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

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

[a]Department of Pharmaceutical Sciences, College of Pharmacy University of Kentucky Lexington, KY 40536 (USA)

[b]Key Laboratory of Biotechnology of Antibiotics, Institute of Medicinal Biotechnology, Chinese Academy of Medicinal Sciences & Peking Union Medical College Beijing (China)

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

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

Abstract

Capuramycins are one of several known natural products that contain an L-Lys-derived L-α-amino-ε-caprolactam (L-ACL). The α-amino group of L-ACL in capuramycins is linked to an unsaturated hexuronic acid component through an amide bond that was previously shown to originate via an ATP-independent enzymatic route. Using 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 found 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 addi-tional domains or proteins and is likely a nonenzymatic event. The results set the stage to examine whether a similar NRPS-mediated mechanism is employed in the biosynthesis of other L-ACL-containing natural products and, just as intriguing, how spontaneous lactamization is avoided in the numerous NRPS-derived peptides that contain an unmodified L-Lys.

Graphical Abstract

The biosynthetic mechanism for the aminocaprolactam that is a component of the capuramycin antimycobacterial antibiotics was delineated using an in vivo and in vitro approach. The process is initiated by adenylation and thioesterification of L-Lys and highlighted by an apparently nonenzymatic intramolecular aminolysis.

[*] svanlanen@uky.edu.

Supporting information for this article is given via a link at the end of the document.
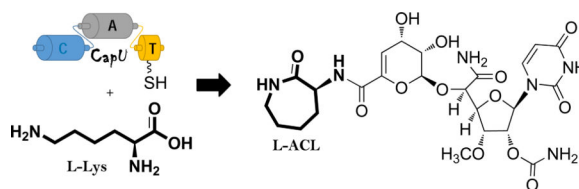
Keywords

Biosynthesis; nonribosomal peptide; antibiotic; natural product; nucleoside

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),[1-5] contain an L-α-amino-ε-caprolactam (L-ACL) that is covalently linked to an uncommon unsaturated hexuronic acid component by an amide bond. Several other natural products are also known that contain an ACL component, including bengamide A (**3**), peritoxin A (**4**), circinatin (**5**), caprolactin A (**6**), and the siderophores nocardimicins (**7**) and mycobactins (**8**) (Scheme 1).[6-10] Similarly to **1** and **2**, the **4-8** ACL, which is the L-isomer with the exception of **4** and **5**, is covalently linked to the rest of the molecule by an amide bond. Often the ACL is further modified such as the stereoselective C-methylation that is observed in the capuramycins A-500359 A (**9**) and A-503083 A (**10**) or the unusual animation of **4**.

The mechanism for the formation, attachment, and modification of ACL in **1-8** remains, for the most part, unknown. Not surprisingly, isotopic enrichment studies using 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**, respectively) encoding nonribosomal peptide synthetases (NRPS) that were proposed to be involved in L-ACL maturation.[11-12] Orf26/CapU was bioinformatically predicted to consist of three domains found in NRPS, a condensation (C), adenylation (A), and thiolation (T) domain (sequential from N- to C-terminus), while 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, A, and 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 posttranslationaly modified with a phosphopantetheine prosthetic group derived from coenzyme A (CoA), 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 oligomerization 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. Following complete elongation, which is dictated by the number of C-A-T modules in the NRPS, the resulting

peptide is typically released by a terminal thioesterase domain by either hydrolysis or intramolecular cyclization, although other mechanisms of chain release are known.[15]

Using the traditional amino acid-dependent ATP-^{32}PPi exchange assay to assess A domain substrate specificity,[16] preliminary characterization of the NRPS CapU was consistent with the A domain having a modest preference for ʟ-Lys.[12] Surprisingly, however, was the discovery that the attachment of the ʟ-ACL to the hexuronic acid 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 **2** biosynthes is appeared to be solely for lactamization of ʟ-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 of CapU nor CapV is necessary for lactamization.

To interrogate the role of the NRPS system in capuramycin biosynthes is, the development of a genetic system within 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 (Supporting Information Results and Figure S1), a phenomenon that has 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 A CL-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 corresponding methyl ester of **11** (Figure 1A and Figure 1B, trace i).[5] The *capU* gene was targeted for inactivation, and the expected double-crossover genotype was confirmed by PCR and Southern Blot analysis (Supporting Information, Figure S2). As expected, HPLC analysis of the *capU* mutant strain revealed that the production of **2** and **10** was abolished with the concomitant increase in the peaks corresponding to **11** and **12** (Figure 1B, trace ii). Upon feeding of ʟ-ACL to the *capU* mutant strain, the production of **2** was restored (Figure 1B, trace iii), demonstrating an essential role for CapU in ʟ-ACL biosynthesis. Moreover, these results are consistent with the previous characterization of CapW as an ʟ-ACL:**12** transacylase.[12] To further support their role, *capU* and *capV* were cloned into pUW201 containing the *ermE** constitutive promoter and expressed in the heterologous host *Streptomyces lividans* TK64. In contrast to the control consisting of the empty vector, expression of *capU* and *capV* in combination with feeding ʟ-Lys resulted in the formation of ʟ-ACL, which was detected by FMOC derivatization and LC-MS analysis (Supporting Information, 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 demonstrated recombinant CapU has a modest preference for ʟ-Lys compared to other proteinogenic amino acids, **11**, or ᴅ-Lys.[12] This was consistent with a mechanism wherein the A domain directly loads ʟ-Lys *in cis* to the holo-T domain, and one of the two available C domains catalyzes lactamization to release ʟ-ACL (Figure 1C). Analysis of model C domains by others has revealed a conserved His is essential for not only intermolecular aminolysis but also 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 whose structures have been solved revealed

