS16002 and ligated into the same sites of pSP72 to afford pBS16008. The 3' region of *kedE* together with *kedE10* was then recovered as an SstI-HindIII fragment from pBS16007 and cloned into the same sites of pBS16008 to yield pBS16009, which contain the complete *kedE-kedE10* cassette. This cassette was moved as a BglII-HindIII fragment into the compatible BamHI-HindIII sites of pETDuet-1 to afford final construct pBS16010 for co-expressing *kedE-kedE10* in *E. coli*.

Co-expression of *kedE-kedE10* in *E. coli* for Heptaene Production

Co-expression of *kedE-kedE10* in *E. coli* was carried out as described previously.^{32,33,53} Thus, pBS16010 was transformed into *E. coli* BL21(DE3) and cultured as described previously, with co-expressions of *sgcE-sgcE10* (pBS1050-pBS1051) as a positive control and of *sgcE* (*C211A*)-*sgcE10* (pBS1065-pBS1051) as a negative control.^{32,33} Briefly, *E. coli* recombinant strains carrying the varying co-expression cassettes were cultured in 50 mL LB medium supplemented with the appropriate antibiotics for selection. The cultures were first grown at 37 °C to an optical density at 600 nm (OD₆₀₀) of ~0.2 and then transfer to 18 °C for continued incubation until they reached OD₆₀₀ ~0.4; upon induction with 0.1 mM IPTG, incubation continued for an additional two days. The cultures were acidified to pH ~3 and harvested by centrifuging to pellet the cells. The cell pellet was extracted by vortexing with 20 mL of acetone. The acetone extract was centrifuged, and the supernatant was concentrated by rotary evaporation to ~1 mL, of which 100 uL was subjected to HPLC analysis. The same HPLC system and Atima-C18 column as described above were used. The column was developed at a flow rate of 1 mL/min with a linear gradient from 40% buffer A (0.1% trifluoroacetic acid in H₂O)/60% buffer B (0.1% trifluoroacetic acid in CH₃OH) to 100% buffer B in 35 min with UV detection at 370 nm. The identity of heptaene was confirmed by comparison with an authentic standard.^{32,33}

Expression of *kedF* in *E. coli* and purification of KedF

The *kedF* gene was amplified by PCR from pBS16002 using Platinum Pfx polymerase (Invitrogen) and the following pair of primers (forward 5'-AAAACCTCTATTTCCAGTCGATGCGCCGCTTCCGCATAGCCG-3'/reverse 5'-TACTTACTTAAATGTTATCAGGCCAGGGAGCGGGCGAACGC-3'). The resultant product was gel-purified and cloned into pBS160011, a variant of pRSFDuet-1 that contains

both a TEV protease recognition site and ligation independent cloning site, to afford the expression construct pBS16012. Under this construct, KedF was overproduced as an N-terminal His₆-tagged fusion protein, whose His₆-tag can be removed upon TEV protease treatment. Introduction of pBS16012 into *E. coli* BL21 (DE3) for *kedF* expression and overproduction and purification of KedF by affinity chromatography using a 5-mL HisTrap HP column (GE Healthcare, Piscataway, NJ) were performed at 4 °C following standard procedures. Immediately following the affinity chromatography, the KedF fraction was diluted to 50 mL with buffer A (50 mM Tris-HCl, pH 8.0, 10 mM NaCl) and loaded on a MonoQ 10/100 column for anion exchange chromatography on an ÄKTA FPLC unit (GE Healthcare). The column was developed at a flow rate of 2 mL/min with a linear gradient from 85% buffer A/15% buffer B (50 mM Tris-HCl, pH 8.0, 1.0 M NaCl) to 40% buffer A/60% buffer B in 40 min. The eluted KedF protein was concentrated with a 30 K MWCO Vivaspın ultrafiltration device (Sartorius, Edgewood, NY) and stored at -80 °C in 100 μ L aliquots. The purified KedF was analyzed by SDS-PAGE on 12% gel. KedF concentration was determined from the absorbance at 280 nm using a molar absorptivity (ϵ 76.89 mM⁻¹ cm⁻¹) calculated according to the deduced KedF amino acid sequence.

