Chain Initiation in the Leinamycin-producing Hybrid Nonribosomal Peptide/Polyketide Synthetase from Streptomyces atroolivaceus S-140

DISCRETE, MONOFUNCTIONAL ADENYLATION ENZYME AND PEPTIDYL CARRIER PROTEIN THAT DIRECTLY LOAD D-ALANINE*

Received for publication, April 3, 2007, and in revised form, April 27, 2007 Published, JBC Papers in Press, May 14, 2007, DOI 10.1074/jbc.M702814200

Gong-Li Tang†, Yi-Qiang Cheng‡, and Ben Shen§*†

From the †Division of Pharmaceutical Sciences, ‡University of Wisconsin National Cooperative Drug Discovery Group, §Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53705

Nonribosomal peptide natural products are biosynthesized from amino acid precursors by nonribosomal peptide synthetases (NRPSs), which are organized into modules. For a typical NRPS initiation module, an adenylation (A) domain activates an amino acid and installs it onto a peptidyl carrier protein (PCP) domain as a thioester; an elongation module, which has a condensation (C) domain located between every consecutive pair of A and PCP domains, catalyzes the formation of the peptide bond between the upstream aminoacyl/peptidyl-\(S\)-PCP and the free amino group of the downstream aminoacyl-\(S\)-PCP. Amino acid constituents in peptide natural products usually arise from the \(L\)-enantiomers through the action of integral epimerization (E) domains of an NRPS. The biosynthetic gene cluster for leinamycin, a hybrid nonribosomal peptide/polyketide containing a \(D\)-alanine moiety, does not encode a typical NRPS initiation module with the expected A-PCP-E domains; instead, it has only an A protein (LnmQ) and a PCP (LnmP), both of which are encoded by separate genes. Here we show the results of biochemical experiments as follows: (i) we demonstrate that LnmQ directly activates \(D\)-alanine as \(D\)-alaninyl-AMP and installs it onto LnmP to generate a \(D\)-alaninyl-\(S\)-PCP intermediate; (ii) we confirm that amination of LnmN by LnmQ in \textit{trans} is the result of specific communication between the separate A and PCP proteins; and (iii) we reveal that leinamycin production can be improved by supplementation of exogenous \(D\)-alanine in the fermentation broth of \textit{Streptomyces atroolivaceous} S-140. These findings unveil an unprecedented NRPS initiation module structure that is characterized by a discrete \(D\)-alanine-specific A protein and a PCP.

Nonribosomal peptide natural products include many clinically important drugs such as vancomycin (antibacterial), bleomycin (anticancer), and cyclosporine (immunosuppressant). They are biosynthesized from amino acid precursors by nonribosomal peptide synthases (NRPSs) that are multifunctional enzymes organized into modules (1). Three groups of NRPS systems are known to date (2). (i) Type A (linear) NRPSs are analogous to type I polyketide synthases (PKSs) (3, 4) in which each module harbors a set of distinct, noniteratively acting activities responsible for the catalysis of one cycle of peptide chain elongation. In this system, the amino acid sequence of the resultant linear peptide chain is co-linear with the number and order of the modules. (ii) Type B (iterative) NRPSs use their modules or domains more than once in the assembly of a single product. This strategy is employed to build up peptide chains containing repeated smaller sequences. (iii) Type C (nonlinear) NRPSs harbor at least one unusual arrangement of the modules or core domains with two primary differences from typical type A NRPSs. One is the unusual internal cyclization or branch point synthase that leads to a deviation from the normal function of a linear NRPS, resulting in a nonlinear peptide product. The other is the specialized condensation or ligase domain that incorporates small soluble molecules, such as amines, into the assembled peptides.

Three enzymatic activities, which normally are specified by a module that contains condensation-adenylation-peptidyl carrier protein (C-A-PCP) domains, are necessary for one complete elongation cycle of peptide synthesis catalyzed by an NRPS. The A domain selects the correct amino acid from the pool of available substrates and activates it as an aminoacyl adenylate, and the activated amino acid is then transferred onto the \(-\text{SH}\) group of the \(4'\)-phosphopantetheinyl (\(4'\) PP) arm attached to the PCP domain. Next, the C domain, usually present between every consecutive pair of A and PCP domains, catalyzes the peptide bond formation between the upstream aminoacyl/peptidyl-\(S\)-PCP and the free amino group of the downstream aminoacyl-\(S\)-PCP, thus facilitating the translocation of the growing peptide chain onto the next module. Correspondingly, a typical initiation module requires two catalytic

*This work was supported in part by University of California BioSTAR Program Grant Bio199–10045 and Kosan Biosciences, Inc. (Hayward, CA) and National Institutes of Health Grants CA106150, CA113297, and AI51687. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†To whom correspondence should be addressed: Division of Pharmaceutical Sciences, School of Pharmacy, University of Wisconsin–Madison, 777 Highland Ave., Madison, WI 53705. Tel.: 608-262-2673; Fax: 608-262-5245; E-mail: bshen@pharmacy.wisc.edu.

The abbreviations used are: NRPS, nonribosomal peptide synthetase; A, adenylation; C or C’, condensation; ESI-MS, electrospray ionization-mass spectrometry; F, formylation; HPLC, high performance liquid chromatography; IPTG, isopropyl \(\beta\)-D-thiogalactopyranoside; LNM, leinamycin; PCP, peptidyl carrier protein; PKS, polyketide synthase; PPi, inorganic pyrophosphate; 4’PP, 4’-phosphopantetheine; DTT, dithiothreitol.

