BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 4159–4168 4159
DOI: 10.1021/bi100150v

Crystal Structures of the Glycopeptide Sulfotransferase Teg12 in a Complex with the Teicoplanin Aglycone[†]

Matthew J. Bick, Jacob J. Banik, Seth A. Darst, and Sean F. Brady*

Howard Hughes Medical Institute, Laboratory of Genetically Encoded Small Molecules, Laboratory of Molecular Biophysics, The Rockefeller University, 1230 York Avenue, New York, New York 10065

Received February 3, 2010; Revised Manuscript Received March 30, 2010

ABSTRACT: The TEG gene cluster, a glycopeptide biosynthetic gene cluster that is predicted to encode the biosynthesis of a polysulfated glycopeptide congener, was recently cloned from DNA extracted directly from desert soil. This predicted glycopeptide gene cluster contains three closely related sulfotransferases (Teg12, -13, and -14) that sulfate teicoplanin-like glycopeptides at three unique sites. Here we report a series of structures: an apo structure of Teg12, Teg12 bound to the desulfated cosubstrate 3′-phosphoadenosine 5′-phosphate, and Teg12 bound to the teicoplanin aglycone. Teg12 appears to undergo a series of significant conformational rearrangements during glycopeptide recruitment, binding, and catalysis. Loop regions that exhibit the most conformational flexibility show the least sequence conservation between TEG sulfotransferases. Site-directed mutagenesis guided by our structural studies confirmed the importance of key catalytic residues as well as the importance of residues found throughout the conformationally flexible loop regions.

During the 20th Century, widespread use of antibiotics significantly reduced the threat of many once lethal infectious diseases. However, the success of these wonder drugs may soon become their Achilles' heel. Bacterial pathogens that have developed resistance to most widely used antibiotics are now regularly seen in clinical settings. Vancomycin and teicoplanin are glycopeptide antibiotics used in the treatment of many Gram-positive bacterial infections, including methicillin-resistant *Staphylococcus aureus* (MRSA). With the appearance of vancomycin-resistant enterococci in the late 1980s and resistant staphylococci in the early 1990s, these traditional antibiotics of last resort are in danger of becoming clinically compromised (*1*−*3*). As with many bacterial natural products, the discovery of additional glycopeptide congeners that might combat the growing problem of antibiotic resistance has slowed as it has become increasingly difficult to identify new biodiversity from which novel molecules might be characterized.

The vast majority of bacteria present in the environment remain recalcitrant to culturing (*4*). This uncultured majority no doubt contains previously inaccessible glycopeptide biosynthetic gene clusters, many of which could encode the biosynthesis of novel glycopeptide congeners. Although metabolites produced by bacteria that are difficult to culture in the laboratory cannot be characterized using standard microbiological methods, it is possible to extract DNA directly from environmental samples and then analyze this DNA for sequences that might encode the biosynthesis of new natural products. In a recent analysis of DNA extracted directly from desert soil, we uncovered a new glycopeptide biosynthetic gene cluster (the TEG[1] gene cluster) that is predicted to encode the biosynthesis of the first polysulfated glycopeptide congeners (*5*). The TEG gene cluster contains three closely related 3′-phosphoadenosine 5′-phosphosulfate (PAPS)-dependent sulfotransferases (Teg12, -13, and -14). In vitro, these three sulfotransferase finishing enzymes sulfate teicoplanin-like glycopeptides at three unique sites and, in combination, can be used to produce seven different glycopeptide sulfation patterns (Figure 1).

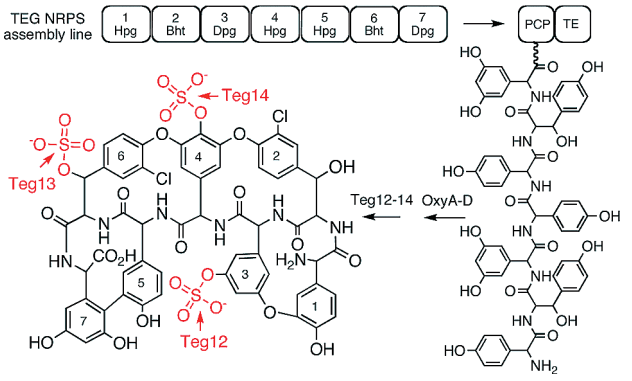

FIGURE 1: Predicted structures for the glycopeptide congeners produced by the TEG gene cluster. The seven nonribosomal peptide synthetase modules encoded by the TEG cluster are predicted to biosynthesize a heptapeptide that is oxidatively cross-linked by OxyA−D homologues and then sulfated by Teg12, -13, and -14.

Vancomycin- and teicoplanin-like glycopeptides are structurally defined by the presence of an oxidatively cross-linked heptapeptide core (*6*, *7*). The peptide core is initially produced as a linear polymer by nonribosmal peptide synthetases, which is then oxidatively cross-linked into either three or four large macrocycles by conserved cytochrome P450 oxidases. During their

[†]This work was supported by National Institutes of Health Grant GM077516, the Howard Hughes Medical Institute, the Beckman Foundation, and the Searle Foundation.

*To whom correspondence should be addressed. E-mail: sbrady@rockefeller.edu. Phone: (212) 327-7478. Fax: (212) 327-8281.

[1]Abbreviations: TEG, teicoplanin-like eDNA-derived gene cluster; PAP, 3′-phosphoadenosine 5′-phosphate; PAPS, 3′-phosphoadenosine 5′-phosphosulfate; Hpg, hydroxyphenylglycine; Bht, betahydroxytyrosine; Dpg, dihydroxyphenylglycine; HEPES, 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid; CHES, 2-(cyclohexylamino)ethanesulfonic acid; DTT, dithiothreitol; IPTG, isopropyl *β*-D-1-thiogalactopyranoside; rmsd, root-mean-square deviation.

Published on Web 04/02/2010

biosynthesis, each member of this family of antibiotics is functionalized by a unique collection of finishing enzymes that includes glycosyl transferases, halogenases, acyl transferases, and sulfotransferases. Although glycopeptides show some variation in the sequence of the heptapeptide core, the bulk of the structural diversity seen within this class of antibiotics arises from the functionality added by finishing enzymes. More than 150 different glycosylated, halogenated, and alkylated glycopeptide congeners have been characterized from cultured bacteria (*8*). Only three naturally occurring sulfated congeners have been identified to date from studies of this same pool of bacteria (*8*, *9*). While anionic glycopeptides have rarely been reported as natural products, telavancin, a semisynthetic phosphono congener that proved to be a very effective antibiotic in clinical trials, was recently approved for use in humans by the Food and Drug Administration (*10*, *11*). Increasing the hydrophilicity via the addition of the negatively charged phosphono group was found to significantly improve the adsorption, distribution, metabolism, and excretion profile of this class of antibiotics (*12*). The enzymatic synthesis of anionic glycopeptides may provide a facile means of accessing additional anionic congeners with improved pharmacological properties.

Here we report a series of Teg12 sulfotransferase structures, including an apo structure, a binary structure in a complex with the teicoplanin aglycone substrate, and a ternary structure containing both PAP and the teicoplanin aglycone. In the binary and ternary structures, the glycopeptide substrate is observed bound at two different locations. Teg12 appears to undergo a series of conformational rearrangements during glycopeptide recruitment, binding, and catalysis. These studies provide insight into the sulfotransferase mechanism, as well as insights into how this rarely seen class of finishing enzymes might be engineered to produce novel anionic glycopeptides.