In Vitro Characterization of KedF

In vitro characterization of KedF as an epoxide hydrolase was carried out as described previously, using styrene oxide as a substrate mimic.^{48,49} Thus, HPLC-based assays were carried out in 200 μ L reaction mixtures containing 2 mM racemic styrene oxide in 50 mM phosphate buffer, pH 8.0.^{48,49} The reaction was initiated by the addition of 50 μ M KedF, incubated at 25 °C for 1 h, and terminated by extracting the assay mixture with 200 μ L of EtOAc for three times. Negative controls were carried out under the identical conditions in the absence of KedF, while positive controls were carried out under the identical conditions with SgcF instead of KedF.⁴⁸ The combined EtOAc extracts were concentrated in vacuum, and the resulting residue was dissolved in 50 μ L of CH₃CN, 25 μ L of which was subjected to HPLC analysis. HPLC was performed with an Alltech Appollo C18 column (5 μ M, 4.6 x 250 mm, Grace Davison Discovery Sciences), developed at a flow rate of 1 mL/min with a linear gradient from 0 to 60% CH₃CN in H₂O in 20 min with UV detection at 254 nm. The enantiomeric analysis of the vicinal diol products was performed on a Waters HPLC system equipped with 600 pumps, a 996 photodiode array detector, and a Chiralcel OD-H column (5 μ M, 4.6 x 250 mm, Grace Davison Discovery Sciences). The column was eluted isocratically, at a flow rate of 0.7 mL/min, with 2.5% isopropanol in hexane.

Determination of the steady-state kinetic parameters of KedF-catalyzed hydrolysis of (*R*)- or (*S*)-styrene oxide followed the continuous spectrophotometric assay⁶² previously adopted for the SgcF and NcsF2 epoxide hydrolase.^{48,49} Thus, the reactions were carried out in 1-mL reaction mixture containing 10 μ L of 300 mM sodium periodate in DMF, 20 μ L of (*R*)- or (*S*)-styrene oxide in DMSO, with varying concentrations between 0.1 mM and 15 mM, in 50 mM phosphate buffer, pH 8.0. The reactions were initiated by the addition of 9.6 or 4.0 μ M KedF, for (*R*)- or (*S*)-styrene oxide, respectively, and these reactions were carried out in triplicate. The absorbance at 290 nm was monitored in a 1-mL quartz cuvette, thermostated at 25 °C, and the velocity was calculated based on the rate of change of absorbance over 5 to 30 sec. Michaelis-Menten equation was fitted to plots of velocity of 1-phenyl-1, 2-ethanediol formation versus substrate concentration to extract the K_m and k_{cat} values.

Nucleotide Sequence Accession Number

The nucleotide sequence reported in this study is available in the GenBank database under accession number JX679499.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Analytical Instrumentation Center of the School of Pharmacy, University of Wisconsin-Madison for support in obtaining MS data. This work is supported in part by NIH grants CA78747 and CA113297. G.P.H. is the recipient of an NSERC (Canada) postdoctoral fellowship.

