JOURNAL OF BACTERIOLOGY, Jan. 2011, p. 441–448 Vol. 193, No. 2
0021-9193/11/$12.00 doi:10.1128/JB.00897-10
Copyright © 2011, American Society for Microbiology. All Rights Reserved.

Identification of Nocobactin NA Biosynthetic Gene Clusters in *Nocardia farcinica*▽§

Yasutaka Hoshino,[1]†* Kazuhiro Chiba,[1]† Keiko Ishino,[1]‡ Toshio Fukai,[2] Yasuhiro Igarashi,[3] Katsukiyo Yazawa,[4] Yuzuru Mikami,[4] and Jun Ishikawa[1]

Department of Bioactive Molecules, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku, Tokyo 162-8640, Japan[1]; Department of Health Pharmacy, Yokohama College of Pharmacy, 601 Matano-cho, Totsuka-ku, Yokohama, Kanagawa 245-0066, Japan[2]; Biotechnology Research Center, Toyama Prefectural University, 5180 Kurokawa, Imizu, Toyama 939-0398, Japan[3]; and Medical Mycology Research Center, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8673, Japan[4]

Received 30 July 2010/Accepted 8 November 2010

We identified the biosynthetic gene clusters of the siderophore nocobactin NA. The *nbt* clusters, which were discovered as genes highly homologous to the mycobactin biosynthesis genes by the genomic sequencing of *Nocardia farcinica* IFM 10152, consist of 10 genes separately located at two genomic regions. The gene organization of the *nbt* clusters and the predicted functions of the *nbt* genes, particularly the cyclization and epimerization domains, were in good agreement with the chemical structure of nocobactin NA. Disruptions of the *nbtA* and *nbtE* genes, respectively, reduced and abolished the productivity of nocobactin NA. The heterologous expression of the *nbtS* gene revealed that this gene encoded a salicylate synthase. These results indicate that the *nbt* clusters are responsible for the biosynthesis of nocobactin NA. We also found putative IdeR-binding sequences upstream of the *nbtA*, *-G*, *-H*, *-S*, and *-T* genes, whose expression was more than 10-fold higher in the low-iron condition than in the high-iron condition. These results suggest that *nbt* genes are regulated coordinately by IdeR protein in an iron-dependent manner. The Δ*nbtE* mutant was found to be impaired in cytotoxicity against J774A.1 cells, suggesting that nocobactin NA production is required for virulence of *N. farcinica*.

The accumulation of genome sequence data has led to a significant understanding of the nature of microorganisms, such as their physiology, evolution, and pathogenicity. The discovery of many natural products from biosynthetic gene clusters in microbial genomes could be counted among these achievements. Recently, genomic sequencing has been used to unveil the hidden production of natural products, and the tide is becoming increasingly accelerated by the emergence of next-generation sequencers. However, despite the identification of hundreds of natural-product biosynthetic gene clusters using genomic sequencing, there are only a few clusters whose metabolites have been specified thus far. In actinomycetes, the discoveries of coelichelin from *Streptomyces coelicolor* M145 (taxonomically belongs to *Streptomyces violaceoruber*), fuscachelin from *Thermobifida fusca*, and erythrochelin from *Saccharopolyspora erythraea* are typical examples of the identification of new natural products by using a genomic approach (5, 16, 26). Since these biosynthetic gene clusters contain nonribosomal peptide synthetase (NRPS) genes, the structure of their products could be predicted using deduced amino acid sequence analysis (32). Thus, in the biosynthesis of nonribosomal peptide natural products, a genomic approach is particularly effective.

Recently, we sequenced the genome of *Nocardia farcinica* IFM 10152 (11), a clinical isolate, and showed that its genome contains many more genes for natural product biosynthesis than were previously estimated from traditional approaches. The *N. farcinica* genome has 14 NRPS genes and 7 polyketide synthase (PKS) genes. Among them, three NRPS genes and two PKS genes are clustered with some other genes. This cluster (cluster I) is predicted to be involved in the biosynthesis of a siderophore because of the significant homologies of genes in the cluster to the biosynthetic genes for mycobactin, a siderophore produced by *Mycobacterium tuberculosis* (Table 1). Cluster I includes 8 genes designated *nbtA*, *-B*, *-C*, *-D*, *-E*, *-F*, *-G*, and *-H* (Fig. 1A). NbtA is homologous to thioesterases, especially to the thioesterase domain of MbtB from *M. tuberculosis*. NbtB and NbtC could form a PKS and contain ketoacyl synthase (KS), acyltransferase (AT), ketoreductase (KR), and acyl carrier protein (ACP) domains. NbtD, NbtE, and NbtF are NRPSs, and each of them contains at least one module defined by adenylation (A), peptidyl carrier protein (PCP), and condensation (C) domains. NbtG is homologous to the lysine *N*-oxygenase of MbtG that catalyzes the N6-hydroxylation of lysine (15). NbtH is homologous to MbtK, which transfers an acyl chain to the ε-amino group of lysine. These facts convinced us that *N. farcinica* produces a siderophore that consists of three amino acids and an acyl group; therefore, we first tried to isolate this siderophore from *N. farcinica*.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Escherichia coli* JM109 was used as the host strain for gene cloning. *E. coli* BL21 *recA*::Tn*10*/pUB307 was used as the donor strain for conjugation into *Streptomyces avermitilis* (both strains were a gift from

* Corresponding author. Mailing address: Department of Bioactive Molecules, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku, Tokyo 162-8640, Japan. Phone: 81 3 5285 1111. Fax: 81 3 5285 1272. E-mail: hyasu@nih.go.jp.

† These authors contributed equally to this work.

‡ Present address: School of Pharmaceutical Sciences, Showa University, 1-5-8 Hatanodai, Shinagawa, Tokyo 142-8555, Japan.

§ Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 19 November 2010.

TABLE 1. Deduced functions of the nocobactin biosynthetic genes

Locus tag[a]	Gene name	Size (aa)[b]	Proposed function	Protein homolog (NCBI accession no.; % identity/% similarity)
nfa6190	*nbtS*	438	Salicylate synthase	MbtI (NP_336935; 52/65), salicylate synthase; Irp9 (YP_001006810; 40/58), salicylate synthase
nfa6200	*nbtT*	536	Salicylate-AMP ligase	MbtA (NP_216900; 57/69), salicyl-AMP ligase/salicyl-*S*-ArCP synthetase; EntE (NP_415126; 47/62), 2,3-dihydroxybenzoate-AMP ligase
nfa7610	*nbtG*	429	Lysine-*N*-oxygenase	MbtG (NP_216894; 56/70), lysine-*N*-oxygenase
nfa7620	*nbtH*	226	Lysine acetyltransferase	Rv1347c, MbtK (NP_215863; 42/54), *N*-acyltransferase
nfa7630	*nbtA*	251	Thioesterase	MbtB; position 1177-1409 (NP_216899; 43/59), NRPS/phenyloxazoline synthase
nfa7640	*nbtB*	436	Polyketide synthase	MbtC (NP_216898; 55/72), PKS
nfa7650	*nbtC*	1,028	Polyketide synthase	MbtD (NP_216897; 34/49), PKS
nfa7660	*nbtD*	1,701	Nonribosomal peptide synthetase	MbtE (NP_216896; 46/61), NRPS
nfa7670	*nbtE*	1,522	Nonribosomal peptide synthetase	MbtF (NP_216895; 43/58), NRPS
nfa7680	*nbtF*	1,167	Nonribosomal peptide synthetase	MbtB (NP_216899; 52/65), NRPS/phenyloxazoline synthase; PchE (NP_252916; 42/55), dihydroaeruginoic acid synthetase

[a] According to the *N. farcinica* genome sequencing project.
[b] aa, amino acids.

