ORGANIC LETTERS

2013
Vol. 15, No. 14
3674–3677

Naphthyridinomycin Biosynthesis Revealing the Use of Leader Peptide to Guide Nonribosomal Peptide Assembly

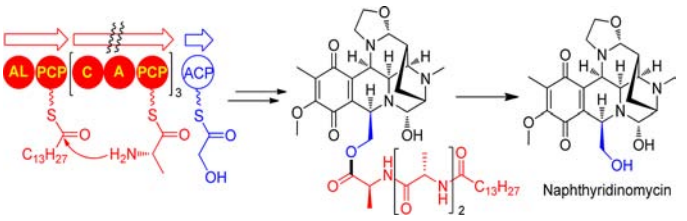
Jin-Yue Pu,[#,†] Chao Peng,[#,†] Man-Cheng Tang,[†] Yue Zhang,[†] Jian-Ping Guo,[†] Li-Qiang Song,[†] Qiang Hua,[‡] and Gong-Li Tang*[,†]

State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Road, Shanghai, 200032, China, and State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai, China

gltang@sioc.ac.cn

Received June 2, 2013

ABSTRACT



Analysis of naphthyridinomycin gene cluster revealed that this antibiotic is generated by nonribosomal peptide synthetase (NRPS) machinery. However, four modules encoded by two genes do not correspond with the structural units in the final product. Genetic and biochemical characterization of the gene cluster suggested that the leader peptide mechanism for the NRPS assembly line was involved in biosynthesis of this tetrahydroisoquinoline alkaloid.

Nonribosomal peptide synthetases (NRPSs) catalyze the biosynthesis of a vast family of natural products including many clinical drugs.[1] They usually use a multienzyme thiotemplate assembly line for the peptide generation and bear a condensation-adenylation-peptidyl carrier protein (C-A-PCP) module organization for one elongation cycle.[2] The substrate selection of A domain in each module, the variation of domains within the modules, and the order of NRPS modules dictate the structure of the final products, which provides the molecular basis for the combinatorial biosynthesis.[2,3]

Naphthyridinomycin (NDM, Figure 1B), first isolated from *Streptomyces lusitanus* (NRRL8034), exhibits a main carbon frame of most six condensed rings, including four six-membered and a five-membered bridged ring system.[4] It belongs to the family of tetrahydroisoquinoline alkaloids incuding safracin (SAC) and saframycin (SFM) (Figure S1A, Supporting Information), and shows strong antitumor and antimicrobial activities.[4c] Extraordinarily, its biological activity against methicillin-resistant *Staphylococcus aureus* has spurred considerable interests.[4c,5] Previously feeding experiments suggested that Tyr, Met, Gly or Ser, and ornithine (Orn) could be used as the precursors for the biosynthesis of core frame;[6] we recently demonstrated that a hydroxyethyl two-carbon (C_2) unit from ketose was incorporated into NDM as a extender unit of NRPS.[7] In this study, we focus on analyzing and characterizing the genetic basis of NDM biosynthesis, which not only provides another example of noncolinear

[†] State Key Laboratory of Bioorganic and Natural Products Chemistry
[‡] State Key Laboratory of Bioreactor Engineering
[#] J.-Y. Pu and C. Peng contributed equally.

(1) Felnagle, E. A.; Jackson, E. E.; Chan, Y. A.; Podevels, A. M.; Berti, A. D.; McMahon, M. D.; Thomas, M. G. *Mol. Pharmaceutics* **2008**, *5*, 191.
(2) (a) Sieber, S. A. M.; Marahiel, M. A. *Chem. Rev.* **2005**, *105*, 715. (b) Condurso, H. L.; Bruner, S. D. *Nat. Prod. Rep.* **2012**, *29*, 1099.
(3) Fischbach, M. A.; Walsh, C. T. *Chem. Rev.* **2006**, *106*, 3468.
(4) (a) Kluepfel, D.; Baker, H. A.; Piattoni, G.; Sehgal, S. N.; Sidorowicz, A.; Singh, K.; Vezina, C. *J. Antibiot.* **1975**, *28*, 497. (b) Siengalewicz, P.; Rinner, U.; Mulzer, J. *Chem. Soc. Rev.* **2008**, *37*, 2676. (c) Scott, J. D.; Williams, R. M. *Chem. Rev.* **2002**, *102*, 1669.
(5) Bernan, V. S.; Montenegro, D. A.; Korshalla, J.; Maiese, W. M.; Steinberg, D. A.; Greenstein, M. *J. Antibiot.* **1994**, *47*, 1417.
(6) Zmijewski, M. J., Jr.; Mikolajczak, M.; Viswanatha, V.; Hruby, V. J. *J. Am. Chem. Soc.* **1982**, *104*, 4969. (b) Zmijewski, M. J., Jr.; Palaniswamy, V. A.; Gould, S. J. *J. Chem. Soc., Chem. Commun.* **1985**, *18*, 1261. (c) Palaniswamy, V. A.; Gould, S. J. *J. Am. Chem. Soc.* **1986**, *108*, 5651.
(7) Peng, C.; Pu, J.-Y.; Song, L.-Q.; Jian, X.-H.; Tang, M.-C.; Tang, G.-L. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 8540.

