Published in final edited form as:

Biochemistry. 2012 February 14; 51(6): 1213–1222. doi:10.1021/bi201860q.

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, The University of Texas at Austin, Austin, Texas 78712, USA

[‡]Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712, USA

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 >60,000-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 through the binding of TDP. The SpnG-AGL complex shows that the binding of AGL is mediated through hydrophobic interactions and that H13, the potential catalytic base, is within 3 Å of the nucleophilic AGL 9-OH. A model for the Michaelis complex was constructed to reveal the features that enable SpnG to transfer diverse sugars; it also revealed that the rhamnosyl moiety is in a skew-boat conformation during the transfer reaction.

Keywords

Glycosyltransferase; spinosyn; tailoring enzyme; enzyme catalysis; enzyme mechanism

Spinosyns A and D (**7** and **8**; 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 ten 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

Dr. Liu – h.w.liu@mail.utexas.edu, 512-232-7811. Dr. Keatinge-Clay – adriankc@mail.utexas.edu, 512-471-2977.

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.

SUPPORTING INFORMATION AVAILABLE

Supplementary Figure 1 shows the glycosyltransferase assay of SpnG and its mutants. This material is available free of charge via the Internet at <http://pubs.acs.org>

sugar appendages to the C9 and C17 positions of the aglycone is catalyzed by the rhamnosyl- and forosaminytransferases, SpnG and SpnP, respectively.

Based on *in vitro* studies, the pathway for spinosyn tailoring has been established (7) with oxidation of 15-OH 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 enable a [4+2] intramolecular cycloaddition yielding **3** and **4** that is catalyzed by SpnF. Subsequent 9-OH-rhamnosylation by SpnG provides **5** and **6**. 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 17-OH. The glycosyltransfer reactions catalyzed by SpnG and SpnP are significant since 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 sequence identities; 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 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 >60,000-fold decrease in k_{cat} (Supplementary Figure 1) (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. 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, 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, Purification of SpnG

The *spnG* gene was amplified from *S. spinosa* NRLL18537 genomic DNA by polymerase chain reaction (PCR) with 5'-AAACATATGCGCGTACTCGTCGTTCCCTTGCCCTATCCGACGCATCTCAT-3' and 5'-ATTAAGCTTCAGGCACGGATGGCCGCAGTGTTCTCCAGTGT-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 NCBI 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 LB with 50 µg/mL kanamycin) that were used to inoculate 6 × 1 L LB (50 mg/L kanamycin). The cultures were incubated with shaking at 37 °C until OD₆₀₀ = 0.6, at which point isopropyl-β-D-thiogalactopyranoside (0.5 mM final concentration) 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, 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 (30,000 ×g, 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, 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 were obtained per liter culture.

Crystallization, Data Processing, and Structure Determination

Crystals grew overnight through the sitting-drop vapor-diffusion method. The crystallization buffer was 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 immediately prior to data collection. SpnG-TDP crystals were obtained by supplying 20 mM TDP to the crystallization drop and incubating 1 hour prior to flash-freezing. SpnG-AGL crystals were obtained by adding 40 mM AGL to the crystallization drop and incubating 1 hour 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 at the Advanced Light Source Beamline 5.0.1, and subsequently processed with HKL2000 (18). The apo structure of SpnG was solved by molecular replacement with the UrdGT2 dimer (PDB code: 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 with the program Coot and Refmac (21–23). The SpnG-TDP and SpnG-AGL structures were solved 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 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'-GGAGCAGACCGCGTCCAATATGGCGCAAAGCTCG-3' and for 5'-CGAGCTTTGCGCCATATTGGACGCGGTCTGCTCC-3' V94M; 5'-CCGGTCCAGTACGTGCCGTACGACGGAAGCGGCGC-3' and 5'-GCGCCGCTTCCGTCGTACGGCAGTACTGGACCGG-3' for N202D; 5'-GTGCTTCCCCAGTACGCCGACCAGTTCGACTAC-3' and 5'-GTAGTCGAAGTGGTCGGCGTACTGGGGAAGCAC-3' for F315A; 5'-GTGCTTCCCCAGTACTGGGACCAGTTCGACTAC-3' and 5'-GTAGTCGAAGTGGTCCCAGTACTGGGGAAGCAC-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 methanol/7.2 N H₂SO₄ (1:1.5) and refluxed for 3 h. The reaction was cooled to room temperature and quenched through the addition of NaHCO₃(s) before the solution was extracted with diethyl ether (×3). The combined organic layers were washed with brine (×2), dried over anhydrous MgSO₄, and concentrated under vacuum. AGL was then purified by flash chromatography using a methanol gradient (0–7.5% methanol:dichloromethane).