References

1. Lam KS, Hesler GA, Gustavson DR, Crosswell AR, Veitch JM, Forenza S, Tomita K. *J Antibiot.* 1991; 44:472–478. [PubMed: 2061190]
2. Hofstead SJ, Matson JA, Malacko AR, Marquardt H. *J Antibiot.* 1992; 45:1250–1254. [PubMed: 1399845]
3. Constantine KL, Colson KL, Wittekind M, Friedrichs MS, Zein N, Tuttle J, Langley DR, Leet JE, Schroeder DR, Lam KS, Farmer BT II, Metzler WJ, Bruccoler RE, Mueller L. *Biochemistry.* 1994; 33:11438–11452. [PubMed: 7918358]
4. Leet JE, Schroeder DR, Hofstead SJ, Golik J, Colson KL, Huang S, Klohr SE, Doyle TW, Matson JA. *J Am Chem Soc.* 1992; 114:7946–7948.
5. Leet JE, Schroeder DR, Langley DR, Colson KL, Huang S, Klohr SE, Lee MS, Golik J, Hofstead SJ, Doyle TW, Matson JA. *J Am Chem Soc.* 1993; 115:8432–8443.
6. Kawata S, Ashizawa S, Hiramata M. *J Am Chem Soc.* 1997; 119:12012–12013.
7. Ren F, Hogan PC, Anderson AJ, Myers AG. *J Am Chem Soc.* 2007; 129:5381–5383. [PubMed: 17417855]
8. Nicolaou KC, Dai WM. *Angew Chem Int Ed.* 1991; 30:1387–1530.
9. Xi, Z.; Goldberg, IG. *Comprehensive Natural Products Chemistry.* Barton, D.; Nakanishi, K.; Meth-Cohn, O., editors. Vol. 7. Elsevier; New York: 1999. p. 553-592.
10. Galm U, Hager MH, Van Lanen SG, Thorson JS, Shen B. *Chem Rev.* 2005; 105:739–758. [PubMed: 15700963]
11. Zein N, Colson KL, Leet JE, Schroeder DR, Solomon W, Doyle TW. *Proc Natl Acad Sci USA.* 1993; 90:2822–2826. [PubMed: 8464895]
12. Kennedy DR, Gawron LS, Ju J, Liu W, Shen B, Beerman TA. *Cancer Res.* 2007; 67:773–781. [PubMed: 17234789]
13. Kennedy DR, Ju J, Shen B, Beerman TA. *Proc Natl Acad Sci USA.* 2007; 104:17632–17637. [PubMed: 17978180]
14. Beerman TA, Gawron LS, Shin S, Shen B, McHugh MM. *Cancer Res.* 2009; 69:593–598. [PubMed: 19147573]
15. Doyle, TW.; Borders, DB., editors. *Enediyne antibiotics as antitumor agents.* Marcel-Dekker; New York: 1995.
16. Maeda, H.; Edo, K.; Ishida, N., editors. *Neocarzinostatin: the past, present, and future of an anticancer drug.* Springer-Verlag; New York: 1997.
17. Thorson JS, Shen B, Whitwam RE, Liu W, Li Y. *Bioorg Chem.* 1999; 27:172–188.
18. Brukner I. *Curr Opin Oncol Endocr Met Invest Drugs.* 2000; 2:344–352.
19. Van Lanen SG, Shen B. *Curr Topics Med Chem.* 2008; 8:448–459.
20. Liang ZX. *Nat Prod Rep.* 2010; 27:499–528. [PubMed: 20336235]
21. Edo K, Mizugaki M, Koide Y, Seto H, Furihata K, Otake N, Ishida N. *Tetrahedron Lett.* 1985; 26:331–340.
22. Lee MD, Dunne TS, Siegel MM, Chang CC, Morton GO, Borders DB. *J Am Chem Soc.* 1987; 109:3464–3466.
23. Nam SJ, Gaudencio SP, Kauffman CA, Jensen PR, Kondratyuk TP, Marler LE, Pezzuto JM, Fenical W. *J Nat Prod.* 2010; 73:1080–1086. [PubMed: 20481500]