MATERIALS AND METHODS

Teg12 Expression and Purification. Teg12 was cloned and expressed as previously described (*5*). Briefly, *teg12* was amplified (30 cycles at 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 90 s; FailSafe system from Epicentre) from eDNA cosmid clone D30 using the following primers: Teg12FWD(BclI), GCGC**TGATCA**ATGAACGGAATTCGATGG; Teg12REV(HindIII), GCGC**AAGCTT**TCCTTAACCGGCATACCCGTA. Restriction enzyme sites added for cloning purposes are shown in bold. The resulting product was doubly digested with BclI and HindIII and subsequently ligated into pET28a, which had been doubly digested with BamHI and HindIII. The resulting construct was then transformed into *Escherichia coli* BL21(DE3) for protein expression. Expression cultures were grown to an OD_{600} of 0.6, followed by IPTG induction, and overnight growth at 20 °C. The culture was pelleted by centrifugation (3200*g* for 30 min), the supernatant discarded, and the cell pellet resuspended in 40 mL of lysis buffer [50 mM HEPES (pH 7.5), 0.5 M NaCl, 5% (v/v) glycerol, 20 mM imidazole (pH 8), 10 mM β-mercaptoethanol, and 0.5% (v/v) Triton X-100]. The resuspended cell pellet was lysed by sonication, and the insoluble portion was removed by centrifugation (15000*g* for 30 min). The cleared cell lysate was incubated with 1 mL of Ni-NTA resin for 15 min. The slurry was loaded onto a column, allowed to empty by gravity flow, washed with 40 mL of lysis buffer, and finally washed with 40 mL of wash buffer [50 mM HEPES (pH 7.5), 0.5 M NaCl, 5% (v/v) glycerol, 20 mM imidazole (pH 8.0), and 10 mM β-mercaptoethanol]. The protein was eluted by the addition of 15 mL of elution buffer [50 mM HEPES (pH 7.5), 0.5 M NaCl, 5% (v/v) glycerol, 125 mM imidazole (pH 8), and 10 mM β-mercaptoethanol]. No attempt was made to remove the vector-derived six-histidine tag, resulting in a Teg12 protein with 34 additional residues N-terminal to the start methionine. Protein was concentrated using Vivascience Vivaspin 30000 MWCO ultrafiltration concentrators and was buffer exchanged three times into protein buffer [200 mM NaCl, 20 mM HEPES (pH 7.5), 5% glycerol, and 1 mM DTT].

Teg12 Crystallization. Concentrated protein was centrifuged at 14000 rpm (4 °C for 30 min) in a microcentrifuge to remove any insoluble material prior to crystallization. All crystals were grown using the hanging drop vapor diffusion method. We obtained initial Teg12-apo crystals by mixing 1 μL of protein (7.5 mg/mL in protein buffer) with 1 μL of reservoir solution [1.0 M sodium citrate and 0.1 M sodium cacodylate (pH 6.5), JCSG core III-48, Qiagen] over a 500 μL reservoir. Bladelike crystals grew overnight at 22 °C and reached a maximal size of 400 μm × 50 μm × 20 μm in approximately 1 week. To improve crystal thickness, Teg12-apo crystals were optimized by microseeding, in addition to mixing 1 μL of protein at 7.5 mg/mL with 0.5 μL of reservoir and 0.5 μL of Silver Bullets reagent 29 (Hampton Research). A component of the Silver Bullets screen, aspartame, was modeled into the PAPS binding site of one of the monomers of the Teg12 dimer. Crystals were soaked in a 20 μL drop containing reservoir solution and 10% ethylene glycol. The drop was allowed to dehydrate by being exposed to open air at room temperature for approximately 5 h before the crystals were flash-cooled in liquid ethane.

Teg12-ternary crystals were cocrystallized at 4 °C in the presence of 2 mM PAP and 1 mM teicoplanin aglycone. Teg12 was first concentrated to 20 mg/mL in protein buffer. The protein was then diluted 1:1 with 50 mM CHES (pH 9.1), 2 mM teicoplanin aglycone, and 4 mM PAP, producing a Teg12 final concentration of 10 mg/mL in 0.5× protein buffer, 25 mM CHES (pH 9.1), 1 mM teicoplanin aglycone, and 2 mM PAP. One microliter of protein was mixed 1:1 with reservoir solution [0.2 M ammonium acetate and 20% (w/v) PEG 3350, JCSG core I-25, Qiagen]. Crystals appeared in 2−3 days and grew to a maximal size of 100 μm × 50 μm × 50 μm in approximately 1 week. These crystals were of an irregular chunk-like morphology and had cracks throughout. Crystals were cryoprotected by being quickly dunked in reservoir solution plus 15% ethylene glycol and were flash-cooled in liquid nitrogen.

Teg12-binary crystals were cocrystallized at 4 °C in the presence of 1 mM teicoplanin aglycone. Similar to Teg12-ternary complex crystallization, the protein was first concentrated to 20 mg/mL in protein buffer and then diluted to 10 mg/mL with 50 mM CHES (pH 9.1) and 2 mM aglycone (final concentrations of 0.5× protein buffer, 25 mM CHES, and 1 mM teicoplanin aglycone). One microliter of protein solution was mixed with 1 μL of reservoir solution [2.0 M sodium formate and 0.1 M sodium acetate (pH 4.6), JCSG core III-85, Qiagen]. Cubic crystals grew in 2−3 weeks and were approximately 50 μm × 50 μm × 50 μm in size. Crystals were soaked in 6.0 M sodium formate, 0.1 M sodium acetate (pH 4.6), and 1 mM teicoplanin aglycone overnight, prior to being flash-cooled in liquid nitrogen.

Data Collection and Structure Determination. All data sets were reduced and scaled using the HKL2000 package (*18*). Data for Teg12-apo crystals were collected at the NSLS, beamline X29A (Table 1). All but one of the crystals screened diffracted poorly to approximately 4 Å resolution. The crystal

Table 1: Diffraction Statistics for Teg12 Structures

	Teg12-apo	Teg12-ternary	Teg12-binary
Data Collection			
space group	$C222_1$	$P2_12_12_1$	$I2_12_12_1$
cell dimensions			
a, b, c (Å)	79.39, 126.09, 145.13	77.12, 78.56, 100.54	66.12, 80.15, 132.94
α, β, γ (deg)	90, 90, 90	90, 90, 90	90, 90, 90
no. of reflections (observed/unique)	55339/15372	177647/38619	66588/16647
resolution (Å)	50.00−2.93 (3.03−2.93)	50.00−2.05 (2.12−2.05)	50.00−2.28 (2.36−2.28)
R_{merge}/R_{sym}	0.063 (0.396)	0.084 (0.464)	0.079 (0.542)
$I/\sigma I$	19.10 (2.04)	16.43 (2.62)	15.50 (2.17)
completeness (%)	95.5 (98.0)	97.7 (96.6)	99.9 (100.0)
redundancy	3.6 (3.6)	4.6 (4.6)	4.0 (4.1)
Refinement			
resolution (Å)	72.57−2.91	32.95−2.04	34.32−2.27
no. of reflections	14562	36951	15761
R_{work}/R_{free}	21.96/27.19	17.12/22.47	17.30/22.61
total no. of atoms	3628	4820	2222
protein	3550	4247	1945
ligand/ion	44	233	170
water	34	340	95
total B factor	36.2	35.7	43.0
protein	35.9	35.6	40.3
ligand/ion	61.7	30.1	69.5
water	39.1	40.5	47.4
solvent content	52.4	40.9	46.5
rmsd for bonds (Å)	0.014	0.007	0.006
rmsd for angles (deg)	1.495	1.183	1.038
rotamer outlier (%)	3.6	0.7	1.0
Ramachandran (favored/outlier) (%)	94.6/0.6	98.7/0.0	100.0/0.0
PDB code	3MGC	3MGB	3MG9

from which the 2.91 Å data set was collected rotated briefly out of the cryostream and thereby had gone through a room-temperature annealing cycle of several seconds. Diffraction from this crystal was dramatically improved compared with that from other crystals taken from the same drop. Data for Teg12-apo were reduced and scaled in space group $C222_1$. Phase information was obtained by molecular replacement using Phaser and StaL [GenBank accession number AAM80529, Protein Data Bank (PDB) entry 2OV8], devoid of all flexible loops, as the search model (*19*). The initial molecular replacement model was refined against the Teg12-apo data set using rigid body refinement in Refmac (*20*). Additional features of the map were enhanced through density modification and 2-fold NCS averaging in CNS (*21*, *22*). The model was rebuilt manually using Coot (*23*). Full restrained refinement was conducted using the translation/libration/screw model in Refmac, with the addition of hydrogen atoms, converging to final R_{work} and R_{free} values of 21.96 and 27.19, respectively (*24*). NCS restraints were not used during refinement. The final model comprises residues 1−129, 136−203, and 251−285 for monomer A and residues 1−27, 42−128, 137−210, and 247−285 for monomer B. The Teg12-apo model was used as a molecular replacement model for all subsequent structures.

Teg12-ternary and Teg12-binary data sets were collected at the APS, microfocus beamline 24-IDE (Table 1). Teg12-ternary data were reduced and scaled in space group $P2_12_12_1$. Teg12-binary was scaled in space group $I2_12_12_1$. The crystal structure of glycopeptide aglycone A-40926 was used as a starting point to generate a restraint definition file for the teicoplanin aglycone using Phenix Elbow (*25*). Geometry optimization was achieved using the semiempirical quantum mechanical AM1 method. Teg12-binary and Teg12-ternary models were refined using the translation/libration/screw model in Phenix Refine to final R_{work} and R_{free} values of 17.30 and 22.61, and 17.12 and 22.47, respectively (*26*). The final Teg12-binary model comprises residues 1−129, 135−216, and 247−285. The final Teg12-ternary model comprises residues 1−215, 220−224, and 231−285 for monomer A and residues 1−129, 136−224, and 240−285 for monomer B. All structures were validated using the Molprobity server from the Richardson laboratory at Duke University (Durham, NC) (*27*).

