The *dhb* Operon of *Bacillus subtilis* Encodes the Biosynthetic Template for the Catecholic Siderophore 2,3-Dihydroxybenzoate-Glycine-Threonine Trimeric Ester Bacillibactin* 

Received for publication, October 6, 2000, and in revised form, December 6, 2000
Published, JBC Papers in Press, December 8, 2000 DOI 10.1074/jbc.M009140200

Jürgen J. May, Thomas M. Wendrich, and Mohamed A. Marahiel‡

From the Department of Chemistry, Philipps-Universität Marburg, D-35032 Marburg, Germany

*Bacillus subtilis* was reported to produce the catecholic siderophore itoic acid (2,3-dihydroxybenzoate (DHB)-glycine) in response to iron deprivation. However, by inspecting the DNA sequences of the genes *dhbE*, *dhbB*, and *dhbF* as annotated by the *B. subtilis* genome project to encode the synthetase complex for the siderophore assembly, various sequence errors within the *dhbF* gene were predicted and confirmed by re-sequencing. According to the corrected sequence, *dhbF* encodes a dimodular instead of a monomodular nonribosomal peptide synthetase. We have heterologously expressed, purified, and assayed the substrate selectivity of the recombinant proteins Dhbb, DhbB, and DhbF. DhbE, a stand-alone adenylation domain of 59.9 kDa, activates, in an ATP-dependent reaction, DHB, which is subsequently transferred to the free thiol group of the cofactor phosphopantetheine of the bifunctional isochorismate lyase/aryl carrier protein DhbB. The third synthetase, DhbF, is a dimodular nonribosomal peptide synthetase of 264 kDa that specifically adenylates threonine and, to a lesser extent, glycine and that covalently loads both amino acids onto their corresponding peptidyl carrier domains. To functionally link the *dhb* gene cluster to siderophore synthesis, we have disrupted the *dhbF* gene. Comparative mass spectrometric analysis of culture extracts from both the wild type and the *dhbF* mutant led to the identification of a mass peak at *m/z* 881 (M+H)* that corresponds to a cyclic trimeric ester of DHB-glycine-threonine.

Iron is an essential trace element. In organisms, most iron is bound intracellularly in heme, iron storage compounds such as ferritin, or iron-sulfur compounds. It can also be complexed to extracellular carrier glycoproteins in the form of transferrin in serum or lactoferrin in mussel surfaces (1–3). Under certain conditions, the level of physiologically available iron in terrestrial as well as in aquatic environments can drop by oxidation of ferrous to ferric iron (2, 4). This leads to concentrations far below 1 μM iron and becomes growth-limiting for bacteria. To survive, many bacteria evolved specialized transport systems to retrace ferric iron ions by utilizing low molecular mass iron-chelating compounds termed siderophores, which also play a crucial role in successful infection of pathogens in their host (5).

The most common siderophores can be classified traditionally into two major groups, catecholic and hydroxamate siderophores, which have recently been extended to a third main group of carboxylic siderophores. Other important subgroups are heterocyclic and mixed-type siderophores (6). Many of the siderophores are produced nonribosomally by large multidomain enzymes termed nonribosomal peptide synthetases (NRPS)† that can assemble peptides of wide structural diversity and broad biological activity (7–9). NRPS have a distinct modular structure in which each module is responsible for the recognition, activation, and, in some cases, modification of a single amino acid residue of the final peptide product. The modules are aligned in a sequence that is collinear with the sequence of the product (10) and can be subdivided into domains that catalyze specific biochemical reactions (7, 8). The adenylation (A) domain recognizes a specific substrate (amino or hydroxy acid) and activates it as (amino) acyl adenylate by hydrolysis of ATP. The activated acyl moiety is then covalently thioesterified to the enzyme-bound cofactor 4′-phosphopantetheine of a peptidyl carrier protein domain (11, 12). This reaction is followed by the direct transfer of the aminoacyl-S-Ppant-PCP intermediate onto the adjacent downstream module, a peptide bond formation mediated by the condensation (C) domain (13). A terminal thioesterase (Te) attached to the C-terminal module releases the peptide from the enzyme by cyclization or hydrolysis. For post-translational modification of all PCP domains in NRPS, a dedicated 4′-phosphopantetheinyltransferase (Ppant-PCP transferase) is needed. It converts all PCP domains from their inactive apo form to the active holo form by covalently linking a phosphopantetheine moiety from coenzyme A to a highly conserved serine residue (14). The crystal structure of Sfp, a Ppant-transferase associated with the biosynthesis of the lipopeptide surfactin, has recently been solved (15, 16), as well as the NMR structure of a prototype PCP domain (17).

