Chemistry & Biology
Article

Rational Design of an Aryl-C-Glycoside Catalyst from a Natural Product O-Glycosyltransferase

Johannes Härle,[1] Stefan Günther,[2] Benjamin Lauinger,[1] Monika Weber,[1] Bernd Kammerer,[3] David L. Zechel,[4] Andriy Luzhetskyy,[1] and Andreas Bechthold[1,*]
[1]Department of Pharmaceutical Biology und Biotechnology
[2]Department of Pharmaceutical Bioinformatics
[3]Zentrum für Biosystemanalytik (ZBSA)
Institute of Pharmaceutical Sciences, Albert-Ludwigs-Universität Freiburg, 79104 Freiburg, Germany
[4]Department of Chemistry, Queens University Kingston, Kingston, OH, K7M 4Z8 Canada
*Correspondence: andreas.bechthold@pharmazie.uni-freiburg.de
DOI 10.1016/j.chembiol.2011.02.013

SUMMARY

Because the sugar moieties of natural products are primarily O-linked, the hydrolytic sensitivity of the glycosidic linkage limits their therapeutic application. One potential solution to this problem is to replace the labile O-glycosidic bond with an enzymatically and chemically stable C-glycosidic bond. In this study, computational analysis of the O-glycosyltransferase LanGT2 and the C-glycosyltransferase UrdGT2 was used to predict the changes necessary to switch the O-glycosylating enzyme to a C-glycosyltransferase. By screening rationally designed LanGT2 mutants a number of LanGT2 variants with C-glycosyltransferase activity were identified. One variant, having 10 amino acid substitutions, revealed the primary region that determines O- versus C-glycosylation. By modeling the active site of this mutant and probing the role of active site residues with alanine substitutions, this work also illuminates the mechanistic features of O- and C-glycosylation.

INTRODUCTION

The field of protein engineering is steadily developing new techniques to create enzymes and proteins with desired features (Schmid et al., 2001). Improving enzyme characteristics such as substrate specificity, stereo- and regioselectivity, thermostability, solvent tolerance, or the enhancement of catalytic efficiency has been successfully applied to many enzymes (Petrounia and Arnold, 2000). Part of this success is due to rapid advances in the creation of large protein libraries through random mutagenesis techniques (particularly with polymerase chain reaction [PCR]) along with efficient screening or selection methods to find proteins with desired features (Williams et al., 2008a; Reidhaar-Olson and Sauer, 1988; Turner, 2009). Likewise, rational design in which targeted mutations are guided by techniques like crystallography and computational modeling, has also proven to be a successful, and more straightforward, approach for generating desired properties in proteins (Love, 2010; Hohne et al., 2010).

Glycosyltransferases (GTs) are a key class of enzyme involved in the biosynthesis and diversification of many natural products (Williams and Thorson, 2008; Thibodeaux et al., 2007). With regio- and stereospecific control these enzymes transfer a wide array of glycosyl moieties onto other sugars or natural product aglycons of polyketide or nonribosomal polypeptide origin (Luzhetskyy et al., 2008; Olano et al., 2010). Often this reaction converts a nonbioactive intermediate into a drug (Weymouth-Wilson, 1997; Heddle and Maxwell, 2002). In the field known as combinatorial biosynthesis (Walsh, 2002), GTs and other so-called post-tailoring enzymes are introduced or deleted within a biosynthetic pathway to produce new compounds (Walsh et al., 2001). Although this approach can be very valuable to gain structural diversity, its application is often limited by the native specificity of an enzyme. The rational design of new tailoring enzymes may be one solution for expanding these boundaries. In addition to the physiologically relevant properties, the stability of a pharmaceutical compound is also a key feature that must be considered. This is a particularly crucial issue concerning glycosylated compounds. For instance, although O-glycosides are acid-sensitive, sugar moieties attached by C-glycosidic bonds are stable toward chemical hydrolysis. This fact offers interesting biological applications because C-glycosides may mimic the bioactivity of the corresponding O-glycosides, but offer considerably more resistance to degradation (Weatherman, et al., 1996). In this context the GTs present a significant opportunity for protein engineering as they catalyze the attachment of a sugar to a sugar-acceptor (Davis, 2007). In general glycosyl transfer occurs from a NDP-activated sugar to the hydroxyl group of an acceptor (Unligil and Rini, 2000). Consequently, GTs divide into two mechanistic classes of glycosyl transfer, inverting and retaining, according to the anomeric configuration of the sugar before and after glycosylation. Both retaining and inverting GTs require general base catalysis to increase the reactivity of the acceptor hydroxyl group. Whereas the retaining GTs likely follow a double displacement mechanism to achieve retention of anomeric configuration, the inverting O-GTs follow what is essentially a single step S_N2 reaction, wherein an acceptor hydroxyl group is deprotonated to directly attack the anomeric C1 of the NDP-sugar to form an O-glycosidic

bond. This reaction is also assisted by the properties of the nucleoside diphosphate that functions as a good leaving-group (Unligil and Rini, 2000). With steadily increasing knowledge about GT structures and mechanisms, these enzymes have also been subjected to rational design to obtain desired features (Williams et al., 2008b).

Namely, sequence and structural-guided mutagenesis has been successfully applied to several bacterial natural product GTs (e.g., OleD, GtfE, KanF, ElmGT) (Williams et al., 2007; Ramos et al., 2009; Park et al., 2009; Truman et al., 2009). Next to others, these articles impressively document how modified GTs expand the diversification of therapeutically important compounds. Generally, the engineering approaches were mainly focused on modifying the enzyme to incorporate other sugars. However, in other studies GTs were designed with new regioselectivities, including an example from our lab involving the urdamycin olivosyltransferase from *Streptomyces fradiae* Tü2717, UrdGT1b. Whereas wild-type UrdGT1b transfers D-olivose to the C4 hydroxyl of L-rhodinose during the biosynthesis of urdamycin, an engineered derivative was created that catalyzed transfer to the C4 hydroxyl of the nearby D-olivose moiety, yielding in an unusual branched oligosaccharide (Hoffmeister et al., 2002). Furthermore, in 2009 we were able to convert the *Streptomyces globisporus* LndGT1 from a GT that natively catalyzes the attachment of a single sugar during landomycin E biosynthesis toward an iteratively acting enzyme (Krauth et al., 2009). This modification is of special interest as an elongated sugar chain is believed to enable better penetration of a molecule through cell membranes (Weymouth-Wilson, 1997).

In this study, we present the process of engineering the O-glycosylating enzyme LanGT2 from *Streptomyces cyanogenus* S136 to a C-glycosylating enzyme. LanGT2 catalyzes formation of an O-glycosidic bond in what is the initial glycosylation step in the biosynthesis of the anticancer agent landomycin A (Figure 1A) (Luzhetskyy et al., 2005; Shaaban et al., 2011). The nucleophilic site for the wild-type glycosyl transfer reaction is the hydroxyl group of a hydroxyanthroquinone ring system. We hypothesized that LanGT2 could be converted to a C-glycosylating enzyme through redesign of the active site to achieve glycosyl transfer to the carbon ortho to the acceptor hydroxyl moiety. We also hypothesized that general base activation of the phenolic oxygen would afford carbanionic character to the ortho-carbon, thereby providing an opportunity for nucleophilic attack of this carbon on the anomeric position of the donor sugar to form the desired aryl-C-glycoside.

The exact mechanism of C-glycosidic bond forming GTs is not well understood and can primarily be suggested based on knowledge from general aromatic substitution chemistry. Two mechanisms, the O- to C-glycosyl rearrangement and a Friedel-Crafts-like mechanism, have been proposed for C-GTs (Bililign et al., 2004; Durr et al., 2004). Early studies by our group on UrdGT2, the C-glycoside forming GT from the Urdamycin A producer *S. fradiae* Tü2717 (Figure 1A), revealed this enzyme to be highly promiscuous with respect to acceptor and sugar donor specificity. However, in all cases, UrdGT2 was regiospecific for glycosyl transfer to the ortho-position of hydroxyanthroquinone acceptors. This was even the case for 1,2-dihydroxyanthraquinone (alizarin), whereby UrdGT2 transferred D-olivose to the ortho-hydroxyl, thereby forming an O-glycoside. The ability of UrdGT2 to form O- and C-glycosidic bonds ortho to the phenolic oxygen of the acceptor strongly suggests a direct Fiedel-Crafts-like alkylation as the mechanism rather than an O- to C-rearrangement (Durr et al., 2004; Baig et al., 2006; Faust et al., 2000; Hoffmeister et al., 2000, 2001, 2003). In 2006, the X-ray crystal structure of UrdGT2 was published, but unfortunately attempts to obtain a complex with glycosyl donor and acceptor substrates, which would have provided more concrete insights into the reaction mechanism, were unsuccessful (Mittler et al., 2007). Nevertheless, considerable structural information exists for related natural product O-GTs, such as GtfA (Protein Data Bank [PDB] code 1PN3), GtfD (PDB code 1RRV), OleD (PDB code 2IYF), and OleI (PDB code 2IYA), all of which had been crystallized with substrates. Moreover, the overall topologies and active sites of these enzymes are highly conserved (Mulichak et al., 2003, 2004; Bolam et al., 2007). This provides important information about the mechanisms and substrate specificities of the GT-B superfamily, to which most natural product GTs belong.