Enzymatic Preparation of TDP-^D-Glucose

Preparation of TDP-^D-glucose was achieved through a reported procedure (26). Specifically, in a 15 mL reaction with 50 mM Tris, pH 7.5, 78 mM phosphoenolpyruvate, 24 mM thymidine, 1.6 mM ATP, and 27 mM MgCl₂ were incubated with 1000 U of pyruvate kinase (1000 U), 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 reaction. The filtrate was incubated with ^D-glucose-1-phosphate (0.51 mmol), MgCl₂ (0.45 mmol) and 41 mM α-^D-glucose-1-phosphate thymidyltransferase (RfbA) for 16 h at 37 °C. The reaction was centrifuged (5000 ×g, 10 min.), and RfbA was removed through filtration. After the flowthrough was frozen and lyophilized, the product residue was passed through a P2 size-exclusion column (Bio-Rad) using water as the eluent. 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 reactions consisted of 2.5 mM AGL, 0.5% (v/v) DMSO, 2.5 mM MgCl₂, 12 mM TDP-^D-glucose, 25 μM SpnG, and 50 mM HEPES, pH 8.5 (100 μL final volume) and 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 equipped with a photodiode array detector (Waters) (Supplementary Figure 1). Samples were analyzed on a C₁₈ column (Varian) using the following method: linear gradient from 1–36% B over 1 min., 36–75% B over 5 min., 75% B for 3 min., 75–1% B over 5 min., and 1% B for 4 min. (solvent A, water; solvent B, acetonitrile; 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 following 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 solved by molecular replacement using the dimeric urdamycin GT UrdGT2 (PDB code: 2P6P) as the search model and was refined to 1.65 Å resolution (19–20). One SpnG dimer crystallized in the asymmetric unit within spacegroup P1. Crystals were soaked with TDP or AGL to yield the SpnG-TDP and SpnG-AGL complex structures at 1.86 Å and 1.70 Å resolution, respectively. Despite our efforts, a ternary complex could not be obtained.

Overall Structure of SpnG

SpnG forms a C₂-symmetric homodimer, an architecture that has been observed previously in the GT1 enzymes UrdGT2 and CalG3 (27–28) (Figure 2). The 1270 Å² 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, hydrogen-bonding with several residues including R78 and D319. A search for structural homologs of the SpnG monomer through the DALI server identified GT1 enzymes that showed high structural similarity despite their low sequence identity: UrdGT2 (Z-score: 42.4, r.m.s.d.: 2.2 Å, identity: 28%, PDB code: 2P6P), CalG1 (Z-score: 39.5, r.m.s.d.: 3.0 Å, identity: 26%, PDB code: 3OTH), CalG3 (Z-score: 39.1, r.m.s.d.: 2.6 Å, identity: 28%, PDB code: 3OTI), OleI (Z-score: 34.4, r.m.s.d.: 2.9 Å, identity: 20%, PDB code: 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, 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 reorder 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 rotameric conformations of F315 and D316. The conformation observed for F315 in the SpnG-TDP complex is also adopted in the SpnG-AGL complex; it enables 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 towards the active site. The binding site is largely formed by an α - β - α 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, the thymine N3 hydrogen-bonding to the V277 main chain carbonyl, the thymine O4 hydrogen-bonding to the V277 main chain NH, 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 structure and with F296 in the GtfD-TDP complex structure) (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 code: 1PN3; GtfD, PDB code 1RRV); specifically, the 3'-OH accepts a hydrogen bond from the side chain NH₂ 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 codes: 1PN3, 1RRV) (33–34) (Figure 4A). A glutamate in UDP-sugar-utilizing GTs typically forms charged hydrogen bonds with the UDP ribose 2'-OH and 3'-OH (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 towards both TDP- and UDP-sugars (17). In the SpnG-TDP complex structure, the α -phosphate forms charged hydrogen bonds with the main chain NHs of G296 and T297 as well as the T297 hydroxyl, while the β -phosphate makes charged hydrogen bonds with the main chain NHs of M227 and V228 (Figure 3A). The charged hydrogen bond to the T297 hydroxyl may contribute significantly to binding since 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-X₃-G-T loop”; however, in SpnG, this histidine is replaced by 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). Similar to 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 sidechain possesses high B-factors, in the SpnG-AGL complex it is ordered as a single rotamer that enables complementary hydrophobic interactions with the acceptor substrate. The AGL 9-OH and H13 N ϵ 2 are separated by less than 3 Å.