In the case of the catechol-type siderophore enterobactin from *Escherichia coli*, biochemical data proved the current model of the NRPS assembly line mechanism. The activation of the first hydroxy acid, 2,3-dihydroxybenzoate (DHB), is catalyzed by the aminoacyl-CoA ligase EntE (18) transferring the activated aryl moiety onto the aryl carrier protein (ArCP) encoded by *entB* (19, 20). Amide bond formation with serine is catalyzed by EntF, a four-domain protein exhibiting the domain organization C-A-PCP-Te (see Fig. 1C) (21, 22). The Te

* This work was supported by the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie, and the Graduiertenkolleg “Prozesschemie mit 18 U.S.C. Section 1734 solely to indicate this fact.

† The abbreviations used are: NRPS, nonribosomal peptide synthetase(s); A domain, adenylation domain; C domain, condensation domain; Te domain, thioesterase domain; Ppant, 4′-phosphopantetheinyl; PCP, peptidyl carrier protein; ArCP, aryl carrier protein domain; DHB, 2,3-dihydroxybenzoate; PCR, polymerase chain reaction; CAS, ferric hexadecyltrimethylammonium bromide-Chrom-Azurol-S; ESI-MS, electrospray ionization mass spectrometry.
domain has been shown to be responsible for trimerization and release of the trilactone enterobactin (DHB-Ser$_3$) (23). The only catecholic siderophore that has been described so far in Bacillus subtilis is the itioic acid (DHB-glycine) encoded by the dhb operon (24, 25). Transcription of the dhb operon was found to be controlled by a single $\sigma^B$-dependent promoter that comprises a fur box-binding site, an iron regulatory element (26). According to the published data derived from B. subtilis genome strain 168, DhbF would end within a C domain. Consequently, no terminal Te domain that is normally fused to the C-terminal NRPS module had been annotated, indicating a nonfunctional enzyme. This unusual truncation of dhbF led us to take a closer look at the neighboring sequences and the entire biosynthetic operon for this siderophore.

In this study, we re-sequenced the dhbF gene from several B. subtilis strains, including B. subtilis genome project strain 168, and have identified several sequence errors within the dhbF gene. We expressed the three NRPS genes dhbB, dhbE, and dhbF and purified and characterized the synthetases DhbB, DhbE, and DhbF in vitro. By comparing the culture broth extracts of a dhbF knockout mutant and several B. subtilis strains, the siderophore could be identified and characterized.

**EXPERIMENTAL PROCEDURES**

**PCR Amplification, Cloning, and Sequencing**—PCR amplifications were carried out using the Expand Long Range PCR system (Roche Molecular Biochemicals, Mannheim, Germany) in accordance with the manufacturer’s protocol. Chromosomal DNA of B. subtilis ATCC 23132 that had been prepared using a DNA preparation kit (QIAGEN, Hilden, Germany) was used as a template. Primers (listed in Table I) were that had been prepared using a DNA preparation kit (QIAGEN, Hilden, Germany) in accordance with the manufacturer’s protocol. Chromosomal DNA of B. subtilis ATCC 23132 was used as a template.

**Siderophore Extraction**—For siderophore extraction, 200 ml of cells from B. subtilis strains ATCC 21332, 168, and JH642 were cultured for 48 h at 37 °C in Spitzien’s minimal medium supplemented with 50 mM FeCl$_3$. After centrifugation (10,000 rpm for 20 min), the cultured broth was extracted three times with equal volumes of ethyl acetate. Subsequently, the volume of the pooled organic extracts was reduced to ~100 ml in a rotary evaporator at 37 °C, and the remainder was washed twice with 0.1 M sodium citrate buffer (pH 5.5) and water, respectively.

**Siderophore Assay**—For siderophore assay, 200 ml of cells from B. subtilis strains ATCC 21332, 168, and JH642 were cultured for 48 h at 37 °C in Spitzien’s minimal medium supplemented with 50 mM FeCl$_3$. After centrifugation (10,000 rpm for 20 min), the cultured broth was extracted three times with equal volumes of ethyl acetate. Subsequently, the volume of the pooled organic extracts was reduced to ~100 ml in a rotary evaporator at 37 °C, and the remainder was washed twice with 0.1 M sodium citrate buffer (pH 5.5) and water, respectively. The organic layers were dried, and the pellets were resuspended in a small volume (200 ml) of methanol. The resulting suspensions were cleared by centrifugation (21 °C at 13,000 rpm for 5 min), and the supernatants were further analyzed.