H. Ikeda, Kitasato Institute for Life Sciences, Kitasato University, Japan). *S. avermitilis* SUKA-1 (14) was used as the host strain for the expression of the *nbtS* gene. *N. farcinica* IFM 10152 was obtained from the Medical Mycology Research Center, Chiba University, Japan, and maintained in our laboratory. *E. coli* cultures were grown in Luria-Bertani (LB) broth or agar medium at 37°C. The *E. coli-Streptomyces* shuttle plasmid pGM160Δ*aac1*::*oriT* (14) and pKU251 were provided by H. Ikeda and used for the heterologous expression of the *nbtS* gene in *S. avermitilis*. pK18*mobsacB* (29) was obtained from the National Institute of Genetics, Japan, and used for the construction of the deletion mutants. pNV18 and pNV19 (2) were used for the expression of the *nbt* genes in *N. farcinica*.

Purification and structural elucidation of nocobactin NA. Nocobactin NA was purified from a total of 2 liters of IFM 10152 culture by silica gel column and thin-layer chromatography. Detailed procedures are described in the supplemental material. Mass spectrometry (MS) and nuclear magnetic resonance (NMR) analyses were carried out using LCT Premier XE (Waters) and JNM EXP-500 (JEOL), respectively. The absolute configuration of nocobactin NA was determined by methanolysis followed by chiral high-performance liquid chromatography (HPLC) (Sumichiral OA-5000 columns, 4 by 150 mm; Sumika Chemical Analysis Service, Ltd.).

Detection of metabolites by HPLC. To analyze salicylate, the *S. avermitilis* strain carrying the *nbtS* gene was incubated in Trypticase soy broth (Becton, Dickinson & Co.) for 5 days at 28°C. The whole culture broth was added to an equal volume of methanol (MeOH) and subjected to reversed-phase HPLC (4.6- by 150-mm Cosmosil C_{18} AR-II columns; Nacalai Tesque) using a linear gradient from 10 to 90% CH_3CN with H_2O (containing 0.1% trifluoroacetic acid [TFA]) for 20 min at a flow rate of 1 ml/min.

To analyze nocobactin NA, *N. farcinica* strains were grown in minimal medium (MM) consisting of 1.5% $Na_2HPO_4 \cdot 2H_2O$, 0.3% KH_2PO_4, 0.5% $Na_3C_6H_5O_7 \cdot 2H_2O$, 0.02% NH_4Cl, 0.1% $MgSO_4 \cdot 7H_2O$, and 0.05% NaCl. After a 4-day incubation at 37°C, the cells were harvested from 5 ml of the culture and extracted with 250 μl of MeOH. The extracts were subjected to reversed-phase HPLC (4.6- by 150-mm Cosmosil C_{18} AR-II column; Nacalai

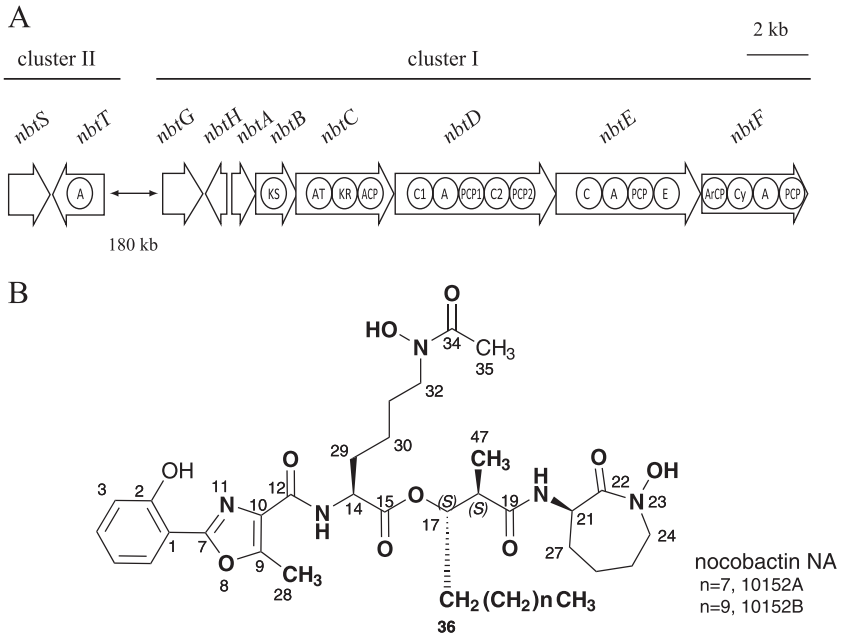

FIG. 1. Gene and domain organization of the *nbt* clusters (A) and structure of nocobactin NA from *N. farcinica* IFM 10152 (B). Abbreviations: A, adenylation; C, condensation; PCP, peptidyl carrier protein; ArCP, aryl carrier protein; E, epimerization; Cy, cyclization; KS, ketoacyl synthase; AT, acyltransferase; KR, ketoreductase; ACP, acyl carrier protein.

Tesque) using a linear gradient from 70 to 90% CH_3CN with H_2O (containing 0.1% TFA) for 20 min at a flow rate of 1 ml/min.

Construction of the *nbtS* gene expression plasmid in *S. avermitilis*. To clone the *nbtS* gene, a 4.1-kb XhoII fragment carrying the *nbtS* gene was prepared from pKNL033_G04, which is a plasmid from the *N. farcinica* IFM 10152 ordered plasmid library (http://nocardia.nih.go.jp), and cloned into the BamHI site of pGM160Δ*aacI*::*oriT* (14). The resulting plasmid, pGMnbtS, was digested with SacI, and the 2-kb SacI fragment was cloned into pKU251. The resulting plasmid, pKUnbtS, was introduced into *E. coli* BL21 *recA*::Tn*10*/pUB307 and transferred to *S. avermitilis* by conjugation (14).

Construction of deletion mutants. We constructed in-frame, unmarked deletions of the *nbtA*, *nbtE*, and *nbtS* genes using a previously reported method (10). For the in-frame deletion of *nbtA*, a 1.2-kb EcoRI-BglII fragment containing the *nbtA* gene was ligated to pBluescript KS(+) digested with EcoRI and BamHI, yielding pBnbtA. To make an in-frame deletion, pBnbtA was digested with NarI followed by self-ligation. A 0.8-kb EcoRI-XbaI fragment carrying the deletion allele was subcloned into pK18*mobsacB* to generate pKDnbtA.