The O-glycosyltransferase LanGT2, the target of our study, accepts the same sugar-donor and a similar sugar-acceptor as the C-GT UrdGT2 (Figure 1A) (Luzhetskyy et al., 2005; Zhu et al., 2005). Additionally, the primary sequences of LanGT2 and UrdGT2 (373 and 378 amino acids, respectively) share 53% sequence identity (Figure 1B). The high sequence similarity shared by these enzymes, along with the observed flexibility of UrdGT2 in catalyzing C- and O-glycosidic bonds, suggested that it may be possible to use specific mutations to transform LanGT2 from an O-GT to a C-GT. An in vivo test system for this approach, the mutant *S. cyanogenus ΔlanGT2*, was available in our laboratory. Expression of *urdGT2* in this mutant results in the production of the C-glycoside 9-C-D-olivosyltetrangulol (Luzhetskyy et al., 2005). In contrast, expression of *lanGT2* in this mutant leads to the production of O-glycoside 8-O-D-olivosyllandomycinone (Luzhetskyy et al., 2005, Zhu et al., 2007) (Figure 2). Following a sequence and structural-based mutagenesis approach we constructed in total 106 variants of LanGT2 and tested for C- or O-glycosylating activity.

The differences in sequence of LanGT2 and UrdGT2 are distributed throughout the polypeptide chains, resulting in uncertain predictions about regions governing the substrate specificity or functional relevance. Despite the lesser extent of sequence identity of LanGT2 to the natural product GTs OleD (18%), OleI (20%), GtfA (18%), and, GtfD (17%), the structures of these enzymes provided a basis for the computational protein modeling experiments described in this study and ultimately the successful creation of a C-glycosylating-LanGT2.

RESULTS

Construction of LanGT2-UrdGT2 Chimeras

To determine which sequence regions of UrdGT2 were responsible for the C-glycosylating function, two routes in parallel were attempted. The first used restriction enzyme digests and ligation reactions to assemble a small library of *urdGT2-lanGT2* chimeric genes. Sequence alignment revealed specific restriction sites convenient for constructing the hybrids (Figure 1B). Additional restriction sites were introduced by PCR-based mutagenesis (see Supplemental Experimental Procedures

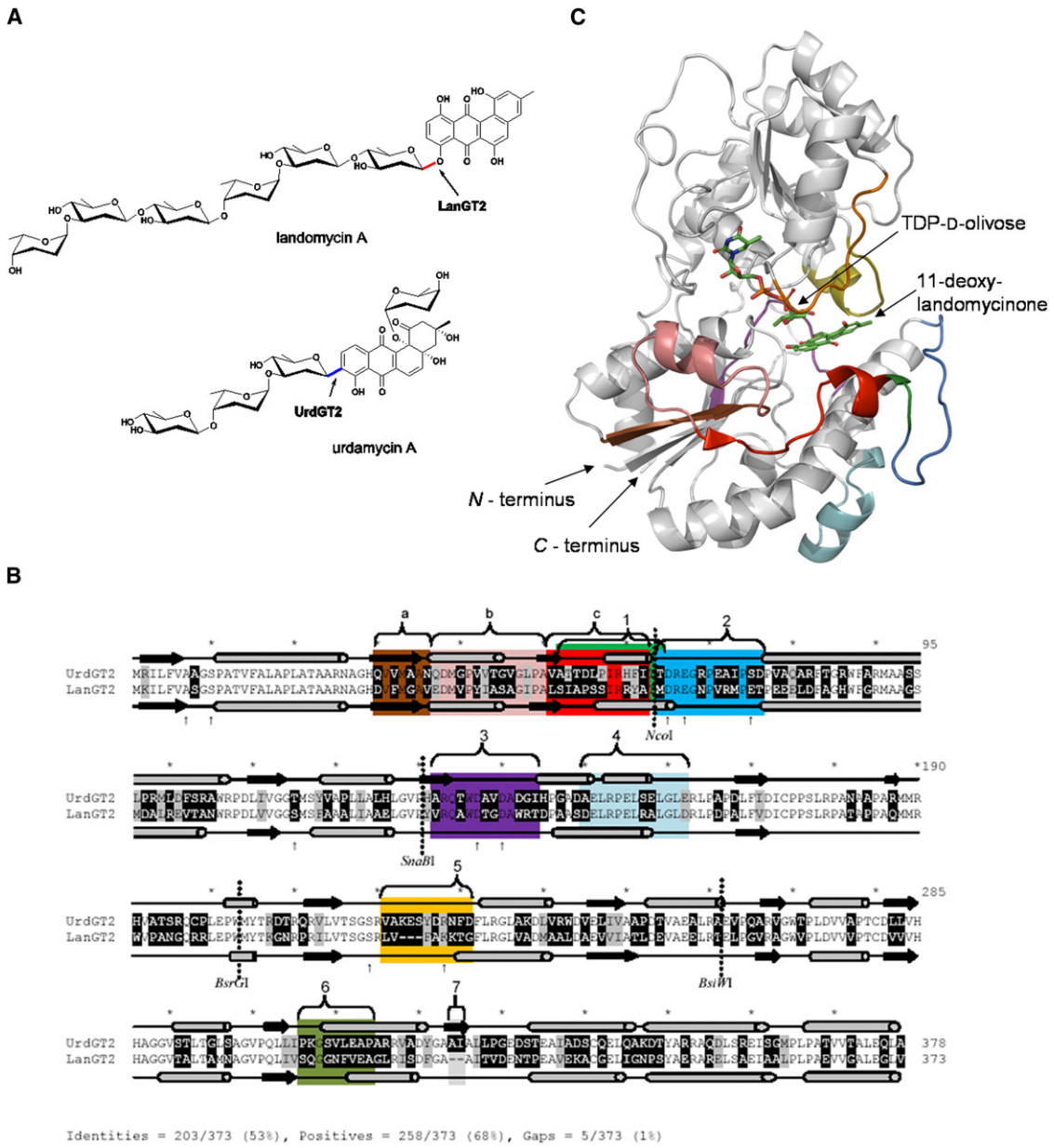

Figure 1. Functions of the Glycosyltransferases LanGT2 and UrdGT2, a Structural-Based Sequence Alignment and a LanGT2 Ternary Model Illustrating the Protein Regions Investigated by Targeted-Amino Acid Replacement Studies

(A) During landomycin A biosynthesis the O-glycosyltransferase LanGT2 from *Streptomyces cyanogenus* S136 catalyzes an O-glycosidic bond (red) between D-olivose and the corresponding landomyin aglycon. The C-glycosyltransferase UrdGT2 from *Streptomyces fradiae* Tü2717 attaches by a C-glycosidic bond (blue) the same deoxy sugar toward a very similar angucycline named UWM6.

(B) The secondary structure is derived from UrdGT2 crystal structure (Protein Data Bank [PDB] code 2P6P). Protein regions of homology between LanGT2 with UrdGT2 used for substitution studies are presented in brackets and segmented in colored blocks: 1 (dark green); 2 (marine blue); 3 (purple); 4 (cyan); 5 (orange); 6 (olive); and 7 (gray) plus a (brown), b (salmon), and c (red). Residues replaced with alanine to probe their role in O- and C- glycosylation are indicated with arrows. The location of restriction sites utilized for creating LanGT2-UrdGT2 hybrids are indicated.

(C) Homology model of LanGT2 with TDP-D-olivose and 11-deoxy-landomycinone based on the modeled UrdGT2 ternary complex described in Mittler et al. (2007). The LanGT2 protein regions that were substituted with homologous sites from UrdGT2 are highlighted with the some colors as presented in the sequence alignment: region 1 (Ser52–Met64) in dark green, region 2 (Met63–Thr76) in marine blue, region 3 (Val132–Thr144) in purple, region 4 (Asp150–Asp162) in cyan, region 5 (Leu221–Gly228) in orange, region 6 (Ser303–Ala311) in olive, block a (Asp30–Val36) in brown, block b (Glu37–Ala50) in salmon, and block c (Leu51–Y61) in red.

See also Figures S1 and S2.

