BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

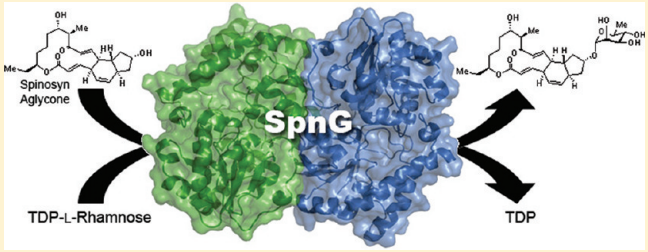
Structural Studies of the Spinosyn Rhamnosyltransferase, SpnG

Eta A. Isiorho,[†,‡] Hung-wen Liu,*[,†] and Adrian T. Keatinge-Clay*[,‡]

[†]Division of Medicinal Chemistry, College of Pharmacy, and [‡]Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712, United States

Supporting Information

ABSTRACT: Spinosyns A and D (spinosad), like many other complex polyketides, are tailored near the end of their biosyntheses through the addition of sugars. SpnG, which catalyzes their 9-OH rhamnosylation, is also capable of adding other monosaccharides to the spinosyn aglycone (AGL) from TDP-sugars; however, the substitution of UDP-D-glucose for TDP-D-glucose as the donor substrate is known to result in a >60000-fold reduction in k_{cat}. Here, we report the structure of SpnG at 1.65 Å resolution, SpnG bound to TDP at 1.86 Å resolution, and SpnG bound to AGL at 1.70 Å resolution. The SpnG−TDP complex reveals how SpnG employs N202 to discriminate between TDP- and UDP-sugars. A conformational change of several residues in the active site is promoted by the binding of TDP. The SpnG−AGL complex shows that the binding of AGL is mediated via hydrophobic interactions and that H13, the potential catalytic base, is within 3 Å of the nucleophilic 9-OH group of AGL. A model for the Michaelis complex was constructed to reveal the features that allow SpnG to transfer diverse sugars; it also revealed that the rhamnosyl moiety is in a skew-boat conformation during the transfer reaction.



Spinosyns A and D (**7** and **8**, respectively; spinosad) are secondary metabolites naturally produced by the rare actinomycete *Saccharopolyspora spinosa* and are the active ingredients of the commercial pesticide Tracer Naturalyte (Figure 1).[1,2] Spinosad possesses a unique mechanism of action, targeting the nicotinic acetylcholine and γ-aminobutyric acid (GABA) receptors to effect the rapid excitation of the nervous systems of such pests as caterpillars, leaf beetles, and lice. The biosynthetic gene cluster of spinosad has been identified; spinosyn is produced through the activities of 10 polyketide synthase (PKS) modules and several tailoring enzymes.[3−5] Among them, SpnJ, SpnM, SpnF, and SpnL each perform a key reaction during the post-PKS modification of the polyketide macrocycle.[6,7] SpnN, SpnQ, SpnR, and SpnS catalyze the construction of the forosamine moiety.[8,9] SpnH, SpnI, and SpnK are responsible for permethylation of the rhamnose unit.[10] Attachment of these two sugar appendages to the C9 and C17 positions of the aglycone is catalyzed by the rhamnosyl- and forosaminyltransferases, SpnG and SpnP, respectively.

On the basis of in vitro studies, the pathway for spinosyn tailoring has been established[7] with oxidation of the 15-OH groups of **1** and **2** by SpnJ as the first step.[6] Then, a double bond is formed between C11 and C12 through a SpnM-mediated 1,4-dehydration to allow a [4+2] intramolecular cycloaddition yielding **3** and **4**, respectively, that is catalyzed by SpnF. Subsequent 9-OH rhamnosylation by SpnG provides **5** and **6**, respectively. The rhamnose residue is then methylated by SpnI, SpnK, and SpnH in sequence[10] to produce the permethylated rhamnosyl tricyclic spinosyn pseudoaglycone (whether rhamnose methylation occurs before or after SpnL has not been determined). The final ring closure reaction between C3 and C14 is catalyzed by SpnL, and SpnP supplies a forosamine to the 17-OH group. The glycosyl transfer reactions catalyzed by SpnG and SpnP are significant because the rhamnose and forosamine residues of spinosad have been shown to be important for insecticidal activity. Derhamnosylated spinosad is 200 times less active than spinosad, while modifying the rhamnose or substituting it with another sugar can also affect activity.[11,12]

Most glycosyltransferases (GTs) transfer monosaccharides activated as nucleotide diphosphate-sugars (NDP-sugars) to nucleophilic oxygens, nitrogens, sulfurs, and carbons on acceptor substrates such as lipids, sugars, secondary metabolites, and proteins, in a regio- and stereospecific manner.[13] They are classified into families based on sequence homology (CAZy Carbohydrate-Active enZymes database, http://www.cazy.org[14,15]). When compared to each other, GTs typically display low levels of sequence identity; however, with only a couple of exceptions, the GTs that have been structurally characterized possess one of only two folds, GT-A and GT-B.[16]

Like most GTs found in the biosynthetic pathways of natural products (e.g., those in the rebeccamycin and calicheamicin pathways), the first GT in the spinosyn tailoring pathway, SpnG, belongs to glycosyltransferase family 1 (GT1), a group of inverting GT-B enzymes. SpnG is known to be capable of transferring sugars to the spinosyn aglycone (AGL, **9**) from TDP-sugar donors other than TDP-L-rhamnose and is thus a

Received: December 16, 2011
Revised: January 21, 2012
Published: January 23, 2012

   dx.doi.org/10.1021/bi201860q | *Biochemistry* 2012, 51, 1213−1222

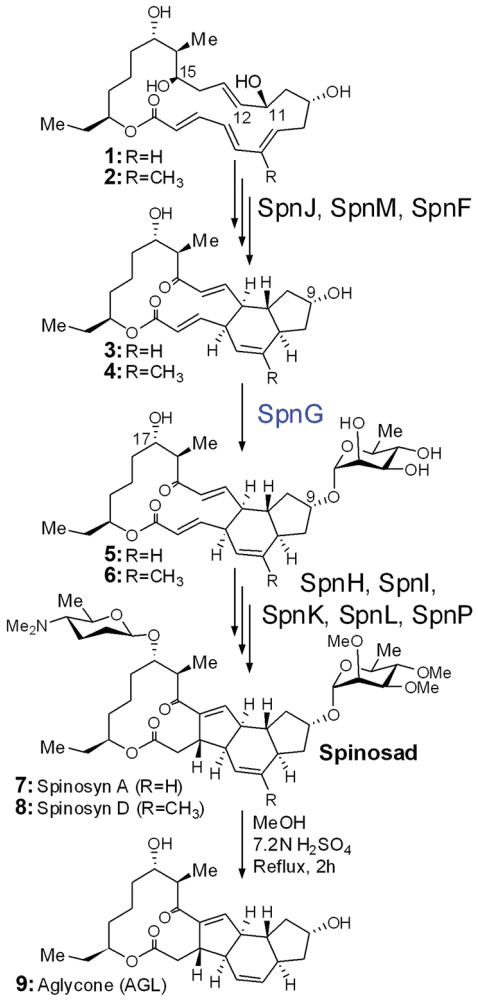


Figure 1. Spinosad biosynthetic pathway. SpnG is proposed to rhamnosylate intermediates **3** and **4** at their 9-OH groups.[7] The rhamnose is important for insecticidal activity.

potential biocatalyst in the generation of spinosyn derivatives with enhanced biological activity. SpnG is also known to be highly selective for TDP-sugars over UDP-sugars; the substitution of UDP-D-glucose for TDP-D-glucose results in a >60000-fold decrease in k_{cat} (Figure 1 of the Supporting Information).[17]

Here we present atomic-resolution structures of SpnG. The apo, TDP-bound, and AGL-bound structures help elucidate how substrates are recognized and catalysis is mediated by this GT. The 1.86 Å resolution structure of TDP-bound SpnG reveals that the loop residue N202 impedes binding of UDP-sugars by sterically occluding the ribose 2′-OH group. The binding of TDP was noted to promote a conformational rearrangement of several active site residues. The 1.70 Å resolution structure of the SpnG−AGL complex shows H13 is within 3 Å of the AGL 9-OH group, suggesting its role as the catalytic base. A model for the Michaelis complex including TDP-L-rhamnose and the acceptor substrate mimic, AGL, was also constructed; it reveals that the L-rhamnose moiety is likely in a skew-boat conformation during the transfer reaction.