For the in-frame deletion of *nbtE*, an *nbtE*-containing DNA fragment was amplified by PCR using the primers MfeI_nbtE_start (GAGGACAATTGCTT CGCCCTCGGCG) and MfeI_nbtE_end (CCGCCAATTGGTGCTCAGCTGC GGC). The 5.3-kb resultant fragment was digested with MfeI and cloned into the EcoRI site of pK18*mobsacB* to generate pKnbtE. To make an in-frame deletion, pKnbtE was digested with EcoRI followed by self-ligation, yielding pKDnbtE.

For the in-frame deletion of *nbtS*, a 2.9-kb XhoII-PstI fragment containing the *nbtS* gene was ligated to pK18*mobsacB* digested with BamHI and PstI. To make an in-frame deletion, the resulting plasmid was digested with StuI followed by self-ligation, yielding pKDnbtS.

Construction of plasmids for complementation experiments. For the *nbtA* gene, a 1.2-kb EcoRI-XbaI fragment containing the *nbtA* gene prepared from pBnbtA was ligated to pNV18 digested with EcoRI and XbaI, yielding pNVnbtA. For the *nbtS* gene, a 2.9-kb XhoII-XhoI fragment containing the *nbtS* gene was cloned into pNV19 digested with BamHI and SalI, yielding pNVnbtS. For the *nbtE* gene, the upstream region of the *nbtE* gene was replaced with the promoter region from the *nbtA* gene, because *nbtE* lacks its own promoter. First, a 6.1-kb XhoI fragment containing *nbtE* was cloned into the SalI site of pNV19, yielding pNVnbtE. Next, the 0.8-kb 5′ region of the *nbtE* fragment was amplified from pKDnbtE by PCR using the primers cross_AE (GTGTAGGGAGACGCAAG ATGACCGGGGCCACGCCGCA) and MfeI_nbtE_end. Thereafter, the 1.2-kb fragment carrying the fusion of the *nbtA* promoter region and the *nbtE* 5′ region was made by PCR, using pBnbtA as a template and the 0.8-kb fragment, M13-RV507 (TCCGGCTCGTATGTTGTGTGGA), and MfeI_nbtE_end as primers, and cloned into pCR-Blunt II-TOPO (Invitrogen), generating pCR_{nbtAp}nbtE. Finally, the 0.4-kb BsrGI-XbaI fragment of pNVnbtE was replaced with the 0.6-kb BsrGI-XbaI fragment of pCR_{nbtAp}nbtE, generating pNV_{nbtAp}nbtE.

qRT-PCR. The wild-type strain grown in brain heart infusion (BHI) broth was inoculated into MM broth containing 2 or 50 μM $FeCl_3$. After 24 h of incubation, cells were harvested and RNA was extracted with Sepasol-RNA I Super G (Nacalai Tesque), and then cDNAs were synthesized with ReverTra Ace (Toyobo Co. Ltd.). Quantitative reverse transcription-PCR (qRT-PCR) experiments were carried out in triplicates by using Thunderbird SYBR qPCR mix (Toyobo Co. Ltd.) and ABI PRISM 7000 (Applied Biosystems). Data were analyzed by 7000 Sequence Detection software version 1.2.3 and normalized to 16S rRNA as an internal control. PCR primers are as follows: nbtA_423F, GGAACTCCTCGACCACATGAC; nbtA_527R, CGGTAGTCGGCCTTCAT CAC; 16S_1248F, TGGAGCGAATCCCTTAAAGC; 16S_1348R, CCGCAGC GTTGCTGATCT; nbtT_1175F, CCTACACCATTCGCGGCTAT; nbtT_1322R, CGGTTGATGACGTCCTTGATC; nbtS_918F, CGCGGTGTCGGACTTCAT; nbtS_1043R, ACCGACGGGAACAGCACTT; nbtH_405F, CGACATGGCCTTC CACATC; nbtH_527R, AGCAGCCTGCGGCATTC; nbtG_765F, CAGCGACC CGACCAAGTG; and nbtG_894R, CAGGTGATGAACCCGGTTGT.

Infection experiment. A 24-well plate was seeded with 1×10^6 murine macrophage-like J774A.1 (ATCC TIB-67) cells propagated in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and incubated for 16 h at 37°C in a 5% CO_2 atmosphere. After a medium change, monolayers were infected with *Nocardia* strains at a multiplicity of infection of 5 and incubated for 24 h at 37°C in a 5% CO_2 atmosphere.

Bioinformatic analysis. Sequence alignment and tree drawing were carried out by using the ClustalX program. Prediction of the IdeR-binding sequence by the HMMER 2.3.2 program package (http://hmmer.org/) was carried out as follows. A profile-hidden Markov model was made from the published consensus sequence (17) by the hmmbuild program, and the genome sequence was searched by the hmmsearch program using the model. After the first-round search, the model was remade from the alignment of the newly identified sequences, and the processes were repeated until no more new sequences were found.

RESULTS

Isolation of a siderophore from *N. farcinica*. On the basis of genomic estimation, we attempted to isolate a siderophore from *N. farcinica* IFM 10152. From 2 liters of culture, two compounds (10152A, 125 mg; 10152B, 92 mg) were purified as compounds with antitumor activity against the HL-60 tumor cell line (20). In MS analysis, the molecular formulas of 10152A and 10152B were determined to be $C_{38}H_{58}N_5O_{10}$ (*m*/*z* of 744.4190 $[M+H]^+$; calcd, 744.4184) and $C_{40}H_{62}N_5O_{10}$ (*m*/*z* of 772.4451 $[M+H]^+$; calcd, 772.4497), respectively. Furthermore, ^1H and ^{13}C NMR analysis (see Table S1 in the supplemental material) revealed that 10152A and 10152B were nocobactin NA (Fig. 1B). Nocobactin NA has been reported to be produced by *Nocardia asteroides* ATCC 3313 (currently renamed *N. farcinica* ATCC 3313) and is a mixture primarily composed of two variants with nonyl and undecyl side chains (23), corresponding to 10152A and 10152B, respectively.

The structure of nocobactin NA has been determined by Ratledge and Snow (23), but its absolute configuration has not yet been determined. We have now, however, determined the absolute configuration of the *N*-hydroxy-*N*-acetyl-lysine and *N*-hydroxy-ε-caprolactam moieties of nocobactin NA. Nocobactic acid NA methyl ester and cobactin NA obtained by the methanolysis of nocobactin NA were analyzed using chiral HPLC. The lysine residues of the *N*-hydroxy-*N*-acetyl-lysine and *N*-hydroxy-ε-caprolactam moieties have been determined as L and D configurations, respectively. On the other hand, NMR data indicated that the two asymmetric carbons of nocobactin NA (C-17 and C-18) have an *S* configuration (see Table S1 in the supplemental material). These results provided crucial information for deducing the function of NRPSs in the biosynthesis of nocobactin NA, as described below.