this His is found in CapV (H119) but is substituted with Gln in CapU (Q184; Supporting Information, 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 *E. coli* using standard expression conditions (Supporting Information, 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 CapW, which transfers L-ACL to **12** and was previously 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 attainable from *Streptomyces lividans* TK24 (Supporting Information, 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 to controls (e.g., Figure 2A, traces i and ii), HPLC analysis of the reactions containing all four proteins starting with substrates L-Lys and **12** revealed a new peak that coeluted with authentic **2** that was 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** (calc. 612.2) (all conditions tested are summarized in the Supporting Information, 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 been reported for T domains of other NRPS systems.[24-26] The heterologous production of some holo-CapU may also explain why L-Lys was only moderately preferred in the amino acid-dependent ATP-^{32}PPi exchange assay, as 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, consistent with the hypothesis that L-Lys is first activated as an acyl-adenylate and transferred to the T domain of holo-CapU. In contrast to our expectations, however, CapV was not essential nor enhanced **2** formation when starting with L-Lys (Figure 2A, trace iv), suggesting that CapU was the sole responsible catalyst for L-Lys activation and lactamization.

We next aimed to directly monitor L-ACL formation by removing **12** from the reaction and derivatizing any product and remaining L-Lys with dansyl chloride prior to analysis (Figure 2B). In comparison to the appropriate controls, HPLC analysis clearly revealed that L-ACL was formed under the condition that CapU and ATP were included in the reaction (Figure 2B, trace i). Moreover, a truncated didomain version of CapU was prepared that consisted of the A and T domains (CapU_AT). Similarly to the native protein, CapU_AT catalyzed the formation of L-ACL from L-Lys (Figure 2B, trace ii), clearly demonstrating a functional A and T domain for this recombinant protein. With hopes to further simplify the interpretation of the results, L-Lys-*S*-*N*-acetylcysteamine **13** and the 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. Following completion of the reaction, a ratio of 62:38 and 90:10 of L-ACL:L-Lys was detected starting from **13** and **14,** respectively.

By monitoring the formation of the free sulfhydryl 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] 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 a potential mixed kinetics with second order at higher concentrations of **14** ($k = 5.5 \times 10^{-3}$ mM^{-1}min^{-1}; Figure 3B) and first order at low er concentrations ($k = 1.7 \times 10^{-3}$ min^{-1}; Supporting Information, Figure S6). These kinetic parameters are comparable to the reported enzymatic N-acetylation of N^ε-OH-Lys using 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 to infer any mechanistic insight from this data, it is nevertheless clear that L-ACL is nonenzymatically formed from L-Lys-thioesters at a significant rate.

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

The revelation that neither C domain of CapU or CapV is essential for lactamization of L-Lys raises the question of the role of these domains in **1** and **2** biosynthesis, if any. 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 *capU*, *V*, and *W* form a subcluster that are flanked by genes encoding putative transposases followed by long intergenic regions (Supporting Information, Figure S7). Perhaps more indicative, however, is that the calculated codon adaptation index for *capV* (0.401) and *capU* (0.495) is substantially lower than the predicted 17 open reading frames involved in **2** biosynthesis (0.521-0.779) when referenced against the codon usage of *Streptomyces griseus* subs. griseus (Table S3).[30] A similar trend is observed when using the housekeeping glucokinase gene from *S. griseus* subs. 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 further interrogate the unexpected lack of requirement of CapV and the C domain of CapU for **2** biosynthesis in vitro.

