ChemPubSoc Europe

DOI: 10.1002/cbic.201500701

CHEMBIOCHEM Communications

The Role of a Nonribosomal Peptide Synthetase in L-Lysine Lactamization During Capuramycin Biosynthesis

Xiaodong Liu,[a] Yuanyuan Jin,[b] Zheng Cui,[a] Koichi Nonaka,[c] Satoshi Baba,[c] Masanori Funabashi,[d] Zhaoyong Yang,*[b] and Steven G. Van Lanen*[a]

Capuramycins are one of several known classes of natural products that contain an L-Lys-derived L-α-amino-ε-caprolactam (L-ACL) unit. The α-amino group of L-ACL in a capuramycin is linked to an unsaturated hexuronic acid component through an amide bond that was previously shown to originate by an ATP-independent enzymatic route. With the aid of a combined in vivo and in vitro approach, a predicted tridomain nonribosomal peptide synthetase CapU is functionally characterized here as the ATP-dependent amide-bond-forming catalyst responsible for the biosynthesis of the remaining amide bond present in L-ACL. The results are consistent with the adenylation domain of CapU as the essential catalytic component for L-Lys activation and thioesterification of the adjacent thiolation domain. However, in contrast to expectations, lactamization does not require any additional domains or proteins and is likely a nonenzymatic event. The results set the stage for examining whether a similar NRPS-mediated mechanism is employed in the biosynthesis of other L-ACL-containing natural products and, just as intriguingly, how spontaneous lactamization is avoided in the numerous NRPS-derived peptides that contain an unmodified L-Lys residue.

The capuramycin-type nucleoside antibiotics are natural products endowed with excellent antimycobacterial activity. Most members of the family, represented by A-500359 B (**1**) from *Streptomyces griseus* SANK 60196 and the 2′-O-carbamoylated derivative A-503083 B (**2**) from *Streptomyces* sp. SANK 62799 (Scheme 1 A),[1–5] contain an L-α-amino-ε-caprolactam (L-ACL) unit that is covalently linked to an uncommon unsaturated hexuronic acid component by an amide bond.

Several other natural products containing ACL components are also known; they include bengamide A (**3**), peritoxin A (**4**), circinatin (**5**), caprolactin A (**6**), and the siderophores the nocardimicins (e.g., nocardimicin A, **7**) and the mycobactins (**8**, Scheme 1 B).[6–10] Similarly to those in **1** and **2**, the ACL units in **4**–**8** (the L isomers except in the cases of **4** and **5**) are each covalently linked to the rest of the molecule by an amide bond. Often the ACL unit is further modified, as in the stereoselective C-methylation that is observed in the capuramycins A-500359 A (**9**) and A-503083 A (**10**) or the unusual amination of **4**.

The mechanisms for the formation, attachment, and modification of the ACL units in **1**–**8** remain, for the most part, unknown. Not surprisingly, isotopic enrichment studies with the **1**-producing strain were consistent with L-Lys as the direct biosynthetic precursor to L-ACL.[4] The subsequent identification of the biosynthetic gene cluster for **1** and **2** revealed an orthologous set of genes (*orf26* and *orf27* for **1**, *capU* and *capV* for **2**) encoding nonribosomal peptide synthetases (NRPSs) that were proposed to be involved in L-ACL maturation.[11,12] Orf26/CapU was bioinformatically predicted to consist of three domains found in NRPSs: a condensation (C), an adenylation (A), and a thiolation (T) domain (sequential from N to C terminus), whereas Orf27/CapV was predicted to be a stand-alone C domain.

The archetypical NRPS system orchestrates the assembly of peptides in a modular fashion, with the addition of each amino acid requiring a C, an A, and a T domain, the last of which is often termed a peptidyl carrier protein domain.[13–14] Prior to NRPS catalysis, a seryl residue on the apo-T domain is first post-translationally modified with a phosphopantetheine prosthetic group derived from coenzyme A (CoA) in a reaction catalyzed by a phosphopantetheinyltransferase (PPTase). In turn, the A domain selects an amino acid substrate, activates the carboxylic acid as the acyl adenylate at the expense of ATP, and attaches the amino acid to the sulfhydryl group of the holo-T domain to generate a thioesterified intermediate. Subsequently, the C domain catalyzes the coupling between adjacent thioester-linked substrates to form an amide bond. The resulting dipeptide, which remains attached to the acyl-accepting T domain, then serves as a donor for aminolysis that is catalyzed by a downstream C domain, if present. After complete elongation, which is dictated by the number of C-A-T modules in the NRPS, the resulting peptide is typically released by a ter-

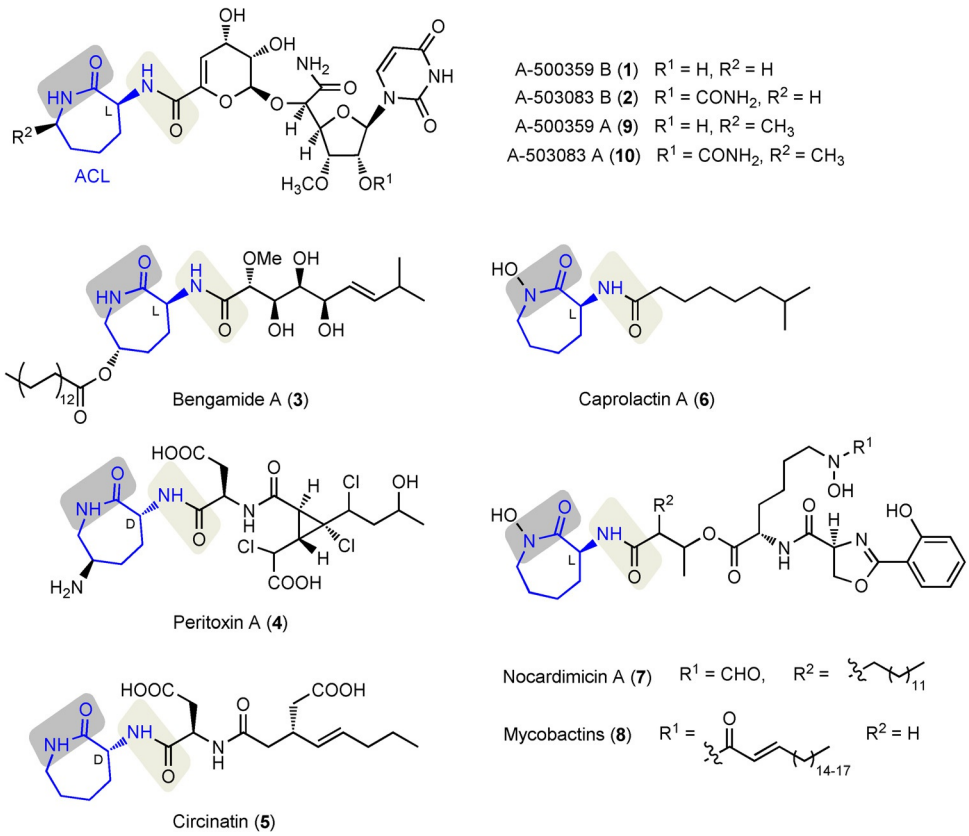
[a] Dr. X. Liu, Z. Cui, Dr. S. G. Van Lanen
Department of Pharmaceutical Sciences, College of Pharmacy
University of Kentucky
Lexington, KY 40536 (USA)
E-mail: svanlanen@uky.edu