Identification of the second cluster. Nocobactin NA consists of hydroxybenzoate, methyloxazoline, *N*-hydroxy-*N*-acetyl-lysine, a fatty acyl chain, and *N*-hydroxy-ε-caprolactam moieties. This is in good agreement with the gene content of cluster I except for the hydroxybenzoate moiety, which is considered to be derived from salicylate. Therefore, we searched the genome to identify the gene(s) responsible for salicylate biosynthesis. The second cluster (cluster II) was found 180 kb from cluster I and included two genes designated *nbtS* and *nbtT* (Fig. 1A). NbtS has high homologies to the bifunctional salicylate synthetases *Yersinia enterocolitica* Irp9 (13) and *M. tuberculosis* MbtI (8). NbtT is presumed to be a discrete NRPS A protein due to its extensive homologies to *E. coli* EntE (2,3-dihydroxybenzoate-AMP ligase) (28) and *M. tuberculosis* MbtA (salicylate-AMP ligase) (22).

Heterologous expression of the *nbtS* gene. To confirm whether the *nbtS* gene is responsible for the production of salicylate, we introduced the *nbtS* gene into *S. avermitilis*, whose genome lacks a salicylate synthase gene, using the pKUnbtS plasmid. As shown in Fig. 2, the production of salicylate was observed only when *S. avermitilis* carried pKUnbtS. This result clearly indicates that the *nbtS* gene encodes a salicylate synthase.

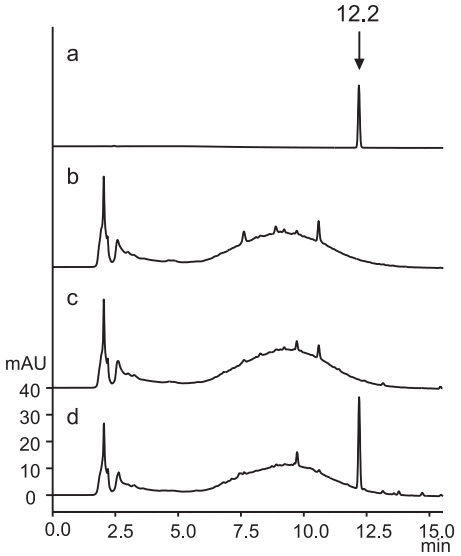

FIG. 2. Production of salicylate by the heterologous expression of the *nbtS* gene in *S. avermitilis*. HPLC conditions and sample preparations are described in Materials and Methods. (a) Authentic standard of salicylate; (b) *S. avermitilis*; (c) *S. avermitilis*/pKU251; (d) *S. avermitilis*/pKUnbtS. The elution peak of salicylate is indicated by an arrow with its retention time.

Inactivation of *nbt* genes. To elucidate the involvement of the clusters in nocobactin NA biosynthesis, we constructed unmarked in-frame deletion mutants of the *nbtA*, *nbtE*, and *nbtS* genes. In BHI medium, no apparent differences in growth or morphology were observed between the wild-type and mutant strains (data not shown); however, significant differences were observed for the production of nocobactin NA. In the Δ*nbtA* and Δ*nbtS* mutants, nocobactin NA production was reduced to 1% or less than that of the wild-type strain (Fig. 3A, b and f). In the Δ*nbtE* mutant, nocobactin NA was not detected, but two new peaks were detected instead (Fig. 3A, d). The retention times of the new peaks (8.2 and 13.2 min) were longer than those of nocobactin NA (7.4 and 12.3 min), but no difference was observed in their UV/visible absorption spectra (Fig. 3B). By MS analysis, the new peaks were determined to correspond to 2 compounds each. The molecular formulas of the new compounds with a retention time of 8.2 min were $C_{32}H_{48}N_3O_9$ (*m/z* of 618.3378 $[M+H]^+$; calcd, 618.3391) and $C_{32}H_{48}N_3O_8$ (*m/z* of 602.3432 $[M+H]^+$; calcd, 602.3441), and those with a retention time of 13.2 min were $C_{34}H_{52}N_3O_9$ (*m/z* of 646.3688 $[M+H]^+$; calcd, 646.3704) and $C_{34}H_{52}N_3O_8$ (*m/z* of 630.3754 $[M+H]^+$; calcd, 630.3754). These data indicate that the new compounds were nocobactin NA derivatives missing the *N*-hydroxy-ε-caprolactam moiety (Fig. 3C). Furthermore, the smaller compounds (*m/z* of 601 and 629) were most likely the 33-dehydroxy forms of the larger compounds.

Complementation analysis of the deletion mutations was performed using pNVnbtA, pNV_{nbtAp}nbtE, and pNVnbtS. The defects in nocobactin NA production in all mutants were restored by expressing the relevant gene in *trans* (Fig. 3A, c, e, and g, respectively). Furthermore, in the Δ*nbtS* mutant, the decrease in nocobactin NA production was also restored by supplementing 100 μg/ml salicylate to the culture broth (Fig. 3A, h).

Sequence analysis of the Nbt proteins. We further analyzed the deduced amino acid sequences of the *nbt* genes, details of which are described in the supplemental material. NbtA is likely to be a thioesterase (TE) and shows higher similarities to type II TE (TEII) than to type I TE (TEI) (see Fig. S1 in the supplemental material). NbtB and NbtC are orthologous to MbtC and MbtD, respectively, and probably form a PKS which condenses the long fatty acid side chain of nocobactin NA. As MbtCD catalyzes the condensation of 3-hydroxybutyrate, NbtBC catalyzes the condensation of 3-hydroxy-2-methyldodecanoate or 3-hydroxy-2-methyltetradecanoate, which were identified in the position corresponding to that of the 3-hydroxybutyrate of mycobactin. NbtD, NbtE, and NbtF are presumed to be NRPSs. Pfam domain searches of these NRPSs identified the domain organizations of NbtD, NbtE, and NbtF as C1-A-PCP1-C2-PCP2, C1-A-PCP-C2, and PCP1-C-A-PCP2, respectively. A phylogenetic analysis of each A domain (see Fig. S2 in the supplemental material) showed that the A domains of NbtD and NbtE were predicted to activate lysine, while that of NbtF was predicted to activate serine, threonine, or cysteine. Among the amino acids, threonine is considered the most reasonable substrate because of the structure of nocobactin NA. Since nocobactin NA contains a methyloxazoline ring, which could be generated by the cyclization of threonine, and a D-lysine as the *N*-hydroxy-ε-caprolactam moiety, it is possible that the cyclization (Cy) and epimerization (E) domains exist somewhere in the Nbt proteins. Alignment of the C domains of Nbt proteins with known Cy and E domains (12, 24) (see Fig. S3 in the supplemental material) and the existence of highly conserved sequences among known Cy (Fig. S4) or E (Fig. S5) domains have suggested that NbtF C and NbtE C2 are the Cy and E domains, respectively. The domain organizations of NbtF and NbtE are revised to PCP1-Cy-A-PCP2 and C-A-PCP-E, respectively. NbtF has two PCP domains; however, the phylogenetic analysis of PCP domains showed that the PCP1 domain of NbtF formed a clade with known aryl carrier proteins (ArCP) (see Fig. S6 in the supplemental material). According to this result, we revised the domain organization of NbtF to ArCP-Cy-A-PCP. NbtG and NbtH could both have tailoring functions. NbtG is highly homologous to the lysine-*N*-oxygenase of MbtG that catalyzes the *N*-hydroxylation of lysine (15). NbtH is homologous to MbtK, which transfers an acyl chain to the ε-amino group of lysine, suggesting that NbtH would transfer an acetyl group, which is an acyl group, to the ε-amino group of lysine. NbtS, as indicated above, is a salicylate synthase and shows homologies to the bifunctional salicylate synthases MbtI (8) and Irp9 (13). These enzymes catalyze the conversion of chorismate to isochorismate and the subsequent elimination of the enolpyruvyl side chain in a lyase reaction to generate salicylate. The A domain of NbtT forms a distinct clade with those of MbtA, YbtE, PchD, and EntE (see Fig. S2 in the supplemental material). It is reasonable to assume, therefore, that NbtT activates salicylate synthesized by NbtS and transfers it to NbtF containing the only ArCP among the Nbt proteins.