10.1021/ol401549y
Published on Web 07/10/2013

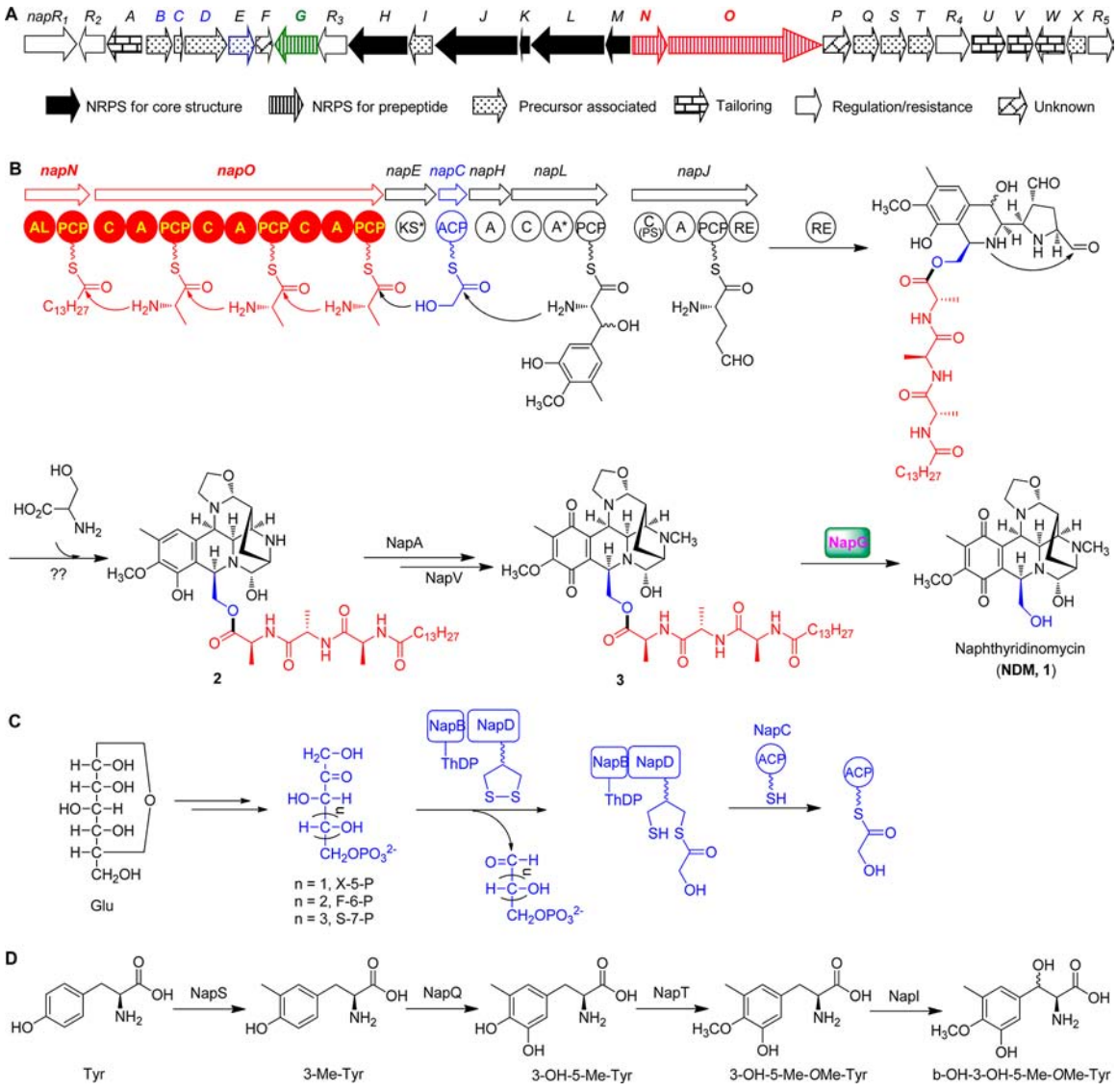


Figure 1. Biosynthetic gene cluster and proposed biosynthetic model of naphthyridinomycin (NDM). (A) Gene organization of NDM cluster. Functional assignments of these genes are summarized in Table S3 (Supporting Information). (B) Hypothesis of the core structure formation. The proposed biosynthetic pathways of precursor: (C) hydroxyacetyl unit and (D) β-hydroxy-3-hydroxy-5-methyl-*O*-methyltyrosine (β-OH-3-OH-5-Me-*O*Me-Tyr).

NRPSs, but also joins the growing family of nonribosomal peptide production guided by the leader peptide mechanism.

We used degenerate primers for the PCR screening of the NRPS system that was postulated to be involved in biosynthesis of NDM based on the known SAC and SFM biosynthetic pathway.[8] Subsequent chromosome walking led to a 92.6 kb DNA region in three overlapping fosmids (Figure S1B, Supporting Information, GenBank under accession No. JQ996389). Bioinformatic analysis and sequential inactivation of ORFs from both ends of this region (Figure 2A, S1C, and Table S3, Supporting Information) led to the assignment of NDM