24. Liu W, Christenson SD, Standage S, Shen B. *Science*. 2002; 297:1170–1173. [PubMed: 12183628]
25. Liu W, Nonaka K, Nie L, Zhang J, Christenson SD, Bae J, Van Lanen SG, Zazopoulos E, Farnet CM, Yang CF, Shen B. *Chem Biol*. 2005; 12:293–302. [PubMed: 15797213]
26. Van Lanen SG, Oh TJ, Liu W, Wendt-Pienkowski E, Shen B. *J Am Chem Soc*. 2007; 129:13082–13094. [PubMed: 17918933]
27. McGlinchey RP, Nett M, Moore BS. *J Am Chem Soc*. 2008; 130:2406–2407. [PubMed: 18232689]
28. Ahlert J, Shepard E, Lomovskaya N, Zazopoulos E, Staffa A, Bachmann BO, Huang K, Fonstein L, Czisny A, Whitwam RE, Farnet CM, Thorson TS. *Science*. 2002; 297:1173–1176. [PubMed: 12183629]
29. Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, Ahlert J, Thorson JS, Shen B, Farnet CM. *Nat Biotechnol*. 2003; 21:187–190. [PubMed: 12536216]
30. Liu W, Ahlert J, Gao Q, Wendt-Pienkowski E, Shen B, Thorson JS. *Proc Natl Acad Sci USA*. 2003; 100:11959–11963. [PubMed: 14528002]
31. Gao Q, Thorson JS. *FEMS Microbiol Lett*. 2008; 282:105–114. [PubMed: 18328078]
32. Zhang J, Van Lanen SG, Ju J, Liu W, Dorrestein PC, Li W, Kelleher NL, Shen B. *Proc Natl Acad Sci USA*. 2008; 105:1460–1465. [PubMed: 18223152]
33. Horsman GP, Chen Y, Thorson JS, Shen B. *Proc Natl Acad Sci USA*. 2010; 107:11331–11335. [PubMed: 20534556]
34. Liu W, Shen B. *Antimicrob Agents Chemoth*. 2000; 44:382–392.
35. Murrell JM, Liu W, Shen B. *J Nat Prod*. 2004; 67:206–213. [PubMed: 14987060]
36. Christenson SD, Liu W, Toney MD, Shen B. *J Am Chem Soc*. 2003; 125:6062–6063. [PubMed: 12785829]
37. Christenson SD, Wu W, Shen B, Toney MD. *Biochemistry*. 2003; 42:12708–12718. [PubMed: 14580219]
38. Christianson CV, Montavon TJ, Van Lanen SG, Shen B, Bruner SD. *Biochemistry*. 2007; 46:7025–7214.
39. Christianson CV, Montavon TJ, Festin GM, Cooke HA, Shen B, Bruner SD. *J Am Chem Soc*. 2007; 129:15744–15745. [PubMed: 18052279]
40. Montavon TJ, Christianson CV, Festin GM, Shen B, Bruner SD. *Bioorg Med Chem Lett*. 2008; 18:3099–3102. [PubMed: 18078753]
41. Van Lanen SG, Dorrestein PC, Christenson SD, Liu W, Ju J, Kelleher NL, Shen B. *J Am Chem Soc*. 2005; 127:11594–11595. [PubMed: 16104723]
42. Van Lanen SG, Lin S, Dorrestein PC, Kelleher NL, Shen B. *J Biol Chem*. 2006; 281:29633–29640. [PubMed: 16887797]
43. Lin S, Van Lanen SG, Shen B. *J Am Chem Soc*. 2007; 129:12432–12438. [PubMed: 17887753]
44. Van Lanen SG, Lin S, Shen B. *Proc Natl Acad Sci USA*. 2008; 105:494–499. [PubMed: 18182490]
45. Lin S, Van Lanen SG, Shen B. *J Am Chem Soc*. 2008; 130:6616–6623. [PubMed: 18426211]
46. Lin S, Van Lanen SG, Shen B. *Proc Natl Acad Sci USA*. 2009; 106:4183–4188. [PubMed: 19246381]
47. Van Lanen SG, Lin S, Horsman GP, Shen B. *FEMS Microbiol Lett*. 2009; 300:237–241. [PubMed: 19817865]
48. Lin S, Horsman GP, Shen B. *J Am Chem Soc*. 2009; 131:16410–16417. [PubMed: 19856960]
49. Lin S, Horsman GP, Shen B. *Org Lett*. 2010; 12:3816–3819. [PubMed: 20704329]
50. Cooke HA, Zhang J, Griffin MA, Nonaka K, Van Lanen SG, Shen B, Bruner SD. *J Am Chem Soc*. 2007; 129:7728–7729. [PubMed: 17539640]
51. Luo Y, Lin S, Zhang J, Cooke HA, Bruner SD, Shen B. *J Biol Chem*. 2008; 283:14694–14702. [PubMed: 18387946]
52. Cooke HA, Guenther EL, Luo Y, Shen B, Bruner SD. *Biochemistry*. 2009; 48:9590–9598. [PubMed: 19702337]
53. Horsman GP, Van Lanen SG, Shen B. *Methods Enzymol*. 2009; 459:97–112. [PubMed: 19362637]