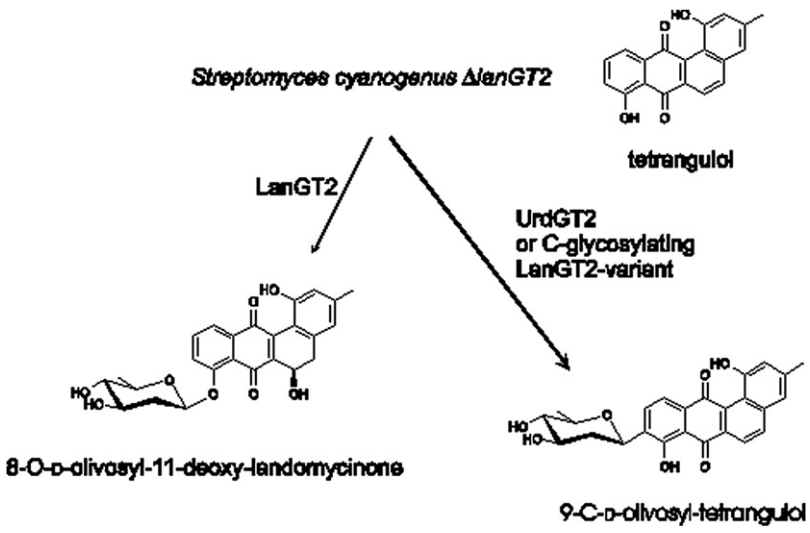

Figure 2. In Vivo Test System

S. cyanogenus ΔlanGT2 served as the expression host. Whereas the mutant itself produces tetrangulol, expression of LanGT2, UrdGT2, or a C-glycosylating LanGT2 variant, the derivatives 8-O-D-olivosyl-11-deoxy-landomycinone or 9-C-D-olivosyltetrangulol are produced.

available online). The restriction fragments encoding the N- and C-terminal halves of each enzyme were ligated to generate 12 different *lanGT2-urdGT2* chimeras (*lanGT2.h1* to *lanGT2.h12*) (Table S1). The second approach utilized site-directed mutagenesis guided by sequence and structure analysis of LanGT2 with UrdGT2 and other natural product GTs.

Generating LanGT2-Variants with UrdGT2 Loops

Sequence alignments of UrdGT2 and LanGT2 with the GTs OleI, OleD, GtfA, and GtfD (Figure S1) revealed two protein regions that are exceptionally variable. These regions can be localized to the N3- and N5-loops, both of which link a β sheet to an α helix. Due to the observed sequence variability, we presumed that these regions may in part govern substrate specificity either by direct or second sphere derived ligand interactions. X-ray crystal structures of the cocrystallized natural product GTs also show that these loops line the sugar acceptor pockets (Unligil and Rini, 2000). Additionally, in GTs crystallized without substrates, the electron density representing these regions were poorly resolved (Unligil and Rini, 2000; Mittler et al., 2007). This is indicative of conformational flexibility and further suggests the relevance of these loops in sculpting the sugar acceptor pocket. The very similar topology of GTs within the GT-B family, plus the high sequence variability in N3- and N5-loops, prompted the idea that enzyme specificity could be altered simply by swapping the loop regions. The fact that the same sugar donor and very similar acceptor molecules are recognized by LanGT2 and UrdGT2 suggested that LanGT2 could be switched to an UrdGT2-like C-glycosylating enzyme mainly by redesigning the loop regions of the acceptor binding site. To experimentally evaluate this idea the loop regions assumed to govern specificity in the LanGT2 acceptor pocket were substituted with the corresponding UrdGT2 loop regions. To avoid a structural clash arising from replacing relatively long stretches of sequence, the N3- and N5-loops were subdivided into blocks 1 + 2 and 3 + 4, respectively (Figure 1B). In addition, the structure of UrdGT2 revealed three residues (Glu66, Glu71, or Glu224) that could serve as a general base to increase the phenolate character of the acceptor and trigger the C-glycosylation reaction. As presented by Mittler et al. (2007) the modeled UrdGT2 enzyme-ligand complex confirmed the close contact of Glu224 to both the sugar donor and acceptor molecules. Interestingly, as shown by the LanGT2-UrdGT2 sequence alignment, LanGT2 displays a gap of three amino acids where this residue is expected (^{223}KES^{225} in UrdGT2). To prove if this region is relevant for C-GT functionality, LanGT2 residues within the region Leu221–Gly228 (block 5) were substituted with the corresponding UrdGT2 amino acids (Val221–Asp231). By applying a PCR megaprimer method (Ho et al., 1989), 30 different combinations of the UrdGT2 protein regions 1–5 (Figure 1B) were substituted into LanGT2, resulting in mutants 13–43 (Table S2).

Functional Screening of the LanGT2-UrdGT2 Hybrids

The 12 *lanGT2-urdGT2* chimeras and 30 PCR derived *lanGT2* point mutants were cloned into the *Escherichia coli-Streptomyces* shuttle vector pUWL-oriT. The expected changes were verified by DNA sequencing. The *lanGT2* mutants were then heterologously expressed in *S. cyanogenus ΔlanGT2* under the control of the *ermE** promotor on the vector. After cultivating the complimented strain under producing conditions, the secondary metabolites were extracted and analyzed by HPLC-MS. Four weakly active LanGT2 variants were identified from this library. The mutants LanGT2.4, LanGT2.5, and LanGT2.45 all formed 8-O-D-olivosyllandomycinone and thus still retained the wild-type O-glycosylating activity (Table S2). However, careful MS analysis of the extract afforded by the LanGT2.1 mutant revealed a minute quantity of the C-glycoside product. In this variant residues 52–64 derived from UrdGT2 appear to contribute to the desired C-glycosylating function (Table S2).

A very similar result was obtained by investigating the 12 LanGT2-UrdGT2 chimeras. Most of these mutants were likely unstable or inactive, as no conversion could be detected. However, the variant LanGT2.h1, in which the N-terminal 62 amino acid residues are derived from UrdGT2, specifically produced the C-glycoside product with 17% of the UrdGT2 conversion factor (Table S1). The product 9-C-D-olivosyltetrangulol displayed the typical ultraviolet (UV)-visible spectrum of tetrangulol, whereas HPLC-MS analysis revealed the expected pseudo-molecular ion (433.1 $[M-H]^-$) representing the mass of D-olivose attached to tetrangulol (Luzhetskyy et al., 2005) (Figure 3B). The structures of the metabolites 8-O-D-olivosyllandomycinone and 9-C-D-olivosyltetrangulol were characterized previously by NMR and MS analysis (Luzhetskyy et al., 2005; Zhu et al., 2007). For instance Luzhetskyy et al. heterologously expressed the C-GT UrdGT2 in *S. cyanogenus ΔlanGT2* and observed a product with the analogous retention time, mass

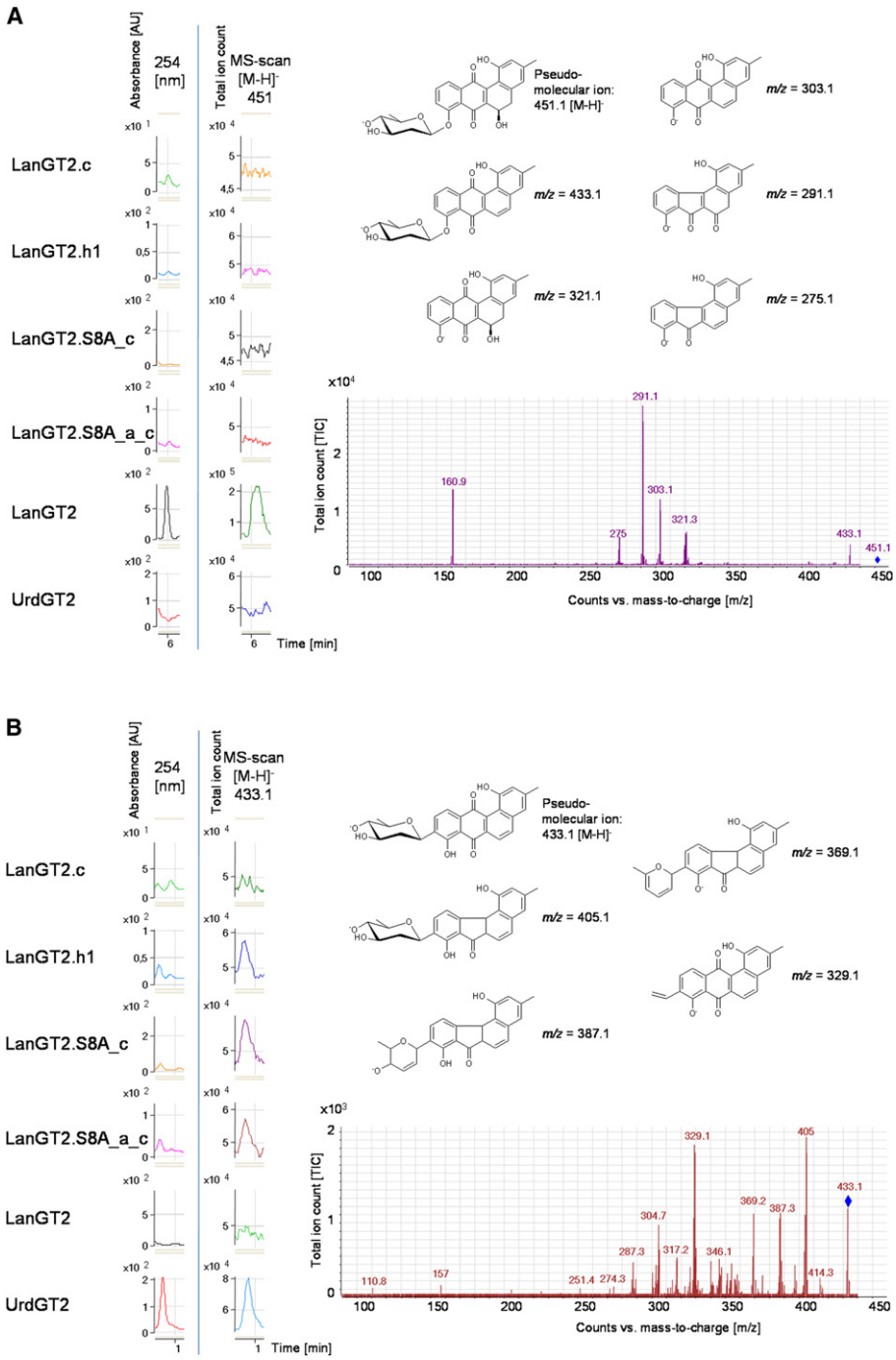

Figure 3. HPLC-MS/MS Analysis of the O- and C-Glycosides Produced by LanGT2, UrdGT2, and Selected C-Glycosylating LanGT2 Variants

