Crystal Structure of Vancosaminyltransferase GtfD from the Vancomycin Biosynthetic Pathway: Interactions with Acceptor and Nucleotide Ligands[†,‡]

Anne M. Mulichak,[§] Wei Lu,[||] Heather C. Losey,[||] Christopher T. Walsh,[||] and R. Michael Garavito*[,§]

Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824-1319, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, 02115

Received November 26, 2003; Revised Manuscript Received March 2, 2004

ABSTRACT: The TDP-vancosaminyltransferase GtfD catalyzes the attachment of L-vancosamine to a monoglucosylated heptapeptide intermediate during the final stage of vancomycin biosynthesis. Glycosyltransferases from this and similar antibiotic pathways are potential tools for the design of new compounds that are effective against vancomycin resistant bacterial strains. We have determined the X-ray crystal structure of GtfD as a complex with TDP and the natural glycopeptide substrate at 2.0 Å resolution. GtfD, a member of the bidomain GT-B glycosyltransferase superfamily, binds TDP in the interdomain cleft, while the aglycone acceptor binds in a deep crevice in the N-terminal domain. However, the two domains are more interdependent in terms of substrate binding and overall structure than was evident in the structures of closely related glycosyltransferases GtfA and GtfB. Structural and kinetic analyses support the identification of Asp13 as a catalytic general base, with a possible secondary role for Thr10. Several residues have also been identified as being involved in donor sugar binding and recognition.

The vancomycin and related glycopeptide antibiotics have been widely used as effective agents against Gram-positive bacterial infections. Moreover, vancomycin has been the drug of last resort for the treatment of life-threatening infections by Gram-positive pathogens, many of which are resistant to most other antibiotics (*1*). However, there has been an explosive rise of vancomycin resistant enterococci in nosocomial infections, resulting in elevated rates of morbidity and mortality. More recently, vancomycin resistance has been reported for the first time in methicillin resistant *Staphylococcus aureus*, one of the most common causes of both hospital- and community-acquired infection, and a pathogen for which vancomycin had been considered the only effective treatment. This serious and growing threat has created an urgent need for novel antibiotics active against resistant bacterial strains (*2*−*4*).

The vancomycin group antibiotics are natural products that act by inhibiting bacterial cell wall biosynthesis. Structurally, these molecules consist of an oxidatively cross-linked heptapeptide core; the vancomycin group antibiotics differ chemically due to variable glycosylation (Figure 1a). The identity and position of the sugar constituents influence antibiotic bioactivity, both by increasing solubility and by mediating important binding interactions. The sugar moieties are added in the final stages of biosynthesis, transferred from an activated NDP-sugar[1] donor by a series of homologous Gtfs, which are highly specific with respect to both substrate and regioselection (Figure 1b). During vancomycin biosynthesis, GtfE is the first of these Gtfs to act on the aglycone core, transferring D-glucose to the phenolic hydroxyl of OH−Phegly$_4$ to form a DVV intermediate. Then, a second enzyme, GtfD, attaches L-vancosamine to the glucose O2′ hydroxyl oxygen, generating the disaccharide chain containing vancomycin. In the pathway of the related antibiotic chloroeremomycin, there are three sites of sugar attachment and three corresponding Gtfs. Glucose is first added by GtfB, an enzyme structurally and functionally homologous to GtfE. In the second step, L-*epi*-vancosamine is transferred to DVV at a third attachment site, the β-OH-Tyr$_6$ position, in a reaction catalyzed by the unique GtfA enzyme. Last, GtfC, the corresponding isoform of GtfD, transfers L-*epi*-vancosamine to form an analogous epi-vancosaminyl-glucose disaccharide.

Variation of these glycosylation sites is one potential avenue for the generation of novel glycopeptide analogues with enhanced potency or novel activity. Even modest structural modifications of vancomycin have been shown to overcome resistance (*5*−*7*). Moreover, Kahne and co-workers (*4*, *8*) demonstrated that carbohydrate-modified derivatives may operate by a different mechanism than vancomycin. Although synthetic vancomycin analogues are prohibitively difficult to produce due to the complex chemistry involved, the Gtfs from these natural biosynthetic pathways are a promising tool for the combinatorial design of new semi-

[†] This work was supported by grants from the National Institute of General Medical Sciences (GM49338 to C.T.W. and GM65501 to R.M.G.) and the Michigan Life Science Corridor (R.M.G.). DND-CAT is supported by E.I. DuPont de Nemours & Co., Dow Chemical Company, NSF Grant DMR-9304725, and a grant from the State of Illinois (IBHE HECA NWU 96).

[‡] Atomic coordinates have been deposited in the Protein Data Bank as entry 1RRV.

* To whom correspondence should be addressed. Tel: 517-355-9724. Fax: 517-353-9334. E-mail: garavito@msu.edu.

[§] Michigan State University.

[||] Harvard Medical School.

[1] Abbreviations: DVV, desvancosaminyl vancomycin; Gtf, glycosyltransferase; NDP, nucleotide diphosphate; Phegly, phenylglycine; rmsd, root-mean-square deviation; TCEP, tris-(2-carboxtethyl)phosphone; BSA, bovine serum albumin.

10.1021/bi036130c CCC: $27.50 © 2004 American Chemical Society
Published on Web 04/10/2004