2 The abbreviations used are: NRPS, nonribosomal peptide synthetase; A, adenylation; C or C’, condensation; ESI-MS, electrospray ionization-mass spectrometry; F, formylation; HPLC, high performance liquid chromatography; IPTG, isopropyl \(\beta\)-D-thiogalactopyranoside; LNM, leinamycin; PCP, peptidyl carrier protein; PKS, polyketide synthase; PPi, inorganic pyrophosphate; 4’PP, 4’-phosphopantetheine; DTT, dithiothreitol.
activities arranged as A-PCP didomain for the formation of a starter aminocacyl-S-PCP; a native elongation module can be biochemically converted into an initiation module by deleting the C domain (5).

One of the hallmarks of nonribosomal peptide natural products is the presence of D-amino acids. For example, three of the four amino acids of glycopeptidolipids (6), three of the seven amino acids of mycosubtilin (7), and four of the seven amino acids of vancomycin (8) have the D-configuration. In general, the A domains of NRPSs select the readily available L-amino acids for activation, and epimerization (E) domains generate a D-configuration, whereas the amino acids is tethered to the PCP (1, 2, 3, 5).

D-Alanine is the only exception to this rule in a few cases. A number of nonribosomal peptide natural products, such as HC toxin (9), cyclosporin (10), daptomycin (11), microcystin (12), dendoamide A (13), and ramoplanin (14), contain D-Ala. Sequence data suggest that the internal C-A-PCP-E module, which activates l-Ala and then epimerizes l-Ala into D-Ala, has been found in the NRPS genes for daptomycin (11) and microcystin (12) biosynthesis. The NRPS module for D-Ala incorporation in cyclosporine or HC toxin biosynthesis, however, does not contain an integrated E domain and cannot activate l-Ala; instead, D-Ala is among the pool of available amino acids and is directly activated. In these two cases, D-Ala is provided by a dedicated Ala racemase that is encoded by a separate gene (15, 16).

Leinamycin (LNM) (Fig. 1A), a hybrid nonribosomal peptide/polyketide natural product containing a D-Ala, exhibits potent antitumor activity, especially against tumors that are resistant to clinically important anticancer drugs (17). We have reported on the identification, localization, sequencing analysis, and genetic and biochemical characterization of the biosynthetic gene cluster for LNM biosynthesis in Streptomyces atroolivaceus S-140 (18–21). Sequencing analysis of the lnmQ and lnmP genes, cells harboring pBS3049 (for LnmQ and LnmP proteins) or pBS3021 (for LnmQ and LnmP proteins, cells harboring pBS3048 (for LnmQ and LnmP proteins, cells harboring pBS3049). The resulting plasmid, as well as pBS3021, an lnmP expression plasmid constructed previously (19), was transformed separately into E. coli BL21(DE3) for gene expression. Overexpression of the LnmQ and LnmP proteins, cells harboring pBS3049 (for lnmQ) or pBS3021 (for lnmP) (19) were grown in LB medium supplemented with 50 μg/ml kanamycin. Cultures (2 × 500 ml) were grown to an A600 of 0.4–0.6 at 37 °C and then cooled to 18 °C for 30 min, and gene expression was induced by the addition of 0.2 mM IPTG. Cultures were grown for an additional 18 h. The purification of the His6- or His8-tagged fusion protein with nickel-nitrilotriacetic acid affinity resin was performed according to the manufacturer’s manual (Qiagen, Valencia, CA), and the resultant proteins were dialyzed against 25 mM Tris-HCl (pH 7.5), 25 mM NaCl, 2 mM DTT, and 10% glycerol and stored at −80 °C.

**Amino Acid Specificity Assay of the LnmQ Protein**—Amino acid-dependent ATP-sodium pyrophosphate (PPi) assays were performed according to literature procedures (18, 23–25). A typical 100-μl assay reaction contained 125 mM LNMQ, 3 mM ATP, 0.1 mM PPi with 0.5 μCi of 32PPi (40.02 Ci/mmol; PerkinElmer Life Sciences), 5 mM MgCl2, 0.1 mM EDTA, and 0.5 mM various amino acids in 75 mM Tris-HCl (pH 7.5) buffer. After 30 min of incubation at 30 °C, the assays were stopped by the addition of 0.5 ml of 1% (w/v) activated charcoal in 4.5% (w/v) tetrasodium pyrophosphate and 3.5% (v/v) perchloric acid. The precipitates were collected on glass fiber filters (2.4 cm, G-4, Fisher), washed successively with 10 ml of 40 mM sodium pyrophosphate plus 1.4% perchloric acid, 10 ml of
allowed to proceed at 30 °C for another 2–10 min. The reactions were quenched with 0.9 ml of acetone at −80 °C for 2 h; the precipitated proteins were collected by centrifugation at 14,000 rpm, 4 °C for 20 min, and the protein pellets were redissolved and loaded on a 4–20% SDS-polyacrylamide gel. After being stained with Coomassie Blue R-250, destained, and dried, the gels were visualized by a Phosphor Imager (low energy screen, GE Healthcare).

**Identification of the Reaction Products of LnmQ and LnmP by HPLC and ESI-MS Analysis**—After complete phosphopantetheinylation of the apoLnmP (20 μM) as described as above, 0.5 μM LnmQ and 60 μM D-Ala or 200 μM Gly was added into the reaction, followed by 3 mM ATP to initiate the aminoaacylation reaction, which was allowed to proceed at 30 °C for another 2–5 min. The reactions were analyzed by injecting a portion into an HPLC equipped with a Jupiter C18 reverse phase column (5 μm, 300 Å, 250 × 4.6 mm) and chromatographing with a linear gradient of 40–75% acetonitrile in water (containing 0.1% trifluoroacetic acid) in 25 min at a flow rate of 1 ml/min with monitoring by a UV detector at 215 nm. Peaks were collected, lyophilized, redissolved in H2O, and subjected to ESI-MS analysis (Agilent 1000 HPLC-MSD SL instrument, Palo Alto, CA).