Reversed phase HPLC-MS/MS analysis of the crude extracts isolated from *S. cyanogenus* Δ*lanGT2* after expressing LanGT2.c, LanGT2.h1, LanGT2.S8A_c, LanGT2.S8A_a_c, LanGT2, or UrdGT2. LanGT2 and UrdGT2 were used as positive controls.

(A) LanGT2 produces the O-glycoside 8-O-D-olivosyl-11-deoxy-landomycinone that has a retention time of 16 min and can be detected by absorbance at 254 nm as well as selective ion monitoring by MS (negative ion mode, m/z = 451.1). Absorbance and MS chromatograms are shown at the far left. Only the wild-type LanGT2 produces the O-glycoside product at detectable levels. The corresponding MS/MS fragmentation ion spectrum for 8-O-D-olivosyl-11-deoxy-landomycinone (molecular ion marked with a diamond) is shown along with the structures of the corresponding putative fragment ions.

(B) The enzymes LanGT2.c, LanGT2.h1, LanGT2.S8A_c, LanGT2.S8A_a_c, and UrdGT2 produce the C-glycoside 9-C-D-olivosyltetrangulol that elutes at 20.8 min and can be detected by absorbance at 254 nm and selective ion monitoring for the molecular ion m/z = 433.1. The MS/MS fragmentation pattern for the pseudo-molecular ion 433.1 $[M-H]^-$ (marked with a diamond), derived from the UrdGT2 and C-glycosylating LanGT2 derivatives product, differs strongly from that for the O-glycoside product of LanGT2. The suggested structures of the fragment ions are shown.

See also Figure S3.

and UV-visible spectrum of the 9-C-D-olivosyltetrangulol gained by the LanGT2 derivative LanGT2.h1.

In addition to the distinct differences in absorbance spectra and masses of the O- and C- glycosylated tetrangulol derivatives, HPLC-MS/MS analysis could also demonstrate the presence of the C-glycosidic bond. This is because gas phase fragmentation of the much stronger C-glycosidic bond requires higher collision energies than that required for O-glycosides (Abad-Garcia et al., 2008). Crude extracts isolated from *S. cyanogenus* Δ*lanGT2* expressing LanGT2, UrdGT2, and selected C-glycosylating LanGT2-derivatives were analyzed by LC-MS/MS. For the glycosylation product formed by the C-GTs the fragmentation pattern for the pseudo-molecular ion 433.1 $[M-H]^-$ was extracted. As this ion represents 9-C-D-olivosyltetrangulol, no fragment corresponding to the deglycosylated state could be detected. In contrast, on fragmentation of the LanGT2 product ion 451.1 $[M-H]^-$, the expected deglycosylated landomycinone ion m/z = 321.1 was clearly observed (Figure 3; Figure S3).

Additional Refinement of N-Terminal Residues to Improve C-Glycosylation

The activity shown by LanGT2.1 and LanGT2.h1 motivated us to investigate the specific sites within the N-terminal 62 amino acids that are relevant for the C-glycosylating activity. For this purpose the most variable sequence regions between LanGT2 and UrdGT2 (positions 30–62) were divided into blocks a, b, and c (Figure 1B). By constructing *lanGT2* derivatives with single and combined substitutions the importance of these regions in influencing the C-glycosylation was investigated. Six derivatives (*lanGT2.a*, *lanGT2.b*, *lanGT2.c*, *lanGT2.a_b*, *lanGT2.a_c*, and *lanGT2.a_b_c*) were constructed by mega-primer PCR. The mutant genes were cloned, verified by sequencing, expressed, and screened for activity as stated above. As expected from the earlier result with LanGT2.1, region c was identified to be important for the C-glycosylating function. The derivative LanGT2.c, carrying 10 residues from UrdGT2, completely lost its O-glycosylating activity and instead gained the desired C-glycosylating capability. However, the enzyme activity was weakly active, affording only 5.9% the yield of C-glycoside product as compared to UrdGT2. By splitting block c into c-loop (Val51–Pro57) and c-helix (Pro57–Ile62) a synergistic effect was observed whereby residues from both parts were essential for C-glycosylation. Moreover, HPLC-MS analysis revealed that LanGT2.c-loop and LanGT2.c-helix also exhibited O-glycosylating functionality (Table S2). The residues in region a or b slightly enhanced the C-glycosylating activity of LanGT2.c, but the region conferring the C-C bond formation could clearly be localized to a 10-amino acid sequence, Val52 to Ile62, located adjacent to the N3-loop region. To further demonstrate the function of this sequence, the corresponding 10 residues in UrdGT2 were substituted with the sequence from LanGT2 and the resulting derivative UrdGT2.c was tested for functionality. As the crude extract of *S. cyanogenus* Δ*lanGT2* x pUWL-*urdGT2.c* exhibited no product formation as compared to a negative control (*S. cyanogenus* Δ*lanGT2* x pUWL) this substitution in UrdGT2 most probably led to an inactive hybrid. As a result of the reduced C-glycosylating activity of LanGT2.c with respect to LanGT2.h1, as well as the lack of significant enhancement on substituting regions a or b, we considered whether engineering other sites within the N-terminal 50 amino acids would lead to an improvement. Based on the LanGT2-UrdGT2 sequence alignment position 8 appeared to be a promising site for modification. Whereas LanGT2 has serine at this position, UrdGT2 and the related C-GTs SaqGT5, SimB7, Med-orf8, Hed-orf13, all of which glycosylate the phenolic ortho carbon of similar acceptor molecules, have alanine or glycine at this position (Figure S2). Indeed, the corresponding *lanGT2* mutant LanGT2.S8A_c had improved efficiency, identical to LanGT2.h1, affording 17% of the yield of C-glycoside product generated by UrdGT2. A further improvement to 19% of the UrdGT2 yield was achieved with the substitutions D30Q, F32V, G34A, and V36N within block a. The resulting hybrid LanGT2.S8A_a_c, with a total of 15 UrdGT2 derived substitutions, was examined for further studies concerning the question of the underlying mechanistic difference between the O- and C-glycosylation.

Computational Docking Studies Suggest that a Modified N5-Loop in a LanGT2-Chimera Favors C-Glycosylation

As structural data is not available for LanGT2 or UrdGT2 bound to substrates, a computational approach was used to identify the sites most likely involved in substrate binding and catalysis. Based on the crystal structure of UrdGT2 (PDB code 2P6P), the natural glycosyl acceptor UWM6 and the glycosyl donor TDP-D-olivose were docked onto the active site using the docking software Glide. Although this model was largely consistent with the modeling studies from Mittler et al. (2007), we observed one additional docking mode for UWM6. In this configuration the C8 hydroxyl group of the acceptor and the C1-atom of the glycosyl donor are located in close proximity (~4.1 Å) and thus might be able to form an O-glycosidic bond. Both substrate binding modes (1 and 2) for C- and O-glycosylation are characterized by a high number of favorable molecular interactions (van der Waals, Coulombic, and hydrophobic contacts) as expressed by the Glide Score (1 = −7.9 and 2 = −8.1), although the number of hydrogen bonds differs (three for mode 1, one for mode 2). Based on the assumption that the LanGT2 active site will be very similar to that calculated for UrdGT2, both docking modes can likewise be expected, and therefore selection of amino acids favoring O- or C-glycosylation might be possible. Using UrdGT2 as a template, structure homology models of LanGT2 were built and, as expected, both binding modes were also observed on docking the putative glycosyl acceptor 11-deoxy-landomycinone within the proposed active site. However, the binding mode favoring O-glycosylation had the better binding energy term as reflected by the Glide Score (−6.9 versus −5.9). In contrast, the model of LanGT2.S8A_a_c clearly agrees with the C-glycosylating activity observed for this variant, as indicated by the proximity of the acceptor C9 to C1 of the activated sugar (Figure 4), as well as an improved Glide Score for this binding mode relative to wild-type LanGT2 (Table S5). Although the two acceptor binding modes can explain each glycosylation outcome, the Glide Scores for each binding mode do not always match the observed glycosylation activity of all LanGT2 variants (Tables S1–S4 versus Table S5). This likely reflects the failure of computational modeling to fully account for the flexibility in the active site loops, particularly the interaction between residues