[b] Y. Jin, Dr. Z. Yang
Key Laboratory of Biotechnology of Antibiotics
Institute of Medicinal Biotechnology
Chinese Academy of Medicinal Sciences & Peking Union Medical College
Beijing (China)

[c] Dr. K. Nonaka, S. Baba
Biologics Technology Research Laboratories, Daiichi Sankyo, Co. Ltd.
Gunma 370-0503 (Japan)

[d] Dr. M. Funabashi
Natural Product Research Group
Discovery Science and Technology Department
Daiichi Sankyo RD Novare Co. Ltd.
Tokyo 134-8630 (Japan)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under http://dx.doi.org/10.1002/cbic.201500701.



Scheme 1. Representative natural products containing α-amino-ε-caprolactam (ACL) units. A) Structures of the members of the capuramycin family of antimycobacterial antibiotics. B) Structures of other known natural products containing ACL components.

minal thioesterase domain either by hydrolysis or by intramolecular cyclization, although other mechanisms of chain release are known.[15]

Preliminary characterization of the NRPS CapU, by use of the traditional amino-acid-dependent ATP-^{32}PPi exchange assay to assess A domain substrate specificity,[16] was consistent with the A domain having a modest preference for L-Lys.[12] Surprisingly, however, it was discovered that the attachment of the L-ACL unit to the hexuronic acid component did not involve the NRPS but was instead mediated by a carboxymethyltransferase (CapS) and an ATP-independent transacylase (CapW).[12] Therefore, the role of the NRPS in the biosynthesis of **2** appeared to be solely the lactamization of L-Lys, yet it was unclear why two C domains were present and which one was important for lactam formation. Using a combination of bioinformatic analysis and in vivo and in vitro approaches, we now demonstrate that neither C domain—neither that of CapU, nor that of CapV—is necessary for lactamization.

To interrogate the role of the NRPS system in capuramycin biosynthesis, the development of a genetic system in the **1**-producing strain was initially explored.[11] However, the results suggested that this strain harbored multiple copies for several—if not all—of the required biosynthetic genes (Results and Figure S1 in the Supporting Information), a phenomenon that has also been observed for a few other natural product biosynthetic gene clusters.[17–19] Thus, we switched to the **2**-producing strain, which makes the two aforementioned ACL-containing **2** and **10** as major congeners along with minor amounts of the deaminocaprolactam precursors A-503083 F (**11**), characterized by a carboxylic acid, and A-503083 E (**12**), the methyl ester of **11** (Figure 1 A and B, trace i).[5] The *capU* gene was targeted for inactivation, and the expected double-crossover genotype was confirmed by PCR and Southern blot analysis (Figure S2). As expected, HPLC analysis of the Δ*capU* mutant strain revealed that the production of **2** and **10** was abolished, with concomitant increases in the peaks corresponding to **11** and **12** (Figure 1 B, trace ii). Upon feeding of L-ACL to the Δ*capU* mutant strain, the production of **2** was restored (Figure 1 B, trace iii), thus demonstrating an essential role for CapU in L-ACL biosynthesis. Moreover, these results are consistent with the previous characterization of CapW as an L-ACL:**12** transacylase.[12]

To provide further support for their role, *capU* and *capV* were cloned into pUW201 containing the *ermE** constitutive promoter and expressed in the heterologous host *S. lividans* TK64. In contrast to the control consisting of the empty vector, expression of *capU* and *capV* in combination with feeding with L-Lys resulted in the formation of L-ACL, which was detected by Fmoc derivatization and LC-MS analysis (Figure S3 and Table S1). The combined in vivo results suggest that, as initially posited, CapU and CapV orchestrate ACL biosynthesis.

As previously noted, prior results had demonstrated that recombinant CapU has a modest preference for L-Lys over other proteinogenic amino acids, **11**, or D-Lys.[12] This was consistent with a mechanism in which the A domain directly loads L-Lys