Regulatory sequence. Due to the limited availability of free iron in the host, iron plays a crucial role in pathogenesis. From this point of view, the genes in clusters I and II must be

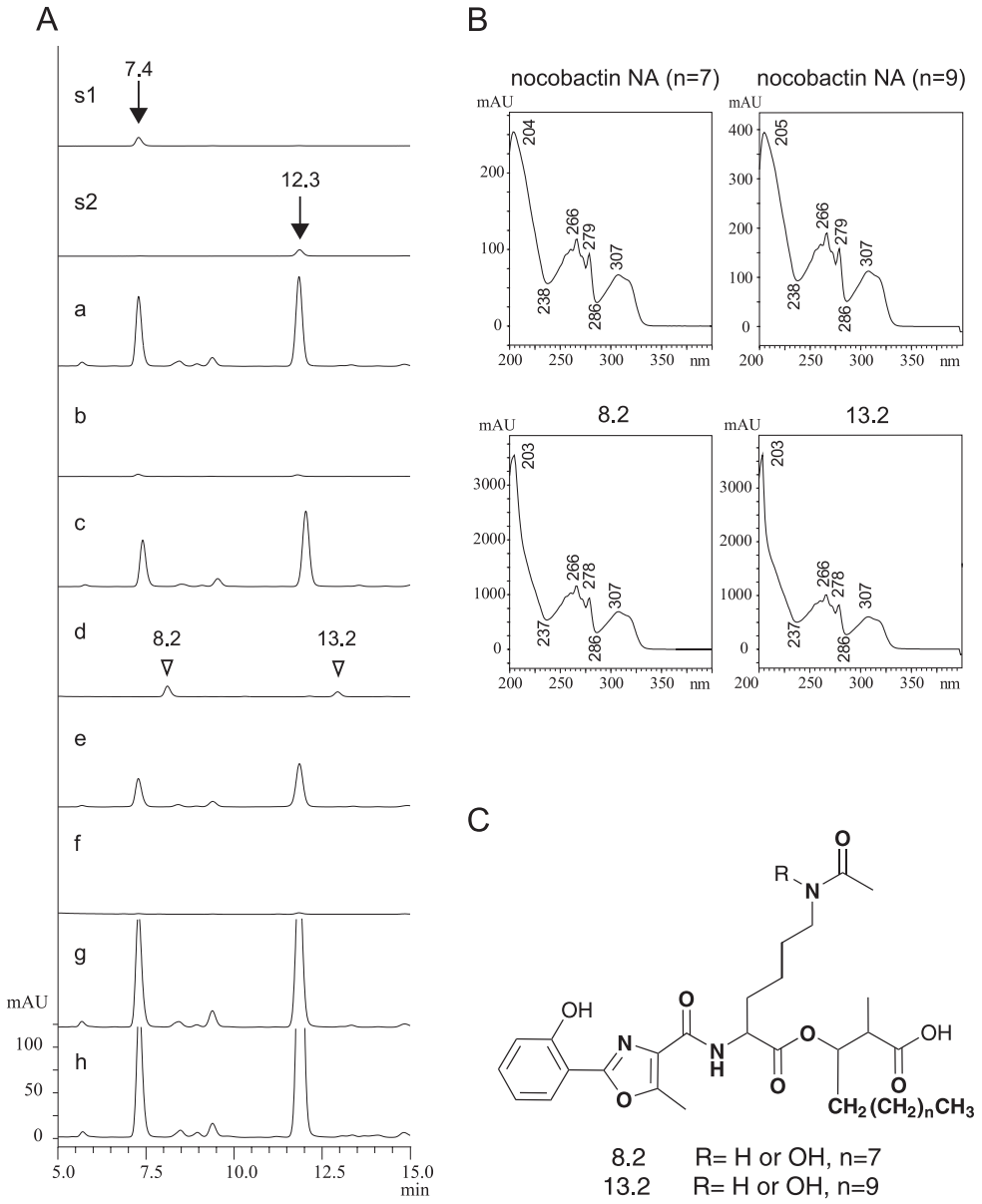

FIG. 3. (A) Analyses of nocobactin NA production in the wild-type and mutant strains. a, wild type; b, Δ*nbtA* mutant; c, Δ*nbtA*/pNVnbtA mutant; d, Δ*nbtE* mutant; e, Δ*nbtE*/pNV mutant with nbtE mutant; f, Δ*nbtS* mutant; g, Δ*nbtS*/pNVnbtS mutant; h, Δ*nbtS*100 μg/ml salicylate $_{nbtAp}$; s1, nocobactin NA ($n = 7$); s2, nocobactin NA ($n = 9$). Nocobactin NA was detected at 267 nm. The elution peaks of the standards are indicated by arrows with their retention times. White arrowheads indicate the elution peaks of the newly identified metabolites (8.2 and 13.2 min) in the Δ*nbtE* mutant. UV spectra (B) and proposed structures (C) of the newly identified metabolites are also indicated.

coordinately regulated. In *M. tuberculosis*, the *mbt* genes, which are clustered at two loci in the genome, are regulated by IdeR in an iron-dependent manner (27). An IdeR-iron complex binds to the upstream regions of genes to be regulated and represses their expression. Nocobactin NA biosynthesis appears to be regulated in a manner similar to mycobactin biosynthesis, because the *N. farcinica* genome contains the ortholog of IdeR (Nfa37790), and the productivity of nocobactin NA in MM (low-iron condition) is higher than that in BHI medium (high-iron condition) (data not shown). We searched the *N. farcinica* genome by using a profile-hidden Markov model made from the consensus sequence of the IdeR-binding site (17). A total of 33 sequences were identified, and 6 of them were located immediately upstream of the *nbtA*, *-G*, *-H*, *-S*, and *-T* genes (Table 2), suggesting that the *nbt* genes are coordinately regulated by IdeR. To confirm this, we carried out qRT-PCR, and the results showed that the expression of the five genes increased more than 10-fold within 24 h after shifting to the low-iron condition (Table 2).