**Detection and Analysis of the Siderophore by Ferric Hexadecyltrimethylammonium Bromide-Chrom-Azurol-S (CAS) Assay and Electrospray Ionization Mass Spectrometry (ESI-MS)**—For the detection of siderophore-producing B. subtilis strains, organic extracts of their cultured broth were applied to a CAS solution assay as described by Schwyn and Neilands (40). Additionally, the B. subtilis strains were streaked out on CAS plates and tested for growth and the ability to breakdown the CAS complex. The corresponding plates were prepared as described (40) and supplemented with 0.2% (w/v) casamino acids. Extracts of siderophore-producing strains were further analyzed by ESI-MS on a Hewlett-Packard 1100 series mass spectrometric detector. 5-μl samples were flow-injected into an autosampler at 0.3 ml/min. Scans were taken in negative-ion mode over the m/z range of 100–1200 at a capillary voltage of 5000 V, a fragmenter voltage of 240 V, drying gas (N$_2$) at 250 °C, and a flow rate of 8 liters/min. A solution of 100% acetonitrile and 0.04% (v/v) formic acid was used as carrier solvent for the detection of the peak at m/z 881 (M-H$^+$). For the detection of fragment ions, a carrier solvent of 50% methanol, water, and 0.04% (v/v) formic acid; a capillary voltage of 4000 V, and varying fragmenter voltages (up to 350 V) were used.

**ATP-PP Exchange Assay for DhbE and DhbF Substrate Selectivity**—The ATP-PP exchange reaction was performed as described (41), and reaction mixtures (100-μl final volume) normally contained 2 mM amino acid, 2 mM ATP, and 500 mM enzyme in assay buffer (50 mM HEPES (pH 7.8) and 20 mM MgCl$_2$). Reactions were initiated by the addition of 0.15 μCi of sodium $^{32}$P-orthophosphate (PerkinElmer Life Sciences) and 0.1 mM ATP, at pH 7.8. ATP-PP exchange reactions with DhbF1-A-PCP were carried out at a slightly higher pH value of 8.8. For DhbE, the amino acid concentration had to be lowered to 0.5 mM to avoid nonspecific techniques (27). After ligation and transformation, the corresponding plates were prepared as described by Schwyn and Neilands (40). Additionally, the B. subtilis strains were streaked out on CAS plates and tested for growth and the ability to breakdown the CAS complex. The corresponding plates were prepared as described (40) and supplemented with 0.2% (w/v) casamino acids. Extracts of siderophore-producing strains were further analyzed by ESI-MS on a Hewlett-Packard 1100 series mass spectrometric detector. 5-μl samples were flow-injected into an autosampler at 0.3 ml/min. Scans were taken in negative-ion mode over the m/z range of 100–1200 at a capillary voltage of 5000 V, a fragmenter voltage of 240 V, drying gas (N$_2$) at 250 °C, and a flow rate of 8 liters/min. A solution of 100% acetonitrile and 0.04% (v/v) formic acid was used as carrier solvent for the detection of the peak at m/z 881 (M-H$^+$). For the detection of fragment ions, a carrier solvent of 50% methanol, water, and 0.04% (v/v) formic acid; a capillary voltage of 4000 V, and varying fragmenter voltages (up to 350 V) were used.

**ATP-PP Exchange Assay for DhbE and DhbF Substrate Selectivity**—The ATP-PP exchange reaction was performed as described (41), and reaction mixtures (100-μl final volume) normally contained 2 mM amino acid, 2 mM ATP, and 500 mM enzyme in assay buffer (50 mM HEPES (pH 7.8) and 20 mM MgCl$_2$). Reactions were initiated by the addition of 0.15 μCi of sodium $^{32}$P-orthophosphate (PerkinElmer Life Sciences) and 0.1 mM ATP, at pH 7.8. ATP-PP exchange reactions with DhbF1-A-PCP were carried out at a slightly higher pH value of 8.8. For DhbE, the amino acid concentration had to be lowered to 0.5 mM to avoid nonspecific
Post-translational Modification Assay of the ArCP Domain of DhbB

Analysis of Covalent A-C-Salicylation of Holo-DhbB by DhbE—To test for covalent aminomethylacetylation of holo-DhbB, 50 pmol of carrier protein were incubated with 10 pmol of DhbE, 2 pmol of [14C]salicylate, 1 mM ATP, and 10 mM MgCl₂ in a total volume of 100 μL. Reactions were incubated for 15 min at 37 °C under moderate shaking and quenched by the addition of 800 μL of 10% (v/v) trichloroacetic acid. The trichloroacetic acid precipitate was washed once with 1 mL of 10% trichloroacetic acid, pelleted and dissolved in 200 μL of formic acid, and analyzed by liquid scintillation counting (Packard TriCarb Model 2300TR).

