A free-standing condensation enzyme catalyzing ester bond formation in C-1027 biosynthesis

Shuangjun Lin[a], Steven G. Van Lanen[a,1], and Ben Shen[a,b,c,2]

[a]Division of Pharmaceutical Sciences, [b]University of Wisconsin National Cooperative Drug Discovery Group, and [c]Department of Chemistry, University of Wisconsin, Madison, WI 53705

Edited by Craig A. Townsend, Johns Hopkins University, Baltimore, MD, and accepted by the Editorial Board January 6, 2009 (received for review September 10, 2008)

Nonribosomal peptide synthetases (NRPSs) catalyze the biosynthesis of many biologically active peptides and typically are modular, with each extension module minimally consisting of a condensation, an adenylation, and a peptidyl carrier protein domain responsible for incorporation of an amino acid into the growing peptide chain. C-1027 is a chromoprotein antitumor antibiotic whose enediyne chromophore consists of an enediyne core, a deoxy aminosugar, a benzoxazolinate, and a β-amino acid moiety. Bioinformatics analysis suggested that the activation and incorporation of the β-amino acid moiety into C-1027 follows an NRPS mechanism whereby biosynthetic intermediates are tethered to the peptidyl carrier protein SgcC2. Here, we report the biochemical characterization of SgcC5, an NRPS condensation enzyme that catalyzes ester bond formation between the SgcC2-tethered (*S*)-3-chloro-5-hydroxy-β-tyrosine and (*R*)-1-phenyl-1,2-ethanediol, a mimic of the enediyne core. SgcC5 uses (*S*)-3-chloro-5-hydroxy-β-tyrosyl-SgcC2 as the donor substrate and exhibits regiospecificity for the C-2 hydroxyl group of the enediyne core mimic as the acceptor substrate. Remarkably, SgcC5 is also capable of catalyzing amide bond formation, albeit with significantly reduced efficiency, between (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and (*R*)-2-amino-1-phenyl-1-ethanol, an alternative enediyne core mimic bearing an amine at its C-2 position. Thus, SgcC5 is capable of catalyzing both ester and amide bond formation, providing an evolutionary link between amide- and ester-forming condensation enzymes.

enediyne | nonribosomal peptide synthetase

Nonribosomal peptide synthetases (NRPSs) are large multifunctional proteins that catalyze the synthesis of many pharmaceutically important peptides, including the anticancer drugs bleomycin and actinomycin and antibacterial antibiotics vancomycin and daptomycin. The prototypical NRPS is composed of loading, extension, and termination modules, with each extension module consisting of minimally 3 domains—an adenylation (A) domain, a peptidyl carrier protein (PCP) domain, and a condensation (C) domain (1–4). The A domain specifically selects an amino acid, activates it by formation of an aminoacyl adenylate, and transfers the activated substrate to the thiol group of the 4′-phosphopantetheinyl arm of the PCP domain to yield a thioester-linked amino acid. The C domain then catalyzes nucleophilic condensation between upstream and downstream PCP-tethered amino acids (also known as donor and acceptor substrates, respectively; ref. 5) to form a new amide bond, thus extending the growing peptide chain by 1 amino acid.

Although NRPS catalysis is often described to follow assembly-line enzymology, wherein the amino acid sequence of the peptide product can be directly predicted from the molecular architecture of NRPS domains and modules (3, 4), recent studies have revealed numerous NRPSs that deviate from this assembly-line molecular logic (6–8). For example, variations to the standard C-A-PCP module were found in the kutzneride (9) and bleomycin NRPSs (10), and iteratively acting domains have been observed for capreomycin (11), saframycin (12), and enterobactin (13) biosynthesis. Moreover, freestanding C, A, or PCP proteins may function *in trans* to form an NRPS module (8), and such freestanding proteins or didomain proteins have also been identified in biosynthetic machinery for natural products that do not even contain a peptide moiety. One such case was identified upon characterization of the biosynthetic gene cluster for the enediyne antitumor antibiotic C-1027, a cytotoxic secondary metabolite isolated from *Streptomyces globisporus* that undergoes an electronic rearrangement that, when in the presence of molecular oxygen, can lead to double-stranded DNA breaks (Fig. 1*B*) (14). In the C-1027 gene cluster, a freestanding C domain (SgcC5), A domain (SgcC1), and PCP (SgcC2) were identified, and together they constitute a minimal NRPS extension module that was predicted to be involved in the biosynthesis of the (*S*)-3-chloro-5-hydroxy-β-tyrosine moiety of C-1027 (Fig. 1) (15).

Although NRPS C domains are known to catalyze amide bond formation between 2 PCP-tethered amino acids (3, 4), examination of the C-1027 chromophore structure suggests that SgcC5 instead catalyzes condensation to generate an ester bond by using a single PCP-tethered amino acid (Fig. 1*B*). C domains have been recently proposed to catalyze ester bond formation with carrier protein-tethered substrates, although they are homologous to the typical amide-forming C enzymes. For instance, the C domains located at the C terminus of the RapP and FkbP proteins from the rapamycin (16) and FK506 (17) NRPS biosynthetic machinery, respectively, were proposed to catalyze intramolecular cyclization via ester bond formation, hence releasing the final products from the carrier proteins. These C domains therefore act in a functional analogy to thioesterase domains, despite having no sequence homology between them. Some C domains embedded in elongation modules have also been proposed to catalyze chain extension via ester bond formation rather than amide bond formation on the basis of comparing the NRPS architecture with the expected product. This includes the nonribosomal peptides kutzneride (9), valinomycin (18), and cereulide (19), and the nonribosomal peptide-polyketide hybrid cryptophycin 1 (20). Recently, the C domain of a PCP-C didomain protein Fum14 in the biosynthesis of the fungal metabolite fumonisin was shown to generate 2 tricarballylic esters by using a thioester of *N*-acetylcysteamine as an acyl donor mimicking the phosphopantetheinyl group of the carrier protein, and in this example, the acceptor substrate was the free alcohol of the fumonisin precursor (21). This study provided the first biochemical evidence for a C domain to catalyze ester bond

Author contributions: B.S. designed research; S.L. performed research; S.L., S.G.V.L., and B.S. analyzed data; and S.L., S.G.V.L., and B.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. C.A.T. is a guest editor invited by the Editorial Board.