Cytotoxicity of mutant strains. Cytotoxicity is exhibited by many pathogens and correlates with virulence. Cytotoxicity is often evaluated by cell detachment from a culture plate as well as by trypan blue staining and/or lactate dehydrogenase release. To estimate the involvement of nocobactin NA produc-

TABLE 2. Putative IdeR-binding sequence and the expression of downstream genes under low- and high-iron conditions

Position	IdeR-binding sequence	Downstream gene	mRNA expression (arbitrary units)[a]		Ratio (low/high)
			Low iron	High iron	
Published consensus	TTAGGTTAGGCTAACCTAA				
822260–822278	TTCGGTAAGGCTAGCCTAT	*nbtA*	10.4 ± 2.55	0.618 ± 0.190	16.8
820079–820097	TAAGGTAAGCCTATGCAAA	*nbtG*	45.9 ± 14.4	3.18 ± 0.952	14.4
822283–822265	TTAGTATAGGCTAGCCTTA	*nbtH*	12.8 ± 2.24	0.720 ± 0.167	17.9
640250–640268	CGAGGTAATGCTAACCTTA	*nbtS*	8.81 ± 1.43	0.778 ± 0.223	11.3
640280–640298	TTAGGTCAGCTTAACCTTT	*nbtS*	8.81 ± 1.43	0.778 ± 0.223	11.3
643373–643355	ATAGGGTTGCCTAACCAAA	*nbtT*	10.1 ± 1.07	0.759 ± 0.268	13.3

[a] mRNA levels were normalized to the 16S rRNA level and shown as the means ± standard deviations.

tion in virulence, a murine macrophage-like cell line, J774A.1, was infected with the Δ*nbtE* and wild-type strains (Fig. 4). After 24 h of infection, the wild-type strain detached and clumped the J774A.1 cells, whereas the Δ*nbtE* mutant did not. Furthermore, the complemented strain was as cytotoxic as the wild-type strain. These results suggest that nocobactin NA plays a role in the virulence of *Nocardia*.

DISCUSSION

Genome-based approach. Identifying biosynthetic gene clusters through a genome-based approach is a successful strategy for the discovery of secondary metabolites (35). We initially predicted that *N. farcinica* IFM 10152 produced a mycobactin-type siderophore, and as expected, the isolated compounds (10152A and 10152B) turned out to be siderophore nocobactin NA (23). The gene organization of the *nbt* clusters is consistent with the structure of nocobactin NA. This fact confirms that a genome-based approach is potent for discovering natural products. Besides nocobactin NA, many mycobactin-type siderophores have already been isolated from *Nocardia* species, for example, formobactin (19), amamistatin (33), brasilibactin (34), and asterobactin (20), but their biosynthetic genes have not yet been identified. It is certain that a genome-based approach will facilitate their identification.

Biosynthetic pathway. On the basis of the results described above, we propose the following biosynthetic pathway for nocobactin NA (Fig. 5). Nocobactin NA biosynthesis is initiated by loading salicylate, which is synthesized from chorismate by NbtS and activated by NbtT, onto the NbtF ArCP domain. NbtF then activates and cyclizes threonine, yielding 2-hydroxyphenyl-(5-methyloxazoline) (asteroidic acid). Subsequently, condensation with lysine and a long fatty acyl chain occurs with NbtD. Thereafter, the epimerization and condensation of lysine are catalyzed by NbtE. The terminal ε-capro-

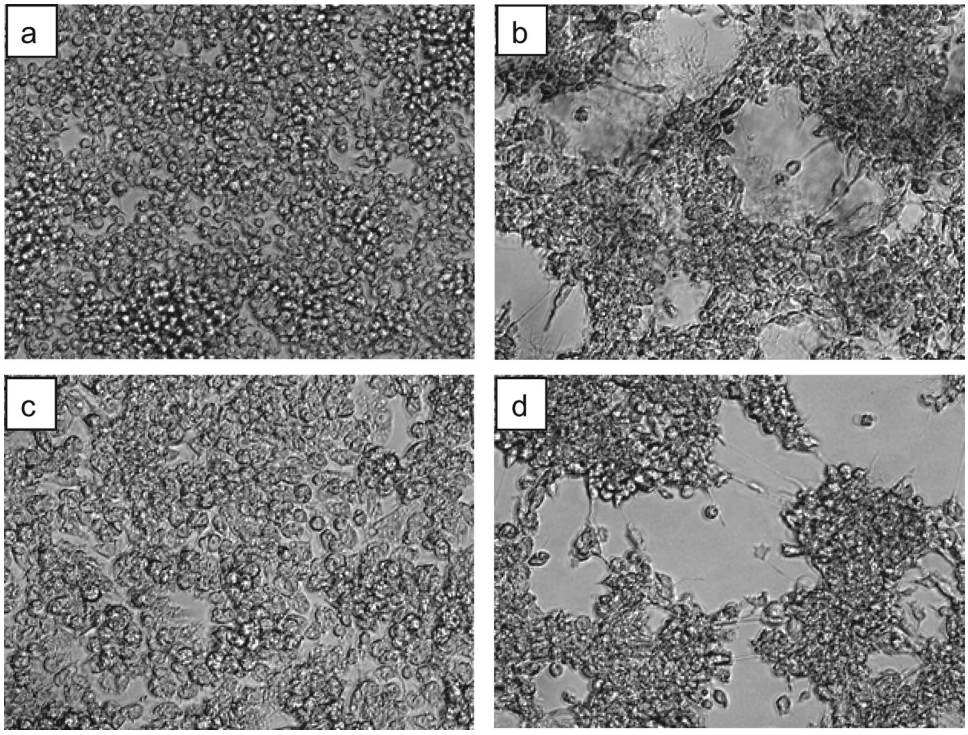

FIG. 4. Cytotoxic activity of *Nocardia* strains. J774A.1 cells were infected with the wild-type (b), Δ*nbtE* (c), and Δ*nbtE*/pNV$_{nbtAp}$nbtE (d) strains, and uninfected J774A.1 cells (a) served as the control. Cells were photographed by phase-contrast microscopy 24 h after infection. Magnification, ×200.