Table 2: Oligonucleotide Primers for Site-Directed Mutagenesis

mutant	primer (5′−3′)
S9A	GTTTCCAGCCTTTGGATA**CGC**TGCGATCCATCGAATTCC
K12A	CCACGTGTTTCCAGC**CGC**TGGATACGATGC
T16A	CACCTGACCCA**CGC**GTTTCCAGCCTTTGG
W17A	CAACATGCACCTGAC**CGC**CGTGTTTCCAGC
K65A	CATCGGCCTTGAGGTGCGT**CGC**CACCAGCACCGGTTC
H67A	CACATCGGCCTTGAG**GGC**CGTCTTCACCAG
R90A	CATATCCCGCGGGTT**CGC**CACGAGATAGAG
S98A	GGCCATGCGATCGA**GGC**GAGCAGCATATC
R101A	TATCGAGGCCAT**GGC**CATCGAGCTGAG
S106A	CTACGTCGTCGCG**CGC**TATCGAGGCCATGCG
R107A	GCTTTTTTCTACGTCGTC**GGC**CGATATCGAGGC
D108A	GCTTTTTTCTACGTC**GGC**GCGCGATATCGAGGCC
Y167A	GTGCTGACGATGCGT**GCT**GAGGACCTGAAGGGC
K171A	CGTTATGAGGACCTG**GCG**GGCGATCCGGTCGCACGG
E206A	GCTGCCTCCACGCTG**GCG**CGGATGCGTGAACTG
R207A	GCTGCCTCCACGCTGGAG**GCG**ATGCGTGAACTG
E210A	GAGCGGATGCGT**GCA**CTGGAGAAACGGAG
E212A	CGGATGCGTGAACTG**GCG**AAACGG
K213A	CGGATGCGTGAACTGGAG**GCA**CGGAGCGAGCAGCAG
R214A	CGGATGCGTGAACTGGAGAAA**GCG**AGCGAGCAGCAG
M232A	GGTGATGCGAGAATG**GCG**AAAGGG
K233A	GATGCGAGAATGATG**GCA**GGGGG
G234A	GCGAGAATGATGAAA**GCG**GGACC
G235A	AGAATGATGAAAGGG**GCA**CCTGGTG
P236A	ATGATGAAAGGGGGA**GCT**GGTGG
G238A	AAAGGGGGACCTGGT**GCC**GCGAGG
A239R	AAAGGGGGACCTGGTGGC**CGG**AGG
R240A	GGACCTGGTGGCGCG**GCG**CCCCAG
P241A	CCTGGTGGCGCGAGG**GCC**CAGTTC
Q242A	GGTGGCGCGAGGCCC**GCG**TTCGTG
R248A	CAGTTCGTGGGCGAGGGC**GCG**TACGACCAGTCCCTG
Q251A	GAGGGCAGGTACGAC**GCG**TCCCTGTCCTTCTTG

Site-Directed Mutagenesis. Teg12 point mutants were generated using the "megaprimer" method, with slight modifications (*28*). Oligonucleotide primers were designed for each mutant (Table 2), and a megaprimer was generated by PCR amplification from the Teg12/pET28a construct using the Pfx Accuprime System (Invitrogen), the relevant mutant oligonucleotide primer, and either the T7 promoter (for mutations of residues 9−108) or the T7 terminator (for mutations of residues 167−251) as the second oligonucleotide primer (30 rounds of amplification at 95 °C for 30 s, 55 °C for 30 s, and 68 °C for 30 s). The full-length mutant Teg12 gene was amplified from the Teg12/pET28a construct, using the megaprimer, which then contained the bases that encode the specific mutant residue, and either the T7 terminator (for mutations of residues 9−108) or the T7 promoter (for mutations of residues 167−251) as the second oligonucleotide primer (30 rounds of amplification at 95 °C for 30 s, 55 °C for 30 s, and 68 °C for 80 s). Full-length mutant amplicons were then sequentially digested with BamHI and HindIII and subsequently ligated into pET28a doubly digested with BamHI/HindIII. Ligated constructs were transformed

into *E. coli* EC100 (Epicentre) and sequenced to identify successfully mutated constructs. Mutant constructs containing the desired point mutation were then transformed into *E. coli* BL21(DE3) for protein expression.

Mutant Teg12 Expression and Purification. Mutant proteins were expressed and purified in a manner similar to that used for native Teg12, except on a reduced scale; 100 mL overnight expression cultures were pelleted and resuspended in 4 mL of lysis buffer. After sonication to lyse the cells, the crude lysates were centrifuged to remove insoluble material (10 min at 15000*g*). The cleared lysates were incubated with 100 μL of Ni-NTA resin for 15 min. The slurry was then loaded onto a column, allowed to empty by gravity flow, and washed with 4 mL of lysis buffer, followed by a second wash with 4 mL of wash buffer. The protein was eluted by the addition of 1.5 mL of elution buffer. All Teg12 mutants used in activity assays appeared to be homogeneous as determined by polyacrylamide gel electrophoresis

Teg12 Activity Assays. All soluble Teg12 mutants were assayed for activity using the teicoplanin aglycone as a substrate. Reactions (50 μL) were run in duplicate [15 mM HEPES (pH 7.5), 1 mM 3′-phosphoadenosine 5′-phosphosulfate (PAPS), 0.1 mM DTT, 1.2 mM teicoplanin aglycone (in DMSO), and 500 ng of purified protein in elution buffer]. Reactions were conducted at 30 °C for either 10, 15, 20 or 25 min, followed by heat inactivation at 99 °C for 10 min, and a further 10 min in an ice/water bath. V_{max} and K_m values were determined under the same reaction conditions using the teicoplanin aglycone as a substrate (5−100 μM). Twenty-five microliters of each reaction mixture was run on a Waters analytical HPLC system [C_{18} (4.6 mm × 150 mm)]. A linear gradient (1.5 mL/min) was run from 95:5 20 mM ammonium acetate/acetonitrile to 70:30 20 mM ammonium acetate/acetonitrile over 20 min. The area under the UV peak (diode array, 240−400 nm) was determined for both the monosulfated product and the teicoplanin aglycone substrate at each time point. The percent substrate conversion for duplicate time points was averaged. The slope of the graph derived from the four time points for Teg12 and each mutant was then determined. The relative activity of each mutant is reported as a percent of the slope for wild-type Teg12.

RESULTS

Teg12-apo Structure. N-Terminally His-tagged Teg12 was affinity purified using nickel NTA resin and crystallized without the need for further purification. The apo structure was determineded to 2.91 Å resolution by molecular replacement using StaL (PDB entry 2OV8) as a search model (*13*). StaL is a related sulfotransferase involved in the biosynthesis of the monosulfated glycopeptide congener A47934 (*9*, *14*). StaL is 52.9% identical to Teg12 in sequence. The Teg12 structure superimposes well on StaL (rmsd of 1.053 Å over 210 Cα atoms). Teg12 crystallizes as a dimer, and the overall structure resembles that of StaL. Running Teg12 through the DALI server shows that its closest eukaryotic sulfotransferase relative is the human cytosolic sulfotransferase 1C1 (SULT1C1, PDB entry 1zhe), with an rmsd of 2.8 Å over all Cα atoms (*15*). As with other sulfotransferases in this family, Teg12 consists of a single globular α/β domain composed of a parallel β sheet core surrounded by α helices (Figure 2A). The β sheet core of both StaL and Teg12 contains four strands. The dimer interface, which resembles that seen in StaL, consists of a symmetrical interaction between a short helix−loop motif from one monomer and the same helix−loop motif from the other monomer. In addition to the hydrophobic contacts that exist between the two helices, there is a hydrogen bond between the carbonyl oxygen of Gly51 at the end of the helix from one monomer and the backbone nitrogen of Val74 of the other monomer. The active site cavity of each monomer faces the dimer interface.

