JACS

COMMUNICATIONS

Biosynthesis of the β-Amino Acid Moiety of the Enediyne Antitumor Antibiotic C-1027 Featuring β-Amino Acyl-S-carrier Protein Intermediates

Steven G. Van Lanen,[†] Pieter C. Dorrestein,[§] Steve D. Christenson,[†] Wen Liu,[†] Jianhua Ju,[†] Neil L. Kelleher,[§] and Ben Shen*[,†,‡]
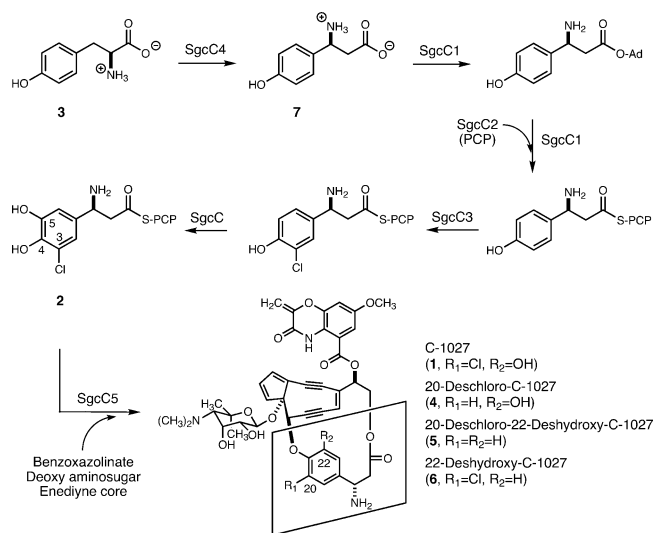
Division of Pharmaceutical Sciences and Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53705, and Department of Chemistry, University of Illinois, Urbana-Champaign, Urbana, Illinois 61801

Received May 2, 2005; E-mail: bshen@pharmacy.wisc.edu

C-1027, a chromoprotein natural product produced by *Streptomyces globisporus* and composed of an apoprotein and the C-1027 chromophore (**1**), belongs to the enediyne family of antitumor antibiotics that have been considered as some of the most potent, highly active anticancer agents ever discovered. We have previously cloned and characterized the C-1027 biosynthetic gene cluster from *S. globisporus* and proposed that the biosynthesis of the (*S*)-3-chloro-4,5-dihydroxy-β-phenylalanine moiety (**2**) from L-tyrosine (**3**) and its incorporation into **1** are catalyzed by six enzymes: SgcC, SgcC1, SgcC2, SgcC3, SgcC4, and SgcC5.[1] We now report the delineation of this pathway by in vivo and in vitro studies of five of the six enzymes. Our findings unveil a novel strategy for β-amino acid modification, featuring β-amino acyl-S-carrier protein intermediates, and shed new light into engineering the C-1027 biosynthetic machinery for the production of novel analogues as exemplified by 20-deschloro-C-1027 (**4**) and 20-deschloro-22-deshydroxy-C-1027 (**5**) (Figure 1).

Functional assignments to SgcC (hydroxylase), SgcC1 (adenylation enzyme), SgcC2 [peptidyl carrier protein, (PCP)], SgcC3 (halogenase), and SgcC4 (aminomutase) for their roles in the biosynthesis of **2** from **3** and to SgcC5 (condensation enzyme) for its role in incorporating **2** into **1** were initially made on the basis of bioinformatics analysis.[1] We subsequently confirmed SgcC4 as a novel L-tyrosine aminomutase in vitro,[2] demonstrated SgcC1 to be essential for the biosynthesis of **1** in vivo,[1] and verified SgcC as the C-5 hydroxylase, inactivation of which resulted in the production of 22-deshydroxy-C-1027 (**6**).[1] While these studies (a) established **3** as the biosynthetic precursor for **2**, (b) indicated **2** to be most likely activated as an aminoacyl-S-carrier protein prior to its incorporation into **1**, and (c) revealed the substrate promiscuity of SgcC5 in incorporating analogues of **2** into **1**, they, however, fell short of establishing the precise timing of each biosynthetic step. In contrast, they presented a dilemma for the nature of substrate for SgcC1—SgcC1 exhibits the signature nonribosomal code of "DPAQLMLIAK", characteristic of activating an α-amino acid according to the current nonribosomal peptide synthetase paradigm,[3] yet SgcC4 is an L-tyrosine aminomutase catalyzing the first step for the biosynthesis of **2** from **3**, a fact that precludes any α-amino acid as possible intermediates beyond (*S*)-β-tyrosine (**7**).[2]

To delineate the biosynthetic pathway for **2** from **3** with the ultimate goal of producing novel C-1027 analogues by rational engineering of the C-1027 biosynthetic machinery, we set out to characterize the substrate specificity of SgcC1 and the timing of tethering the biosynthetic intermediate to the SgcC2 carrier protein by both in vivo and in vitro experiments. Inspired by the fact that inactivation of *sgcC* resulted in the production of **6**,[1] we reasoned

† Division of Pharmaceutical Sciences, University of Wisconsin-Madison.
‡ Department of Chemistry, University of Wisconsin-Madison.
§ University of Illinois, Urbana-Champaign.