Several NRPS-derived peptides contain a Lys wherein the ε-amine is either unmodified, hydroxylated and acylated, or undergoes lactamization to form a macrocyclic peptide or ACL. A domains of the NRPS involved in the biosynthesis of **8**, which contains two Lys-derived amino acids (Figure 1), are the only enzymes for these ACL-containing compounds that have been biochemically interrogated in vitro.[31] When using the amino acid-dependent

ATP-^{32}PP$_i$ exchange assay, the results 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 per final product, have revealed that in both cases the ε-amine is modified prior to A domain activation. MbtE is most active with $N^ε$-acyl-$N^ε$-OH-L-Lys, suggesting it is responsible for incorporating the internal L-Lys unit, and the modifications prior to thioesterification potentially protect against lactamization. On the other hand, MbtF is most active with $N^ε$-OH-L-Lys, suggesting that this NRPS component is responsible for incorporating the terminal L-Lys unit that is converted to L-ACL. Similarly to the capuramycin NRPS system, there is an extra C domain at the terminus of MbtF that was proposed to catalyze 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 into NRPS-derived peptides. 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 possibly a natural protection/deprotection strategy is used analogous to the biosynthesis of the 4-amino-2-hydroxybutyrate component found in the aminoglycoside butirosin.[32]

Experimental Section

Gene inactiv ation of *capU*

The *capU* gene was inactivated 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 following primer pair: (forward) 5′-CCTCGCGGCGCGCACCTCCGCGGCAACGTCTCTGAGGTGCCATTCCGGGGATCCGTCGACC-3′ / (reverse) 5′-GCGGTACGCCGAACTTCTCCACCACCCCTCCGACGATTTTGTAGGCTGGAGCTGCTTC-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 of either wild-type or the *capU* mutant strain was isolated, digested with *Xho*I, and applied to agarose gel electrophoresis for detection with a DIG labeled probe generated from PCR using primers complementary to regions flanking the *capU* gene.

Biotransformation with CapU and CapV

The *capU* and *capV* genes were amplified by PCR as a single fragment using Expand Long Template PCR System from KOD PLUS Neo with supplied Buffer, 200 μM dNTPs, 5 % DMSO, 10 ng DNA template pN-4,[12] 5 U DNA polymerase, and 400 nM each of the following primer pairs: (forward) 5′-GGGAATTCCATATGGTGGACGCCCCGCGTCACTGGA-3′ (*Nde*I site underlined)/ (reverse) 5′-CCCAAGCTTTCACCAACGGAATGTCCCGGCGATT-3′ (*Hind*III site

underlined). 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, 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 using PEG-mediated protoplast transformation.[34] After 20 h at 28 °C, plates were overlaid with 1 mL of soft nutrient agar supplemented with 200 μg of thiostrepton. After 3 additional days at 28 °C, single colonies were transferred to fresh R2YE plates supplemented with 50 μg/mL thiostrepton. After 4 days at 28 °C, positive transformants were confirmed by colony PCR using InstaGene Matrix from Bio-Rad (Hercules, CA) and LA-Taq polymerase with GC buffer II from Takara Bio Inc. (Shiga, Japan). Positive transformants were used to inoculate 50 mL R2YE containing 25 μg/mL thiostrepton and grown for 2 days at 28 °C at 250 rpm, at which time 2 mL was transferred to fresh 100 mL R2YE containing 25 μg/mL thiostrepton. Following growth for 3 days at 28 °C at 250 rpm, L-Lys was added (25 mg/mL 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 1 M NaOH. The mixture was applied to a column containing XAD16 resin and sequentially washed with 5 volumes each of water, 20 % acetone, 50 % acetone, 80 % acetone. Fractions eluting with 20 % acetone were collected and dried under vacuum. The dried solid containing L-ACL from a 100 mL culture was dissolved in 500 μL of sodium bicarbonate solution (20 mM), and an equal volume of FMOC-OSU acetonitrile solution (2 mM) was added. After the mixture was stirred for 5 min, samples were analyzed using an Agilent 1200 Series Quaternary LC system equipped with an Eclipse XDB-C18 column (150 mm × 4.6 mm, 5 μm, 80Å). A series of linear gradients was developed from 0.1 % TFA in water (A) to 0.1 % in 90 % acetonitrile (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, 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 using Expand Long Template PCR System from Roche (Indianapolis, IN, USA) with supplied Buffer 2, 200 μM dNTPs, 5 % DMSO, 10 ng pN-4,[12] 5 U DNA polymerase, and 400 nM each of the following primer pair: *capU* (forward) 5′-GGTATTGAGGGTCGCATGGACGCCCCGCGTCACTG-3′/(reverse) 5′-AGAGGAGAGTTAGAGCCTCAGGGCGAGGAGTCGACATAG-3′; *capV* (forward) 5′-GGTATTGAGGGTCGCATGCCCGGACCGCAGAATG-3′/(reverse) 5′-AGAGGAGAGTTAGAGCCTCACCAACGGAATGTCCCG-3′; and *capU*_AT (forward) 5′-AAAAAACATATGGGTGATGTCATCGGCCCC-3′/(reverse) 5′-AAAAAAGGATCCTCAGGGCGAGGAGTCGACA-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 using 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 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 using pET30-*capV* as the template and the following primer pair: (forward) 5′-GATAGG<u>CATATG</u>CCCGGACCGCAGAATG-3′ (*NdeI* site underlined)/(reverse) 5′-CGAGTT<u>AAGCTT</u>TCCCCAACGGAATGTCCCGG-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