Minor differences exist between the two monomers of the apo structure. The pET28a vector-derived N-terminal tail, including the six-His, thrombin cleavage site, and T7 tags, from monomer A is ordered and involved in crystal packing interactions. The arrangement of these residues within the crystal causes distortions in the region from Gly28 to Ser41 when compared with StaL, and subsequent Teg12 complex structures. In monomer B, residues Gly28−Ser41 could not be modeled in the electron density map, further underscoring the conformational flexibility of this region. From our cocrystal structures and that of StaL, it is known that residues Glu216−Asp250 constitute a largely disordered loop. In the Teg12-apo structure, this disorder extends even further N-terminally into what is observed as a helix in the Teg12-binary and -ternary structures and StaL. Electron density falls off in monomer A after Ser203 but extends through Glu210 in monomer B. Ser203 corresponds to Cys196 in StaL. In StaL, Cys196 makes a disulfide bond with Cys20 (residue 20 is also a cysteine in Teg12). The absence of a disulfide in Teg12 could impart a greater flexibility to this region and explain the additional observed disorder. In subsequent Teg12 cocrystal structures, portions of this helix−loop region (Thr204−Asp250) play key roles in binding of the glycopeptide substrate. Because of its role in binding the glycopeptide, we have termed this entire conformationally flexible region the GHL (glycopeptide−helix−loop).

ClustalW alignment of the three TEG sulfotransferases shows that most of the sequence variability seen within this family of sulfotransferases is concentrated on three short loops (variable regions V1−V3) that surround the predicted active site (Figure 3). The large disordered loop from the GHL corresponds to the longest of these three variable regions, V3. V2 encompasses residues Gly127−Gly137. Much of V2 (Ala130−Gly135 in monomer A and Asn129−Gly136 in monomer B) is disordered in the Teg12-apo structure. V1, the shortest of the three variable sequences, corresponds to an ordered region (though disordered in monomer B) that spans Ile37−Thr43.

Electron density was observed in the PAPS binding site of monomer A (this binding site was empty in monomer B). However, neither PAP nor PAPS could be modeled into the density. We were able to model the dipeptide aspartame [*N*-(L-α-aspartyl)-L-phenylalanine, 1-methyl ester] into this density (Figure 4). Aspartame is a component of the Silver Bullets screen (Hampton) used during crystallization. The orientation of aspartame in the structure mimics that of PAP. Its phenylalanine ring stacks with Trp17, as does the adenine ring of PAP, and its N-terminal nitrogen forms a hydrogen bond with the hydroxyl of Ser98. Eukaryotic sulfotransferases are known to tightly bind ribose, adenine, and other nucleotides in the PAPS binding pocket (*16*).

Teg12 Cocrystal Structures. The TEG gene cluster is predicted to encode the biosynthesis of a heptapeptide [hydroxyphenylglycine (Hpg)-betahydroxytyrosine (Bht)-dihydroxyphenylglycine (Dpg)-Hpg-Hpg-Bht-Dpg] that is oxidatively cross-linked into the four macrocycles seen in teicoplanin-like glycopeptides (Figure 1). This TEG-derived, oxidatively cross-linked heptapeptide skeleton differs from the teicoplanin aglycone only in the

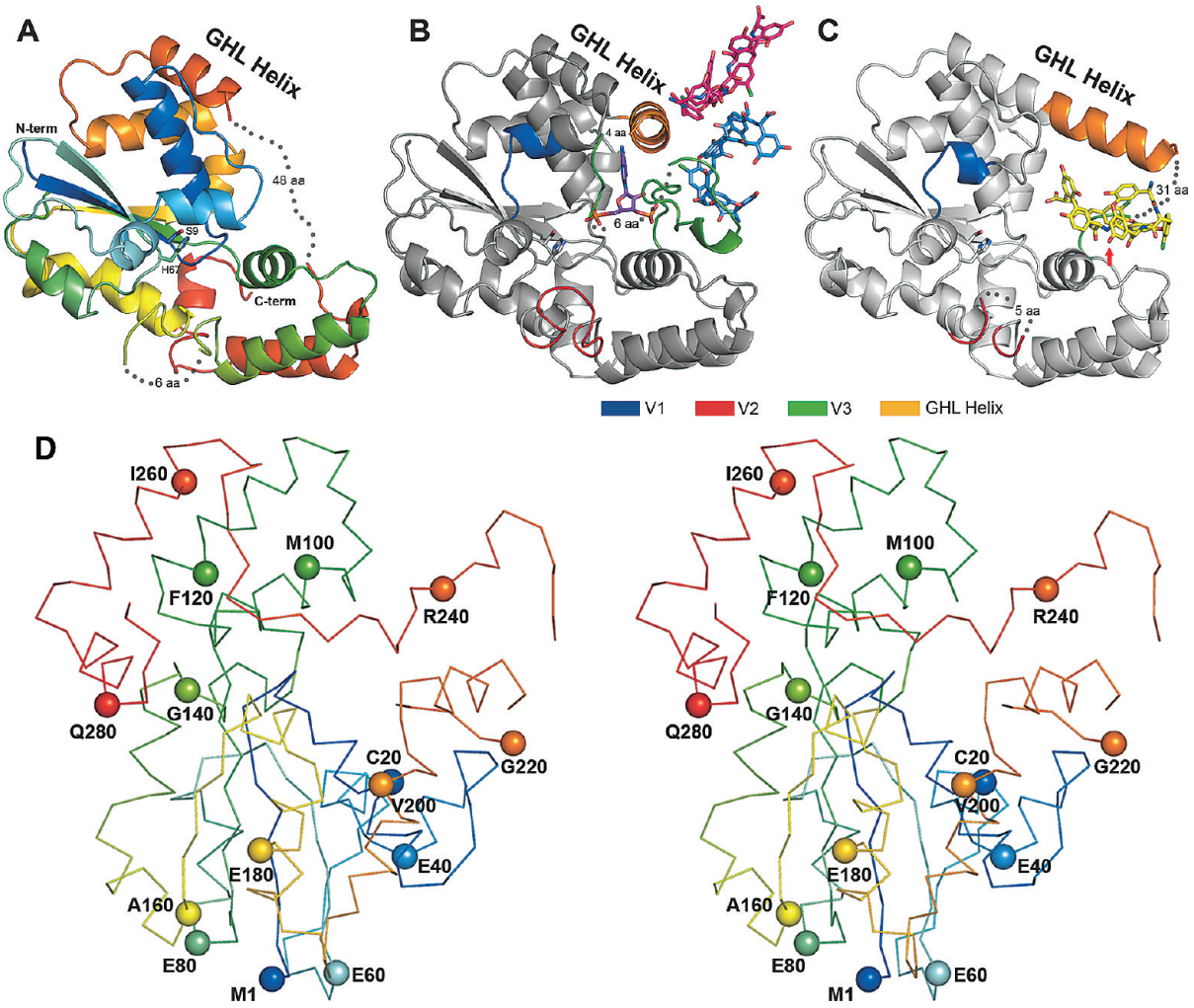

FIGURE 2: Three Teg12 structures. Only a single monomer from the dimer is shown for the sake of clarity. (A) Teg12-apo is colored using a rainbow scheme, from blue (N-terminus) to red (C-terminus). Regions of disorder are connected by gray dots, with the number of disordered residues indicated. The flexible GHL helix is colored orange. Side chains for the proposed active site residues His67 and Ser9 are also shown. (B) Teg12-ternary complex containing PAP and the teicoplanin aglycone. The protein is colored gray to accentuate regions of the structure that differ from Teg12-apo. Specifically, these regions are variable loops V1−V3, colored blue, red, and green, respectively. Also shown are a molecule of PAP in the active site, a molecule of teicoplanin aglycone that interacts with the V3 loop (sky blue), and an additional molecule of teicoplanin aglycone involved in crystal packing interactions (hot pink). (C) Teg12-binary structure in a complex with teicoplanin aglycone in the active site cavity. Again, the protein is colored gray, with V1−V3 colored as in the ternary structure. The active site is shown as a stereo close-up image in Figure 6. Teg12 sulfates the hydroxyl of residue 3 of the teicoplanin aglycone, which has been denoted with a red arrow. Omitted from the binary structure is a second molecule of teicoplanin aglycone bound to the outside of the protein on the opposite side of the GHL helix. (D) Stereoview of a Cα trace of the Teg12-ternary structure represented as a ribbon diagram with the Cα atom of every 20th amino acid shown as a sphere.

substitution of Bht for Tyr at the second position in the peptide core. All three TEG sulfotranferases can use the teicoplanin aglycone as a substrate. We therefore used this molecule in Teg12 cocrystallization experiments. The V_{max} and K_m values for the teicoplanin aglycone were determined to be 215.3 ± 25.2 nmol min^{-1} mg^{-1} and 59.6 ± 1.1 μM, respectively. Attempts to obtain cocrystals by soaking Teg12-apo crystals with either the teicoplanin aglycone or the cosubstrate PAP were unsuccessful. The addition of PAP to drops containing apo crystals caused the crystals to rapidly dissolve into the mother liquor. Therefore, we cocrystallized Teg12 and the teicoplanin aglycone with and without PAP. PAP is the desulfated byproduct of PAPS from the sulfonation reaction. Our cocrystallization experiments led to a Teg12-binary structure bound to teicoplanin, and a ternary structure bound to PAP and the teicoplanin aglycone. These structures represent the first examples of sulfotransferase structures containing glycopeptide substrates.

