

Deoxysugars in glycopeptide antibiotics: Enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis

Huawei Chen*, Michael G. Thomas, Brian K. Hubbard, Heather C. Losey, Christopher T. Walsh[†], and Michael D. Burkart*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115

Contributed by Christopher T. Walsh, August 17, 2000

The 2,3,6-trideoxysugar L-epivancosamine is the terminal sugar added to the aglycone scaffold in chloroeremomycin, a member of the vancomycin family of glycopeptide antibiotics. Five proteins from the chloroeremomycin biosynthetic cluster, ORF14 and ORF23 to ORF26, have been expressed heterologously in *Escherichia coli* and purified to near homogeneity, and each has been characterized for an enzymatic activity. These five enzymes reconstitute the complete biosynthesis of TDP-L-epivancosamine from TDP-4-keto-6-deoxy-D-glucose. This process involves C-2 deoxygenation, C-3 amination and methylation, C-5 epimerization, and C-4 ketoreduction. Intermediates and the final product of this pathway have been identified by mass spectrometry and NMR. The pathway established here represents the complete *in vitro* reconstitution of an unusual sugar for an important class of antibiotics and sets the groundwork for future combinatorial biosynthesis for new bioactive compounds.

Several categories of natural products contain deoxysugar moieties that are essential for their biological activities. Within these natural product categories are the glycosylated macrolactones of the erythromycin and tylosin families as well as the vancomycin family of glycopeptide antibiotics (1). Examples of the latter group include vancomycin (compound 1), balhimycin (compound 2), and chloroeremomycin (compound 3), each containing an identical crosslinked heptapeptide scaffold and the same monosaccharide, glucose, attached to the phenolic oxygen at residue 4 of the aglycone (Fig. 1). The feature that distinguishes these three molecules is the further glycosylation with an unusual deoxysugar at one or more sites of the glycopeptide. In the case of compound 1, 3-amino-2,3,6-trideoxy-3-C-methyl-L-arabino-hexose, or vancosamine, is attached by an α -1,2 linkage to the glucose moiety. In contrast, compound 2 contains the 4-oxo form of 3-methyl-3-amino trideoxyhexose attached to the β -hydroxytyrosine at position 6 (β -OH-Tyr-6) of the heptapeptide. Finally, compound 3 has the 4-epi isomer of vancosamine, epivancosamine (3-amino-2,3,6-trideoxy-3-C-methyl-L-arabino-hexopyranose), attached to both the glucose and β -OH-Tyr-6 of the glycopeptide.

Insight into the biosynthetic pathway for one of these deoxysugars, epivancosamine, was recently obtained by the sequencing of the gene cluster for biosynthesis of compound 3 from *Amycolatopsis orientalis* (2). Of the 39 *orf*s reported for the cluster, *orf23*–*orf26* were proposed to encode the enzymes catalyzing TDP-L-epivancosamine biosynthesis on the basis of amino acid homology to known sugar biosynthetic enzymes. ORF14, however, was assigned an unknown hydroxylase function. On the basis of our own homology studies, we reinterpreted the TDP-L-epivancosamine biosynthetic pathway to be as shown in Scheme 1 and attempted to validate this proposal by reconstituting the pathway *in vitro*.

Here we report on the purification of ORF14 and ORF23 to ORF26 that were heterologously expressed in *Escherichia coli*. We identify the role each ORF plays in the conversion of TDP-D-glucose, via TDP-6-deoxy-4-keto-D-glucose (compound 4), to TDP-L-epivancosamine (compound 13). This is a

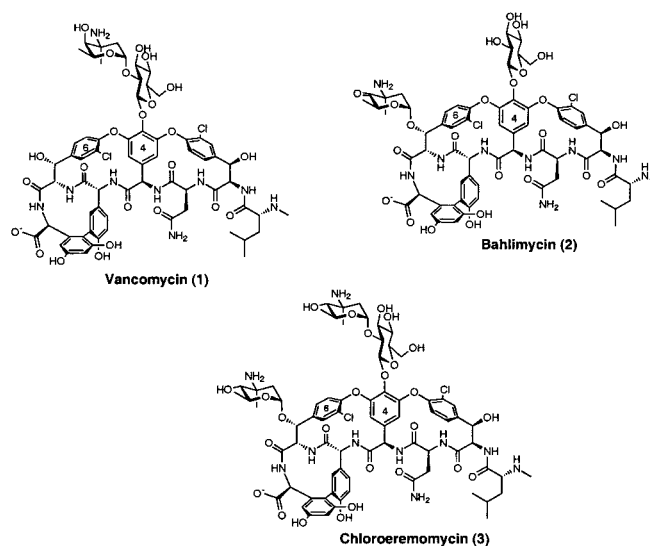


Fig. 1. Glycopeptide antibiotics of the vancomycin family.

complete *in vitro* reconstitution of an unusual deoxysugar biosynthetic pathway for an important class of antibiotics. The pathway established here sets the groundwork for future combinatorial biosynthetic approaches to generate novel bioactive compounds.

Materials and Methods

Cloning, Overexpression, and Purification of ORF14 and ORF23 to ORF26. The proposed epivancosamine biosynthetic genes, *orf14* and *orf23*–*orf26*, were cloned into pET vectors by using standard cloning procedures (Table 1). Each protein was overexpressed in *E. coli* BL21(DE3) harboring the appropriate plasmid, and the cell culture was grown in LB medium supplemented with a suitable antibiotic. For ORFs 14, 25, and 26, the cells were grown at 37°C until an OD₆₀₀ of 0.6 was reached, induced with 1 mM, 50 μ M, and 1 mM isopropyl β -D-thiogalactoside, respectively, and allowed to grow at 24°C for 4 additional hours. For ORF23, the cells were grown at 37°C for 18 h without induction. For ORF24, the cells were grown at 15°C for 72 h without induction.