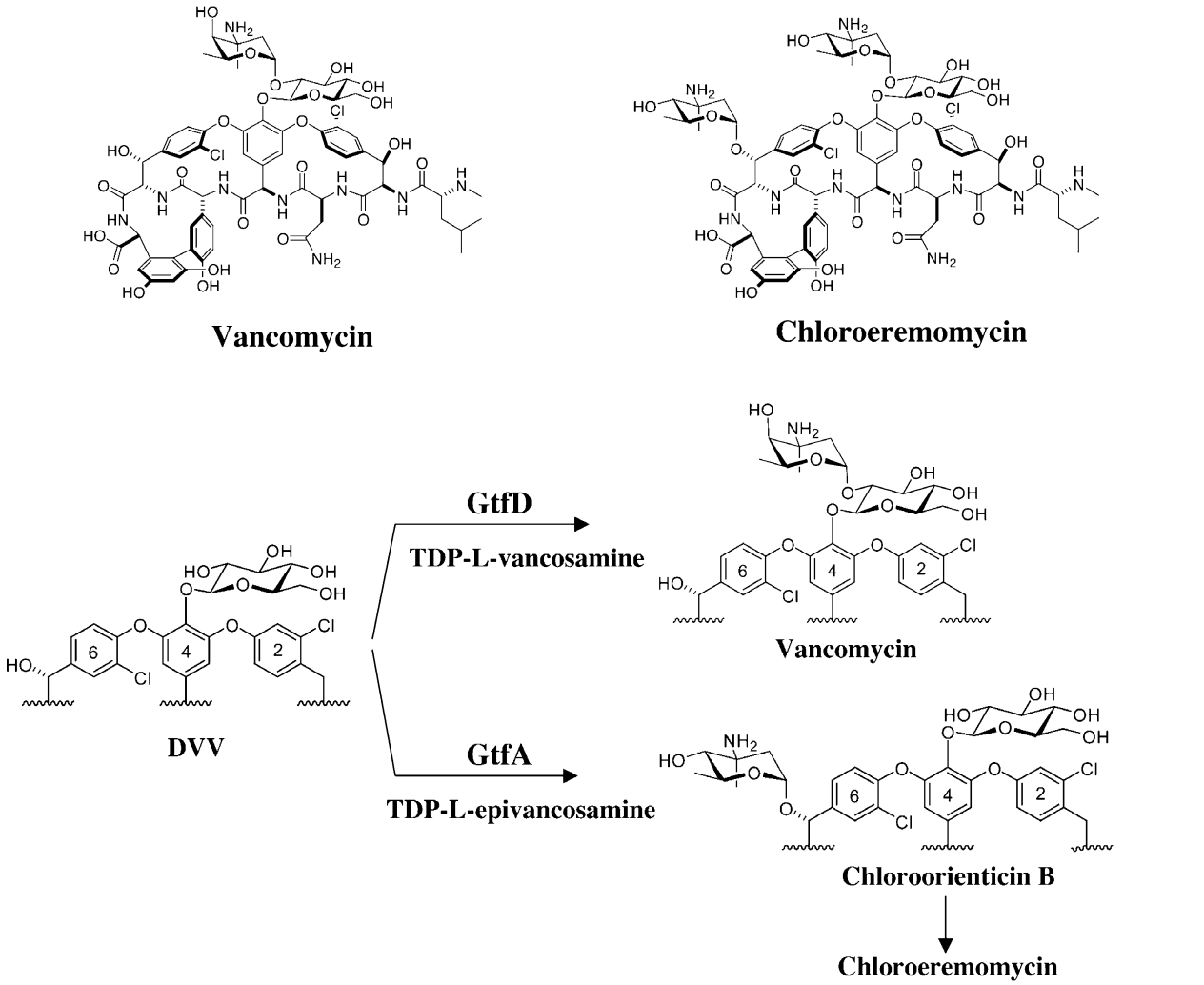

FIGURE 1: (a) Structure and glycosylation pattern of vancomycin and chloroeremomycin. Gtfs that catalyze each sugar attachment are shown in bold. (b) Comparison of sugar transfer reactions catalyzed by GtfD and GtfA on an identical DVV intermediate in the two antibiotic pathways.

synthetic antibiotics, if they can be reprogrammed to accept alternative aglycone scaffolds or NDP-sugar donors. Previous studies have demonstrated that some Gtfs from the vancomycin and chloroeremomycin pathways can effectively utilize nonnatural substrates to generate new compounds (*9*, *10*). Moreover, if the structural determinants of substrate specificity can be understood, the directed genetic modification of these enzymes could be exploited to yield a broader diversity of products.

To better understand the underlying issues of catalysis and substrate specificity, we have undertaken the structural analysis of the vancomycin group Gtfs. Previously, we reported X-ray crystal structures of two Gtfs: the unliganded form of GtfB (*11*), which transfers D-glucose to the 4-position of the heptapeptide aglycone, the initial glycosylation step for both vancomycin and chloroeremomycin, as well as acceptor and donor ligand complexes of GtfA (*12*), which transfers L-*epi*-vancosamine to the 6-position of DVV (Figure 1). These structures established the vancomycin group Gtfs as members of a growing superfamily of enzymes that share a common GT-B fold (*13*) despite having low to undetectable amino acid sequence similarity. The vancomycin group Gtfs are further classified according to their reaction mechanism as part of a subgroup or "clan", which includes two other Gtfs with known three-dimensional structures: the UDP-N-acetylglucosaminyltransferase MurG from *Escherichia coli*, which transfers N-acetyl glucosamine to the peptidogycan precursor during cell wall biosynthesis (*14*, *15*), and the more distantly related T4 phage β-glucosyltransferase, which transfers glucose to the hydroxymethylcytosine of DNA (*16*−*18*). Thus, this clan of enzymes has evolved to accommodate a diverse range of acceptor substrates. Moreover, the discovery of a structurally homologous UDP-N-acetylglucosamine epimerase revealed that the structural family extends to diverse functionalities as well (*19*). The bidomain architecture of the Gtfs, in which the N- and C-terminal domains are independently responsible for recognition and specificity of the acceptor and donor substrates, respectively, suggests the engineering of chimeric enzymes as yet an additional tool toward the generation of novel compounds.

Here, we report the X-ray crystal structure of the third and last unique Gtf from the vancomycin group series: GtfD, which transfers vancosamine from a TDP-L-vancosamine donor to the monoglycosylated acceptor DVV to form a disaccharide moiety (Figure 1b). The structure was determined as a ternary complex with both TDP and the DVV acceptor substrate and thus provides insights into the mechanisms of substrate recognition and binding. With examples of each unique Gtf of the vancomycin/chloroeremomycin pathways and in multiple binding states, we can

begin to understand the structural basis for their unique substrate specificities and regiospecificity of sugar transfer.

EXPERIMENTAL PROCEDURES

Plasmid Construction. The construction of pET22b-GtfD-His_6 has been reported previously (*9*). The GtfD mutant constructs were generated using pET22b-GtfD-His_6 as a template by a QuikChange site-directed mutagenesis kit (Stratagene). The T10A mutant was prepared using the following two primers: 5′-GTTGTCGGTGTGCGGAGCT-CGCGGGGACGTCGAA-3′ and 5′-TTCGACGTCCCCGC-GAGCTCCGCACACCGACAAC-3′. The D13A mutant was generated using the following two primers: 5′-GTGCGGAA-CCCGCGGGGCTGTCGAAATCGGGGTG-3′ and 5′-CAC-CCCGATTTCGACAGCCCCGCGGGTTCCGCAC-3′; the sites mutated are underlined. The site-directed mutagenesis experiments were carried out based on the protocol provided by Stratagene. All of the constructs were transformed into *E. coli* BL21(DE3) after the GtfD open reading frames had been confirmed by DNA sequencing.

Protein Expression and Purification. All of the GtfD proteins were produced and purified using procedures described previously (*9*). Briefly, all three proteins were purified using Ni-NTA resin (Qiagen) and were used without further purification for the measurement of their enzymatic activities. For protein crystallization, wild-type GtfD was further purified by size exclusion chromatography on a Hi-load 26/60 Superdex 200 column (Pharmacia) using equilibration buffer (20 mM Tris [pH 7.5] and 100 mM NaCl) at the flow rate of 3 mL/min. The protein was dialyzed against dialysis buffer (25 mM Tris [pH 8.0] and 2 mM DTT) and concentrated to 10 mg/mL.

GtfD Assays. In a total volume of 50 μL, DVV and UDP-L-*epi*-vancosamine prepared as previously reported (*9*) were incubated with wild-type GtfD (50−500 nM), GtfD-T10A (2.5 μM), or GtfD-D13A (5−30 μM) in the reaction buffer (75 mM Tris [pH 7.0], 2.5 mM TCEP, 8 mM $MgCl_2$, and 1 mg/mL BSA). The reactions were maintained at 37 °C for 1−60 min and were quenched by adding 100 μL of methanol. The reaction mixtures were analyzed by HPLC using a Vydac small pore C_{18} column (10−30% CH_3CN in 0.1% TFA/H_2O over 20 min, 1 mL/min). The peaks, monitored at A_{280} for product (t_R = 11.0 min) and remaining DVV (t_R = 12.0 min), were integrated, and the concentration of product was calculated from its percentage of the total peak area. The molecular weight of the product was confirmed by MALDI mass spectrometry.

Crystallization and Data Collection. The purified GtfD enzyme was crystallized using the hanging drop vapor diffusion method; the drop was formed by mixing equal parts protein solution and a reservoir solution of 1.3 M sodium potassium phosphate, pH 7.0, in the presence of ∼2 mM DVV and 5 mM TDP-glucose or TDP (Sigma-Aldrich, Inc.). The diamond-shaped plates belong to the monoclinic space group $P2_1$, (a = 50.7 Å, b = 64.1 Å, c = 144.1 Å, β = 91.7°) and have two protein molecules in the asymmetric unit. For data collection, crystals were transferred first to a 1.7 M sodium potassium phosphate storage buffer and then in a stepwise manner to buffers having increasing concentrations of glycerol as a cryo-protectant, to a final concentration of 30%. The crystals were then flash-frozen in liquid propane. From a crystal of GtfD cocrystallized with DVV and TDP, 2.1 Å resolution X-ray diffraction data used for phasing were measured at the DND-CAT beamline 5-ID at the Advanced Photon Source, Argonne National Laboratory. Subsequently, higher quality data were collected on a GtfD crystal grown in the presence of DVV and TDP-glucose for which diffraction was observed beyond 1.9 Å. All data were processed and scaled using XDS software (*20*).