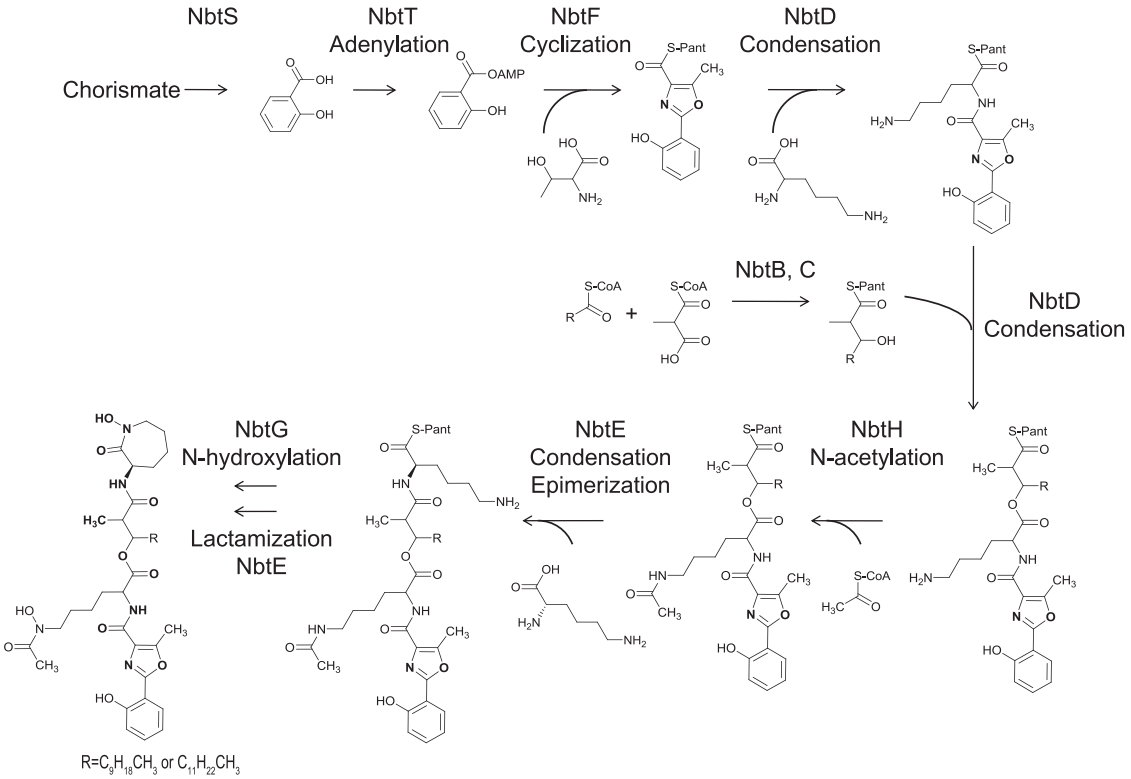

FIG. 5. Proposed biosynthetic pathway of nocobactin NA.

lactam moiety is generated by the intramolecular lactamization of lysine, resulting in the release of the product from the enzymes. However, it is not clear when NbtG and NbtH play their roles in nocobactin NA biosynthesis. In aerobactin biosynthesis, free lysine is hydroxylated, acetylated, and then incorporated into the assembly line as *N*-hydroxy-*N*-acetyl-lysine (3). In contrast, in mycobactin biosynthesis, the N6-hydroxylation of lysine residues occurs after the synthesis of the mycobactin backbone because didehydroxymycobactin, whose lysine residues are not N6-hydroxylated, has been isolated from *M. tuberculosis* (18). In nocobactin NA biosynthesis, we propose that the N6-hydroxylation of lysine residues by NbtG occurs subsequently to the synthesis of the nocobactin NA backbone. This hypothesis would be supported, in part, by the identification of 33-dehydroxy derivatives in the Δ*nbtE* mutant (Fig. 3C). Identification of such derivatives may also indicate that the 33-acetylation by NbtH occurs at least before the 33-hydroxylation by NbtG.

NbtBC and MbtCD are orthologous to each other, and both contain only one PKS module (KS-AT-KR-ACP). In MbtCD, this module organization is consistent with the existence of a 3-hydroxybutyrate spacer between the two hydroxylysine residues of mycobactin. However, 3-hydroxy-2-methyldodecanoate or 3-hydroxy-2-methyltetradecanoate exists at the corresponding position in nocobactin NA. Since such long fatty acids are unable to be synthesized by a single PKS module, NbtC appears to carry long-chain fatty acids.

Heterologous expression of the *nbtS* gene in *S. avermitilis* resulted in the production of salicylate. This observation indicates that the *nbtS* gene encodes a salicylate synthase. In such context, disruption of the *nbtS* gene was expected to lead to the loss of nocobactin NA production. However, unexpectedly, the Δ*nbtS* mutant produced a small amount of nocobactin NA (less than 1% of that produced by the wild-type strain). Therefore, *N. farcinica* might be able to produce salicylate through an unknown biosynthetic pathway.

On the basis of sequence analysis, NbtA was presumed to be a TEII. TEI is responsible for the release of the acyl chain from the PKS or NRPS, whereas TEII is proposed to rid aberrantly loaded substrates from the PKS or NRPS (1, 6, 9, 21, 25, 31). From this point of view, NbtA possibly plays such an editing role in the biosynthesis of nocobactin NA. It is known from many cases that the disruption of TEII genes results in a significant decrease in product yield. For example, disruption of *angT* (36), *srfA*-TE (30), or *ybtT* (7) reduced anguibactin, surfactin, or yersiniabactin production to 30 to 40%, 16%, or 6% of the wild-type level, respectively. In contrast, the Δ*nbtA* mutant produces nocobactin NA levels less than 1% of the wild-type level. This result implies that NRPSs for nocobactin NA biosynthesis might be error prone and that NbtA could play a critical role in the biosynthesis of nocobactin NA. However, the biosynthetic steps edited by NbtA remain unclear.

We found putative IdeR-binding sequences upstream of the *nbtA*, *-G*, *-H*, *-S*, and *-T* genes that demonstrated more than 10-fold-higher expression of the downstream genes in the low-iron condition than in the high-iron condition (Table 2). This finding suggests that nocobactin NA biosynthesis is regulated by IdeR in an iron-dependent manner. Such regulation could be important for the coordinated expression of *nbt* genes located at two different regions distant from each other.

Virulence. Iron is an essential nutrient for the growth of most organisms and is especially important for the survival of pathogens within a host. Mycobactin is a known virulence factor of *M. tuberculosis*, and mycobactin-deficient mutants have attenuated virulence (4, 17). In the present study, we showed the decreased cytotoxicity of the deletion mutant, suggesting the involvement of nocobactin NA in the virulence of *N. farcinica*. Detailed studies are currently in progress to investigate the role of nocobactin NA in the virulence of *Nocardia* and will be published in the near future.

ACKNOWLEDGMENT

This work was supported in part by KAKENHI (grant no. 17019009 and 19710170).

