THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 42, Issue of October 19, pp. 38370–38377, 2001

Printed in U.S.A.

A Polyketide Synthase in Glycopeptide Biosynthesis

THE BIOSYNTHESIS OF THE NON-PROTEINOGENIC AMINO ACID (*S*)-3,5-DIHYDROXYPHENYLGLYCINE*

Received for publication, July 13, 2001, and in revised form, August 2, 2001
Published, JBC Papers in Press, August 8, 2001, DOI 10.1074/jbc.M106580200

Volker Pfeifer‡, Graeme J. Nicholson§, Johannes Ries‡, Jürgen Recktenwald‡, Alexandre B. Schefer§, Riham M. Shawky‡, Joachim Schröder¶, Wolfgang Wohlleben‡‖, and Stefan Pelzer‡

From the ‡Universität Tübingen, Institut für Mikrobiologie and Biotechnologie, Auf der Morgenstelle 28 and §Universität Tübingen, Institut für Organische Chemie, Auf der Morgenstelle 18, 72076 Tübingen and ¶Universität Freiburg, Institut für Biologie II, Schänzlestr. 1, 79104 Freiburg, Germany

Balhimycin, a vancomycin-type antibiotic from *Amycolatopsis mediterranei*, contains the unusual amino acid (*S*)-3,5-dihydroxyphenylglycine (Dpg), with an acetate-derived carbon backbone. After sequence analysis of the biosynthetic gene cluster, one gene, *dpgA*, for a predicted polyketide synthase (PKS) was identified, sharing 20–30% identity with plant chalcone synthases. Inactivation of *dpgA* resulted in loss of balhimycin production, and restoration was achieved by supplementation with 3,5-dihydroxyphenylacetic acid, which is both a possible product of a PKS reaction and a likely precursor of Dpg. Enzyme assays with the protein expressed in *Streptomyces lividans* showed that this PKS uses only malonyl-CoA as substrate to synthesize 3,5-dihydroxyphenylacetic acid. The PKS gene is organized in an operon-like structure with three downstream genes that are similar to enoyl-CoA-hydratase genes and a dehydrogenase gene. The heterologous co-expression of all four genes led to accumulation of 3,5-dihydroxyphenylglyoxylic acid. Therefore, we now propose a reaction sequence. The final step in the pathway to Dpg is a transamination. A predicted transaminase gene was inactivated, resulting in abolished antibiotic production and accumulation of 3,5-dihydroxyphenylglyoxylic acid. Interestingly, restoration was only possible by simultaneous supplementation with (*S*)-3,5-dihydroxyphenylglycine and (*S*)-4-hydroxyphenylglycine, indicating that the transaminase is essential for the formation of both amino acids.

Vancomycin-group glycopeptides are important antibiotics that are frequently used as a last line of defense against methicillin-resistant *Staphylococcus aureus* (1). However, the recent emergence of vancomycin-resistant strains raises the possibility that vancomycin may soon be rendered ineffective over time and that other therapeutic agents must be developed (2). One obvious approach to obtain new antibiotics is the chemical modification of the existing structures to yield semi-synthetic glycopeptides (3). Another attractive alternative is the specific manipulation of the biosynthetic pathway of the producing organism. This, however, requires detailed knowledge of glycopeptide biosynthesis.

Recently, the gene cluster for balhimycin (4) biosynthesis has been identified in *Amycolatopsis mediterranei* DSM5908 by gene inactivation experiments (5), and the previously established genetic system (6) now permits the precise identification of gene functions and also the construction of new strains able to produce modified glycopeptides (5, 7–9).

It has been shown that the aromatic amino acids (2*S*,3*R*)-*m*-chloro-β-hydroxytyrosine and (*R*)-4-hydroxyphenylglycine in the aglycone of vancomycin group antibiotics are derived from tyrosine (10), and all enzymes in the biosynthesis of the latter amino acid were identified recently (11, 12). The backbone of a third unusual amino acid in these antibiotics, (*S*)-3,5-dihydroxyphenylglycine, was shown to be derived from acetate, and a polyketide synthase (PKS)[1] reaction was proposed (10).

Therefore, it was of considerable interest that the biosynthetic gene cluster of balhimycin does not contain genes for typical bacterial PKS (type I or II) but a gene for a protein related to the plant-specific polyketide synthases of the chalcone synthase family (reviewed in Ref. 13). Putative proteins with 20–30% identity with those plant proteins have been discovered in several different bacteria (reviewed in Ref. 14), but the precise enzyme function is known in only one case (15). In this study we describe the function of the chalcone synthase-related polypeptide encoded in the balhimycin biosynthetic gene cluster and the genes involved in the post PKS modification steps, which lead to (*S*)-3,5-dihydroxyphenylglycine.

EXPERIMENTAL PROCEDURES

Chemicals and Radiochemicals—Chemicals were obtained routinely from Sigma-Aldrich or from Merck and were of analytical grade. [2-^{14}C]malonyl-CoA (1.85 GBq/mmol) and [2-^{14}C]acetyl-CoA (1.85 GBq/mmol) were obtained from Amersham Pharmacia Biotech. (*S*)-3,5-dihydroxyphenylglycine was obtained from ICN Biomedicals GmbH, Eschwege, Germany), and 3,5-dihydroxyphenylacetic acid was obtained by enzymatic hydrolysis of 3,5-dihydroxyphenylacetic acid methyl ester with pig liver esterase (Roche Molecular Biochemicals, Mannheim, Germany). Media constituents were obtained from Life Technologies, Inc. and OXOID (Wesel, Germany). Restriction enzymes were from New England Biosystems (Frankfurt on Main, Germany) and MBI Fermentas (St. Leon-Rot, Germany).

Bacterial Strains, Plasmids, and Culture Conditions—The bacterial strains and plasmids used in this study are listed in Table I. Cloning experiments were performed in *Escherichia coli* XL1Blue by standard

* This work was supported in part by Bundesministerium für Bildung und Forschung Grant ZSP 0316500C and European Community Grant QLK3-1999-00650). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) Y16952.

‖ To whom correspondence should be addressed. Tel.: 49-7071-2976944; Fax: 49-7071-295979; E-mail: wowo@biotech.uni-tuebingen.de.

[1] The abbreviations used are: PKS, polyketide synthase; bp, base pair(s); PCR, polymerase chain reaction; HPLC, high performance liquid chromatography; GC-MS, gas chromatography-coupled mass spectrometry; contig, group of overlapping clones; ORF, open reading frame; ID, identity; SM, similarity.