MATERIALS AND METHODS

Cloning, Expression, and Purification of SpnG. The *spnG* gene was amplified from *S. spinosa* NRLL18537 genomic DNA by polymerase chain reaction (PCR) with 5′-AAA***CATATG***CGCGTACTCGTCGTTCCCTTGCCCTATCCGACGCATCTCAT-3′ and 5′-ATT***AAGCTT***TCAGGCACGGATGGCCGCAGTGTTCTCCAGTGT-3′ as described previously (NdeI and HindIII restriction sites in bold, stop codon underlined).[10] The amplicon was digested with NdeI and HindIII and ligated with a pET28b(+) backbone cut with the same enzymes. *Escherichia coli* TOP10 cells (Invitrogen) were transformed with the ligation reaction and plated on LB-agar (50 μg/mL kanamycin). One colony was mini-prepped to obtain the expression plasmid that was then transformed into the *E. coli* BL21(DE3) expression host. A discrepancy was noted from sequencing the expression plasmid at the codon that encodes residue 360 (an alanine replaces a serine reported in the National Center for Biotechnology Information database); this residue difference does not affect the glycosyltransferase activity of SpnG (data not shown).

The transformed expression host was grown in overnight cultures (5 mL of LB with 50 μg/mL kanamycin) that were used to inoculate 6 × 1 L of LB (50 mg/L kanamycin). The cultures were incubated with shaking at 37 °C until the OD_{600} reached 0.6, at which point isopropyl β-D-thiogalactopyranoside (final concentration of 0.5 mM) was added. The temperature was lowered to 17 °C, and the cultures were grown for an additional 18 h. Cells were harvested via centrifugation (5000*g* for 10 min), and the pelleted cells were resuspended in lysis buffer [5% (v/v) glycerol, 500 mM NaCl, and 30 mM HEPES (pH 7.5)] for sonication. The lysed cells were centrifuged (30000*g* for 30 min), and the lysate was loaded onto a Ni-NTA column (Qiagen). The column was washed with 15 column volumes of 20 mM imidazole in lysis buffer before SpnG was eluted with 200 mM imidazole in lysis buffer. SpnG was further purified over a gel filtration column (Superdex 200, GE Healthcare Life Sciences) equilibrated with 5% (v/v) glycerol, 150 mM NaCl, and 10 mM HEPES (pH 7.5). A protein concentrator (Amicon, YM10 membrane) was used to exchange the protein into 5% (v/v) glycerol, 25 mM NaCl, and 10 mM HEPES (pH 7.5) and achieve a concentration of 20−30 mg/mL. Aliquots were flash-frozen in liquid nitrogen and stored at −78 °C until further use. Two milligrams of pure SpnG was obtained per liter of culture.

Crystallization, Data Processing, and Structure Determination. Crystals grew overnight via the sitting-drop vapor-diffusion method. The crystallization buffer consisted of 18% (w/v) PEG 3350, 12.5−13.5% (w/v) glucose, 1% (v/v) glycerol, 0.1 M magnesium formate, and 0.1 M sodium cacodylate (pH 6.9). Each drop consisted of 2 μL of SpnG (20−30 mg/mL) and 1 μL of crystallization buffer. Crystals were flash-frozen in liquid nitrogen prior to data collection. SpnG-TDP crystals were obtained by supplying 20 mM TDP to the crystallization drop and incubating for 1 h prior to flash-freezing. SpnG-AGL crystals were obtained by adding 40 mM AGL to the crystallization drop and incubating for 1 h prior to flash-freezing. Data for apo-SpnG were collected at the Macromolecular Crystallography Facility (Rigaku MicroMax 007HF) of the University of Texas at Austin. Data for SpnG complexed with TDP and AGL were collected on Advanced Light Source Beamline 5.0.1 and subsequently processed with HKL2000 (Table 1).[18] The apo structure of SpnG was determined by molecular replacement with the UrdGT2 dimer [Protein Data Bank (PDB) entry 2P6P] as a search model by Phaser.[19,20] The model was initially refined with ARP/wARP and then built and refined through several cycles

Table 1. Data Collection and Refinement Statistics for SpnG and the SpnG−TDP and SpnG−AGL Complexes

	SpnG	SpnG−TDP	SpnG−AGL
	Crystallization		
resolution (Å)	25.00−1.65 (1.73−1.70)	25.00−1.86 (1.89−1.86)	29.83−1.70 (1.73−1.70)
wavelength (Å)	1.54	0.98	0.98
space group	*P*1	*P*1	*P*1
α, γ, β (deg)	81.5, 73.8, 86.0	81.6, 73.8, 85.7	82.0, 74.1, 86.1
no. of molecules per asymmetric unit	2	2	2
no. of measured intensities	131164	230234	314137
no. of unique reflections	85110	66093	84558
completeness (%)	89.5 (85.8)	93.1 (95.2)	96.9 (95.0)
R_{sym}	0.043 (0.328)	0.098 (0.567)	0.069 (0.512)
redundancy	1.7 (1.3)	3.7 (3.8)	3.8 (3.5)
average $I/\sigma I$	12.1 (2.9)	18.5 (2.6)	22.6 (2.0)
	Refinement		
no. of reflections	72,312	57,921	78,024
R_{cryst}/R_{free}	0.179/0.227	0.232/0.291	0.189/0.232
no. of protein atoms	5577	5535	5527
no. of ligand atoms	12	48	42
no. of solvent atoms	614	435	528
average *B* factor (Å^2)			
monomer A	18.7	19.4	19.9
monomer B	19.0	20.5	19.8
water	30.2	28.4	39.4
glucose	24.6	28.6	28.7
TDP	–	29.8	–
AGL	–	–	36.2
root-mean-square deviation			
bond lengths (Å)	0.025	0.019	0.023
bond angles (deg)	2.166	2.144	2.256
Ramachandran plot (%)			
preferred regions	97.83	97.14	97.00
allowed regions	2.04	2.72	2.86
outliers	0.14	0.14	0.14

with Coot and Refmac.[21−23] The structures of the SpnG−TDP and SpnG−AGL complexes were determined by molecular replacement using the apo-SpnG structure. The model for AGL was obtained by removing the sugars from the spinosyn A crystal structure.[24] The configuration of the 12-membered ring was slightly adjusted, using the appropriate regularization parameters within Coot, so that it fit the electron density map.[21]

Site-Directed Mutagenesis. SpnG mutants (Y10F, V94M, N202D, F315G, F315A and F315W) were engineered via site-directed mutagenesis (QuikChange, Stratagene) using primers 5′-GTTCCCTTGCCCTTTCCGACGCATCTCATG-3′ and 5′-CATGAGATGCGTCGGAAAGGGCAAGGGAAC-3′ for Y10F, 5′-GGAGCAGACCGCGTCCAATATGGCG-CAAAGCTCG-3′ and 5′-CGAGCTTTGCGCCATATTG-GACGCGGTCTGCTCC-3′ for V94M, 5′-CCGGTCCAG-TACGTGCCGTACGACGGAAGCGGCGC-3′ and 5′-GCGCCGCTTCCGTCGTACGGCACGTACTGGACCGG-3′ for N202D, 5′-GTGCTTCCCCAGTACGCCGAC-CAGTTCGACTAC-3′ and 5′-GTAGTC-GAACTGGTCGGCGTACTGGGGAAGCAC-3′ for F315A, and 5′-GTGCTTCCCCAGTACTGGGACCAGTTCGAC-TAC-3′ and 5′-GTAGTCGAACTGGTCCCAGTACTGGG-GAAGCAC-3′ for F315W. Mutations were confirmed through sequencing. SpnG mutants were purified as described for wild-type SpnG.