Phasing and Structure Refinement. The positions of the two GtfD molecules in the asymmetric unit were determined by AMoRe molecular replacement programs (*21*) in the CCP4 crystallographic program suite (*22*), using a search model based on atomic coordinates of the previously determined GtfA structure in the closed conformation. Good electron density was observed for the protein, as well as for DVV and TDP molecules bound to both independent protein molecules. The model was completed and ligands were added using CHAIN interactive graphics software (*23*), alternated by cycles of refinement with CNS Solve v 1.1 software (*24*). In the early stages of refinement, simulated annealing protocols were used, followed by simple positional refinement and individual *B* factor refinement. After initial refinement, higher quality data at 2.0 Å measured from a second crystal grown in the presence of TDP-glucose were used; however, electron density was observed only for the NDP moiety of the TDP-glucose ligand. In later stages of refinement, water molecules were added to the model, as well as three glycerol molecules, where $|F_o - F_c|$ electron density maps showed peaks at least 3σ above background and in a position appropriate for hydrogen bonding. Two large spheres of difference density associated with each monomer were modeled as potassium cations based on relative peak height, presence of the ion in the crystallization conditions, and coordination distances with Asp and Thr side chains and water molecules. Noncrystallographic symmetry restraints between the two independent GtfD molecules were used initially during refinement, but restraints were released in later stages. The final GtfD model (R_{factor} = 0.212, R_{free} = 0.253) includes residues 1−400. The terminal eight amino acids and hexa-histidine tag are unobserved. The structure also includes DVV and TDP molecules bound to both protein monomers, as well as 460 water molecules, two potassium ions, and three glycerol molecules. Statistics of data collection and structure refinement are summarized in Table 1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank as entry 1RRV.

RESULTS AND DISCUSSION

Overall Structure. The two independent protein molecules in the GtfD asymmetric unit are virtually identical, with a rmsd on Cα atoms of 0.32 Å. As expected, the enzyme displays the characteristic GT-B fold previously observed for GtfA, GtfB, and other related enzymes (Figure 2). This fold is comprised of two distinct N- and C-terminal domains having similar topology and connected by a linker peptide and a long C-terminal helix. The core of both domains is a Rossmann type motif of parallel β-strands and interleaving α-helices, which is typically associated with mono- and dinucleotide binding. A deep cleft between the two domains forms the NDP-sugar donor binding site and is occupied by TDP in the crystal structure. On the adjacent surface of the N-terminal domain, the two loops (N3, N5) connecting the

Table 1: Statistics of Data Collection and Structure Refinement

Data Collection	
resolution	30−2.0 Å
R_{sym}[a] (last shell)	7.6% (35.3)
measured reflections	265 444
unique reflections	65 256
completeness	98% (92)
I/σ	15.5 (8.5)
Structure Refinement	
reflns, working set	52 661
reflns, test set	2668
R factor[b]	21.2%
R_{free}[c]	25.3%
atoms, total	6509
protein	5797
ligands	232
waters	460
glycerol, K+	20
$\langle B \rangle$, protein	36.9 Å^2
$\langle B \rangle$, ligands	33.6 Å^2
rmsd	
bonds	0.007 Å
angles	1.3°
dihedrals	23.0°
impropers	1.1°

[a] $R_{sym} = \sum|I - \langle I \rangle|/\sum I$, where I is the observed intensity of a measured reflection and $\langle I \rangle$ is the mean intensity of that reflection. [b] R factor $= \sum|F_o - F_c|/\sum F_o$, where F_o and F_c are observed and calculated structure factors, respectively. [c] F_{free} is the cross-validation R factor computed for a test set of reflections (5% of total).

third and fifth β-strands with the subsequent helices are greatly elaborated relative to the classical Rossmann motif; a deep crevice between them is the DVV acceptor substrate binding site.

The GtfD structure has a "cleft closed" conformation, similar to that observed previously for the GtfA-TDP complex. Revealed in previous work to be induced by nucleotide binding, the closed conformation allows the NDP to interact directly with both N- and C-terminal domains simultaneously (*12*). However, whereas in the GtfA complex the two domains remain largely independent in the closed form, the GtfD complex shows significantly more extensive interdomain contacts. In the GtfA complex, two interdomain hydrogen bonds are introduced by cleft closure: one hydrogen bond between Arg11 and Glu277 side chains, which is part of an intricate salt bridge network, and a second hydrogen bond between the Arg11 main chain amide and Ser230 side chain, both near the site of TDP binding. Both interactions are conserved in the GtfD structure, but numerous additional hydrogen bonds are also formed between residues of the N3 (60−76) and N5 (139−144) loops and the C-terminal domain. These interactions include hydrogen bonds of residues Gln62, Glu63, Ser127, Ser248, Gly249, Arg330, and Asp333, involving both main chain and side chain atoms, as well as an antiparallel β-sheet type interaction between the main chain atoms of Tyr140 and Thr332. In addition to the hydrogen bonding interactions, the Trp276 indole ring is tucked between the two domains and makes van der Waals contacts with the side chains of Gln27, Met60, and Gln62 on the N3 loop. Comparison with the previous GtfA and GtfB crystal structures reveals that the conserved Trp276 residue is part of a flexible loop that folds down toward the cleft to bury the indole ring in the closed conformation of the enzyme. In the GtfA-TDP complex, the Trp side chain is also inserted between the two domains in a similar but nonidentical manner.

Acceptor Binding Site. Well-defined electron density clearly identifies the location of a DVV molecule bound to the enzyme (Figure 3a). The largely rigid ligand binds edgewise in a deep crevice between the N3 and the N5 loops of the N-terminal domain. The glucosyl moiety is buried at the innermost edge of the crevice, abutting against the N1 and N4 turns that follow the first and fourth β-strands, residues 8−13 and 102−103, respectively. The hexose reactive O2′ hydroxyl hydrogen bonds to both Thr10 and Asp13 side chains, with the adjacent O3′ hydroxyl hydrogen bonding to the Asp 13 side chain as well (Figure 3b). The flexible O6′ hydroxyl is tethered by a hydrogen bond to the Asp103 side chain. Immediately preceding the Nα4 helix, the Asp103 side chain is in turn stabilized by additional hydrogen bonds to the main chain amide groups of residues 105 and 106 at the positive helical dipole. The convex surface of the cup-shaped aglycone core packs against one wall of the crevice formed by the N3 loop and half of the subsequent α-helix, residues 55−80. The binding interactions of the heptapeptide core are predominantly hydrophobic in nature, involving the surrounding side chains of residues Leu55, Met60, Leu61, Pro67, Pro68, Leu76, and Met79. However, the aglycone makes a small number of hydrogen bonds along this surface of the enzyme, particularly via the heptapeptide main chain atoms. The DVV residues 1 and 3 carbonyl oxygen atoms hydrogen bond to the Thr80 side chain hydroxyl and Leu61 amide groups, respectively, while the residue 2 carbonyl and residue 4 amide groups make solvent-bridged binding interactions. In addition, the hydroxyl group of the β-OH-Tyr$_2$ side chain is in hydrogen bonding distance of both the Cys8 and the Thr80 side chains, and the OH-Phegly$_5$ side chain hydrogen bonds to the Met65 main chain carbonyl groups.

