Biosynthesis of L-*p*-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics*

Brian K Hubbard **, Michael G Thomas ** and Christopher T Walsh

Background: The non-proteinogenic amino acid *p*-hydroxyphenylglycine is a crucial component of certain peptidic natural products synthesized by a non-ribosomal peptide synthetase mechanism. In particular, for the vancomycin group of antibiotics *p*-hydroxyphenylglycine plays a structural role in formation of the rigid conformation of the central heptapeptide aglycone in addition to being the site of glycosylation. Initial labeling studies suggested tyrosine was a precursor of *p*-hydroxyphenylglycine but the specific steps in *p*-hydroxyphenylglycine biosynthesis remained unknown. Recently, the sequencing of the chloroeremomycin gene cluster from *Amycolatopsis orientalis* gave new insights into the biosynthetic pathway and allowed for the prediction of a four enzyme pathway leading to L-*p*-hydroxyphenylglycine from the common metabolite prephenate.

Results: We have characterized three of the four proposed enzymes of the L-*p*-hydroxyphenylglycine biosynthetic pathway. The three enzymes are encoded by open reading frames (ORFs) 21, 22 and 17 (ORF21: [PCZA361.1, O52791, CAA11761]; ORF22: [PCZA361.2, O52792, CAA11762]; ORF17: [PCZA361.25, O52815, CAA11790]), of the chloroeremomycin biosynthetic gene cluster and we show they have *p*-hydroxymandelate synthase, *p*-hydroxymandelate oxidase and L-*p*-hydroxyphenylglycine transaminase activities, respectively.

Conclusions: The L-*p*-hydroxyphenylglycine biosynthetic pathway shown here is proposed to be the paradigm for how this non-proteinogenic amino acid is synthesized by microorganisms incorporating it into peptidic natural products. This conclusion is supported by the finding of homologs for the four L-*p*-hydroxyphenylpyruvate biosynthetic enzymes in four organisms known to synthesize peptidic natural products that contain *p*-hydroxyphenylglycine. Three of the enzymes are proposed to function in a cyclic manner in vivo with L-tyrosine being both the amino donor for L-*p*-hydroxyphenylglycine and a source of *p*-hydroxyphenylpyruvate, an intermediate in the biosynthetic pathway.

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

*Supported in part by NIH grant GM 49338.
**These authors contributed equally to this work.

Correspondence: Christopher T Walsh
E-mail: walsh@walsh.med.harvard.edu

Keywords: Antibiotic; Hydroxyphenylglycine; Hydroxyphenylpyruvate; Tyrosine; Vancomycin

Received: **20 July 2000**
Revisions requested: **30 August 2000**
Revisions received: **18 September 2000**
Accepted: **21 September 2000**

Published: **17 October 2000**

Chemistry & Biology 2000, **7**:931–942

1074-5521/00/$ – see front matter

PII: S1074-5521(00)00043-0

Introduction

One of the hallmarks of peptidic natural products synthesized by non-ribosomal peptide synthetases (NRPSs) is the incorporation of non-proteinogenic amino acids into the antibiotic products. Selection, activation and incorporation of these unusual amino acids is a central mechanism by which NRPS synthesized peptides achieve the remarkable structural diversity that contributes to their wide assortment of biological activities [1–3]. These activities range from antimicrobial, antiviral and immunosuppressive activities to specialized functions as iron chelators and biosurfactants. The unusual, non-proteinogenic amino acid moieties may arise from the incorporation of a natural, proteinogenic amino acid into the growing peptidyl chain by the NRPS, followed by downstream modification of the residue by epimerization, *N*-methylation, cyclization or oxidation to name a few of the observed modifications [1–3]. Alternatively, the non-proteinogenic amino acid may be synthesized by a novel, dedicated biosynthetic pathway to provide the monomer to the particular NRPS assembly line. Once incorporated into the growing peptidyl chain, additional modifications may be catalyzed by the NRPS.

The non-proteinogenic amino acid *p*-hydroxyphenylglycine is found in several peptidic natural products including the vancomycin group of antibiotics (e.g. vancomycin [4], chloroeremomycin [5], A47934 [6], and complestatin [7,8]) as well as other antimicrobial compounds such as ramoplanin [9] and calcium-dependent antibiotic [10] (Figure 1). For members of the vancomycin group of antibiotics, *p*-hydroxyphenylglycine residues play crucial roles in the structure and function of the final glycopeptide antibiotic. In vancomycin and chloroeremomycin, the D-*p*-hydroxyphenylglycine at residue four of the heptapeptide is oxidatively cross-linked with the aryl rings of the β-hydroxytyrosines at positions two and six, creating part of the rigid

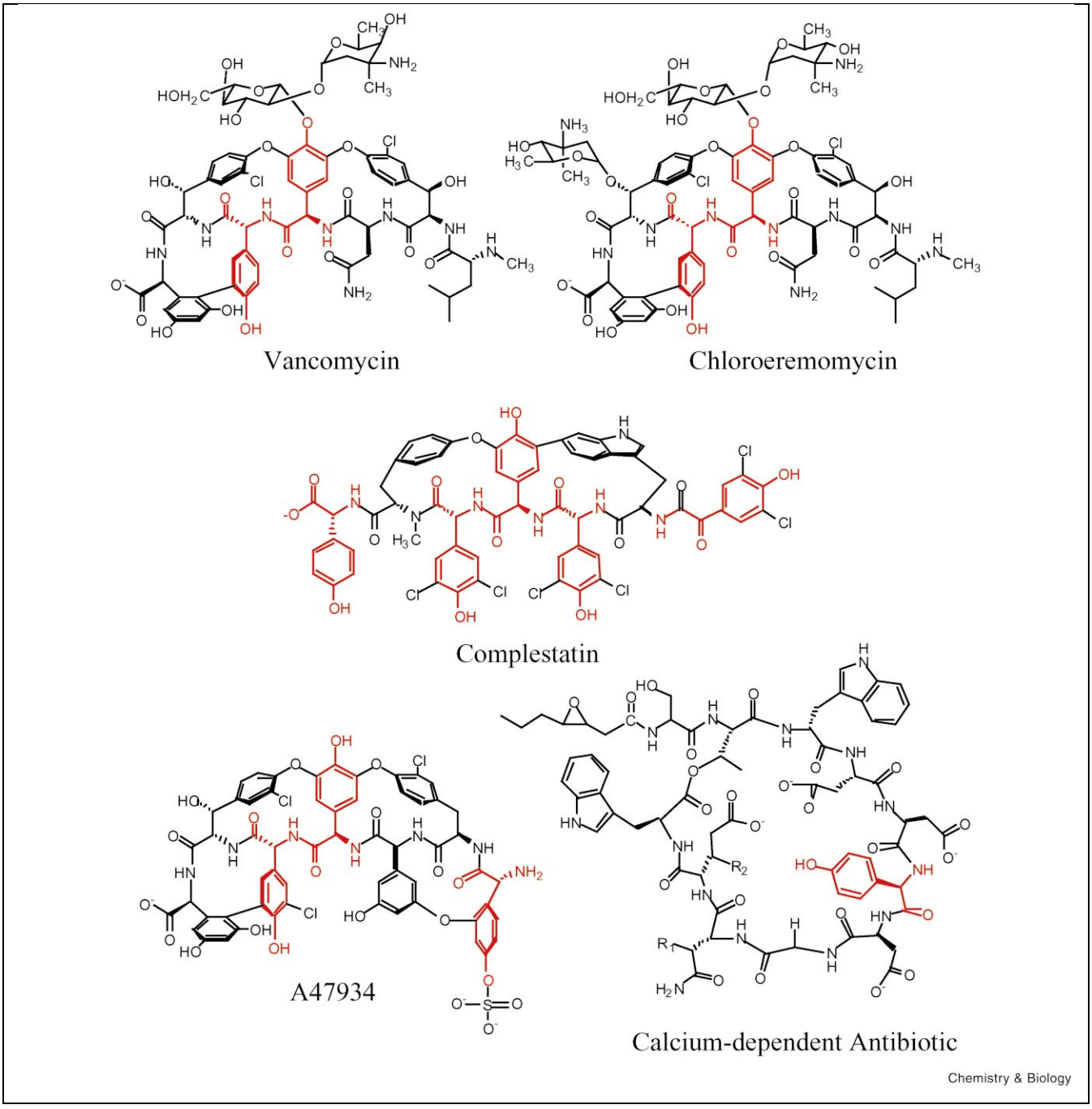

Figure 1. Chemical structures of representative *p*-hydroxyphenylglycine-, highlighted in red, containing peptidic natural products. For calcium-dependent antibiotic (CDA): CDA1: $R_1 = OPO_3H_2$, $R_2 = H$; CDA2: $R_1 = OPO_3H_2$, $R_2 = CH_3$; CDA3b: $R_1 = OH$, $R_2 = H$; CDA4b: $R_1 = OH$, $R_2 = CH_3$) [10].

scaffold characteristic of these antibiotics. Analogously, the D-*p*-hydroxyphenylglycine at residue five is oxidatively crosslinked to the 3,5-dihydroxy-L-phenylglycine at residue seven, completing the rigid scaffold (Figure 1). Finally, the D-*p*-hydroxyphenylglycine of residue four of the crosslinked aglycone is also the site of glycosylation as sugars are added to create the final antibiotics. The committed synthesis of *p*-hydroxyphenylglycine and clustering the genes encoding enzymes for its proposed biosynthetic pathway with other enzymes in vancomycin assembly suggests this residue provides reactivity and/or architecture not available in the proteinogenic homolog tyrosine. The

essential nature of this hydroxyaryl amino acid in the core structure of vancomycin group antibiotics suggests knowledge of the biosynthetic pathway is important for future efforts in combinatorial biosynthetic manipulation.