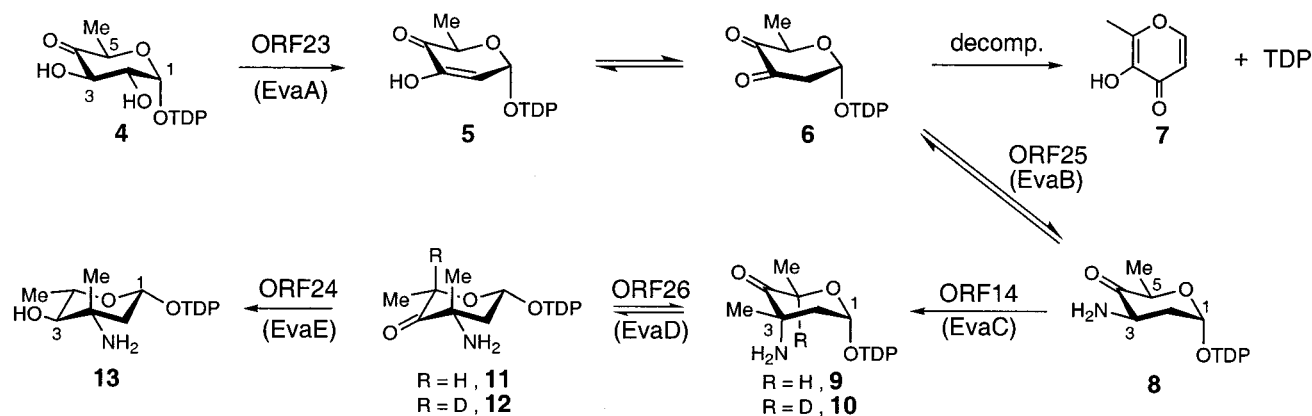
Abbreviations: SAM, S-adenosylmethionine; PLP, pyridoxal 5-phosphate; SAX, strong anion exchange; MALDI-MS, matrix-assisted laser desorption/ionization mass spectrometry; rt, retention time; NOE, nuclear Overhauser effect.

*H.C. and M.D.B. contributed equally to this work.

[†]To whom reprint requests should be addressed. E-mail: walsh@walsh.med.harvard.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.210395097. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.210395097



Scheme 1. Biosynthetic pathway of TDP-L-epivancosamine.

Cells from each 3 × 1-liter culture were harvested by centrifugation (10 min at 5,000 × g) and resuspended in 30 ml of 1 × bind buffer (20 mM Tris-HCl, pH 7.9 at 4°C/0.5 M NaCl/5 mM imidazole). Resuspended cells were broken by a French press (three passes), and cell debris was removed by centrifugation (30 min at 16,000 × g). The soluble protein was purified with a single Ni-affinity chromatographic step by using HisBind (Novagen), Ni-NTA (Qiagen, Chatsworth, CA), or HiTrap (Amersham Pharmacia) resin and following the manufacturer's instructions.

HPLC Methods. All HPLC assays were conducted with a Beckman System Gold HPLC by using an Adsorbosphere strong anion-exchange (SAX) analytical column (5 μm, 4.6 × 250 mm, Alltech Associates), with a linear gradient from 50 to 500 mM potassium phosphate buffer (pH 3.5). Samples for matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) analysis were isolated with a linear gradient between 50 and 500 mM ammonium formate buffer (pH 3.5). A semipreparative Adsorbosphere SAX column (5 μm, 25 × 250 mm, Alltech Associates) was used to separate material for NMR experiments with the same elution profile as that for purifying the MS samples. The pH of the collected samples was immediately adjusted to ≈7.0 with 1 M NH₄OH.

ORF23 Assays. A reaction mixture (100 μl) containing 50 mM Tris-HCl (pH 7.5), 2.5 mM substrate (compound 4), and 10 μM ORF23 was incubated at 24°C for 2 h, and then the reaction was terminated by heat denaturation of the protein at 90°C for 3 min. Substrate consumption and TDP release (Scheme 1) were assessed by SAX-HPLC analysis. To identify maltol (compound 7) formation, the reaction mixture was extracted twice with ethyl acetate (200 μl), the organic solvent was removed by Speedvac, and the residue was analyzed by reverse-phase HPLC with a

Vydac C₁₈ small-pore analytical column (5 μm, 4.6 × 250 mm), using a linear gradient of 0 to 50% (vol/vol) acetonitrile in water over 30 min. A peak with the same retention time (rt = 12.1 min) as the authentic sample of maltol (Aldrich) was collected and subjected to MALDI-MS analysis.

ORF25 Assays. Because of very low hydrolytic stability of the putative substrate of ORF25 (compound 6), a coupled assay was performed with ORF23 included in the reaction to generate compound 6 *in situ*. A typical reaction mixture of 100 μl contained 50 mM Tris-HCl (pH 7.5), 1 mM compound 4, 0.1 mM pyridoxal 5-phosphate (PLP), 20–50 mM L-glutamate, 10 μM ORF23, and 50 μM ORF25. After a 2-h incubation at 24°C, the reactions were terminated by heat treatment and subjected to SAX-HPLC analysis. The new peak with rt = 8.72 min was collected, lyophilized, and analyzed by MALDI-MS. To determine possible amino donors, six L-amino acids, Glu, Gln, Asp, Asn, Ala, and Val, and Gly were screened for activity in the coupled assay above. The reactions were terminated at 30 min, and products were assayed by SAX-HPLC. Addition of exogenous PLP from 0.01 to 1.0 mM was examined for enhancing enzymatic activity of ORF25 in 30-min incubations at 24°C.