54. Ling J, Horsman GP, Huang SX, Luo Y, Lin S, Shen B. *J Am Chem Soc.* 2010; 132:12534–12536. [PubMed: 20718468]
55. Lin S, Huang T, Horsman GP, Huang SX, Guo X, Shen B. *Org Lett.* 2012; 14:2300–2303. [PubMed: 22519717]
56. Lohman JR, Shen B. *Methods Enzymol.* 2012; 516:299–319. [PubMed: 23034235]
57. Lin S, Huang T, Shen B. *Methods Enzymol.* 2012; 516:321–343. [PubMed: 23034236]
58. Myers AG, Hurd AR, Hogan PC. *J Am Chem Soc.* 2002; 124:4583–4585. [PubMed: 11971704]
59. Ogawa K, Koyama Y, Ohashi I, Sato I, Hirama M. *Angew Chem Int Ed.* 2009; 48:1110–1113.
60. Kieser, T.; Bibb, MJ.; Buttner, MM.; Chater, KF.; Hopwood, DA. *Practical Streptomyces Genetics.* The John Innes Foundation; Norwich, UK: 2000.
61. Hang VTT, Kim TS, Oh TJ, Sohng K. *Biotechnol Bioproc Eng.* 2011; 16:462–469.
62. Mateo C, Archelas A, Furstoss R. *Anal Biochem.* 2003; 314:135–141. [PubMed: 12633612]
63. Thibodeaux CJ, Melancon CE III, Liu H-w. *Nature.* 2007; 446:10081016.
64. Thibodeaux CJ, Melancon CE III, Liu H-w. *Angew Chem Int Ed.* 2008; 47:9814–9858.
65. Tao M, Wang L, Wendt-Pienkowski E, George NP, Galm U, Zhang G, Coughlin JM, Shen B. *Mol Biosyst.* 2007; 3:60–74. [PubMed: 17216057]
66. Inouye S, Shomura T, Tsuruoka T, Ogawa Y, Watanabe H, Yoshida J, Niida T. *Chem Pharm Bull.* 1975; 23:2669–2677. [PubMed: 1218439]
67. Kwon HJ, Smith WC, Scharon AJ, Hwang SH, Kurth MJ, Shen B. *Science.* 2002; 297:1327–1330. [PubMed: 12193782]
68. Kusebauch B, Busch B, Scherlach K, Roth M, Hertweck C. *Angew Chem Int Ed.* 2009; 48:5001–5004.
69. Shen B. *Curr Opinion Chem Biol.* 2003; 7:285–295.
70. Rachid S, Scharfe M, Bloecker H, Weissman KJ, Mueller R. *Chem Biol.* 2009; 16:70–81. [PubMed: 19171307]
71. Rajgarhia VB, Strohl WR. *J Bacteriol.* 1997; 179:2690–2696. [PubMed: 9098068]
72. Lukashin AV, Borodovsky M. *Nucleic Acids Res.* 1998; 26:1107–1115. [PubMed: 9461475]
73. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJA, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C. *Nucleic Acids Res.* 2009; 37:D211–215. [PubMed: 18940856]
74. Sambrook, J.; Fritsch, EF.; Maniatis, T. *Molecular Cloning: A Laboratory Manual.* 3. Cold Spring Harbor Laboratory; Cold Spring Harbor, NY: 2000.

Fig. 1. Structures of the KED enediyne chromophore and the proposed aromatized product^{58,59}

Fig. 2.

The *ked* biosynthetic gene cluster from *Streptoalloteichus* sp. ATCC 53650. (A) The sequenced 135-kb DNA region encompassed by five overlapping cosmids pBS16002, pBS16003, pBS16004, pBS16005, and pBS16006. Probe-1, -2, -3, -4, and -5 were used to isolate the overlapping cosmids from a *Streptoalloteichus* sp. ATCC 53650 genomic library. (B) Genetic organization of the *ked* biosynthetic gene cluster. Solid black indicates the region whose gene products are predicted to be involved in KED biosynthesis (~105 kb). Proposed functions for individual orfs are pattern-coded and summarized in Table 1.

Fig 3.

Proposed biosynthetic pathway for the KED chromophore: (A) L-mycarose and kedarsamine from D-glucose-1-phosphate; (B) (*R*)-2-aza-3-chloro- β -tyrosine from 2-aza-phenylalanine, (C) 3-hydroxy-7,8-dimethoxy-6-isopropoxy-2-naphthoic acid from 3,6,8-trihydroxy-2-naphthoic acid; and (D) the enediyne core from acetate and a convergent assembly of the four components to yield the KED chromophore.

Fig. 4. HPLC chromatograms with UV detection at 370 nm showing production of heptaene (●) upon co-expression of *kedE-kedE10* in *E. coli*: (I) *sgcE-sgcE10* as a positive control; (II) *sgcE(C211A)-sgcE10* as a negative control; and (iii) *kedE-kedE10*.