REFERENCES

1. **Butler, A. R., N. Bate, and E. Cundliffe.** 1999. Impact of thioesterase activity on tylosin biosynthesis in *Streptomyces fradiae*. Chem. Biol. **6:**287–292.
2. **Chiba, K., et al.** 2007. Construction of a pair of practical *Nocardia-Escherichia coli* shuttle vectors. Jpn. J. Infect. Dis. **60:**45–47.
3. **De Lorenzo, V., and J. B. Neilands.** 1986. Characterization of *iucA* and *iucC* genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli*. J. Bacteriol. **167:**350–355.
4. **De Voss, J. J., et al.** 2000. The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. Proc. Natl. Acad. Sci. U. S. A. **97:**1252–1257.
5. **Dimise, E. J., P. F. Widboom, and S. D. Bruner.** 2008. Structure elucidation and biosynthesis of fuscachelins, peptide siderophores from the moderate thermophile *Thermobifida fusca*. Proc. Natl. Acad. Sci. U. S. A. **105:**15311–15316.
6. **Doi-Katayama, Y., et al.** 2000. Thioesterases and the premature termination of polyketide chain elongation in rifamycin B biosynthesis by *Amycolatopsis mediterranei* S699. J. Antibiot. **53:**484–495.
7. **Geoffroy, V. A., J. D. Fetherston, and R. D. Perry.** 2000. *Yersinia pestis* YbtU and YbtT are involved in synthesis of the siderophore yersiniabactin but have different effects on regulation. Infect. Immun. **68:**4452–4461.
8. **Harrison, A. J., et al.** 2006. The structure of MbtI from *Mycobacterium tuberculosis*, the first enzyme in the biosynthesis of the siderophore mycobactin, reveals it to be a salicylate synthase. J. Bacteriol. **188:**6081–6091.
9. **Heathcote, M. L., J. Staunton, and P. E. Leadlay.** 2001. Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units. Chem. Biol. **8:**207–220.
10. **Ishikawa, J., K. Chiba, H. Kurita, and H. Satoh.** 2006. Contribution of *rpoB2* RNA polymerase β subunit gene to rifampin resistance in *Nocardia* species. Antimicrob. Agents Chemother. **50:**1342–1346.
11. **Ishikawa, J., et al.** 2004. The complete genomic sequence of *Nocardia farcinica* IFM 10152. Proc. Natl. Acad. Sci. U. S. A. **101:**14925–14930.
12. **Keating, T. A., D. A. Miller, and C. T. Walsh.** 2000. Expression, purification, and characterization of HMWP2, a 229 kDa, six domain protein subunit of yersiniabactin synthetase. Biochemistry **39:**4729–4739.
13. **Kerbarh, O., A. Ciulli, N. I. Howard, and C. Abell.** 2005. Salicylate biosynthesis: overexpression, purification, and characterization of Irp9, a bifunctional salicylate synthase from *Yersinia enterocolitica*. J. Bacteriol. **187:**5061–5066.
14. **Komatsu, M., T. Uchiyama, S. Omura, D. E. Cane, and H. Ikeda.** 2010. Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism. Proc. Natl. Acad. Sci. U. S. A. **107:**2646–2651.
15. **Krithika, R., et al.** 2006. A genetic locus required for iron acquisition in *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U. S. A. **103:**2069–2074.
16. **Lautru, S., R. J. Deeth, L. M. Bailey, and G. L. Challis.** 2005. Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. Nat. Chem. Biol. **1:**265–269.
17. **Manabe, Y. C., B. J. Saviola, L. Sun, J. R. Murphy, and W. R. Bishai.** 1999. Attenuation of virulence in *Mycobacterium tuberculosis* expressing a constitutively active iron repressor. Proc. Natl. Acad. Sci. U. S. A. **96:**12844–12848.
18. **Moody, D. B., et al.** 2004. T cell activation by lipopeptide antigens. Science **303:**527–531.
19. **Murakami, Y., et al.** 1996. Formobactin, a novel free radical scavenging and neuronal cell protecting substance from *Nocardia* sp. J. Antibiot. **49:**839–845.
20. **Nemoto, A., et al.** 2002. Asterobactin, a new siderophore group antibiotic from *Nocardia asteroides*. J. Antibiot. **55:**593–597.
21. **Pazirandeh, M., S. S. Chirala, and S. J. Wakil.** 1991. Site-directed mutagenesis studies on the recombinant thioesterase domain of chicken fatty acid synthase expressed in *Escherichia coli*. J. Biol. Chem. **266:**20946–20952.
22. **Quadri, L. E., J. Sello, T. A. Keating, P. H. Weinreb, and C. T. Walsh.** 1998. Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. Chem. Biol. **5:**631–645.
23. **Ratledge, C., and G. A. Snow.** 1974. Isolation and structure of nocobactin NA, a lipid-soluble iron-binding compound from *Nocardia asteroides*. Biochem. J. **139:**407–413.
24. **Rausch, C., I. Hoof, T. Weber, W. Wohlleben, and D. H. Huson.** 2007. Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. BMC Evol. Biol. **7:**78.
25. **Reimmann, C., H. M. Patel, C. T. Walsh, and D. Haas.** 2004. PchC thioesterase optimizes nonribosomal biosynthesis of the peptide siderophore pyochelin in *Pseudomonas aeruginosa*. J. Bacteriol. **186:**6367–6373.
26. **Robbel, L., T. A. Knappe, U. Linne, X. Xie, and M. A. Marahiel.** 2010. Erythrochelin—a hydroxamate-type siderophore predicted from the genome of *Saccharopolyspora erythraea*. FEBS J. **277:**663–676.
27. **Rodriguez, G. M., M. I. Voskuil, B. Gold, G. K. Schoolnik, and I. Smith.** 2002. *ideR*, an essential gene in *Mycobacterium tuberculosis*: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. Infect. Immun. **70:**3371–3381.
28. **Rusnak, F., W. S. Faraci, and C. T. Walsh.** 1989. Subcloning, expression, and purification of the enterobactin biosynthetic enzyme 2,3-dihydroxybenzoate-AMP ligase: demonstration of enzyme-bound (2,3-dihydroxybenzoyl)adenylate product. Biochemistry **28:**6827–6835.
29. **Schäfer, A., et al.** 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene **145:**69–73.
30. **Schneider, A., and M. A. Marahiel.** 1998. Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in nonribosomal peptide biosynthesis in *Bacillus subtilis*. Arch. Microbiol. **169:**404–410.
31. **Schwarzer, D., H. D. Mootz, U. Linne, and M. A. Marahiel.** 2002. Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. Proc. Natl. Acad. Sci. U. S. A. **99:**14083–14088.
32. **Stachelhaus, T., H. D. Mootz, and M. A. Marahiel.** 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. **6:**493–505.
33. **Suenaga, K., S. Kokubo, C. Shinohara, T. Tsuji, and D. Uemura.** 1999. Structures of amamistatins A and B, novel growth inhibitors of human tumor cell lines from an actinomycete. Tetrahedron Lett. **40:**1945–1948.
34. **Tsuda, M., et al.** 2005. Brasilibactin A, a cytotoxic compound from actinomycete *Nocardia brasiliensis*. J. Nat. Prod. **68:**462–464.
35. **Van Lanen, S. G., and B. Shen.** 2006. Microbial genomics for the improvement of natural product discovery. Curr. Opin. Microbiol. **9:**252–260.
36. **Wertheimer, A. M., et al.** 1999. Characterization of the *angR* gene of *Vibrio anguillarum*: essential role in virulence. Infect. Immun. **67:**6496–6509.