Proposals for *p*-hydroxyphenylglycine biosynthesis for vancomycin group antibiotics have been based on labeling studies where either [^{13}C]tyrosine or [^{2}H,^{13}C]tyrosine was fed to growing cultures and analysis of the final glycopeptide product showed incorporation of label into the D-*p*-hydroxyphenylglycine residues leading to the conclusion that *p*-hydroxyphenylglycine is derived from tyrosine [11–14]. Recently, new insights into *p*-hydroxyphenylglycine biosynthesis came from the sequencing of the chloroeremomycin biosynthetic gene cluster from *Amycolatopsis orientalis* [15] (Figure 2a). In this gene cluster, four open reading frames (ORFs) could be assigned to genes encoding enzymes with putative functions for *p*-hydroxyphenylglycine biosynthesis. These four ORFs (1, 17, 21, 22) (ORF1: [PCZA363.1, 052817, CAA11792.1]; ORF21: [PCZA361.1, O52791, CAA11761]; ORF22: [PCZA361.2, O52792, CAA11762]; ORF17: [PCZA361.25, O52815,

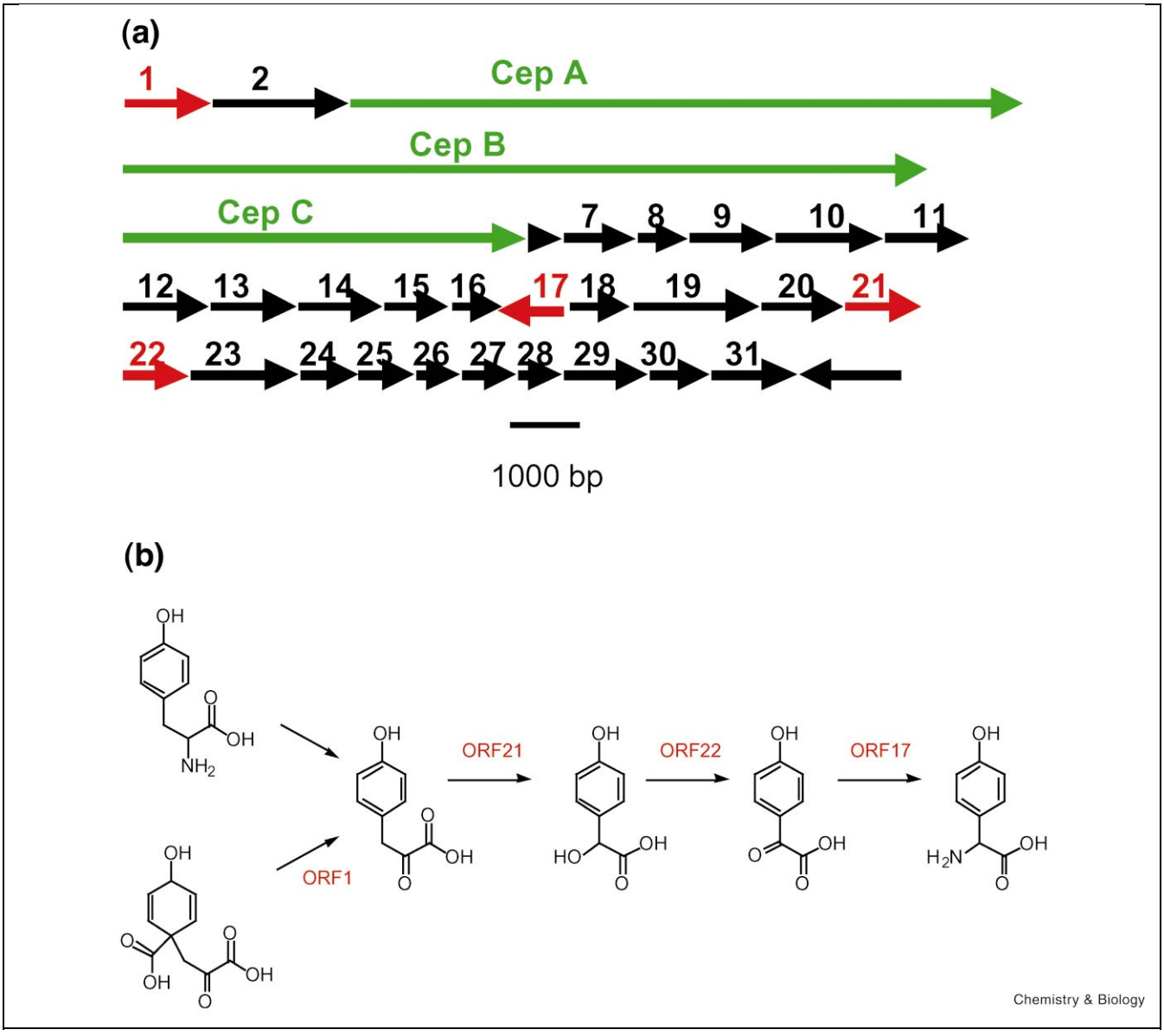

Figure 2. (a) Diagram of the genes encoded in the chloroeremomycin biosynthetic cluster. Numbers indicate the ORF starting with the first gene of the gene cluster. Cep A, B and C are the NRPS encoding genes [15]. **(b)** Outline of proposed biosynthetic pathway for the non-proteinogenic amino acid L-*p*-hydroxyphenylglycine.

CAA11790]) can be organized into a biosynthetic pathway, whereby *p*-hydroxyphenylglycine biosynthesis can begin not only with tyrosine as previously proposed, but also with prephenate, a common precursor for L-phenylalanine and L-tyrosine biosynthesis (Figure 2b).

Here, we present the purification and initial biochemical characterization of ORFs 17, 21 and 22 of the chloroeremomycin biosynthetic gene cluster and define their enzymatic roles in the biosynthesis of L-*p*-hydroxyphenylglycine.

Results

Expression and purification of the enzymes of the L-*p*-hydroxyphenylglycine biosynthetic pathway

Analysis of the chloroeremomycin biosynthetic cluster of *A. orientalis* identified four ORFs that could be organized into a putative L-*p*-hydroxyphenylglycine biosynthetic pathway based on amino acid sequence homologies. The candidates for this pathway were ORFS 1, 17, 21 and 22, which have the proposed functions of prephenate dehydrogenase, L-tyrosine:*p*-hydroxybenzoylformate transaminase, *p*-hydroxyphenylpyruvate dioxygenase and L-*p*-hydroxymandelate oxidase, respectively (Figure 2b). To assess the validity of this proposed pathway, each of the candidate ORFs was cloned and the encoded protein was overexpressed in *Escherichia coli* for initial biochemical characterization. ORF1 is homologous to prephenate dehydrogenases and is anticipated to generate *p*-hydroxyphenylpyruvate from prephenate. Attempts to obtain soluble protein for biochemical characterization have been unsuccessful to date. We therefore focused on the three remaining ORFs of the pathway and as shown in Figure 3, all three could be purified with affinity tags to near homogeneity based on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) analysis.

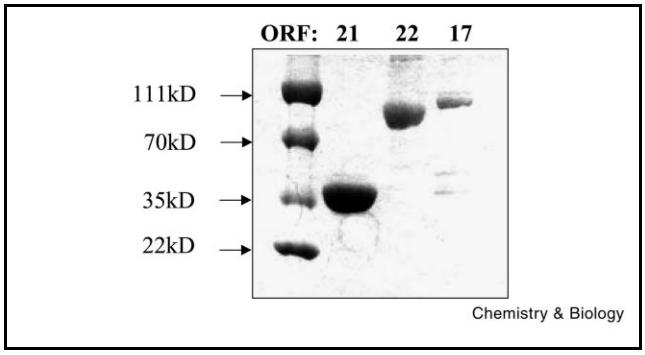

Figure 3. SDS–PAGE gel showing the purified protein preparations of ORF21/His, ORF22/MBP and ORF17/MBP. Lane 1: Molecular weight markers, lane 2: ORF21/HIS, lane 3: ORF22/MBP, lane 4: ORF17/MBP.