[1]Present address: Department of Pharmaceutical Sciences, University of Kentucky, Lexington, KY 40536.

[2]To whom correspondence should be addressed. E-mail: bshen@pharmacy.wisc.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808880106/DCSupplemental.

BIOCHEMISTRY

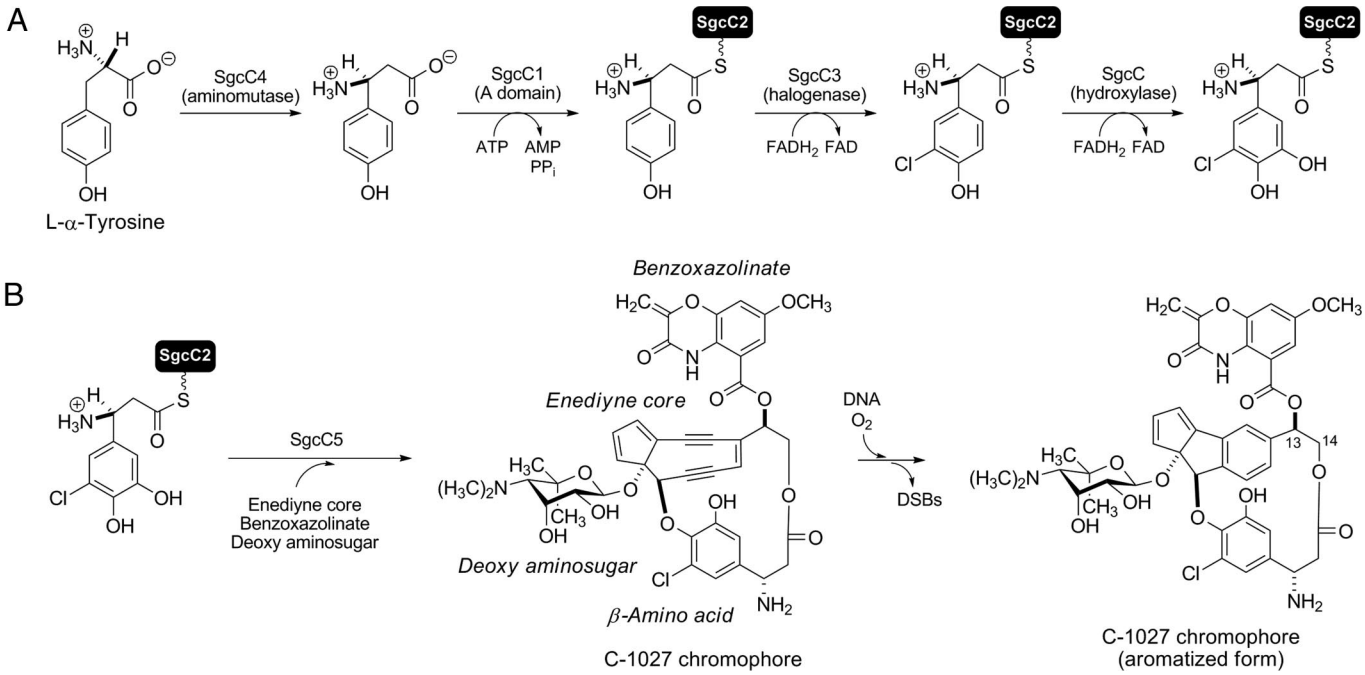

Fig. 1. Biosynthesis of (*S*)-3-chloro-5-hydroxy-β-tyrosine and its incorporation into the C-1027 chromophore. (*A*) Biosynthesis of the β-amino acid moiety from L-α-tyrosine featuring 4 functionally assigned enzymes. (*B*) Attachment of (*S*)-3-chloro-5-hydroxy-β-tyrosine onto the enediyne core by SgcC5, leading to the C-1027 chromophore. The timing for each of the coupling steps is unknown. C-1027 undergoes an electronic rearrangement that generates a benzenoid biradical species that can abstract hydrogen atoms from DNA, leading to DNA breaks and an aromatized C-1027 chromophore. DSBs indicates double-strand breaks.

formation. The results suggested, as observed for the freestanding C enzyme VibH (22), that C domains can also use a free substrate as the acceptor nucleophile to couple with a carrier protein-tethered acyl donor for ester (as with Fum14) or amide (as with VibH) bond formation.