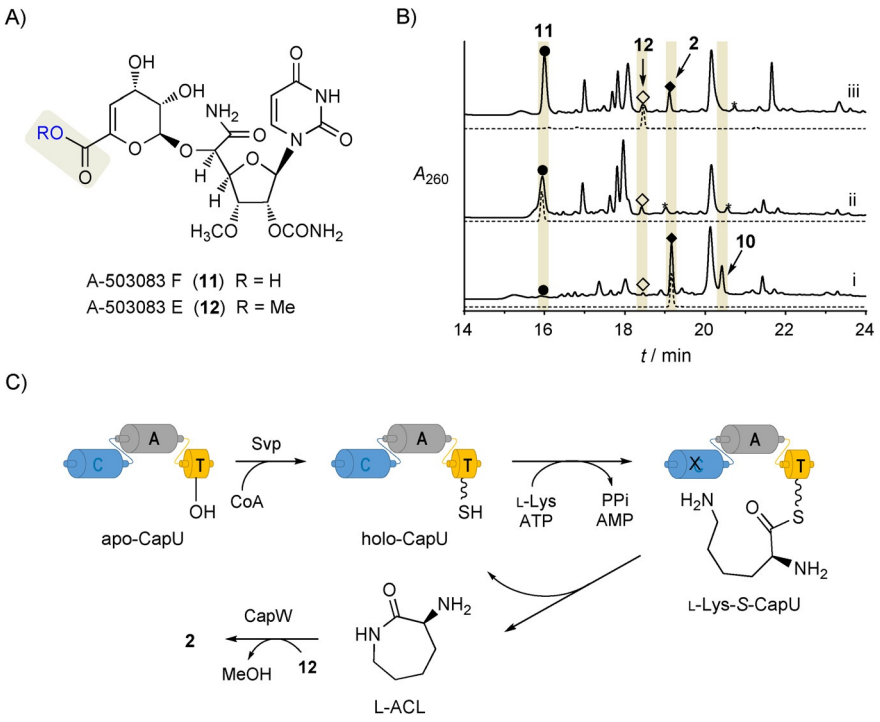


Figure 1. Role of the NRPS system in L-ACL biosynthesis. A) Structures of the deaminocaprolactam congeners. B) HPLC chromatograms of the methanol extracts from i) the wild-type strain, ii) the *ΔcapU* mutant strain, and iii) the Δ*capU* mutant strain cultured with exogenously supplied L-ACL. A_{260} is the absorbance at 260 nm; asterisks (*) each denote an unidentified metabolite with a UV/Vis spectrum different from those of capuramycins. C) In vitro reconstitution of L-ACL biosynthesis starting from L-Lys and apo-CapU in the presence of Svp with or without additional enzymes.

in *cis* to the holo-T domain, and one of the two available C domains catalyzes lactamization to release L-ACL (Figure 1C). Analysis of model C domains by others has revealed a conserved His residue to be essential not only for intermolecular aminolysis but also for an unnatural intramolecular macrolactamization catalyzed by an excised C domain of the tyrocidin NRPS.[20–22] Amino acid sequence alignments of CapU and CapV with C domains of previously solved structure revealed that this His residue is also found in CapV (H119), but that in CapU it is substituted with Gln (Q184, Figure S4). Therefore, CapV was initially predicted to be the amide-bond-forming catalyst and was therefore targeted for in vitro analysis along with CapU. Similarly to CapU, soluble CapV was obtained from *Escherichia coli* by using standard expression conditions (Figure S5).

To simplify the detection of the expected L-ACL product, activity was initially assessed by enzymatically coupling the reactions of CapU and/or CapV with that of CapW, which transfers L-ACL to **12** and had previously been determined to have no activity with L-Lys as an acyl acceptor (Figure 2A).[12] Unlike CapU and CapV, however, soluble, recombinant CapW was only obtainable from *S. lividans* TK24 (Figure S5). Svp, a well-characterized and promiscuous PPTase,[23] was also included for the in situ generation of holo-CapU from CoA and apo-CapU (Figure 1C). In comparison with controls (e.g., Figure 2A, traces i and ii), HPLC analysis of the reaction mixtures containing all four proteins starting with substrates L-Lys and **12** revealed a new peak that coeluted with authentic **2** generated by CapW or purified from the producing strain (Figure 2A, traces iii–v). LC-MS analysis of the product yielded an $[M+H]^+$ ion at *m/z* 612.6, which is consistent with the molecular formula of **2** (calcd 612.2; all conditions tested are summarized in Table S2). When CoA and Svp were omitted, **2** was still detected, albeit at a significantly reduced level. This suggested that some CapU is produced in the holo form upon heterologous expression in *E. coli*, a phenomenon that has also been reported for T domains of other NRPS systems.[24–26] The heterologous production of some holo-CapU might also explain why L-Lys was only moderately preferred in the amino-acid-dependent ATP-^{32}PPi exchange assay, because turnover to L-ACL (vide infra) would counteract the incorporation of the radiolabel into ATP. As expected, no product was detected when ATP or CapU was excluded, consistently with the hypothesis that L-Lys is first activated as an acyladenylate and transferred to the T domain of holo-CapU. In contrast to our expectations, however, CapV was not essential and nor did it enhance formation of **2** in the presence of L-Lys (Figure 2A, trace iv), thus suggesting that CapU was the sole catalyst responsible for L-Lys activation and lactamization.

We next aimed to monitor L-ACL formation directly by removing **12** from the reaction mixture and derivatizing any product and remaining L-Lys with dansyl chloride prior to analysis (Figure 2B). By comparison with the appropriate controls, HPLC analysis clearly revealed that L-ACL was formed under the condition that CapU and ATP were included in the reaction mixture (Figure 2B, trace i). Moreover, a truncated didomain version of CapU consisting of the A and T domains (CapU_AT) was prepared. Similarly to the native protein, CapU_AT cata-