Characterization of ORF21 as an iron-dependent *p*-hydroxyphenylpyruvate decarboxylating hydroxylase that produces L-*p*-hydroxymandelate

ORF21, affinity-tagged with 10 histidine residues at its amino terminus, was purified in the apo form with a protein yield of 10 mg/l. The apo form of ORF21 was inactive, but holo-ORF21 could be reconstituted by preincubation of the enzyme with excess $FeCl_2$ immediately before use. This resulted in Fe^{3+} holo-enzyme that was inactive until the iron was reduced to Fe^{2+} by ascorbic acid that was present in the reaction mixture.

Based on amino acid sequence homology, ORF21 was anticipated to encode a *p*-hydroxyphenylpyruvate dioxygenase, a member of the α-keto acid-dependent non-heme iron-dependent dioxygenase family of enzymes [16]. ORF21 shows 28% identity and 44% similarity, at the amino acid level, with the *p*-hydroxyphenylpyruvate dioxygenase from *Pseudomonas fluorescens*, the structure of which has recently been reported [17]. Additionally, ORF21 retains 42 of the 54 amino acid residues conserved in all *p*-hydroxyphenylpyruvate dioxygenases from both eucaryotes and procaryotes [17]. *p*-Hydroxyphenylpyruvate dioxygenases are non-heme iron dioxygenases that catalyze the conversion of *p*-hydroxyphenylpyruvate to homogentisate during oxidative tyrosine degradation (reviewed in [16]).

It was our expectation that ORF21 would use *p*-hydroxyphenylpyruvate, putatively provided by the action of ORF1, and convert it not to homogentisate, but rather to *p*-hydroxymandelate (Figure 4a). To test this hypothesis of novel regiospecificity, holo-ORF21 was incubated with *p*-hydroxyphenylpyruvate and the product of the reaction was analyzed by high performance liquid chromatography (HPLC) (Figure 4b). This analysis revealed the complete conversion of the α-keto acid substrate to a product that was distinct from homogentisate. Furthermore, the product eluted from the HPLC with a retention time equal to that seen for authentic *p*-hydroxymandelate (Figure 4b). Coinjection of the reaction product with commercially available *p*-hydroxymandelate resulted in the elution of a single peak from the HPLC (data not shown), providing further support for the conclusion that ORF21 catalyzes the conversion of *p*-hydroxyphenylpyruvate to *p*-hydroxymandelate. The hydroxylation is likely to be stereospecific based on the specificity of ORF22 for L-*p*-hydroxymandelate (see below) and validation of this proposal was provided by in situ coupling of ORF21 and ORF22 (discussed below). While this manuscript was being submitted, Choroba et al. published a communication on ORF21 [18], independently establishing some of the results discussed here. To stay consistent with their recommended nomenclature, ORF21 is referred to as 4-hydroxymandelate synthase (HmaS).

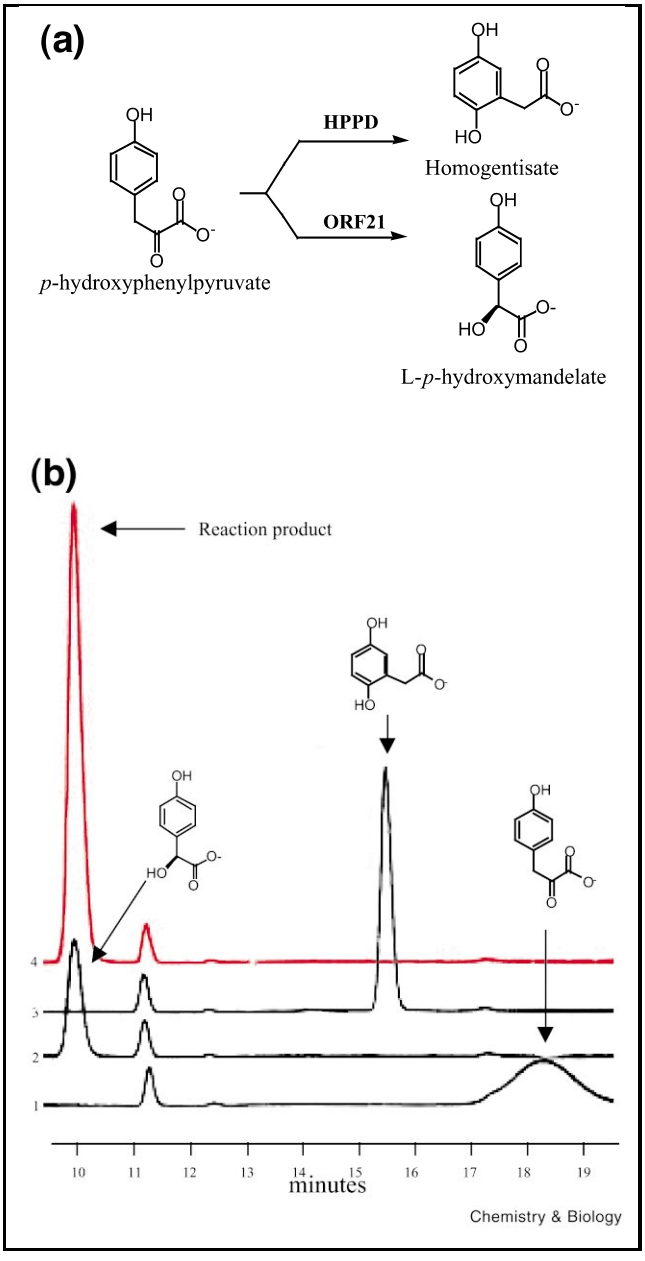

Figure 4. (a) Comparison of the reaction catalyzed by *p*-hydroxyphenylpyruvate dioxygenase versus the reaction proposed to be catalyzed by ORF21. **(b)** HPLC traces of products from the following reactions: 1, *p*-hydroxyphenylpyruvate standard; 2, *p*-hydroxymandelate standard; 3, homogentisate standard and 4, complete reaction mixture (ORF21 plus *p*-hydroxyphenylpyruvate), highlighted in red. Each peak is identified by the chemical structure of the corresponding compound. For each of the standards, ORF21 was inactivated by TFA addition prior to addition of standard compound. The peak at 11.3 min is due to ascorbic acid, the reductant used in all ORF21 assays (data not shown).

Characterization of ORF22 as a hydroxymandelate oxidase (Hmo), an FMN containing enzyme

Expression of ORF22

The predicted reaction for Hmo is the oxidation of *p*-hydroxymandelate to give *p*-hydroxybenzoylformate. Hmo shares amino acid sequence homology to the FMN containing glycolate oxidase (GO) and mandelate dehydrogenase (MDH) (50 and 52% amino acid sequence identity, respectively). MDH catalyzes the oxidation of L-mandelate to yield benzoylformate using FMN as a cofactor. Both GO and MDH contain FMN when purified from overexpression in *E. coli* [19].

The gene encoding Hmo was initially cloned into the protein expression plasmid pET-16b, resulting in an amino-terminal decahistidyl-tagged Hmo. Expression of Hmo in pET-16b yielded large quantities of soluble protein. However, after purification of the soluble, affinity-tagged Hmo by nickel chelate chromatography using standard conditions, the purified enzyme did not show UV/Vis absorbance at 375 or 480 nm (the usual absorbance of FMN; data not shown), suggesting the FMN cofactor was removed during purification. Consistent with that conclusion, no activity could be detected for the oxidation of D,L-*p*-hydroxymandelate by this purified enzyme. Furthermore, the addition of exogenous FMN to the enzyme did not result in detectable oxidation of *p*-hydroxymandelate.

It was assumed that purification of this enzyme under the conventional high salt conditions for poly-his-tagged proteins resulted in the irreversible release of the FMN from the enzyme. As an alternative purification procedure, Hmo was cloned into a vector that introduced an amino-terminal maltose binding protein (MBP) affinity tag that allowed purification of the MBP–Hmo fusion protein under low salt conditions (see Materials and Methods). The resulting protein was yellow and showed a UV/Vis absorbance spectrum typical of FMN containing enzymes (data not shown). The MBP–Hmo protein was used for all assays listed below.

Activity of ORF22

The Hmo catalyzed oxidation of *p*-hydroxymandelate was detected by UV/Vis assays using potassium ferricyanide as electron acceptor [19]. HPLC assays of Hmo activity, after exhaustive incubation with substrate, showed that half of the D,L-*p*-hydroxymandelate was converted to a new product (Figure 5) that decomposed with time. Mass spectral analysis of this product showed the compound to have a mass of 166 (data not shown), consistent with the product being *p*-hydroxybenzoylformate, as predicted (Figure 5). The secondary products of the reaction were not analyzed but are presumed to result from decarboxylation of *p*-hydroxybenzoylformate.