We have previously characterized the first 4 enzymatic steps of the biosynthetic pathway for the (*S*)-3-chloro-5-hydroxy-β-tyrosine moiety of C-1027 leading to (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 (Fig. 1*A*), which includes generation of (*S*)-β-tyrosine by the aminomutase SgcC4 (23, 24), activation and loading of (*S*)-β-tyrosine to SgcC2 by the A enzyme SgcC1 (25, 26), and halogenation and hydroxylation of (*S*)-β-tyrosyl-S-SgcC2 by the FAD-dependent SgcC3 halogenase and SgcC hydroxylase, respectively (27, 28). We now report that SgcC5 does indeed catalyze ester bond formation between the SgcC2-tethered (*S*)-3-chloro-5-hydroxy-β-tyrosine and (*R*)-1-phenyl-1,2-ethanediol, an enediyne core analog that mimics the aromatized C-1027 structure (structure **1a**; Figs. 1*B* and 2*A*), thereby supporting its proposed role in C-1027 biosynthesis (15). SgcC5 absolutely requires carrier protein-tethered molecules as the donor substrates and is regiospecific for the C-2 hydroxyl group of the enediyne core mimic as the acceptor substrate (C-2 of **1a** corresponds to C-14 of the enediyne core; Figs. 1*B* and 2*A*). Remarkably, SgcC5 is also competent and regiospecific in catalyzing amide bond formation between the SgcC2-tethered (*S*)-3-chloro-5-hydroxy-β-tyrosine and (*R*)-2-amino-1-phenyl-1-ethanol, an enediyne core mimic bearing an amine as a nucleophile at its C-2 position (**2a**; Fig. 2*A*), as an alternative acceptor substrate. The broad acceptor tolerance of SgcC5 contrasts with the strict substrate specificity of other previously characterized C domains (3, 5, 29), although SgcC5 does exhibit a strong regiospecificity for the C-2 group as a nucleophile (i.e., C-2 hydroxyl or amino group) of the acceptor substrate. SgcC5 is an example of a C enzyme that catalyzes ester bond formation but retains the canonical amide-forming activity, albeit with significantly reduced efficiency. These findings provide biochemical evidence supporting an evolutionary link between amide- and ester-forming C enzymes that has been implied from structural similarities between C domains and chloramphenicol acetyltransferases (CATs) (30). Further mechanistic and structural characterization of SgcC5 should therefore allow us to engineer the ester-forming activity into other NRPS C domains or enzymes, thereby expanding the size and diversity of the novel peptide library by combinatorial biosynthesis or chemoenzymatic methods.

Results and Discussion

Bioinformatics Analysis of SgcC5. The biosynthesis of the β-amino acid moiety and its incorporation into C-1027 are catalyzed by 6 proteins: SgcC, SgcC1, SgcC2, SgcC3, SgcC4, and SgcC5 (Fig. 1) (15). Among them, SgcC1 and SgcC2 have been identified as homologs to A domains and PCP domains in NRPSs, respectively, and the activity of SgcC1 to activate and load (*S*)-β-tyrosine to SgcC2 has been demonstrated (25, 26). SgcC5 features the highly conserved $HHXXXDX_{14}Y$ motif known to all modular NRPS C domains [supporting information (SI) Fig. S1] (22, 31–34), clearly supporting its functional assignment as a C enzyme but putatively catalyzing an ester bond formation (Fig. 1*B*). However, the overall sequence homology between SgcC5 and other well-studied NRPS C domains is low—SgcC5 shows only 19% identity/36% similarity to the SrfA-C C domain, 20% identity/38% similarity to the TycC C domain, and 16% identity/27% similarity to the freestanding C enzyme VibH, whose structures have all been solved by X-ray crystallography (22, 33, 34).

Overproduction of SgcC5 and Chemoenzymatic Preparation of Donor Substrate. SgcC5 was overproduced in *Escherichia coli* BL21 (DE3), purified to homogeneity as an N-terminal, His_6-tagged fusion protein (≈100 mg/L), and showed a single protein band upon SDS/PAGE analysis consistent with the predicted molecular weight (55 kDa) (Fig. S2).

Three recombinant proteins—SgcC1 (25), SgcC2 (27), and

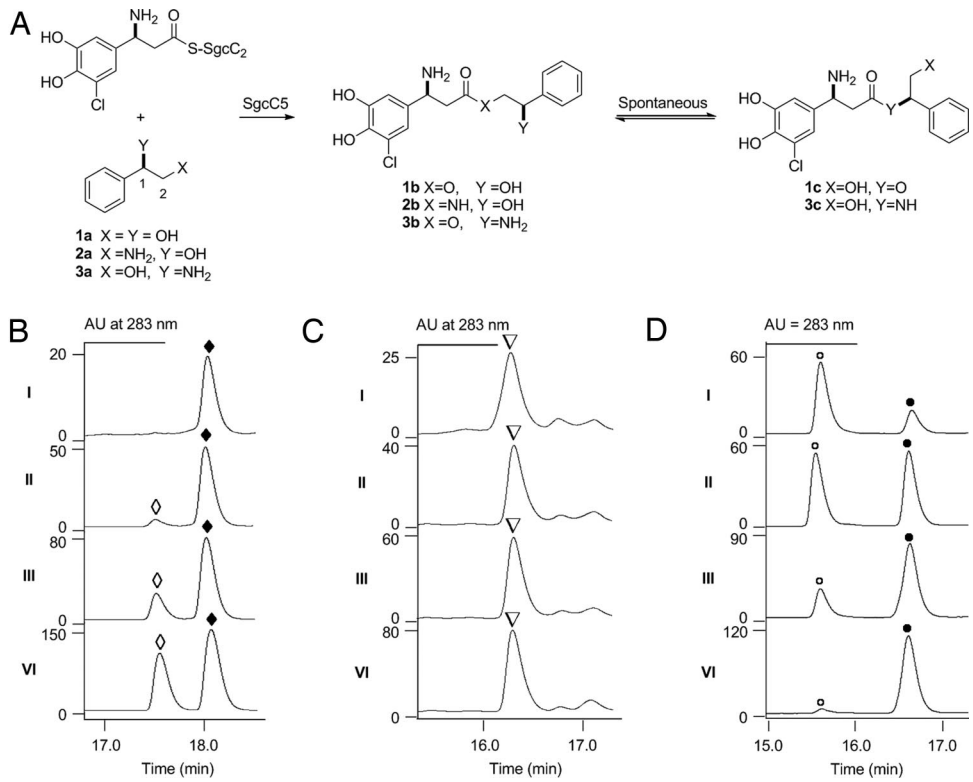

Fig. 2. Characterization of SgcC5 as a condensation enzyme catalyzing both ester bond and amide bond formation. (*A*) SgcC5-catalyzed condensation reaction between SgcC2-tethered (*S*)-3-chloro-5-hydroxy-β-tyrosine and enediyne core mimics (*R*)-1-phenyl-1,2-ethanediol (**1a**), (*R*)-2-amino-1-phenyl-1-ethanol (**2a**), or (*R*)-2-phenylglycinol (**3a**) and the products established by spectroscopic analysis. HPLC profiles of time course analysis for SgcC5-catalzyed condensation between (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 incubated with (*B*) **1a** and 1 μM SgcC5 for 1 min (*i*), 4 min (*ii*), 15 min (*iii*), and 90 min (*iv*), (*C*) **2a** and 100 μM SgcC5 for 5 min (*i*), 15 min (*ii*), 60 min (*iii*), and 180 min (*iv*), and (*D*) **3a** and 20 μM SgcC5 for 1 min (*i*), 2.5 min (*ii*), 5 min (*iii*), and 10 min (*iv*). **1b** (◆), **1c** (◇), **2b** (▽), **3b** (○), and **3c** (●).