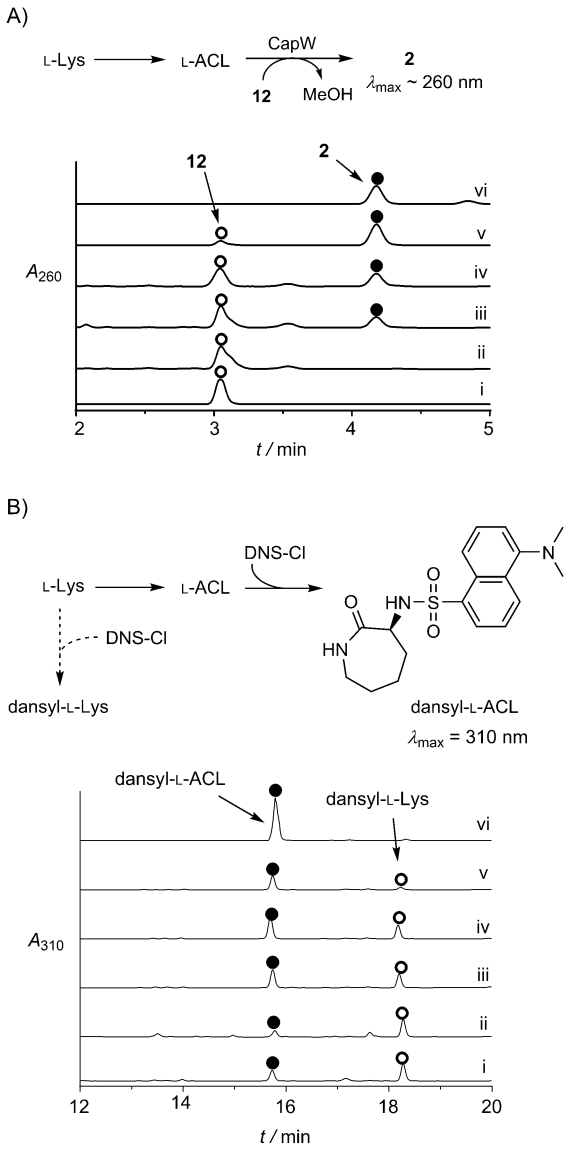


Figure 2. Detection of L-ACL formation. A) Expected amide-bond-forming reactions starting from L-Lys and enzymatic coupling with CapW. Representative HPLC traces for the four-enzyme reaction: i) authentic **12**, ii) reaction mixture without CapU, CapV, and Svp, thus confirming that CapW is unable to incorporate L-Lys directly, iii) reaction mixture containing all of the necessary enzymes (CapU, CapV, CapW, and Svp) and substrates/cofactors (L-Lys, ATP, **12**, CoA, and $MgCl_2$), iv) reaction mixture without CapV, v) positive control confirming CapW-catalyzed transacylation starting from L-ACL and **12**, and vi) authentic **2**. A_{260} is absorbance at 260 nm. B) Expected amide-bond-forming reactions starting from L-Lys and post-reaction derivatization with dansyl chloride. Representative HPLC traces of L-ACL formation by dansyl chloride modification: i) reaction mixture without CapV, thus confirming that CapV is not necessary for lactamization, ii) reaction mixture without CapV but with truncated CapU_AT (A + T domains), iii) reaction mixture with L-Lys-SNAC (**13**) as the substrate in place of L-Lys, iv) reactions starting from **13** without inclusion of any enzymes, demonstrating nonenzymatic lactamization and hydrolysis in buffered aqueous conditions; v) reactions with **14** as the substrate, which favors lactamization over hydrolysis; and vi) control of L-ACL modified with dansyl chloride. A_{310} is absorbance at 310 nm. Dansyl-L-Lys is L-Lys modified with two dansyl groups (based on MS analysis).

lyzed the formation of L-ACL from L-Lys (Figure 2B, trace ii), thus clearly demonstrating a functional A and T domain for this recombinant protein.

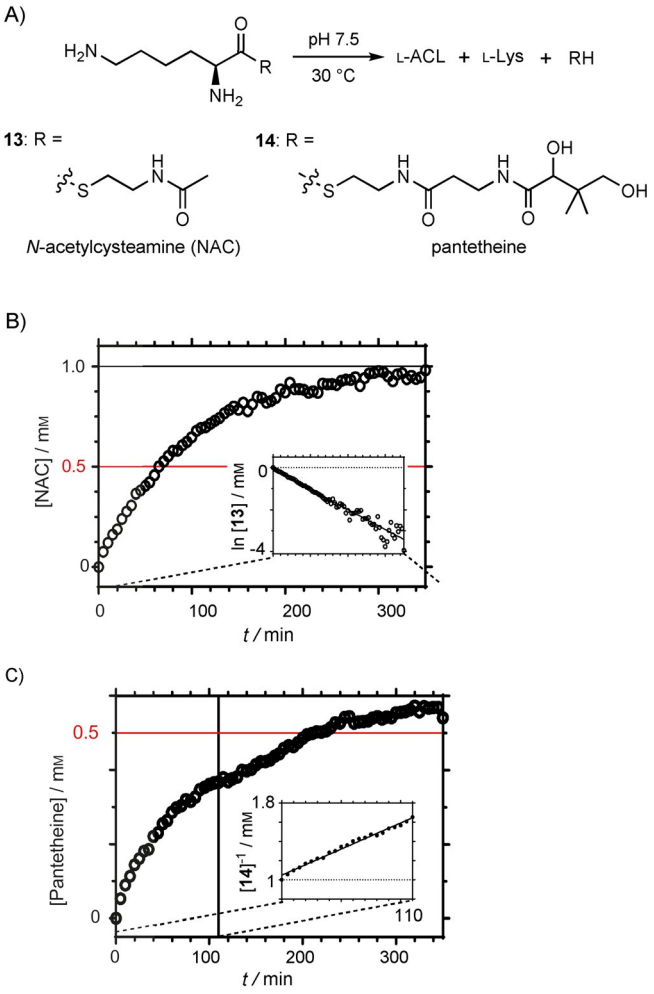


Figure 3. Nonenzymatic lactamization/hydrolysis. A) Structures of the surrogate substrates **13** and **14**. B) Time dependence of the nonenzymatic reaction of **13**; the inset is suggestive of first-order kinetics. C) Time dependence of the nonenzymatic reaction of **14**; the inset depicts the initial stages of the reaction ($t < 110$ min) that are suggestive of second-order kinetics.

In the hope of further simplifying the interpretation of the results, L-Lys-*S*-*N*-acetylcysteamine (**13**) and L-Lys-*S*-pantetheine thioester **14** (Figure 3A) were synthesized and used as potential surrogate substrates for L-Lys loaded onto the pantetheine group of the T domain (L-Lys-*S*-CapU in Figure 1C). However, when **13** and **14** were transferred to aqueous Tris or phosphate buffer (pH 7.5), L-ACL was almost immediately detected in the absence of enzyme. After completion of the reaction, L-ACL/L-Lys ratios of 62:38 and 90:10 were detected when starting from **13** and **14**, respectively. Monitoring of the formation of the free sulfhydryl system upon modification with 5,5′-dithiobis(2-nitrobenzoate), as previously described during the characterization of N^{ε}-OH-Lys:acetyl-CoA N^{ε}-transacetylase,[27] showed that the time dependence of substrate consumption for the parallel reactions displayed apparent first-order kinetics for **13** ($k = 1.0 \times 10^{-2}$ min^{-1}; Figure 3B). The time dependency for **14** appeared more complex under the reaction conditions, displaying potential mixed kinetics, of second order at higher concentrations of **14** ($k = 5.5 \times 10^{-3}$ mM^{-1} min^{-1}; Figure 3B) and first order at lower concentrations ($k = 1.7 \times 10^{-3}$ min^{-1}; Fig-

ure S6). These kinetic parameters are comparable to those of the reported enzymatic N-acetylation of N^{ε}-OH-Lys with acetyl-CoA ($k_{cat} = 2.9 \times 10^{-3}$ min^{-1} and $k_{cat}/K_m = 52 \times 10^{-3}$ mM^{-1} min^{-1}).[27] Although additional kinetic investigations will be required in order to allow any mechanistic insight to be inferred from these data, it is nevertheless clear that L-ACL is nonenzymatically formed from L-Lys thioesters at a significant rate.