Reactions (100 μL) consisted of 50 mM Tris-Cl (pH 7.5), 5 mM L-Lys, 0.05 mM CoA, 5 mM ATP, 20 mM MgCl$_2$, 2 mM **12**, 3.5 μM Svp, 3.5 μM CapU, 3.5 μM CapV, and 3.5 μM CapW at 30 °C for 8 h. Following removal of protein by ultrafiltration, the reaction components were analyzed using reverse-phase chromatography using 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 0.1 % TFA in 2.5 % acetonitrile (C) to 0.1 % TFA in 90 % acetonitrile (D) was developed in the following manner (beginning time and ending time with linear increase to % D): 0-5 min, 0 % D; 5-10 min, 0-50 % D; 10-15 min, 50-100 % D; 15-18 min, 100 % D. The flow rate was kept constant at 1 mL/min, and elution was monitored at 260 nm. Peaks were identified by retention time and MS by comparison to control reactions or authentic standards.

Formation and analysis of L-ACL

Reactions (100 μL) consisted of 50 mM Tris-Cl or phosphate (pH 7.5), 5 mM L-Lys, 5 mM ATP, 20 mM MgCl$_2$, 3.5 μM Svp, 3.5 μM CapU/Truncated CapU (A+T domains), and 3.5μM CapV or CapW at 30 °C for 2 h. Following removal of the protein by ultrafiltration, the reaction components (50 μL) were treated with 1 M sodium bicarbonate (50 μL) and dansyl chloride (25 μL of 10 mg/mL in DMF). After incubating 1 h at 42 °C, reaction components were analyzed using the Agilent system described above. A series of linear gradients from 0.1 % formic acid in water (E) to 0.1 % formic acid in acetonitrile (F) was developed in the following manner (beginning time and ending 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, and elution was monitored at 310 nm. Peaks were identified by comparison to 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 DCM (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 (50–100 % ethyl acetate/hexane mixture) as a

white solid. *N*-Boc-ʟ-Lys-SNAC was dissolved in DCM (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 removed under vacuum. **13** was purified by silica column chromatography as a colorless solid in 62 % yield over two steps. Chemical characterization: ^{1}H NMR (400 MHz, Methanol-d_4) δ ppm 4.18 (t, *J*=6.3 Hz, 1 H), 3.32-3.43 (m, 2 H), 3.03-3.19 (m, 2 H), 2.93 (t, *J*=7.6 Hz, 2 H), 1.91-2.06 (m, 2 H), 1.90 (s, 3 H), 1.65-1.75 (m, 2 H), 1.43-1.58 (m, 2 H); ^{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. HRMS (ESI): $C_{10}H_{21}N_3O_2S+H^+$, Calc: 248.1427, Found: 248.1424.

Synthesis of ʟ-Lys-PANT (14)

a) 2,2-dimethoxypropane (4.88 mL, 40 mmol) was added to a solution of ᴅ-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 time the solvent was removed under vacuum. Isopropyl-ᴅ-pantethine was purified by silica column chromatography (5-10 % methanol/ ethyl acetate mixture) as a white powder. b) Isopropyl-ᴅ-pantethine (0.635 g, 1.0 mmol) was dissolved into a 1 M sodium bicarbonate solution (5 mL) containing dithiothreitol (0.185 g, 1.2 mmol) and stirred at room temperature for 2 h. The mixture was freeze-dried and the crude powder containing isopropyl-ᴅ-pantetheine was used in the next step without further purification. c) Boc-ʟ-Lys(Boc)-OH (0.415g, 1.2 mmol), *N,N*-diisopropylethylamine (0.7 mL, 4.0 mmol), and PyBOP (1.04g, 2.0 mmol) were sequentially added to a solution of isopropyl-ᴅ-pantetheine (0.318 g, 1.0 mmol) in DCM (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 dissolved in ethyl acetate (30 mL). The organic layer was washed with 5 % $KHSO_4$ (2 × 10 mL), 5 % $NaHCO_3$ (2 × 10 mL), and brine (10 mL) and dried to give a colorless oil. *N*-Boc-ʟ-Lys-PANT was purified by silica column chromatography (50-100 % ethyl acetate/hexane mixture) as a hygroscopic white solid. d) *N*-Boc-ʟ-Lys-PANT (0.323g, 0.5 mmol) was dissolved in 6 M hydrogen chloride (2 mL) and stirred at room temperature for 30 min. The reaction was quenched with 1 M sodium bicarbonate until pH = 5 was reached. **14** was purified by semi-preparative HPLC (water/methanol) as a white powder in 24 % yield over four steps. Chemical characterization: ^{1}H NMR (400 MHz, Methanol-d_4) δ ppm 4.33 (t, *J*=6.5 Hz, 1 H), 3.98-4.06 (m, 2 H), 3.42-3.49 (m, 2 H), 3.19-3.26 (m, 5 H), 2.99-3.03 (m, 3 H), 2.65-2.71 (m, 3 H), 2.04 - 2.15 (m, 2 H), 1.92-2.04 (m, 2 H), 1.78 (dt, *J*=15.1, 7.3 Hz, 3 H), 1.50-1.67 (m, 4 H), 1.20 (s, 3 H), 1.03 (s, 3 H); ^{13}C NMR (100 MHz, D_2O with methanol as standard) δ 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. HRMS (ESI): $C_{17}H_{34}N_4O_5S+H^+$, Calc: 407.2323, Found: 407.2329