The opposite wall of the binding site is formed by the N5 loop, residues 130−148, where minimal hydrophobic contacts with the bound ligand are contributed by the side chains of Asp141 and Pro143. However, this side of the site is not entirely filled by the substrate, leaving a sizable solvent cavity approximately 12 Å in diameter. This extra room within the binding site is consistent with the observed flexibility in GtfD acceptor specificity. The enzyme is also able to accept as a substrate the glucosylated teicoplanin aglycone intermediate, which differs from DVV by the presence of bulkier bridged OH-Phegly side chains at residues 1 and 3 (*9*). Contiguous with the solvent cavity, a network of 3−4 well-ordered water molecules is found within the concave bowl of the DVV molecule. One water molecule bridges the main chain amide and carbonyl groups of heptapeptide residues 1 and 4, respectively, while others hydrogen bond with the main chain amide nitrogen atom of residue 7, as well as the O6′ hydroxyl of the glucose moiety.

Both the N3 and the N5 loops that comprise the aglycone binding site are partially disordered in the previously determined unliganded structure of the GtfB homologue. This suggests that in GtfD as well these loops may be disordered when substrate is not bound. However, the variable loops of GtfD may benefit from stabilization not present in GtfB. In the N3 loop, an extended sequence of five consecutive Pro residues immediately preceding the Nα3 helix is not present in the GtfB sequence and may therefore render the peptide inherently less flexible. In addition, hydrogen bonds

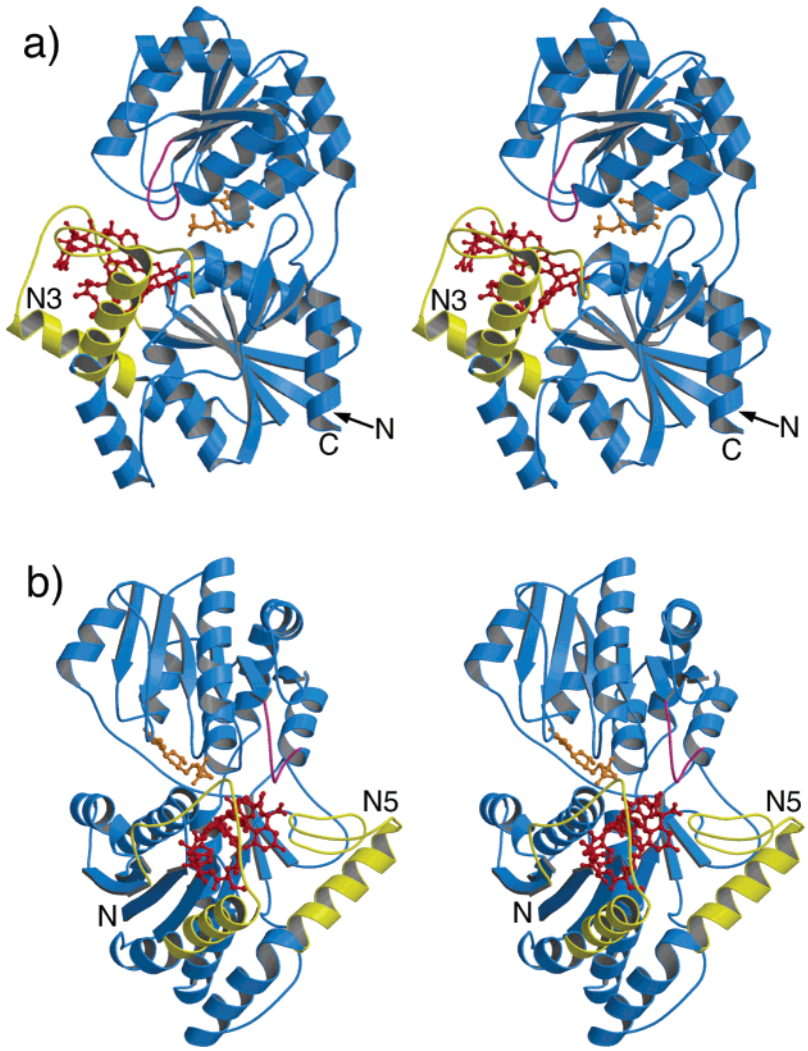

FIGURE 2: Stereoviews showing perpendicular views of the GtfD complex. Bound TDP and DVV ligands are shown as ball-and-stick models in orange and red, respectively. Hypervariable N3 and N5 loops that form the acceptor binding site are highlighted in yellow. The N3 loop can be seen more clearly in side view (a); the N5 loop can be seen more clearly in front view (b). The figure was prepared using Molscript (*29*) and Raster3D (*30*).

between Glu63 and the C-terminal domain side chains of Ser248 and Arg330 stabilize the N3 loop in the GtfD complex, whereas none of these side chains are conserved in the GtfB sequence. Furthermore, the GtfB structure, having no nucleotide bound, does not adopt the fully cleft-closed conformation. These interdomain interactions may be important for the formation of a stable surface to act as the acceptor binding site, once TDP-sugar binding and cleft closure have occurred. In the GtfD complex, the N5 loop is also stabilized by several hydrogen bonds with the adjacent Nα5 helix: Asp141-Arg158, Glu142-Arg161, and Thr145-Glu157. Many of these side chains are also not conserved in GtfB.

In addition to the N3 and N5 loops, a few additional hydrophobic binding contacts are made with the adjacent Asn331 and Thr332 side chains of the C-terminal domain, further burying the bound DVV molecule. However, the C-terminal edge of the heptapeptide is exposed at the mouth of the crevice, along with the residue 6 main chain carbonyl and side chain hydroxyl groups. The accessibility of the β-OH-Tyr$_6$ side chain is particularly significant with regard to GtfC, the GtfD analogue in the chloroeremomycin pathway. GtfC catalyzes the last of three glycosylation events (*25*), converting chloroorienticin B into chloroeremomycin (Figure 1b). An analysis of the GtfD complex shows that the crevice mouth could provide room for the additional sugar moiety appended at position 6 of chloroorienticin B without difficulty. Moreover, the surrounding enzyme side chains, including conserved Glu63, as well as the substitutions Asn331His and Tyr140Glu in GtfC, may offer hydrogen bonding opportunities for the sugar.

Interestingly, DVV binding by GtfD distorts the aglycone core significantly as compared to the crystal structures of both free vancomycin dimers and GtfA-bound DVV, the heptapeptide scaffolds superimposing with a rmsd of 1.3 Å (Figure 4). In contrast, aglycone observed in the GtfA complex is distorted only slightly to optimize binding interactions, resulting in a rmsd of 0.5 Å when compared to the crystal structure of free vancomycin. The distortion in the GtfD complex occurs primarily in the main chain torsional angles of residues 1−3 of the heptapeptide, which are skewed by up to 75° in the GtfD complex. The overall result is a more pronounced curvature of the cup-shaped DVV molecule, with the nearest atoms of opposite residues 1 and 7 compressed together from a distance of 7−7.5 Å to less than 4 Å.