Knockout Mutant—For generation of a dhb knockout mutant, an integration plasmid was generated by replacement of the coding region of the second A-PCP module in dhbB by a kanamycin (kan) resistance cassette in plasmid pJJM305 (Table II). For this purpose, the kan gene was amplified from pDG782 (36) using primers Kan_EcoI and Kan_ApoI. The kan gene fragment was cut with EcoI and AplI and cloned into ApaI/EcoRI-digested pJJM305, resulting in plasmid pJJM201.

Subsequently, a lincomycin/erythromycin resistance-conferring gene (erm) (36) was cloned at the 3' end of dhbB in pJJM201 to facilitate a differentiation between single and double crossover events after integration into the B. subtilis chromosome. The erm cassette was prepared by cutting pDG780 (36) with BamHI, and it was cloned into BamHI-digested pJJM201. The resulting plasmid (pJJM202) was used to disrupt the dhbB gene in B. subtilis ATCC 21332 (srfA⁻) following the protocol of Klein et al. (44). The integration was checked by Southern blot analysis (45) using the PCR fragment of dhbF2-A-PCP and the kan gene as probes (data not shown).

RESULTS

Sequencing of the dhb locus in Different B. subtilis Strains and Sequence Analysis—We first thought that the truncated sequence of dhbF is the reason for the failure of some B. subtilis strains to produce a siderophore. Analysis of the neighboring sequences in B. subtilis strain 168 (46) prompted us to assure that two other reading frames, i.e. yukL and yukM (Fig. 1A), resulted from sequencing errors and might originally be part of a larger dhbF gene.

We therefore sequenced this region from different Bacillus strains. Despite a minor sequence divergence, we found three major sequence errors; these critical regions of the published sequence of B. subtilis strain 168 are aligned in Fig. 2 against our sequence data. For the first error region, we found that one thymidine base was missing at position 3828 (the ATG start codon of dhbB refers to position 1), whereas two adenosine base insertions were annotated at positions 3831 and 3834. This leads to a premature TGA stop codon of dhbB at position 3838. The second error consists of a missing thymidine at position 3921, which leads to a stop codon (TGA) at position 3927. The third error is a C → T transition at position 4406, resulting in the end of yukM.

The corrected sequence results in one larger gene (7146 base pairs) (Fig. 1B) that comprises the entire former dhbF, yukL, and yukM genes as well as the intergenic regions and that encodes a dimodular peptide synthetase of 2378 amino acids (264 kDa). We propose to maintain the designation dhbF for this gene. The sequence has been deposited in the GenBank™/EBI Data Bank under accession number AF184977. From this corrected sequence, we postulate a new dhb operon organization. The genes involved in siderophore biosynthesis would now be more similar to those of the ent cluster from E. coli, as shown in Fig. 1 (B and C). dhbE and dhbB are similar to entE and entB, respectively; but dhbF now encodes a dimodular peptide synthetase with two modules composed of C, A, and PCP domains and a terminal Te domain.

A BLAST search (47) with DhbE revealed the greatest similarity to the stand-alone DHB-AMP ligases EntE from E. coli (47% identity), VibE from Vibrio cholerae (68% identity), PchD from Pseudomonas aeruginosa (52% identity), and YshE from Yersinia pestis (49% identity). DhbB shows the greatest similarity to the bifunctional isochorismatase lyase/ArcCP EntB from E. coli (48% identity). This similarity is also extended to other ArCP domain-containing proteins such as VibB from V. cholerae (46% identity), PchE from P. aeruginosa (38% identity), and VenB from Vibrio vulnificus (45% identity). DhbF shares
the highest similarity with the dimodular actinomycin synthetase II from Streptomyces chrysomallus (45% identity) and the two last modules of the tetramodular pristinamycin III synthetase from Streptomyces pristinae spiralis (41% identity).

Cloning and Expression of dhbB, dhbE, and dhbF—To analyze the activity and substrate selectivity of the proteins encoded by dhbB, dhbE, and dhbF, the genes were amplified with the primers listed in Table I and cloned into C-terminal His 6-tagged vectors pQE60 and pQE70 (see “Experimental Procedures”). Expression of dhbE, dhbB, and dhbF was carried out in E. coli, and all proteins were purified as described under “Experimental Procedures” (Fig. 3). The overproduction of the 264-kDa DhbF protein substantiates the newly determined sequence of dhbF (Fig. 3D). The two A-PCP modules encoded by dhbF were also cloned as single modules (DhbF1-A-PCP and DhbF2-A-PCP) to unambiguously assign their substrate selectivity.