**Determination of the Kinetic Parameters of LnmP Aminoaacylation by LnmQ**—To determine the kinetic parameters for holo-LnmP aminoaacylation by LnmQ, reactions were carried out at 30 °C in total volume of 100 μl that contained 75 mM Tris-HCl (pH 7.5), 10 mM MgCl2, 5 mM DTT, 3 mM ATP, 5 nM LnmQ, 1 mM PPi, with 0.8 μCi of 32PPi (12.1Ci/mmol; PerkinElmer Life Sciences), and increasing concentrations of D-Ala (0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 5.0, 7.5, 10, and 15 mM) for 5 min. For Gly, the concentrations were 2, 5, 10, 20, 30, 40, 60, 80, 100, 120, 150, 200, 300, and 400 mM, and the reaction time was 10 min. The initial rate conditions were maintained by computing the substrate amino acid conversion at less than 10% conversion, on the assumption that these conditions were within the linear range of enzyme concentration and enzyme turnover. The data were fitted with the Michaelis-Menten equation to extract the kinetic constants.

**In Vitro Aminoaacylation Assays of PCP (LnmP)**—The **in vitro** phosphopantetheinylation and aminoaacylation of LnmP were carried out according to literature methods (17, 23–25). A typical reaction (100 μl) contained 75 mM Tris (pH 7.5), 10 mM MgCl2, 5 mM DTT, 0.3 mM CoA, 10 μM LnmP, and 100 mM Svp, a known promiscuous phosphopantetheinyl transferase (23), and was incubated for 30 min at 30 °C to allow complete phosphopantetheinylation of the apoLnmP. Then 0.2 μM LnmQ and 1.5 μCi of 3-[14]CAla or 14C Gly (55 mCi/mmOL, ICN Biomedical Inc., Irvine, CA) were added into the reaction, followed by 3 mM ATP to initiate the aminoaacylation reaction, which was with a linear gradient of 40–75% acetonitrile in water (containing 0.1% trifluoroacetic acid) in 25 min at a flow rate of 1 ml/min with monitoring by a UV detector at 215 nm. Peaks were collected, lyophilized, redissolved in H2O, and subjected to ESI-MS analysis (Agilent 1000 HPLC-MSD SL instrument, Palo Alto, CA).

**LNM Production, Isolation, and Analysis**—LNM production and isolation from wild type S. atroolivaceus and HPLC analysis were carried out as reported previously (18–21), the only difference was feeding D-Ala to the fermentation broth. Seed culture (5 ml) was inoculated into the fermentation medium (50 ml), which was allowed to grow at 28 °C, 300 rpm for 10–15 h, and then different concentrations of D-Ala (5, 10, 25 mM) were added separately. After 3–5 h, moisturized Diaion HP-20 resin...
was added to 5% (w/v), and the fermentation was continued for up to 48 h (18–21).

RESULTS

Characterization of LnmQ as an NRPS A Domain with a Substrate Specificity for D-Ala—According to the LNM structure (Fig. 1A), the widely accepted bacterial NRPS/PKS paradigm (4, 26), sequencing analysis, and the result of in vivo gene replacement (20), we have proposed that LnmQ serves as the NRPS A domain of the LNM biosynthetic initiation (or loading) module, to activate D-Ala as an aminoacyl-AMP and transfer it onto LnmP, the PCP partner of LnmQ (Fig. 1A). To validate this hypothesis, we amplified the lnmQ gene by PCR and cloned it into the T7 overexpression vector pET37b. This construction allows the production of LnmQ as a C-terminal His8-tagged protein. LnmQ was overproduced in E. coli BL21(DE3) and purified by nickel affinity chromatography to homogeneity, as visualized by SDS-PAGE and Coomassie Blue staining (Fig. 1B).

The yield of purified fusion protein was ~20 mg/liter. Substrate-dependent ATP-32PPi exchange assays showed that, among D-Ala, β-Ala, and the 20 proteinogenic L-amino acids tested, D-Ala is the preferred substrate of LnmQ (Fig. 2A). LnmQ also demonstrated activity, albeit much weaker (11% relatively to D-Ala under the specific assay condition), toward Gly. In addition, kinetic parameters for D-Ala and Gly as substrates of LnmQ (Fig. 2, B and C, and Table 1) were comparable with other A domains of NRPSs (27–30). The $K_m$ value for D-Ala is 1.62 mM, which is 22 times lower than that of Gly; and the catalytic efficiency of D-Ala is 264 mM$^{-1}$min$^{-1}$, which is about 50 times higher than that of Gly. These results support D-Ala as the natural substrate of LnmQ.