Structures of several GTs bound to their acceptor substrates are known: OleI+UDP+oleandomycin (PDB code: 2IYA), OleD+UDP+erythromycin (PDB code: 2IYF), GtfD+TDP+desvancosaminyl vancomycin (PDB code: 1RRV), GtfA+TDP+vancomycin (PDB code: 1PN3), vvGT1+UDP-2F-Glucose+UDP-2F (PDB code: 2C1Z), CalG1+calicheamicin+TDP (PDB code: 3OTH), CalG2+TDP+calicheamicin (PDB code: 3RSC), CalG3+calicheamicin+TDP (PDB code: 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 in other known GTs (another long helix is present in UrdGT2, PDB code: 2P6P) (27). AGL is sandwiched between L8 and F315 in the SpnG-AGL structure and makes several other hydrophobic contacts. In the AGL- as well as the TDP-liganded structures, F315 adopts a rotameric 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 (Supplementary Figure 1). The glycosyltransferase activity 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 showed 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, 565.0; calculated mass, 565.7).

DISCUSSION

SpnG catalyzes the first glycosylation in the synthesis of the insecticide spinosad through the rhamnosylation of the tricyclic aglycones **3** and **4** (Figure 1). Since 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 specificity, accepting diverse TDP-sugars as donor substrates as well as various aglycones as acceptor substrates (17, 35). That SpnG can glycosylate the tetracyclic AGL is convenient for assaying SpnG glycosyltransferase activity since 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) (Supplementary Figure 1). V94M and F315W showed 37% and 94% reduced catalytic activity, respectively. That F315 is quite sensitive to mutation could be predicted from the close contact it makes with AGL in the SpnG-AGL complex structure (Figure 3B).

The structure of the VvGT1 non-productive 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 AGL 9-OH to be in an equivalent position as the kaempferol hydroxyl, less than 3 Å from H13, which is activated by D120. AGL 9-OH 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 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 codes 1PN3 and 1RRV) (33–34). UDP-sugar-

utilizing GTs directly contact the 2'-OH and 3'-OH 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 solved TDP-sugar-utilizing GTs in that part of the linker that connects the N- and C-terminal domains of SpnG is shifted towards the donor substrate-binding site, enabling linker residue N202 to help form the TDP-L-rhamnose binding site (Figure 4A) (25, 31). T300 forms a hydrogen bond with N202 NH₂, orienting it towards the TDP-L-rhamnose binding site. The SpnG-TDP complex structure shows that the N202 side chain directly forms a hydrogen bond with the deoxyribose 3'-OH. Modeling UDP in place of TDP reveals that the UDP 2'-OH sterically clashes with N202 NH₂. This clash may explain the >60,000-fold reduction in k_{cat} for UDP-D-glucose relative to TDP-D-glucose (17). From our observation that the N202D mutant is catalytically inactive, N202 NH₂ likely donates a hydrogen bond to deoxyribose 3'-OH.