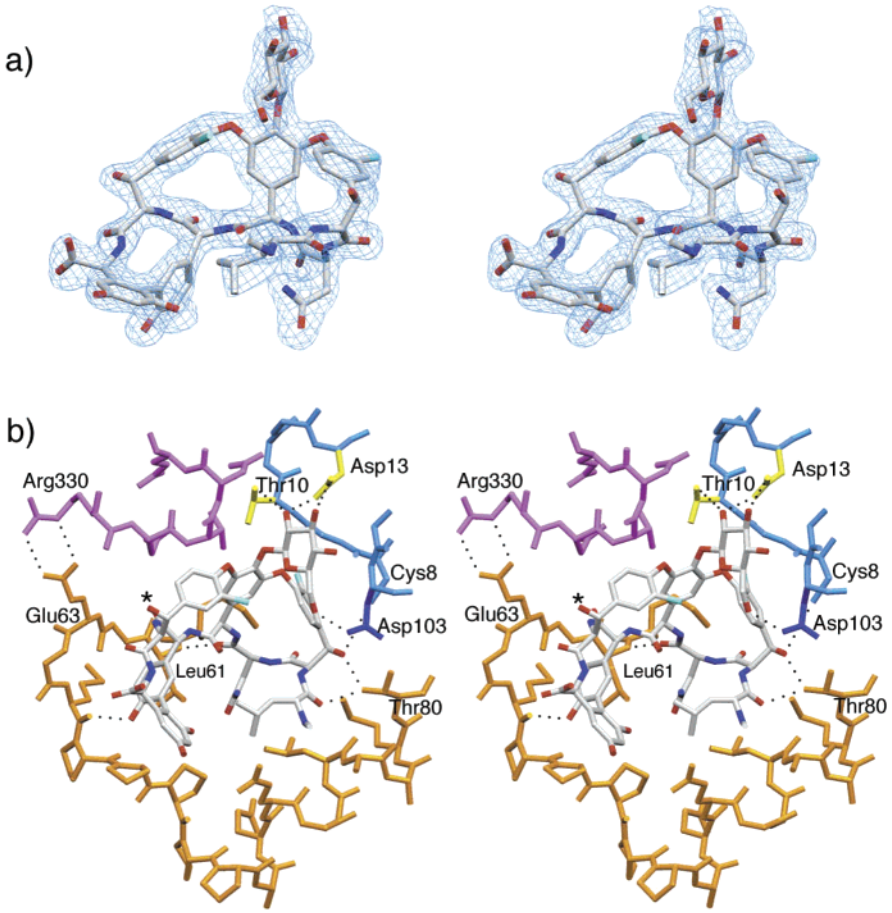

FIGURE 3: (a) Stereoview showing difference electron density (3σ level) for DVV bound in the GtfD acceptor binding site. (b) Binding interactions of DVV with GtfD. The hypervariable N3 loop (gold), sugar recognition loop (magenta), and potential catalytic residues (yellow) are highlighted. Dashed lines indicate hydrogen bonds, including interdomain ion pair between Glu63 and Arg330 side chains. The asterisk indicates sugar attachment position on the DVV intermediate for GtfA in the chloroeremomycin pathway. Figures 3−6 were prepared using SETOR (*31*).

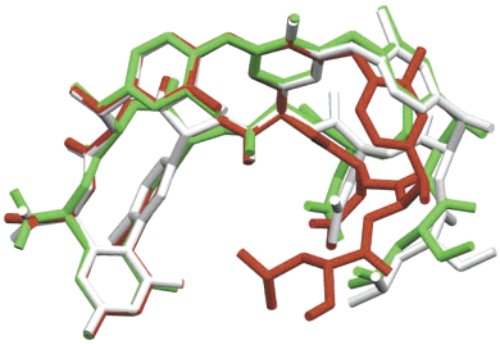

FIGURE 4: Comparison of heptapeptide core structures as observed in crystal structures of the DVV complex with GtfD (red), the DVV complex with GtfA (green), and free vancomycin (gray).

Donor Substrate Binding Site. Although GtfD was cocrystallized with TDP-glucose, electron density is observed only for the TDP portion of the ligand (Figure 5a). Similarly, when GtfD is cocrystallized with the natural product, vancomycin, clear electron density is observed for the DVV portion of the ligand but not for the second vancosamine sugar (data not shown). In both cases, the missing sugar may be hydrolyzed by the enzyme during the one week period necessary for crystals to form. This phenomenon was previously reported for T4 phage β-GT, which was only successfully cocrystallized with the UDP-glucose substrate when an inactive mutant of the enzyme was used (*26*).

The TDP moiety binds to GtfD in an almost identical fashion as that observed previously in the GtfA complex, and this binding state is assumed to closely represent binding of the natural TDP-vancosamine donor substrate (Figure 5b). The pyrophosphate moiety sits between the positive dipoles of the Nα1 and Cα4 helices where the presence of Gly residues allows direct hydrogen bonding with the main chain in a manner analogous to nucleotide binding by the more classical Rossmann motif. The α-phosphate hydrogen bonds to the main chain amide groups of Gly313 and Thr314, as well as the Thr314 side chain from the C-terminal domain; the β-phosphate interacts analogously with the Nα1 helix via the Gly12 main chain amide and Thr10 side chain. In addition, the β-phosphate moiety has hydrogen bonding interactions to the C-terminal domain, via the side chains of His309 and Ser311, as well as both amide and hydroxyl groups of Ser246. These residues are highly conserved among antibiotic Gtfs, and identical interactions are observed in the GtfA complex. One minor difference is the conformation of the Ser311 side chain, which adopts a different rotomer orientation in the GtfD structure, so that it may simultaneously hydrogen bond to the adjacent Gln334 side chain.

These highly conserved C4 and N1 loops responsible for binding the pyrophosphate moiety display some sequence conservation in the broader GT-B family as well. Notably, the conserved ^{309}HXXAGT loop at the C4 helix dipole

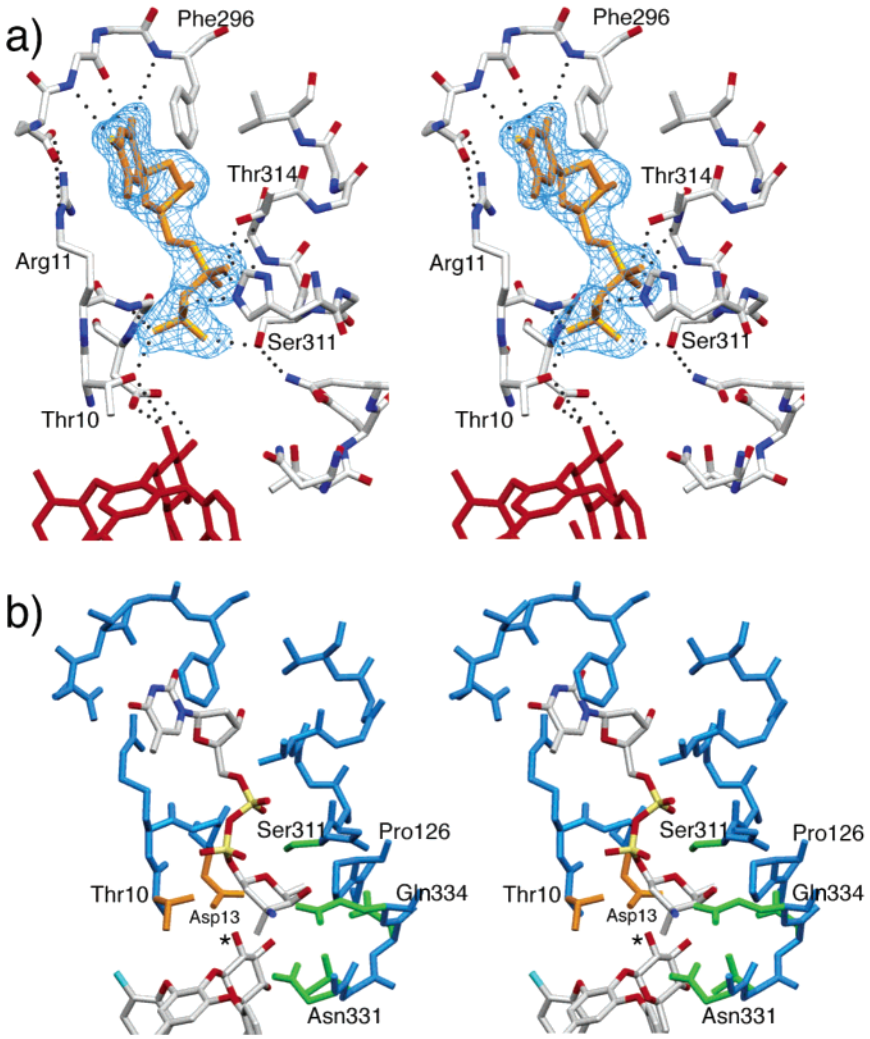

FIGURE 5: (a) Electron density of TDP bound in the GtfD cleft, with hydrogen bonding interactions indicated by dashed lines. Position of adjacent bound DVV molecule shown in red. (b) Model for TDP-vancosamine substrate docked in the GtfD binding cleft based on observed TDP position. Catalytic Thr10 and Asp13 side chains are highlighted in orange; side chains likely to make sugar binding interactions are highlighted in green. The Ser311 side chain has been rotated from observed position to another low energy rotamer, as seen in GtfA crystal structure, to avoid steric clash with ligand. The asterisk indicates reactive hydroxyl of DVV.