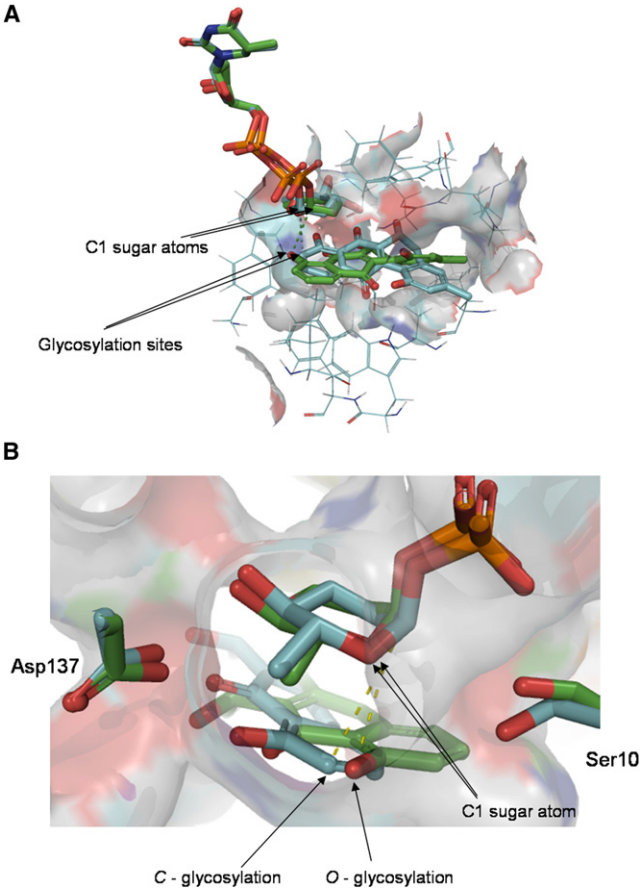

Figure 4. LanGT2 Modeling with a Focus on the Acceptor Positioning as the Key for O- or C-Glycosidic Bond Formation

LanGT2 modeling was based on the modeled UrdGT2 ternary complex described in Mittler et al. (Mittler et al., 2007).

(A and B) Overlay of the ligands docked onto LanGT2 and its C-glycosylating variant LanGT2.S8A_a_c_KES. Highlighted in lines and in surface view are hydrophobic amino acids involved in shaping the acceptor pocket. The LanGT2 redocking experiments present LanGT2 favoring O-glycosylation toward C8 hydroxyl group (shown in green), whereas the LanGT2.S8A_a_c_KES re-docking resulted in favoring C-glycosylation onto C9 (shown in cyan). (B) A view of (A) rotated 90° anti-clock wise. Modes for O- and C-glycoside formation are shown in green and cyan and are slightly angulated to each other. The carboxylic amino acid aspartate in position 137 and the serine residue 10 are in appropriate proximity to mediate the glycosylation reaction.

See also Table S5.

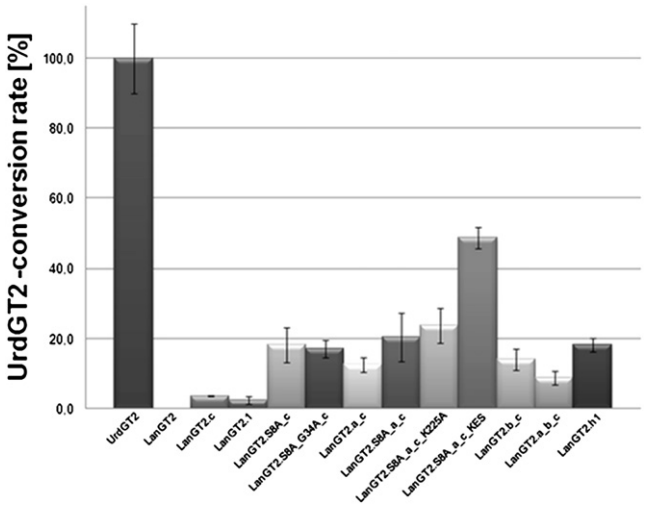

Figure 5. Relative C-Glycosylation Yields of Selected LanGT2 Derivatives

Relative yields (in triplicate) were determined by HPLC and calculated by normalizing the peak area for the C-glycoside produced with the LanGT2 variant with the peak area of the 9-C-D-olivosyltetrangulol produced by UrdGT2. Error bars represent the standard deviation.

See also Tables S1–S4.

in block c (residues 51–61) and the highly flexible loop consisting of residues 47–76.

Optimizing C-Glycosylation Activity in LanGT2 Variants by Rational Engineering

UrdGT2 possesses five additional amino acids when compared to LanGT2, three of which are in positions 223 to 225 of block 5 and two in positions 324 and 325 of block 7. In silico analysis suggested that the amino acids ^{223}KES^{225} would induce an alternative loop conformation that would help position the acceptor within the binding pocket for C-glycosylation. To determine the functional relevance of ^{223}KES^{225}, these residues were introduced into wild-type LanGT2 and the C-GT variant LanGT2.S8A_a_c, yielding LanGT2.KES and LanGT2.S8A_a_c_KES. HPLC-MS analysis of crude extracts produced by LanGT2.KES revealed that these additional residues did not change functionality, but rather diminished the O-glycosylation yield to 62% of wild-type LanGT2 (Table S2). However, variant LanGT2.S8A_a_c_KES, derived from the C-glycosylating derivative LanGT2.S8A_a_c, had dramatically improved activity, producing 45% of the yield of C-glycoside obtained with the native C-GT UrdGT2 (Figure 5; Table S2). In total, variant LanGT2.S8A_a_c_KES has 18 substitutions derived from UrdGT2. Furthermore, LanGT2 variants with modifications in block 6 and block 7 were constructed, but HPLC-MS analysis revealed no improvement in the C-glycosylation activity (for more details, see Supplemental Experimental Procedures).

Features that Dictate O- or C-Glycosylation in LanGT2

The variant LanGT2.c demonstrates that a sequence of 10 amino acids within block c is critical to convert LanGT2 from an O-GT to a C-GT. The underlying factors are not fully understood, as the modeled LanGT2 ternary structure shows the major part of block c to have no direct ligand interactions. Nevertheless, in silico analysis of LanGT2 bound to substrates showed Ala62 to be in close proximity to the sugar acceptor. Interestingly, in UrdGT2 and the other C-GTs considered above a bulkier isoleucine residue is located at this position (Figure S2). Based on our modeled ternary structures we predicted that in an interplay between the residues at positions 8 and 62 would shift the sugar acceptor toward the appropriate location and thereby direct glycosylation at the C8 hydroxyl or C9 carbon (Figure 4). However, as shown by HPLC-MS analysis, the substitutions S8A and A62I in LanGT2, producing the variant LanGT2.S8A_A62I, was not sufficient to produce C-glycosylation activity. Apparently the UrdGT2 amino acids upstream of position 62 are also necessary for