Preparation of Spinosyn A Aglycone (9). AGL (**9**) was prepared as previously described.[25] Briefly, spinosad (0.85 mmol, Dow AgroSciences) was dissolved in 70 mL of a methanol/7.2 N H_2SO_4 mixture (1:1.5) and refluxed for 3 h. The mixture was cooled to room temperature and the reaction quenched through the addition of solid $NaHCO_3$ before the solution was extracted with diethyl ether (three times). The combined organic layers were washed with brine (twice), dried over anhydrous $MgSO_4$, and concentrated under vacuum. AGL was then purified by flash chromatography using a methanol gradient (0 to 7.5% methanol/dichloromethane gradient).

Enzymatic Preparation of TDP-D-Glucose. Preparation of TDP-D-glucose was achieved through a reported procedure.[26] Specifically, in a 15 mL volume with 50 mM Tris (pH 7.5), 78 mM phosphoenolpyruvate, 24 mM thymidine, 1.6 mM ATP, and 27 mM $MgCl_2$ were incubated with 1000 units of pyruvate kinase, 75 mM thymidine kinase, 74 mM thymidylate kinase, and 113 mM nucleoside diphosphate kinase at 37 °C for 4 h. The mixture was then filtered through a YM-10 filter (Millipore) to remove enzymes from the mixture. The filtrate was incubated with D-glucose 1-phosphate (0.51 mmol), $MgCl_2$ (0.45 mmol), and 41 mM α-D-glucose-1-phosphate thymidylyltransferase (RfbA) for 16 h at 37 °C. The reaction mixture was centrifuged (5000*g* for 10 min), and RfbA was removed via filtration. After the flowthrough had been frozen and lyophilized, the product residue was passed through a P2 size-exclusion column (Bio-Rad) using water as the eluent.

dx.doi.org/10.1021/bi201860q | *Biochemistry* 2012, 51, 1213−1222

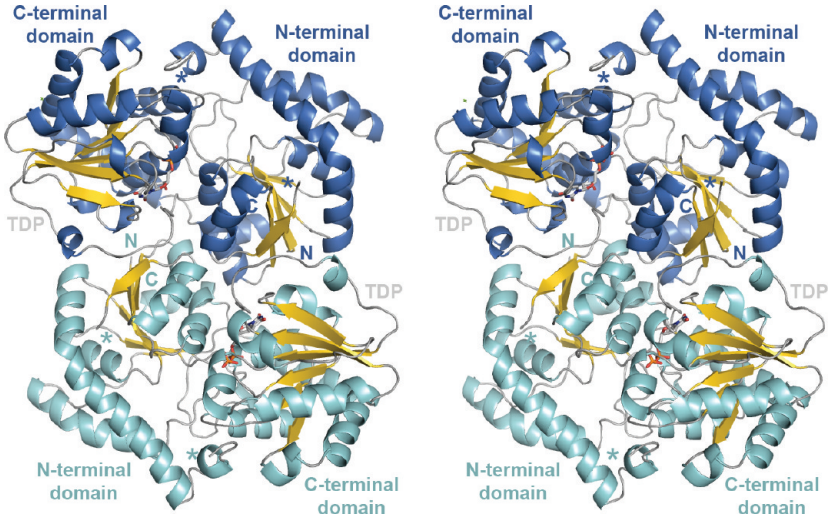


Figure 2. SpnG dimer. A stereodiagram shows the SpnG dimer that comprises the asymmetric unit. Asterisks mark the beginning and end of a disordered loop (residues 56−67). TDP is represented as sticks. The N- and C-termini of each monomer are indicated.