Nonenzymatic hydrolysis or lactamization of 13 and 14

Reactions (100 μL) consisted of 50 mM Tris-Cl or phosphate (pH 7.5), 1 mM **13** or **14**, and 2 mM 5,5-dithiobis-(2-nitrobenzoic acid) at 30 °C. Identical reactions were performed with the inclusion of 2 mM **12** and 10 μM CapW. The production of 2-nitro-5-thiobenzoate was monitored at 412 nm using ε = 14,150 $M^{-1}cm^{-1}$. To validate the assay and 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 derivatized with dansyl chloride as described in the previous section. The reported data are an average of three replicates. To obtain rate constants, data were fitted to equations describing simple parallel/competitive reactions where 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 2 biosynthesis using http://genomes.urv.es/CAIcal/ when using the codon usage from *Streptomyces griseus* subs. griseus as a reference or http://www.umbc.edu/codon/cai/cais.php when using the codon usage of *Streptomyces griseus* subs. griseus glucokinase gene as a reference.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Muramatsu Y, Muramatsu A, Ohnuki T, Ishii M, Kizuka M, Enokita R, Tsutsumi S, Arai M, Ogawa Y, Suzuki T, Takatsu T, Inukai M. J. Antibiot. 2003; 56:243–252. [PubMed: 12760680]

2. Muramatsu Y, Ishii MM, Inukai M. J. Antibiot. 2003; 56:253–258. [PubMed: 12760681]

3. Muramatsu Y, Miy akoshi S, Ogawa Y, Ohnuki T, Ishii MM, Arai M, Takatsu T, Inukai M. J. Antibiot. 2003; 56:259–267. [PubMed: 12760682]

4. Ohnuki T, Muramatsu Y, Miy akoshi S, Takatsu T, Inukai M. J. Antibiot. 2003; 56:268–279. [PubMed: 12760683]

5. Muramatsu Y, Ohnuki T, Ishii MM, Kizuka M, Enokita R, Shunichi M, Takatsu T, Inukai M. J. Antibiot. 2004; 57:639–646. [PubMed: 15638324]

6. Adamczeski M, Quiñoà E, Crews PJ. Am. Chem. Soc. 1989; 111:647–654.

7. Ikeda Y, Nonaka H, Furumai T, Onaka H, Igarashi Y. J. Nat. Prod. 2005; 68:1061–1065. [PubMed: 16038549]

8. Ratledge C, Snow GA. Biochem. J. 1974; 139:407–413. [PubMed: 4614794]

9. Macko V, Stimmel MB, Wolpert TJ, Dunkle LD, Acklin W, Bänateli R, Juan B, Arigoni D. Proc. Natl. Acad. Sci. 1992; 89:9574–9578. [PubMed: 1409667]

10. Dav idson BD, Schumacher RW. Tetrahedron. 1993; 49:6569–6574.

11. Funabashi M, Nonaka K, Yada C, Hosobuchi M, Masuda N, Shibata T, Van Lanen SG. J. Antibiot. 2009; 62:325–332. [PubMed: 19478828]

12. Funabashi M, Yang Z, Nonaka K, Hosobuchi M, Fujita Y, Shibata T, Chi X, Van Lanen SG. Nat. Chem. Biol. 2010; 6:581–586. [PubMed: 20562876]

13. Felnagle EA, Jackson EE, Chan YA, Podev els AM, Berti AD, McMahon MD, Thomas MG. Mol. Pharm. 2008; 5:191–211. [PubMed: 18217713]

14. Koglin A, Walsh CT. Nat. Prod. Rep. 2009; 26:987–1000. [PubMed: 19636447]

15. Du, L. L, Lou. Nat. Prod. Rep. 2010; 27:255–278. [PubMed: 20111804]

16. Itoh H, Yamada M, Tomino S, Kurahashi K. J. Biochem. 1969; 64:259–261. [PubMed: 5708329]

17. Peschke U, Schmidt H, Ahang HZ, Piepersberg W. Mol. Microbiol. 1995; 16:1137–1156. [PubMed: 8577249]
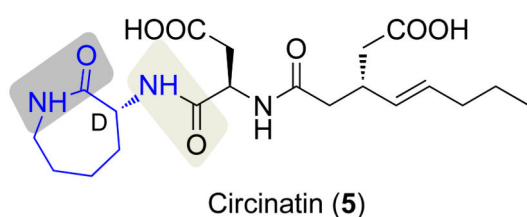
18. Liu W, Nonaka K, Nie L, Zhang J, Christenson SD, Bae J, Van Lanen SG, Zazopoulos E, Farnet CM, Yang CF, Shen B. Chem. Biol. 2005; 12:293–302. [PubMed: 15797213]