Svp, a promiscuous 4-phosphopantetheinyl transferase from the bleomycin biosynthetic pathway (35)—were produced as described previously to enzymatically prepare the SgcC2-tethered donor substrate. Apo-SgcC2 was first converted to holo-SgcC2 by Svp (35). The free substrate 3-chloro-5-hydroxy-β-tyrosine was chemically synthesized (27), activated, and loaded onto holo-SgcC2 by SgcC1 to yield (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2. SgcC1 activated and loaded only the (*S*)-3-chloro-5-hydroxy-β-tyrosine enantiomer to holo-SgcC2, despite using racemic 3-chloro-5-hydroxy-β-tyrosine, as determined by a chiral HPLC analysis in comparison with authentic (*S*)-3-chloro-5-hydroxyl-β-tyrosine and (*R*)-3-chloro-5-hydroxyl-β-tyrosine, respectively (Fig. S3 and *SI Methods*).

In Vitro Characterization of SgcC5 as an Ester-Forming Condensation Enzyme. SgcC5 was initially assayed with 3-chloro-5-hydroxy-β-tyrosine as the donor substrate and (*R*)-1-phenyl-1,2-ethanediol (**1a**; Fig. 2*A*) as the acceptor substrate, the latter of which serves as an enediyne core mimic representing the aromatized version of the C-1027 enediyne chromophore (the exact structure of the native substrate for SgcC5 remains unknown; hence, it is unavailable) (Fig. 1*B*). HPLC analysis of the assay solution, however, showed no product formation, suggesting that SgcC5 may require carrier protein-tethered substrate. SgcC5 was next assayed with (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and (*R*)-1-phenyl-1,2-ethanediol as the donor and acceptor substrates, respectively, and the reaction was followed by HPLC analysis (Fig. 2*B*). After 15 min of incubation, 2 products formed, with the major one at retention time (rt) 18.2 min and the minor one at rt 17.6 min, and the 2 products reached a 4:5 ratio after 90 min of incubation (Fig. 2*B iii* and *iv*). Each of the 2 products was isolated and subjected to electrospray ionization mass spectroscopy (ESI-MS) analysis, yielding the same pair of $[M + H]^+$ ions at 352.1 and 354.1 in a 3:1 ratio, characteristic of monochlorinated species, which was consistent with the molecular weight of the predicted ester products **1b** and **1c** (calculated $[M + H]^+$ ions at m/z = 352.1 and 354.1; Fig. 2*A*). A large-scale reaction was then carried out to produce both products (Fig. S4 and *SI Methods*), and ^{1}H, ^{13}C, and 2D NMR spectroscopic analyses (Table S1) confirmed them as the ester products arising from 3-chloro-5-hydroxy-β-tyrosine coupled to the C-2 and C-1 hydroxyl groups of (*R*)-1-phenyl-1,2-ethanediol, respectively (**1b** and **1c**; Fig. 2*A*).

Kinetic Analysis of Esterification by SgcC5. HPLC analysis following the SgcC5-catalyzed ester formation between (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and (*R*)-1-phenyl-1,2-ethanediol showed that the enzymatic reaction was time-dependent, enabling the determination of single-substrate kinetic constants (Fig. 2*B*). (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 showed substrate inhibition at concentrations >200 μM, and the kinetics constants obtained therefore are only estimates. The formation of product with constant (*R*)-1-phenyl-1,2-ethanediol (5 mM) and variable (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 exhibited Michaelis–Menten kinetics, with a K_m = 71 ± 9 μM and a k_{cat} = 22 ± 2 min^{-1} (Fig. 3*A*), and assays with 200 μM (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and variable (*R*)-1-phenyl-1,2-ethanediol yielded a K_m = 1.2 ± 0.1 mM and k_{cat} = 27 ± 2 min^{-1} (Fig. 3*B*).

In Vitro Characterization of SgcC5 as an Amide-Forming Condensation Enzyme. Inspired by the canonical amide bond-forming activity of NRPS C domains and enzymes, SgcC5 was next assayed as an