TABLE I
Strains and plasmids used in this study

	Relevant characteristics	Reference
Strains		
E. coli		
XL1Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lac*IqZΔM15Tn*10(Tet^r)*]	16
A. mediterranei		
DSM5908	Balhimycin-producing wild type	4
VP1–2	Balhimycin mutant with an in-frame deletion in the *dpgA* gene	This study
JR1	Balhimycin mutant with an in-frame deletion in the *pgat* gene	This study
S. lividans		
T7	*S. lividans* TK23 (22), *tsr, ble*, harboring the T7-RNA polymerase gene downstream of *tipA*p	J. Altenbucher, Stuttgart, Germany
VP1	*S. lividans* T7 transformed with pVP1	This study
VP2	*S. lividans* T7 transformed with pVP2	This study
VP5	*S. lividans* T7 transformed with pVP5	This study
Plasmids		
cosmid16.1	pOJ446 (22) derivative, contains a fragment balhimycin biosynthetic gene cluster	5
pSP1	pT7/T3-α19, *ermE*, gene replacement vector for transformation of *A. mediterranei*	6
pIFCS	Gene replacement vector pSP1, harboring flanking regions of *dpgA*	This study
pIFbat	Gene replacement vector pSP1, harboring flanking regions of *pgat*	This study
pRSET$_B$	*bla*, T7-expression system	17, 18
pSLE61	pUC19, *tsr*	19
pGM9	*Streptomyces* vector, *aphII*, *ble*, *tsr*	20
pRSETdpgA	pRSET$_B$ harboring the *dpgA* gene	This study
pSLEK7	pSLE61, contains a *Cla*I/*Xba*I fragment with a C-terminal fragment of *dpgA* and *dpgBCD*	This study
pRSETK7	pRSET$_B$ including the *dpgABCD* genes	This study
pVP1	*Hin*dIII fusion of pRSET with pGM9	This study
pVP2	*Hin*dIII fusion of pRSETdpgA with pGM9	This study
pVP5	*Hin*dIII fusion of pRSETK7 with pGM9	This study

procedures (21). Heterologous expression experiments were carried out with *Streptomyces lividans* T7, kindly provided by J. Altenbuchner (Stuttgart, Germany). Enzyme activity was determined after cultivation in CRM medium (103 g of saccharose, 20 g of tryptic soy broth, 10 g of yeast extract, and 10.12 g of $MgCl_2 \cdot 6H_2O$ in 1.0 liter of distilled water) for 24 h at 30 °C and 180 rpm in baffled shake flasks. Thiostrepton and kanamycin were added to a final concentration of 25 mg/liter, where appropriate. *A. mediterranei* strains were grown in R5 medium (22) at 30 °C. Liquid/solid media were supplemented with 50 μg/ml erythromycin to select for strains carrying the integrated antibiotic resistance gene. Supplementation of mutants was carried out with 100 μM 3,5-dihydroxyphenylacetic acid, 41 μM (*S*)-3,5-dihydroxyphenylglycine, and 41 μM (*S*)-4-hydroxyphenylglycine.

Construction of In-frame Deletion Mutants of A. mediterranei DSM5908—For in-frame deletion of *dpgA* in *A. mediterranei*, the fragment IFCS1 (1091 bp) and the fragment IFCS2 (1044 bp) were amplified by PCR. The PCR mixture (100 μl) contained 200 μM each primer, 1.0 μg of template DNA (cosmid16.1), desoxyribonucleoside 5′-triphosphates at a final concentration of 200 μM each (DNA polymerization mix; Amersham Pharmacia Biotech), 10× reaction buffer, and 2.5 units of Herculase™ enhanced DNA polymerase (Stratagene, Amsterdam, The Netherlands). To decrease secondary structures in template DNA, dimethyl sulfoxide at a final concentration of 3% (v/v) was added to the reaction mixture. The primer pairs were: IFCS1.1, 5′-GGC GAG GAA TTC TCG CTG CCG ATG TAC-3′, and IFCS1.2, 5′-GTC GAG GAT ATC GAG GAC CTC GGA CTG-3′; and IFCS2.1, 5′-GGC TCC GAT ATC GAA ATG GCG CTG ATC-3′, and IFCS2.1, 5′-CGC GCA TCT AGA TGC CCT GGT CGA TC-3′. Restriction sites introduced into the sequence are underlined in the primer sequences. After restriction, the amplified DNA fragments were fused over their common *Eco*RV restriction site, resulting in an in-frame transition between the two PCR products. The 5′-*Eco*RI and 3′-*Xba*I restriction sites of the fusion product were used to clone it into the gene replacement vector pSP1, resulting in the plasmid pIFCS (Fig. 1). For transformation of *A. mediterranei* DSM5908, a modified direct transformation method was used as described previously (6). After transformation, the clones harboring the erythromycin selection marker had integrated the pIFCS plasmid over a single cross-over into the chromosome. To obtain an in-frame deletion mutant, a second cross-over event at the other site of the deleted fragment was required. To increase the probability of detecting such a second cross-over event, ultrasound treatment (Branson Sonifier 250, Danbury, CT) was performed with the transformed mycelium in liquid medium. A clone (*A. mediterranei* VP1–2) that had lost erythromycin resistance was selected, and the correct second cross-over event was verified by PCR and Southern hybridization.

For in-frame deletion of *pgat* in *A. mediterranei*, the fragment IFbat1 (1070 bp) and the fragment IFbat2 (1029 bp) were amplified by PCR (same reaction mixture as described above). The following primer pairs were used: IFbat1.1, 5′-TCT AGA TCG CCC GGT AGG CCT-3′, and IFbat1.2, 5′-A GAT CTC CGT ACA AGA GCG TG-3′; and IFbat2.1, 5′-A GAT CTC CGG GAA GCT GCT CCT C-3′ and IFbat2.1, 5′-GCA TGC GCC GGC TGC CCT TCC-3′.

As described for construction of pIFCS, the two fragments IFbat1 and IFbat2 were fused over their common restriction site *Bgl*II and cloned into the gene replacement vector pSP1 over the 5′-*Xba*I and the 3′-*Sph*I site of the fusion product of IFbat1 and IFbat2, resulting in the plasmid pIFbat (Fig. 1). After transformation of *A. mediterranei*, the same selection procedure was used to obtain a double cross-over event in the *A. mediterranei* chromosome with the plasmid pIFbat, resulting in the strain *A. mediterranei* JR1.

Construction of S. lividans T7 Expression Strains—To obtain plasmids replicating in *E. coli* as well as in *S. lividans*, all expression plasmids were generally constructed by *Hin*dIII fusion of pRSET$_B$ derivatives (replicative in *E. coli*) with the streptomycetes vector pGM9. To prevent plasmid instability, all the genes in the resulting fusion plasmids were oriented in the same direction. The control plasmid pVP1 was constructed by fusing just the pRSET$_B$ vector with pGM9. For construction of the expression plasmid pVP2, PCR was performed with the primer pair DPGA1 (5′-CAT **ATG** GGG GTG GAT GTA TCG-3′) and DPGA2 (5′-AAGCTT **TCA** CCA TTG GAT CAG CGC-3′) and with cosmid16.1 as template (Fig. 1). The deduced start and stop codons of *dpgA* are bold face in the primer sequences. The amplified 1128 bp fragment, coding for *dpgA*, was cloned into the pRSETb vector over the primer-derived *Nde*I-*Hin*dIII restriction sites. The fusion of the resulting plasmid pRSETdpgA with pGM9 yielded pVP2.

The co-expression plasmid pVP5 was constructed by cloning a 7-kilobase pair *Cla*I-*Xba*I fragment containing *dpgBCD* and a C-terminal *dpgA* fragment from cosmid16.1 into pSLE61 (*Cla*I-*Xba*I), resulting in pSLEK7. The 7-kilobase pair *Cla*I-*Hin*dIII fragment from pSLEK7 was then cloned into pRSETdpgA (*Cla*I-*Hin*dIII), replacing 1102 bp of the *dpgA* PCR insert and creating a full *dpgABCD* fragment. *S. lividans* T7 was transformed with the resulting plasmids pVP1, pVP2, and pVP5 by protoplast transformation as described (22).