Figure 1. Biosynthetic pathway for the (*S*)-3-chloro-4,5-dihydroxy-β-phenylalanine moiety (**2**) of C-1027 (**1**, boxed) from L-tyrosine (**3**) and structures of **1** and the engineered analogues 20-deschloro-C-1027 (**4**), 20-deschloro-22-deshydroxy-C-1027 (**5**), and 22-deshydroxy-C-1027 (**6**).

that inactivation of *sgcC4* could result in the production of a C-1027 analogue with the corresponding α-amino acid variant in place of the **2** moiety if **3** could be similarly processed in the subsequent steps catalyzed by SgcC1, SgcC2, SgcC, SgcC3, and SgcC5, respectively (Figure 1). However, the resultant Δ*sgcC4* mutant failed to produce any metabolite,[4] raising the question that SgcC1 in fact is a β-amino acid specific adenylation enzyme and only the intermediates tethered to the SgcC2 carrier protein can be recognized, hence, incorporated by SgcC5 into the C-1027 chromophore.

We next expressed *sgcC1* and *sgcC2* in *Escherichia coli* and purified the overproduced SgcC1 and SgcC2 protein to homogeneity to investigate their activity and substrate specificity in vitro. SgcC1 was first examined by amino acid-dependent exchange of the radiolabel from [32P]pyrophosphate (PPi) to ATP.[4] Efficient exchange was only observed with **7**, and no exchange was detected with all the α-amino acids tested, including **3**, D-tyrosine, L-phenylalanine, L-3-chlorotyrosine, and L-3,4-dihydroxyphenylalanine. These findings contradict the bioinformatics-based prediction that SgcC1 is specific for an α-amino acid and clearly suggest SgcC1 as a β-amino acid-activating adenylation enzyme that stereospecifically prefers **7** as a substrate.

Adenylation enzymes such as SgcC1 have two activities: (a) ATP-dependent activation of the amino acid substrate into an aminoacyl adenylate and (b) aminoacylation of the phosphopantetheinyl prosthetic group of the carrier protein to afford the aminoacyl-S-carrier protein intermediate.[3] To further verify the activity and substrate specificity of SgcC1 and the relevance of
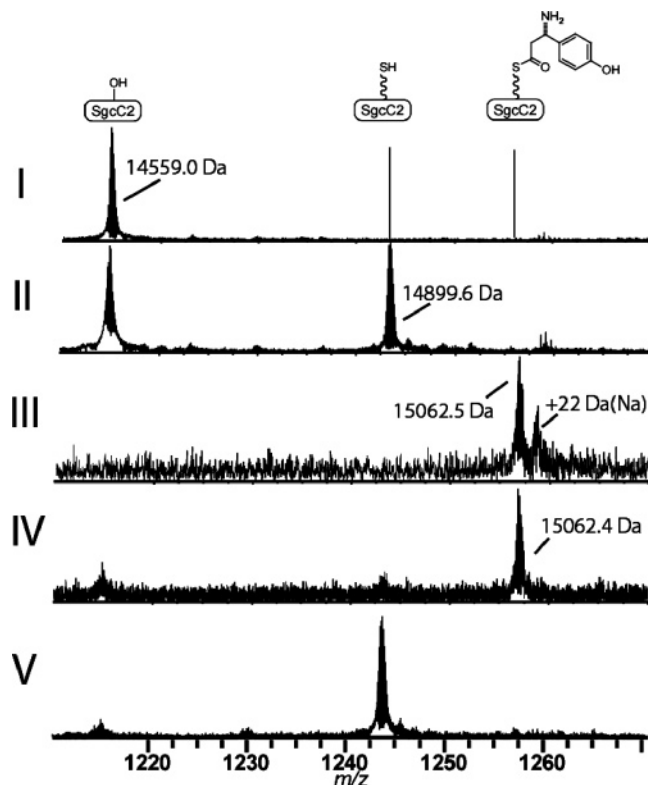
Figure 2. Detection of SgcC1 catalyzed activation of (S)-β-tyrosine (**3**) and (S)-β-tyrosinylation of SgcC2 by FT-MS analysis. I, apo-SgcC2; II, apo- and holo-SgcC2; III, holo-SgcC2 upon incubation of SgcC1 with **7** and ATP; IV, holo-SgcC2 upon incubation of SgcC1 with **3**, SgcC4, and ATP; V, holo-SgcC2 upon incubation with **3** and ATP as a negative control. The charge state for all ions is 12$^+$, and the numbers reported are the neutral monoisotopic masses.