Substrate Selectivity of DhbE—All proteogenic amino acids as well as DHB and salicylate were tested as substrates for DhbE in the ATP-PPi exchange reaction. The protein was highly specific for DHB and salicylate, and no side specificity was detected (Fig. 4). The $K_m$ and $k_{cat}$ values of 7.6 $\mu$M and 167 min$^{-1}$ (Table III) (18, 49), respectively, clearly indicate that DHB is the preferred substrate. These data are in accordance with those reported for EntE and VibE for enterobactin and vibriobactin, respectively (18).2

Post-translational Modification of the ArCP Domain of DhbB by Sfp with [3H]Coenzyme A and Covalent Thioesterification of Holo-DhbB by DhbE—We assumed that DhbB contains the ArCP domain, which covalently binds the activated DHB moiety as a thioester. We first checked the ability of the recombinant apo-DhbB isolated from E. coli to be phosphorylated by the Ppant-transferase Sfp. In fact, Sfp stoichiometrically modified apo-DhbB to holo-DhbB (Fig. 5A, B). The activity was measured by incorporation of [3H]CoASH. In the control experiment without Sfp, no phosphopantetheine was incorporated. By this approach, we proved that DhbB contains an ArCP domain and that it is recognized by the Ppant-transferase Sfp. Using holo-DhbB isolated from E. coli BL21(pREP4-gsp), we assayed the transfer of [14C]salicylate from DhbE to DhbB (radiolabeled DHB is not commercially available). Fig. 5B shows that 98% of DhbB could be acylated.

1 Keating, T. A., Marshall, C. G., and Walsh, C. T. (2000) Biochemistry 39, 15522–15530.
by DhbE after 10 min as judged by the thioster binding assay and scintillation counting of the trichloroacetic acid-precipitable material. Thus, using recombinant DhbE and holodhbB, the initial reactions of the Dhb synthetases could be reconstituted.

**Biochemical Analysis of DhbF**—The entire dimodular NRPS DhbF was produced as a 264-kDa protein, supporting the revised dhbF sequence (Fig. 3D). The separate DhbF modules DhbF1-A-PCP and DhbF2-A-PCP were also produced as individual proteins and were assayed for substrate selectivity using the ATP-PPi exchange reaction. For the DhbF1-A-PCP construct (Fig. 6), a low but specific glycine-dependent ATP-PPi exchange was observed. This reaction was dependent on a slightly higher pH value (pH 8.8) than the standard assay condition of pH 7.8. In contrast, for the second module, DhbF2-A-PCP, a highly specific ATP-PPi exchange dependent on l-threonine was observed (Fig. 7). l-Threonine was not recognized as a substrate, whereas the stereoisomeric 3-allo-Thr covered 47% activity. Upon testing all other proteogenic amino acids, only side specificity for phenylalanine (9%) could be observed. The activation pattern of full-length DhbF in the ATP-PPi exchange assay was, as expected, a sum of the activities from the single domains (data not shown).

**dhbF Disruption Mutant**—To identify the siderophore and to correlate the dhb gene cluster with its synthesis, we generated a dhbF disruption mutant from *B. subtilis* ATCC 21332 (see “Experimental Procedures”). The integration and the double crossover event in the resulting *B. subtilis* strain, JMM405, were proved by Southern blot analysis (data not shown). JMM405 was subsequently investigated for siderophore production.

**Cross-reaction of Surfactin in the CAS Assay**—As described above, most of the known siderophores are able to retract ferric ions from the CAS complex, a process that is monitored by the activity of the siderophore to be retracted ferric ions from the CAS complex, a process that is monitored by the activity. Thus, using recombinant DhbF and holodhbB, the initial reactions of the Dhb synthetases could be reconstituted.

**Biochemical Analysis of DhbF**—The entire dimodular NRPS DhbF was produced as a 264-kDa protein, supporting the revised dhbF sequence (Fig. 3D). The separate DhbF modules DhbF1-A-PCP and DhbF2-A-PCP were also produced as individual proteins and were assayed for substrate selectivity using the ATP-PPi exchange reaction. For the DhbF1-A-PCP construct (Fig. 6), a low but specific glycine-dependent ATP-PPi exchange was observed. This reaction was dependent on a slightly higher pH value (pH 8.8) than the standard assay condition of pH 7.8. In contrast, for the second module, DhbF2-A-PCP, a highly specific ATP-PPi exchange dependent on l-threonine was observed (Fig. 7). l-Threonine was not recognized as a substrate, whereas the stereoisomeric 3-allo-Thr covered 47% activity. Upon testing all other proteogenic amino acids, only side specificity for phenylalanine (9%) could be observed. The activation pattern of full-length DhbF in the ATP-PPi exchange assay was, as expected, a sum of the activities from the single domains (data not shown).