HPLC Analysis of Culture Filtrates—Culture filtrates of induced expression and control strains of *S. lividans* or *A. mediterranei* (wild type or mutant strain JR1) were extracted with ethyl acetate at pH 1.0. After evaporation of the solvent, the residual material was redissolved in methanol to get a 25-fold concentration relative to the original culture volume. The concentrated extracts (20 μl) were separated via HPLC (Thermo Request System, Egelsbach, Germany) using a Nucleosil C18 column at a flow rate of 2 ml/min. The solvents used were solvent A (double distilled H_2O, 0.1% phosphoric acid) and solvent B

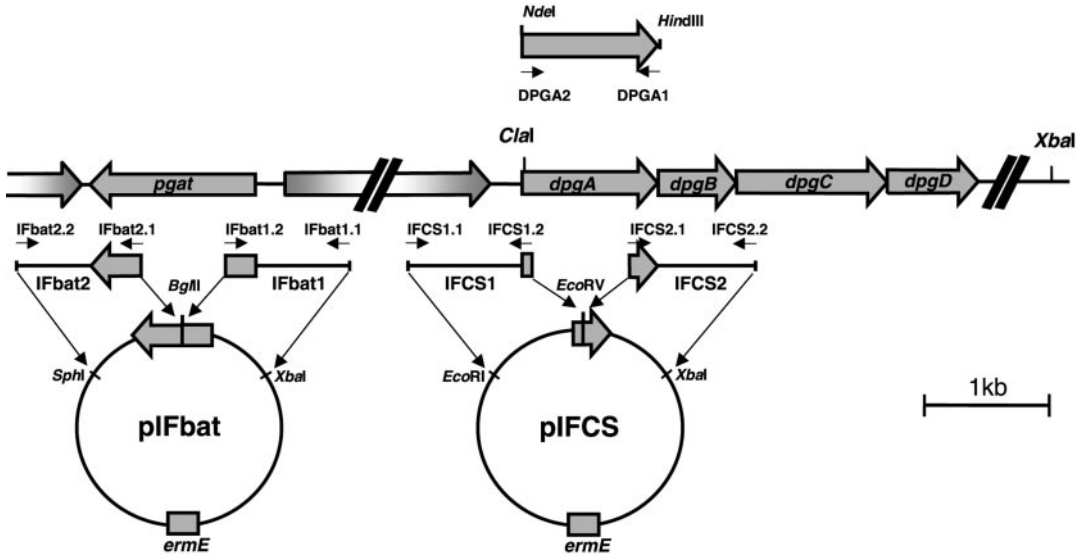


FIG. 1. **Genetic organization of the genes described in this study and construction of gene replacement plasmids.** The start codon of *pgat* is located 11,769 bp upstream of the start codon of *dpgA*. The gene replacement plasmids pIFbat and pIFCS were constructed by cloning two flanking PCR fragments into the gene disruption vector. *kb*, kilobase.

(acetonitrile). The profile for separation was a linear gradient from 100% A/0% B to 90% A/10% B at run-time 13.3 min, a linear gradient to 0% A/100% B at run-time 16 min, and a re-equilibration time of 4 min at 100%A/0% B. The elution was monitored with a multi-wavelength detector between 200 and 350 nm.

Bioassay for Detection of Antibiotically Active Balhimycin—Balhimycin production was determined by bioassays with *Bacillus subtilis* ATCC6633 as the test organism. Filter disks of 5-mm diameter soaked with liquid medium were placed on balhimycin test medium (3.0 g of KH_2PO_4, 7.0 g of K_2HPO_4, 0.5 g of sodium citrate 2-hydrate, 0.1 g of $MgSO_4 \cdot 7\,H_2O$, 1.0 g of $(NH_4)_2SO_4$, 2.0 g of glucose, and 17.0 g of agar in 1.0 liters of distilled water, inoculated with a spore suspension of *B. subtilis* after sterilization and cooling).

Protein Analysis—Protein concentrations were determined by the Bradford method (23) using bovine serum albumin as a standard. SDS-polyacrylamide gel electrophoresis was carried out according to the method of Laemmli (24), and protein bands were stained with Coomassie Brilliant Blue G-250.

DpgA Activity Determination in Vitro—Induced cells were washed and resuspended in assay buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM dithioerythritol, final pH 7.5) on ice. Resuspended cells (4 ml of assay buffer/1 g wet weight of cells) were broken by passage through a French press cell, and the cell debris was removed by centrifugation (20 min at 25,000 × *g*). A 100-μl reaction mixture contained 20 μl (150 μg of protein) and additions as detailed in the figure legends. Reactions mixtures were incubated at 30 °C for 1 h and stopped by the addition of 5 μl of concentrated HCl. The mixtures were then extracted 2 times with 200 μl of ethyl acetate. After evaporation of the solvent, the resultant material was dissolved in 10 μl of methanol and analyzed by thin layer chromatography (Merck TLC aluminum sheets Silica gel 60 F_{254}, solvent, ethyl acetate:petrol ether:acetic acid, 9:9:1, v/v/v). The dried plates were developed by 1-h exposure with a phosphorimaging plate to the thin layer plate followed by a quantitative scan with a phosphorimaging plate reader (Fujifilm Bio-imaging Analyzer System BAS-1800 II, Raytest GmbH, Straubenhardt, Germany).

Structure Analysis of the Enzyme Products—Gas chromatography-coupled mass spectrometry (GC-MS) analysis was carried out with a GC-MS system of Hewlett-Packard (Waldborn, Germany; HP6890 (GC) coupled with HP5973 (MS); column, H4–5MS (30 m × 250 μm, 0.25-μm film thickness)) The substances to be analyzed were derivatized by trimethylsilylation with *N*,*O*-bis-(trimethylsilyl)trifluoracetamide in pyridine (30 min, 60 °C). ^{1}H NMR of the purified component with the retention time of 2.6 min was carried out with a AMX600 NMR spectrometer (Bruker, Karlsruhe, Germany).

Nucleic Acid Sequence Data and Sequence Analysis—Sequence data of the genes described in this study were obtained by standard sequencing procedures (25) using cosmid16.1 as template. Alignment of sequence contigs and examination for open reading frames (ORFs) were performed by applying the programs gap4 and nip4 (26). The NCBI BLASTP (27) net service for homology search was used with the following parameters: non-redundant protein data base; expect threshold 10; word size 3; blosum62 matrix; gap costs, existence 11 and extension 1.