Characterization of LnmP as an NRPS PCP—The lnmP gene, encoding a discrete PCP, is located upstream of lnmQ and is transcriptionally coupled with the lnmQ open reading frame (Fig. 1A). LnmP contains a modestly conserved signature motif of GX(HD)S, in which the Ser residue (underlined) can be posttranslationally modified by the covalent attachment of the 4’PP group (20, 25). To test LnmP as the PCP partner of LnmQ, we expressed this small gene using the pET28a vector in E. coli and purified the resultant PCP as an N-His$_6$ and C-His$_6$-tagged fusion protein to homogeneity (Fig. 1C), with a yield of ~50 mg/liter (19). To our surprise, two distinguishable protein bands were visualized on SDS-PAGE. HPLC and ESI-MS analysis (Fig. 3C, panel I, and Table 2) confirmed that one band was the apo-form, and the other band was the holo-PCP that has a 4’PP group added by E. coli ACPS phosphopantetheinyl transferase (24, 25). By HPLC analysis, the holo-form represented ~40% of the total PCP after 18 h of induced expression at 18 °C (Fig. 3C, panel I). When LnmP was incubated with CoA and the Svp phosphopantetheinyl transferase (25), the apoPCP was completely converted into the holo-form (Fig. 3C, panel II).

Aminoacylation of LnmP by LnmQ—We previously proposed that LnmQ and LnmP constitute the initiation module that is responsible for chain initiation of LNM biosynthesis by incorporating a D-Ala moiety (Fig. 1A) (19, 20). To verify whether holo-LnmP can be aminoacylated by LnmQ, we incubated the holo-LnmP and LnmQ with D-$[^{14}$C]Ala or $[^{14}$C]Gly and ATP to directly test amino acid loading (Fig. 3A). Reaction mixtures were subjected to SDS-PAGE and phosphorimaging to detect specific loading of the $[^{14}$C]-amino-acid onto the 4’PP group of PCP (Fig. 3B). Reaction mixtures with nonradioactive substrates were subjected to HPLC (Fig. 3C), and the resolved apoPCP, holo-PCP, and aminoacyl-S-PCP species were collected and subjected to ESI-MS analysis to confirm the predicted modifications (Table 2). The collective results showed that both D-Ala and Gly can be effectively activated by LnmQ as
kinetic parameters of holo-LnmP aminoacylation by LnmQ were subsequently determined (Table 3). The $K_m$ values for holo-LnmP as determined with fixed concentrations of D-Ala (at 2 mM) or Gly (at 40 mM) are about 6.2 and 8.0 $\mu$M, respectively, whereas the $k_{cat}$ values as determined with D-Ala and Gly are similar, 99 and 39 min$^{-1}$, respectively. These values are comparable with those of the monofunctional A and PCP pair converting l-proline to pyrrolyl-2-carboxyl-S-PCP during undecylprodigiosin and pyoluteorin biosynthesis (30); these constants are also comparable with those observed in typical NRPS systems where the aminoacyl-AMP ligase must aminoacylate the cognate holo-PCP in trans (31, 32).

Confirming the Protein-Protein Interaction between LnmQ and LnmP—The results above confirmed that LnmQ, a monofunctional A domain, activates D-Ala and transfers it to the LnmP PCP. Next we asked whether there is a specific protein-protein interaction between this discrete A:PCP pair, and if so, does the aminoacylation of LnmP depend upon this specific communication. To address these questions, we first chose to study the natural A-PCP didomain of Lnlm and used l-Cys as substrate for an orthogonal experiment. The LnmI A-PCP was originally used to show specific loading of l-Cys to the PCP (18). In the present work the LnmI-A and LnmI-PCP domains were expressed as individual mono-domain proteins then used to establish that LnmI-A can still specifically activate l-Cys and load it onto the 4'PP group of LnmI-PCP (data not shown). In contrast, neither the LnmQ (for D-Ala):LnmI-PCP pair nor the LnmI-A (for l-Cys):LnmP pair could achieve effective aminoacylation. These results suggest that LnmQ and LnmP interact with each other specifically, to serve as the initiation module for LNM biosynthesis.

For a pair of NRPS A and PCP proteins, there are two possible mechanisms of aminoacylation in trans. One mechanism results from specific communication between the A and PCP domains; alternatively, aminoacylation could happen if dissociation of the aminoacyl-AMP is followed by adventitious thiolation by the free-standing PCP in solution (Fig. 4A). To address this issue, we designed the following experiment. LnmQ, with its substrates D-Ala and ATP together in assay buffer, were put into a dialysis tube with a molecular cutoff of 8 kDa. This dialysis tube was then submerged into the same assay buffer containing holo-LnmP (Fig. 4B). If the aminoacylation

**TABLE 2**

ESI-MS analysis of apo-, holo-, and aminoacyl-S-LnmP

| LnmP                  | [M + H]$^+$ calculated | [M + H]$^+$ found |
|-----------------------|------------------------|-------------------|
| apo-LnmP              | 12,335                 | 12,333            |
| holo-LnmP             | 12,675                 | 12,673            |
| d-Alaninyl-S-LnmP     | 12,746                 | 12,744            |
| Glycinyl-S-LnmP       | 12,732                 | 12,730            |

**TABLE 3**

Apparent Michaelis-Menten constants for LnmQ as an aminoacylation enzyme loading selected amino acids to LnmP as determined at fixed concentration of D-Ala or Gly

| Substrate               | $K_m$ (mM) | $k_{cat}$ (s$^{-1}$) | $k_{cat}/K_m$ (s$^{-1}$ mM$^{-1}$) | Relative efficiency |
|-------------------------|------------|----------------------|-----------------------------------|---------------------|
| holo-LnmP at 2 mM D-Ala | 6.2 ± 0.3  | 98.9 ± 5.2           | 16                                | 1.0                 |
| holo-LnmP at 40 mM Gly  | 8.0 ± 0.8  | 39.0 ± 7.2           | 4.9                               | 0.3                 |