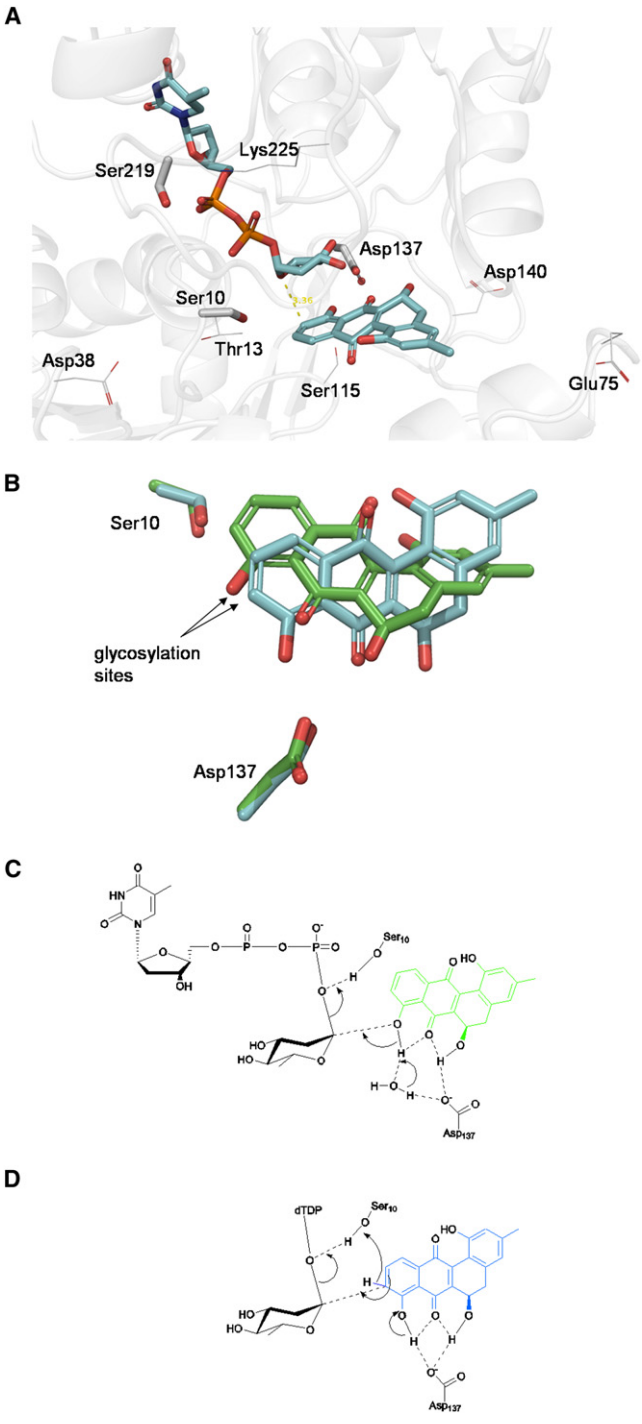

Figure 6. A Mechanism for O- and C-Glycoside Formation Based on Acceptor Positioning

Residues shown to be involved in the catalytic reactions were confirmed by Ala substitutions.

(A) Shown in gray cartoon mode is the active site of the modeled ternary structure of the C-glycosylating LanGT2 variant LanGT2.S8A_a_c_KES with the ligands TDP-D-olivose and 11-deoxy-landomycinone after accomplished redocking experiments. The residues Ser10, Asp137, and Ser219 identified to be essential for catalysis are presented in stick view. Ala substitutions performed onto residues presented in lines were shown to be nonlethal on glycoside formation.

(B) Overlay of the two positioning modes of 11-deoxy-landomycinone favoring O- and C-glycosylation, respectively. Active site residues Asp137 and Ser10 are shown as sticks. The mode favoring O-glycosylation toward C8 hydroxyl group is presented in green colored stick view. The acceptor mode favoring glycosylation of C9 is colored in cyan. The glycosylation site of each enzyme is highlighted with an arrow.

(C) Mechanism of LanGT2 catalyzed O-glycosylation of 11-deoxy-landomycinone with TDP-D-olivose illustrating the proposed roles of Asp137 and Ser10. A water molecule is thought to mediate proton transfer to Asp137.

(D) Mechanism for C-glycosylation of C9 of 11-deoxy-landomycinone (see text for details).

C-glycosylation. This additional contribution can be calculated by substrate docking experiments onto LanGT2 and its C-glycosylating derivatives. In the modeled LanGT2.c ternary complex the sugar acceptor favored the UrdGT2-like position, consistent with the weak C-glycosylation activity of this mutant. Likewise, acceptor docking with LanGT2.S8A_a_c_KES revealed Glide Scores that were comparable to those observed with UrdGT2 (Table S5). This calculation supports the superior activity of this mutant, which affords 45% of the product yield of wild-type UrdGT2.

Proposed Reaction Mechanisms for LanGT2 and UrdGT2

Following the assumption that LanGT2 and UrdGT2 execute a very similar reaction mechanisms, and that the positioning of the sugar acceptor largely determines O- or C-glycosidic sugar attachment, there are likely similar active site residues that are used to stabilize the glycosylation transition state. To experimentally identify these critical residues we performed a series of individual alanine substitutions in LanGT2, the C-glycosylating variant LanGT2.S8A_a_c, and the native C-glycosylating enzyme UrdGT2. These substitutions targeted all aspartate, glutamate, and serine amino acids observed within a distance of 10 Å around the anomeric carbon of the sugar donor as well as C8 and C9 of the acceptor molecules. Selection of these sites was assisted by the modeled LanGT2 and UrdGT2 ternary complexes, leading to a total of 26 point mutations (Tables S3 and S4). As shown by HPLC-MS analysis, the activity of all three enzymes was abolished with the substitutions S10A, D137A, or S219A, indicating that these residues are critical for enzyme folding and/or catalysis. For all other mutants the Ala substitution resulted in reduced glycosylation yields (Tables S3 and S4).

As previously proposed in Mittler et al. (2007) for UrdGT2, Asp137 likely serves as a general base by stabilizing development of a phenolate ion and thereby increase the carbanionic character of the ortho carbon (C9) through resonance. This would increase the reactivity of the acceptor in a Friedel-Crafts like reaction with the sugar donor. Likewise, models of the C-glycosylating variants of LanGT2 (Figure 4B) reveal the carboxylate group of Asp137 to be within hydrogen bonding distance of the phenolic (C8) hydroxyl of the acceptor. The importance of this residue for activity suggests a similar role for Asp137 as a general base. Indeed, the models indicate that no other residue is appropriately positioned to fulfill this critical role (Figure 6). Interestingly, the role of Asp137 in wild-type LanGT2 may be slightly different due to the position of the acceptor in the active site. As shown in Figure 6C, the distance between the carboxylate

of Asp137 and the C8 hydroxyl group (~5 Å) is too large for direct proton transfer during general base catalysis. Nevertheless, transfer of a proton through a water molecule or slight shifting of the Asp137 side chain to within hydrogen bonding distance of the acceptor may allow this residue to fulfill this role.

The reaction with the glycosyl donor most likely follows the typical S_N2 mechanism as often described for inverting GTs. As illustrated in Figure 4 and Figure 6, the active site residue Ser10 in UrdGT2 and LanGT2 is in close proximity to the acceptor nucleophilic atoms. As this residue is within hydrogen bonding distance of the acceptor C8 hydroxyl group (in LanGT2) and C9 carbon (in UrdGT2 and C-glycosylating LanGT2 variants), Ser10 possibly serves to stabilize the increase in electron density expected at these positions in the transition state. After acceptor activation the anomeric carbon of sugar donor is attacked to form the glycosidic bond, which is assisted by the good leaving group properties of the nucleoside diphosphate. As revealed by the modeled C-glycosylating LanGT2 ternary complex (Figure 6A), the dramatic effect of the S219A substitution is most likely due to a disrupted interaction with the sugar-donor. As shown in the sequence alignment (Figure S1) this site is highly conserved within GTs. The modeled ternary complex shows Ser219 to be within hydrogen bonding distance of the oxygen of the sugar nucleotide phosphate (Figure 6A). Clearly this interaction is essential for catalysis, possibly in stabilizing charge development on the phosphate oxygen in the transition state. The proximity of the basic residue Lys225 in LanGT2 (Arg228 in UrdGT2) to the phosphates might additionally stabilize such charge development (Figure 6A).

DISCUSSION

Microorganisms produce a plethora of natural products that hold great promise for use as therapeutic agents. But before clinical applications are feasible, properties influencing the pharmacological or pharmacokinetic qualities of a molecule often need to be improved. This can be achieved by adding or deleting functional groups on the molecule through synthetic or enzyme based chemistry. However, the application of conventional chemistry is often thwarted by the stereochemical and functional group complexity of natural products, which typically demand multiple reaction steps to achieve one simple modification. Although enzymes in theory offer a solution to this problem, the inherently high specificity of most enzymes involved in natural product biosynthesis restricts their use in generating modified molecules. To circumvent this limitation, enzymes may be evolved or rationally engineered to create a variant that performs the desired biocatalytic reaction. In this article we present the engineering process to convert an O-glycosylating transferase to a C-glycosylating enzyme. To the best of our knowledge, the rational transformation of an O-GT toward a C-GT has not been documented before. To circumvent the considerable challenge of synthesizing the required NDP-deoxy sugar donor and acceptors for assaying LanGT2 variants, functional screening in this study was performed in vivo. Likewise, the use of computational design reduced the number of potential mutants to a manageable number. Visualized by modeled ternary structures we could show that the architecture of the LanGT2 acceptor-binding sites primarily determines the outcome of O- versus C-glycosylation. Furthermore, we were able to propose the underlying catalytic mechanisms, which in both reactions is driven by identical amino acids located in similar positions. The mechanism follows an S_N2 like reaction where Asp137 in LanGT2 (Asp137 in UrdGT2) is the general base that increases the phenolic character in C8 hydroxyl group. This increases the nucleophilicity of the C8 hydroxyl group, and through resonance likewise increases the nucleophilic character of the ortho carbon (C9) of the acceptor. However, the C8 hydroxyl is expected to be the more reactive site as reaction at C9 transiently destroys aromaticity. This is most simply illustrated by preference for O-alkylation of phenol in all solvents except those that strongly hydrogen bond to the hydroxyl group (Carey and Sundberg, 2007). Therefore GT catalyzed O-glycosylation and C-glycosylation is ultimately dependent on positioning the C8 hydroxyl or the C9 carbon of the acceptor relative to anomeric carbon of the glycosyl donor, as opposed to the relative nucleophilicities of the two sites. Interestingly, none of the LanGT2 variants produced a mixture of O- and C-glycosides. This supports the hypothesis that GT catalyzed C-glycosylation follows the direct Friedel-Crafts alkylation mechanism and excludes the possibility of an underlying O- to C-rearrangement. The success in converting LanGT2 to a strictly C-glycosylating enzyme can be attributed to replacing block C of LanGT2 with the analogous sites from UrdGT2. A model of this LanGT2 variant complexed with the glycosyl donor and acceptor predicts that these 10 amino acid substitutions collectively position the acceptor for C-glycosylation. Residues at positions 8 and 62 are the most crucial as they appear to have the most influence over glycosylation of the C8 hydroxyl or the C9 carbon. However, in the absence of an LanGT2 X-ray crystal structure, postulating an exact reason for the observed switch in regioselectivity is not possible, especially because block c is comprised of "second sphere" residues that do not directly interact with the substrates. Likewise, it is unclear why LanGT2.c and LanGT2.1, which differ only in positions 51 and 64, produce such dramatically different C-glycoside yields. Presumably the valine or leucine in position 51 is not as critical as the residue in position 64, because in LanGT2.c position 64 is occupied by threonine with a hydroxyl group available to form a hydrogen bond, whereas LanGT2.1 cannot do so as this position is occupied by methionine.