Teg12-ternary Structure. We obtained a 2.05 Å structure of a Teg12−PAP−teicoplanin aglycone ternary complex by cocrystallization (Figure 2B,D). Several regions of the Teg12 sequence not seen in the other Teg12 structures could be modeled into the electron density. The loop region of Gly127−Val137, which is primarily disordered in the other Teg12 structures, is fully ordered in one of the monomers of the ternary complex. This short loop contains the V2 region, one of the three regions that are highly variable within the three TEG sulfotransferase sequences. The V2 loop in Teg12 is extended by four amino acids compared with StaL, and by two amino acids when compared with Teg13 and Teg14. A much larger portion of the GHL helix and loop could be modeled in the ternary structure. The GHL helix extends through Ser15 in one monomer. In the opposite monomer, the helix extends to Gln218 and transitions uninterrupted to the GHL loop.

Both monomers of the ternary structure contain a single molecule of PAP bound in the predicted active site location.

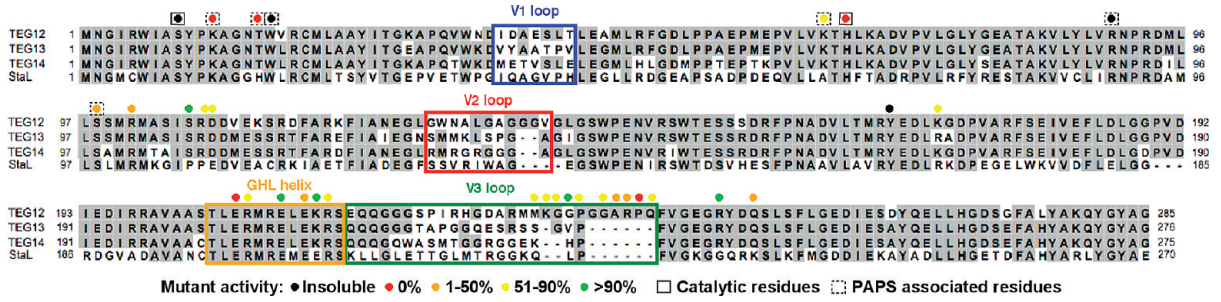

FIGURE 3: ClustalW alignment of glycopeptide sulfotransferases. Highly variable sequences (V1−V3) and the flexible part of the GHL helix appear in colored boxes that match the coloration seen in panels B and C of Figure 2. Results from alanine exchange mutagenesis experiments are color-coded by percent activity. Results for catalytic and PAPS-associated residues are shown in solid and dashed boxes, respectively. All other mutated residues are thought to interact with the glycopeptides. The H67A mutant was insoluble, and therefore, H67Q data are reported.

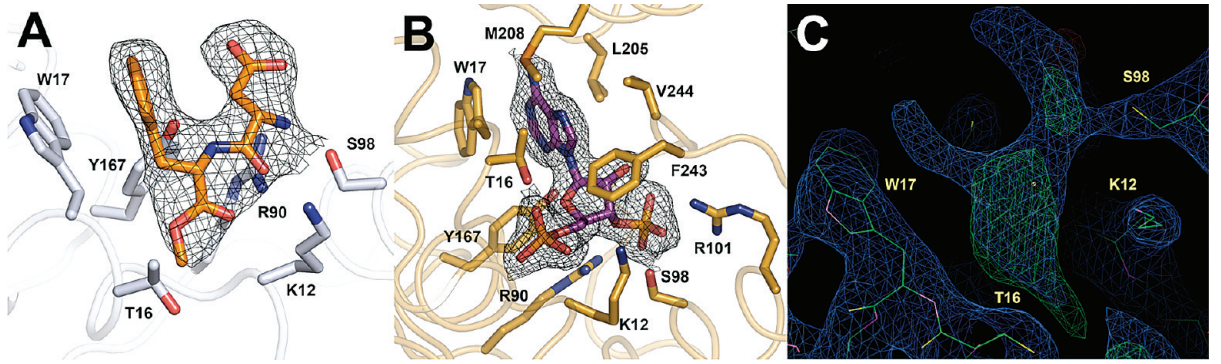

FIGURE 4: Comparison of the ligands bound in the PAPS binding sites from Teg12-apo and Teg12-ternary. Electron density from the $2F_o - F_c$ map surrounding each ligand and contoured at 1σ is shown. (A) The PAPS binding site of monomer A from Teg12-apo contained a molecule of the dipeptide aspartame [*N*-(L-α-aspartyl)-L-phenylalanine, 1-methyl ester]. Several Teg12 residues important for coordinating PAP also make contacts with aspartame. Their side chains are shown. (B) PAP bound in the Teg12-ternary complex. Leu205 and Met208 from the GHL helix and Phe243 and Val244 from the V3 loop make significant van der Waals contacts with PAP. (C) Electron density in the PAPS binding site of Teg12-apo monomer A, before the addition of aspartame to the model. Both the $2F_o - F_c$ (contoured at 1σ) (blue) and the $F_o - F_c$ (contoured at 3σ) (green) maps are shown. Aspartame was added to the model only after the protein had been fully built, at the second to last step of refinement (before the addition of waters).

Three highly conserved structural motifs in sulfotransferases are known to play a role in binding PAPS, all of which are seen in Teg12. Lys12 from the 5′-phosphate binding loop (PSB) (Pro11−Thr16) forms a hydrogen bond with the 5′-phosphate of PAP; Ser98 from the 3′-phosphate binding loop (PB) (Val89−Ser99) forms a hydrogen bond with the 3′-phosphate of PAP, and Trp17 stacks in parallel with the adenine base. A number of additional residues from the PSB loop region, the PB loop region, and the conformationally flexible GHL region are within hydrogen bonding distance of PAP (Figure 4B). Thr16 is only ~2.6 Å from the 5′-phosphate, to which it hydrogen bonds. This residue corresponds to a histidine in StaL, although threonine is the more common side chain at this position in eukaryotic sulfotransferases (*13*). A general hydrogen bond donor at position 16 is likely to be important for properly coordinating PAPS in the active site. Indeed, an alanine replacement mutant of Thr16 was inactive for the sulfation reaction. A K12A mutant was also inactive, while S98A was partially active (27% compared with that of the wild type). A W17A mutant did not express as a soluble protein, indicating a role not only in binding PAP but also in protein folding and stability.

PAPS-dependent sulfotransferases are believed to use a histidine as a general base to activate the hydroxyl or amine used in the sulfate exchange reaction. Teg12 contains three histidine residues, His67, His226, and His269. Of these, only His67 is located near the cosubstrate binding site. The 5′-phosphate of PAP is ~5 Å from the imidazole of His67, and when PAPS is modeled in its place, the sulfate is positioned directly adjacent to His67 (Figure 7). In this model, both Lys12 and Lys65 are sufficiently close to coordinate with the sulfate. In place of the sulfate from PAPS, a water molecule occupies this space in the ternary structure. Ser9, which is within hydrogen bonding distance of the imidazole ring, was proposed to help activate the catalytic histidine in StaL. Attempts to mutate the predicted catalytic histidine in StaL resulted in an insoluble protein, and therefore, its role in catalysis had not been previously confirmed (*13*). Alanine exchange mutants were generated for His67, Ser9, and Lys65. H67A and S9A did not express as soluble proteins, and K65A was only partially active (65%). Additional glutamate and glutamine mutants of His67 were generated to potentially provide better expression. H67E also expressed as an insoluble protein. H67Q, which was soluble, produced a completely inactive sulfotransferase.