19. Yanai K, Murakami T, Bibb M. Proc. Natl. Acad. Sci. USA. 2006; 103:9661–9666. [PubMed: 16766657]

20. Bergendahl V, Linne U, Marahiel MA. Eur. J. Biochem. 2002; 269:620–629. [PubMed: 11856321]

21. Samel SA, Schoenaf inger G, Knappe TA, Marahiel MA, Essen L-O. Structure. 2007; 15:781–792. [PubMed: 17637339]

22. Tanov ic A, Samel SA, Essen L-O, Marahiel, M. A. MA. Science. 2008; 321:659–663. [PubMed: 18583577]

23. Sánchez C, Du L, Edwards DJ, Toney MD, Shen B. Chem. Biol. 2001; 8:725–738. [PubMed: 11451672]

24. Stachelhaus T, Hüser A, Marahiel MA. Chem. Biol. 1996; 3:913–921. [PubMed: 8939706]

25. Gocht M, Marahiel MA. J. Bacteriol. 1994; 176:2654–2662. [PubMed: 8169215]

26. Stachelhaus T, Marahiel MA. J. Biol. Chem. 1995; 318:927–936.

27. Coy C, Paw BH, Bindereif A, Nielands JB. Biochemistry. 1986; 25:2485–2489. [PubMed: 3521734]

28. Simic Z, Weiwad M, Schierhorn A, Steegborn C, Schutkowski M. ChemBioChem. 2015; 16:2337–2347. [PubMed: 26382620]

29. Kuhn ML, Zemaitaitis B, Hu LI, Sahu A, Sorensen D, Minasov G, Lima BP, Scholle M, Mrksich M, Anderson WF, Gibson BW, Schilling B, Wolfe AJ. PLOS ONE. 2014; 9:e94816. [PubMed: 24756028]

30. Sharp PM, Li W-H. Nucl. Acid Res. 1987; 15:1281–1295.

31. McMahon MD, Rush JS, Thomas MG. J. Bacteriol. 2012; 194:2809–2818. [PubMed: 22447909]

32. Llewelly n NM, Li Y, Spencer JB. Chem. Biol. 2007; 14:379–386. [PubMed: 17462573]

33. Gust B, Chandra G, Jakimowicz D, Yuqing T, Bruton CJ, Chater KF. Adv. Appl. Microbiol. 2004; 54:107–128. [PubMed: 15251278]

34. Kieser, T.; Bibb, M.; Buttner, M.; Chater, KF.; Hopwood, DA. Practical Streptomyces Genetics. The John Innes Foundation; Norwich, UK: 2000.
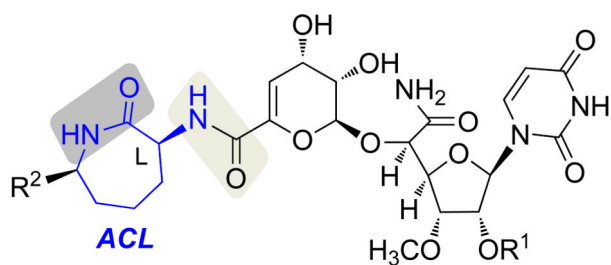
A-500359 B (**1**) R^1 = H, R^2 = H
A-503083 B (**2**) R^1 = CONH$_2$, R^2 = H
A-500359 A (**9**) R^1 = H, R^2 = CH$_3$
A-503083 A (**10**) R^1 = CONH$_2$, R^2 = CH$_3$

Bengamide A (**3**)

Caprolactin A (**6**)

Peritoxin A (**4**)
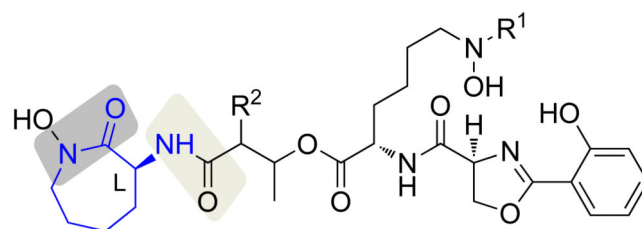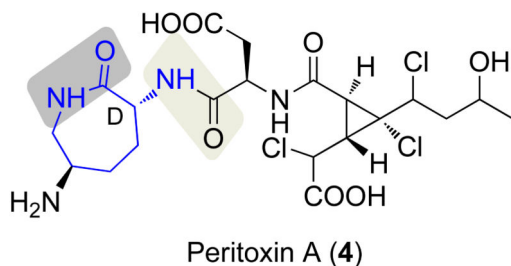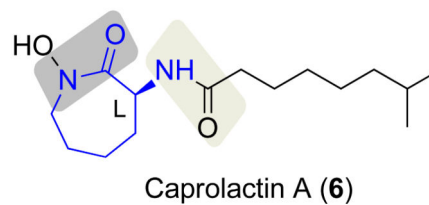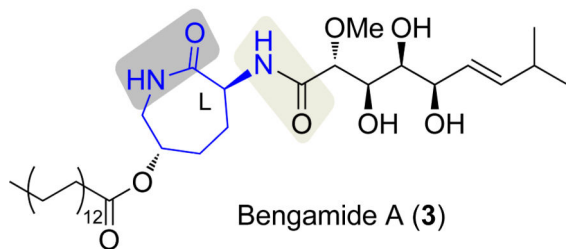
Nocardimicin A (**7**) R^1 = CHO, R^2 = $\underset{11}{\overbrace{}}$