BIOCHEMISTRY

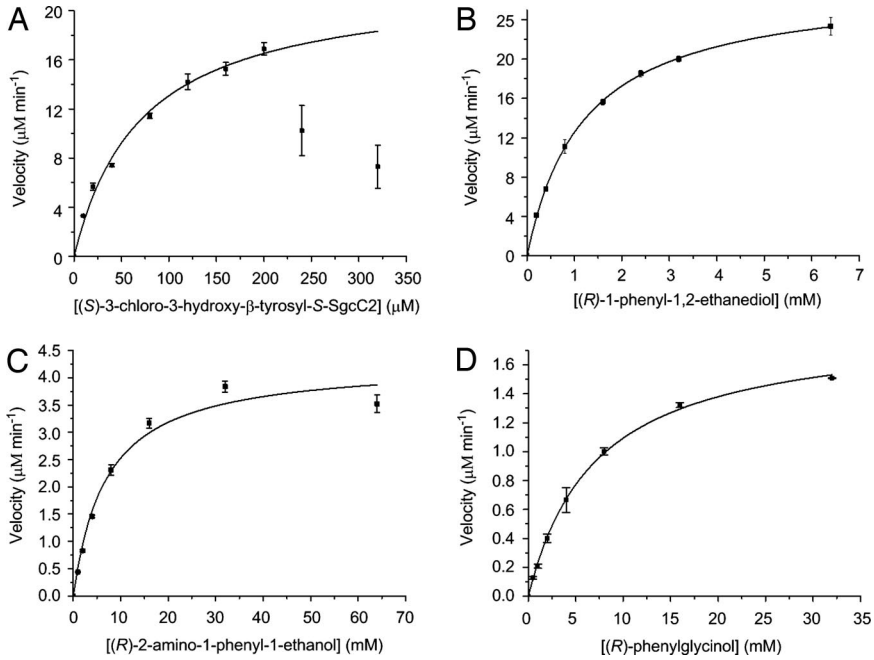

Fig. 3. Single-substrate kinetic analysis of SgcC5-catalyzed ester or amide formation. Reactions with variable donor substrate (*A*) (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and with variable acceptor substrates (*B*) (*R*)-1-phenyl-1,2-ethanediol (**1a**), (*C*) (*R*)-2-amino-1-phenyl-1-ethanol (**2a**), and (*D*) (*R*)-2-phenylglycinol (**3a**), respectively. See Fig. 2*A* for **1a**, **2a**, and **3a** structures.

amide-forming C enzyme with (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 as the donor substrate and (*R*)-2-amino-1-phenyl-1-ethanol (**2a**; Fig. 2*A*) as the acceptor substrate. Remarkably, HPLC analysis of the assay solution showed time-dependent formation of a new product with an rt of 16.3 min (Fig. 2*C*). This new peak was collected and subjected to HR-MALDI-MS analysis, yielding a pair of $[M + Na]^+$ ions at m/z = 373.0926 and 375.0900 and with a 3:1 ratio characteristic of the monochlorinated amide product (**2b**; Fig. 2*A*) (calculated $[M + Na]^+$ ions at m/z = 373.0931 and 375.0897 in a 3:1 ratio). Production and purification of **2b** from a scale-up reaction and subsequent ^{1}H, ^{13}C, and 2D NMR analyses unambiguously established its amide structure (Table S2). The apparent kinetic parameters were determined with 200 μM (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and variable (*R*)-2-amino-1-phenyl-1-ethanol to give K_m = 7.0 ± 1.3 mM and k_{cat} = 0.043 ± 0.002 min^{-1} (Fig. 3*C*). Under the conditions examined, SgcC5 was ≈3,700-fold more efficient at catalyzing ester bond formation with **1a** than catalyzing amide bond formation with **2a** as the acceptor substrate on the basis of k_{cat}/K_m value comparison (Fig. 2*A*).

Regiospecificity of SgcC5 Toward Acceptor Substrates. The regiospecificity of SgcC5 was first examined by following the time course of **1b** and **1c** formation from (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and (*R*)-1-phenyl-1,2-ethanediol (**1a**) as substrates (Fig. 2*A*). The formation of **1b** is SgcC5-dependent; **1b** is produced exclusively in the early stage, but **1c** appears upon longer incubation (Fig. 2*B*), indicating that **1c** originates from **1b**. Both products were subsequently purified, and incubation of each of them under the assay conditions in the absence of SgcC5 indeed confirmed the spontaneous conversion between **1b** and **1c**, with an equilibrium constant of **1c**/**1b** = 0.80 ± 0.10 (Fig. 2*A* and *SI Methods*). These observations support that the SgcC5-catalyzed ester bond formation with (*R*)-1-phenyl-1,2-ethanediol as the acceptor substrate is regiospecific for the C-2 hydroxyl group.

(*R*)-2-amino-1-phenyl-1-ethanol and (*R*)-2-phenylglycinol (**2a** and **3a**; Fig. 2*A*) were next used as acceptor substrates to further confirm the regiospecificity of SgcC5 with (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 as the donor substrate by following their reaction time courses, respectively. (*R*)-2-amino-1-phenyl-1-ethanol (**2a**) afforded the corresponding **2b** exclusively, as described earlier (Fig. 2*C*). (*R*)-2-phenylglycinol (**3a**) yielded 2 products initially, but the first product at rt of 15.5 min completely converted to the second product at rt 16.6 min after prolonged incubation (Fig. 2*D*). The second product at rt of 16.6 min was collected and analyzed by HR-ESI-MS, giving a pair of $[M + H]^+$ ions at m/z = 351.1106 and 353.1089 in the 3:1 ratio expected for the monochlorinated coupling product (**3c**; Fig. 2*A*) (calculated $M + H]^+$ ions at m/z = 351.1112:353.1082 ≈ 3:1). Scale-up reactions afforded purified **3c**, whose amide structure was confirmed by ^{1}H NMR analysis (*SI Methods*). The first product at rt 15.5 min, predicted to be the ester product (**3b**; Fig. 2*A*), was unstable, and thus no spectroscopic analysis was performed. Steady-state kinetic analysis using 200 μM (*S*)-3-choro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 and variable (*R*)-2-phenylglycinol was also carried out, yielding a K_m = 7.1 ± 0.3 mM and k_{cat} = 1.87 ± 0.03 min^{-1} (Fig. 3*D*). SgcC5 is ≈100-fold less efficient using **3a** compared with **1a** as the acceptor substrate on the basis of their k_{cat}/K_m value comparison. Taken together, these observations support that SgcC5 catalyzes either ester bond formation with (*R*)-1-phenyl-1,2-ethanediol (**1a**) or (*R*)-2-phenylglycinol (**3a**) or amide bond formation with (*R*)-2-amino-1-phenyl-1-ethanol (**2a**) as the acceptor substrate and is regiospecific for the C-2 hydroxyl or amino group, respectively. Of the 3 substrates examined, the preferred substrate for SgcC5 is a 1,2-diol enediyne mimic **1a**; the nascent product **1b**, however, readily undergoes a 1,2-migration to afford **1c** with an equilibrium constant of **1c**/**1b** = 0.80 ± 0.10 (Fig. 2*A*).