Fig. 5. In vitro characterization of KedF as an epoxide hydrolase using (*R*)- and (*S*)-styrene epoxide as substrate mimics, preferring (*S*)-styrene epoxide with a 4.0-fold greater specificity. (A) Regio- and stereoselectivity of KedF-catalyzed hydrolysis of (*R*)- and (*S*)-styrene epoxide and HPLC chromatograms with UV detection at 254 nm showing (*R*)- and (*S*)-1-phenyl-1,2-ethanediol, respectively, as the major product: (I) (*S*)-1-phenyl-1,2-ethanediol standard, (II) KedF with (*S*)-styrene epoxide, (III) (*R*)-1-phenyl-1,2-ethanediol standard, (IV) KedF with (*R*)-styrene epoxide, (V) (*R*)- and (*S*)-1-phenyl-1,2-ethanediol standards, (VI) KedF with racemic styrene epoxide; (*R*)- (◆) and (*S*)-1-phenyl-1,2-ethanediol (◇). (B) Steady-state kinetic analysis of KedF-catalyzed hydrolysis of (*R*)- and (*S*)-styrene epoxide showing single substrate kinetic plots for (I) (*R*)-styrene epoxide and (II) (*S*)-styrene epoxide.

Fig. 6. Genetic organization of the five 9- and three 10-membered enediyne biosynthetic gene clusters known to date highlighting the five-gene enediyne PKS cassettes (black) that are absolutely conserved among both 9- and 10-membered enediyne clusters and the varying number of conserved genes that encode the 9-membered enediyne pathway specific accessory enzymes (gray). 9-Membered enediynes including: C-1027; NCS, neocarzinostatin; MDP, maduropeptin, KED, kedarcidin, SPO, sporolide. 10-Membered enediynes including: CAL, calicheamicin; ESP, esperamicin; DYN, dynemicin.

Table 1

Deduced functions of open reading frames in the *ked* biosynthetic gene cluster

Gene	Amino acids ^a	Protein homologs ^b	Identity%/Similarity%	Deduced function	Proposed roles in KED biosynthesis
<i>orf(-33) to orf(-1)</i>			ORFs predicted to be beyond the upstream boundary		
<i>kedE11</i>	267	SgcE11 (AAL06691)	59/70	Unknown	Enediynes core biosynthesis
<i>kedM</i>	335	SgcM (AAL06686)	46/55	Unknown	Enediynes core biosynthesis
<i>kedU1</i>	221	Amir_4121 (ACU37976)	32/44	Hypothetical protein	Unknown
<i>kedS</i>	182	SgcS (AAL06705)	60/72	Unknown	Enediynes core biosynthesis
<i>kedE9</i>	546	SgcE9 (AAL06693)	76/86	Ketoreductase	Enediynes core biosynthesis
<i>kedE8</i>	190	SgcE8 (AAL06694)	62/74	Unknown	Enediynes core biosynthesis
<i>kedR2</i>	259	SgcR2 (AAL06696)	50/63	AraC-like transcriptional regulator	Regulation
<i>kedE7</i>	447	SgcE7 (AAL06697)	63/74	P-450 monooxygenase	Enediynes core biosynthesis
<i>kedU2</i>	123	None	--/--	Hypothetical protein	Unknown
<i>kedE5</i>	351	SgcE5 (AAL06700)	61/70	Unknown	Enediynes core biosynthesis
<i>kedE4</i>	649	SgcE4 (AAL06701)	58/73	Unknown	Enediynes core biosynthesis
<i>kedE3</i>	328	SgcE3 (AAL06702)	54/62	Unknown	Enediynes core biosynthesis
<i>kedE2</i>	342	SgcE2 (AAL06703)	59/70	Unknown	Enediynes core biosynthesis
<i>kedE1</i>	320	SgcE1 (AAL06710)	41/65	Unknown	Enediynes core biosynthesis
<i>kedE</i>	1919	SgcE (AAL06703)	55/66	Polyketide synthase	Enediynes core biosynthesis
<i>kedE10</i>	148	SgcE10 (AAL06692)	65/77	Thioesterase	Enediynes core biosynthesis
<i>kedE6</i>	164	SgcE6 (AAL06698)	51/64	Flavin reductase	Enediynes core biosynthesis
<i>kedL</i>	391	SgcL (AAL06685)	62/75	Enoyl reductase	Enediynes core biosynthesis
<i>kedI</i>	141	SgcI (AAL06676)	62/72	Unknown	Enediynes core biosynthesis
<i>kedD2</i>	464	SgcD2 (AAL06699)	62/75	FAD dependent monooxygenase	Enediynes core biosynthesis
<i>kedN3</i>	409	NcsB3 (AAM77997)	49/64	P-450 monooxygenase	Naphthoic acid biosynthesis
<i>kedF</i>	385	SgcF (AAL06662)	64/77	Epoxide hydrolase	Enediynes core biosynthesis
<i>kedU3</i>	240	CalU12 (AAM94790)	30/40	Unknown (thioredoxin-like)	Unknown
<i>kedX2</i>	756	SgcB2 (AAL06654)	52/70	Efflux pump	Resistance
<i>kedS2</i>	458	LanS (AAD13549)	65/77	NDP-hexose 2,3-dehydratase	Sugar biosynthesis
<i>kedY</i>	514	SgcC (AAL06674)	58/70	FAD-dependent monooxygenase	β -Azatryptophan biosynthesis
<i>kedN1</i>	334	SgcD4 (AAL06683)	56/75	O-methyltransferase	Naphthoic acid biosynthesis
<i>kedS7</i>	383	MdpA5 (ABY66023)	66/75	aminotransferase	Sugar biosynthesis