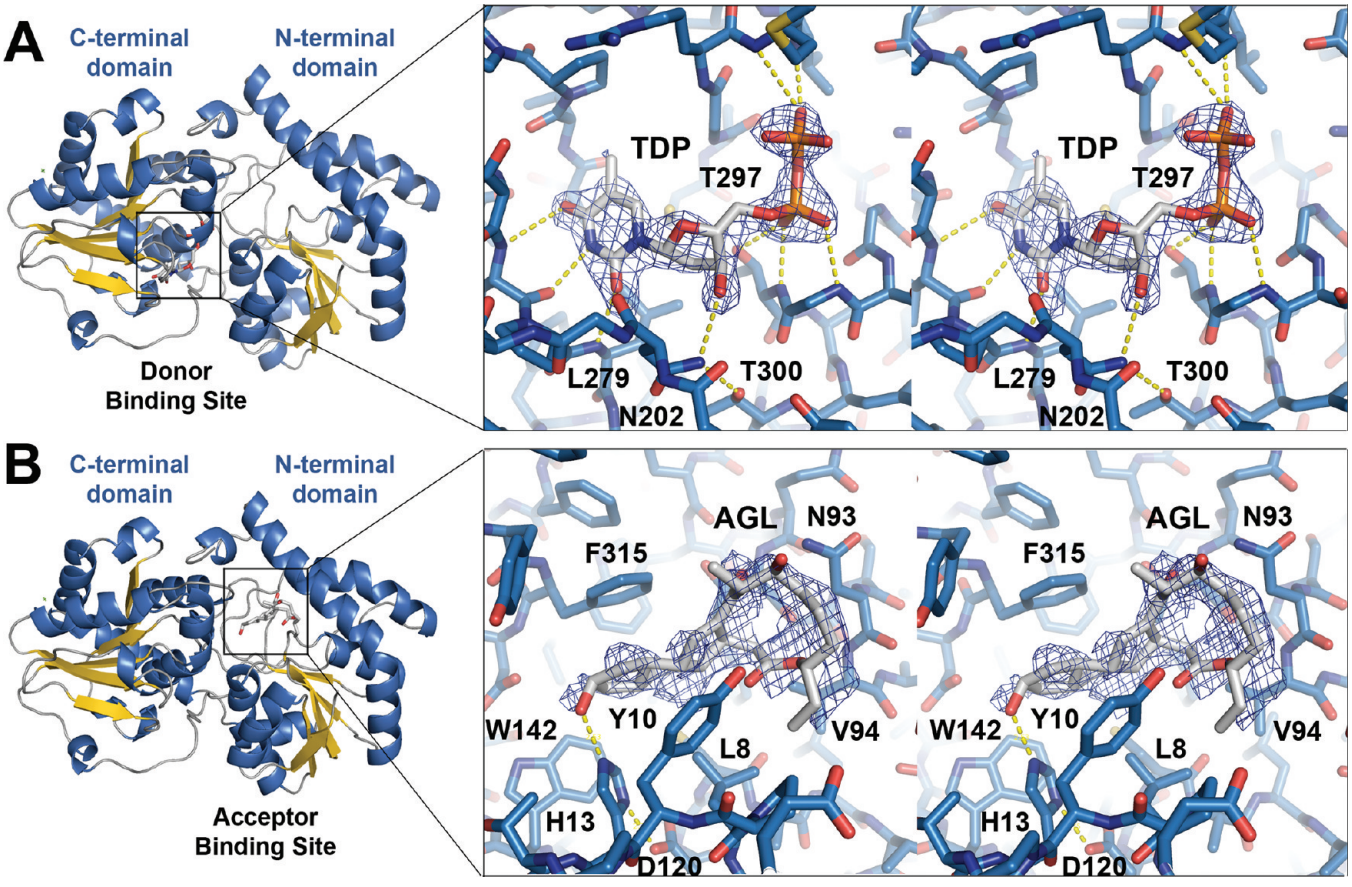


Figure 3. Donor and acceptor binding sites. (A) Stereodiagram of the donor binding site of the SpnG−TDP complex showing the simulated annealing $F_o - F_c$ omit map of TDP (contoured at 2.5σ). The thymine base, deoxyribose, and the pyrophosphate of TDP form hydrogen bonds with the donor binding site. (B) Stereodiagram of the acceptor binding site of the SpnG−AGL complex showing the simulated annealing $F_o - F_c$ omit map of AGL (contoured at 2.5σ). Though AGL is a substrate mimic, the putative natural substrates **3** and **4** likely bind similarly via hydrophobic contacts.

Fractions that absorbed at 267 nm were pooled, frozen, and lyophilized. The resulting TDP-D-glucose was stored at −20 °C until further use.

Glycosyltransferase Assay. The SpnG glycosyltransferase mixtures consisted of 2.5 mM AGL, 0.5% (v/v) DMSO, 2.5 mM $MgCl_2$, 12 mM TDP-D-glucose, 25 μM SpnG, and 50 mM HEPES (pH 8.5) (final volume of 100 μL), and reactions were performed at 37 °C for 45 min. Samples (25 μL) withdrawn at 0, 15, 30, and 45 min were quenched with ethyl acetate (100 mL). The resulting mixtures were spun down, and the extracted glucosylated product was quantified by an HPLC system equipped with a photodiode array detector (Waters) (Figure 1