The L-rhamnose moiety of TDP-L-rhamnose must dock similar to the D-glucose moiety in the flavonoid UDP-glucosyltransferase VvGT1 non-productive 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 enable the nucleophilic displacement reaction (36). However, a ring-flip is not possible since the TDP, the 3'-OH, and the 5'-methyl substituents would sterically clash in axial positions; thus, L-rhamnose likely adopts a skew-boat conformation. TDP-L-rhamnose was modeled into the SpnG-AGL complex structure based on the position of TDP in the SpnG-TDP complex structure (Figure 5). The modeled L-rhamnose residue was guided by the VvGT1 non-productive Michaelis complex. The D/E-Q motif of VvGT1 interacts with the equatorial D-glucose hydroxyls - the aspartate makes charged hydrogen bonds with 3'-OH and 4'-OH while the glutamine forms a hydrogen bonds with the 2'-OH and 3'-OH. 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. While a VvGT1 threonine forms a hydrogen bond with the D-glucose 6'-OH, W142 of SpnG is in position to make hydrophobic contact with the equivalent L-rhamnosyl 5'-methyl group. That several TDP-sugars (e.g. TDP-L-rhamnose, TDP-D-glucose, TDP-6-deoxy-D-glucose, TDP-L-olivose, 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 the valuable guidance in the combinatorial biosynthesis of novel glycosylated natural products.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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

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. and F-1511 to H.-w.L.).

Abbreviations

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

REFERENCES

1. Kirst H. A83543A-D, unique fermentation-derived tetracyclic macrolides. *Tetrahedron Lett.* 1991; 32:4839–4842.
2. Sparks T, Crouse G, Durst G. Natural products as insecticides: the biology, biochemistry and quantitative structure-activity relationships of spinosyns and spinosoids. *Pest Manag. Sci.* 2001; 57:896–905. [PubMed: 11695182]
3. Huang, K-xue; Xia, L.; Zhang, Y.; Ding, X.; Zahn, J. Recent advances in the biochemistry of spinosyns. *Appl. Microbiol. Biotechnol.* 2009; 82:13–23. [PubMed: 19082588]
4. Waldron C, Madduri K, Crawford K, Merlo DJ, Treadway P, Broughton MC, Baltz RH. A cluster of genes for the biosynthesis of spinosyns, novel macrolide insect control agents produced by *Saccharopolyspora spinosa*. *Anton. Leeuw. Int. J. G.* 2000; 78:385–390.
5. Waldron C, Matsushima P, Rosteck PR, Broughton MC, Turner J, Madduri K, Crawford KP, Merlo DJ, Baltz RH. Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*. *Chem. Biol.* 2001; 8:487–499. [PubMed: 11358695]
6. Kim HJ, Pongdee R, Wu Q, Hong L, Liu H-w. The biosynthesis of spinosyn in *Saccharopolyspora spinosa*: synthesis of the cross-bridging precursor and identification of the function of SpnJ. *J. Am. Chem. Soc.* 2007; 129:14582–14584. [PubMed: 17985910]
7. Kim H, Ruszczycky M, Choi S-H, Liu Y-N, Liu H-w. Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature.* 2011; 473:109–112. [PubMed: 21544146]
8. Hong L, Zhao Z, Liu H-w. 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.* 2006; 128:14262–14263. [PubMed: 17076492]
9. Hong L, Zhao Z, Melançon CE III, Zhang H, Liu H-w. *In vivo* characterization of the enzymes involved in TDP-d-ferosamine biosynthesis in the spinosyn pathway of *Saccharopolyspora spinosa*. *J. Am. Chem. Soc.* 2008; 130:4954–4967. [PubMed: 18345667]
10. Kim HJ, White-Phillip J, Ogasawara Y, Shin N, Isiorho E, Liu H-w. Biosynthesis of spinosyn in *Saccharopolyspora spinosa*: synthesis of permethylated rhamnose and characterization of the