Mycobactins (**8**) R^1 = $\underset{14-17}{\overbrace{}}$ R^2 = H

Circinatin (**5**)

Scheme 1.
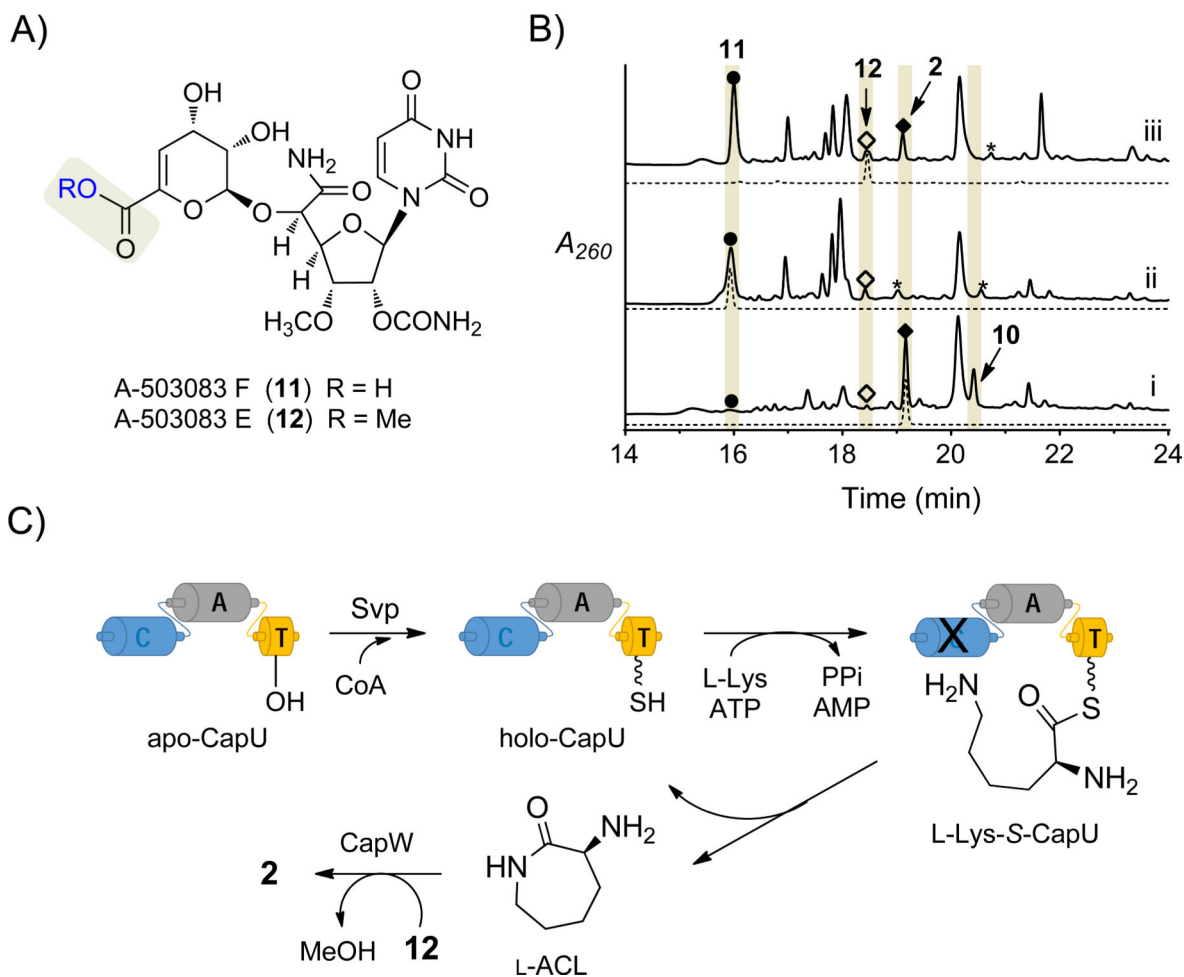Representative natural products that contain an α-amino- ε-caprolactam (ACL).