these activities in **2** biosynthesis, SgcC1-catalyzed (S)-β-tyrosinylation of SgcC2 was examined directly by Fourier transform mass spectrometry (FT-MS).[4,5] SgcC2 overproduced in *E. coli* is predominantly in its apo-form as evidenced upon FT-MS analysis, showing a single peptide species with a observed mass of 14 559.0 Da (calcd 14 559.1 Da) (Figure 2, panel I). Incubation of apo-SgcC2 with either Sfp or Svp (phosphopanteinyl transferases known to exhibit broad carrier protein specificity)[6] in the presence of CoA resulted in a mass shift of the peptide to 14 899.6 Da. This is in agreement with phosphopantetheinylation of the active site Ser observed during the conversion of SgcC2 from apo- to its holo-form (calcd 14 899.2 Da) (Figure 2, panel II). Incubation of holo-SgcC2 with either **7** directly or a mixture of **3** and SgcC4, which generates **7** in situ,[2] in the presence of ATP afforded a new peptide species with a mass of 15 062.5 Da, in complete agreement with the predicted (S)-β-tyrosinyl-S-SgcC2 intermediate (calcd 15 062.3 Da) (Figure 2, panels III and IV). In contrast, no new peptide species was detected with **3** as a substrate under the identical condition, serving as a negative control (Figure 2, panel V). These results, together with the amino acid-dependent ATP-[^{32}P]PPi exchange data, support the proposal that SgcC1 is a **7**-specific adenylation enzyme and both SgcC and SgcC3 utilize β-aminoacyl-S-SgcC as substrates. This places the steps for **7** activation and (S)-β-tyrosinylation of SgcC2 before that of **7** chlorination and hydroxylation (Figure 1). They should also serve as precaution in predicting substrate specificity for adenylation enzymes according to the nonribosomal codes alone.[3]

Covalent tethering of an amino acid onto a carrier protein has been speculated previously as a general strategy for diverting a fraction of a proteinogenic amino acid into secondary metabolism.[7] Since **7** is a nonproteinogenic acid, competition for **7** as a substrate by other metabolic pathways might not be a major concern. However, tethering of **7** to SgcC2 could be advantageous in optimizing molecular recognition between downstream enzymes and the β-amino acid intermediates by providing structural features from the SgcC2 carrier protein. The downstream enzymes therefore present an outstanding opportunity for metabolic pathway engineering since these enzymes most likely will tolerate the substrate variants as long as they are tethered to SgcC2, hence leading to novel C-1027 analogue production. This is consistent with the findings that inactivation of *sgcC* results in the production of **6** while that of *sgcC1* only abolishes C-1027 production. To further exploit SgcC2-tethered intermediates in engineering novel C-1027 analogues, we finally inactivated *sgcC3*.[4] The resultant Δ*sgcC3* mutant produced two new compounds, the identity of which as **4** and **5** were confirmed by MALDI-MS analyses: **4** showed an [M + H]$^+$ ion at m/z = 810.3, consistent with the molecular formula $C_{43}H_{43}N_3O_{13}$, and **5** showed an [M + H]$^+$ ion at m/z = 794.3, consistent with the molecular formula $C_{43}H_{43}N_3O_{12}$.[4] The fact that the Δ*sgcC* mutant produced exclusively **6**[1] while the Δ*sgcC3* mutant produced both **4** and **5** supports the proposal that SgcC3 acts prior to SgcC. Thus, hydroxylation of (S)-β-tyrosinyl-S-SgcC2 by SgcC followed by incorporation of the resultant hydroxylated intermediate into the C-1027 chromophore by SgcC5 afforded **4**, but direct incorporation of the (S)-β-tyrosinyl-S-SgcC2 by SgcC5 yielded **5** (Figure 1).

Supporting Information Available: Full experimental details and FTMS data and assignment for SgcC2 and its tethered intermediate. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Liu, W.; Christenson, S. D.; Standage, S.; Shen, B.; Nakae, K. *Science* **2002**, *297*, 1170−1173. (b) Takemoto, Y.; Nakae, K.; Kawatani, M.; Takahashi, Y.; Naganawa, H.; Imoto, M. *J. Antibiot.* **2001**, *54*, 1104−1107.

(2) (a) Christenson, S. D.; Liu, W.; Toney, M. D.; Shen, B. *J. Am. Chem. Soc.* **2003**, *125*, 6062−6063. (b) Christenson, S. D.; Wu, W.; Spies, M. A.; Shen, B.; Toney, M. D. *Biochemistry* **2003**, *42*, 112708−12718.

(3) (a) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493−505. (b) Challis, G. L.; Ravel, J.; Townsend, C. A. *Chem. Biol.* **2000**, *7*, 211−224.

(4) See Supporting Information for full experimental details.

(5) (a) McLoughlin, S. M.; Kelleher, N. L. *J. Am. Chem. Soc.* **2004**, *126*, 13265−13275. (b) Gatto, G. J.; McLoughlin, S. M.; Kelleher, N. L.; Walsh, C. T. *Biochemistry* **2005**, *44*, 2770−2780.

(6) Sanchez, C.; Du, L.; Edwards, D. J.; Toney, M. D.; Shen, B. *Chem. Biol.* **2001**, *8*, 725−738.

(7) Chen, H.; Hubbard, B. K.; O'Connor, S. E.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 103−112.

JA052871K