Teicoplanin Aglycone Bound in the Ternary Structure. One of the monomers in the ternary structure, in addition to PAP, contains two molecules of teicoplanin aglycone. Both molecules of the aglycone are located outside of the predicted active site pocket and appear to be organized primarily by crystallographic forces (Figure 5A). One of the molecules of aglycone makes numerous contacts with the GHL loop (Figure 5B). Beginning with the N-terminal section of the GHL loop, residues Gly220−Ile224 fold back toward the dimer interface. Pro223 stacks with Trp134 and helps to stabilize this region. The six subsequent amino acids (Arg225−Arg230) are disordered, before

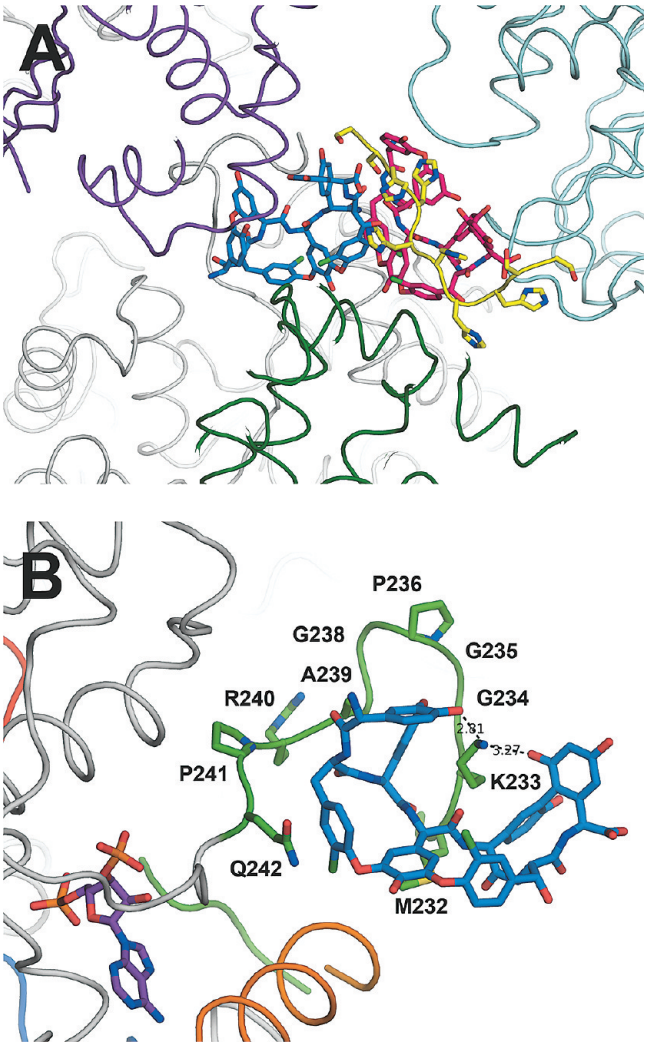

FIGURE 5: (A) Two molecules of teicoplanin aglycone were modeled in the Teg12-ternary complex, colored sky blue and hot pink. Their location within the crystal appears to mediate the packing of several Teg12 monomers, as shown in purple, cyan, dark green, and gray. The two teicoplanins belong to the gray Teg12 in the crystal. Also colored yellow is the six-His portion of the N-terminally derived tag from the pET28a expression vector. The tag−aglycone interaction is proposed to mimic the binding of the aglycone to its cellular target, D-Ala-D-Ala of the bacterial cell wall. (B) Close-up of the GHL loop−aglycone complex from the Teg12-ternary structure. Side chains for all alanine replacement experiments are shown. Lys233 intercalates into the glycopeptide and makes hydrogen bond contacts with residues 1 and 7 of the aglycone. For the sake of orientation, the molecule of PAP is colored purple. Also colored orange is the GHL helix.

density picks up again at Met231. While disordered, these residues occupy a region of space where they could easily make additional contacts with the glycopeptides. Lys233−Gln242 closely trace the outer surface of the three N-terminal amino acids of the aglycone. These 10 amino acids immediately precede the five residues (Phe243−Gly247) that interact with PAP. Lys233 intercalates between the Dpg at position 3 and the Hpg at position 5, and it forms a hydrogen bond with hydroxyls from the Hpg at position 1 and the Dpg at position 7. The carboxylate from Asp271 of a symmetry mate makes several contacts with backbone nitrogens in the concave cleft of the glycopeptides. The C-terminal end of the teicoplanin aglycone extends out from the surface of the protein and interacts with a second glycopeptide molecule that is bound to the His tag from an adjacent monomer.

Table 3: Relative Activities of TEG12 Mutant Sulfotransferases

catalytic or PAP binding	activity (%)	V3 aglycone binding	activity (%)	non-V3 aglycone binding	activity (%)
S9A	N/A[a]	M232A	74	R101A	29
K12A	0	K233A	79	S106A	92
T16A	0	G234A	82	R107A	82
W17A	N/A[a]	G235A	100	D108A	82
K65A	61	P236A	73	Y167A	N/A[a]
H67A	N/A[a]	G238A	56	K171A	69
H67E	N/A[a]	A239R	37	E206A	0
H67Q	0	R240A	48	R207A	79
R90A	N/A[a]	P241A	0	E210A	97
S98A	27	Q242A	64	E212A	26
				K213A	99
				R214A	67
				R248A	196
				Q251A	41

[a]The mutation produced an insoluble protein.

Despite the assumption that these aglycone−GHL interactions are mediated primarily by crystal packing, we generated a series of mutants in GHL loop residues that contact the glycopeptides. All mutations in this loop region, with the exception of G235A, had deleterious effects on the rate of substrate conversion, supporting the notion that the V3 loop interacts with the substrate and is involved in substrate recruitment. The P241A mutation abolished Teg12's activity completely. While the glycopeptide−GHL interactions observed in the ternary structure are clearly stabilized by crystal packing interactions, our mutagenesis experiments make it so we cannot rule out the possibility that this represents a snapshot of the substrate recruitment path. A summary of all mutations introduced into Teg12 and their corresponding activities is given in Table 3.

Teg12-binary Structure and Teg12 Active Site. We determined a Teg12−teicoplanin aglycone structure to 2.27 Å resolution (Figure 2C). A single monomer constitutes the asymmetric unit, and the dimer is generated by crystallographic symmetry. The active site contains a single molecule of teicoplanin aglycone. Two sets of side chain hydrogen bonds play key roles in creating the glycopeptide binding cavity. The base of the binding pocket is largely defined by the outstretched side chain of Arg107 coordinating with the side chain of Gln251. The back of the cavity is largely defined by two outstretched arginine side chains (Arg214 and Arg248) that interact with each other via a mediating water molecule. Numerous hydrogen bonds exist between amino acid side chains from residues found in the GHL helix (Glu206, Arg207, Glu210, and Lys213) and the glycopeptide (Figure 6). Several of the residues that make contacts with PAP also interact with the teicoplanin aglycone (Lys12, Arg90, Arg101, and Tyr167). Side chains from Tyr167 and Lys12 coordinate the C-terminal carboxylate, while a backbone carbonyl oxygen from Lys213 and water hydrogen bonded to Glu216 coordinate the N-terminal amine.

The GHL helix and the helix immediately N-terminal to it (residues Ile193−Ser203) form a single contiguous helix that runs along the top surface of the aglycone (Figure 2C). In the ternary structure with an empty active site pocket, these two helices adopt a bent conformation that results from unwinding of the long helix at Thr204 [compare the orange GHL helix (Figure 2B,C)]. This bent conformation is also observed in StaL. In the bent