overlaps with a common H/R-X_7-E motif found among more distantly related NDP-sugar transferases. The conserved His has been proposed to stabilize the NDP leaving group during catalysis (*27*); hydrogen bonding between His309 and the TDP β-phosphate observed in the GtfD complex is consistent with this prediction. At the Nα1 helical dipole, a strictly conserved G(T/S)RGD sequence found in the antibiotic Gtfs is also somewhat conserved in related enzymes. However, the corresponding GTGGH loop of MurG does not make similar interactions with bound NDP ligand. Thus, the simultaneous interaction of the NDP group with both domains may be a feature that is unique to the antibiotic Gtfs within the broader GT-B structural family. In both MurG and β-GT complexes, the interdomain cleft does not close as tightly and the NDP-sugar substrate binds only to the C-terminal domain (*15*, *18*).

In the region of the TDP base, the edge of the thymine moiety makes hydrogen bonds via the carbonyl oxygen and ring nitrogen atoms with the main chain amide and carbonyl groups of Val294, respectively. Interestingly, this interaction is conserved not only in the antibiotic Gtfs but appears to be the only common feature in the binding of TDP and UDP moieties by the larger family of GT-B enzymes. Equivalent interactions are also seen in ligand complexes of GtfA, MurG, β-GT, and even the UDP-N-acetylglucosamine 2-epimerase (*12*, *15*, *18*, *19*). In the GtfD complex, an additional hydrogen bond may occur with the adjacent thymine carbonyl, which is within 3.5 Å of the Phe296 main chain amide nitrogen atom. This latter interaction is not seen in the complexes of other related enzymes, however.

On either side, the thymine ring is sandwiched by parallel π-stacking interactions with the phenyl ring of Phe296 and the interdomain salt bridge between Arg11 and Glu293 side chains. The interdomain salt bridge, which caps the ligand at the surface of the cleft, is strictly conserved among the antibiotic Gtfs. However, the Phe side chain is not conserved but is replaced by a Leu residue in the GtfA complex. This substitution may account for a tilt in the relative orientation of the base by approximately 25° in the GtfD complex, which allows the additional hydrogen bond with the second thymine carbonyl group.

The TDP deoxyribose moiety makes no direct hydrogen bonds with the surrounding protein; the single hydroxyl interacts only with an adjacent buried water molecule. Whereas in more distantly related GT-B enzymes, the Glu side chain of the conserved H-X_7-E motif interacts with both

ribose hydroxyls of UDP ligands, this residue is substituted by Val in GtfD. In fact, in all vancomycin group Gtfs, a Val or Leu residue is found at this position, offering no hydrogen bonding opportunity for a bound ribose. This structural feature is likely a determinant in the observed preference of these enzymes for the deoxyribose of TDP donor substrates.

Although no donor sugar moiety is observed in the GtfD crystal structure, the location of the binding subsite for the hexose moiety is evident by a small solvent cavity between the β-phosphate of TDP and the reactive glucosyl 2′-hydroxyl of DVV. Extending from the observed position of bound TDP, a model for the bound vancosamine sugar can be easily built into the cavity such that the C1 center is within approximately 3.7 Å of the reactive DVV hydroxyl oxygen (Figure 5b). Moreover, this site corresponds well with the position of N-acetyl glucosamine in the MurG complex with UDP-GlcNAc (*15*). The pocket is covered by a β-hairpin turn formed by residues ^{331}DTDQ^{334}, part of the loop that connects the β4 strand and α4 helix of the C-terminal domain. The side chains of these residues line the binding pocket, presenting potential hydrogen bonding opportunities for a bound sugar. Of particular note, the Asp-Gln sequence is strictly conserved among the vancomycin group Gtfs and is highly conserved among more distantly related GT-B enzymes, suggesting that these residues are likely to be important in donor sugar binding and recognition. Moreover, in the MurG complex, both residues, as a Gln-Gln sequence variant, hydrogen bond to the bound UDP-GlcNAc sugar moiety (*15*). In the GtfD structure, hydrogen bonds of Asp333 and Gln334 side chains with the Thr332 side chain and amide groups of residues 330 and 331 may additionally serve to stabilize the conformation of this sugar binding loop. While residues Asn331 and Thr332 may have a role in donor sugar binding, the side chains are also in van der Waals contact with the adjacent bound aglycone. These residues are conserved in the GtfC isoform from the chloroeremomycin pathway but are substituted by hydrophobic side chains in the other vancomycin group Gtfs. Thus, they may also be important to accommodate or complement the acceptor substrate.

Two additional side chains within the binding pocket may also impact the binding and recognition of the donor sugar. Ser311 is conserved in GtfD, GtfC, and GtfA, all of which transfer vancosamine or *epi*-vancosamine sugars, but is replaced by Gly in GtfB and GtfE, which are specific for glucose. When a model for the donor sugar is positioned in the pocket, this Ser side chain is directly adjacent to the hexose O2′ ring oxygen atom, where it may also hydrogen bond the substrate (Figure 5b). In contrast, Pro126, which is on the adjacent surface of the N-terminal domain, could provide a favorable hydrophobic contact with the C6′ methyl group of L-vancosamine and thus may play a role in donor sugar specificity. In GtfA, the conservative replacement of this proline by a leucine would also maintain this hydrophobic interaction with L-*epi*-vancosamine. In GtfB and GtfE, enzymes specific for glucosyl donors, this residue is substituted by His, which might instead offer hydrogen bonding to an O6′ hydroxyl group. Crystallographic analysis of GtfD complexed with the TDP-L-vancosamine donor substrate or close analogue will be needed to better understand the structural determinants of sugar specificity and to identify any additional conformational adjustments that may accompany NDP-sugar binding.

Table 2: Steady State Analysis of Wild-Type and Mutant GtfD Enzymes

enzyme	K_m (DVV) (μM)	K_m (UDP-L-*epi*-vancosamine) (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (UDP-L-*epi*-vancosamine) ($min^{-1}\,mM^{-1}$)
wild-type	16.3 ± 4.1	2.0 ± 0.2	67.6 ± 2.0	33.8
T10A	27.4 ± 4.5	ND	ND	0.197
D13A	ND	ND	0.012	ND

Implications for Catalysis and Acceptor Substrate Recognition. According to the catalytic mechanism proposed for other GT-B enzymes, a general base initiates catalysis by abstracting a proton from the reactive hydroxyl group of the aglycone. Sugar transfer then proceeds by direct nucleophilic attack of the hydroxyl oxygen at the donor sugar C1 carbon center (*28*). In the GtfD complex, two residues are in a suitable position to serve as this catalytic base: Thr10 and Asp13. Both side chains are within hydrogen bonding distance of the reactive glucosyl O2′ hydroxyl of DVV, but the interaction with Asp13 is significantly closer (2.5 Å) than that with Thr10 (3.6 Å). Both residues are highly conserved among the vancomycin group Gtfs as part of a common (S/T)-RGD sequence forming the N1 turn. In the GtfA complex, the corresponding Ser and Asp side chains also hydrogen bond the reactive β-OH group of DVV residue 6 in a similar manner (*12*). To clarify the role of these residues, we prepared single site mutations at Thr10 and Asp10 to generate the GtfD mutants, T10A and D13A, respectively.