- functions of SpnH, SpnI, and SpnK. *J. Am. Chem. Soc.* 2010; 132:2901–2903. [PubMed: 20158237]
11. Creemer L, Kirst H, Paschal J, Worden T. Synthesis and insecticidal activity of spinosyn analogs functionally altered at the 2'-,3'- and 4'- positions of the rhamnose moiety. *J. Antibiot.* 2000; 53:171–178. [PubMed: 10805578]
 12. Sparks T, Thompson G, Kirst H, Hertlein M, Larson L, Worden T, Thibault S. Biological activity of the spinosyns, new fermentation derived insect control agents, on tobacco budworm (Lepidopter: Noctuidae) larvae. *J. Econ. Entomol.* 1998; 91:1277–1283.
 13. Lairson L, Henrissat B, Davies G, Withers S. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* 2008; 77:521–555. [PubMed: 18518825]
 14. Cantarel B, Coutinho P, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 2009; 37:D233–D238. [PubMed: 18838391]
 15. Coutinho P, Deleury E, Davies Gideon J, Henrissat B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* 2003; 328:307–317. [PubMed: 12691742]
 16. Breton C, Snajdrova L, Jeanneau C, Koca J, Imberty A. Structures and mechanisms of glycosyltransferases. *Glycobiology.* 2006; 16:29R–37R. [PubMed: 16049187]
 17. Chen Y-L, Chen Y-H, Lin Y-C, Tsai K-C, Chiu H-T. Functional characterization and substrate specificity of spinosyn rhamnosyltransferase by *in vitro* reconstitution of spinosyn biosynthetic enzymes. *J. Biol. Chem.* 2009; 284:7352–7363. [PubMed: 19126547]
 18. Otwinowski Z, Minor W. Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology.* 1997; 276:307–326.
 19. Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D.* 2003; 59:1131–1137.
 20. McCoy A, Grosse-Kunstleve R, Adams P, Winn M, Storoni L, Read R. Phaser crystallographic software. *J. Appl. Cryst.* 2007; 40:658–674. [PubMed: 19461840]
 21. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D.* 2004; 60:2126–2132. [PubMed: 15572765]
 22. Vagin A, Steiner R, Lebedev A, Potterton L, McNicholas S, Long F, Murshudov G. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D.* 2004; 60:2184–2195. [PubMed: 15572771]
 23. Cohen S, Morris R, Fernandez F, Jelloul M, Kakaris M, Parthasarathy V, Lamzin V, Kleywegt G, Perrakis A. Towards complete validated models in the next generation of ARP/wARP. *Acta Crystallogr. D.* 2004; 60:2222–2229. [PubMed: 15572775]
 24. De Amicis CV, Graupner PR, Erickson JA, Paschal JW, Kirst HA, Creemer LC, Fanwick PE. The stereochemical outcome of electrophilic addition reactions on the 5,6-double bond in the spinosyns. *J. Org. Chem.* 2001; 66:8431–8435. [PubMed: 11735521]
 25. Creemer L, Kirst H, Paschal J. Conversion of spinosyn A and spinosyn D to their respective 9- and 17-pseudoaglycones and their aglycones. *J. Antibiot.* 1998; 51:795–800. [PubMed: 9766471]
 26. Borisova S, Zhang C, Takahashi H, Zhang H, Wong A, Thorson J, Liu H-w. Substrate specificity of the macrolide-glycosylating enzyme pair DesVII/DesVIII: opportunities, limitations, and mechanistic hypotheses. *Angew. Chem. Int. Ed. Engl.* 2006; 45:2748–2753. [PubMed: 16538696]
 27. Mittler M, Bechthold A, Schulz GE. Structure and action of the C-C bond-forming glycosyltransferase UrdGT2 involved in the biosynthesis of the antibiotic urdamycin. *J. Mol. Biol.* 2007; 372:67–76. [PubMed: 17640665]
 28. Zhang C, Bitto E, Goff R, Singh S, Bingman C, Griffith B, Albermann C, Phillips G, Thorson J. Biochemical and structural insights of the early glycosylation steps in calicheamicin biosynthesis. *Chem. Biol.* 2008; 15:842–853. [PubMed: 18721755]
 29. Chang A, Singh S, Helmich KE, Goff RD, Bingman CA, Thorson JS, Phillips GN Jr. Complete set of glycosyltransferase structures in the calicheamicin biosynthetic pathway reveals the origin of regiospecificity. *Proc. Natl. Acad. Sci. U.S.A.* 2011; 108:17649–17654. [PubMed: 21987796]
 30. Bolam D, Roberts S, Proctor M, Turkenburg J, Dodson E, Martinez-Fleites C, Yang M, Davis B, Davies G, Gilbert H. The crystal structure of two macrolide glycosyltransferases provides a