dx.doi.org/10.1021/bi201860q | *Biochemistry* 2012, 51, 1213−1222

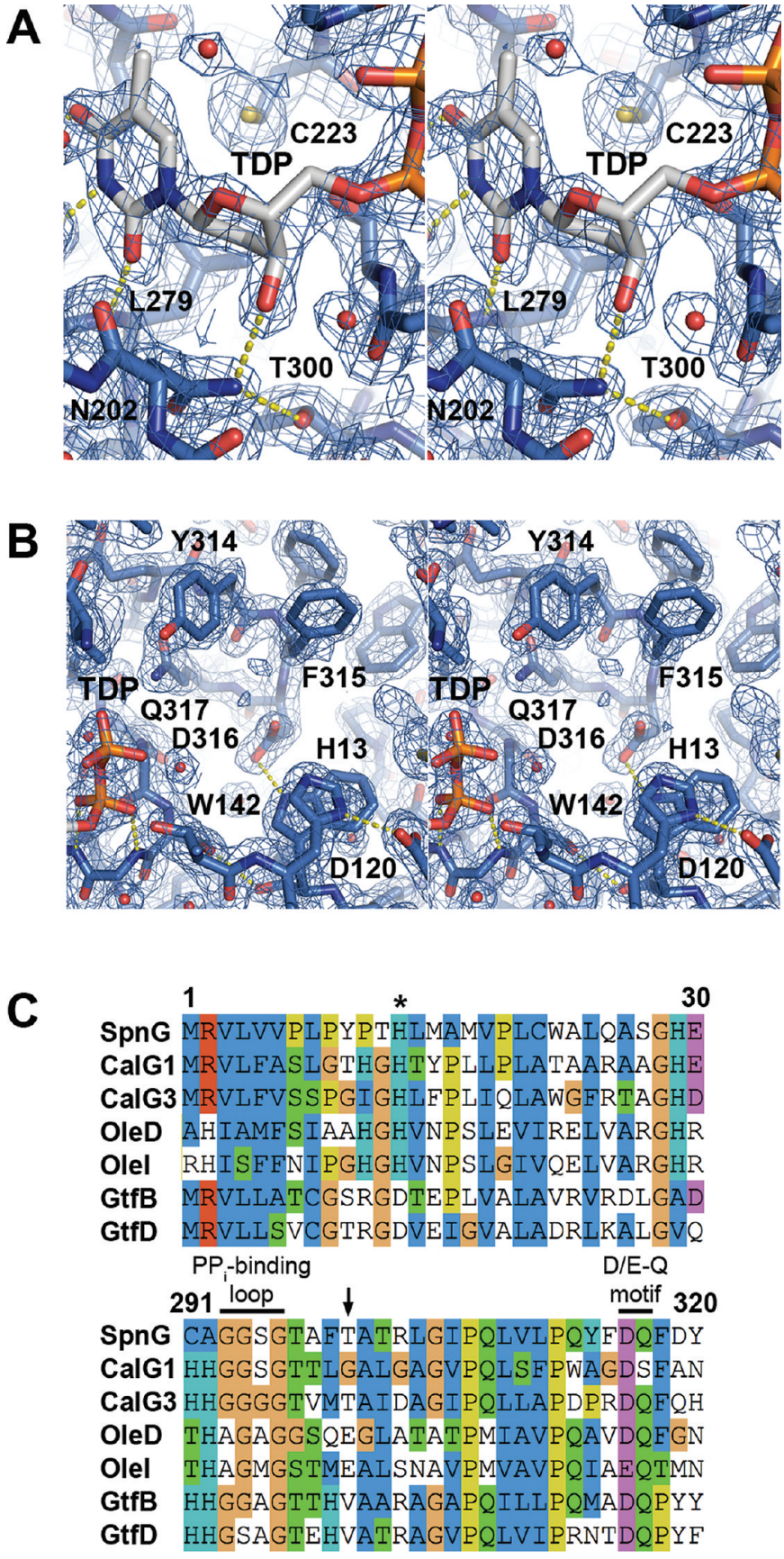


Figure 4. continued

dx.doi.org/10.1021/bi201860q | *Biochemistry* 2012, 51, 1213−1222

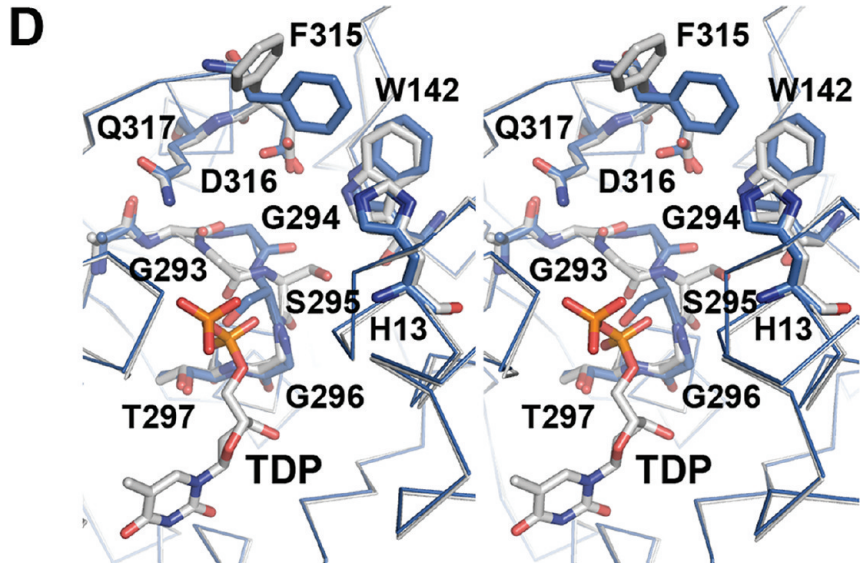


Figure 4. Active site residues. (A) Stereodiagram of the $2F_o - F_c$ electron density maps of the SpnG−TDP complex (contoured at 1.3σ) illustrating the role of N202 in selecting TDP-sugars. The N202 side chain NH_2 group forms a hydrogen bond with the TDP 3′-OH group and also prevents a ribose-containing NDP-sugar from binding because of clashes with the 2′-OH group. (B) Stereodiagram of the $2F_o - F_c$ electron density maps of the SpnG−TDP complex (contoured at 1.3σ) showing the SpnG active site. H13 may deprotonate the 9-OH group of aglycone substrates **3** and **4** during catalysis. D316 and Q317 form the D/E-Q motif that in other GTs helps bind the sugar portion of the donor substrate. (C) Sequence alignment of GTs. An asterisk indicates the position of the general base (H13 in SpnG). The pyrophosphate-binding loop (G293−G296 in SpnG) and D/E-Q motif (D316 and Q317 in SpnG) help bind NDP-sugars. The arrow points to the residue that in UDP-sugar-utilizing GTs is a glutamate that directly interacts with the 2′-OH and 3′-OH groups of ribose but in SpnG is a threonine (T300) that coordinates N202 to help select for NDP-sugars containing a deoxyribose: SpnG (spinosyn rhamnosyltransferase, AAG23268.1, *S. spinosa*), CalG1 (calicheamicin 3-*O*-methyl-rhamnosyltransferase, AAM70336.1, *Micromonospora purpurea*), CalG3 (calicheamicin glucosyltransferase, AAM70351.1, *M. purpurea*), OleD (oleandomycin glucosyltransferase, CAA80301.1, *Streptomyces antibioticus*), OleI (oleandomycin glucosyltransferase, ABA42118.2, *St. antibioticus*), GtfB (vancomycin glucosyltransferase, AAB49293.1, *Amycolptopsis orientalis*), and GtfD (vancomycin vancosaminyltransferase, AAK31352.1, *A. orientalis*). (D) Stereodiagram showing a SpnG active site in the absence of TDP (blue) and in the presence of TDP (gray). The TDP α-phosphate stabilizes one conformation of the pyrophosphate-binding loop (G293−G296), causing several residues to shift, notably F315, which contacts the acceptor substrate.

of the Supporting Information). Samples were analyzed on a C_{18} column (Varian) using the following method: linear gradient from 1 to 36% B over 1 min, from 36 to 75% B over 5 min, 75% B for 3 min, from 75 to 1% B over 5 min, and 1% B for 4 min (solvent A, water; solvent B, acetonitrile; rate of 1 mL/min). Each reaction was performed and analyzed in duplicate. The concentration of product at each time point, $[P]_t$, was calculated using the equation $[P]_t = [A_P/(A_S + A_P)][S]_0$, where $[S]_0$ is the initial substrate concentration and A_S and A_P are the areas of the substrate and product peaks in the HPLC trace, respectively. Formation of the glucosylated product was confirmed by mass spectrometry. Mutants V94M, N202D, and F315W were assayed under the same conditions as wild-type SpnG.

RESULTS

Crystallization and Structure Determination. Crystals of SpnG were readily obtained from sparse matrix screens; however, the generated crystals possessed high mosaicity and were not suitable for structure determination. Through an additive screen, the presence of D-glucose was found to significantly decrease mosaicity. The apo structure was determined by molecular replacement using the dimeric urdamycin GT UrdGT2 (PDB entry 2P6P) as the search model and was refined to 1.65 Å resolution.[19,20] One SpnG dimer crystallized in the asymmetric unit within space group *P*1. Crystals were soaked with TDP or AGL to yield the structures of the SpnG−TDP and SpnG−AGL complexes at 1.86 and 1.70 Å resolution, respectively. Despite our efforts, a ternary complex could not be obtained.

Overall Structure of SpnG. SpnG forms a C_2-symmetric homodimer, an architecture that has been observed previously in the GT1 enzymes UrdGT2 and CalG3[27,28] (Figure 2). The 1270 Å^2 interface is largely comprised of hydrophobic interactions with a major contact created by the stacking of two W22 residues, analogous to a dimeric contact made in CalG3.[28] Each SpnG monomer contains two domains, both possessing a Rossmann-like fold. The N-terminal domain (residues 1−178) consists of a six-stranded, parallel β-sheet surrounded by six α-helices; the C-terminal domain (residues 218−371) contains a six-stranded, parallel β-sheet surrounded by nine α-helices. A long loop (residues 183−209) connects the two domains. Residues 56−67 are invisible in the electron density maps. By analogy with other GT1 enzymes, SpnG is predicted to bind the donor substrate (TDP-L-rhamnose) in the C-terminal domain and the acceptor substrate (**3** or **4**) in the N-terminal domain. In each of the structures, D-glucose is bound on the surface, forming hydrogen bonds with several residues, including R78 and D319. A search for structural homologues of the SpnG monomer through the DALI server identified GT1 enzymes that showed a high degree of structural similarity despite their low level of sequence identity: UrdGT2 (*Z* score of 42.4, rmsd of 2.2 Å, identity of 28%, PDB entry 2P6P), CalG1 (*Z* score of 39.5, rmsd of 3.0 Å, identity of 26%, PDB entry 3OTH), CalG3 (*Z* score of 39.1, rmsd of 2.6 Å,

identity of 28%, PDB entry 3OTI), and OleI (Z score of 34.4, rmsd of 2.9 Å, identity of 20%, PDB entry 2IYA).[27,29−31]

The active site is located between the two domains in a deep cleft (Figure 3). Two residues known to be critical for SpnG activity, H13 and D316, are in this region (Figure 4).[17] Most other GT1 enzymes possess a histidine residue equivalent to H13; however, in each of the vancomycin GTs (GtfA, GtfB, and GtfD), an aspartate replaces the histidine.[32−34] D316 helps comprise a two-residue motif (D/E-Q) that mediates the binding of the sugar moiety in related GTs.[36]

Within the apo-SpnG dimer, many of the active site residues are differently positioned in the two monomers, yet in the SpnG−TDP complex, the TDP pyrophosphate interacts with residues G293−G296 to stabilize one of these active site configurations (Figure 4D). The equivalent loops within the UrdGT2 dimer also exist in two different conformations, and the equivalent loops in the CalG3 dimer (comprised of four glycine residues) are known to become reordered upon substrate binding (Figure 4C).[27−29] In the stabilized conformation, the S295 side chain forms a hydrogen bond with the backbone carbonyl of W142, which may influence the rotomeric conformations of F315 and D316. The conformation observed for F315 in the SpnG−TDP complex is also adopted in the SpnG−AGL complex; it allows the phenyl ring to make significant contact with AGL (Figure 3B).