Despite the nonenzymatic lactamization/hydrolysis, lastly we tested whether any enzyme or combination thereof would accelerate the intramolecular aminolysis of **14**. Whether with CapU and/or CapV, the rate of product formation was the same as that of control reactions without any enzyme (Figure S6C). Additionally, CapW had no effect on the rate except for a slight increase under the condition that **12** was included (Figure S6D). Given the lack of additional candidates within the biosynthetic gene cluster, the results suggest that lactamization is a nonenzymatic process during the biosynthesis of **2**. This phenomenon is not unlike the recent reports of nonenzymatic acylation of Lys residues in bacterial peptides by acyl-CoA thioesters[28] and acetyl phosphate.[29]

The revelation that neither C domain—neither that of CapU nor that of CapV—is essential for lactamization of L-Lys raises the question of the role of these domains, if any, in **1** and **2** biosynthesis. In this particular case, we propose that they indeed have no role and are evolutionary remnants of horizontal gene transfer. In support of this, close bioinformatic inspection of the **2** biosynthetic locus reveals that *capU*, *-V*, and *-W* form a subcluster flanked by genes encoding putative transposases followed by long intergenic regions (Figure S7). Perhaps more indicative, however, is the fact that the calculated codon adaptation indexes for *capV* (0.401) and *capU* (0.495) are substantially lower than the predicted 17 open reading frames involved in the biosynthesis of **2** (0.521–0.779) when referenced against the codon usage of *S. griseus* subsp. *griseus* (Table S3).[30] A similar trend is observed when using the housekeeping glucokinase gene from *S. griseus* subsp. *griseus* as a reference (Table S3). The observation of a low codon adaptation index is strongly indicative of genes of heterologous origin and that *capV*, and to a lesser extent *capU*, have undergone minimal selection for optimizing the codon usage in a *Streptomyces* host, as might be expected for a nonfunctional gene. The genetic and heterologous biotransformation systems that were developed here can now be employed to interrogate further the unexpected lack of requirement for CapV and the C domain of CapU for the biosynthesis of **2** in vitro.

Several NRPS-derived peptides contain Lys residues wherein the ε-amine is either unmodified, hydroxylated and acylated, or has undergone lactamization to form a macrocyclic peptide or ACL. A domains of the NRPSs involved in the biosynthesis of **8**, which contains two Lys-derived amino acids (Scheme 1), are the only enzymes for these ACL-containing compounds that have been biochemically interrogated in vitro.[31] The results of the amino-acid-dependent ATP-$^{32}PP_i$ exchange assay with the full-length mycobactin NRPS proteins MbtE (C-A-T-C-T domains) and MbtF (C-A-T-C domains), which are responsible for incorporating the two L-Lys units into the final product, revealed that in both cases the ε-amine group is modified prior to A domain activation. MbtE is most active with N^{ε}-acyl-N^{ε}-OH-L-Lys; this suggests that it is responsible for incorporating the internal L-Lys unit and that the modifications prior to thioesterification potentially protect against lactamization. On the other hand, MbtF is most active with N^{ε}-OH-L-Lys, and this suggests that this NRPS component is responsible for incorporating the terminal L-Lys unit that is converted into L-ACL. Similarly to the capuramycin NRPS system, there is an extra C domain at the terminus of MbtF, and it was proposed that this catalyzes lactamization. Whether lactamization with concomitant peptide release is indeed a catalyzed process or nonenzymatic as revealed here remains unknown. Juxtaposed with L-Lys lactamization is the process for introducing an unmodified L-Lys residue into an NRPS-derived peptide. Currently, this mechanism has not been elucidated, and it will be of general interest to decipher how Lys is incorporated and whether this process involves a high degree of kinetic control during peptide assembly or whether, possibly, a natural protection/deprotection strategy is used, analogous to that seen in the biosynthesis of the 4-amino-2-hydroxybutyrate component found in the aminoglycoside butirosin.[32]

Experimental Section

Gene inactivation of *capU*: The *capU* gene was inactivated by using REDIRECT technology.[33] Briefly, cosmid pN-1[12] containing *capU* was introduced into *E. coli* BW25141/pKD78 to generate *E. coli* BW-N-1. A linear PCR fragment, obtained with template pIJ773 and the primer pair (forward) 5′-CCTCG CGGCG CACCT CCGCG GCAAC GTCTC TGAGG TGCCA TTCCG GGGAT CCGTC GACC-3′/(reverse) 5′-GCGGT ACGCC GAACT TCTCC ACCAC CCCTC CGACG ATTTT GTAGG CTGGA GCTGC TTC-3′ (underlined sequence corresponds to target gene), was introduced into *E. coli* BW-N-1 by electroporation. The double-crossover was selected by apramycin resistance and confirmed by PCR. The modified cosmid was transformed into *E. coli* ET12567(pUZ8002) and introduced into *Streptomyces* sp. SANK 62799 by conjugation. The double-crossover event was selected by apramycin resistance on ISP4 agar. The genotype was subsequently confirmed by PCR and Southern blot analysis. For Southern blot analysis, genomic DNA either of the wild type or of the Δ*capU* mutant strain was isolated, digested with *Xho*I, and subjected to agarose gel electrophoresis for detection with a DIG-labeled probe generated by PCR with use of primers complementary to regions flanking the *capU* gene.

Biotransformation with CapU and CapV: The *capU* and *capV* genes were amplified by PCR, each as a single fragment, with the Expand Long Template PCR System from KOD PLUS Neo (Toyobo, Osaka, Japan) with the supplied buffer, dNTPs (200 μM), DMSO (5%), DNA template pN-4 (10 ng),[12] DNA polymerase (5 U), and the primer pairs (forward) 5′-GGGAA TTCCA TATGG TGGAC GCCCC GCGTC ACTGG A-3′ (NdeI site underlined)/(reverse) 5′-CCCAA GCTTT CACCA ACGGA ATGTC CCGGC GATT-3′ (HindIII site underlined, 400 nM each). The PCR program included an initial hold at 94 °C for 4 min, followed by 30 cycles of 96 °C for 30 s, 63 °C for 30 s, and 68 °C for 5 min, and then hold at 68 °C for 7 min. The gel-purified PCR product was ligated to the identical sites of pUWL201pw to yield pUWL201pw-*capUV*.