gene cluster to consist of 28 ORFs (Figure 1A), which include nine genes encoding NRPS and associated enzymes (*napG*, *napH*, *napJ*–*napO*, and *napX*), eight genes encoding enzymes for precursor biosynthesis (*napB*–*napE*, *napI*, *napQ*, *napS*, and *napT*), four genes encoding tailoring enzymes (*napA*, *napU*–*napW*), three resistance genes (*napR1*, *R2*, and *R4*), two regulator genes (*napR3* and *R5*) and two genes (*napP* and *napF*) whose function cannot be predicted on the basis of sequence analysis (Table S3, Supporting Information).

The gene cluster contains a three-gene cassette, *napB*–*napC*–*napD*, encoding a two-component transketolase (NapB/NapD) and an acyl carrier protein (ACP, NapC).

(8) (a) Pospiech, A.; Cluzel, B.; Bietenhader, J.; Schupp, T. *Microbiology* **1995**, *141*, 1793. (b) Velasco, A.; Acebo, P.; Gomez, A.; Schleissner, C.; Rodriguez, P.; Aparicio, T.; Conde, S.; Munoz, R.; De La Calle, F.; Luis Garcia, J.; Sanchez-Puelles, J. M. *Mol. Microbiol.* **2005**, *56*, 144. (c) Li, L.; Deng, W.; Song, J.; Ding, W.; Zhao, Q.-F.; Peng, C.; Song, W.-W.; Tang, G.-L.; Liu, W. *J. Bacteriol.* **2008**, *190*, 251.