Gene	Amino acids ^a	Protein homologs ^b	Identity%/Similarity%	Deduced function	Proposed roles in KED biosynthesis
<i>kedS8</i>	249	SgcA5 (AAL06660)	45/56	N-methyltransferase	Sugar biosynthesis
<i>kedS9</i>	244	SgcA5 (AAL06660)	53/68	N-methyltransferase	Sugar biosynthesis
<i>kedS10</i>	428	SgcA6 (AAL06670)	36/51	Glycosyltransferase	Sugar moiety coupling
<i>kedR1</i>	1088	Strop_2737 (ABP55181)	51/63	Transcriptional regulator	Regulation
<i>kedN5</i>	627	Sare_2941 (ABY98767)	55/70	Radical SAM C-methyltransferase	Naphthoic acid biosynthesis
<i>kedN4</i>	425	SAV_4024 (Q82G74)	34/46	Acyl-CoA N-acyltransferase	Naphthoic acid biosynthesis
<i>kedS6</i>	417	SgcA6 (AAL06670)	34/50	Glycosyltransferase	Sugar moiety coupling
<i>kedU4</i>	73	RHA1_ro00868 (ABG92701)	52/72	Hypothetical protein	Unknown
<i>kedU11</i>	367	Strop_2833 (ABP55274)	68/79	Monooxygenase	Unknown type II PKS locus ^d
<i>kedU12</i>	328	Strop_2814 (ABP55255)	48/55	Enoyl reductase	
<i>kedU13</i>	383	Strop_2813 (ABP55254)	61/73	Enoyl reductase	
<i>kedU14</i>	397	Strop_2801 (ABP55242)	71/81	Acyltransferase	
<i>kedU15</i>	407	Strop_2800 (ABP55241)	68/78	CoA transferase	
<i>kedU16</i>	247	Strop_2799 (ABP55240)	71/81	Ketoreductase	
<i>kedU17</i>	152	Strop_2798 (ABP55239)	43/53	Dehydratase	
<i>kedU18</i>	164	Strop_2797 (ABP55238)	59/72	Dehydratase	
<i>kedU19</i>	78	Strop_2796 (ABP55237)	76/93	ACP	
<i>kedU20</i>	402	Strop_2795 (ABP55236)	70/80	Ketosynthase	
<i>kedU21</i>	353	Strop_2794 (ABP55235)	55/64	Ketosynthase ^c	
<i>kedU22</i>	375	Strop_2793 (ABP55234)	66/76	Ketosynthase	
<i>kedU23</i>	290	Strop_2792 (ABP55233)	54/62	Ketosynthase ^c	
<i>kedU24</i>	268	Strop_2791 (ABP55232)	50/61	Unknown	
<i>kedU25</i>	307	Strop_2780 (ABP55231)	64/75	Thioesterase	
<i>kedU26</i>	252	Strop_2789 (ABP55230)	66/73	Isomerase	
<i>kedU27</i>	246	Strop_2788 (ABP55229)	73/82	Aldolase	
<i>kedU28</i>	1042	Strop_2787 (ABP55228)/Strop_2786 (ABP55227)	59/71 65/77	Acyl-CoA synthetase/P-450 monooxygenase	
<i>kedS5</i>	326	SgcA (AAL06671)	25/37	NDP-hexose oxidoreductase	Sugar biosynthesis
<i>kedS4</i>	427	SgcA3 (AAL06661)	31/49	C-methyltransferase	Sugar biosynthesis
<i>kedN2</i>	553	NesB2 (AAM77987)	54/64	Acyl-CoA synthetase	Naphthoic acid biosynthesis
<i>kedA</i>	146	CaqA (AAL06658)	41/58	Apoptein	Resistance
<i>kedX</i>	561	SgcB (AAL06672)	49/68	Efflux pump	Resistance