After sequencing to confirm PCR fidelity, pUWL201pw-*capUV* was transformed into *S. lividans* TK64 by PEG-mediated protoplast transformation.[34] After 20 h at 28 °C, plates were overlaid with soft nutrient agar (1 mL) supplemented with thiostrepton (200 μg). After three additional days at 28 °C, single colonies were transferred to fresh R2YE plates supplemented with thiostrepton (50 μg mL^{-1}). After 4 d at 28 °C, positive transformants were confirmed by colony PCR with InstaGene Matrix from Bio-Rad (Hercules, CA) and LA-Taq polymerase with GC buffer II from Takara Bio (Shiga, Japan). Positive transformants were used to inoculate R2YE (50 mL) containing thiostrepton (25 μg mL^{-1}) and grown for 2 d at 28 °C at 250 rpm, at which point 2 mL was transferred to fresh R2YE (100 mL) containing thiostrepton (25 μg mL^{-1}). After growth for 3 d at 28 °C and at 250 rpm, L-Lys was added (25 mg mL^{-1} final concentration), and the culture was grown for an additional 72 h. The culture was separated by centrifugation, and the pH of the supernatant was adjusted to 7 with NaOH (1 M). The mixture was applied to a column containing XAD16 resin and sequentially washed with 5 volumes each of water, acetone (20%), acetone (50%), and acetone (80%). Fractions eluting with acetone (20%) were collected and dried under vacuum. The dried solid containing L-ACL from a 100 mL culture was dissolved in sodium bicarbonate solution (500 μL, 20 mM), and an equal volume of Fmoc-OSu acetonitrile solution (2 mM) was added. After the mixture had been stirred for 5 min, samples were analyzed with an Agilent 1200 Series Quaternary LC system equipped with an Eclipse XDB-C18 column (150 mm × 4.6 mm, 5 μm, 80 Å). A series of linear gradients was developed from TFA (0.1%) in water (A) to TFA (0.1%) in acetonitrile (90%, B) in the following manner (beginning time and ending time with linear increase to % B): 0–20 min, 50–90% B; 20–25 min, 90% B. The flow rate was kept constant at 1.0 mL min^{-1}, and elution was monitored at 254 nm.

Cloning and expression of *capU*, *capV*, and *capU_AT*: The genes for CapU, CapV, and CapU_AT (truncated CapU consisting of the A and T domains) were amplified by PCR with use of an Expand Long Template PCR System from Roche with supplied buffer 2, dNTPs (200 μM), DMSO (5%), pN-4 (10 ng),[12] DNA polymerase (5 U), and the primer pairs (400 nM each) *capU* (forward) 5′-GGTAT TGAGG GTCGC ATGGA CGCCC CGCGT CACTG-3′/(reverse) 5′-AGAGG AGAGT TAGAG CCTCA GGGCG AGGAG TCGAC ATAG-3′, *capV* (forward) 5′-GGTAT TGAGG GTCGC ATGCC CGGAC CGCAG AATG-3′/(reverse) 5′-AGAGG AGAGT TAGAG CCTCA CCAAC GGAAT GTCCC G-3′, and *capU*_AT (forward) 5′-AAAAA ACATA TGGGT GATGT CATCG GCCCC-3′/(reverse) 5′-AAAAA AGGAT CCTCA GGGCG AGGAG TCGAC A-3′. The PCR program included an initial hold at 94 °C for 2 min, followed by 30 cycles of 94 °C for 10 s, 56 °C for 15 s, and 68 °C for 60 s. The gel-purified PCR product of *capU* and *capV* was inserted into pET-30 Xa/LIC by ligation-independent cloning as described by Novagen (Madison, WI, USA) to yield pET30-*capU* and pET30-*capV*, which were subsequently sequenced to confirm PCR fidelity. The gel-purified PCR product of *capU_AT* was digested with NdeI and BamHI and inserted into similarly digested Pdb·His·MBP by using T4 DNA ligase (New England Biolabs) to yield Pdb·His·MBP-*capU_AT*, which was sequenced to confirm PCR fidelity. The *capV* gene, engineered to be expressed as an N-terminal His$_6$-protein, was subcloned by PCR with pET30-*capV* as the template and the primer pair (forward) 5′-GATAG GCATA TGCCC GGACC GCAGA ATG-3′ (NdeI site underlined)/(reverse) 5′-CGAGT TAAGC TTTCC CCAAC GGAAT GTCCC GG-3′ (HindIII site underlined). After restriction digestion the PCR product was ligated into a pET30 DNA fragment originating from pET30-*capV* that was digested with the same enzymes. The resulting plasmid, pET30-C*capV*, was used for production of the C-terminal His$_6$-tagged CapV.

One-pot reaction with CapU, CapV and CapW: Reaction mixtures (100 μL) consisted of Tris·Cl (pH 7.5, 50 mM), L-Lys (5 mM), CoA (0.05 mM), ATP (5 mM), $MgCl_2$ (20 mM), **12** (2 mM), Svp (3.5 μM), CapU (3.5 μM), CapV (3.5 μM), and CapW (3.5 μM), at 30 °C for 8 h. After removal of protein by ultrafiltration, the reaction components were analyzed by reversed-phase chromatography with use of the Dionex Ultimate 3000 system equipped with an Acclaim 120 C-18 column (4.6 mm × 100 mm, 3 μm). A series of linear gradients from TFA (0.1%) in acetonitrile (2.5%, C) to TFA (0.1%) in acetonitrile (90%, D) was developed in the following manner (start time and end time with linear increase to % D): 0–5 min 0% D, 5–10 min 0–50% D, 10–15 min 50–100% D, and 15–18 min 100% D. The flow rate was kept constant at 1 mL min^{-1}, and elution was monitored at 260 nm. Peaks were identified by retention time and MS by comparison with control reactions or authentic standards.

Formation and analysis of L-ACL: Reaction mixtures (100 μL) consisted of Tris·HCl or phosphate (50 mM, pH 7.5), L-Lys (5 mM), ATP (5 mM), $MgCl_2$ (20 mM), Svp (3.5 μM), CapU/truncated CapU (A + T domains, 3.5 μM), and CapV or CapW (3.5 μM), at 30 °C for 2 h. After removal of the protein by ultrafiltration, the reaction components (50 μL) were treated with sodium bicarbonate (1 M, 50 μL) and dansyl chloride (25 μL of 10 mg mL^{-1} in DMF). After incubation for 1 h at 42 °C, reaction components were analyzed by use of the Agilent system described above. A series of linear gradients from formic acid (0.1%) in water (E) to formic acid (0.1%) in acetonitrile (F) was developed in the following manner (start time and end time with linear increase or decrease of % F): 0–18 min 5–90% F and 18–20 min 90–5% F. The flow rate was kept constant at 0.4 mL min^{-1}, and elution was monitored at 310 nm. Peaks were identified by comparison with the control reaction and authentic standards.