**FIGURE 3. In vitro aminoacylation of holo-LnmP by LnmQ.** A, chemical reaction mechanism. B, SDS-polyacrylamide analysis on a 4–20% gradient gel (panel I) and its autoradiogram (panel II); lane 1, molecular weight markers; lanes 2 and 5, complete assay with no ATP; lanes 3 and 6, complete assay with no Svp; and lanes 4 and 7, complete assay; lanes 2–4, [14C]D-Ala as a substrate; lanes 5–7, [14C]Gly as a substrate. B, HPLC analysis: panel I, purified LnmP; panel II, holo-PCP modified by Svp; panel III, aminoacylation products with D-Ala; or panel IV, Gly, panel IV, as a substrate, respectively. C, apoLnmP; D, holo-LnmP; 224, LnmQ; E, Svp; F, d-alaninyl-S-LnmP; G, glycyl-S-LnmP.

aminoacyl-AMP and transferred onto holo-LnmP to form the D-Ala or Gly aminoacyl-S-PCP intermediate, and that formation of aminoacyl-S-LnmP is dependent on ATP (Fig. 3B, lanes 2 and 5) and holo-LnmP (Fig. 3B, lanes 3, 4, 6, and 7). The
between LnmQ and LnmP depends upon specific communication, the outside holo-PCP should not be aminoacylated, because the dialysis tube prevents direct protein-protein interaction; in contrast, if the aminoacylation results from the dissociation of the aminoacyl-AMP followed by adventitious thiolation by the free-standing PCP in solution, the outside holo-PCP could be aminoacylated, because the D-alaninyl-AMP catalyzed by LnmQ can dissociate and permeate through the dialysis tube. By HPLC analysis of the holo-PCP mixture (Fig. 4C), no aminoacylated product, D-alaninyl-S-LnmP, was formed, even when reaction was allowed to last for up to 1 h; once the dialysis tube was broken, nearly all holo-PCP was converted into aminoacylated D-alaninyl-S-LnmPs in 2 min. These results suggest that there is specific communication between the discrete A and PCP, which contributes to the aminoacylation in trans.

Improving LNM Production by in Vivo Feeding of d-Ala—One of the well characterized examples of d-Ala incorporation in nonribosomal peptide natural products is HC toxin, a secondary metabolite produced by a filamentous fungi (9). In this case, an NRPS A domain directly activates d-Ala, which is produced by a dedicated Ala racemase. Disruption of the Ala racemase gene abolished the production of d-Ala-containing isoforms of HC toxin. Supplementation of exogenous d-Ala to the growth medium restored the normal HC toxin production profile of the Ala racemase null mutant. Higher d-Ala concentrations resulted in higher levels of the d-Ala-containing isoforms of HC toxin (9). This knowledge inspired us to ask whether LNM production in S. atroolivaceus S-140 can be improved by feeding additional d-Ala in the fermentation medium. Indeed, when the fermentation medium was supplemented with different concentrations of d-Ala, LNM production was significantly improved. Supplementation of 25 mM d-Ala resulted in 2–3-fold increase in LNM production, compared with S. atroolivaceus S-140 fermentation under the identical condition without d-Ala supplementation, as analyzed by HPLC.

DISCUSSION

LNM, an antitumor antibiotic, represents a new class of microbial hybrid nonribosomal peptide/polyketide scaffold featuring a macrolactam with a spiro-linked 1,3-dioxo-1,2-dithiolane 5-membered ring structure that has not been found in any other natural products. Gene cloning and sequencing analysis, plus genetic and biochemical characterization work of the LNM biosynthetic gene cluster, have revealed not only a hybrid NRPS/PKS with a novel AT-less type I PKS and several other unprecedented features (18–21) but also a novel type of NRPS initiation (or loading) module with the following features. (i) It is composed of discrete A (LnmQ) and PCP (LnmP) proteins (Fig. 1). (ii) LnmQ directly selects and activates d-Ala to form D-alaninyl-AMP, which is the first characterized A enzyme with d-Ala as a substrate in prokaryotic NRPSs (Fig. 2). (iii) The d-Ala transfer in trans from LnmQ onto LnmP to generate the D-alaninyl-S-PCP intermediate depends upon spe-
specific communication between the discrete A and PCP proteins (Figs. 3 and 4).

NRPSs assemble their peptide products via a thiotemplate process (1, 2, 4, 5), in which A domains activate amino acid substrates as the aminoacyl-AMP before their transfer onto the –SH group of 4\(^{\text{H11032}}\)PP on adjacent PCP domains. Subsequently, C domains catalyze the formation of peptide bonds between the upstream peptidyl-S-PCP and the free amino group of the downstream aminoacyl-S-PCP. An archetypical NRPS contains a didomain (A-PCP) module for chain initiation with L-amino acid and repeated modules of C-A-PCP for chain elongation (Fig. 5A, panel I), as seen in the biosynthetic systems for vanco-

![Diagram of NRPS assembly](image)