RESULTS

Identification of a Polyketide Synthase Gene—To identify the genes responsible for the biosynthesis of the 3,5-dihydroxyphenylglycine residue in balhimycin, we focused our search in the balhimycin biosynthetic gene cluster on polyketide synthase-like genes. Typical PKSs of either type I or II were absent: BLASTP searches with the used parameters resulted in no significant similarities over the first 1000 hits of every detected ORF. But we detected a gene predicting a protein of significant similarity with the plant polyketide synthases of the chalcone synthase type. This gene was called *dpgA* because it is the first gene in an operon-like structure that plays the key role in dihydroxyphenylglycine biosynthesis (*dpg*). The highest identity score with functionally characterized proteins was obtained with the plant PKS GHCHS2 from *Gerbera hybrida* (26% identity (ID)/44% similarity (SM)), which uses acetyl-CoA and two malonyl-CoA to synthesize the backbone of the plant secondary metabolite gerberin, a 2-hydroxy-pyrone derivative (28). In addition to the significant overall similarity, DpgA contained in the expected positions the key amino acids necessary for activity of these plant-specific type PKSs (29), suggesting that bacterial protein shared basic similarities in the function with the plant enzymes. DpgA also was very similar (95% ID/97% SM) to the functionally not-characterized gene product of *orf27* from the biosynthetic gene cluster of the glycopeptide chloroeremomycin (30), which has the same backbone as balhimycin. High similarities (45–49%) were also obtained to functionally uncharacterized sequences from *Mycobacterium tuberculosis* (Pks11 and Pks18) (31) and *Streptomyces coelicolor* (32); these proteins are also considered as chalcone synthase-related PKSs (14).

Analysis of Genes Located Downstream of dpgA—Downstream of *dpgA* the biosynthetic gene cluster of balhimycin contains three open reading frames named *dpgB* (654 bp), *dpgC* (1299 bp), and *dpgD* (804 bp). The start (GTG) and stop (TGA) codons of the four genes overlap with the sequence GTGA. The translational coupling suggests that these genes are arranged in an operon structure (Fig. 1). The alignment of the DpgB sequence with protein sequence data bases showed similarity (24% ID/43% SM) to an enoyl-CoA hydratase with

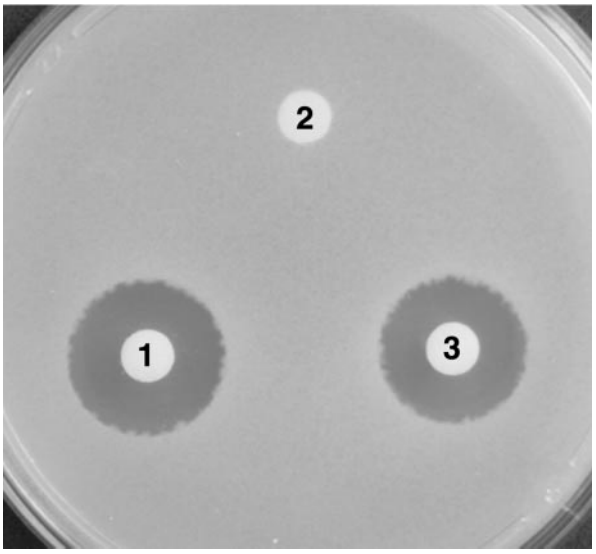


FIG. 2. **Bioassay of *A. mediterranei* strains.** Each 40-μl culture filtrate (60 h incubation at 30 °C) was analyzed on agar plates with balhimycin-sensitive *B. subtilis*. *1*, wild type strain; *2*, mutant strain VP1–2; *3*, mutant strain VP1–2 supplemented with 3,5-dihydroxyphenylacetic acid (100 μM).

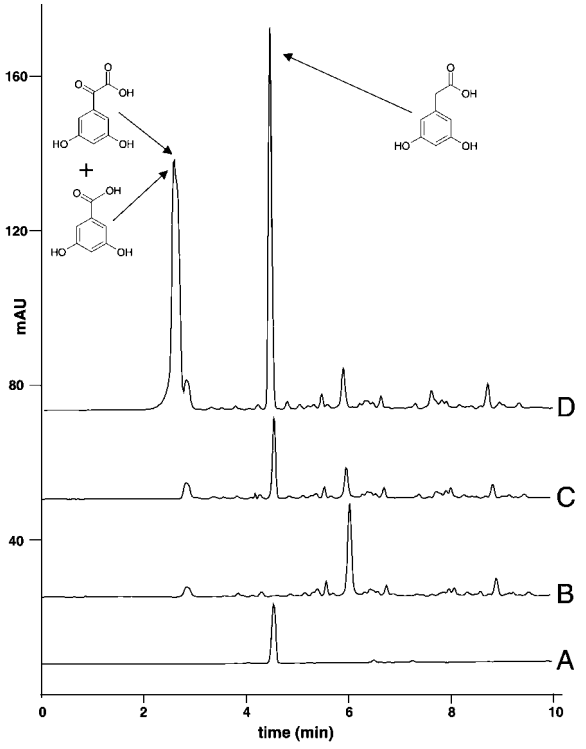


FIG. 3. **HPLC analysis of culture filtrate extracts of *S. lividans*.** *A*, authentic 3,5-dihydroxyphenylacetic acid; *B*, control strain VP1; *C*, DpgA expression strain VP2; *D*, co-expression strain VP5. HPLC traces are recorded at 280 nm. *mAU*, milliabsorbance units.

3-hydroxybutryl-CoA dehydratase activity from *Clostridium acetobutylicum* (33).

The N-terminal 160 amino acid residues of DpgC revealed no similarity to described protein sequences. Its C terminus (from amino acid residue 161), however, displayed a high degree of similarity (46% ID/62% SM) to Orf6, a 3- and 4-hydroxybutyrate dehydrogenase from an uncultured bacterium (34), using NAD^+ to yield 3- or 4-oxobutyrate. Lower scores of similarity (≤49% SM) were obtained to enoyl-CoA hydratases (highest score obtained for a crotonase from *Thermoanaerobacterium thermosaccharolyticum*, accession number P97087).

The alignment of DpgD resulted in a significant similarity (40% ID/54% SM) to an enoyl-CoA hydratase, as well namely, to a carnitine racemase from *E. coli* (35). Furthermore, DpgB and DpgD resemble each other (32% SM). Again, very high similarities (>91% SM) were found between DpgBCD and the corresponding deduced proteins Orf28, Orf29, and Orf30 of the biosynthetic gene cluster of the glycopeptide chloroeremomycin (30).

Analysis of a dpgA In-frame Deletion Mutant—A *dpgA* deletion mutant of *A. mediterranei* was used to test whether the gene was necessary for balhimycin formation. We constructed an in-frame deletion because the analysis of the operon suggested a transcriptional and translational coupling (see "Discussion"). The genotype of the mutant was verified by PCR and Southern hybridization (data not shown). The mutant strain *A. mediterranei* VP1–2 has a 954-bp chromosomal deletion in the 1119-bp-coding region, resulting in the deletion of amino acid residues 48–365 of the DpgA protein. The balhimycin phenotype of the VP1–2 mutant strain was analyzed in a bioassay, with the glycopeptide-sensitive *B. subtilis* as the indicator strain. In contrast to the wild-type strain, the mutant was unable to produce a biologically active substance, demonstrating that DpgA is involved in balhimycin biosynthesis. Supplementation of the VP1–2 mutant strain with 3,5-dihydroxyphenylacetic acid restored antibiotic formation (Fig. 2), indicating that this compound is either a direct intermediate of balhimycin biosynthesis or a derivative of an intermediate.