After succeeding in this O-GT to C-GT transformation the question arises if this engineering process can be adopted to other GTs. Some of the issues faced in aryl-C-bond chemistry have been summarized by Bililign et al. (2005). In particular, aryl C-glycosylation likely requires electron rich aromatic systems, such as those bearing hydroxyl, methoxy, or amine substituents, of which there are many examples. Our results suggest that glycosylation can be directed to the electron rich ortho and para positions by adjusting the position of the acceptor in the GT active site. Accordingly, further potential applications for engineering might be the conversion of the antipsoriatic agents anthralin and its 1,2 dihydroxy derivative into the corresponding 2-C- and 2-O-glycosides, which could reduce their inflammatory side effects. Our work suggests that even in the absence of a experimentally determined structural data, a combination of homology modeling and sequence analysis can be successfully applied to solving such engineering problems.

SIGNIFICANCE

Based on the C-glycosylating enzyme UrdGT2 we generated a structural LanGT2 model, followed by computational docking of substrates in the predicted LanGT2 active site. By comparing these predicted structural features with UrdGT2 and other C-glycosylating enzymes, 106 variants of LanGT2 were iteratively generated and tested for C-glycosylating activity. In addition to generating a highly active C-glycosylating variant of LanGT2, we have also gained considerable insight into the mechanistic features that control O- and C-glycosidic bond formation. A very notable feature of this effort is that the screening of C- versus O-glycosylating activity was conducted in vivo within the context of a complex biosynthetic pathway, which we believe is a timely contribution to the rapidly emerging field of combinatorial biosynthesis. Due to the difficulty in chemically synthesizing C-glycosides, as well as the potential of these compounds in drug development, enzymes involved to catalyze C-glycosidic bonds are a promising alternative to conventional chemistry.

EXPERIMENTAL PROCEDURES

Construction of *lanGT2-urdGT2* Chimeric Genes

As described previously the wild-type genes *lanGT2* and *urdGT2* were cloned into the *Hind*III and *Xba*I sites of pMun2 (Trefzer et al., 2000, 2001). In general, for generating the *lanGT2-urdGT2* chimeras *lanGT2.h1* to *lanGT2.h12* we made use of the gene flanking sites *Hind*III and *Xba*I as well as the restriction sites *Nco*I, *Sna*BI, *Bsr*GI, *Bsi*WI within *lanGT2* and *urdGT2*. PCR was used to incorporate silent mutations encoding the restriction sites *Nco*I, *Sna*BI, and *Bsi*WI at positions matching the *lanGT2* sequence. The required PCR primers are listed in Supplemental Experimental Procedures.

Generating *lanGT2-urdGT2* Hybrids and Alanine Substitutions by the PCR Megaprimer Method

In general the constructs were generated as described previously (Krauth et al., 2009). The applied protocol as well as the primers are listed in Supplemental Experimental Procedures.

Genetic Manipulation of *S. Cyanogenus* ΔlanGT2

Before transforming the mutant with the LanGT2-UrdGT2 hybrides encoded on pUWL (Doumith et al., 2000) expression vector, they were passed through *E. coli* ET 12567 cells (Flett et al., 1997) to obtain unmethylated DNA. The conjugation of *Streptomyces* strains as well as growth conditions are described in Supplemental Experimental Procedures.

Functionality Screening

Exconjugants were cultured for 72 hr in 100 ml baffled flasks (250 rpm, 28°C), each containing 20 ml SG medium (1% w/v soybean meal, 2% glucose, 0.2% $CaCO_3$, 0.2% valine) supplemented with hygromycin (50 μg/ml), spectinomycin (50 μg/ml), and phosphomycin (200 μg/ml). Landomycins and tetrangulol derivatives were extracted from liquid cultures with 1 volume of ethyl acetate. The O- and C-glycosides products were analyzed on a Agilent 1100 HPLC-ESI-MS system (Agilent Technologies, Waldbronn, Germany) equipped with UV-visible and a single quadrupole detector. HPLC separation was carried out on a Zorbax XDB-C8 column (5 μm, 4.6 × 150 mm) with a Zorbax SB-C8 precolumn (5 μm, 4.6 × 12.5 mm) from Agilent Technologies. Samples were dissolved in MeOH and resolved at flow rate of 0.7 ml/min with the following step gradient: 0–9 min 30% B, 9–16 min 35% B, 16–19 min 60% B, 19–23 min 85% B, 23–26 min 95% B (solvent A: 99.5% water, 0.5% acetic acid; solvent B: acetonitrile). The UV detection wavelength was 254 nm. ESI-MS (negative mode) was carried out using the following settings: drying gas flow, 12 ml/min; drying gas temperature, 350°C; nebulizer pressure, 50 psig; capillary voltage, 3000 V, mass range set to: 300–1150 Da. Fragmentation analysis of O- and C-glycosidic linkages was performed on an Agilent 1100 HPLC-ESI-MS/MS system equipped with UV-visible and a triple quadrupole detector. Analysis was performed in negative ion mode with a mass range from 100 to 600 m/z. Samples were resolved on a LiChrodphor 100RP 18 column (5 μm particle size, 2 × 125 mm) at 0.3 ml/min with the following step gradient: 0–5 min 0% B, 5–20 min 100% B, 20–25 min 100% B, 25–27 min 0% B, 27–35 min 0% B (solvent A: 99.5% water, 0.5% acetic acid; solvent B: acetonitrile).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and five tables and can be found with this article online at doi:10.1016/j.chembiol.2011.02.013.

ACKNOWLEDGMENTS

The authors thank all people of the laboratory and the Albert-Ludwigs-University Freiburg for funding.

Received: January 10, 2011
Revised: February 16, 2011
Accepted: February 22, 2011
Published: April 21, 2011

REFERENCES

Abad-Garcia, B., Garmon-Lobato, S., Berrueta, L.A., Gallo, B., and Vicente, F. (2008). New features on the fragmentation and differentiation of C-glycosidic flavone isomers by positive electrospray ionization and triple quadrupole mass spectrometry. Rapid Commun. Mass Spectrom. *22*, 1834–1842.

Baig, I., Kharel, M., Kobylyanskyy, A., Zhu, L., Rebets, Y., Ostash, B., Luzhetskyy, A., Bechthold, A., Fedorenko, V.A., and Rohr, J. (2006). On the acceptor substrate of C-glycosyltransferase UrdGT2: three prejadomycin C-Glycosides from an engineered mutant of Streptomyces globisporus 1912 DeltalndE(urdGT2). Angew. Chem. Int. Ed. Engl. *45*, 7842–7846.

Bililign, T., Hyun, C.G., Williams, J.S., Czisny, A.M., and Thorson, J.S. (2004). The hedamycin locus implicates a novel aromatic PKS priming mechanism. Chem. Biol. *11*, 959–969.

Bililign, T., Griffith, B.R., and Thorson, J.S. (2005). Structure, activity, synthesis and biosynthesis of aryl-C-glycosides. Nat. Prod. Rep. *22*, 742–760.

Bolam, D.N., Roberts, S., Proctor, M.R., Turkenburg, J.P., Dodson, E.J., Martinez-Fleites, C., Yang, M., Davis, B.G., Davies, G.J., and Gilbert, H.J. (2007). The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. Proc. Natl. Acad. Sci. USA *104*, 5336–5341.

Carey, F.A., and Sundberg, R.J. (2007). Advanced Organic Chemistry-Part A: Structure and Mechanisms, Fifth Edition (New York: Springer-Verlag).

Davis, B.G. (2007). Enzymatic glycosynthesis GeTs better. Nat. Chem. Biol. *3*, 604–605.