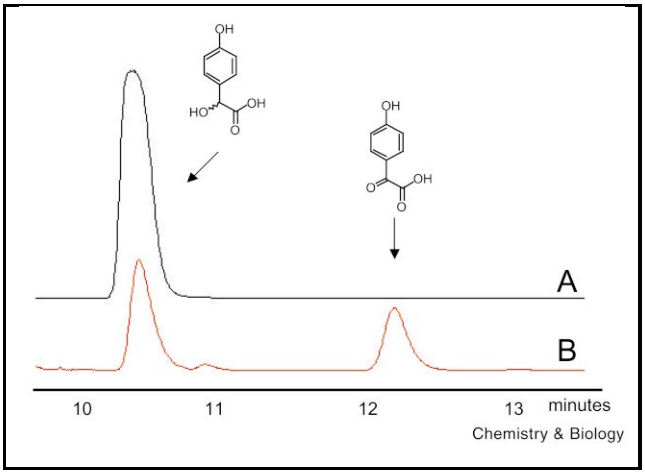

Figure 5. HPLC trace of Hmo catalyzed reaction. A: Authentic D,L-*p*-hydroxymandelate. B: Hmo catalyzed reaction with D,L-*p*-hydroxymandelate showing the production of *p*-hydroxybenzoylformate.

Since half of the D,L-*p*-hydroxymandelate was converted to product in the Hmo catalyzed reaction, we sought to determine the enzyme specificity for the D- or L-isomer (the presumed product of the HmaS reaction). Both authentic D- and L-mandelate were tested as substrates and HPLC traces showed that only L-mandelate was converted to a product that eluted from the HPLC with a retention time consistent with benzoylformate (Figure 6). The product of the Hmo/L-mandelate reaction coeluted from the HPLC with authentic benzoylformate. By extension we assigned L-*p*-hydroxymandelate as the substrate for Hmo.

When Hmo was assayed in conjunction with HmaS, the product of the HmaS reaction (*p*-hydroxymandelate) was converted to *p*-hydroxybenzoylformate by Hmo (Figure 7). Since L-*p*-hydroxymandelate was the substrate for Hmo and nearly all the HmaS generated product was converted by Hmo, the product of HmaS is L-*p*-hydroxymandelate as seen in Figure 7a.

Characterization of ORF17 as a *p*-hydroxyphenylglycine transaminase (HpgT)

Expression of ORF17

ORF17 shows protein sequence homology to a number of amino acid transaminases and was hypothesized to be a HpgT catalyzing the conversion of *p*-hydroxybenzoylformate to L-*p*-hydroxyphenylglycine. The gene was cloned and the enzyme was expressed as an amino-terminal MBP fusion protein, similar to the procedure used for Hmo. Following purification, the protein did not contain the proposed coenzyme, pyridoxal phosphate (PLP). Exogenous PLP was required for HpgT catalytic activity and was added to the reaction mixtures for the assays described below.

Characterization of the HpgT reaction

The conversion of L-*p*-hydroxyphenylglycine (the presumed product of the enzyme) to *p*-hydroxybenzoylformate in the presence of various common α-keto acids as amino acceptors was used as the initial characterization of HpgT. In this 'reverse direction' assay L-*p*-hydroxyphenylglycine was used as the amino donor in conjunction with the common α-keto acids, glutamate, pyruvate and oxaloacetate, as well as *p*-hydroxyphenylpyruvate, an intermediate from the *p*-hydroxyphenylglycine biosynthetic route, as amino acceptors. The reactions demonstrated that only *p*-hydroxyphenylpyruvate was able to function as a cosubstrate for HpgT catalyzed conversion of L-*p*-hydroxyphenylglycine to *p*-hydroxybenzoylformate. Previous feeding experiments demonstrated that radiolabeled L-*p*-hydroxyphenylglycine and D-*p*-hydroxyphenylglycine were incorporated into mature vancomycin [11,12]. To determine if L-*p*-hydroxyphenylglycine and/or D-*p*-hydroxyphenylglycine were the product of the amino transferase reaction, L- and D-*p*-hydroxyphenylglycine were used in an assay with glutamate, pyruvate, oxaloacetate or *p*-hydroxyphenylpyruvate as amino acceptors. Only the combination of L-*p*-hydroxyphenylglycine and *p*-hydroxyphenylpyruvate gave two new products, *p*-hydroxybenzoylformate and tyrosine. This demonstrated that *p*-hydroxybenzoylformate and tyrosine were interconverted with L-*p*-hydroxyphenylglycine and *p*-hydroxyphenylpyruvate by HpgT (Figure 8a,b).

To demonstrate that the product of the Hmo reaction was converted to L-hydroxyphenylglycine, the 'forward direction' assay, a reaction with *p*-hydroxymandelate, L-tyrosine,

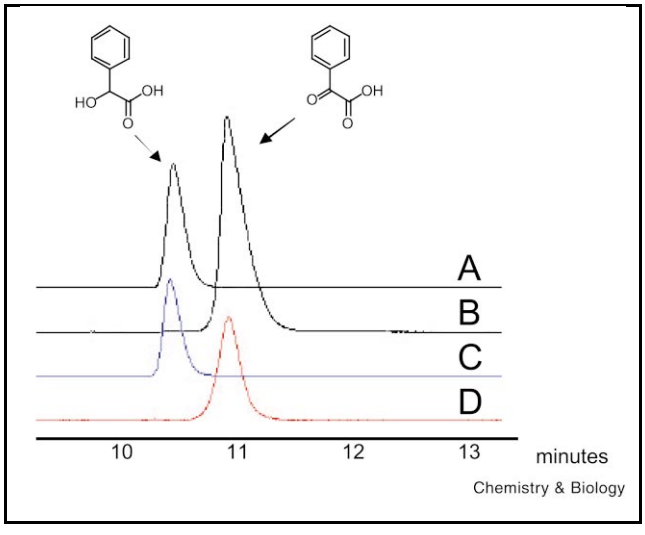

Figure 6. HPLC trace of the Hmo catalyzed reaction with D- and L-mandelate. A: Authentic mandelate. B: Authentic benzoylformate. C: Reaction of Hmo with D-mandelate. D: Reaction of Hmo with L-mandelate showing the formation of benzoylformate.

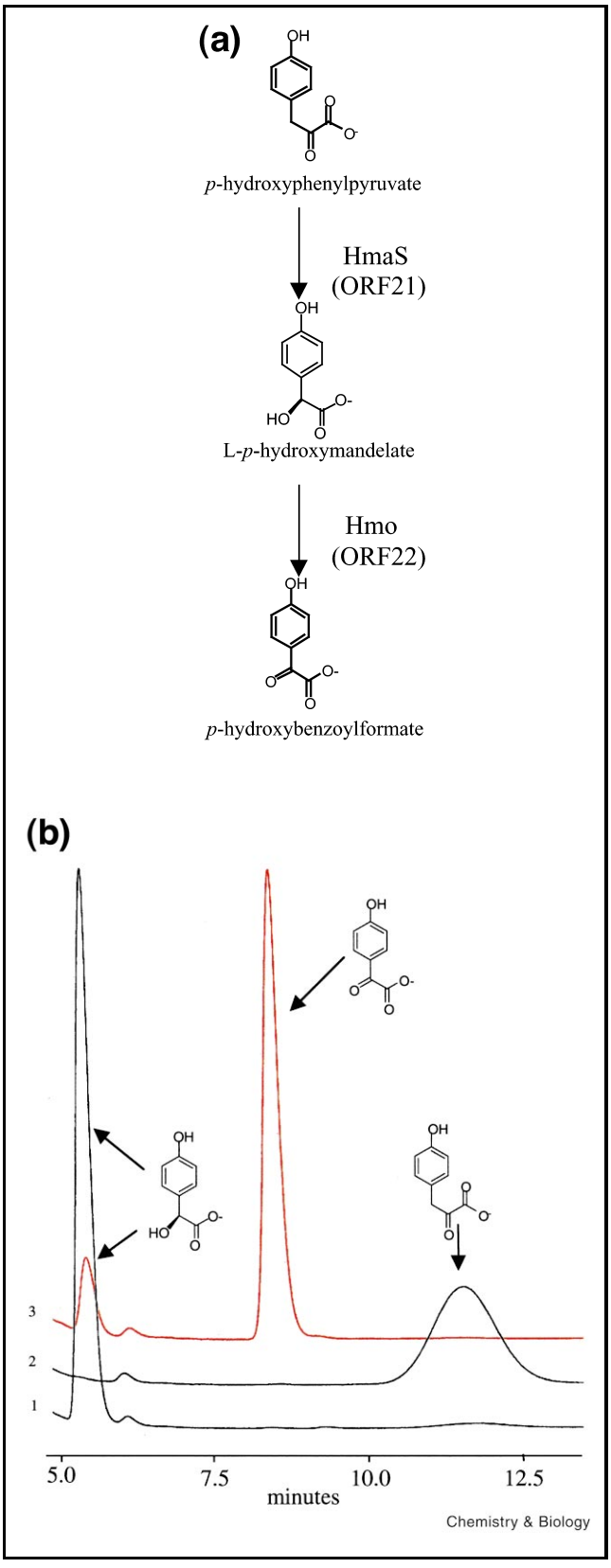

Figure 7. (a) Schematic representation of the first two steps of the L-*p*-hydroxyphenylglycine biosynthetic pathway. **(b)** HPLC traces of the following reactions: 1, *p*-hydroxyphenylpyruvate plus HmaS (ORF21); 2, *p*-hydroxyphenylpyruvate plus Hmo (ORF22) and 3, *p*-hydroxyphenylpyruvate plus HmaS and Hmo (red). Each peak is identified by the chemical structure of the corresponding compound.
←

Hmo and HpgT, was carried out and monitored by HPLC. Half of the *p*-hydroxymandelate was converted to *p*-hydroxyphenylglycine, as determined by co-migration with authentic L-*p*-hydroxyphenylglycine and *p*-hydroxyphenylpyruvate (Figure 9). The conversion of the Hmo product to the substrate for HmaS establishes a three enzyme cycle, catalytic in *p*-hydroxyphenylpyruvate (Figure 10).