Heterologous Expression of dpgA in S. lividans—Several attempts to express DpgA in *E. coli* yielded insoluble protein. We therefore decided to express DpgA in *S. lividans* T7, which contains a chromosomal copy of a thiostrepton-inducible T7-RNA polymerase gene. The bacteria were transformed with the DpgA expression plasmid pVP2, which carries the PCR-amplified coding region of *dpgA* downstream of the T7-promoter (details under "Experimental Procedures"). By SDS-polyacrylamide gel electrophoresis analysis of *S. lividans* VP2, a band was detected at 39.5 kDa that corresponded to the calculated molecular mass of DpgA. This protein was absent in the control strain *S. lividans* VP1 (not shown). Ethyl acetate extracts of the culture filtrates were prepared and analyzed by HPLC to investigate whether the DpgA expression resulted in a changed pattern of excreted compounds. Comparison of the chromatograms revealed a new peak in the culture extract of *S. lividans* VP2 (Fig. 3*C*, retention time 4.5 min) at the same position as authentic 3,5-dihydroxyphenylacetic acid (Fig. 3*A*). The identity was confirmed by UV-visible spectroscopy and GC-MS analysis in comparison with the authentic substance.

DpgA Activity Determination in Vitro—3,5-Dihydroxyphenylacetic acid biosynthesis can be explained by a PKS reaction that involves acetyl-CoA or malonyl-CoA as starter substrates and three condensation reactions with malonyl-CoA (Fig. 4).

We used extracts from *S. lividans* VP2 (DpgA expression) and VP1 (control) to test whether the protein synthesized 3,5-dihydroxyphenylacetic acid and whether acetyl-CoA or malonyl-CoA or both were used in its biosynthesis. Representative results are summarized in Fig. 5. Incubations with radioactive malonyl-CoA led to 3,5-dihydroxyphenylacetic acid formation with VP2 but not with VP1, indicating that DpgA was responsible for the activity. More importantly, the result showed that the addition of other CoA esters was not necessary. This did not exclude the possibility that residual acetyl-CoA in the extracts participated in the reaction, and therefore, a series of additional experiments were carried out to investigate that possibility. Experiments with radioactive acetyl-CoA and unlabeled malonyl-CoA led to barely detectable product formation (Fig. 5, experiment *2*), indicating that acetyl-CoA was not used to any

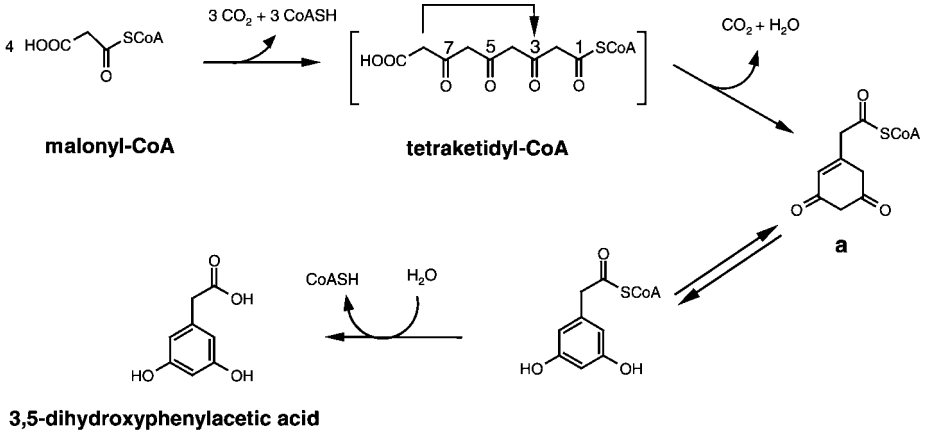
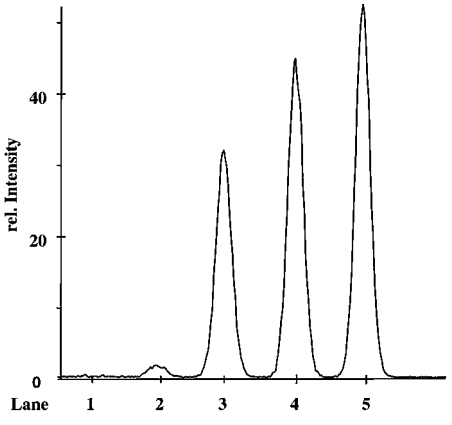


FIG. 4. **Proposed reactions of the DpgA expression strain *S. lividans* VP2.** *a*, see "Discussion."



Assay no.	1	2	3	4	5
Protein	VP1	VP2	VP2	VP2	VP2
Acetyl-CoA	-	-	-	+	-
[2-^{14}C]Acetyl-CoA	+	+	+	-	-
Malonyl-CoA	-	+	-	-	-
[2-^{14}C]Malonyl-CoA	+	-	+	+	+

FIG. 5. **Synthesis of 3,5-dihydroxyphenylacetic acid *in vitro* from the possible starter substrates acetyl-CoA and malonyl-CoA.** The data indicate the amount of radioactivity quantified at the position of 3,5-dihydroxyphenylacetic acid in TLC plates, analyzing the products of the *in vitro* reactions. *VP1*, control extract; *VP2*, extract from *S. lividans* cells expressing DpgA. The concentrations of the CoA esters were 10 μM for acetyl-CoA and 20 μM for malonyl-CoA. The radioactive substrates supplied 1235 Bq to the incubations.

appreciable extent, even considering that only one molecule (compared with three malonyl-CoA) would be incorporated into the product. Moreover, supplementation of incubation mixtures containing radioactive malonyl-CoA with additional acetyl-CoA (either unlabeled, experiment *3*, or labeled, experiment *4*) did not lead to increased formation of radioactive product, but even to a more or less pronounced reduction. The sum of these results demonstrated that DpgA is a PKS synthesizing 3,5-dihydroxyphenylacetic acid solely from malonyl-CoA.

Co-expression of dpgABCD in S. lividans T7—3,5-Dihydroxyphenylacetic acid was shown to be the product of the enzymatic conversion of four molecules of malonyl-CoA with DpgA. To yield the α-amino acid 3,5-dihydroxyphenylglycine, a further modification of the α-position of 3,5-dihydroxyphenylacetic acid is necessary. From our alignment studies we deduce that DpgB and DpgD may catalyze different isomerization reactions by first adding and then eliminating water and that DpgC may possess a dehydrogenase activity. Therefore, a biosynthetic pathway was postulated as depicted in Fig. 6 from 3,5-dihydroxyphenylacetic acid to 3,5-dihydroxyphenylglyoxylic acid, the corresponding 2-oxo acid of 3,5-dihydroxyphenylglycine. To examine the hypothesized role of DpgBCD, the co-expression plasmid pVP5 was used. pVP5 contains *dpgABCD* under control of the T7 promoter and was constructed as described under "Experimental Procedures." *S. lividans* T7 was transformed with pVP5, resulting in expression strain *S. lividans* VP5. This strain was cultivated and induced as described above. HPLC analysis of the ethyl acetate extracts prepared from induced cultures demonstrated the presence of 3,5-dihydroxyphenylacetic acid and a new compound at a retention time of 2.6 min (Fig. 3*D*). To elucidate the chemical structure of the corresponding compound, the substance was purified by preparative HPLC and analyzed by GC-MS and ^{1}H NMR. GC-MS of the HPLC fraction revealed two substances, identified as dihydroxyphenylglyoxylic acid and dihydroxybenzoic acid (probably a shunt product, see "Discussion") by the MS fragmentation pattern. ^{1}H NMR analysis of the purified substances confirmed the predicted structures and revealed a 3,5 substitution pattern for the aromatic hydroxyl groups of both compounds (data not shown).