Synthesis of L-Lys-SNAC (13): Diisopropylethylamine (1.4 mL, 8.0 mmol) and *N*-acetylcysteamine (0.22 mL, 2.2 mmol) were sequentially added to a solution of *N*-Boc-L-Lys (0.692 g, 2.0 mmol) and PyBOP (2.0 g, 4.0 mmol) in CH_2Cl_2 (10 mL). The mixture was stirred at room temperature, and the reaction was monitored by TLC until completion (≈2 h). The solvent was removed under vacuum, and the residue was dissolved in ethyl acetate (50 mL). The organic layer was washed with brine (2 × 10 mL) and dried to give a colorless oil. *N*-Boc-L-Lys-SNAC was purified by silica column chromatography (isocratic EtOAc/hexane 1:1) as a white solid. *N*-Boc-L-Lys-SNAC was dissolved in CH_2Cl_2 (2 mL) at room temperature, and TFA (2 mL) was added dropwise. The resulting mixture was stirred for 1 h at room temperature, and the solvent was removed under vacuum. Compound **13** was purified by silica column chromatography as a colorless solid in 62% yield over two steps. ^{1}H NMR (400 MHz, [D$_4$]methanol): δ = 4.18 (t, J = 6.3 Hz, 1H), 3.32–3.43 (m, 2H), 3.03–3.19 (m, 2H), 2.93 (t, J = 7.6 Hz, 2H), 1.91–2.06 (m, 2H), 1.90 (s, 3H), 1.65–1.75 (m, 2H), 1.43–1.58 ppm (m, 2H); ^{13}C NMR (100 MHz, D_2O with methanol as standard): δ = 197.92, 173.65, 60.33, 49.99, 40.29, 39.71, 32.27, 29.82, 28.07, 22.76 ppm; HRMS (ESI): *m/z* calcd for $C_{10}H_{21}N_3O_2S + H^+$:248.1427; found: 248.1424.

Synthesis of L-Lys-PANT (14): 1) 2,2-Dimethoxypropane (4.88 mL, 40 mmol) was added to a solution of D-pantethine (1.526 g, 2 mmol) and PTSA (0.038 g, 0.2 mmol) in acetone (10 mL). The mixture was stirred at room temperature for 8 h, at which point the solvent was removed under vacuum. Isopropylidene-D-pantethine was purified by silica column chromatography (5–10% methanol/ethyl acetate mixture) as a white powder. 2) Isopropylidene-D-pantethine (0.635 g, 1.0 mmol) was dissolved in sodium bicarbonate solution (1 M, 5 mL) containing dithiothreitol (0.185 g, 1.2 mmol),

and the mixture was stirred at room temperature for 2 h. The mixture was freeze-dried, and the crude powder containing isopropyl-D-pantetheine was used in the next step without further purification. 3) Boc-L-Lys(Boc)-OH (0.415 g, 1.2 mmol), *N*,*N*-diisopropylethylamine (0.7 mL, 4.0 mmol), and PyBOP (1.04 g, 2.0 mmol) were sequentially added to a solution of isopropyl-D-pantetheine (0.318 g, 1.0 mmol) in CH_2Cl_2 (10 mL). The mixture was stirred at room temperature, and the reaction was monitored by TLC until completion (≈3 h). The solvent was removed under vacuum, and the residue was dissolved in ethyl acetate (30 mL). The organic layer was washed with $KHSO_4$ (5%, 2×10 mL), $NaHCO_3$ (5%, 2×10 mL), and brine (10 mL) and dried to give a colorless oil. *N*-Boc-L-Lys-PANT was purified by silica column chromatography (50–100% ethyl acetate/hexane mixture) as a hygroscopic white solid. 4) *N*-Boc-L-Lys-PANT (0.323 g, 0.5 mmol) was dissolved in hydrogen chloride (6 M, 2 mL), and the mixture was stirred at room temperature for 30 min. The reaction was quenched with sodium bicarbonate (1 M) until pH 5 was reached. Compound **14** was purified by semipreparative HPLC (water/methanol) as a white powder in 24% yield over four steps. ^{1}H NMR (400 MHz, [D_4]methanol): $\delta = 4.33$ (t, $J = 6.5$ Hz, 1H), 3.98–4.06 (m, 2H), 3.42–3.49 (m, 2H), 3.19–3.26 (m, 5H), 2.99–3.03 (m, 3H), 2.65–2.71 (m, 3H), 2.04–2.15 (m, 2H), 1.92–2.04 (m, 2H), 1.78 (dt, $J = 15.1$, 7.3 Hz, 3H), 1.50–1.67 (m, 4H), 1.20 (s, 3H), 1.03 ppm (s, 3H); ^{13}C NMR (100 MHz, D_2O with methanol as standard): $\delta = 197.66$, 179.47, 172.55, 77.35, 76.70, 60.32, 41.95, 40.38, 39.73, 37.25, 33.11, 32.21, 29.86, 28.16, 22.98, 22.93, 19.36 ppm; HRMS (ESI): *m/z* calcd for $C_{17}H_{34}N_4O_5S + H^+$: 407.2323; found: 407.2329.

Nonenzymatic hydrolysis or lactamization of 13 and 14: Reaction mixtures (100 μL) consisted of Tris·HCl or phosphate (50 mM, pH 7.5), **13** or **14** (1 mM), and 5,5-dithiobis-(2-nitrobenzoic acid) (2 mM) at 30 °C. Identical reactions were performed with the inclusion of **12** (2 mM) and CapW (10 μM). The production of 2-nitro-5-thiobenzoate was monitored at 412 nm with use of $\varepsilon = 14150\ M^{-1}\,cm^{-1}$. To validate the assay and to determine the relative amounts of amine-containing products, identical reactions were performed with phosphate buffer and without 5,5′-dithiobis-(2-nitrobenzoic acid), and the components were derivatized with dansyl chloride as described in the previous section. The reported data are averages of three replicates. To obtain rate constants, data were fit to equations describing simple parallel/competitive reactions in which *k* is equal to the sum of the individual rate constants for the formation of L-ACL and L-Lys.

Codon adaptation index: The codon adaptation index was calculated for the genes involved in the biosynthesis of **2** with the aid of http://genomes.urv.es/CAIcal/ when the codon usage from *S. griseus* subsp. *griseus* was used as a reference or of http://www.umbc.edu/codon/cai/cais.php when the codon usage of *S. griseus* subsp. *griseus* glucokinase gene was used as a reference.