Gene	Amino acids ^a	Protein homologs ^b	Identity%/Similarity%	Deduced function	Proposed roles in KED biosynthesis
<i>kedY4</i>	539	SgcC4 (AAL06680)	66/82	Tyrosine aminomutase	β-Azatyrosine biosynthesis
<i>kedY1</i>	1172	SgcC1 (AAL06681)	36/44	NRPS adenylation enzyme	β-Azatyrosine biosynthesis
<i>kedY5</i>	452	SgcC5 (AAL06678)	40/55	NRPS condensation enzyme	β-Azatyrosine moiety coupling
<i>kedU31</i>	497	SSHG_05343 (EFE84901)	32/44	Enoyl reductase	Unknown type I PKS locus ^d
<i>kedU32</i>	389	M23134_01012 (EAY30688)	30/48	Enoyl reductase	
<i>kedU33</i>	495	lcfB (O07610)	25/44	Acyl-CoA synthetase	
<i>kedU34</i>	92	SSHG_05345 (EFE84903)	30/58	ACP	
<i>kedU35</i>	297	Celf_2318 (AEE46445)	26/35	Unknown	
<i>kedU36</i>	247	DapB (Q97G18)	29/47	Dihydrodipicolinate reductase	
<i>kedU37</i>	635	LcfB (O07610)	30/48	Acyl-ACP synthetase	
<i>kedU38</i>	1803	Sros_2423 (ACZ85401)	44/56	Type I PKS (KS-AT-DH-KR-ACP) ^e	
<i>kedU39</i>	291	Haur_3968 (ABX06600)	58/76	Enoyl reductase	
<i>kedU40</i>	84	Haur_3969 (ABX06601)	49/74	ACP	
<i>kedU41</i>	398	Haur_3970 (ABX06602)	54/70	Enoyl reductase	
<i>kedU42</i>	356	Hoch_2947 (ACY15461)	59/75	Unknown	
<i>kedU43</i>	248	GrsT (P14686)	43/57	Thioesterase	
<i>kedU44</i>	183	Sare_0361 (ABV96290)	37/47	Hypothetical protein	
<i>kedU45</i>	409	CYP107B1 (P33271)	48/62	P-450 monooxygenase	
<i>kedR3</i>	259	SgcR2 (AAL06696)	35/49	Transcriptional regulator	Regulation
<i>kedY2</i>	90	SgcC2 (AAL06679)	59/67	NRPS PCP	β-Azatyrosine biosynthesis
<i>kedY3</i>	492	SgcC3 (AAL06656)	69/81	FAD dependent halogenase	β-Azatyrosine biosynthesis
<i>kedS3</i>	333	KijD10 (ACB46498)	58/72	NDP-hexose oxidoreductase	Sugar biosynthesis
<i>kedS1</i>	194	SgcS1 (AAL06668)	42/60	NDP-hexose epimerase	Sugar biosynthesis
<i>orf1 and orf2</i>			ORFs predicted to be beyond the downstream boundary		

^aNumbers are in amino acids.

^bGiven in parentheses are NCBI accession numbers. Homologues from the C-1027 pathway were selected from comparison. If no homologue was found within the C-1027 cluster, homologues from NCS and MDP clusters were preferred over others in the GeneBank.

^cNonfunctional on the basis of the mutated P-D-A (for KedU21) or A-D-G (KedU23) active site triad C-H-H of acyl-ACP ketosynthases.

^dWhile the overall organization of and the genes within the KED biosynthetic gene cluster show high homology to other known 9-membered enediyne biosynthetic gene clusters, including, C-1027, ²³NCS, ²⁴and MDP, ²⁵there are two loci, *kedU11-kedU28*, termed type II PKS locus, and *kedU31-kedU45*, termed type I PKS locus, within the KED cluster (shaded in grey) whose roles in KED biosynthesis cannot be proposed on the basis of bioinformatics.

^eThe KedU38 type I PKS consists of five domains (KS, ketosynthase; AT, acyltransferase; DH, dehydratase; KE, ketoreductase; ACP, acyl carrier protein).