Analysis of a pgat In-frame Deletion Mutant of A. mediterranei—For the synthesis of the amino acid (*S*)-3,5-dihydroxyphenylglycine, the product of DpgA, DpgB, DpgC, and DpgD, 3,5-dihydroxyphenylglyoxylic acid, has to be transaminated. Therefore, the presence of a corresponding transaminase gene within the balhimycin biosynthetic gene cluster has to be postulated. Indeed, two genes similar to aminotransferases are localized in the cluster. One aminotransferase gene is probably involved in the amination of dehydrovancosamine (36), the second one, *pgat* (phenylglycine aminotransferase), corresponds to the 4-hydroxyphenylglycine aminotransferase gene *hpgT* (12) of the chloroeremomycin biosynthetic gene cluster. We, therefore, assumed that Pgat may function as the dihydroxyphenylglycine aminotransferase. *Pgat* was inactivated by an in-frame deletion in *A. mediterranei*, as described under "Experimental Procedures." PCR and Southern hybridization verified the correct mutation (data not shown). The resulting mutant strain *A. mediterranei* JR1 lacks a chromosomal DNA fragment of 657 bp coding for the amino acid residues 94–312 of Pgat. The remaining open reading frame encodes an internally truncated protein of 209 amino acid residues. In the bioassay, the mutant strain *A. mediterranei* JR1 was not able to produce an antibiotically active compound, confirming the participation of Pgat in balhimycin biosynthesis (Fig. 7, filter *1* and *2*).

To elucidate the specific function of Pgat, the biosynthetic intermediates accumulated in the mutant were analyzed by HPLC of the *A. mediterranei* JR1 culture filtrate (Fig. 8). In the wild type strain a peak at 4.5 min could not be verified as 3,5-dihydroxyphenylacetic acid by UV-visible spectrum. In contrast, 3,5-dihydroxyphenylacetic acid was detected in the mu-

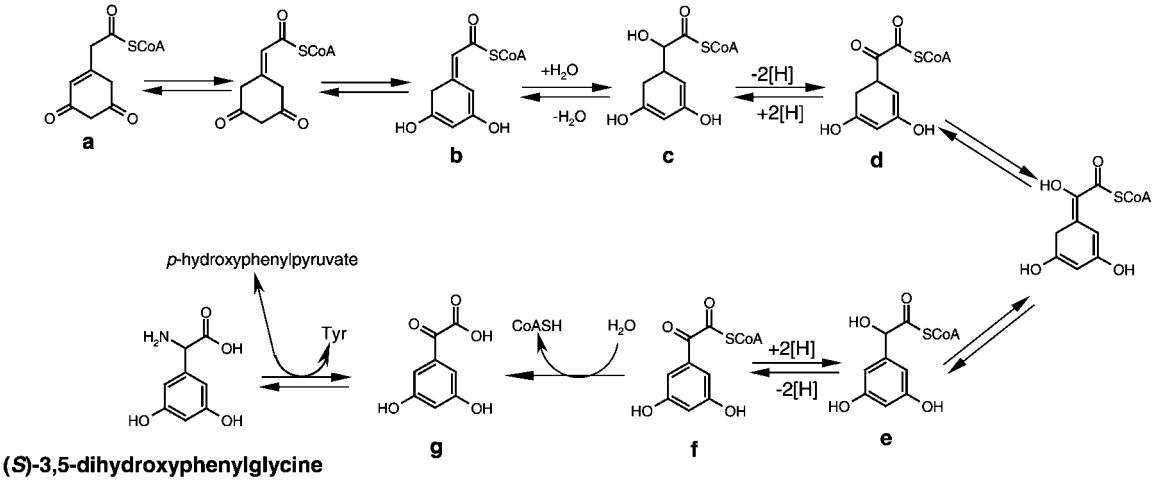


FIG. 6. **Putative biosynthetic pathway of the non-proteinogenic amino acid (*S*)-3,5-dihydroxyphenylglycine.** *a–g*, see "Discussion."

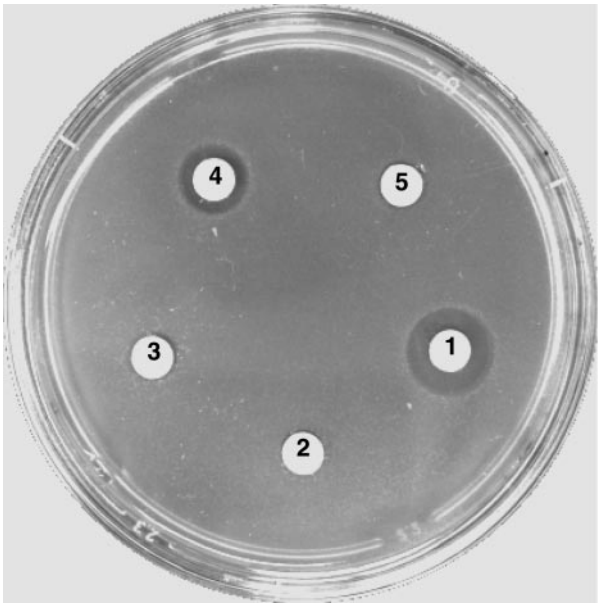


FIG. 7. **Bioassay of *A. mediterranei* culture filtrates (20 μl) after 60 h of incubation at 30 °C.** *1*, wild type strain, 2–5 mutant strain JR1; *2*, without supplementation; *3*, supplemented with (*S*)-3,5-dihydroxyphenylglycine; *4*, simultaneously supplemented with (*S*)-3,5-dihydroxyphenylglycine and (*S*)-4-hydroxyphenylglycine; *5*, supplemented with (*S*)-4-hydroxyphenylglycine. The final concentration of supplemented amino acids was 41 μM.

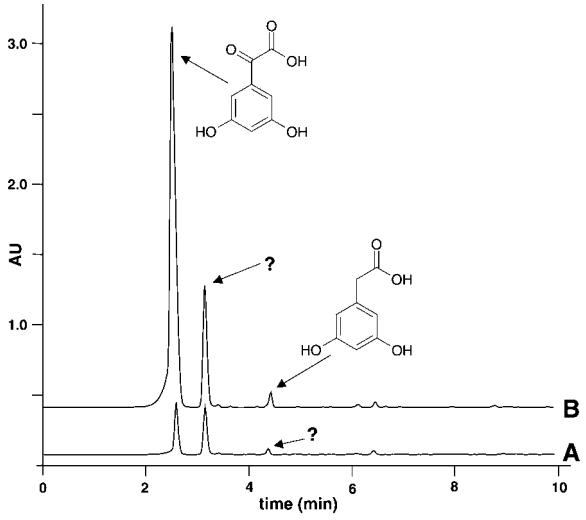


FIG. 8. **HPLC analysis of *A. mediterranei* culture filtrates.** *A*, wild type; *B*, mutant strain JR1. HPLC traces were recorded at 280 nm. *AU*, absorbance units.

tant strain (verified by GC-MS). Furthermore, the *pgat* mutant strain accumulated approximately a 10-fold excess of the dihydroxyphenylglycine precursor 3,5-dihydroxyphenylglyoxylic acid compared with the wild type strain, indicating a block in the subsequent conversion of this intermediate. This clearly proves the participation of Pgat in dihydroxyphenylglycine biosynthesis. On the other hand, overexpressed HpgT was shown to catalyze the conversion of 4-hydroxyphenylglyoxylic acid as well as phenylglyoxylic acid into their corresponding amino acids 4-hydroxyphenylglycine and phenylglycine, respectively (12). Because of this lack of specificity, we assumed that both HpgT and Pgat are capable of transaminating hydroxyphenylglycine and dihydroxyphenylglycine. Therefore, an accumulation of the described precursors of 4-hydroxyphenylglycine, 4-hydroxymandelic acid and 4-hydroxyphenylglyoxylic acid (12), can be supposed for a Pgat mutant. However, by HPLC analysis of the culture filtrate of the mutant JR1, none of the precursors were detectable. Only in the wild-type strain were very small amounts of 4-hydroxymandelic acid detected by HPLC (peak not visible in Fig. 8).