Substrate specificity for HpgT

To determine substrate specificity of HpgT, the reactions listed in Table 1 were performed. These results imply the reactivities shown in Figure 9, where the *p*-hydroxyl group on the amino donor or amino acceptor is not a stringent requirement, and 3-Cl-L-tyrosine is also accepted as an amino donor. This implies a promiscuity by the HpgT. Since many of the glycopeptide antibiotics contain mono- and dichloro-hydroxyphenylglycine, it is possible that this HpgT is involved in the biosynthesis of the chlorinated analogues.

Discussion

The non-proteinogenic amino acid *p*-hydroxyphenylglycine is incorporated into various peptidic natural products such as the vancomycin group of antibiotics, ramoplanin and calcium-dependent antibiotic. The presence of a non-proteinogenic amino acid, such as *p*-hydroxyphenylglycine, in a peptidic natural product strongly suggests that the molecule has an NRPS component to its biosynthetic pathway. This is clearly the case for chloroeremomycin from *A. orientalis* [15] and calcium-dependent antibiotic from *S. coelicolor* A3(2) (ftp://ftp.sanger.ac.uk/pub/S_coelicolor/). While both the L- and D-isomers of *p*-hydroxyphenylglycine are found in natural products, it is anticipated that the L-isomer is the natural substrate of the NRPS. Epimerization of L-*p*-hydroxyphenylglycine to D-*p*-hydroxyphenylglycine would occur cosynthetically as the growing peptide translocates down the NRPS assembly line [15]. Thus, the *p*-hydroxyphenylglycine biosynthetic pathway should and does generate L-*p*-hydroxyphenylglycine, not the D-isomer. An outline of how L-*p*-hydroxyphenylglycine biosynthesis is likely to occur was recently suggested by the sequencing of the biosynthetic cluster for a vancomycin group antibiotic, chloroeremomycin [15]. Encoded within this biosynthetic cluster were four ORFs (1, 17, 21, 22) that predict a four enzyme biosynthetic pathway beginning with the aromatic metabolite prephenate and proceeding to L-*p*-hydroxyphenylglycine (Figure 2). A portion of this prediction has been validated here and inde-

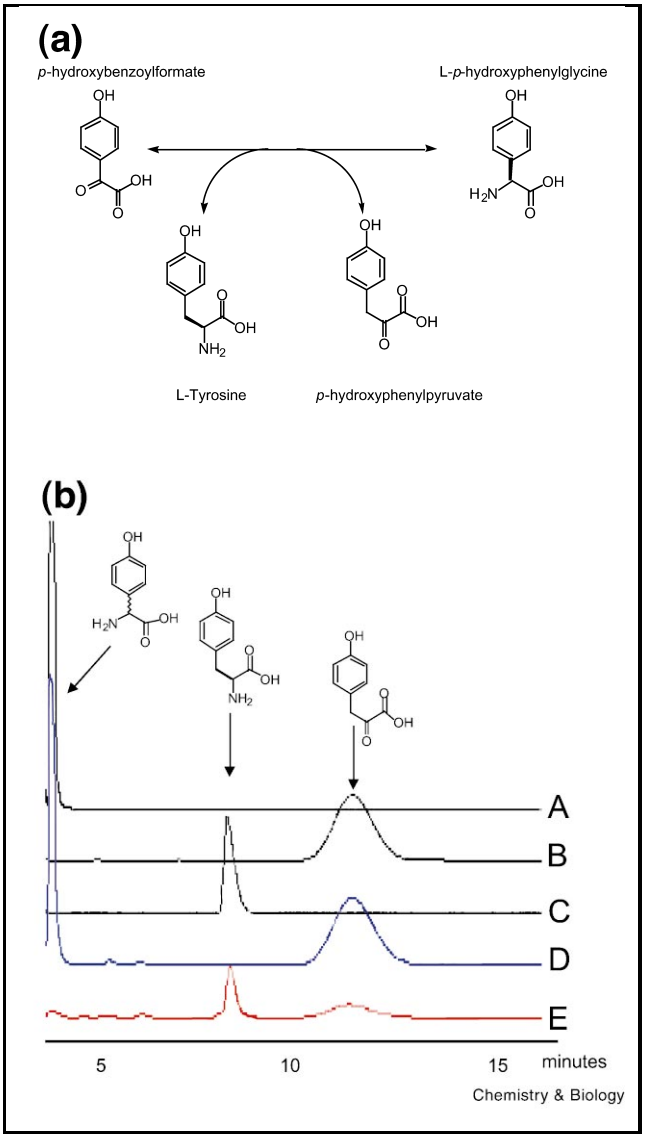

Figure 8. (a) HpgT catalyzed interconversion of *p*-hydroxybenzoylformate and L-tyrosine with *p*-hydroxyphenylpyruvate and L-*p*-hydroxyphenylglycine. **(b)** HPLC of the HpgT catalyzed reaction. A: Authentic *p*-hydroxyphenylglycine. B: Authentic *p*-hydroxyphenylpyruvate. C: Authentic L-tyrosine. D: Reaction of HpgT with D-*p*-hydroxyphenylglycine and *p*-hydroxyphenylpyruvate showing no reaction. E: Reaction of HpgT with L-*p*-hydroxyphenylglycine and *p*-hydroxyphenylpyruvate showing the formation of L-tyrosine.

pendently by Choroba et al. in a communication which appeared as this manuscript was being submitted [18], demonstrating the reaction of HmaS.

HmaS

The generation of L-*p*-hydroxymandelate from *p*-hydroxyphenylpyruvate is the key enzymatic transformation of the L-*p*-hydroxyphenylglycine biosynthetic pathway. The decarboxylation and hydroxylation activity of HmaS shows novel and distinct regioselectivity, compared to all other known *p*-hydroxyphenylpyruvate dioxygenases, by hydroxylating the benzylic position of the substrate instead of the phenyl ring, attendant with acetate side chain migration. The independent results of Choroba et al. have confirmed ORF21 is in fact a dioxygenase by $^{18}O_2$ incorporation studies [18]. It is likely that the *p*-hydroxyphenylpyruvate substrate is oriented subtly different towards the Fe^{2+}-O_2 in the active site. Consistent with this proposal, alignment of HmaS with the *P. fluorescens* *p*-hydroxyphenylpyruvate dioxygenase, which has been crystallized and the structure solved [17], suggests that residues surrounding the opposing face of the 2-his-1-carboxylate iron-coordinating residues are not conserved between the two enzymes (data not shown). We have also established that the benzylic hydroxylation is stereospecific and yields the L-hydroxyisomer of *p*-hydroxymandelate by coupling the HmaS reaction with the next enzyme of the biosynthetic pathway, Hmo, to generate the biosynthetic intermediate *p*-hydroxybenzoylformate.

Hmo

We have established that Hmo (previously ORF22) is the L-*p*-hydroxymandelate oxidase, catalyzing the formation of *p*-hydroxybenzoylformate. Hmo shares amino acid sequence homology with GOs and MDHs. MDH is a membrane bound enzyme while GO is found in the cytosol of the cell. Sequence comparison of Hmo with GO and MDH show that Hmo is missing the membrane anchoring sequence of MDH and should therefore be present in the cytosol.

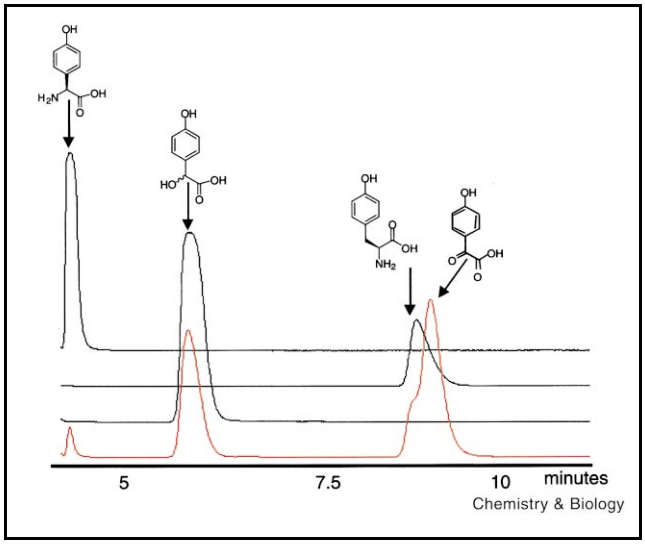

Figure 9. HPLC of the HpgT catalyzed reaction in the biosynthetic direction. A: Authentic *p*-hydroxyphenylglycine. B: Authentic L-tyrosine. C: Authentic *p*-hydroxymandelate. D: Reaction consisting of D,L-*p*-hydroxymandelate, Hmo, L-tyrosine and HpgT showing the formation of *p*-hydroxybenzoylformate and L-*p*-hydroxyphenylglycine.