Thymidine-5'-diphospho-3-amino-2,3,6-trideoxy-D-threo-hexopyranose-4-ulose (Compound 8). A 1.8-ml incubation mixture contained 50 mM Tris-HCl (pH 7.5), 3 mM compound 4, 50 mM glutamate, 10 μM ORF23, and 80 μM ORF25, and was incubated for 5 h at 24°C. After passing through a Centricon 10 to remove proteins, the product was isolated with semipreparative HPLC and lyophilized to give 0.86 mg of compound 8 (30% yield). ¹H NMR (D₂O, 400 MHz) δ 1.05 (d, *J* = 6.4 Hz, 3 H, 5-Me), 1.73 (d, *J* = 1.2, 3 H, 5''-Me), 1.82–1.90 (m, 1 H, 2-H_a), 2.07–2.24 (m, 3 H, 2-H_b, 2'-H_s), 3.52 (dd, *J* = 12.8, 4.8, 1 H, 3-H), 3.94–4.04

Table 1. Cloning of the proposed epivancosamine biosynthetic genes

Gene	5' primer (5'-3')	3' primer (5'-3')	Vector	Restriction sites	Predicted mass, kDa
<i>orf14</i>	GAATTC CATATG TCGGTCACGTCCCA	TGATG TCTCGAGG CTCACGTGCACGTT	pET22b(+)	<i>Nde</i> I/ <i>Xho</i> I	45
<i>orf23</i>	GGGAATTT CATATG TCGTCCTTCGTCGTTCCATCA	AAACCG TCTCGAGT CATGCACCTCCCCGAGGCTGGGC	pET24b(+)	<i>Nde</i> I/ <i>Xho</i> I	53
<i>orf24</i>	GGTCGAT CATATGA AGCTGATCACCG	ATACAC TCTCGAGT CATGCAGCGAGC	pET16b(+)	<i>Nde</i> I/ <i>Xho</i> I	34
<i>orf25</i>	GGAATTC CATATG ACCACGCGTGTATGGGAT	GGGTA ACTCGAGG GTGCAGCACTTCGCG	pET16b(+)	<i>Nde</i> I/ <i>Xho</i> I	41
<i>orf26</i>	GGAGC ACATATG CAAGCACGCAAACCTCG	CAG AAGCTT GGTCGAGACGGGGACCG	pET22b(+)	<i>Nde</i> I/ <i>Hind</i> III	23

The proposed epivancosamine biosynthetic genes were amplified by PCR from *Amycolatopsis orientalis* genomic DNA with listed primer pairs. The restriction site introduced by each primer is shown in boldface. The PCR products were gel purified, digested with the appropriate restriction enzymes, and ligated into the designated vector with matching restriction sites.

(m, 4 H, 5, 4', 5'-Hs), 4.40–4.46 (m, 1 H, 3'-H), 5.48 (m, 1 H, 1-H), 6.16 (dd, $J = 7.2, 6.4$, 1 H, 1'-H), 7.57 (d, $J = 1.2$, 6''-H). MALDI-MS calculated for $C_{16}H_{24}N_3O_{13}P_2^-$ ($M - H^+$) 528.078, found 528.095.

ORF14 Assays. Incubation mixtures containing 50 mM Tris-HCl (pH 7.5), 2.5 mM compound **4**, 50 mM L-glutamate, 10 mM S-adenosylmethionine (SAM), 10 μ M ORF23, 80 μ M ORF25, and 70 μ M ORF14 in a total volume of 100 μ l were incubated overnight at 24°C. The reactions were terminated by heat treatment and the products were analyzed by SAX-HPLC. A new peak with $rt = 7.50$ min was collected, lyophilized, and analyzed by MALDI-MS.

Thymidine-5'-diphospho-3-amino-2,3,6-trideoxy-3-C-methyl-D-erythro-hexopyranos-4-ulose (Compound 9). A 1-ml solution with 3.6 mM compound **4**, 50 mM Tris-HCl (pH 7.5), 10 mM SAM, 50 mM glutamate, 100 μ M PLP, 70 μ M ORF14, 14 μ M ORF23, and 80 μ M ORF25 was incubated at 24°C for 12 h and subsequently passed through a Centricon 10 to remove proteins. Semipreparative SAX-HPLC purification followed by lyophilization yielded pure compound **9** (816 μ g, 42% yield). Proton NMR decoupling and nuclear Overhauser effect (NOE) experiments were performed to verify the structural assignments of compound **9**. 1H NMR (D_2O , 400 MHz) δ 1.11 (d, $J = 6.3$ Hz, 3 H, 5-Me), 1.14 (s, 3 H, 3-Me), 1.78 (d, $J = 1.1$, 3 H, 5''-Me), 2.04 (m, 2 H, 2-Hs), 2.22 (m, 2 H, 2'-Hs), 4.02 (m, 3 H, 4', 5'-Hs), 4.25 (q, $J = 6.3$, 5-H), 4.44 (m, 1 H, 3'-H), 5.53 (ddd, $J = 7.0, 2.6, 2.5$, 1-H), 6.19 (dd, $J = 6.0, 7.0$, 1 H, 1'-H), 7.57 (d, $J = 1.1$, 6''-H). MALDI-MS calculated for $C_{17}H_{26}N_3O_{13}P_2^-$ ($M - H^+$) 542.094, found 542.153.