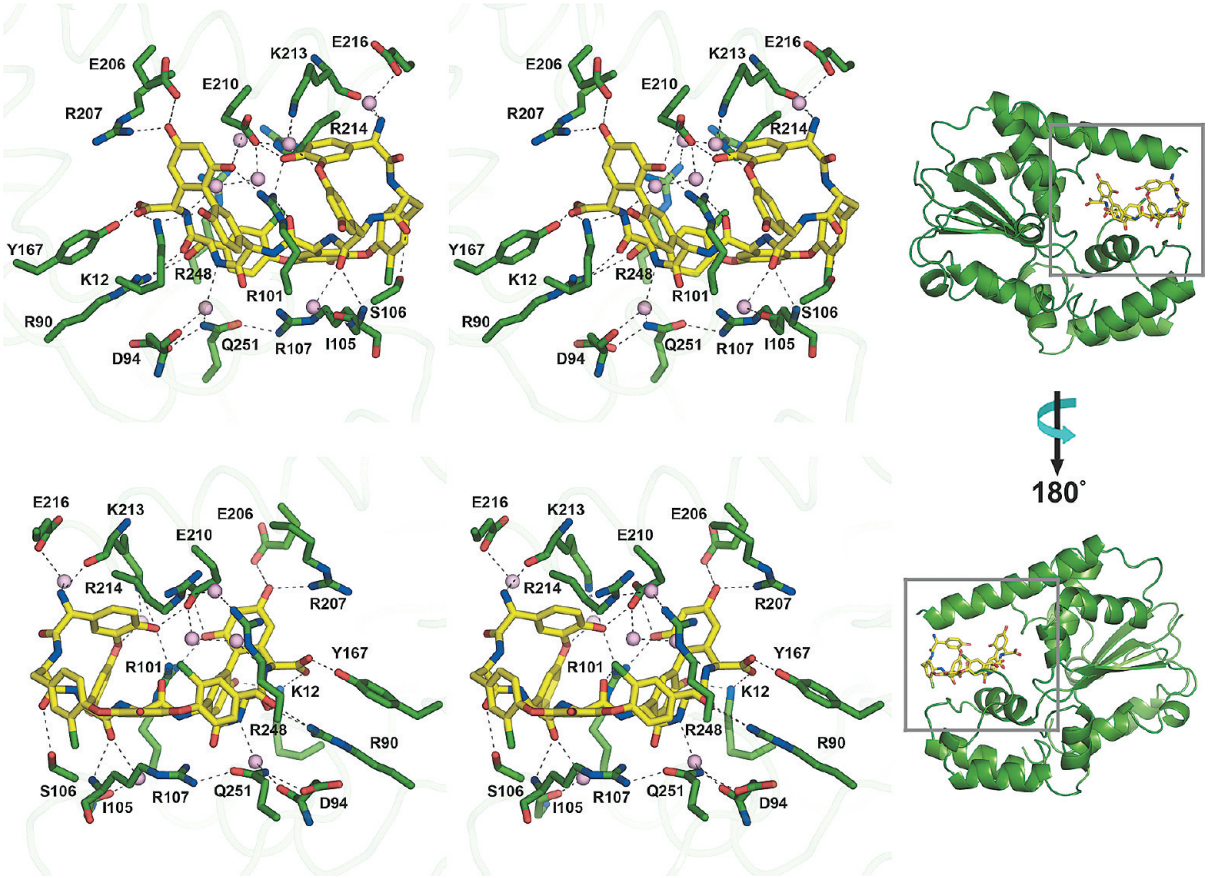

FIGURE 6: Stereoviews of the front and back sides of the Teg12−aglycone complex in the active site of the Teg12-binary structure. Protein residues are colored green, and the aglycone is colored yellow. Hydrogen bonds are represented as dashed lines. Water molecules that make important hydrogen bonds are shown as pink spheres.

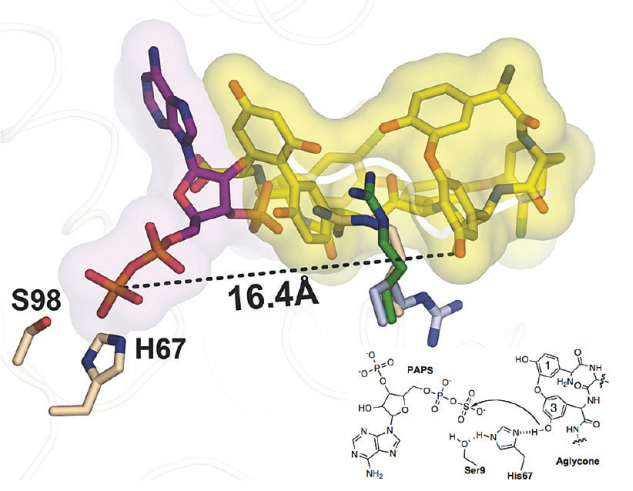

FIGURE 7: Composite of the active site from the Teg12-binary and -ternary structures. PAPS was modeled into the ternary structure in place of PAP. In this aligned view, PAPS and teicoplanin aglycone would not be able to occupy their respective positions simultaneously, as there would be strong clashes between the two substrates. Residue 3 of the teicoplanin aglycone, the site of sulfation, is more than 16 Å from the sulfate of PAPS when the ternary and binary structures are overlapped. A substantial rearrangement of teicoplanin aglycone within the active site would have to occur for sulfation to proceed by the proposed in-line attack mechanism. The positions of Arg101 in the binary (green) and ternary (tan) structures appear to preclude the movement of teicoplanin aglycone toward His67. Arg101's position in the context of the apo structure (gray) creates a more open active site, where the aglycone could pivot toward His67 more easily.

conformation, the helices are at a nearly perpendicular angle to one another. The long, largely disordered loop region of the GHL that interacts with PAP in the ternary structure is completely displaced by the bound glycopeptide in the binary structure. This disordered loop is now positioned at the back of Teg12, behind the bound glycopeptide. The front surface of the glycopeptide, where the sulfotransferase reaction is predicted to occur on residue 3 (see the red arrow in Figure 2C), is largely exposed to solvent. With the exception of an additional small compensatory movement of a short loop (Gln31−Ile37) into the void left by the straightening of the flexible helix from the GHL, the remainder of this Teg12-binary structure is essentially identical to that seen in the other Teg12 structures.

The C-terminus of the glycopeptide is buried in the PSB loop from the strand−loop−helix motif, and the helix from the PB loop containing the strand−turn−helix motif runs along the underside of the glycopeptide. While most residues contact only the outer surface of the glycopeptide, one residue, Arg101, extends into the core of the molecule. The side chain of Arg101 intercalates between the Dpg at position 3 and the Hpg at position 5. The guanidinium from Arg101 forms a hydrogen bond with the backbone carbonyl from the Hpg at position 4 as well as with hydroxyls from the Dpg at position 3 and the Dpg at position 7. In this respect, Arg101 behaves in much the same way as Lys233 in the ternary structure. We generated a series of alanine replacement mutants for residues contacting the aglycone to determine which are important for substrate conversion. The results are summarized in Table 3. R101A exhibited a

reduction in activity to 27% of the wild-type activity. Interestingly, an E212A mutation also showed a reduction of activity to 26%. Glu212 is part of the GHL helix but does not make contacts with the aglycone in the binary structure. Instead, this residue is found on the upper, solvent-exposed face of the helix. Arg248 along with Arg214 forms the back of the active site cavity. An alanine placed at position 248 resulted in a 2-fold increase in the degree of sulfation. Removing the bulky side chain at this position may impart more flexibility to the active site, allowing the glycopeptide to shift more easily toward His67 (see Discussion).

Although the glycopeptide substrate interacts with different sets of GHL residues in the binary and ternary structures, the general glycopeptide binding motif seen in these structures is very similar. In both cases, a long positively charged amino acid side chain (Lys233 in the ternary complex and Arg101 in the binary complex) extends into the core of the glycopeptide, where it coordinates with the same set of side chain hydroxyls. All other glycopeptide−protein contacts observed in these structures involve the outer surface of the glycopeptide.

In addition to the glycopeptide bound in the active site cavity, there is a second glycopeptide molecule bound on the surface of the protein (not shown in Figure 2C). The location of the second teicoplanin aglycone in the crystal is likely not biologically relevant. Instead, it appears to mediate crystal contacts between symmetry partners. A stretch of electron density, large enough to contain four residues, is located in the D-Ala-D-Ala binding cleft of the second aglycone. Although it is presumed that this stretch of residues comes from the N-terminal vector-derived His tag of an adjacent symmetry mate, we were not able to determine its sequence, and hence, it was modeled as polyalanine.

DISCUSSION

Teg12 sulfates the hydroxyl on the dihydroxyphenyl glycine at position 3 of the teicoplanin aglycone (Figure 2C, red arrow). In an alignment of the binary and ternary structures, this hydroxyl is 16.4 Å from where the sulfate of PAPS would be and there is significant overlap between the C-terminus of the glycopeptide and PAP (Figure 7). Before the sulfation reaction can proceed, there must be a substantial conformational rearrangement in the active site. In the binary structure, the interaction of the glycopeptide with Arg101 appears to preclude the substrate from shifting to accommodate PAPS or from its shifting into the vicinity of His67. Arg101 makes mutually exclusive contacts with the substrates present in the binary and ternary structures. In the binary structure, it is inserted into the peptide core and forms a hydrogen bond with residues 1 and 7, while in the ternary structure, it forms a key salt bridge with the 3-phosphate of PAP. If, when both substrates are present in the active site, Arg101 were to adopt the conformation seen in the apo structure, the glycopeptide would largely be free to move out of the PAPS binding pocket and toward His67. By pivoting toward His67, the glycopeptide would occupy the same general location to which substrates bind in eukaryotic sulfotransferases (*17*), and it would also be positioned so that it could interact with all three variable loops.