**B. subtilis** strain 6051 and found ethyl acetate suitable for extraction of the putative siderophores.

**Mass Spectrometric Analysis of the Siderophore**—As shown in Fig. 8 (50), ESI-MS in negative-ion mode allowed us to identify the mass of the siderophore by comparing the culture broth extracts from the ΔdhbF mutant JMM405 with those made from the parental strain ATCC 21332 (see “Experimental Procedures” for details). The mass peak at m/z 881.2 ([M-H]-) found in the wild-type extract matches three DHB-Gly-Thr units minus 3H2O (\([M-(C_{99}H_{4}N_{6}O_{18})^{-}]^{1-}\)), which corresponds to a cyclic trimeric ester, (DHB-Gly-Thr)3. The calculated exact mass of \([C_{99}H_{4}N_{6}O_{18}]^{1-}\) is 881.25.

Variations of the fragmenter voltage and carrier solvent induced the following fragmentation patterns (data not shown). (i) m/z 745.2 can be attributed to a single loss of DHB at the amide bond \([M-(DHB)]^{-}\), i.e. \([C_{32}H_{32}N_{2}O_{15}]^{1-}\), with a calculated mass of 745.23. (ii) m/z 587.3 corresponds to (DHB-Gly-Thr)-2 minus 2H2O \([M-(DHB-Gly-Thr-H)]^{-}\), i.e. \([C_{20}H_{24}N_{2}O_{12}]^{1-}\), with a calculated mass of 587.16. Note that no m/z 605, i.e. (DHB-Gly-Thr)-2 minus 1H2O, was found. (iii) m/z 293.1 corresponds to (DHB-Gly-Thr)-1 minus 2H2O \([M-(DHB-Gly-Thr-H)]^{-}\), i.e. \([C_{14}H_{14}N_{2}O_{1}]^{1-}\), with a calculated mass of 293.08. No m/z 311, i.e. (DHB-Gly-Thr)-1 minus 1H2O, was detected. (iv) m/z 249.1 can be attributed to the deacylation of (DHB-Gly-Thr)-2 minus 3CO2. The resulting three fragments have the same mass \((C_{12}H_{14}N_{2}O_{1})^{1-}\) with a calculated mass of 249.09.

Summarizing these mass spectrometric data and the biochemistry of the DhbEBF synthetases, we deduce the siderophore to be a trilactone ((DHB-Gly-Thr)3). In analogy with the enterobactin biosynthesis, we suggest that the function of the terminal thioesterase domain (DhbF-Te) is to catalyze the terminal thioesterase domain (DhbF-Te) is to catalyze the
decarboxylation of (DHB-Gly-Thr)3 minus 3CO2. The resulting three fragments have the same mass \((C_{12}H_{14}N_{2}O_{1})^{1-}\) with a calculated mass of 249.09.

**DISCUSSION**

In this report, we show that the Gram-positive bacterium *B. subtilis* produces the catecholic siderophore bacillibactin, which is structurally related to enterobactin isolated from the Gram-negative bacterium *E. coli*. Both siderophores contain three 2,3-dihydroxybenzoate moieties for octahedral iron complexation that are coupled to a cyclic amino acid core synthe-
sized by multimodular nonribosomal peptide synthetases.

The three genes entE, entB, and entF encoding the EntEBF synthetase complex for enterobactin assembly in E. coli were initially used to identify putative homologs named dhbE, dhbB, and dhbF in B. subtilis. We revised the erroneous dhbF sequence and characterized DhbF as dimodular NRPS in contrast to the monomodular EntF synthetase.

Although the three dhb genes encoding the synthetases for the assembly of bacillibactin reside in one operon, this is not the case for the enterobactin synthetase genes (Fig. 1), where entF is located ~7 kilobases upstream of the entCEBA operon. The homologous vibriobactin synthetase genes are furthermore divided into two gene clusters (51, 52) separated by 1 megabase on the larger of the two chromosomes of V. cholerae (53). Nevertheless, the dhb genes and homologs support par excellence the idea of horizontal gene transfer between Gram-negative and Gram-positive bacteria and genetic rearrangement of NRPS, leading to a variety of structurally related siderophores in different organisms (5). Although the sequence of the DhbEBF modules is collinear to the sequence of the product, some domains need further attention.