Wild-type GtfD catalyzes the transfer of L-*epi*-vancosamine to DVV with a k_{cat} of 67.6 min^{-1}. Disruption of the Asp13 residue leads to a 5500-fold drop in k_{cat} (Table 2), supporting the hypothesis that Asp13 serves as a catalytic base. Likewise, a His residue at the corresponding position in MurG has also been found to be critical for catalysis by that enzyme (*15*). However, the Asp13 side chain provides two of only nine direct hydrogen bonds between GtfD and the bound DVV, and the D13A mutant may diminish the ability of the enzyme to bind the aglycone binding as well. The mutation of Thr10 to Ala reveals that this residue is also important for catalysis, albeit less critical than Asp13, since disruption of Thr10 results in only a 170-fold drop in k_{cat}/K_m (Table 2). Interestingly, the T10A activity does not saturate even when 8 mM UDP-L-*epi*-vancosamine was used, consistent with the observation that the Thr side chain hydrogen bonds not only the aglycone, but the TDP ligand as well, and may be important for substrate binding or stabilization of reaction intermediates.

A particularly intriguing aspect of the vancomycin group Gtfs is that they have evolved not only to accept different acceptor substrates but to act with different regiospecifity as well. Notably, GtfD and GtfA both recognize the identical DVV substrate but transfer L-vancosamine or L-*epi*-vancosamine moieties to distinct positions 11 Å apart on the aglycone. This requires that the two enzymes utilize different binding modes for the same ligand in order to present unique attachment points at the catalytic site. With crystal structures of both enzyme complexes now in hand, we can begin to understand how the enzymes are designed to carry out their unique roles. Comparison of the DVV complexes of GtfD

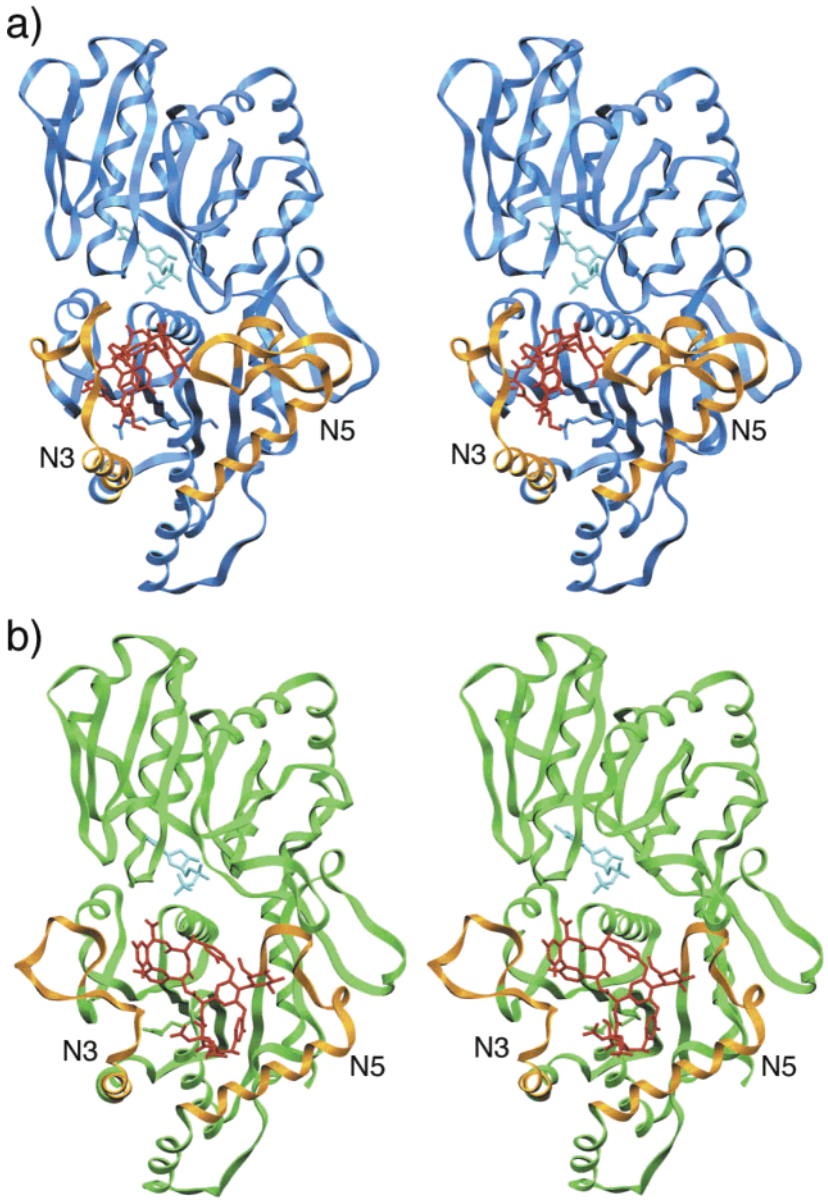

FIGURE 6: Comparison of DVV binding by (a) GtfD and (b) GtfA, with DVV and TDP ligands displayed as CPK models. Hypervariable loops N3 and N5 (gold) adopt dramatically different conformations in the two enzymes.

and GtfA reveals that ligand binding by the two enzymes is markedly different (Figure 6), as the cross-linked heptapeptides bind in nearly perpendicular orientations. In the GtfD complex, the aglycone binds edgewise in a deep crevice; in the GtfA complex, the aglycone instead lies flat within a broad pocket on the protein surface. Thus, in this case, changes in regioselectivity are achieved not by subtle adjustments in binding, such as a simple planar rotation of the heptapeptide, but by entirely different binding modes. Moreover, whereas in the GtfD complex the substrate is almost completely buried within the protein, in the GtfA complex, the aglycone remains highly exposed on the surface of the enzyme.

In both enzymes, the N3 and N5 loops define the shape and chemical nature of the binding surface. These loops are hypervariable portions of the enzyme structure, varying greatly in sequence and size among antibiotic Gtfs and differing markedly in conformation in the two enzyme complexes. In GtfD, the two elongated loops extend directly outward from the surface of the protein to form a deep crevice. In contrast, in GtfA, the N5 loop is truncated by 15 residues, which essentially eliminates the long flexible portion of the loop, and the N3 loop is splayed outward. These features thus create a broad, shallow binding pocket on the surface on the N-terminal domain of GtfA. The difference in N3 loop conformations between GtfA and GtfD may arise in part from the positioning of Pro residues, such as a sequence of five consecutive Pro residues that are present only in GtfD. Despite these differences in binding mode (Figure 6), the two regiospecific Gtf enzymes share some common features, which drive ligand binding: recognition and positioning of the substrate glucosyl moiety, hydrogen bonding of main chain atoms on the convex surface of the heptapeptide, as well as van der Waals interactions with a hydrophobic surface of complementary shape. These elements are satisfied differently as a result of both compensatory changes in the N3 and N5 loop conformations, as well as selected amino acid substitutions.

The two enzymes also differ in the presence of localized structure caused by intraloop interactions. The N3 loop of

GtfD contains one β-hairpin turn at residues Pro56-Met59 but no other internal hydrogen bonding stabilization. The loop is instead primarily stabilized by hydrogen bonds to the C-terminal domain, promoting an extended conformation. In contrast, the N3 loop of GtfA has numerous internal interactions. The peptide adopts a kinked conformation, resulting from three consecutive β-hairpin turns at residues Arg58-Ala61, Glu63-Glu66, and Pro68-Ala71, which is further stabilized by hydrogen bonding of main chain atoms by the Arg62 and Glu63 side chains.

The two enzymes also provide distinctly unique sites for recognition of the DVV glucosyl moiety. In GtfD, the subsite that binds the glucose of DVV is the catalytic center itself. In addition to the catalytic Asp13 and Thr10 side chains, the hydrophilic pocket is expanded by the presence of Asp103 on the adjacent N4 turn, a residue substituted by Leu in the other antibiotic Gtfs. In GtfA, the glucose subsite is instead formed at a perpendicular bend in the N5 loop, lined by the side chains of His128, Gln133, and Tyr141. The larger N5 loop of GtfD occupies this corner and eliminates the constellation of available hydrogen bonding partners.