To clarify whether Pgat is involved in the transamination of 4-hydroxyphenylglycine and 3,5-dihydroxyphenylglycine or only in the transamination of one of both amino acids, feeding experiments with *A. mediterranei* JR1 were carried out (Fig. 7, filter *3–5*). Only simultaneous feeding of the mutant strain with both amino acids resulted in the production of balhimycin. In contrast, the supplementation solely with either (*S*)-3,5-dihydroxyphenylglycine or (*S*)-4-hydroxyphenylglycine did not restore the balhimycin production. Therefore a double function can be supposed for this enzyme.

DISCUSSION

The balhimycin biosynthetic gene cluster contains the *dpgA* gene, which is similar to plant polyketide synthase genes and is a member of the new type of bacterial polyketide synthases (14). The genetic organization of *dpgA* with *dpgB*, *dpgC*, and *dpgD* and the translational coupling suggested a functional connection of the four genes. This suggestion was supported by the similarity of all four deduced proteins to enzymes that utilize coenzyme A derivatives as substrates.

Since we are able to genetically manipulate the balhimycin biosynthesis in the producing *A. mediterranei* strain, the function of DpgA could be shown by two independent experiments;

i) an in-frame deletion mutant of *dpgA*, which was not able to produce any antibiotically active compound, was complemented with 3,5-dihydroxyphenylacetic acid, and ii) heterologous expression of *dpgA* in *S. lividans* resulted in the synthesis of 3,5-dihydroxyphenylacetic acid *in vivo* and *in vitro*.

The *in vitro* assays with heterologously expressed DpgA and with ^{14}C-labeled acetyl- and malonyl-CoA demonstrated that only malonyl-CoA was incorporated into 3,5-dihydroxyphenylacetic acid. To obtain 3,5-dihydroxyphenylacetic acid, two different ring closure reactions (C8 → C3 or C2 → C7 of the tetraketidyl-CoA in Fig. 4) are possible, but only the C8 → C3 way leads to the CoA-activated 3,5-dihydroxyphenylacetic acid, which is likely to be the substrate of the subsequent reactions (see below). DpgA is the second enzyme of this new type of prokaryotic polyketide synthases (14) that has now been biochemically characterized. DpgA as well as the previously described RppA of *Streptomyces griseus* (15) use the same substrate, malonyl-CoA, to synthesize the completely different reaction products 3,5-dihydroxyphenylacetic acid and 1,3,6,8-tetrahydroxynaphthalene. In contrast to RppA, DpgA uses only four instead of five malonyl-CoA molecules to synthesize its product. In addition, the ring closure mechanism is completely different; RppA builds up a naphthalene ring system via one aldol condensation and one decarboxylating Claisen ester condensation, whereas DpgA performs only one ring closure reaction via a decarboxylating aldol condensation. Thus, the low similarity (22%) of DpgA to RppA on amino acid level is not surprising.

Heterologous expression of the gene cassette *dpgABCD* in *S. lividans* led to the production of 3,5-dihydroxyphenylglyoxylic acid, indicating that the three gene products DpgB-D are required to convert the reaction product of DpgA into the 2-oxo acid, the direct precursor of the amino acid 3,5-dihydroxyphenylglycine. Since all four enzymes DpgA-D show similarity to enzymes using CoA derivatives as substrates, a biosynthetic pathway with CoA-activated intermediates is more likely than a pathway with unactivated carboxylic acids. The free carboxylic acid intermediates detected in our experiments are probably derived from hydrolysis of accumulated CoA intermediates (enzymatically or spontaneously).

The presence of 3,5-dihydroxybenzoic acid in addition to 3,5-dihydroxyphenylglyoxylic acid in the culture filtrate of *S. lividans* VP2 was probably the result of 3,5-dihydroxyphenylglyoxylic acid degradation, either by an oxidative decarboxylation or hydrolysis reaction. As postulated in Fig. 6, the additionally produced 3,5-dihydroxyphenylacetic acid in *S. lividans* VP5 supports the assumption of equilibrium reactions from (Fig. 6*a*) to 3,5-dihydroxyphenylglyoxylic acid, and it is tempting to speculate on a pathway channeled in a multienzyme complex.

The final step of the 3,5-dihydroxyphenylglycine biosynthetic pathway is a transamination. The biosynthesis of the non-proteinogenic amino acids of the vancomycin-type backbone requires two different transaminations, one to the amino acid discussed here and the second to 4-hydroxyphenylglycine. However, the biosynthetic gene cluster contains only one gene for a predicted aminotransferase (*pgat*), and other experiments with a closely related gene cluster indicated that this gene is involved in 4-hydroxyphenylglycine biosynthesis (12). Our experiments now show that Pgat is essential for 3,5-dihydroxyphenylglycine formation, based on the results that the precursor 3,5-dihydroxyphenylglyoxylic acid accumulated after *pgat* deletion, and that supplementation with the amino acid was necessary to restore antibiotic formation. Restoration also required the addition of 4-hydroxyphenylglycine, indicating that Pgat serves in the biosynthesis of both amino acids. Remarkably, no accumulation of 4-hydroxyphenylglyoxylic acid, the precursor expected for 4-hydroxyphenylglycine was observed. This can be explained by the reaction sequence proposed for the biosynthesis of that amino acid; the transaminase reaction with 4-hydroxyphenylglyoxylic acid uses tyrosine as the amino group donor, and the deaminated product 4-hydroxyphenylpyruvic acid is the precursor source for the formation of 4-hydroxyphenylglyoxylic acid. The Pgat inactivation therefore would be expected to block the entire pathway and, thus, also the accumulation of the direct precursor.