ORF26 Assays. To assay ORF26, an incubation mixture (100 μ l) containing 20 mM Tris-HCl (pH 7.5), 0.5 mM methylated aminosugar (compound **9**), and 7 μ M ORF26 was performed at 24°C. After a 1-h incubation, the reaction was quenched by heating at 90°C for 3 min, and both the substrate peak ($rt = 7.50$ min) and a product peak ($rt = 9.03$ min) were isolated by SAX-HPLC. Their mass was determined by MALDI-MS after lyophilization. To support that ORF26 functions as a C-5 epimerase with a protonation/deprotonation mechanism, a D_2O incubation was performed. The reaction mixture (50 μ l) contained 0.72 mM compound **9**, 30 mM ORF26, and 20 mM Tris-DCl (pD 7.8) and was incubated at 24°C for 4 h. Both the substrate and product peaks were collected from SAX-HPLC and subjected to MALDI-MS analysis.

ORF24 Assays. The reaction catalyzed by ORF26 is reversible, and the epimerized product (compound **11**) accounted for $\approx 15\%$ by HPLC integration. Hence, a coupled reaction was performed in which ORF26 was included to generate compound **11** *in situ*. A reaction mixture (100 μ l) containing 20 mM Tris-HCl (pH 7.5), 1.0 mM compound **9**, 2 mM NADPH, 5 μ M ORF26, and 5 μ M ORF24 was incubated at 24°C for 2 h. The reaction was terminated by diluting 20 μ l of the incubation mixture into 100 μ l of 50 mM potassium phosphate buffer (pH 3.5). The new product ($rt = 9.63$ min) was collected from SAX-HPLC and the mass was determined by MALDI-MS.

Thymidine-5'-diphospho-3-amino-2,3,6-trideoxy-3-C-methyl-L-arabino-hexose (Thymidine-5'-diphosphate-4-epi-L-vancosamine) (Compound 13). A 2-ml reaction mixture of pure compound **9** (0.9 mM) in 100 mM Tris-HCl (pH 7.5)/20 μ M ORF26/50 mM ORF24/100 μ M PLP/10 mM NADPH was incubated at 24°C for 2 h and subsequently passed through a Centricon 10 to remove proteins. Semipreparative SAX-HPLC purification followed by lyophilization yielded pure compound **13** (0.85 mg, 87% yield). 1H NMR (D_2O , 400 MHz) δ 1.11 (d, $J = 6.4$ Hz, 3 H, 5-Me), 1.25 (s, 3 H,

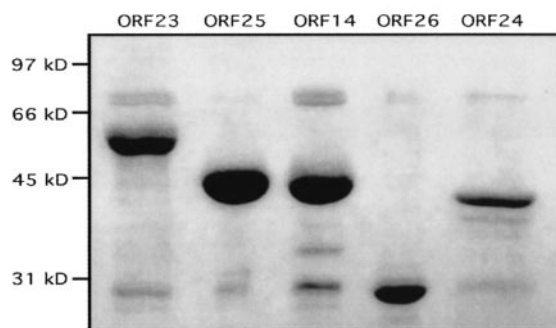


Fig. 2. SDS/10% PAGE of purified proteins heterologously expressed in *E. coli*.

3-Me), 1.73 (d, $J = 1.2$, 3 H, 5''-Me), 2.14–2.22 (m, 4 H, 2-Hs, 2'-Hs), 3.21 (d, $J = 9.6$, 1 H, 4-H), 3.54 (m, 1 H, 5-H), 3.94–4.02 (m, 3 H, 4', 5'-Hs), 4.42 (m, 1 H, 3'-H), 5.24 (ddd, $J = 9.0, 8.7, 2.0$, 1 H, 1-H), 6.16 (dd, $J = 7.0, 6.5$, 1 H, 1'-H), 7.57 (d, $J = 1.2$, 6''-H). MALDI-MS calculated for $C_{17}H_{28}N_3O_{13}P_2^-$ ($M - H^+$) 544.110, found 544.115.

Results

Production of Epivancosamine Biosynthetic Enzymes. ORF14 and ORFs 23–26 were overexpressed in *E. coli* and purified in soluble form with the expected molecular mass (Table 1), as determined by SDS/PAGE (Fig. 2). Each enzyme was purified by a single nickel-affinity column chromatography to near homogeneity. The yields for each enzyme are as follows in mg/liter: 8, ORF14; 10, ORF23; 20, ORF25; 8, ORF24; and 8, ORF26. Only ORF25 contained an organic chromophore attributed to PLP, with two absorption maxima at 335 and 418 nm, and was in partial holo-form as judged by UV-visible absorbance. ORF24 was insoluble when expressed above 15°C.