**FIGURE 5. Summary of representative chain initiation mechanisms of NRPS.** A, classification of NRPS initiation according to the nature of the starter unit as follows: panel I, typical module with L-amino acids (L-amino acids) as start units; panel II, typical module with D-amino acids (D-amino acids) as a start unit (with an additional E domain epimerizing the L-amino acids into the D-enantiomers once they are installed as aminoacyl-S-PCP intermediates by the A domain; panel III, special module for N-acylated amino acid starter units (with an additional N-terminal C\(^{\text{C1032}}\) domain responsible for the N-acylation); panel IV, a special module for N-formylated amino acid starter units (with an additional formylation (F) domain for the N-formylation); panel V, special module for aromatic acids instead of amino acids (with the A domains are acyl-CoA ligase-like domains that are also known as acyl-CoA ligase-like domain and PCP domains also known as aryl-carrier proteins); and panel VI, the special module as exemplified by LnmQ and LnmP here with D-Ala as a start unit. Ox, oxidization; C or C\(^{\text{C1032}}\), condensation; Cy, cyclization. B, classification of NRPS initiation according to the organization of the start (or loading) module as follows: panel I, the starter module fused together with the extension modules as the N-terminal part of a large NRPS protein (1); panel II, the starter module existed as a separated protein from the extension modules (46); panel III, only the A domain of the starter module exists as a discrete protein with the PCP domain of start module fused with the extension modules as the N-terminal part of the NRPS protein (38); panel IV, the starter module composed of discrete A and ICL-PCP didomain proteins (42); and the starter module composed of discrete, monofunctional A and PCP proteins for panel V an aromatic acid (43), and panel VI for a D-Ala as exemplified by LnmQ and LnmP reported here. Sal, salicylic acid; Dhb, 2,3-dihydroxybenzoic acid; Hqa, 3-hydroxyquinaldic acid.
mycin family antibiotics (8, 33, 34). A variation of an initiation module, organized as A-PCP-E, where the E domain epimerizes the L-amino acid into D-configuration once they are installed as aminoacyl-S-PCP intermediates, represents an initiation module for D-amino acid incorporation (Fig. 5A, panel II), as exemplified by tyrocidine (28). For biosynthesis of more complex moieties, such as the N-acylated amino acid starter units of surfactin (35) or daptomycin (11), the initiation module starts with an additional N-terminal condensation (C') domain (Fig. 5A, panel III). Similarly for the N-formylated amino acid starters, the initiation module starts with an additional formylation (F) domain (Fig. 5A, panel IV) as has been recently characterized for gramicidin (36). Additionally, when the starter unit is something other than an amino acid, such as the aromatic acids in mycobactin (37), yersiniabactin (38), myxochelin (39), pyochelin (40), vibriobactin (41), bacillibactin (42), enterobactin (31), and thiocoraline (43) or amine group in bleomycin (44) or tallysmycin (45), the starter unit is activated by an acyl-CoA ligase-like domain with the consumption of ATP, a process mechanistically similar to NRPS A catalysis (therefore these acyl-CoA ligase-like domains sometimes have also been named A domains) (Fig. 5A, panel V).

NRPS initiation modules can also be classified by the organizational patterns of domains. Besides a typical initiation module fused together with the domain(s) of an elongation module (Fig. 5B, panel I), as seen in the bacitracin (1) and vancomycin (8) biosynthetic systems, the initiation module can also exist as a discrete protein, separated from the elongation module (Fig. 5B, panel II), as exemplified by tyrocidine (28) and gramicidin S (46). These two types of NRPS initiation modules are commonly observed in type A (linear) NRPS system (2), in which the sequence of the resulting linear peptide chain is co-linear with the number and order of the modules. A third type initiation module is known, in which only the A domain exists as a discrete protein, the PCP domain of which is fused with elongation modules. This rare case has been found in yersiniabactin (38) or pyochelin (40) biosynthesis (Fig. 5B, panel III). Another type of initiation module, as seen in enterobactin (31), myxochelin (39), vibriobactin (41), and bacillibactin biosynthetic systems (42), has a discrete A domain and a bifunctional protein with a PCP domain fused with an isochorismate lyase (ICL) domain (Fig. 5B, panel IV). Finally, an initiation module consisting of a discrete acyl-CoA ligase-like domain and PCP (Fig. 5B, panel V) has also been proposed for the marine natural producy of thiodepsipeptide thiocoraline; the 3-hydroxyquinoline acid starter is first activated as an acyl-AMP and subsequently loaded to the PCP to initiate its biosynthesis (43). A variation of this organization has also been reported for the quinoxa-line antibiotics whose initiation module consisted of a discrete A enzyme and an acyl carrier protein (47, 48). These three types use aromatic acids as starter units; the discrete A proteins are also called acyl-AMP ligases and the PCP domains also as aryl-carrier proteins (49). They are only found in type B (iterative) or type C (nonlinear) NRPS systems (2), in which the modules or domains can be used more than once in the assembly of a single product, resulting in repeated sequences in peptide chains or unusual internal cyclization or branch point syntheses.

A sixth type of NRPS initiation module is described here from the LNM biosynthetic system. It uses an NRPS initiation module comprised of two discrete monofunctional proteins: LnmQ, acting as A domain for activation of D-Ala to form D-alanyl-AMP, and LnmP, as the cognate PCP partner of LnmQ (Fig. 5, A, panel VI, and B, panel VI). This type of A:PCP pair had been postulated as a strategy to activate and sequester amino acids as enzyme-bound aminoacyl thioesters for specific modification (30, 50, 51). In the biosynthesis of pyrrole moieties in pyoluteorin, undecylprodigiosin and coumermycin A1, an A domain activates L-proline as L-prolyl-AMP and transfers the L-prolyl moiety to the –SH group of 4’PP of the respective holo-PCP partner, and then the L-prolyl-S-PCP acts as substrate for further desaturation and oxidation to form pyrrole-2-carboxyl-S-PCP, which serves as the activated monomer for incorporation into the final product (30). Another example is barbamide, a hybrid peptide-polyketide from a marine source (51), in which the A domain activates L-leucine as L-leucinyl-AMP and transfers it to the holo-PCP partner. The L-leucinyl-S-PCP then acts as the substrate for chlorination to form trichloroleucinyl-S-PCP, which is incorporated into the final product. In either case, the A:PCP pair seems to have evolved from a typical NRPS assembly line to serve a distinct purpose, i.e. not to catalyze peptide bond formation but to create a pool of covalently tethered amino acids dedicated to subsequent modification (50).