SgcC5 Is an Ester Bond-Forming and Amide Bond-Forming Condensation Enzyme. We have experimentally demonstrated that SgcC5 is indeed a C enzyme-catalyzing ester bond formation to incorporate the SgcC2-tethered (*S*)-3-chloro-5-hydroxy-β-tyrosine moiety into C-1027, concluding functional assignment of the minimal NRPS module, composed of 3 freestanding SgcC1, SgcC2, and SgcC5

proteins, in C-1027 biosynthesis (Fig. 1). This was accomplished by taking advantage of the chemoenzymatically prepared (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 as the donor substrate and the enediyne core mimic (*R*)-1-phenyl-1,2-ethanediol as the acceptor substrate for SgcC5. SgcC5 absolutely requires an SgcC2-tethered donor substrate and efficiently and regiospecifically couples it, in an ester linkage, with the C-2 hydroxy of the free acceptor substrate that mimics the C-14 hydroxy of the enediyne core (Figs. 1*B* and 2*A*). The fact that SgcC5 recognizes the enediyne core mimics, such as **1a**, **2a**, and **3a** (Fig. 2) is indicative of relaxed specificity for the acceptor substrate, and this is in contrast to previous investigations that suggest C domains in general have strict specificity for acceptor substrates (3, 5, 29).

Kinetic analysis was performed with (*R*)-1-phenyl-1,2-ethanediol (**1a**) as a surrogate acceptor substrate, yielding a K_m = 1.2 mM and k_{cat} = 27 min^{-1}. Despite using a nonnative substrate, the K_m is comparable to VibH with respect to norspermidine as an acceptor substrate, but the turnover of SgcC5 is substantially lower when compared with VibH (k_{cat} of 6,010 min^{-1}) (22) or when compared with the ester bond-forming CAT with chloramphenicol as an acceptor (k_{cat} of 3,600 min^{-1}) (36). In contrast, dihydrolipoamide acetyltransferases, which also belong to the same superfamily as condensation enzymes and CAT, have similar K_m with respect to acceptor substrates but significantly lower k_{cat} compared with SgcC5, as exemplified with yeast dihydrolipoamide acetyltransferase (K_m of 4.0 mM and k_{cat} = 0.13 min^{-1}) (37). The significance of these kinetic constants—particularly in the absence of kinetic data with natural substrate—is currently uncertain. It is apparent, however, that SgcC5-catalyzed amide bond formation occurs at a substantially lower efficiency, manifested by a large decrease in k_{cat}. Thus, amide bond formation clearly has no physiological relevance, but instead may present the opportunity to apply combinatorial biosynthetic approaches for structural diversification.

NRPS C domains, including SgcC5, consist of ≈450 amino acids and contain an $HHXXXDX_{14}Y$ motif, with the Asp and second His essential for activity (excluding only VibH, where catalysis was unaffected by mutation to the His residue) (3) (*SI Methods* and Fig. S1). Recent structural investigations with VibH (22), SrfA-C C domain (34), and TycC C domain (33) have revealed that C domains are composed of 2 subdomains, each of which has a similar fold to a monomer of CATs. CATs also contain an essential $HHXXXDX_{14}Y$ motif and are catalytically active as trimers, forming 3 active sites at the monomer interface (30). The structural similarities of C domains and CATs suggest that a similar strategy is used in amide bond and ester bond formation, respectively, and alludes to the possibility that C domains may be able to catalyze ester bond formation.

Here, we have shown that SgcC5 not only catalyzes ester bond formation, as was predicted from the structure of C-1027 (15), but also retains the canonical amide-forming activity typical of NRPS condensation enzymes, albeit with significantly reduced efficiency. SgcC5 now joins Fum14 (21) in the category of C domains that have been biochemically demonstrated to catalyze ester bond formation, but to our knowledge, SgcC5 is the first C domain to be shown to catalyze both ester and amide bond formation. SgcC5, therefore, may represent a rare event of evolution caught in action, providing an opportunity to investigate how to evolve an amide-forming C enzyme to an ester-forming enzyme or vice versa. Interestingly, the biosynthetic gene cluster for the enediyne maduropeptin contains an SgcC5 homolog, MdpC5, that has high sequence homology (45/58% identity/similarity), and hence likely similar function; however, in contrast to SgcC5, MdpC5 is predicted to catalyze amide bond formation based on examination of the maduropeptin structure (38). The relative simplicity of SgcC5 with respect to architecture and substrate specificity of typical C domains makes SgcC5 an excellent candidate to further explore the mechanistic details of C domains in general, and structural analysis of SgcC5 should provide critical insights into the selectivity for a hydroxy or amine nucleophile of the acceptor substrate. In view of the central role of C domains in nonribosomal peptide biosynthesis, and given the technical difficulties associated with studying the prototypical C domains, such studies should pave the way for new opportunities to generate novel natural products by combinatorial biosynthesis or chemoenzymatic methods.

Materials and Methods

General. Details for chemicals and instruments used; bioinformatics analysis of SgcC5; preparation of SgcC, SgcC1, SgcC2, SgcC3, SgcC5, SgcE6, and Svp; enzymatic synthesis of the donor substrate (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2; purification of the acceptor substrate (*R*)-2-amino-1-phenyl-1-ethanol (**2a**); time courses of SgcC5-catalyzed ester and amide bond formation and HPLC analyses; and the large-scale preparation of the ester and amide products **1b**, **2b**, **1c**, and **3c** and their structural characterization of these products by MS and NMR analyses are provided in the *SI Methods*.