ORF23 and ORF25 Catalyze the Formation of Compound 8. The anticipated substrate for ORF23, TDP-6-deoxy-4-keto-D-glucose (compound **4**) was generated enzymatically from TDP-D-glucose with the enzyme RfbB from the rhamnose pathway (a gift from H.-w. Liu, University of Minnesota) (1). Pure ORF23, a putative 2,3-dehydratase, eliminated the C-2 hydroxyl of compound **4** to yield the 2,6-dideoxy-3,4-diketo-TDP sugar, compound **6**, as judged by its facile decomposition to compound **7**, a route previously described (3, 4). Further characterization of compound **6** was not attempted, but the decomposition product compound **7** was compared with an authentic sample with RP-HPLC and MALDI-MS analysis. The coupling of ORF23 with ORF25, a proposed transaminase, in the presence of compound **4**, with PLP as a coenzyme and L-glutamate as a cosubstrate, produced a new compound, with a rt significantly shorter than that for compound **4** (Fig. 3, *b* vs. *a*), suggesting a positively charged amino group was introduced by ORF25. The new compound was purified by SAX-HPLC and identified as thymidine-5'-diphospho-3-amino-2,3,6-trideoxy-D-threo-hexopyranos-4-ulose (compound **8**) by MALDI-MS (Fig. 4a) and proton NMR analysis. From the proton NMR spectrum of compound **8**, the C-2 deoxygenation was confirmed by the presence of two C-2 methylene protons (δ 1.84 and 2.13 ppm) from the hexose moiety. Additionally, the apparent *trans* coupling constant ($J = 12.6$ Hz) of 3-H and 2-H_a indicated that 3-H assumes an axial position, thus the C-3 amino group is equatorial.

Experiments with six L-amino acids as amine donors, Glu, Gln, Asp, Asn, Ala, and Val, and Gly for ORF25 identified L-Glu as the preferred cosubstrate. The only other competent amine

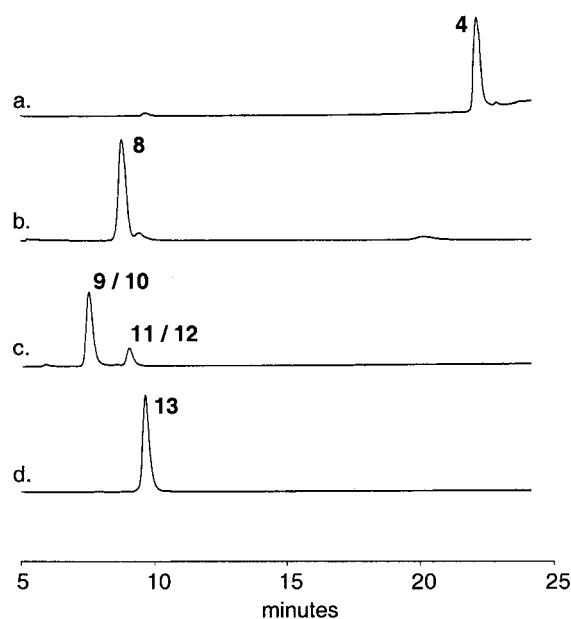


Fig. 3. SAX-HPLC traces of enzymatic products. Elution profile: linear gradient from 0 to 2.5% B over 10 min, followed by a steeper gradient of 2.5 to 100% B over 10 min. Buffer A = 50 mM potassium phosphate buffer (pH 3.5), and B = 500 mM potassium phosphate buffer (pH 3.5)

donor was L-Ala, which showed 40% activity compared with L-Glu (data not shown). Optimal product formation by ORF25 occurred when 0.1 mM PLP was included in the reaction mixture.

In the ORF23- and ORF25-coupled incubation, there was still a significant amount of decomposition of compound **6** even with a 7-fold excess of ORF25. This decomposition was minimized by

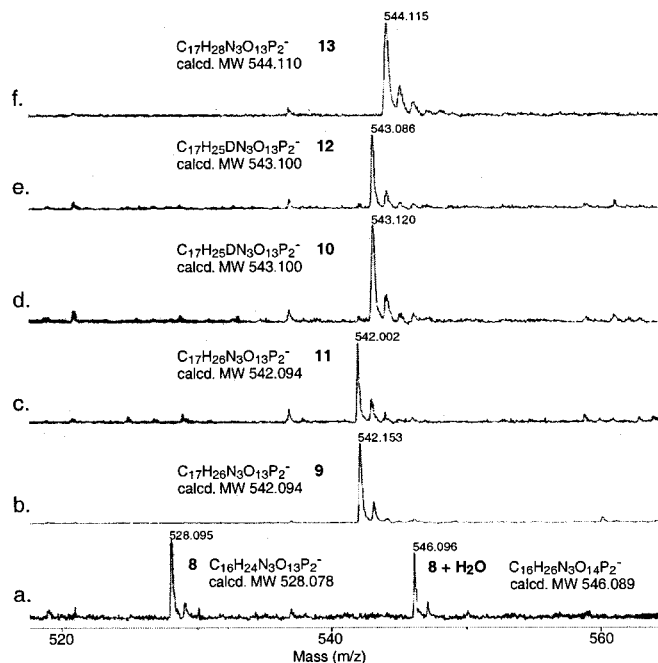


Fig. 4. MALDI-MS spectra of enzymatic products in TDP-L-epivancosamine biosynthetic pathway [PerSeptive Biosystems (Framingham, MA) Voyager 4036, negative reflector mode, 3000 laser intensity, and 2,5-dihydroxybenzoic acid as matrix].

reducing the incubation time to only 5 h. Longer incubation resulted in poor yield of compound **8**, because decomposition of compound **6** shifted the equilibrium of the reversible transamination reaction catalyzed by ORF25, leading to the net consumption of product (Scheme 1).