In LNM biosynthesis, the unique A:PCP pair could have two functions as follows: to act as an initiation module for oligopeptide biosynthesis or to divert a portion of the cellular D-Ala pool into secondary metabolism. This notion is supported by the following three observations. (i) The $K_{m}$ value of D-Ala activation by the A enzyme (LnmQ) is about 1.6 mM; once D-Ala is activated as D-alanyl-AMP form, the $K_{m}$ value of holo-PCP (LnmP) aminoacylation by A domain is about 6 μM. The free-standing, monofunctional D-alanyl-S-PCP thus may balance kinetic sequestration and thermodynamic activation. (ii) In the other two systems directly using D-Ala as an NRPS substrate for the synthesis of the fungal products cyclosporin (9) and HC toxin (10), D-Ala is provided by pathway-specific Ala racemases (15, 16). In the LNM biosynthetic gene cluster, however, no pathway-specific Ala racemase gene is found, indicating that the LNM pathway may compete with the primary bacterial cell wall metabolism for a source of D-Ala. (iii) LNM production can be improved by supplementation of exogenous D-Ala in the fermentation broth.

In most NRPS systems, the A domain activates an amino acid substrate as its aminoacyl-AMP and then transfers the aminoacyl moiety to the cognate PCP in cis. When the A and PCP domains are artificially separated, in trans aminoacylation between the A domain and its cognate PCP also can work effectively, yet it has been reported that aminoacylation in trans failed with its cognate PCP but succeeded with heterologous PCP protein fragments (29). In addition, a free standing PCP of the bleomycin biosynthetic pathway has no apparent cognate A domain partner but can be successfully aminoacylated in trans (24). These observations raise the question whether paired A:PCP proteins communicate with
onto LnmI-PCP but not LnmP suggest that the in step of NRPS biosynthesis largely remain at the experiments support this conclusion.

acylation is specific in this case. The results of the dialysis tube experiments support this conclusion.

Manipulation of chain initiation in PKS assembly lines has proven very useful for generating novel natural products (52, 53); in contrast, attempts at engineering the chain initiation step of NRPS biosynthesis largely remain at the in vitro biochemical level (54). The A domain is the critical fidelity-controlling unit for amino acid monomer selection, activation, and transfer to the paired PCP; at the same time, the C domain shows low selectivity toward the upstream donor residue and higher selectivity toward the downstream receptor residue (55). Thus, the chain initiation module should be the most promising candidate for engineering production of novel metabolites because it is the only module without a C domain in the NRPS assembly line. Results reported herein underscore the flexibility and versatility of NRPS initiation module both structurally and mechanistically for natural product biosynthesis and provide another opportunity for engineered biosynthesis of novel LNM analogs by manipulating the initiation module.

Acknowledgments—We thank Kyowa Hakko Kogyo Co., Ltd., Tokyo, Japan, for the wild-type S. atroclavicus S-140 strain, the Analytic Instrumentation Center of the School of Pharmacy, University of Wisconsin, Madison, for support in obtaining MS data, and Yong Huang for assistance in preparing the manuscript.