- blueprint for host cell antibiotic immunity. *Proc. Natl. Acad. Sci. USA*. 2007; 204:5336–5341. [PubMed: 17376874]
31. Holm L, Rosentrom P. Dali server: conservation mapping in 3D. *Nucl. Acids Res*. 2010; 38:W545–W549. [PubMed: 20457744]
32. Mulichak A, Losey H, Walsh C, Garavito R. Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure*. 2001; 9:547–557. [PubMed: 11470430]
33. Mulichak A, Losey Heather C, Lu W, Wawrzak Z, Walsh C, Garavito R. Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc. Natl. Acad. Sci. USA*. 2003; 100:9238–9243. [PubMed: 12874381]
34. Mulichak A, Lu W, Losey H, Walsh C, Garavito R. Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. *Biochemistry*. 2004; 43:5170–5180. [PubMed: 15122882]
35. Hu Y, Walker S. Remarkable structural similarities between diverse glycosyltransferases. *Chem. Biol*. 2002; 9:1287–1296. [PubMed: 12498881]
36. Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis B, Tarling C, Ford C, Bowles D, Davies G. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *Embo J*. 2006; 25:1396–1405. [PubMed: 16482224]
37. Hu Y, Chen L, Ha S, Gross B, Falcone B, Walker D, Mokhtarzadeh M, Walker S. Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc. Natl. Acad. Sci. USA*. 2003; 100:845–849. [PubMed: 12538870]

Figure 1.

The spinosad biosynthetic pathway. SpnG is proposed to rhamnosylate intermediates **3** and **4** at 9-OH (**7**). The rhamnose is important for insecticidal activity.

Figure 2. The SpnG dimer. A stereodisplay 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 in sticks. The N- and C-termini of each monomer are indicated.

Figure 3. The donor and acceptor binding sites. a) A stereodivision of the donor binding site of the SpnG-TDP complex shows 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) A stereodivision of the acceptor binding site of the SpnG-AGL complex shows 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.

Figure 4.

Active site residues. a) A stereodiagram of the $2F_o - F_c$ electron density maps of the SpnG-TDP complex (contoured at 1.3σ) illustrates the role of N202 in selecting TDP-sugars. The N202 side chain NH_2 forms a hydrogen bond with the TDP 3'-OH and also prevents a ribose-containing NDP-sugar from binding due to clashes with 2'-OH. b) A stereodiagram of the $2F_o - F_c$ electron density maps of the SpnG-TDP complex (contoured at 1.3σ) shows the SpnG active site. H13 may deprotonate the 9-OH of the 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 ribose 2'-OH and 3'-OH, but in SpnG is a threonine (T300) that coordinates N202 to help select for NDP-sugars containing a deoxyribose. SpnG (spinosyn rhamnosyltransferase, AAG23268.1, *Saccharopolyspora 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, *S. antibioticus*), GtfB (vancomycin glucosyltransferase, AAB49293.1, *Amycolptopsis orientalis*), GtfD (vancomycin vancosaminyltransferase, AAK31352.1, *A. orientalis*). d) A stereodiagram shows a SpnG active site in the absence of TDP (blue) and the presence of TDP (grey). 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.

Figure 5. Putative mechanism of SpnG. a) TDP-L-rhamnose was modeled into the SpnG-AGL structure. 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 9-OH for attack at the TDP-L-rhamnose anomeric carbon.

Table 1

Data collection and refinement statistics for SpnG, SpnG+TDP, and SpnG+AGL.

Crystallization Data	SpnG	SpnG+TDP	SpnG+AGL
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	<i>P1</i>	<i>P1</i>	<i>P1</i>
α, γ, β (°)	81.5, 73.8, 86.0	81.6, 73.8, 85.7	82.0, 74.1, 86.1
Molecules per AU	2	2	2
Measured Intensities	131,164	230,234	314,137
Unique Reflections	85,110	66,093	84,558
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)
Avg I/ σ I	12.1 (2.9)	18.5 (2.6)	22.6 (2.0)
Refinement Statistics			
Reflections	72,312	57,921	78,024
$R_{\text{cryst}}/R_{\text{free}}$	0.179/0.227	0.232/0.291	0.189/0.232
No. Protein Atoms	5577	5535	5527
No. Ligand Atoms	12	48	42
No. Solvent Atoms	614	435	528
Avg <i>B</i> -factor			
Monomer A (Å ²)	18.7	19.4	19.9
Monomer B (Å ²)	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			
Bonds Lengths (Å)	0.025	0.019	0.023
Bond Angles (°)	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