ORF14 Catalyzes the Formation of Compound 9. Incubations of compound **4** with ORF23, -25, and -14, and cosubstrate SAM, produced a compound that was identified as thymidine-5'-diphospho-3-amino-2,3,6-trideoxy-3-C-methyl-D-erythro-hexopyranos-4-ulose by MALDI-MS (Fig. 4b) and NMR analysis of the HPLC-purified compound. From the proton NMR spectrum of compound **9**, two C-2 methylene protons (δ 2.04 ppm) from the hexose moiety were clearly present. A methyl peak at 1.14 ppm confirmed the C-methylation of compound **8** by ORF14. Additionally, the H-3 signal was absent from the spectrum, indicating the location of the methylation. Negative NOE between the C-3 methyl and H-5 indicated no through-space coupling of these protons, consistent with an equatorial C-3 methyl group.

ORF26 Catalyzes the Formation of Compound 11. Incubation of compound **9** with ORF26 generated a new species (rt = 9.03 min) as judged by SAX-HPLC analysis. This species has the same molecular mass as compound **9** (Fig. 4c), but is clearly distinct from compound **9** by chromatography (Fig. 3c). It was anticipated that ORF26 functions as a C-5 epimerase in the TDP-L-epivancosamine pathway. To verify the epimerase activity, ORF26 was incubated with compound **9** in deuterated buffer. Because the epimerase reaction should reach an equilibrium after sufficient incubation time in deuterated buffer, both compounds **9** and **11** should have incorporated one deuterium atom at C-5 to form compounds **10** and **12**, respectively. Both substrate and product from a fully equilibrated reaction in D_2O were isolated by HPLC, and MALDI-MS confirmed that both molecules incorporated one deuterium atom and increased their mass by one unit (Fig. 4 b–e), consistent with ORF26 functioning as a C-5 epimerase.

ORF26 + ORF24 Catalyze the Formation of Compound 13 from Compound 9. Incubations of compound **9** with ORF26, ORF24, and cosubstrate NADPH generated a product that has a molecular mass of 544.115 Da (Fig. 4f) equivalent to TDP-L-epivancosamine, compound **13**, after HPLC purification (Fig. 3d). Proton NMR spectral analysis verified the structure assignment of compound **13**.

From the proton NMR spectrum of compound **13** (Fig. 5), the H-1 proton displays a ddd splitting pattern ($J = 9.0, 8.7,$ and 2.0 Hz). The *trans* coupling constant of 1-H and the axial 2-H indicated that the conformation of compound **13** was now in the 1C_4 chair with a *trans*-diaxial relationship between 1-H and 2-H_a. The *trans* coupling constant of 8.7 Hz was resolved from a decoupling experiment by irradiating at the H-2 signals, which simplified the H-1 signal to a doublet of $J = 9$ Hz, corresponding to the 1-H and phosphorus coupling. Epimerization at C-5 induces this ring-flip from 4C_1 to release strain from unfavorable 1,3,5-triaxial steric interactions. A new signal at 3.21 ppm represented the 4-H proton introduced by NADPH reduction of compound **11** with ORF24, which couples to H-5 in a *trans*-diaxial configuration, with $J = 9$ Hz. Positive NOE among 1-H, 5-H, and C-3 methyl demonstrates the triaxial through-space interaction among these groups, further confirming the overall conformation of compound **13** to be the 1C_4 chair.

Discussion

Both the macrolide class of antibiotics and the vancomycin group of glycopeptide antibiotics have aglycone moieties assembled first, followed by decoration with sugars to produce the final