Acknowledgements

This work was supported by National Institutes of Health grants AI087849 and UL1TR000117 (S.V.L.) and National Natural Science Foundation of China grants 81261120417, 81321004, and 81273414 (Z.Y.).

Keywords: antibiotics · biosynthesis · natural products · nonribosomal peptides · nucleosides

[1] Y. Muramatsu, A. Muramatsu, T. Ohnuki, M. Ishii, M. Kizuka, R. Enokita, S. Tsutsumi, M. Arai, Y. Ogawa, T. Suzuki, T. Takatsu, M. Inukai, *J. Antibiot.* **2003**, *56*, 243 – 252.
[2] Y. Muramatsu, M. M. Ishii, M. Inukai, *J. Antibiot.* **2003**, *56*, 253 – 258.
[3] Y. Muramatsu, S. Miyakoshi, Y. Ogawa, T. Ohnuki, M. M. Ishii, M. Arai, T. Takatsu, M. Inukai, *J. Antibiot.* **2003**, *56*, 259 – 267.
[4] T. Ohnuki, Y. Muramatsu, S. Miyakoshi, T. Takatsu, M. Inukai, *J. Antibiot.* **2003**, *56*, 268 – 279.
[5] Y. Muramatsu, T. Ohnuki, M. M. Ishii, M. Kizuka, R. Enokita, M. Shunichi, T. Takatsu, M. Inukai, *J. Antibiot.* **2004**, *57*, 639 – 646.
[6] M. Adamczeski, E. Quiñoà, P. Crews, *J. Am. Chem. Soc.* **1989**, *111*, 647 – 654.
[7] Y. Ikeda, H. Nonaka, T. Furumai, H. Onaka, Y. Igarashi, *J. Nat. Prod.* **2005**, *68*, 1061 – 1065.
[8] C. Ratledge, G. A. Snow, *Biochem. J.* **1974**, *139*, 407 – 413.
[9] V. Macko, M. B. Stimmel, T. J. Wolpert, L. D. Dunkle, W. Acklin, R. Bänateli, B. Juan, D. Arigoni, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9574 – 9578.
[10] B. D. Davidson, R. W. Schumacher, *Tetrahedron* **1993**, *49*, 6569 – 6574.
[11] M. Funabashi, K. Nonaka, C. Yada, M. Hosobuchi, N. Masuda, T. Shibata, S. G. Van Lanen, *J. Antibiot.* **2009**, *62*, 325 – 332.
[12] M. Funabashi, Z. Yang, K. Nonaka, M. Hosobuchi, Y. Fujita, T. Shibata, X. Chi, S. G. Van Lanen, *Nat. Chem. Biol.* **2010**, *6*, 581 – 586.
[13] E. A. Felnagle, E. E. Jackson, Y. A. Chan, A. M. Podevels, A. D. Berti, M. D. McMahon, M. G. Thomas, *Mol. Pharmaceutics* **2008**, *5*, 191 – 211.
[14] A. Koglin, C. T. Walsh, *Nat. Prod. Rep.* **2009**, *26*, 987 – 1000.
[15] L. Du, L. Lou, *Nat. Prod. Rep.* **2010**, *27*, 255 – 278.
[16] H. Itoh, M. Yamada, S. Tomino, K. Kurahashi, *J. Biochem.* **1969**, *64*, 259 – 261.
[17] U. Peschke, H. Schmidt, H. Z. Ahang, W. Piepersberg, *Mol. Microbiol.* **1995**, *16*, 1137 – 1156.
[18] W. Liu, K. Nonaka, L. Nie, J. Zhang, S. D. Christenson, J. Bae, S. G. Van Lanen, E. Zazopoulos, C. M. Farnet, C. F. Yang, B. Shen, *Chem. Biol.* **2005**, *12*, 293 – 302.
[19] K. Yanai, T. Murakami, M. Bibb, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9661 – 9666.
[20] V. Bergendahl, U. Linne, M. A. Marahiel, *Eur. J. Biochem.* **2002**, *269*, 620 – 629.
[21] S. A. Samel, G. Schoenafinger, T. A. Knappe, M. A. Marahiel, L.-O. Essen, *Structure* **2007**, *15*, 781 – 792.
[22] A. Tanovic, S. A. Samel, L.-O. Essen, M. A. Marahiel, *Science* **2008**, *321*, 659 – 663.
[23] C. Sánchez, L. Du, D. J. Edwards, M. D. Toney, B. Shen, *Chem. Biol.* **2001**, *8*, 725 – 738.
[24] T. Stachelhaus, A. Hüser, M. A. Marahiel, *Chem. Biol.* **1996**, *3*, 913 – 921.
[25] M. Gocht, M. A. Marahiel, *J. Bacteriol.* **1994**, *176*, 2654 – 2662.
[26] T. Stachelhaus, M. A. Marahiel, *J. Biol. Chem.* **1995**, *318*, 927 – 936.
[27] C. Coy, B. H. Paw, A. Bindereif, J. B. Nielands, *Biochemistry* **1986**, *25*, 2485 – 2489.
[28] Z. Simic, M. Weiwad, A. Schierhorn, C. Steegborn, M. Schutkowski, *ChemBioChem* **2015**, *16*, 2337 – 2347.
[29] M. L. Kuhn, B. Zemaitaitis, L. I. Hu, A. Sahu, D. Sorensen, G. Minasov, B. P. Lima, M. Scholle, M. Mrksich, W. F. Anderson, B. W. Gibson, B. Schilling, A. J. Wolfe, *PLoS One* **2014**, *9*, e94816.
[30] P. M. Sharp, W.-H. Li, *Nucleic Acid Res.* **1987**, *15*, 1281 – 1295.
[31] M. D. McMahon, J. S. Rush, M. G. Thomas, *J. Bacteriol.* **2012**, *194*, 2809 – 2818.
[32] N. M. Llewellyn, Y. Li, J. B. Spencer, *Chem. Biol.* **2007**, *14*, 379 – 386.
[33] B. Gust, G. Chandra, D. Jakimowicz, T. Yuqing, C. J. Bruton, K. F. Chater, *Adv. Appl. Microbiol.* **2004**, *54*, 107 – 128.
[34] T. Kieser, M. Bibb, M. Buttner, K. F. Chater, D. A. Hopwood in *Practical Streptomyces Genetics*, The John Innes Foundation, Norwich, **2000**.

Manuscript received: December 30, 2015
Accepted article published: February 3, 2016
Final article published: March 18, 2016