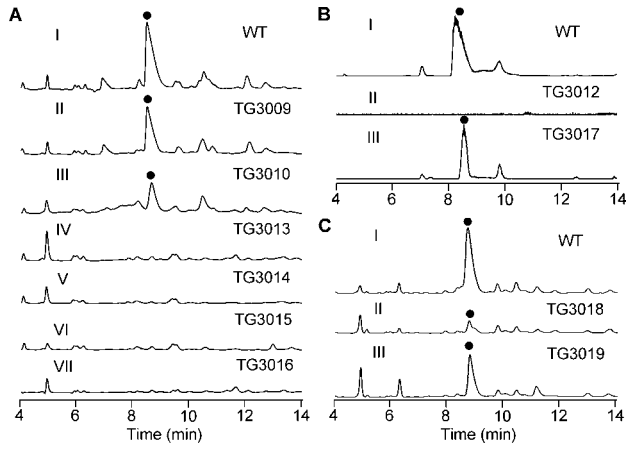


Figure 2. Characterization of the genes involved in NDM biosynthesis in vivo. (A) HPLC analysis (UV absorbance monitored at 270 nm): (I) wild-type *S. lusitanus* NRRL 8034; (II−VII) mutant *S. lusitanus* TG3009 (Δ*orf-1*), TG3010 (Δ*orf+1*), TG3013 (Δ*napO*), TG3014 (Δ*napH*), TG3015 (Δ*napL*), TG3016 (Δ*napJ*). (B) LC−MS analysis: (I) wild-type; (II) mutant *S. lusitanus* TG3012 (Δ*napN*); (III) mutant *S. lusitanus* TG3017 (TG3012 harboring the *napN* expression plasmid pTG3024). (C) HPLC analysis: (I) wild-type; (II) mutant *S. lusitanus* TG3018 (Δ*napG*); (III) mutant *S. lusitanus* TG3019 (TG3018 harboring the *napG* expression plasmid pTG3025). (●) NDM.

It was demonstrated that they are responsible for the generation of a C_2-*S*-ACP extender unit for NRPS from ketose through the lipoyl tethered intermediate (Figure 1C).[7] NapE exhibits sequence homology to 3-oxoacyl-ACP synthase III (KS*) but is characterized by the mutated catalytic triad of C−H−H into C−L−H, so it is supposed to catalyze the formation of a C−O bond rather than the usual C−C bond (Figure 1B). A similar mechanism was first proposed in a nonactin pathway, and biochemically confirmed in tetronate and cervimycin biosynthesis.[9] NapS/Q/T are homologous to SacF/D/G or SfmM2/D/M3, which were biochemically confirmed to be *C*-methyltransferase/peroxidase/*O*-methyltransferase and responsible for the biosynthesis of 3-hydroxy-5-methyl-*O*-methyltyrosine (3-OH-5-Me-*O*Me-Tyr).[10] NapI shows sequence similarity to MppO or VioC, a family of nonheme iron hydroxylase involved in β-hydroxylation of L-enduracididine and Arg,[11] and was subsequently assigned as β-hydroxylase. Thus, the biosynthesis of a Tyr derivative precusor was proposed as follows: Tyr is first methylated to yield 3-Me-Tyr by NapS; then, NapQ catalyzes the hydroxylation of 3-Me-Tyr into 3-OH-5-Me-Tyr; next, 3-OH-5-Me-Tyr is transformed to 3-OH-5-Me-*O*Me-Tyr by NapT; finally, NapI catalyzes the formation of β-OH-3-OH-5-Me-*O*Me-Tyr, which serves as another extender unit for NRPS (Figure 1D). This hypothesis is also consistent with the previous precusor feeding results: Tyr, 3-Me-Tyr, and 3-OH-5-Me-Tyr, but not 3-OH-Tyr, could be incorporated into NDM.[6]

According to the current NRPS machinery and the known precursors, the core scaffold of NDM should be generated by NRPS assembly line using four modules to incorporate the hydroxyacetyl C_2 unit, β-OH-3-OH-5-Me-*O*Me-Tyr, an Orn derivative, and Gly or Ser. However, the gene cluster encodes far more NRPS modules than expected, which makes the biosynthetic pathway complex and unpredictable. NapN bears an acyl-CoA ligase (AL) and a PCP domain, which are also present in SFM-Mx1 and SFM-A biosynthesis.[8] Inspired by Oikawa's finding that the cryptic roles of the fatty acyl chain were necessary for Pictet−Spengler (PS) reactions in SFM-A pathway (Figure S1D, Supporting Information),[12] we propose that this AL-PCP didomain also recongizes a long-chain fatty acid, most likely myristoyl acid (Figure 1B). NapO is a three-module NRPS with domains organized into $(C\text{-}A\text{-}PCP)_3$. Importantly, three modules are highly similar to each other (~90% identify), and the A domains are all predicted to activate Ala with substrate specificity-conferring code of DLFNNALTSK (100% of Ala). Correspondingly, the myristoyl-Ala-Ala-Ala structure is definitely not observed in NDM (Figure 1B); it seems like propeptide or the leader peptide frequently used in ribosomally synthesized and post-translationally modified peptides.[13] This type of leader peptide used in NRPSs has been recently discovered in biosynthesis of xenocoumacin from *Xenorhabdus nematophila*, didemnin from *Tridemnum solidum*, nocardicin from *Nocardia uniformis*, and colibactin from *Escherichia coli*.[14] NapL has domains organized into C-A-PCP, while the A domain is likely inactive for the deficiency of the conserved motifs (Figure S2, Supporting Information).[15]