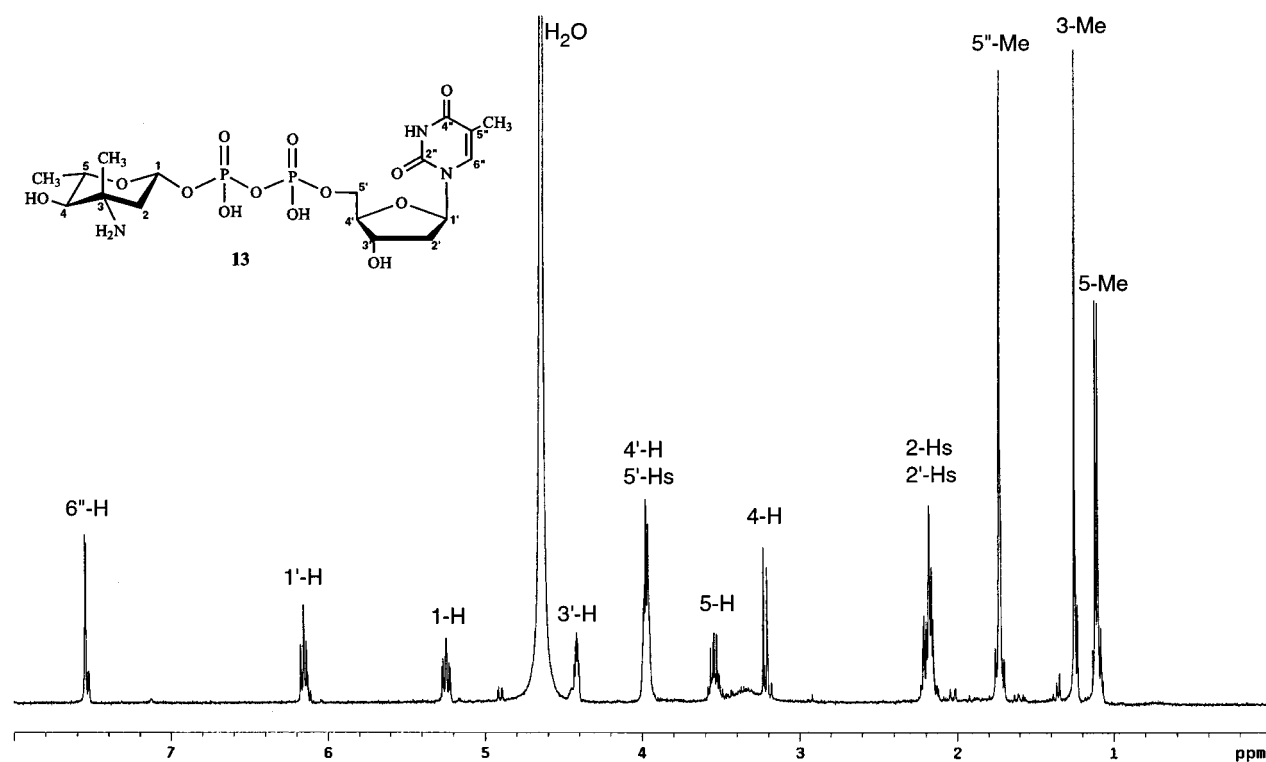


Fig. 5. Proton NMR (D_2O) spectrum of purified TDP-L-epivancosamine (13).

compound (2, 5–7). Many of the sugars are deoxyhexoses (1), which are anticipated to create polar moieties with hydrophobic surfaces that may be important for architecture and conformation, solubility, and receptor recognition. For example, in the vancomycin family of antibiotics, the deoxysugars, vancosamine in vancomycin (compound 1), and 4-epivancosamine in chloroeremomycin (compound 3), have been shown to affect back-to-back dimerization constants of the antibiotic molecules (8), correlated with some attributes of potency against sensitive bacteria (9, 10).

The vancosamine/epivancosamine isomers are both 2,3,6-trideoxy-L-hexoses in which C-3 is substituted with both a C-methyl and an amino group. This amino group is reductively alkylated synthetically in LY333328 (9) to increase membrane association. The enzymes dedicated to producing the deoxysugars decorating these antibiotics are often clustered together with the biosynthetic genes for the aglycone, presumably enabling the coordinated production of the complementary moieties of the antibiotic. The sequence of the 39 *orf*s in the chloroeremomycin biosynthetic cluster identified several ORFs as candidates for the conversion of UDP-D-glucose or TDP-D-glucose to the U(T)DP-L-epivancosamine (2). In the work presented here, five of these ORFs (14 and 23–26) have been heterologously expressed in *E. coli*, purified to near homogeneity, and shown to convert TDP-4-keto-6-deoxy-D-glucose (compound 4) to TDP-L-epivancosamine (compound 13). This process involves the formation of two additional deoxy centers at C-2 and C-3, amination and methylation at C-3, epimerization at C-5, and ketoreduction at C-4.

The introduction of the three deoxy centers appears to follow precedents in C-6-deoxysugar, C-2,6-dideoxysugar, and C-3,6-dideoxysugar biosynthesis (3, 4, 11–14). By analogy, the enzyme for conversion of TDP-D-glucose to the 4-keto-6-deoxyglucose nucleotide is not encoded within the *Cep* gene cluster but is likely present elsewhere in the chromosome. Therefore, we used an

enzyme from the rhamnose biosynthetic pathway, RfbB (15), as a reagent to produce the 4-keto-6-deoxysugar nucleotide compound 4 as the starting material. The 2-deoxygenation is typically affected by a 2,3-dehydratase (3, 4), a function fulfilled here by ORF23. The initial enol product ketonizes to the 3,4-dioxo-6-deoxysugar nucleotide compound 6, which is known to be unstable toward elimination of the TDP substituent, with 1,2-olefin formation (3, 4). This degradation product, compound 7, was detected by reverse-phase HPLC and MALDI-MS analysis when ORF23 was the only enzyme in the incubation and the intermediate compound 6 had no other outlet. When ORF25 and the PLP cofactor were added as the second enzyme in the presence of the cosubstrate L-glutamate, a regio- and stereospecific transamination was effected at C-3 to produce the 2,3,6-trideoxy skeleton with a C-3 amino substituent. This is a net transamination because L-glutamate is oxidized to α -ketoglutarate, and it draws off compound 6 to accumulate as compound 8, which can be purified by SAX-HPLC. Precedent for other transamination enzymes in deoxysugar biosynthesis has been reported for the biosynthesis of mycaminose in the tylosin pathway (14).

The remaining enzymes, ORFs 14, 26, and 24, may be the most intriguing of the five studied here. ORF14 shows signature motifs of a SAM-binding enzyme (16), and results here show that ORF14, when added together with its cosubstrate SAM, as the third enzyme to incubations starting from compound 4, or used on its own starting from compound 8, cleanly produces compound 9. This C-methylation undoubtedly is enabled by the 4-keto group of compound 8 that facilitates a low-energy C-3 carbanion/enolate as an attacking nucleophile on the electrophilic CH_3 moiety of SAM.

At this point in the biosynthesis, compound 9 has quaternary substitution at C-3, making this center no longer enolizable. Thus, the configuration of the C-3 methyl and amino is now set to that found in vancosamine and epivancosamine. The only

remaining enolizable center is at C-5, and this is the site acted on by the next enzyme. When ORF26 was added as the fourth enzyme starting from compound **4** or as the only enzyme starting from isolated samples of compound **9**, a new peak was detected by HPLC which had the identical mass spectrum as compound **9** and was assigned the C-5-isomerized product compound **11** in Scheme 1. Epimerization of compound **9** at C-5 by ORF26 will take the C-5-CH₃ substituent from equatorial to axial, an energetically disfavored transformation, which is followed by a flipping of the pyranose ring, releasing strain, and producing the conformer shown as compound **11**. This is now an L-sugar nucleotide rather than a D-sugar nucleotide and is the β -anomer rather than the α -, as in all previous metabolites. We verified the epimerization activity of ORF26 by incubating the enzyme and compound **9** in D₂O for several hours to allow full deuterium incorporation at C-5 for both compounds **9** and **11**. The deuterated molecules, compounds **10** and **12**, were verified by MALDI-MS analysis after HPLC purification (Fig. 4 *b–e*).