In conclusion, the results of this study support the above postulated biosynthetic pathway for the non-proteinogenic amino acid 3,5-dihydroxyphenylglycine. The first step is the DpgA-catalyzed condensation of four molecules of malonyl-CoA via a hypothetical tetraketide intermediate to 3,5-dihydroxyphenylacetyl-CoA (Fig. 4). By the following hydratase reaction (catalyzed by DpgB or DpgD), a hydroxy group is inserted at the α position of the molecule (Fig. 6*b*) by simultaneous loss of the aromatic status of the ring structure, resulting in *c*. The oxidation of the hydroxy group, probably catalyzed by the dehydrogenase activity of the C-terminal part of DpgC, possibly with $NAD(P)^+$ as cofactor, leads to *d*. The aromatic status can now be restored with an isomerization of *d* via a hydratase/dehydratase activity (DpgB or DpgD), leading to the CoA thioester of 3,5-dihydroxymandelic acid (*e*), which is structurally very similar to *c* and, therefore, could also serve as the substrate of DpgC. The resulting CoA-activated 3,5-dihydroxyphenylglyoxylic acid (*f*) can be hydrolyzed to the free acid (*g*), perhaps by the N-terminal part of DpgC. Finally, transamination by Pgat leads to (*S*)-3,5-dihydroxyphenylglycine.

The biosynthetic pathway unraveled in this study is the first example for the participation of a polyketide synthase mechanism in the biosynthesis of an α-amino acid. The described participation of four co-translated enzymes may suggest an organization of the four proteins as a multienzyme complex as previously shown for prokaryotic type II polyketide synthases (37).

Acknowledgments—Riham Shawky acknowledges a scholarship from the Egyptian Ministry of Higher Education. We thank J. Altenbuchner for providing the expression strain *S. lividans* T7. We also thank R. Süssmuth and co-workers for valuable discussions and D. Fink for critical reading of the manuscript.

REFERENCES

1. Foldes, M., Monro, R., Sorrell, T. C., Shankar, S., and Toohey, M. (1983) *J. Antimicrob. Chemother.* **11,** 21–26
2. Tenover, F. C. (1999) *J. Hosp. Infect.* **43,** 3–7
3. Cooper, R. D. G., Snyder, N. J., Zweifel, M. J., Staszak, M. A., Wilkie, S. C., Nicas, T. I., Mullen, D. L., Butler, T. F., Rodriguez, M. J., Huff, R. C., and Thompson, R. C. (1996) *J. Antibiot.* (*Tokyo*) **49,** 575–581
4. Nadkarni, S. R, Patel, M. V., Chatterjee, S., Vijayakumar, E. K. S., Desikan, K. R., Blumbach, J., and Ganguli, B. N. (1994) *J. Antibiot.* (*Tokyo*) **47,** 334–341
5. Pelzer, S., Süssmuth, R. D., Heckmann, D., Recktenwald, J., Huber, P., Jung, G., and Wohlleben, W. (1999) *Antimicrob. Agents Chemother.* **43,** 1565–1573
6. S. Pelzer, M. Huppert, W. Reichert, D. Heckmann und W. Wohlleben. (1997) *J. Biotechnol.* **56,** 115–128
7. Süssmuth, R. D., Pelzer, S., Nicholson, G. J., Walk, T., Wohlleben, W., and Jung, G. (1999) *Angew. Chem. Int. Ed. Engl.* **111,** 2096–2099
8. Bischoff, D., Pelzer, S., Höltzel, A., Nicholson, G. J., Stockert, S., Wohlleben, W., Jung, G., and Süssmuth, R. D. (2001) *Angew. Chem. Int. Ed. Engl.* **40,** 1693–1696
9. Pelzer, S., Huber, P., Süssmuth, R. D., Recktenwald, J., Heckmann, D., and Wohlleben, W. (December 14, 2000) International Patent WO0077182 A1
10. Hammond, S. J., Williamson, M. P., Williams, D. H., Boeck, L. D., and Marconi, G. G. (1982) *J. Chem. Soc. Chem. Commun.* 344–346
11. Choroba, O., Williams, D. H., and Spencer, J. B. (2000) *J. Am. Chem. Soc.* **122,** 5389–5390
12. Hubbard, B. K., Thomas, M. G., and Walsh, C. T. (2000) *Chem. Biol.* (*Lond.*) **42,** 1–12
13. Schröder, J. (1997) *Trends Plant Sci.* **2,** 373–378
14. Moore, B. S. and Hopke, J. N. (2001) *Chem. Biochem.* **2,** 35–38
15. Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y., and Horinouchi, S. (1999) *Nature* **400,** 897–899
16. Bullock, W. O., Fernandez, J. M., and Short, J. M. (1987) *Focus* (*Idaho*) **5,** 376–378

17. Olins, P. O., Devine, C. S., Rangwala, S. H., and Kavka, K. S. (1988) *Gene* **73,** 227–235
18. Rosenberg, A. H., and Studier, F. W. (1987) *Gene* **56,** 125–135
19. Muth, G., Wohlleben, W., and Pühler, A. (1988) *Mol. Gen. Genet.* **211,** 424–429
20. Muth, G., Nuβbaumer, B., Wohlleben, W., and Pühler, A. (1989) *Mol. Gen. Genet.* **219,** 341–348
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M., and Schrempf, H. (1985) *Genetic Manipulations of Streptomyces: A Laboratory Manual*, The John Innes Foundation, Norwich, UK
23. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
24. Laemmli, U. K. (1970) *Nature* **227,** 680–685
25. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci.* U. S. A. **74,** 5463–5467
26. Staden, R. (1996) *Mol. Biotechnol.* **5,** 233–241
27. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
28. Eckermann, S., Schröder, G., Schmidt, J., Strack, D., Edrada, R. A., Helariuatta, Y., Elomaa, P., Kotilainen, M., Kilpeläinen, I., Proksch, P., Teeri, T. H., and Schröder, J. (1998) *Nature* **396,** 387–390
29. Ferrer, J. L., Jez, J. M., Bowman, M. E., Dixon, R. A., and Noel, J. P. (1999) *Nat. Struct. Biol.* **6,** 775–784
30. Van Wageningen, A. M. A., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., Jones, M., Jones, S. J. M., and Solenberg, P. J. (1998) *Chem. Biol.* (*Lond.*) **5,** 155–162
31. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Sqares, R., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393,** 537–544
32. Redenbach, M., Kieser, H. M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H., and Hopwood, D. A. (1996) *Mol. Microbiol.* **21,** 77–96
33. Boynton, Z. L., Bennet, G. N., and Rudolph, F. B. (1996) *J. Bacteriol.* **178,** 3015–3024
34. Henne, A., Daniel, R., Schmitz, R. A., and Gottschalk, G. (1999) *Appl. Environ. Microbiol.* **65,** 3901–3907
35. Eichler, K., Bourgis, F., Buchet, A., Kleber, H. P., and Mandrand-Berthelot, M. A. (1994) *Mol. Microbiol.* **13,** 775–786
36. Chen, H., Thomas, M. G., Hubbard, B. K., Losey, H. C., Walsh, C. T., and Burkart, M. D. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 11942–11947
37. Hopwood, D. A., and Sherman, D. H. (1990) *Annu. Rev. Genet.* **24,** 37–66

A Polyketide Synthase in Glycopeptide Biosynthesis: THE BIOSYNTHESIS OF THE NON-PROTEINOGENIC AMINO ACID (S)-3,5-DIHYDROXYPHENYLGLYCINE

Volker Pfeifer, Graeme J. Nicholson, Johannes Ries, Jürgen Recktenwald, Alexandre B. Schefer, Riham M. Shawky, Joachim Schröder, Wolfgang Wohlleben and Stefan Pelzer

J. Biol. Chem. 2001, 276:38370-38377.
doi: 10.1074/jbc.M106580200 originally published online August 8, 2001