Figure 1.
Role of the NRPS system in L-ACL biosynthesis. A) Structures of the deaminocaprolactam congeners. B) HPLC chromatograms of the methanol extract from the i) wild-type strain; ii) *capU* mutant strain; and iii) *capU* mutant strain cultured with exogenously supplied L-ACL. A_{260}, absorbance at 260 nm; asterisk (*) denotes an unidentified metabolite with a unique UV/Vis spectrum relative to capuramycins. C) In vitro reconstitution of L-ACL biosynthesis starting from L-Lys and apo-CapU using Svp with or without additional enzymes.
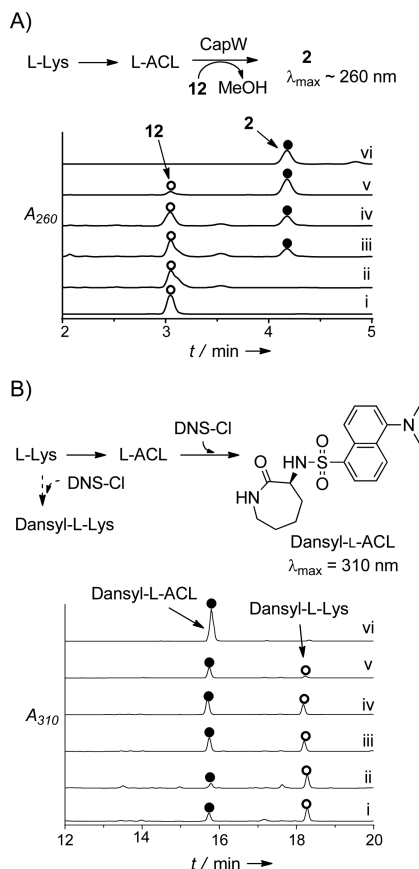
Figure 2.

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 including i) authentic **12**; ii) reactions omitting CapU, CapV, and Svp, thus confirming CapW is unable to incorporate L-Lys directly; iii) reactions containing all of the necessary enzymes (CapU, CapV, CapW, and Svp) and substrates/cofactors (L-Lys, ATP, **12**, CoA, and MgCl$_2$); iv) reaction omitting CapV; v) positive control confirming CapW-catalyzed transacylation starting from L-ACL and **12**; vi) authentic **2**. A_{260}, 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 including i) reactions omitting CapV, thus confirming CapV is not necessary for lactamization; ii) reactions omitting CapV and using truncated CapU_AT (A+T domains); iii) reactions using L-Lys-SNAC (**13**) as the substrate in place of L-Lys; d) reactions starting from **13** without inclusion of any enzymes, demonstrating nonenzymatic lactamization and hydrolysis in buffered aqueous conditions; v) reactions using **14** as the substrate, which favors lactamization over hydrolysis; vi) reaction omitting all of enzymes only with the L-Lys, and vi) control of L-ACL modified with dansyl chloride. A_{310}, absorbance at 310 nm. Dansyl-L-Lys is L-Lys modified with two dansyl groups based on MS analysis.
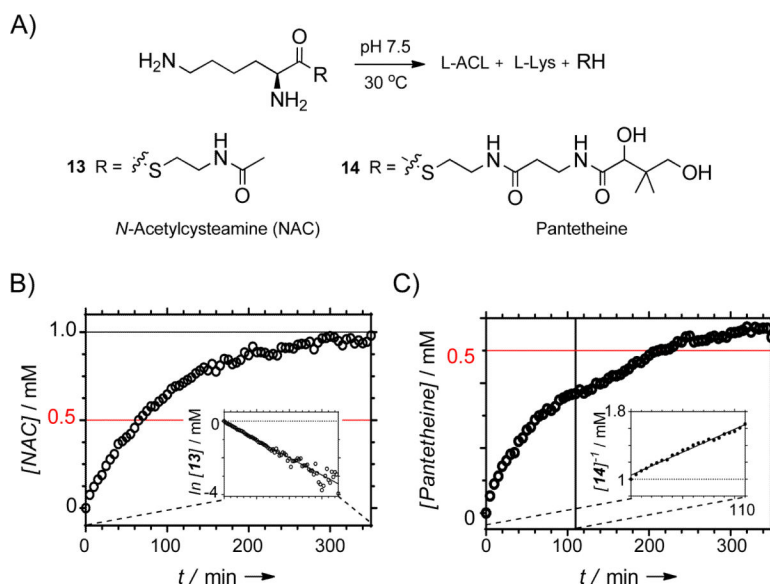
Figure 3.
Nonenzymatic lactamization/hydrolysis. A) Structures of the surrogate substrates **13** and **14**.
B) Time dependence of the nonenzymatic reaction of **13;** the inset is suggestive of first order
kinetics. C) Time dependence of the nonenzymatic reaction of **14;** the inset depicts the
initial stages of the reaction ($t < 110$ min) that is suggestive of second order kinetics.