Meanwhile, NapH, which contains an A domain and a ~300 amino acid unknown N-terminal, is supposed to complement the inactive A domain of NapL. It may select, activate, and load β-OH-3-OH-5-Me-*O*Me-Tyr onto the PCP for peptide chain elongation (Figure 1B). Another

(9) (a) Kwon, H.-J.; Smith, W. C.; Scharon, A. J.; Hwang, S. H.; Kurth, M. J.; Shen, B. *Science* **2002**, *297*, 1327. (b) Sun, Y.; Hahn, F.; Demydchuk, Y.; Chettle, J.; Tosin, M.; Osada, H.; Leadlay, P. F. *Nat. Chem. Biol.* **2010**, *6*, 99. (c) He, H.-Y.; Pan, H.-X.; Wu, L.-F.; Zhang, B.-B.; Chai, H.-B.; Liu, W.; Tang, G.-L. *Chem. Biol.* **2012**, *19*, 1313. (d) Bretschneider, T.; Zocher, G.; Unger, M.; Scherlach, K.; Stehle, T.; Hertweck, C. *Nat. Chem. Biol.* **2012**, *8*, 154.

(10) (a) Fu, C.-Y.; Tang, M.-C.; Peng, C.; Li, L.; He, Y.-L.; Liu, W.; Tang, G.-L. *J. Microbiol. Biotechnol.* **2009**, *19*, 439. (b) Tang, M.-C.; Fu, C.-Y.; Tang, G.-L. *J. Biol. Chem.* **2012**, *287*, 5112.

(11) (a) Haltli, B.; Tan, Y.; Magarvey, N. A.; Wagenaar, M.; Yin, X.; Greenstein, M.; Hucul, J. A.; Zabriskie, T. M. *Chem. Biol.* **2005**, *12*, 1163. (b) Yin, X.; Zabriskie, T. M. *ChemBioChem* **2004**, *5*, 1274. (c) Ju, J.; Ozanick, S. G.; Shen, B.; Thomas, M. G. *ChemBioChem* **2004**, *5*, 1281.

(12) (a) Koketsu, K.; Watanabe, K.; Suda, H.; Oguri, H.; Oikawa, H. *Nat. Chem. Biol.* **2010**, *6*, 408. (b) Koketsu, K.; Minami, A.; Watanabe, K.; Oguri, H.; Oikawa, H. *Curr. Opin. Chem. Biol.* **2012**, *16*, 142.

(13) Oman, T. J.; van der Donk, W. A. *Nat. Chem. Biol.* **2010**, *6*, 9.

(14) (a) Reimer, D.; Pos, K. M.; Thines, M.; Grün, P.; Bode, H. B. *Nat. Chem. Biol.* **2011**, *7*, 888. (b) Xu, Y.; Kersten, R. D.; Nam, S. J.; Lu, L.; Al-Suwailem, A. M.; Zheng, H.; Fenical, W.; Dorrestein, P. C.; Moore, B. S.; Qian, P.-Y. *J. Am. Chem. Soc.* **2012**, *134*, 8625. (c) Davidsen, J. M.; Bartley, D. M.; Townsend, C. A. *J. Am. Chem. Soc.* **2013**, *135*, 1749. (d) Brotherton, C. A.; Balskus, E. P. *J. Am. Chem. Soc.* **2013**, *135*, 3359.

(15) Schwarzer, D.; Finking, R.; Marahiel, M. A. *Nat. Prod. Rep.* **2003**, *20*, 275.

NRPS, NapJ, bearing a domain organization of C-A-PCP-reductase (RE) and showing similarity to SfmC, is then proposed to activate the Orn derivative, possible Glu semialdehyde according to Zmijewski and Gould's proposal,[6b] and carry out multiple enzymatic steps including reductions and PS reactions following the SFM biosynthesis (Figure 1B).[12] However, how Gly or Ser is incorporated and how the five-membered bridged ring is formed still need to be further investigated.

To demonstrate whether these NRPSs are necessary for NDM generation, the NRPS genes were inactivated by gene replacement, and the genotype of the respective mutants were confirmed by PCR analysis (Figure S3, Supporting Information). Each of the resultant mutants, *S. lusitanus* TG3012 (Δ*napN*), TG3013 (Δ*napO*), TG3014 (Δ*napH*), TG3015 (Δ*napL*), and TG3016 (Δ*napJ*) did not produce NDM any more (Figure 2A). Moreover, the gene complementation of *napN* in *S. lusitanus* TG3012 (Δ*napN*) mutant restored the biosynthesis of NDM (Figure 2B). Thus, the results unambiguously confirmed that all these NRPSs are essential to the biosynthesis of NDM.