Activity Assay for SgcC5. Standard assays were carried out in 200 μL of reaction mixture containing 200 μM (*S*)-3-chloro-5-hydroxy-β-tyrosyl-(*S*)-SgcC2 as the donor substrate, 10 mM acceptor substrates [(*R*)-1-phenyl-1, 2-ethanediol (**1a**) or (*R*)-2-amino-1-phenyl-1-ethanol (**2a**) or (*R*)-2-phenylglycinol (**3a**)], 1 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 50 mM phosphate buffer (pH 7.5), and SgcC5 (1.0 μM for **1a**, 100 μM for **2a**, and 20 μM for **3a**). HPLC was used to monitor the reaction, with detection at 254 nm (for acceptor substrates) and 283 nm (for products). The reactions were quenched by addition of trifluoroacetic acid to a final concentration of 16%, and after centrifugation the clarified supernatant was subjected to HPLC analysis (*SI Methods*).

Kinetics of Activity of SgcC5. To determine the kinetics for the donor substrate (*S*)-3-chloro-5-hydroxy-β-tyrosyl-SgcC2 in the production of **1b** and **1c**, 195 μL of reaction mixtures contained 100 mM Tris·HCl (pH 7.5), 5 mM ATP, 2 mM TCEP, 12.5 mM $MgCl_2$, 5 mM 3-chloro-5-hydroxy-β-tyrosine, 10 μM Svp, 10 μM SgcC1, and varying concentrations of apo-SgcC2 (10–320 μM) and CoA (50 μM to 1.6 mM; final concentration for 200 μL of total assay solution upon addition of SgcC5). The reactions were incubated at 25 °C for 45 min to allow for the loading of 3-chloro-5-hydroxy-β-tyrosine to holo-SgcC2 in situ before the addition of 5 mM **1a**. This loading was determined to be complete based on HPLC analysis, and the reaction components were shown not to inhibit SgcC5 activity. The reaction was initiated by the addition of 1 μM SgcC5 and carried out in triplicate. The reactions were quenched with 16% trifluoroacetic acid after being incubated at 25 °C for 5 min. After centrifugation, the resulting clarified supernatant was subjected to HPLC analysis. To quantify the amount of products formed from each reaction, a calibration curve based on the HPLC peak area with UV detection at 283 nm was generated with a known amount of synthetic 3-chloro-5-hydroxy-β-tyrosine, and all of these compounds exhibited identical UV absorption at this wavelength. The Michaelis–Menten equation was fit to plots of initial rate of product formation versus substrate concentration to extract values for the K_m and k_{cat} parameters.

To determine the kinetic parameters for acceptor substrates (*R*)-1-phenyl-1,2-ethanediol (**1a**), (*R*)-2-amino-1-phenyl-1-ethanol (**2a**), and (*R*)-2-phenylglycinol (**3a**), 180 μL of reactions contained 100 mM Tris·HCl (pH 7.5), 200 μM apo-SgcC2, 5 mM ATP, 1 mM CoA, 2 mM TCEP, 12.5 mM $MgCl_2$, 5 mM 3-chloro-5-hydroxy-β-tyrosine, 10 μM Svp, and 10 μM SgcC1 (final concentration for 200 μL of total assay solution upon addition of SgcC5). The reactions were similarly incubated at 25 °C for 45 min to allow for complete loading of 3-chloro-5-hydroxy-β-tyrosine to holo-SgcC2 in situ, and the reaction components were shown not to inhibit SgcC5 activity. The reactions were initiated by the addition of varied concentrations of acceptor substrates (0.2–6.4 mM for **1a**, 1–64 mM for **2a**, and 0.5–32 mM for **3a**), followed by the addition of SgcC5 (1 μM SgcC5 for **1a** and **3a**, and 100 μM SgcC5 for **2a**), and each reaction was performed under initial velocity conditions and carried out in triplicate. The reactions were quenched with 16% trifluoroacetic acid after being incubated at 25 °C for different times (5 min for **1a**, 15 min for **2a**, and 30 min for **3a**). Product formation was quantified by HPLC peak area to obtain velocity data for single-substrate kinetic analysis. Data were fit to the Michaelis–Menton equation to determine the K_m and k_{cat}.

ACKNOWLEDGMENTS. We thank Dr. Y. Li, Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences (Beijing, China), for the wild-type *S. globisporus* strain and the Analytical Instrumentation Center of the School of Pharmacy, University of Wisconsin, Madison for support in obtaining MS and NMR data. This work is supported in part by National Institutes of Health (NIH) Grants CA78747 and CA113297. S.V.L. is the recipient of an NIH postdoctoral fellowship (CA1059845).