TDP Binding Site. The SpnG−TDP complex structure reveals that the TDP portion of TDP-L-rhamnose binds to the C-terminal domain of SpnG with its diphosphate pointing toward the active site. The binding site is largely formed by an $\alpha-\beta-\alpha$ motif as in related GT-B enzymes; the pocket is created by P11, T12, M15, Y201, N202, C223−V228, A255−P257, S276−L279, and S295−T297 (Figure 3A). Interactions between TDP and SpnG include the thymine O2 hydrogen bonding with the L279 main chain NH group, the thymine N3 group hydrogen bonding to the V277 main chain carbonyl, the thymine O4 hydrogen bonding to the V277 main chain NH group, the thymine methyl making hydrophobic contact with the P257 side chain, and the thymine ring making hydrophobic contact with L279 (similar interactions occur with L280 in the GtfA−TDP complex and with F296 in the GtfD−TDP complex).[33,34] In GtfA, a salt bridge formed by an arginine and a glutamate caps the thymine ring; an equivalent interaction is not present in SpnG.[33]

Greater contact is made between the TDP deoxyribose in SpnG than in other TDP-sugar-utilizing GTs (e.g., GtfA, PDB entry 1PN3; GtfD, PDB entry 1RRV); specifically, the 3′-OH group accepts a hydrogen bond from the side chain NH_2 group of N202.[33,34] The loop containing N202 is much closer to the donor substrate-binding site in SpnG than in related TDP-sugar-utilizing GTs (PDB entries 1PN3 and 1RRV)[33,34] (Figure 4A). A glutamate in UDP-sugar-utilizing GTs typically forms charged hydrogen bonds with the 2′-OH and 3′-OH groups of UDP ribose,[30] while an aliphatic residue replaces this glutamate in TDP-sugar-utilizing GTs (Figure 3B). Mutation of this residue in SpnG (T300) to aspartate or valine is known to abolish activity toward both TDP- and UDP-sugars.[17] In the SpnG−TDP complex, the α-phosphate forms charged hydrogen bonds with the main chain NH groups of G296 and T297 as well as the T297 hydroxyl, while the β-phosphate makes charged hydrogen bonds with the main chain NH groups of M227 and V228 (Figure 3A). The charged hydrogen bond to the T297 hydroxyl may contribute significantly to binding because replacing it with alanine is known to abolish activity.[17] In most other GT-B enzymes, the phosphoryl groups make contact with a histidine from a neighboring "H-X_3-G-T loop"; however, in SpnG, this histidine is replaced with A292, the equivalent SpnG loop being formed by residues 292−297.[34]

Acceptor Binding Site. AGL, a mimic of the natural substrate, was soaked into crystals of SpnG. The resulting density, though weak, was sufficient to model AGL into the acceptor binding site of one SpnG monomer.[21] In a manner similar to that for the interactions observed between oleandomycin and OleI, AGL interacts with SpnG predominantly through hydrophobic contacts (with L8, Y10, H13, T90, N93, V94, W142, and F315) (Figure 3B).[30] The 12-membered ring of AGL is located near a glycine- and threonine-rich loop (residues 56−67) that is invisible in the electron density maps but positioned to contact the 12-membered ring through an induced fit motion. The rigid 5-6-5-*cis*-*anti*-*trans*-tricyclic ring system of AGL is sandwiched between L8 and F315. While in the apo structure the F315 side chain possesses high B factors, in the SpnG−AGL complex it is ordered as a single rotamer that allows complementary hydrophobic interactions with the acceptor substrate. The 9-OH group of AGL and Nε2 of H13 are separated by <3 Å.

Structures of several GTs bound to their acceptor substrates are known: OleI−UDP−oleandomycin (PDB entry 2IYA), OleD−UDP−erythromycin (PDB entry 2IYF), GtfD−TDP−desvancosaminyl vancomycin (PDB entry 1RRV), GtfA−TDP−vancomycin (PDB entry 1PN3), vvGT1−UDP-2F-glucose−UDP-2F (PDB entry 2C1Z), CalG1−calicheamicin−TDP (PDB entry 3OTH), CalG2−TDP−calicheamicin (PDB entry 3RSC), and CalG3−calicheamicin−TDP (PDB entry 3OTI)[29,30,33,34,36] (Glyco3D). The residues between the third and fourth β-strands of the N-terminal domain comprise a "specificity loop".[30] While the specificity loop of SpnG is invisible in the electron density maps of the SpnG−AGL complex, the equivalent loops from OleI, OleD, GtfA, GtfB, and VvGT1 are well-ordered, and most make contact with their respective substrate. Residues 80−100 of SpnG form a helix of six turns, longer than those in other known GTs (another long helix is present in UrdGT2, PDB entry 2P6P).[27] AGL is sandwiched between L8 and F315 in the SpnG−AGL complex and makes several other hydrophobic contacts. In the AGL- and TDP-liganded structures, F315 adopts a rotomeric conformation that helps form the acceptor substrate-binding site.

Activity Assays of Mutants. The SpnG mutants V94M, N202D, and F315W were expressed and purified, whereas the SpnG mutants Y10F, F315G, and F315A were insoluble (Figure 1 of the Supporting Information). The glycosyltransferase activities of V94M, N202D, and F315W were compared to that of wild-type SpnG in an assay that follows the transfer of glucose from TDP-D-glucose to AGL. V94M and F315W exhibited decreased k_{cat} values (37 and 94% decreases, respectively), while N202D was inactive. The glucosylated product was confirmed by mass spectrometry (M + H^+, observed mass of 565.0, calculated mass of 565.7).

DISCUSSION

SpnG catalyzes the first glycosylation in the synthesis of the insecticide spinosad through the rhamnosylation of tricyclic aglycones **3** and **4** (Figure 1). Because variations of the rhamnose group have been observed to generate insecticides with improved activities, SpnG may serve as a valuable biocatalyst.[11] SpnG naturally exhibits relaxed substrate

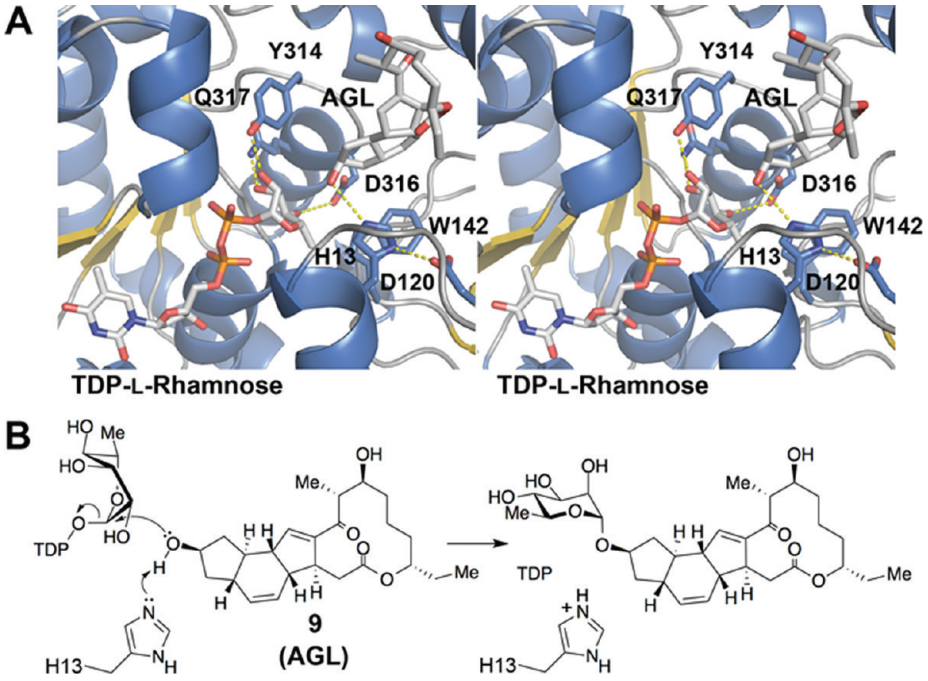


Figure 5. Putative mechanism of SpnG. (A) TDP-L-rhamnose was modeled into the structure of the SpnG−AGL complex. With L-rhamnose in the skew-boat conformation, the equatorial hydroxyl groups form hydrogen bonds with Y314, D316, and Q317, and the equatorial methyl substituent makes a hydrophobic contact with W142. (B) Consistent with related GTs, the proposed nucleophilic displacement mechanism shows H13 activating the 9-OH group for attack at the anomeric carbon of TDP-L-rhamnose.

specificity, accepting diverse TDP-sugars as donor substrates as well as various aglycones as acceptor substrates.[17,35] The fact that SpnG can glycosylate the tetracyclic AGL is convenient for assaying SpnG glycosyltransferase activity because AGL is more readily available than **3** or **4**.[17]

To test which residues are most important in acceptor substrate binding, Y10F, V94M, F315G, F315A, and F315W mutants were constructed; V94M and F315W were tested for activity (Y10F, F315G, and F315A did not produce soluble protein) (Figure 1 of the Supporting Information). V94M and F315W exhibited 37 and 94% reduced catalytic activity, respectively. The fact that F315 is quite sensitive to mutation could be predicted from the close contact it makes with AGL in the SpnG−AGL complex (Figure 3B).