The final step in TDP-L-epivancosamine biosynthesis is the reduction of the C-4 ketone in a stereospecific manner to yield an equatorial hydroxyl of the epivancosamine moiety. Presumably, it is the directionality of hydride transfer by 4-ketoreductase that differentiates the epivancosamine biosynthetic pathway from that of vancosamine, which displays opposite stereochemistry of the C-4 hydroxyl group. Incubation of compound **9** with ORF26, ORF24, and NADPH resulted in the end product of the five-enzyme pathway, TDP-L-epivancosamine, compound **13**. To better reflect the role each of these *orfs* plays in epivancosamine

biosynthesis, we recommend the nomenclature of *evaA*, *evaB*, *evaC*, *evaD*, and *evaE*, for *orfs* 23, 25, 14, 26, and 24, respectively.

Given the structural similarities of the carbohydrate moieties in the vancomycin family, it is likely that the carbohydrate biosynthetic enzymes will bear considerable homology to each other. Vancomycin-producing strains should contain the same set of enzymes to generate compound **11**, but with a distinct C-4 ketoreductase to produce the C-4 epimer. In the bahlmicycin cluster, it is likely that the enzyme corresponding to ORF24 is defective or absent, suggesting that a glycosyltransferase would take compound **11** as a substrate in this particular *Amycolocap-tosis* species.

Finally, the *in vitro* reconstitution of the TDP-L-epivancosamine biosynthetic pathway with five pure enzymes illuminates the molecular logic for the creation of this highly tailored amino-methyl-trideoxy sugar series. Furthermore, it sets the groundwork for investigation into the enzymatic formation of amino-deoxy-hexose analogs and their transfer to different aglycones and partially glycosylated aglycone scaffolds by dedicated glycosyltransferases, ultimately leading to the generation of new antibiotics.

Note Added in Proof. After submission of this manuscript, a report on expression and assay of ORF26 as an epimerase appeared (17).

We thank Prof. Hung-wen Liu, University of Minnesota, for the generous gift of RfbB enzyme. This work was sponsored by National Institutes of Health Grant GM 49338.

1. Johnson, D. A. & Liu, H.-w. (1999) in *Comprehensive Natural Products Chemistry*, ed. Meth-Cohn, O. (Elsevier Science, Oxford), Vol. 3, pp. 311–365.
2. 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) *Chem. Biol.* **5**, 155–162.
3. Draeger, G., Park, S.-H. & Floss, H. G. (1999) *J. Am. Chem. Soc.* **121**, 2611–2612.
4. Chen, H., Agnihotri, G., Guo, Z., Que, N. L. S., Chen, X. H. & Liu, H.-w. (1999) *J. Am. Chem. Soc.* **121**, 8124–8125.
5. Gaisser, S., Bohm, G. A., Cortes, J. & Leadlay, P. F. (1997) *Mol. Gen. Genet.* **256**, 239–251.
6. Gaisser, S., Bohm, G. A., Doumith, M., Raynal, M. C., Dhillon, N., Cortes, J. & Leadlay, P. F. (1998) *Mol. Gen. Genet.* **258**, 78–88.
7. Katz, L. & Donadio, S. (1995) *Genetics and Biochemistry of Antibiotic Production* (Butterworth-Heinemann, Boston), pp. 385–420.
8. Williams, D. H., Maguire, A. J., Tsuzuki, W. & Westwell, M. S. (1998) *Science* **280**, 711–714.
9. Cooper, R. D., 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, B. E., *et al.* (1996) *J. Antibiot.* **49**, 575–581.
10. Ge, M., Chen, Z., Onishi, H. R., Kohler, J., Silver, L. L., Kerns, R., Fukuzawa, S., Thompson, C. & Kahne, D. (1999) *Science* **284**, 507–511.
11. Snipes, C. E., Brillinger, G. U., Sellers, L., Mascaro, L. & Floss, H. G. (1977) *J. Biol. Chem.* **252**, 8113–8117.
12. Yu, Y., Russell, R. N., Thorson, J. S., Liu, L. D. & Liu, H.-w. (1992) *J. Biol. Chem.* **267**, 5868–5875.
13. Liu, H.-w. & Thorson, J. S. (1994) *Annu. Rev. Microbiol.* **48**, 223–256.
14. Chen, H., Yeung, S.-M., Que, N. L. S., Muller, T., Schmidt, R. R. & Liu, H.-w. (1999) *J. Am. Chem. Soc.* **121**, 7166–7167.
15. Jiang, X. M., Neal, B., Santiago, F., Lee, S. J., Romana, L. K. & Reeves, P. R. (1991) *Mol. Microbiol.* **5**, 695–713.
16. Kagan, R. M. & Clarke, S. (1994) *Arch. Biochem. Biophys.* **310**, 417–427.
17. Kirkpatrick, P. N., Scaife, W., Hallis, T. M., Liu, H.-w., Spencer, J. B. & Williams, D. H. (2000) *Chem. Commun.*, 1565–1566.