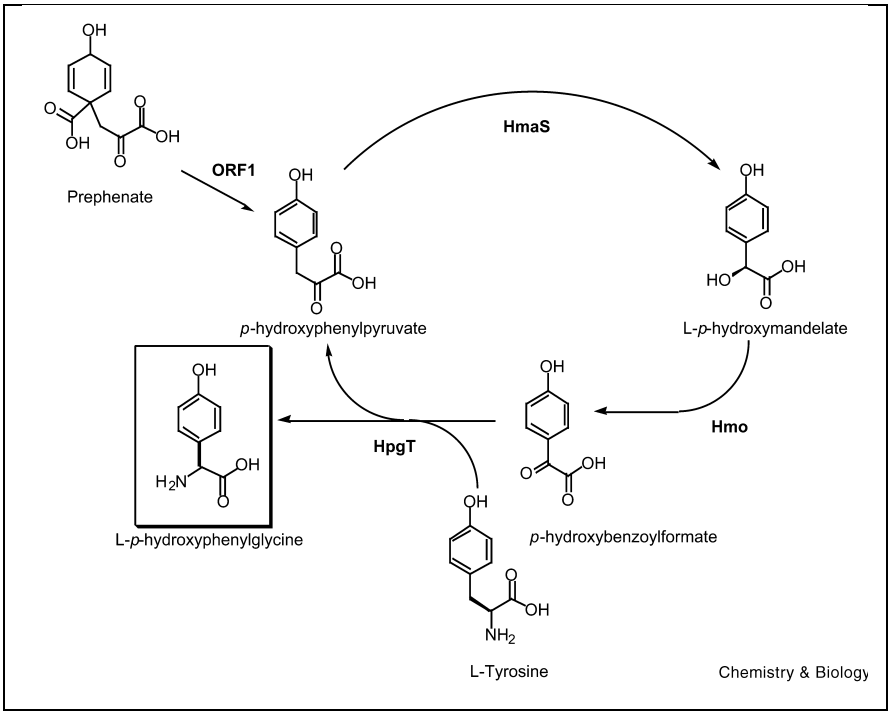

Figure 10. Catalytic cycle for the biosynthesis of L-*p*-hydroxyphenylglycine.

The FMN present in Hmo was susceptible to displacement from the enzyme in high salt conditions. Either purification of Hmo in high salt or incubation of the enzyme in high salt buffers gives irreversible loss of FMN and concomitant loss of catalytic activity. We were able to overcome this problem using a single step, amylose resin, purification of the MBP fused protein. The protein was frozen immediately and retained its activity for at least 3 months.

HpgT

We have demonstrated that *p*-hydroxybenzoylformate (the product of the Hmo catalyzed reaction) is converted to L-*p*-hydroxyphenylglycine by HpgT. HpgT shows amino acid sequence homology to other amino acid transaminases. Amino transferases usually use glutamate, aspartate or alanine as amino donors. In a sequence comparison of known amino transferases, the proteins cluster into three groups depending on the amino donor they utilize (glutamate, aspartate or alanine), however, HpgT appears to be distinct from these enzymes (data not shown). This finding gives insight into the possibility of an unusual amino donor for the conversion of *p*-hydroxybenzoylformate to L-*p*-hydroxyphenylglycine. L-Tyrosine was assayed and found to be the amino donor for the HpgT catalyzed reaction.

The use of L-tyrosine as a cosubstrate in the reductive amination of *p*-hydroxybenzoylformate to L-*p*-hydroxyphenylglycine makes an economical pathway: as *p*-hydroxybenzoylformate is transaminated to the desired non-proteinogenic amino acid L-*p*-hydroxyphenylglycine, the L-tyrosine yields one equivalent of *p*-hydroxyphenylpyruvate to serve as a substrate for HmaS and primes the three enzyme cycle of HmaS, Hmo and HpgT for another

Table 1
Assay for the substrate specificity elicited by HpgT.

Amino acceptor	Amino Donor	Keto acid product	Amino product
p-Hydroxybenzoylformate	L-Tyr	*p*-hydroxyphenylpyruvate	Hpg
p-Hydroxyphenylpyruvate	L-Hpg	*p*-hydroxybenzoylformate	Tyr
p-Hydroxyphenylpyruvate	D-Hpg	no reaction	no reaction
Benzoylformate	L-Tyr	*p*-hydroxyphenylpyruvate	Phe
p-Hydroxyphenylpyruvate	L-Phg	benzoylformate	Tyr
p-Hydroxybenzoylformate	3-Cl-L-Tyr	3-Cl-HPP	Hpg
p-Hydroxybenzoylformate	L-Phe	phenylpyruvate	Hpg

HPP = hydroxyphenylpyruvate; Hpg = *p*-hydroxyphenylglycine.

turn. This proposal is in agreement with the prior tyrosine labeling studies that suggested tyrosine is a *p*-hydroxyphenylglycine precursor for vancomycin group antibiotics [11–14]. Additionally, it is presumed that prephenate dehydrogenase also serves to prime the biosynthetic pathway by generating *p*-hydroxyphenylpyruvate from the aromatic amino acid precursor prephenate.

The strongest in vivo support for this proposed *p*-hydroxyphenylglycine biosynthetic pathways comes from [^{14}C]tyrosine incorporation into the antibiotic nocardicin A [20]. In this β-lactam antibiotic, two molecules of *p*-hydroxyphenylglycine are incorporated into the final product. Through the use of [^{14}C]tyrosine feeding experiments, Hosoda et al. isolated [^{14}C]-labeled *p*-hydroxymandelate, *p*-hydroxybenzoylformate and *p*-hydroxyphenylglycine and proposed a *p*-hydroxyphenylglycine biosynthetic pathway that is nearly identical to the pathway proposed here [20]. Therefore, it is anticipated homologs for the enzymes discussed here will also be found in the biosynthetic clusters of nocardicin antibiotics.

Significance

We have presented the initial biochemical analysis of three ORFs from the chloroeremomycin biosynthetic cluster of *A. orientalis*. Analysis of ORF21 (HmaS) shows the enzyme is a novel non-heme iron dioxygenase that catalyzes the synthesis of L-*p*-hydroxymandelate from *p*-hydroxyphenylpyruvate. The L-*p*-hydroxymandelate generated by HmaS is subsequently converted to L-*p*-hydroxyphenylglycine via ORF22 (Hmo) catalyzed conversion of L-*p*-hydroxymandelate to L-*p*-hydroxybenzoylformate followed by ORF17 (HpgT) catalyzed transamination of *p*-hydroxybenzoylformate to *p*-hydroxyphenylglycine. The biosynthesis of L-*p*-hydroxyphenylglycine is proposed to be a catalytic cycle initiated by the production of *p*-hydroxyphenylpyruvate. ORF1, a homolog to prephenate dehydrogenase, present in the gene cluster could prime the first turn of the catalytic cycle (Figure 8). The conversion of prephenate to *p*-hydroxyphenylpyruvate initiates the cycle and starts the net flux of L-tyrosine to L-*p*-hydroxyphenylglycine. It is also noted that homologs of the ORFs 1, 17, 21 and 22 are also found in the biosynthetic clusters of complestatin from *S. lavendulae* (Chiu and Khosla, personal communication), A47934 from *S. toyocaensis* (Wright, personal communication), and calcium-dependent antibiotic from *S. coelicolor* A3(2) (ftp://ftp.sanger.ac.uk/pub/S_coelicolor/), all synthesizing natural products that contain *p*-hydroxyphenylglycine. This finding suggests this four gene cluster is an L-*p*-hydroxyphenylglycine biosynthetic 'cassette' that will be found in all microorganisms incorporating *p*-hydroxyphenylglycine into an NRPS synthesized natural product.