The movement the glycopeptide must undergo to reach the PAPS sulfate traces a path similar to that which the GHL helix undergoes when it bends. This conformational rearrangement not only may serve as a mechanism to allow large glycopeptide substrates into the active site but also could play a role in positioning a glycopeptide for catalysis once it enters the active site. Alanine exchange mutagenesis of residues found on both the lower glycopeptide-associated face (E206A, R207A, and R214A) and the upper solvent-exposed face of the helix (E212A) led to decreases in the degree of substrate conversion. Residues around the entire helix likely make key contacts with the glycopeptide at different points in time as a result of the rotation in the helix as it bends. In fact, all residues that contact the glycopeptide in the binary structure face out from the active site cavity in the ternary structure. One alanine exchange mutant, R248A, resulted in a 2-fold increase in activity. R248, along with R214 of the GHL helix, and a mediating water molecule form a cagelike enclosure over the top of the glycopeptide (Figure 6). Introduction of an alanine side chain at this position would open up the back side of the active site to the solvent and could provide substrates in the active site more room for positional rearrangement. The sequence of the GHL helix is conserved throughout the TEG sulfotransferases, and a large portion of it is also conserved in StaL. The mechanics of the conformational change from the straight to bent structures may therefore be conserved throughout this family of finishing enzymes.

Sequence and structural similarities between eukaryotic and prokaryotic sulfotransferases indicate that all members have likely arisen from one common ancestor. In eukaryotic structures, the loop that would correspond to the V3 loop of the GHL behaves in much the same way as it does in Teg12; it is disordered in apo structures and associated with PAP and the substrate in cocrystal structures (*17*). The open active site conformation that results from the straightening of the GHL helix was not observed in previous StaL structures. While a GHL-like helix−loop region is present in eukaryotic sulfotransferases, the helix from these structures does not appear to exhibit the same conformational rearrangements from a single straight helix to two helices with a bent conformation during substrate binding as it does in Teg12. At >1000 Da, glycopeptides are significantly larger than the substrates used by most eukaryotic sulfotransferases. The increased conformational plasticity that results from the ability of the GHL helix to easily flex may help TEG-like sulfotransferases accommodate larger substrates.

In addition to the conformational flexibility seen in the V3 loop, both the V1 and V2 loops adopt different conformations in this series of structures, suggesting they could easily reorganize to accommodate an incoming glycopeptide substrate. The sequence differences among Teg12, -13, and -14 suggest that these three short variable regions likely control the selection and orientation of the glycopeptide substrate bound in the active site. Systematically altering the residues found in the three TEG variable regions may provide a means of generating new glycopeptide finishing enzymes that sulfate a broader collection of glycopeptide congeners than is currently possible with the small number of sulfotransferases that have been identified naturally.

The sulfated teicoplainin aglycone derivatives produced by the native TEG sulfotransferase retain potent in vitro antibacterial activity (*5*). The cloning and characterization of biosynthetic gene clusters derived from uncultured bacteria provide a means to access both novel small molecules and new biosynthetic enzymes. Teg12 is one of the first enzymes discovered using culture-independent methodologies to be characterized structurally, and the Teg12−teicoplanin aglycone cocrystal structures are the first examples of a substrate complexed with a member of this family of glycopeptide finishing enzymes. This series of Teg12 structures provides key insights into how sulfotransferases might be engineered to

generate additional anionic glycopeptides that could be evaluated against clinically relevant drug resistant bacteria.

ACKNOWLEDGMENT

We thank Wuxian Shi of beamline X29 at the NSLS, Kanagalaghatta Rajashankar of the NE-CAT 24-ID beamline at the APS, and Deena Oren of the Rockefeller University Structural Biology Resource Center. We also thank Nigel Moriarty for his assistance with Phenix.elbow and generation of a teicoplanin restraints file for use in refinement. The teicoplanin aglycone was generously provided by Supelco.

REFERENCES

1. Linden, P. K. (2008) Vancomycin resistance: Are there better glycopeptides coming? *Expert Rev. Anti-Infect. Ther. 6*, 917–928.
2. Uttley, A. H., Collins, C. H., Naidoo, J., and George, R. C. (1988) Vancomycin-resistant enterococci. *Lancet 1*, 57–58.
3. Hiramatsu, K., Aritaka, N., Hanaki, H., Kawasaki, S., Hosoda, Y., Hori, S., Fukuchi, Y., and Kobayashi, I. (1997) Dissemination in Japanese hospitals of strains of *Staphylococcus aureus* heterogeneously resistant to vancomycin. *Lancet 350*, 1670–1673.
4. Rappe, M. S., and Giovannoni, S. J. (2003) The uncultured microbial majority. *Annu. Rev. Microbiol. 57*, 369–394.
5. Banik, J. J., and Brady, S. F. (2008) Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. *Proc. Natl. Acad. Sci. U.S.A. 105*, 17273–17277.
6. Kahne, D., Leimkuhler, C., Lu, W., and Walsh, C. (2005) Glycopeptide and lipoglycopeptide antibiotics. *Chem. Rev. 105*, 425–448.
7. Hubbard, B. K., and Walsh, C. T. (2003) Vancomycin assembly: Nature's way. *Angew. Chem., Int. Ed. 42*, 730–765.
8. Nicolaou, K. C., Boddy, C. N. C., Brase, S., and Winssinger, N. (1999) Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem., Int. Ed. 38*, 2096–2152.
9. Boeck, L. D., and Mertz, F. P. (1986) A47934, a novel glycopeptide-aglycone antibiotic produced by a strain of *Streptomyces toyocaensis*: Taxonomy and fermentation studies. *Jpn. J. Antibiot. 39*, 1533–1540.
10. Higgins, D. L., Chang, R., Debabov, D. V., Leung, J., Wu, T., Krause, K. M., Sandvik, E., Hubbard, J. M., Kaniga, K., Schmidt, D. E., Jr., Gao, Q., Cass, R. T., Karr, D. E., Benton, B. M., and Humphrey, P. P. (2005) Telavancin, a multifunctional lipoglycopeptide, disrupts both cell wall synthesis and cell membrane integrity in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother. 49*, 1127–1134.
11. Stryjewski, M. E., O'Riordan, W. D., Lau, W. K., Pien, F. D., Dunbar, L. M., Vallee, M., Fowler, V. G., Jr., Chu, V. H., Spencer, E., Barriere, S. L., Kitt, M. M., Cabell, C. H., and Corey, G. R. (2005) Telavancin versus standard therapy for treatment of complicated skin and soft-tissue infections due to Gram-positive bacteria. *Clin. Infect. Dis. 40*, 1601–1607.
12. Leadbetter, M. R., Adams, S. M., Bazzini, B., Fatheree, P. R., Karr, D. E., Krause, K. M., Lam, B. M., Linsell, M. S., Nodwell, M. B., Pace, J. L., Quast, K., Shaw, J. P., Soriano, E., Trapp, S. G., Villena, J. D., Wu, T. X., Christensen, B. G., and Judice, J. K. (2004) Hydrophobic vancomycin derivatives with improved ADME properties: Discovery of telavancin (TD-6424). *Jpn. J. Antibiot. 57*, 326–336.
13. Shi, R., Lamb, S. S., Bhat, S., Sulea, T., Wright, G. D., Matte, A., and Cygler, M. (2007) Crystal structure of StaL, a glycopeptide antibiotic sulfotransferase from *Streptomyces toyocaensis*. *J. Biol. Chem. 282*, 13073–13086.
14. Lamb, S. S., Patel, T., Koteva, K. P., and Wright, G. D. (2006) Biosynthesis of sulfated glycopeptide antibiotics by using the sulfotransferase StaL. *Chem. Biol. 13*, 171–181.
15. Dombrovski, L., Dong, A., Bochkarev, A., and Plotnikov, A. N. (2006) Crystal structures of human sulfotransferases SULT1B1 and SULT1C1 complexed with the cofactor product adenosine-3′-5′-diphosphate (PAP). *Proteins 64*, 1091–1094.
16. Lin, E. S., and Yang, Y. S. (2000) Nucleotide binding and sulfation catalyzed by phenol sulfotransferase. *Biochem. Biophys. Res. Commun. 271*, 818–822.
17. Allali-Hassani, A., Pan, P. W., Dombrovski, L., Najmanovich, R., Tempel, W., Dong, A., Loppnau, P., Martin, F., Thornton, J., Edwards, A. M., Bochkarev, A., Plotnikov, A. N., Vedadi, M., and Arrowsmith, C. H. (2007) Structural and chemical profiling of the human cytosolic sulfotransferases. *PLoS Biol. 5*, e97.
18. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol. 276A*, 307–326.
19. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658–674.
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. 53*, 240–255.
21. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. 54*, 905–921.
22. Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat. Protoc. 2*, 2728–2733.
23. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. 60*, 2126–2132.
24. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. 57*, 122–133.
25. Schafer, M., Schneider, T. R., and Sheldrick, G. M. (1996) Crystal structure of vancomycin. *Structure 4*, 1509–1515.
26. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr. 58*, 1948–1954.
27. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., III, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res. 35*, W375–W383.
28. Sarkar, G., and Sommer, S. S. (1990) The "megaprimer" method of site-directed mutagenesis. *BioTechniques 8*, 404–407.