The first step in nonribosomal peptide synthesis is the adenylation of the cognate substrate (amino) acid. Determination of the substrate selectivity of the catalyzing A domains can be accomplished either by the ATP-PPi exchange assay or by analysis of the selectivity-conferring residues as guided by the nonribosomal code of A domains from NRPS (41). In the case of DhbEF, both studies independently led to the determination of DHB, Gly, and Thr selectivity for DhbE, DhbF1, and DhbF2, respectively. Intriguingly, regarding activation of DHB and salicylate by DhbE, the selectivity code cannot distinguish between both aryl substrates (Table IV) (41), although the various adenylating enzymes (EntE, VibE, and YbtE) (see Table III) reveal a clear preference for one or the other substrate. Likewise, we were able to show that DhbE is selective for DHB, activating salicylate with a 6-fold lower catalytic efficiency. The same kind of analysis revealed Gly and Thr selectivity for the first and second modules of DhbF, respectively. The activation pattern indicates a stereospecific selectivity for L-Thr as proposed before. L-allo-Thr showed an unexpected high specificity, and Phe only a minor side specificity. All cognate amino acids were tested, but only representative amino acids above background activation are shown. The highest activity was set at 100%.

The second step in nonribosomal peptide synthesis is the condensation of the activated amino acid with an already adenylated amino acid. The 3′-phospho-5-aminoimidazole ribonucleotide (5′-PP(NH)2-ribose-3′-phosphate, PPI) is the phosphate-diester moiety of the active-site adenylated amino acids, and this step is catalyzed by the AC domain.

The third step of nonribosomal peptide synthesis is the transacetylation reaction. This reaction is catalyzed by the transacylase domain of the synthetase. The ADP-acyl moiety is transferred from the enzyme-bound PPI to the acyl-enzyme intermediate, where the acyl-enzyme is then activated by the transacylase domain.

The fourth step in nonribosomal peptide synthesis is the cyclization of the nascent peptide. This reaction is catalyzed by the cyclization domain, which is also part of the synthetase.

The fifth step is the post-translational modification of the nonribosomal peptide. This step can involve the addition of terminal modifications such as esterification, sulfoxidation, or carboxylation. These modifications are catalyzed by the terminal modifying enzymes.

The sixth step in nonribosomal peptide synthesis is the secretion of the final product. This process is mediated by the adjacent secretion system, which allows the export of the siderophore outside the cell into the extracellular environment.

The seventh step is the recognition of the siderophore by the receptor protein. This protein recognizes the terminal modification of the siderophore and binds it, allowing the iron-containing siderophore to be internalized into the cell.

The eighth step in nonribosomal peptide synthesis is the release of the iron-containing siderophore from the receptor protein. This is mediated by the efflux system, which allows the iron to be released into the cytoplasm of the cell.
In contrast, the ATP-PP\textsubscript{i} exchange assay with DhbF2-A-PCP revealed a distinct selectivity for the predicted substrate, l-Thr, and some nonspecific activation of the miscognate substrates l-allo-Thr and l-Phe (47 and 9\%, respectively). These moderately differ from the activation pattern as determined for the homologous VibF (48), and both enzymes show a clear stereospecificity for the activation of an l-configured amino acid (D-Thr, 2\%).

The second step in nonribosomal peptide biosynthesis is the transfer of the activated aminoacyl moiety from the adenylate to the thiol group of the adjacent carrier protein. In this study, this so-called thiolation reaction could be reconstituted \textit{in vitro} for the formation of DHB-S-Ppant-DhbB, which represents the initial step in bacillibactin biosynthesis (Fig. 9) (1). Here, DhbE activates the substrate carboxyl acid DHB as DHB-O-AMP; and subsequently, the aryl moiety is transferred.
to holo-DhbB, yielding the DHB-S-Ppant thioester. The reaction necessarily relies on the presence of a functional HS-Ppant holo-carrier protein, which is usually generated by the action of a dedicated Ppant-transferase. We demonstrated that Sfp, the Ppant-transferase known to be associated with lipopeptide antibiotic surfactin biosynthesis (54), is capable of converting the ArCP domain of DhbB in vitro from the inactive apo form to the HS-Ppant holo form. The obvious presumption that Sfp may also represent the dedicated Ppant-transferase of the bacillibactin system is corroborated by the fact that its in vivo biosynthesis in B. subtilis inevitably depends on the presence of a functional Sfp protein. Notably, the genome project B. subtilis strain 168 carries the complete NRPS gene clusters for the synthesis in B. subtilis to now, the following signature sequences could be observed: the signature of their highly conserved Ppant-binding site. Upstream, it is worthwhile to notice that carrier proteins can be distinguished not only by the moiety they have to carry (acyl, amino acid, peptide), the two carrier protein domains of DhbF were determined by sequence homology. In this classification, the carrier protein of DhbB is clearly of the first class of PCP domains. The significance of having ArCP, whereas the two carrier proteins of DhbF belong to the second class of PCP domains. The ArCP domain of EntB (29) and the PCP domain of EntF (21). Likewise, Sfp has been noticed to be rather nonselective, modifying all kinds of carrier protein domains with high catalytic efficiency (54).