REFERENCES
1. Sieber, S. A., and Marahiel, M. A. (2005) Chem. Rev. 105, 715–738
2. Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2002) Chem. Biol. Chem. 3, 490–540
3. Staunton, J., and Weissman, K. J. (2001) Nat. Prod. Rep. 18, 380–416
4. Fischbach, M. A., and Walsh, C. T. (2005) Chem. Rev. 106, 3468–3496
5. Doekel, S., and Marahiel, M. A. (2000) Chem. Biol. 7, 373–384
6. Billman-Jacobe, H., McConville, M. J., Haites, R. E., Kovacevic, S., and Cop pel, R. L. (1999) Mol. Microbiol. 33, 1244–1253
7. Duitman, E. H., Hama en, L. W., Rembold, M., Venema, G., Seitz, H., Saenger, W., Bernhard, F., Reinhardt, R., Schmidt, M., Ullrich, C., Stein, T., Leenders, F., and Vater, J. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 13294–13299
8. Hubbard, B. K., and Walsh, C. T. (2003) Angew. Chem. Int. Ed. 42, 731–765
9. Scott-Craig, J. S., Panaccione, D. G., Poccard, J. A., and Walton, J. D. (1992) J. Biol. Chem. 267, 26044–26049
10. Weber, G., Schöndorf er, K., Schneider-Scherzer, E., and Leitner, E. (1994) Curr. Genet. 26, 120–125
11. Miao, V., Coephett-LeGal, M.-F., Brian, P., Bost, R., Whiting, A., Penn, I., Martin, S., Ford, R., Parr, J., Bouchard, M., Silva, C. I., Wrigley, S. K., and Baltz, R. H. (2005) Microbiology 151, 1507–1523
12. Christiansen, G., Fastner, J., Erhardt, M., Börner, T., and Dittmann, E. (2003) J. Bacteriol. 185, 564–572
13. Xia, Z., and Smith, C. D. (2001) J. Org. Chem. 66, 3459–3466
14. Jiang, W., Wanner, J., Lee, R. J., Bounaud, P.-Y., and Boger, D. L. (2003) J. Am. Chem. Soc. 125, 1877–1887
15. Hoffmann, K., Schneider-Scherzer, E., Kleikauf, H., and Zecher, R. (1994) J. Biol. Chem. 269, 12710–12714
16. Cheng, Y.-Q., and Walton, J. D. (2000) J. Biol. Chem. 275, 4906–4911
17. Miao, V., Takahashi, I., Yoshida, M., Asano, K., Kamimoto, M., and Nakano, H. (1998) J. Antibiot. (Tokyo) 42, 333–335
18. Cheng, Y.-Q., Tang, G.-L., and Shen, B. (2002) J. Bacteriol. 184, 7013–7024
19. Cheng, Y.-Q., Tang, G.-L., and Shen, B. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 3149–3154
20. Tang, G.-L., Cheng, Y.-Q., and Shen, B. (2004) Chem. Biol. 11, 33–45
21. Tang, G.-L., Cheng, Y.-Q., and Shen, B. (2006) J. Nat. Prod. 69, 387–393
22. Schomberg, J., Fritsch, E. F., and Maniatis, T. (2001) Molecular Cloning A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
23. Linne, U., and Marahiel, M. A. (2004) Methods Enzymol. 388, 293–315
24. Du, L., and Shen, B. (1999) Chem. Biol. 6, 507–517
25. Sánchez, C., Du, L., Edwards, D. J., Toney, M. D., and Shen, B. (2001) Chem. Biol. 8, 725–738
26. Du, L., Sánchez, C., and Shen, B. (2001) Metab. Eng. 3, 78–95
27. Chen, H., O’Connor, S., Cane, D. E., and Walsh, C. T. (2001) Chem. Biol. 8, 899–912
28. Mootz, H. D., and Marahiel, M. A. (1997) J. Bacteriol. 179, 6843–6850
29. Ehmann, D. E., Shaw-Reid, C. A., Losey, H. C., and Walsh, C. T. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 2509–2514
30. Thomas, M. G., Burkart, M. D., and Walsh, C. T. (2002) Chem. Biol. 9, 171–184
31. Gehring, A. M., Bradley, K. A., and Walsh, C. T. (1997) Biochemistry 36, 8495–8503
32. Keating, T. A., Bradley, K. A., and Walsh, C. T. (2000) Biochemistry 39, 15522–15530
33. Chiu, H.-T., Hubbard, B. K., Shah, A. N., Eide, J., Freedburg, R. A., Walsh, C. T., and Khosla, C. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 8548–8553
34. Pootoolal, J., Thomas, M. G., Marshall, C. G., Neu, J. M., Hubbard, B. K., Walsh, C. T., and Wright, G. D. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 8962–8967
35. Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., and van Sinderen, D. (1993) Mol. Microbiol. 8, 821–831
36. Schoenafinger, G., Schracker, N., Linne, U., and Marahiel, M. A. (2006) J. Am. Chem. Soc. 128, 7406–7407
37. Quadri, L. E., Sello, J., Keating, T. A., Weinreb, P. H., and Walsh, C. T. (1998) Chem. Biol. 5, 631–645
38. Miller, D. A., Luo, S., Hillson, N., Keating, T. A., and Walsh, C. T. (2002) Chem. Biol. 9, 333–344
39. Gattatzis, N., Kunze, B., and Müller, R. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 11136–11141
40. Patel, H. M., and Walsh, C. T. (2001) Biochemistry 40, 9023–9031
41. Keating, T. A., Marshall, C. G., and Walsh, C. T. (2000) Biochemistry 39, 15513–15530
42. May, J. I., Wendrich, T. M., and Marahiel, M. A. (2001) J. Biol. Chem. 276, 7209–7217
43. Lombo, F., Velasco, A., Castro, A., de la Calle, F., Brana, A. F., Sanchez-Puelles, J. M., Mendez, C., and Salas, J. A. (2006) Chem. Biol. 13, 366–376
44. Du, L., Sánchez, C., Chen, M., Edwards, D. J., and Shen, B. (2000) Chem. Biol. 7, 623–642
45. Tao, M., Wang, L., Wendl-Pienkoski, E., George, N. P., Galm, U., Zhang, G., Coughlin, J. M., and Shen, B. (2007) Mol. Biol. Syst. 3, 60–74
46. Turug, K., Krause, M., and Marahiel, M. A. (1992) Mol. Microbiol. 6, 529–546
47. Schmoeck, G., Pfennig, F., Jewiarcz, J., Schlunbomh, W., Laubinger, W., Schauwecker, F., and Keller, U. (2005) J. Biol. Chem. 280, 4339–4349
Chain Initiation for Leinamycin Biosynthesis

48. Watanabe, K., Hotta, K., Praseuth, A. P., Koketsu, K., Migita, A., Boddy, C. N., Wang, C. C., Oguri, H., and Oikawa, H. (2006) *Nat. Chem. Biol.* 2, 423–428
49. Rusnak, F., Stephen Faraci, W., and Walsh, C. T. (1989) *Biochemistry* 28, 6827–6835
50. Chen, H., Thomas, M. G., O’Connor, S. E., Hubbard, B. K., Burkart, M. D., and Walsh, C. T. (2001) *Biochemistry* 40, 11651–11659
51. Chang, Z., Flatt, P., Gerwick, W. H., Nguyen, V.-A., Willis, C. L., and Sherman, D. H. (2002) *Gene (Amst.)* 296, 235–247
52. Jacobsen, J. R., Hutchinson, C. R., Cane, D. E., and Khosla, C. (1997) *Science* 277, 367–369
53. Marsden, A. F. A., Wilkinson, B., Cortés, J., Dunster, N. J., Staunton, J., and Leadlay, P. F. (1998) *Science* 279, 199–202
54. Walsh, C. T. (2002) *Chem. Biol. Chem.* 3, 124–134
55. Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) *Science* 284, 486–489