Finally, although the aglycone binding site is predominantly formed by the N-terminal domain, DVV interacts directly with the donor sugar loop, demonstrating that this portion of the C-terminal domain must also complement the bound acceptor. In fact, we previously suggested that a four residue insertion found only in the sugar binding loop of GtfA may be required to accommodate the acceptor glucosyl binding position unique to this enzyme (*12*). Both the observed interactions of DVV and the predicted interactions of the donor sugar with residues from both GtfD domains reveal that the determinants of acceptor and donor specificity are not strictly isolated in the N- and C-terminal domains. Moreover, direct interdomain interactions, which were observed for the first time in the GtfD complex, may also be important for the formation or stabilization of the aglycone binding site. Thus, the two Gtf domains are not as independent and potentially interchangeable as our earlier structures suggested (*11*, *12*). These issues will have to be considered in the effort to design modified or chimeric Gtfs for the production of novel antibiotic products. Further structural studies on the ternary complexes of other antibiotic Gtfs, as well as direct observation of donor sugar binding, will be important in better understanding the intricacies of substrate recognition for this family of enzymes.

ACKNOWLEDGMENT

Portions of this work were performed at the DuPont−Northwestern−Dow Collaborative Access Team (DND-CAT) Synchrotron Research Center located at Sector 5 of the Advanced Photon Source (APS). We thank the staff of the DND-CAT beamline for their assistance. We thank Dr. Daniel Kahne and colleagues (Princeton University) for providing DVV and UDP-L-*epi*-vancosamine samples.

REFERENCES

1. Williams, D. H., and Bardsley, B. (1999) The vancomycin group of antibiotics and the fight against resistant bacteria, *Angew. Chem., Int. Ed. 38*, 1172−1193.
2. Malabarba, A., Nicas, T., and Thompson, R. C. (1997) Structural modifications of glycopeptide antibiotics, *Med. Res. Rev. 17*, 69−137.
3. Malabarba, A., and Ciabatti, R. (2001) Glycopeptide derivatives, *Curr. Med. Chem. 8*, 1759−1773.
4. Ge, M., Chen, Z., Onishi, H. R., Kohler, J., Silver, L. L., Kerns, R., Fukuzawa, S., Thompson, C., and Kahne, D. (1999) Vancomycin derivatives that inhibit peptidoglycan biosynthesis without binding D-Ala-D-Ala, *Science 284*, 507−511.
5. Nicas, T., Mullen, D., Flokowitsch, J., Preston, D., Snyder, N., Zweifel, M., Wilkie, S. C., Rodriguez, M., Thompson, R. C., and Cooper, R. D. G. (1996) Semisynthetic glycopeptide antibiotics derived from LY264826 active against vancomycin-resistant enterococci, *Antimicrob. Agents Chemother. 40*, 2194−2199.
6. Cooper, R. D. G., Snyder, N., Zweifel, M., Staszak, M., Wilkie, S. C., Nicas, T., Mullen, D., Butler, T., Rodriguez, M., Huff, B., and Thompson, R. C. (1996) Reductive alkylation of glycopeptide antibiotics: synthesis and antibacterial activity, *J. Antibiot. 49*, 575−581.
7. Rodriguez, M., Snyder, N., Zweifel, M., Wilkie, S. C., Stack, D. R., Cooper, R. D. G., Nicas, T., Mullen, D., Butler, T., and Thompson, R. C. (1998) Novel glycopeptide antibiotics: N-alkylated derivative active against vancomycin-resistant enterococci, *J. Antibiot. 51*, 560−569.
8. Eggert, U. S., Ruiz, N., Falcone, B., Branstrom, A., Goldman, R., Silhavy, T., and Kahne, D. (2001) Genetic basis for activity differences between vancomycin and glycolipid derivatives of vancomycin, *Science 294*, 361−364.
9. Losey, H. C., Peczuh, M. W., Chen, Z., Eggert, U. S., Dong, S. D., Pelczer, I., Kahne, D., and Walsh, C. T. (2001) Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides, *Biochemistry 40*, 4745−4755.
10. Losey, H. C., Jiang, J., Biggins, J. B., Oberthur, M., Ye, X.-Y., Dong, S. D., Kahne, D., Thorson, J. S., and Walsh, C. T. (2002) Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides, *Chem. Biol. 9*, 1305−1314.
11. Mulichak, A. M., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2001) Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics, *Structure 9*, 547−557.
12. Mulichak, A. M., Losey, H. C., Lu, W., Wawrzak, Z., Walsh, C. T., and Garavito, R. M. (2003) Structure of the TDP-epivancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway, *Proc. Natl. Acad. Sci. U.S.A. 100*, 9238−9243.
13. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases, *J. Mol. Biol. 328*, 307−317.
14. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000) The 1.9 Å crystal structure of E. coli MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis, *Protein Sci. 9*, 1045−1052.
15. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases, *Proc. Natl. Acad. Sci. U.S.A. 100*, 845−849.
16. Vrielink, A., Ruger, W., Driessen, H. P. C., and Freemont, P. S. (1994) Crystal structure of the DNA modifying enzyme β-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose, *EMBO J. 13*, 3413−3422.
17. Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W., and Freemont, P. S. (1999) T4 phage β-glucosyltransferase: substrate binding and proposed catalytic mechanism, *J. Mol. Biol. 292*, 717−730.
18. Morera, S., Lariviere, L., Kurzeck, J., Aschke-Sonnenborn, U., Freemont, P. S., Janin, J., and Ruger, W. (2001) High-resolution crystal structures of T4 phage β-glucosyltransferase: induced fit and effect of substrate and metal binding, *J. Mol. Biol. 311*, 569−577.
19. Campbell, R. E., Mosimann, S. C., Tanner, M. E., and Strynadka, N. C. J. (2000) The structure of UDP-*N*-acetylglucosamine 2-epimerase reveals homology to phosphoglycosyltransferases, *Biochemistry 39*, 14993−15001.
20. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants, *J. Appl. Crystallogr. 26*, 795−800.

21. Navaza, J. (1994) AMoRe: an automated package for molecular replacement, *Acta Crystallogr. A50*, 157−163.
22. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
23. Sack, J. S. (1988) CHAIN—a crystallographic modeling program, *J. Mol. Graphics 6*, 224−225.
24. Brunger, A. T. (1998) Crystallography and NMR system (CNS): a new software system for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
25. Lu, W., Oberthur, M., Leimkuhler, C., Tao, J., Kahne, D., and Walsh, C. T. (2004) Characterization of a regiospecific epivancosaminyl transferase GtfA and enzymatic reconstitution of the antibiotic chloroeremomycin, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
26. Lariviere, L., Gueguen-Chaignon, V., and Morera, S. (2003) Crystal structures of the T4 phage β-glucosyltransferase and the D100A mutant in complex with UDP-glucose: Glucose binding and identification of the catalytic base for a direct displacement mechanism, *J. Mol. Biol. 330*, 1077−1086.
27. Ha, S., Gross, B., and Walker, S. (2001) E. coli MurG: a paradigm for a superfamily of glycosyltransferases, *Curr. Drug Targets Infect. Disord. 1*, 201−213.
28. Unligil, U. M., and Rini, J. M. (2000) Glycosyltransferase structure and mechanims, *Curr. Opin. Struct. Biol. 10*, 510−517.
29. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
30. Merritt, E. A., and Murphy, M. E. (1994) RASTER3D version 2.0. A program for photorealistic molecular graphics, *Acta Crystallogr. D50*, 869−873.
31. Evans, S. V. (1993) SETOR: hardware lighted three-dimensional solid model representations of macromolecules, *J. Mol. Graphics 11*, 134−138.

BI036130C