The last steps in nonribosomal peptide biosynthesis consist of substrate condensation, followed by product release catalyzed by the Te domain. The peptide chain is then transferred from the last PCP domain onto a Ser residue within the highly conserved core motif TE (G(HY)S(XG) (10)). Two ways of Te-mediated product release have been described so far: ester hydrolysis by water or nucleophilic attack by amino or hydroxyl groups of the peptide itself, leading to a cyclic lactame or lactone. In the case of enterobactin biosynthesis, mutational analysis of EntF-Te led to reduced product release and enabled Shaw-Reid et al. (23) to identify intermediate reaction products such as (DHB-Ser)₂ and enzyme-bound species such as (DHB-Ser)-S-PCP, (DHB-Ser)₁-O-Te, and (DHB-Ser)₁-O-Te by ESI-MS. These data provide evidence that EntF-Te additionally catalyzes the trimerization of the DHB-Ser units prior to cyclization.

In our case, due to rapid loss of DhbF activity, we were unable to prove these last steps for bacillibactin biosynthesis in vitro. A clue in the determination of the catalytic role of DhbF-Te was the product analysis by ESI-MS (Fig. 8). The ion at m/z 881.2 ([M-H]⁻) and daughter ions clearly define bacillibactin as (DHB-Gly-Thr)₃ tri lactone. Taking our mass spectrometric data together with the assembly line enzymology of EntEBF, we propose the following model for bacillibactin biosynthesis as depicted in Fig. 9. The synthetase holo-DhbEBF activates the substrates DHB (1), Gly, and Thr, which are bound as acyl-S-Ppant intermediates. Subsequent condensation leads to the products DHB-Gly (2) and DHB-Gly-Thr (3) on the corresponding PCP domains. The nucleophilic attack of the Ser hydroxyl group of the Te domain (G(HY)S) leads to an (DHB-Gly-Thr)-O-Te ester (4). The DhbF2-PCP, set free for a second round of DHB-Gly-Thr unit synthesis. Subsequently, the Thr hydroxyl group of 4 attacks a second (DHB-Gly-Thr)-S-PCP (3), leading to a (DHB-Gly-Thr)₁-O-Te intermediate (5). Iteration of this reaction produces the linear (DHB-Gly-Thr)₃-O-Te (6), which is released by cyclization as trilactone (DHB-Gly-Thr)₃ (7) with a corresponding mass of 882. An identical siderophore structure was reported for corynebactin isolated from Corynebacterium glutamicum (48), but neither the genes nor proteins involved in the biosynthesis of this product have been reported.

Acknowledgments—We are indebted to N. Kessler for excellent practical help and to T. Stachenhau and H. Mootz for critical reading of the manuscript. Additionally, we thank Veit Bergendahl for the initial steps in solving some analytical problems.

REFERENCES
1. Griffiths, E. (1991) Biol. Met. 4, 7–13
2. Guerinot, M. L. (1994) Annu. Rev. Microbiol. 48, 743–772
3. Payne, S. M. (1993) Trends Microbiol. 1, 66–69
4. Neilands, J. B. (1983) Adv. Inorg. Biochem. 5, 137–166
5. Quadri, L. E. N. (2000) Mol. Microbiol. 37, 1–12
6. Winkelmann G., C. C. J. (1998) Transition Metals in Microbial Metabolism,

### TABLE IV

| Adenylation domain | Substrate specificity | Residue position according to GrsA Phe⁶ |
|--------------------|----------------------|----------------------------------------|
| DhbE               | DHB/salicylate       | 235 P L P A Q G V V N K                |
| EntE               | DHB/salicylate       | 236 A M P L Q G V V N K                |
| VibE               | DHB/salicylate       | 239 P L P L Q G V V N K                |
| YtbE               | DHB/salicylate       | 278 P L P A Q G V L C K                |
| Consensusb         |                      | 299 P L P A Q G V N K                  |
| DhbF1              | Gly                  | 301 D I L Q L G L I W K                |
| SafA               | Gly                  | 312 D I L Q L G L I W K                |
| CDA2.2             | Gly                  | 322 D I L Q L G L T A K                |
| Consensusb         |                      | 330 D I L Q L G L I W K                |
| DhbF2              | Thr                  | 331 D F W N I G M V H K                |
| SnbC1              | Thr                  | 349 D F W N V G M V H K                |
| CDA1.2             | Thr                  | 357 D F W N V G M H F K                |
| Consensusb         |                      | 374 D F W N I G M V H K                |

⁶ See Ref. 41.

Italics indicate variable residues.