To characterize the function of the starter module, we first tried to express the AL domain and the PCP domain independently in *E. coli*. Unfortunately, the purified AL domain did not show any activity in vitro. Then the NapN (AL-PCP didomain) was overexpressed, purified, and used to characterize the acylation activity of the AL domain. As expected, a new PCP derivative could be detected by HPLC and LC−MS analysis; however, the production level is very low (data not shown). We speculated that the AL domain prefers to acylate the PCP domain of NapN (in-cis) rather than the PCP domain standing free in solution (in-trans). To overcome this problem, NapN S598A was constructed and used in the same assay (Figure 3A and S4, Supporting Information), in which the active site of the PCP domain of NapN, Ser-598, was mutated to Ala. As shown in Figure 3B, the myristoyl acid could be effectively activated by the AL domian and loaded onto PCP to form myristoyl-*S*-PCP in an ATP-dependent manner. And the 210 Da shift in the Q-TOF-MS analysis is in agreement with addition of a myristoyl acid to the holo-PCP (Figure 3C).

After generation of **2**, the next post-NRPS tailoring steps including oxygenation and *N*-methylation may be performed by NapA and NapV to yield intermediate **3** (Figure 1B). Finally, the fully modified propeptide **3** will be matured through hydrolysis of an ester bond to remove the leader peptide and exported from the cell to give NDM, which most likely will be catalyzed by NapG (Figure 1B), a putative peptidase with nine predicted transmembrane helices (Figure S5, Supporting Information). To

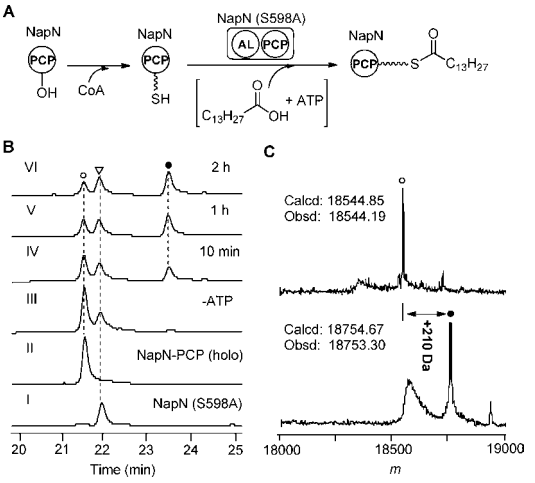


Figure 3. Biochemical characterization of the starter module. (A) Enzymatic reaction. (B) HPLC analysis (UV absorbance monitored at 220 nm): (I) holo-PCP; (II) NapN (S598A); (III) reaction without ATP; Time-course of enzymatic reaction for 10 min (IV), 1 h (V), and 2 h (VI). (C) Q-TOF-MS analysis of holo-PCP and enzymatic product. (▽) NapN (S598A); (○) holo-PCP; (●) myristoyl-*S*-PCP.

support this proposal, *napG* was also inactivated by gene replacement. Although the mutant *S. lusitanus* TG3018 (Δ*napG*) still produced NDM,[16] the yield decreased significantly, and gene complementation could restore NDM biosynthesis (Figure 2C). So, we concluded that *napG* is important for NDM biosynthesis, and it is also consistent with the cleavage of a precursor by a membrane-bound peptidase.

In summary, genetic and biochemical characterization of NDM gene cluster revealed an interesting biosynthetic machinery. This study presented here paves the way for further biochemical investigations into the novel mechanisms and sets the stage for the next combinatorial biosynthesis.

Acknowledgment. We thank Prof. Zixin Deng's laboratory at Shanghai Jiao Tong University for support in obtaining MS data of proteins. Supporting grants include 973 Program 2010CB833200 and 2009CB118901, the National Natural Science Foundation of China (21072214), and the Open Funding Project of the State Key Laboratory of Bioreactor Engineering.

Supporting Information Available. Experimental details. This material is available free of charge via the Internet at http://pubs.acs.org.

(16) The mutant still producing NDM means the function of napG might be complementated in part by peptidases in other metabolic pathways.

The authors declare no competing financial interest.