Materials and methods

Cloning of ORF1 and ORF21

The proposed coding regions for ORF1 and ORF21 of the chloroeremomycin cluster were amplified by polymerase chain reaction (PCR) from *A. orientalis* genomic DNA with primers ORF1Nterm*Nde*I (5′ GGGAATTC**CATATG**GAGAAGGTGCTCGTTGTCGGC 3′ and ORF1Cterm*Xho*I (5′ GGGCCG**CTCGAG**TCAGCGGGGACAACCCTGGGACCG 3′) for ORF1 and ORF21Nterm*Nde*I (5′ GGGAATTC**CATATG**CAGAATTTCGAGATCGACTAC 3′) and ORF21Cterm*Xho*I (5′ GGGCCG**CTCGAG**TCATCGCCGAGCGGCGCCGAACTC 3′) for ORF21. The introduced *Nde*I and *Xho*I restriction sites of each primer are identified by boldface type. The amplified products were gel purified, digested with *Nde*I and *Xho*I and ligated, independently, into the *Nde*I/*Xho*I cloning sites of pET16b. The resulting plasmids, pH10-ORF1 and pH10-ORF21, introduced an in-frame His 10 affinity tag onto the amino terminus of each ORF. For overexpression of ORF1 and ORF21, the affinity-tagged plasmids were transformed into *E. coli* strain BL21(DE3). Affinity-tagged ORF1 overexpressed at 30, 25 and 15°C but was not soluble under these expression conditions.

Overexpression and purification of ORF21

The affinity-tagged ORF21 was overexpressed by inoculating 1 l of Luria–Bertani (LB) broth containing 100 μg/ml ampicillin (in a 3 l Belco baffled flask) with 10 ml of an overnight culture of BL21(DE3)/pH10-ORF21 and growing the culture for 16 h at 30°C. Cells were harvested by centrifugation (10 min at 5000 × *g*) and resuspended in 40 ml of 1 × Bind Buffer (20 mM Tris–HCl pH 7.9; 0.5 M NaCl; 5 mM imidazole). Resuspended cells were broken by sonication (Fisher 550 Sonic Dismembrator, power = 5, 10 min sonication with 2 s on, 2 s off) and cell debris was removed by centrifugation (30 min at 95 800 × *g*). ORF21 was purified by nickel chelate chromatography using His·Bind® Resin (Novagen). After 2 h incubation with cell-free extract, resin was packed into a column and washed with 10 column volumes of 1 × Bind Buffer followed by 10 column volumes of 1 × Wash Buffer (20 mM Tris–HCl pH 7.9; 0.5 M NaCl; 60 mM imidazole). Both column washes and ORF21 elution were performed at a flow rate of 1 ml/min. ORF21 was eluted with a linear gradient (12 column volumes) from 1 × Wash Buffer to 1 × Elution Buffer (20 mM Tris–HCl pH 7.9; 0.5 M NaCl; 0.5 M imidazole). The elution of ORF21 was followed via SDS–PAGE and fractions containing ORF21 were pooled and dialyzed against Buffer A (50 mM Tris–HCl, pH 7.9, at 4°C; 0.1 M NaCl; 5% (v/v) glycerol; 1 mM EDTA) followed by dialysis against Buffer B (50 mM Tris–HCl, pH 7.9, at 4°C; 0.1 M NaCl; 5% (v/v) glycerol; 1 mM DTT). Dialyzed protein was concentrated to 2.5 mg/ml using a Centriprep 10 (Amicon), flash-frozen in liquid nitrogen and stored at −80°C. The protein concentration was determined spectrophotometrically at 280 nm using the calculated molar extinction coefficient of 18 610 M^{-1} cm^{-1}. The yield of ORF21, after purification, was 10 mg/l.

HPLC assay of ORF21

Reconstitution of holo-ORF21 involved diluting apo-ORF21 into Buffer B containing 1 mM $FeCl_2$. Standard reconstitution involved a 20-fold molar excess of $FeCl_2$ compared to enzyme. Reaction conditions were a minor modification of those described for HPPD analysis [21]. The reaction mixture (150 μl), at a pH of 7.5, contained 0.8 M potassium phosphate, 0.2 M Tris–HCl, 25 mM ascorbic acid, 0.1 mg/ml bovine liver catalase, 2 μM ORF21 and additions as indicated in Figure 4. Reactions were incubated at 25°C for 1 h and stopped by the addition of 15 μl of trifluoroacetic acid (TFA). Samples were then centrifuged at 13 000 rpm for 10 min to remove denatured protein.

Products and reactants of the reaction mixtures were separated via HPLC (Beckman System Gold) using a Vydac™ Protein and Peptide C18 column at a flow rate of 1 ml/min. The solvents used were as follows: solvent A (ddH_20, 0.1% TFA) and solvent B (acetonitrile, 0.1% TFA). 50 μl of each reaction mixture was injected for each run. The profile for separation was 1 min isocratic development at 100% A/0% B; 30 min linear gradient from 100% A/0% B to 80% A/20% B. The elution of reaction components was monitored at 230 nm.

Cloning, expression and purification of ORF22/pET-16b

The proposed coding regions for ORF22 of the chloroeremomycin cluster was amplified by PCR from *A. orientalis* genomic DNA with primers ORF22Nterm*Nde*I (5′ GGGAATTC**CATATG**ACGTACGTTTCCCTGGCCGAC 3′) and ORF22Cterm*Xho*I (5′ GGGCCG**CTCGAG**TCAAACAACCCCCAGTTTCGTGTT 3′). The introduced *Nde*I and *Xho*I restriction sites of each primer are identified by bold-face type. The amplified products were gel purified, digested with *Nde*I and *Xho*I, and ligated, into the *Nde*I/*Xho*I cloning sites of pET16b, which encodes for an N-terminal deca-histidyl tag. Growth of ORF22/pET-16b involved inoculation of 1.5 l of LB with 50 mg/ml ampicillin (in a 3 l Belco baffled flask) with 5 ml of an overnight culture. The culture was grown at 37°C for 18 h. The cells were harvested by centrifugation (20 min at 5000 × g). Cell pellets were resuspended in 20 ml of Binding Buffer (20 mM Tris–HCl pH 8.0, 500 mM NaCl, 10 mM imidazole) and sonicated on a 550 Sonic Dismembrator (Fisher Scientific) (power = 5, 15 min sonication, 2 s on, 2 s off).

The cell lysates were centrifuged (45 min at 10 000 × *g*) and loaded on a nickel chelating Sepharose fast flow (Pharmacia) column preequilibrated with Charge Buffer (20 mM Tris–HCl pH 8.0, 100 mM NaCl, 50 mM nickel sulfate) (150 ml) and Binding Buffer (20 mM Tris–HCl pH 8.0, 500 mM NaCl, 10 mM imidazole) (150 ml). Cell lysates were loaded onto the column at 1 ml/min. The column was then washed at 5 ml/min with 200 ml of Binding Buffer, 200 ml of Wash Buffer (20 mM Tris–HCl pH 8.0, 500 mM NaCl, 80 mM imidazole) and 150 ml of Binding Buffer. The protein was eluted at 5 ml/min with 300 ml of Strip Buffer (40 mM Tris–HCl pH 8.0, 500 mM NaCl, 100 mM EDTA). The protein was dialyzed in 10 ml Tris–HCl pH 7.5, 10 mM NaCl and assayed for activity.

Cloning, expression and purification of ORF22/MBP

ORF22 was cloned via PCR into the *Nde*I and *Xho*I sites of pIADL-16 (Walsh lab vector encoding an N-terminal MBP fusion under the control of the T7 promoter; a derivative of pET-28b). Cloning was performed analogous to that for ORF22/pET-16b. Growth of ORF22/MBP involved inoculation of 1.5 l of LB with 50 μg/ml kanamycin (in a 3 l Belco baffled flask) with 5 ml of an overnight culture. The culture was grown at 37°C for 18 h. The cells were harvested by centrifugation (20 min at 5000 × *g*). Cell pellets were resuspended in 20 ml of Amylose A Buffer (10 mM Tris–HCl pH 7.5, 10 mM NaCl) and sonicated on a 550 Sonic Dismembrator (Fisher Scientific) (power = 5, 15 min sonication, 2 s on, 2 s off).

Cell lysates were centrifuged (45 min at 10 000 × *g*) and loaded onto an Amylose resin column, preequilibrated with 150 ml 10 mM Tris–HCl pH 7.5, 10 mM NaCl. The cell lysate was loaded at 1 ml/min. The column was washed at 2 ml/min with 300 ml Amylose A Buffer. The protein was eluted from the column at 2 ml/min with 200 ml Amylose B Buffer (10 mM Tris–HCl pH 7.5, 10 mM NaCl, 10 mM maltose).

Spectrophotometric assay of ORF22

A typical reaction of ORF22 consisted of 15 mM Tris pH 7.5, 10 mM NaCl, 0.1 mM potassium ferricyanide and 1 mM substrate of interest. The decrease in the UV/Vis absorbance at 400 nm, resulting from electron transfer to the potassium ferricyanide was monitored over time.

HPLC assay of ORF22

All HPLC reactions were analyzed on a Beckman System Gold HPLC using a Vydac™ C18 Small Pore column. The solvent system for the analysis of the Hmo reaction used solvent A and solvent B (as above). The profile for separation was 9 min isocratic development at 100% A/0% B, 15 min linear gradient from 100% A/0% B to 85% A/15% B. The column was then washed with 100% B and reequilibrated with 100% A. The flow rate for the profile was 1 ml/min.