Doumith, M., Weingarten, P., Wehmeier, U.F., Salah-Bey, K., Benhamou, B., Capdevila, C., Michel, J.M., Piepersberg, W., and Raynal, M.C. (2000). Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in Saccharopolyspora erythraea. Mol. Gen. Genet. *264*, 477–485.

Durr, C., Hoffmeister, D., Wohlert, S.E., Ichinose, K., Weber, M., Von Mulert, U., Thorson, J.S., and Bechthold, A. (2004). The glycosyltransferase UrdGT2 catalyzes both C- and O-glycosidic sugar transfers. Angew. Chem. Int. Ed. Engl. *43*, 2962–2965.

Faust, B., Hoffmeister, D., Weitnauer, G., Westrich, L., Haag, S., Schneider, P., Decker, H., Kunzel, E., Rohr, J., and Bechthold, A. (2000). Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the angucycline antibiotic urdamycin A in Streptomyces fradiae Tu2717. Microbiology *146*, 147–154.

Flett, F., Mersinias, V., and Smith, C.P. (1997). High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes. FEMS Microbiol. Lett. *155*, 223–229.

Heddle, J., and Maxwell, A. (2002). Quinolone-binding pocket of DNA gyrase: role of GyrB. Antimicrob. Agents Chemother. *46*, 1805–1815.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51–59.

Hohne, M., Schatzle, S., Jochens, H., Robins, K., and Bornscheuer, U.T. (2010). Rational assignment of key motifs for function guides in silico enzyme identification. Nat. Chem. Biol. *6*, 807–813.

Hoffmeister, D., Ichinose, K., Domann, S., Faust, B., Trefzer, A., Drager, G., Kirschning, A., Fischer, C., Kunzel, E., Bearden, D., et al. (2000). The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster. Chem. Biol. *7*, 821–831.

Hoffmeister, D., Ichinose, K., and Bechthold, A. (2001). Two sequence elements of glycosyltransferases involved in urdamycin biosynthesis are responsible for substrate specificity and enzymatic activity. Chem. Biol. *8*, 557–567.

Hoffmeister, D., Wilkinson, B., Foster, G., Sidebottom, P.J., Ichinose, K., and Bechthold, A. (2002). Engineered urdamycin glycosyltransferases are broadened and altered in substrate specificity. Chem. Biol. *9*, 287–295.

Hoffmeister, D., Drager, G., Ichinose, K., Rohr, J., and Bechthold, A. (2003). The C-Glycosyltransferase UrdGT2 is unselective toward d- and l-configured nucleotide-bound rhodinoses. J. Am. Chem. Soc. *125*, 4678–4679.

Krauth, C., Fedoryshyn, M., Schleberger, C., Luzhetskyy, A., and Bechthold, A. (2009). Engineering a function into a glycosyltransferase. Chem. Biol. *16*, 28–35.

Love, J.C. (2010). Making antibodies from scratch. Nat. Biotechnol. *28*, 1176–1178.

Luzhetskyy, A., Taguchi, T., Fedoryshyn, M., Durr, C., Wohlert, S.E., Novikov, V., and Bechthold, A. (2005). LanGT2 Catalyzes the First Glycosylation Step during landomycin A biosynthesis. ChemBioChem *6*, 1406–1410.

Luzhetskyy, A., Mendez, C., Salas, J.A., and Bechthold, A. (2008). Glycosyltransferases, important tools for drug design. Curr. Top. Med. Chem. *8*, 680–709.

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.

Mulichak, A.M., Losey, H.C., Lu, W., Wawrzak, Z., Walsh, C.T., and Garavito, R.M. (2003). Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. Proc. Natl. Acad. Sci. USA *100*, 9238–9243.

Mulichak, A.M., Lu, W., Losey, H.C., Walsh, C.T., and Garavito, R.M. (2004). Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. Biochemistry *43*, 5170–5180.

Olano, C., Mendez, C., and Salas, J.A. (2010). Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis. Nat. Prod. Rep. *27*, 571–616.

Park, S.H., Park, H.Y., Sohng, J.K., Lee, H.C., Liou, K., Yoon, Y.J., and Kim, B.G. (2009). Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening. Biotechnol. Bioeng. *102*, 988–994.

Petrounia, I.P., and Arnold, F.H. (2000). Designed evolution of enzymatic properties. Curr. Opin. Biotechnol. *11*, 325–330.

Ramos, A., Olano, C., Brana, A.F., Mendez, C., and Salas, J.A. (2009). Modulation of deoxysugar transfer by the elloramycin glycosyltransferase ElmGT through site-directed mutagenesis. J. Bacteriol. *191*, 2871–2875.

Reidhaar-Olson, J.F., and Sauer, R.T. (1988). Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. Science *241*, 53–57.

Shaaban, K.A., Stamatkin, C., Damodaran, C., and Rohr, J. (2011). 11-Deoxylandomycinone and landomycins X-Z, new cytotoxic angucyclin (on)es from a Streptomyces cyanogenus K62 mutant strain. J. Antibiot. (Tokyo) *64*, 141–150.

Schmid, A., Dordick, J.S., Hauer, B., Kiener, A., Wubbolts, M., and Witholt, B. (2001). Industrial biocatalysis today and tomorrow. Nature *409*, 258–268.

Thibodeaux, C.J., Melancon, C.E., and Liu, H.W. (2007). Unusual sugar biosynthesis and natural product glycodiversification. Nature *446*, 1008–1016.

Trefzer, A., Hoffmeister, D., Kunzel, E., Stockert, S., Weitnauer, G., Westrich, L., Rix, U., Fuchser, J., Bindseil, K.U., Rohr, J., and Bechthold, A. (2000). Function of glycosyltransferase genes involved in urdamycin A biosynthesis. Chem. Biol. *7*, 133–142.

Trefzer, A., Fischer, C., Stockert, S., Westrich, L., Kunzel, E., Girreser, U., Rohr, J., and Bechthold, A. (2001). Elucidation of the function of two glycosyltransferase genes (lanGT1 and lanGT4) involved in landomycin biosynthesis and generation of new oligosaccharide antibiotics. Chem. Biol. *8*, 1239–1252.

Truman, A.W., Dias, M.V., Wu, S., Blundell, T.L., Huang, F., and Spencer, J.B. (2009). Chimeric glycosyltransferases for the generation of hybrid glycopeptides. Chem. Biol. *16*, 676–685.

Turner, N.J. (2009). Directed evolution drives the next generation of biocatalysts. Nat. Chem. Biol. *5*, 567–573.

Unligil, U.M., and Rini, J.M. (2000). Glycosyltransferase structure and mechanism. Curr. Opin. Struct. Biol. *10*, 510–517.

Walsh, C.T. (2002). Combinatorial biosynthesis of antibiotics: challenges and opportunities. ChemBioChem *3*, 125–134.

Walsh, C.T., Chen, H., Keating, T.A., Hubbard, B.K., Losey, H.C., Luo, L., Marshall, C.G., Miller, D.A., and Patel, H.M. (2001). Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. Curr. Opin. Chem. Biol. *5*, 525–534.

Weatherman, R.V., Mortell, K.H., Chervenak, M., Kiessling, L.L., and Toone, E.J. (1996). Specificity of C-glycoside complexation by mannose/glucose specific lectins. Biochemistry *35*, 3619–3624.

Weymouth-Wilson, A.C. (1997). The role of carbohydrates in biologically active natural products. Nat. Prod. Rep. *14*, 99–110.

Williams, G.J., and Thorson, J.S. (2008). A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. Nat. Protoc. *3*, 357–362.

Williams, G.J., Zhang, C., and Thorson, J.S. (2007). Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. Nat. Chem. Biol. *3*, 657–662.

Williams, G.J., Gantt, R.W., and Thorson, J.S. (2008a). The impact of enzyme engineering upon natural product glycodiversification. Curr. Opin. Chem. Biol. *12*, 556–564.

Williams, G.J., Goff, R.D., Zhang, C., and Thorson, J.S. (2008b). Optimizing glycosyltransferase specificity via "hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization. Chem. Biol. *15*, 393–401.

Zhu, L., Ostash, B., Rix, U., Nur-e-Alam, M., Mayers, A., Luzhetskyy, A., Mendez, C., Salas, J.A., Bechthold, A., Fedorenko, V., and Rohr, J. (2005). Identification of the function of gene lndM2 encoding a bifunctional oxygenase-reductase involved in the biosynthesis of the antitumor antibiotic landomycin E by Streptomyces globisporus 1912 supports the originally assigned structure for landomycinone. J. Org. Chem. *70*, 631–638.

Zhu, L., Luzhetskyy, A., Luzhetska, M., Mattingly, C., Adams, V., Bechthold, A., and Rohr, J. (2007). Generation of new landomycins with altered saccharide patterns through over-expression of the glycosyltransferase gene lanGT3 in the biosynthetic gene cluster of landomycin A in Streptomyces cyanogenus S-136. ChemBioChem *8*, 83–88.