The structure of the VvGT1 nonproductive Michaelis complex revealed the position of the hydroxyl nucleophile prior to the displacement reaction.[36] A histidine, positioned 2.7 Å from this kaempferol hydroxyl, is hypothesized to be the catalytic base, activated by a neighboring aspartate. The SpnG−AGL complex shows the 9-OH group of AGL to be in a position equivalent to that of the kaempferol hydroxyl, <3 Å from H13, which is activated by D120. The 9-OH group of AGL is also in the same position as the glycosidic oxygens of oleandomycin and erythromycin in the OleI and OleD ternary complexes, respectively.[30]

Many UDP-sugar-utilizing enzymes are capable of utilizing TDP-sugars, and likewise, many TDP-sugar-utilizing enzymes are capable of utilizing UDP-sugars; however, SpnG selects against UDP-sugars.[17] Compared to TDP-sugars, UDP-sugars lack a C5 methyl group on the base and possess a 2′-OH group on the ribose. One or both of these features must be selected against, yet the selectivity of SpnG against UDP-sugars has been challenging to explain through previous structures of NDP-bound GTs because the structures of TDP-sugar-utilizing GTs complexed with TDP show little contact between the donor substrate binding pocket and the deoxyribose (e.g., PDB entries 1PN3 and 1RRV).[33,34] UDP-sugar-utilizing GTs directly contact the 2′-OH and 3′-OH groups of a UDP-sugar substrate through a glutamate; the equivalent residue in TDP-sugar-utilizing GTs is small and aliphatic (T300 in SpnG) and would not be expected to prevent the binding of UDP-sugars[37] (Figure 4). The TDP-L-rhamnose binding site of SpnG differs from the donor substrate-binding sites of other determined TDP-sugar-utilizing GTs in that part of the linker that connects the N- and C-terminal domains of SpnG is shifted toward the donor substrate-binding site, allowing linker residue N202 to help form the TDP-L-rhamnose binding site (Figure 4A).[25,31] T300 forms a hydrogen bond with the NH_2 group of N202, orienting it toward the TDP-L-rhamnose binding site. The SpnG−TDP complex structure shows that the N202 side chain directly forms a hydrogen bond with the 3′-OH group of deoxyribose. Modeling UDP in place of TDP reveals that the 2′-OH group of UDP sterically clashes with the NH_2 group of N202. This clash may explain the >60000-fold reduction in k_{cat} for UDP-D-glucose relative to that of TDP-D-glucose.[17] From our observation that the N202D mutant is catalytically inactive, the NH_2 group of N202 likely donates a hydrogen bond to the 3′-OH group of deoxyribose.

The L-rhamnose moiety of TDP-L-rhamnose must dock in manner similar to that of the D-glucose moiety in the flavonoid UDP-glucosyltransferase VvGT1 nonproductive Michaelis complex.[36] While the lowest-energy conformation of TDP-L-rhamnose in solution is with TDP in an equatorial position, within the Michaelis complex of SpnG the TDP substituent must be in an axial position to permit the nucleophilic displacement reaction.[36] However, a ring flip is not possible because the TDP, the 3′-OH group, and the 5′-methyl substituents would sterically clash in axial positions; thus, L-

dx.doi.org/10.1021/bi201860q | *Biochemistry* 2012, 51, 1213−1222

rhamnose likely adopts a skew-boat conformation. TDP-L-rhamnose was modeled into the SpnG−AGL complex structure on the basis of the position of TDP in the SpnG−TDP complex (Figure 5). The modeled L-rhamnose residue was guided by the VvGT1 nonproductive Michaelis complex. The D/E-Q motif of VvGT1 interacts with the equatorial D-glucose hydroxyls; the aspartate makes charged hydrogen bonds with the 3′-OH and 4′-OH groups, while the glutamine forms hydrogen bonds with the 2′-OH and 3′-OH groups. D316 and Q317 from SpnG are in position to form equivalent hydrogen bonds with the rhamnosyl 4′-OH and 3′-OH groups, respectively, and Y314 is in position to form a hydrogen bond with the 2′-OH group. While a VvGT1 threonine forms a hydrogen bond with the 6′-OH group of D-glucose, W142 of SpnG is in position to make hydrophobic contact with the equivalent L-rhamnosyl 5′-methyl group. The fact that several TDP-sugars (e.g., TDP-L-rhamnose, TDP-D-glucose, TDP-6-deoxy-D-glucose, TDP-L-olivose, and TDP-D-xylose) are transferred by SpnG demonstrates that precise contacts are not required within the sugar binding site for catalysis; however, the geometry of nucleophilic attack at the anomeric carbon may not be optimal for these alternate substrates.[17] The substrate promiscuity of SpnG may be enhanced through site-directed mutagenesis, so that transfer from NDP-sugars such as TDP-6-deoxy-L-talose, TDP-D-mannose, and TDP-D-galactose is possible.[17] Thus, structural insights into the selectivities of SpnG will provide valuable guidance in the combinatorial biosynthesis of novel glycosylated natural products.

■ ASSOCIATED CONTENT

Supporting Information

Glycosyltransferase assay of SpnG and its mutants (Figure 1). This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

Atomic coordinates for the apo form of SpnG and its binary complexes with TDP and AGL have been deposited in the Protein Data Bank as entries 3TSA, 3UYL, and 3UYK, respectively.

■ AUTHOR INFORMATION

Corresponding Author

*H.-w.L.: e-mail, h.w.liu@mail.utexas.edu; phone, (512) 232-7811. A.T.K.-C.: e-mail, adriankc@mail.utexas.edu; phone, (512) 471-2977.

Funding

This work is supported by grants provided by the National Institutes of Health (GM35906 to H.-w.L.) and the Robert A. Welch Foundation (F-1712 to A.T.K.-C. and F-1511 to H.-w.L.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Hak Joong Kim and Steven Mansoorabadi for discussions and advice. We also thank the Advanced Light Source for beamtime and Arthur Monzingo for his technical assistance in the Macromolecular Crystallography Facility at the University of Texas at Austin.

■ ABBREVIATIONS

AGL, aglycone; CAZy, Carbohydrate-Active-enZymes; GABA, γ-aminobutyric acid; GAGL, glucosylated aglycone; GT, glycosyltransferase; GT1, glycosyltransferase family 1; HPLC, high-performance liquid chromatography; NDP, nucleotide diphosphate; PCR, polymerase chain reaction; PEP, phosphoenoylpyruvate; PKS, polyketide synthase; PSA, 17-pseudoaglycone; TDP, thymidine diphosphate; TDPG, thymidine diphosphate-D-glucose; UDP, uridine diphosphate.

■ REFERENCES

(1) Kirst, H. (1991) A83543A-D, unique fermentation-derived tetracyclic macrolides. *Tetrahedron Lett. 32*, 4839−4842.