Reactions of Hmo were performed as follows. 100 μl of Hmo (1 mg/ml) was added to a reaction consisting of 1 mM of the desired substrate in 20 mM Tris–HCl, pH 7.5. The reactions were allowed to incubate at 25°C for 24 h. The samples were acidified with TFA to precipitate the protein and protonate the carboxylate containing compounds. The reactions were then loaded (100 μl) using a Beckman System Gold 640 autosampler and run using the above conditions.

HPLC assays for the coupling of the ORF21 and ORF22 reactions

The reaction mixture (200 μl), at a pH of 7.5, contained 0.8 M potassium phosphate, 0.2 M Tris–HCl, 25 mM ascorbic acid, 0.1 mg/ml bovine liver catalase, 500 μM *p*-hydroxyphenylpyruvate, and 2 μM ORF21 and/or 85 μg of MBP-ORF22. Reactions were incubated at 25°C for either 1 h in the presence of ORF21 or ORF22 alone, or 2 h when both enzymes were included in the reaction mixtures. HPLC separation of products and reactants was as described for ORF22 characterization.

Cloning, expression and purification of ORF17/MBP

The proposed coding regions for ORF217 of the chloroeremomycin cluster was amplified by PCR from *A. orientalis* genomic DNA with primers ORF17Nterm*Nde*I (5′ GGGAATTC**CATATG**GAAATCCTAGTATTCATGGAT 3′) and ORF17Cterm*Nde*I (5′ GGGCCG**CATATG**TTATGCCCGGGGCCACCGCAGACG 3′). The introduced *Nde*I restriction sites of each primer are identified by bold-face type. The amplified products were gel purified, digested with *Nde*I, and ligated, into the *Nde*I cloning sites of pIADL-16 (Walsh lab vector encoding an N-terminal MBP fusion under the control of the T7 promoter; a derivative of pET-28b). The growth of ORF17/MBP involved inoculation of 1.5 l of LB with 50 μg/ml kanamycin (in a 3 l Belco baffled flask) with 5 ml of an overnight culture. The culture was grown at 37°C for 18 h. The cells were harvested by centrifugation (20 min at 5000 × g). Cell pellets were resuspended in 20 ml of Amylose A Buffer and sonicated on a 550 Sonic Dismembrator (Fisher Scientific) (power = 5, 15 min sonication, 2 s on, 2 s off).

Cell lysates were centrifuged (45 min at 10 000 × *g*) and loaded onto an Amylose resin column, preequilibrated with 150 ml 10 mM Tris–HCl pH 7.5, 10 mM NaCl. The cell lysate was loaded at 1 ml/min. The column was washed at 2 ml/min with 300 ml Amylose A Buffer. The protein was eluted from the column at 2 ml/min with 200 ml Amylose B Buffer.

HPLC assay of ORF17

All HPLC reactions were analyzed on a Beckman System Gold HPLC using a Vydac™ C18 small pore column. The solvent system for the analysis of the Hmo reaction used solvent A and solvent B. The profile for separation was 9 min isocratic development at 100% A/0% B, 15 min linear gradient from 100% A/0% B to 85% A/15% B. The column was then washed with 100% B and reequilibrated with 100% A. The flow rate for the profile was 1 ml/min.

Reactions of HpgT were performed as follows. One hundred μl of HpgT (1 mg/ml) was added to a reaction consisting of 1 mM of the amino donor and 1 mM of the amino acceptor and 10 μM of PLP in 20 mM Tris pH 7.5. (In the case where *p*-hydroxybenzoylformate was the desired amino acceptor, *p*-hydroxymandelate was used as the amino acceptor and 100 μl of Hmo (1 mg/ml) was included in the reaction.) The reactions were allowed to incubate at 25°C for 24 h. The samples were acidified with TFA to precipitate the protein and protonate the carboxylate containing compounds. The reactions were then loaded using an autosampler using the above conditions.

References

1. Cane, D.E., Walsh, C.T. & Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations and mutations. *Science* **282**, 63–68.
2. Marahiel, M.A. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **97**, 2651–2673.

3. Konz, D. & Marahiel, M.A. (1999). How do peptide synthetases generate structural diversity? *Chem. Biol.* **6**, R39–48.
4. Sheldrick, G.M., Jones, P.G., Kennard, O., Williams, D.H. & Smith, G.A. (1978). Structure of vancomycin and its complex with acetyl-D-alanyl-D-alanine. *Nature* **271**, 223–225.
5. Williams, D.H. (1996). The glycopeptide story-how to kill the deadly 'superbugs'. *Nat. Prod. Rep.* **13**, 469–477.
6. Zmijewski Jr., M.J., Briggs, B., Logan, R. & Boeck, L.D. (1987). Biosynthetic studies on antibiotic A47934. *Antimicrob. Agents Chemother.* **31**, 1497–1501.
7. Seto, J., Fujioka, T., Furihata, K., Kaneko, I. & Takahashi, S. (1989). Structure of complestatin, a very strong inhibitor of protease activity of complement in the human complement system. *Tetrahedron Lett.* **30**, 4987–4990.
8. Kaneko, I., Kamoshida, K. & Takahashi, S. (1989). Complestatin, a potent anti-complement substance produced by *Streptomyces lavendulae*. I. Fermentation, isolation and biological characterization. *J. Antibiot. (Tokyo)* **42**, 236–241.
9. Ciabatti, R., Kettenring, J.K., Winters, G., Tuan, G., Zerilli, L. & Cavalleri, B. (1989). Ramoplanin (A-16686), a new glycolipodepsipeptide antibiotic. III. Structure elucidation. *J. Antibiot. (Tokyo)* **42**, 254–267.
10. Kempter, C., Kaiser, D., Haag, S., Nicholson, G., Gnau, V., Walk, T., Gierling, K.H., Decker, H., Zahner, H., Jung, G. & Metzger, J.W. (1997). CDA: Calcium-dependent antibiotics from *Streptomyces coelicolor* A3 containing unusual residues. *Angew. Chem. Int. Ed. Engl.* **36** (2), 498–501.
11. Hammond, S.J., Williamson, M.P., Williams, D.H., Boeck, L.D. & Marconi, G.G. (1982). On the biosynthesis of the antibiotic vancomycin. *J. Chem. Soc. Chem. Commun.* 344–346.
12. Hammond, S.J., Williamson, D.H. & Nielsen, R.V. (1983). The biosynthesis of ristocetin. *J. Chem. Soc. Chem. Commun.* 116–117.
13. Chung, S.K., Taylor, P., Oh, Y.K., DeBrosse, C. & Jeffs, P.W. (1986). Biosynthetic studies of aridicin antibiotics. I. Labeling patterns and overall pathways. *J. Antibiot. (Tokyo)* **39**, 642–651.
14. Nicas, T.I. & Cooper, R.D.G. (1997). Vancomycin and other glycopeptides. In *Biotechnology of Antibiotics*. (Strohl, W.R., editor), pp. 363–392, Marcel Dekker, New York.
15. van Wageningen, A.M., Kirkpatrick, P.N., Williams, D.H., Harris, B.R., Kershaw, J.K., Lennard, N.J., Jones, M., Jones, S.J. & Solenberg, P.J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Biol.* **5**, 155–162.
16. Que, L.J. & Ho, R.Y.N. (1996). Dioxygen activation by enzymes in mononuclear non-heme iron active sites. *Chem. Rev.* **96**, 2607–2624.
17. Serre, L., Sailland, A., Sy, D., Boudec, P., Rolland, A., Pebay-Peyroula, E. & Cohen-Addad, C. (1999). Crystal structure of *Pseudomonas fluorescens* 4-hydroxyphenylpyruvate dioxygenase: an enzyme involved in the tyrosine degradation pathway. *Struct. Fold. Des.* **7**, 977–988.
18. Choroba, O.W., Williams, D.H. & Spencer, J.B. (2000). Biosynthesis of the vancomycin group of antibiotics: Involvement of an unusual dioxygenase in the pathway to (*S*)-4-hydroxyphenylglycine. *J. Am. Chem. Soc.* **122**, 5389–5390.
19. Lehoux, I.E. & Mitra, B. (1999). (*S*)-Mandelate dehydrogenase from *Pseudomonas putida*: mechanistic studies with alternate substrates and pH and kinetic isotope effects. *Biochemistry* **38**, 5836–5848.
20. Hosoda, J., Tani, N., Konomi, T., Ohsawa, S., Aoki, H. & Imanaka, H. (1977). Incorporation of ^{14}C-amino acids into nocardicin A by growing cells. *Agric. Biol. Chem.* **41**, 2007–2012.
21. Pascal Jr., R.A., Oliver, M.A. & Chen, Y.C. (1985). Alternate substrates and inhibitors of bacterial 4-hydroxyphenylpyruvate dioxygenase. *Biochemistry* **24**, 3158–3165.