1. Felnagle EA, et al. (2008) Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Mol Pharm* 5:191–211.
2. Walsh CT (2008) The chemical versatility of natural-product assembly lines. *Acc Chem Res* 41:4–10.
3. Finking R, Marahiel MA (2004) Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol* 58:453–488.
4. Fischbach MA, Walsh CT (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem Rev* 106:3468–3496.
5. Belshaw PJ, Walsh CT, Stachelhaus T (1999) Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* 284:486–489.
6. Wenzel SC, Müller R (2005) Formation of novel secondary metabolites by bacterial multimodular assembly lines: Deviations from textbook biosynthetic logic. *Curr Opin Chem Biol* 9:447–458.
7. Haynes SW, Challis GL (2007) Non-linear enzymatic logic in natural product modular mega-synthases and -synthetases. *Curr Opin Drug Disc Develop* 10:203–218.
8. Donadio S, Monciardini P, Sosio M (2007) Polyketide synthases and nonribosomal peptide synthetases: The emerging view from bacterial genomics. *Nat Prod Rep* 24:1073–1109.
9. Fujimori DG, et al. (2007) Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proc Natl Acad Sci USA* 104:16498–16503.
10. Du L, Sánchez C, Chen M, Edwards DJ, Shen B (2000) The biosynthetic gene cluster for antitumor drug bleomycin from Streptomyces verticillus ATCC15003 supporting functional interactions between nonribosomal peptides synthetase and a polyketide synthase. *Chem Biol* 7:623–642.
11. Felnagle EA, Rondon MR, Berti AD, Crosby HA, Thomas MG (2007) Identification of the biosynthetic gene cluster and an additional resistance gene for the antituberculosis drug capreomycin. *Appl Environ Microbiol* 73:4162–4170.
12. Li L, et al. (2008) Characterization of the saframycin A gene cluster from *Streptomyces lavendulae* NRRL 11002 revealing a nonribosomal peptide synthetase system for assembling the unusual tetrapeptidyl skeleton in an iterative manner. *J Bacteriol* 190:251–263.
13. Ehmann DE, Shaw-Reid CA, Losey HC, Walsh CT (2000) The EntF and EntE adenylation domains of *Escherichia coli* enterobactin synthetase: Sequestration and selectivity in acyl-AMP transfers to thiolation domain cosubstrates. *Proc Natl Acad Sci USA* 97:2509–2514.
14. Shao RG, Zhen JM (2008) Enediyne anticancer antibiotic lidamycin: Chemistry, biology and pharmacy. *Anticancer Agents Med Chem* 8:292–304.
15. Liu W, Christenson SD, Standage S, Shen B (2002) Biosynthesis of the enediyne antitumor antibiotic C-1027. *Science* 297:1170–1173.
16. Schwecke T, et al. (1995) The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc Natl Acad Sci USA* 92:7839–7843.
17. Motamedi H, Shafiee A (1998) The biosyntheic gene cluster for macrolactone ring of the immunsuppressant FK506. *Eur J Chem* 256:528–534.
18. Cheng YQ (2006) Deciphering the biosynthetic codes for the potent Anti-SARS-CoV cyclodepsipeptide valinomycin in *Streptomyces tsusimaensis* ATCC 15141. *Chembiochem* 7:471–477.
19. Ehling-Schulz M, et al. (2005) Identification and partial characterization of the nonribosomal peptide synthetase gene responsible for cereulide production in emetic *Bacillus cereus*. *Appl Environ Microbiol* 71:105–113.
20. Magarvey NA, et al. (2006) Biosynthetic characterization and chemoenzymatic assembly of the cryptophycins. Potent anticancer agents from *Nostoc* cyanobionts. *ACS Chem Biol* 1:766–779.
21. Zaleta-Rivera K, et al. (2006) A bidomain nonribosomal peptide synthetase encoded by *FUM14* catalyzes the formation of tricarballylic esters in the biosynthesis of fumonisins. *Biochemistry* 45:2561–2569.
22. Keating TA, Marshall CG, Walsh CT, Keating AE (2002) The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Nat Struct Mol Biol* 9:522–526.
23. Christenson SD, Wu W, Spies MA, Shen B, Toney MD (2003) Kinetic analysis of the 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis. *Biochemistry* 42:12708–12718.
24. Christianson CV, Montavon TJ, VanLanen SG, Shen B, Bruner SD (2007) The structure of L-tyrosine 2,3-aminomutase from the C-1027 enediyne antitumor antibiotic biosynthetic pathway. *Biochemistry* 46:7205–7214.
25. Van Lanen SG, et al. (2005) Biosynthesis of the β-amino acid moiety of the enediyne antitumor antibiotic C-1027 featuring β-amino acyl-*S*-carrier protein intermediates. *J Am Chem Soc* 127:11594–11595.
26. Van Lanen SG, Lin S, Dorrestein PC, Kelleher NL, Shen B (2006) Substrate specificity of the adenylation enzyme SgcC1 involved in the biosynthesis of the enediyne antitumor antibiotic C-1027. *J Biol Chem* 281:29633–29640.
27. Lin S, VanLanen SG, Shen B (2007) Regiospecific chlorination of (*S*)-β-tyrosyl-*S*-carrier protein catalyzed by SgcC3 in the biosynthesis of the enediyne antitumor antibiotic C-1027. *J Am Chem Soc* 129:12432–12438.
28. Lin S, Van Lanen SG, Shen B (2008) Characterization of the 2-component, FAD-dependent monooxygenase SgcC that requires carrier protein-tethered substrates for the biosynthesis of the enediyne antitumor antibiotic C-1027. *J Am Chem Soc* 130:6616–6623.
29. Mootz HD, Schwarzer D, Marahiel MA (2000) Construction of hybrid peptide synthetases by module and domain fusions. *Proc Natl Acad Sci USA* 97:5848–5853.
30. Shaw WV, Leslie AGW (1991) Chloramphenicol acetyltransferase. *Annu Rev Biophys Biophys Chem* 20:363–386.
31. Stachelhaus T, Mootz HD, Bergendahl V, Marahiel MA (1998) Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J Biol Chem* 273:22773–22781.
32. Rausch C, Hoof I, Weber T, Wohlleben W, Huson D (2007) Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol Biol* 7:78.
33. Samel SA, Schoenafinger G, Knappe TA, Marahiel MA, Essen LO (2007) Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase. *Structure* 15:781–792.
34. Tanovic A, Samel SA, Essen LO, Marahiel MA (2008) Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* 321:659–663.
35. Sánchez C, Du L, Edwards DJ, Toney MD, Shen B (2001) Cloning and characterization of a phosphopantetheinyl transferase from *Streptomyces verticillus* ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin. *Chem Biol* 8:725–738.
36. Lewendon A, Murray IA, Shaw WV (1994) Replacement of catalytic histidine-195 of chloramphenicol acetyltransferase: Evidence for a general base role for glutamate. *Biochemistry* 33:1944–1950.
37. Niu X, Stoops JK, Reed LJ (1990) Overexpression and mutagenesis of the catalytic domain of dihydrolipoamide acetyltransferase from *Saccharomyces cerevisiae*. *Biochemistry* 29:8614–8619.
38. Van Lanen SG, Oh T, Liu W, Wendt-Pienkowski E, Shen B (2007) Characterization of the maduropeptin biosynthetic gene cluster from *Actinomadura madurae* ATCC 39144 supporting a unifying paradigm for enediyne biosynthesis. *J Am Chem Soc* 129:13082–13094.