(2) Sparks, T., Crouse, G., and Durst, G. (2001) Natural products as insecticides: The biology, biochemistry and quantitative structure-activity relationships of spinosyns and spinosoids. *Pest Manag. Sci. 57*, 896−905.

(3) Huang, K.-x., Xia, L., Zhang, Y., Ding, X., and Zahn, J. (2009) Recent advances in the biochemistry of spinosyns. *Appl. Microbiol. Biotechnol. 82*, 13−23.

(4) Waldron, C., Madduri, K., Crawford, K., Merlo, D. J., Treadway, P., Broughton, M. C., and Baltz, R. H. (2000) A cluster of genes for the biosynthesis of spinosyns, novel macrolide insect control agents produced by *Saccharopolyspora spinosa*. *Antonie van Leeuwenhoek 78*, 385−390.

(5) Waldron, C., Matsushima, P., Rosteck, P. R., Broughton, M. C., Turner, J., Madduri, K., Crawford, K. P., Merlo, D. J., and Baltz, R. H. (2001) Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*. *Chem. Biol. 8*, 487−499.

(6) Kim, H. J., Pongdee, R., Wu, Q., Hong, L., and Liu, H.-w. (2007) The biosynthesis of spinosyn in *Saccharopolyspora spinosa*: Synthesis of the cross-bridging precursor and identification of the function of SpnJ. *J. Am. Chem. Soc. 129*, 14582−14584.

(7) Kim, H., Ruszczycky, M., Choi, S.-H., Liu, Y.-N., and Liu, H.-w. (2011) Enzyme-catalysed [4 + 2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature 473*, 109−112.

(8) Hong, L., Zhao, Z., and Liu, H.-w. (2006) Characterization SpnQ from the spinosyn biosynthetic pathway of *Saccharopolyspora spinosa*: Mechanistic and evolutionary implications for C-3 deoxygenation in deoxysugar biosynthesis. *J. Am. Chem. Soc. 128*, 14262−14263.

(9) Hong, L., Zhao, Z., Melançon, C. E. III, Zhang, H., and Liu, H.-w. (2008) *In vivo* characterization of the enzymes involved in TDP-D-forosamine biosynthesis in the spinosyn pathway of *Saccharopolyspora spinosa*. *J. Am. Chem. Soc. 130*, 4954−4967.

(10) Kim, H. J., White-Phillip, J., Ogasawara, Y., Shin, N., Isiorho, E., and Liu, H.-w. (2010) Biosynthesis of spinosyn in *Saccharopolyspora spinosa*: Synthesis of permethylated rhamnose and characterization of the functions of SpnH, SpnI, and SpnK. *J. Am. Chem. Soc. 132*, 2901−2903.

(11) Creemer, L., Kirst, H., Paschal, J., and Worden, T. (2000) Synthesis and insecticidal activity of spinosyn analogs functionally altered at the 2′-, 3′- and 4′-positions of the rhamnose moiety. *J. Antibiot. 53*, 171−178.

(12) Sparks, T., Thompson, G., Kirst, H., Hertlein, M., Larson, L., Worden, T., and Thibault, S. (1998) Biological activity of the spinosyns, new fermentation derived insect control agents, on tobacco budworm (Lepidopter: Noctuidae) larvae. *J. Econ. Entomol. 91*, 1277−1283.

(13) Lairson, L., Henrissat, B., Davies, G., and Withers, S. (2008) Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem. 77*, 521−555.

(14) Cantarel, B., Coutinho, P., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res. 37*, D233−D238.

(15) Coutinho, P., Deleury, E., Davies, G. J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol. 328*, 307−317.
(16) Breton, C., Snajdrova, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology 16*, 29R−37R.
(17) Chen, Y.-L., Chen, Y.-H., Lin, Y.-C., Tsai, K.-C., and Chiu, H.-T. (2009) Functional characterization and substrate specificity of spinosyn rhamnosyltransferase by *in vitro* reconstitution of spinosyn biosynthetic enzymes. *J. Biol. Chem. 284*, 7352−7363.
(18) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307−326.
(19) Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D59*, 1131−1137.
(20) McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L., and Read, R. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.
(21) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.
(22) Vagin, A., Steiner, R., Lebedev, A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. (2004) REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D60*, 2184−2195.
(23) Cohen, S., Morris, R., Fernandez, F., Jelloul, M., Kakaris, M., Parthasarathy, V., Lamzin, V., Kleywegt, G., and Perrakis, A. (2004) Towards complete validated models in the next generation of ARP/wARP. *Acta Crystallogr. D60*, 2222−2229.
(24) De Amicis, C. V., Graupner, P. R., Erickson, J. A., Paschal, J. W., Kirst, H. A., Creemer, L. C., and Fanwick, P. E. (2001) The stereochemical outcome of electrophilic addition reactions on the 5,6-double bond in the spinosyns. *J. Org. Chem. 66*, 8431−8435.
(25) Creemer, L., Kirst, H., and Paschal, J. (1998) Conversion of spinosyn A and spinosyn D to their respective 9- and 17-pseudoaglycones and their aglycones. *J. Antibiot. 51*, 795−800.
(26) Borisova, S., Zhang, C., Takahashi, H., Zhang, H., Wong, A., Thorson, J., and Liu, H.-w. (2006) Substrate specificity of the macrolide-glycosylating enzyme pair DesVII/DesVIII: Opportunities, limitations, and mechanistic hypotheses. *Angew. Chem., Int. Ed. 45*, 2748−2753.
(27) Mittler, M., Bechthold, A., and Schulz, G. E. (2007) Structure and action of the C-C bond-forming glycosyltransferase UrdGT2 involved in the biosynthesis of the antibiotic urdamycin. *J. Mol. Biol. 372*, 67−76.
(28) Zhang, C., Bitto, E., Goff, R., Singh, S., Bingman, C., Griffith, B., Albermann, C., Phillips, G., and Thorson, J. (2008) Biochemical and structural insights of the early glycosylation steps in calicheamicin biosynthesis. *Chem. Biol. 15*, 842−853.
(29) Chang, A., Singh, S., Helmich, K. E., Goff, R. D., Bingman, C. A., Thorson, J. S., and Phillips, G. N. Jr. (2011) Complete set of glycosyltransferase structures in the calicheamicin biosynthetic pathway reveals the origin of regiospecificity. *Proc. Natl. Acad. Sci. U.S.A. 108*, 17649−17654.
(30) Bolam, D., Roberts, S., Proctor, M., Turkenburg, J., Dodson, E., Martinez-Fleites, C., Yang, M., Davis, B., Davies, G., and Gilbert, H. (2007) The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. *Proc. Natl. Acad. Sci. U.S.A. 204*, 5336−5341.
(31) Holm, L., and Rosentrom, P. (2010) Dali server: Conservation mapping in 3D. *Nucleic Acids Res. 38*, W545−W549.
(32) Mulichak, A., Losey, H., Walsh, C., and Garavito, R. (2001) Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure 9*, 547−557.
(33) Mulichak, A., Losey, H. C., Lu, W., Wawrzak, Z., Walsh, C., and Garavito, R. (2003) Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc. Natl. Acad. Sci. U.S.A. 100*, 9238−9243.
(34) Mulichak, A., Lu, W., Losey, H., Walsh, C., and Garavito, R. (2004) Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: Interactions with acceptor and nucleotide ligands. *Biochemistry 43*, 5170−5180.
(35) Hu, Y., and Walker, S. (2002) Remarkable structural similarities between diverse glycosyltransferases. *Chem. Biol. 9*, 1287−1296.
(36) Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B., Tarling, C., Ford, C., Bowles, D., and Davies, G. (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J. 25*, 1396−1405.
(37) 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.