Chemistry & Biology
Article

Engineering a Function into a Glycosyltransferase

Christine Krauth,[1] Marta Fedoryshyn,[1] Christian Schleberger,[1] Andriy Luzhetskyy,[1] and Andreas Bechthold[1,*]
[1]Albert-Ludwigs-Universität, Institut für Pharmazeutische Wissenschaften, Stefan-Meier-Strasse 19, 79104 Freiburg, Germany
*Correspondence: andreas.bechthold@pharmazie.uni-freiburg.de
DOI 10.1016/j.chembiol.2008.12.003

SUMMARY

As glycosyltransferases found in nature often show distinct substrate specificity, glycosyltransferase engineering is an important research field. In this work, we were able to introduce an activity into a glycosyltransferase involved in natural product (landomycin E) biosynthesis. This was achieved by recognizing hot spot amino acids in glycosyltransferases which are strongly involved in determining substrate specificity.

INTRODUCTION

The pharmaceutical activity of many natural products, among them valuable antibiotics and anticancer therapeutics, depends on regio- and stereospecifically attached sugar moieties (Newman et al., 2003; Luzhetskyy et al., 2008). The attachment of these sugars is catalyzed by glycosyltransferases (GTs). The acceptor substrates of GTs vary widely. The donor substrate is almost always an activated sugar, with the most common activated species being NDP sugars (Méndez et al., 2008; Salas and Méndez, 2007). GTs display a high level of sequence diversity (Liang and Qiao, 2007) but also great structural similarity. More than 16,000 GT sequences have been described in the databases. They have been classified into 91 families by amino acid sequence similarities. GTs are all globular proteins falling into two structural superfamilies termed GT-A and GT-B which each have an N-terminal and a C-terminal domain (Breton et al., 2006). The enzymes of the GT-A fold have two dissimilar domains. The N-terminal domain consisting of several β strands that are each flanked by α helices (Rossmann folds) recognizes the sugar nucleotide donor. The C-terminal domain, which consists largely of mixed β sheets, contains the acceptor-binding site. Enzymes of the GT-B fold contain two similar Rossmann-like folds. The N-terminal domain provides the acceptor-binding site, whereas the C-terminal domain is responsible for binding the sugar donor. In both types of fold, domains are connected by a linker region and the active site is located between the two domains.

Most GTs involved in natural product biosynthesis belong to the GT-B family. As these GTs are often involved in late biosynthetic steps, they are ideal candidates for biotech approaches to generate novel unnatural products. The substrate specificity of GTs provides a critical issue in natural product diversification, and scientists have started recently to broaden the specificity by genetic engineering.

One of the first experiments on the engineering of natural product GTs was performed with UrdGT1b and UrdGT1c, both involved in urdamycin biosynthesis. The enzymes show different specificities for both nucleotide sugar and acceptor substrates, but share a surprisingly high number (91%) of identical amino acids. A region of 31 amino acids within the N-terminal part of UrdGT1b and UrdGT1c differing in 18 positions was identified to control the substrate specificity of both enzymes. It was possible to alter selectivity for both the donor and the acceptor by mutating this region. Finally, a novel compound carrying an unusual branched saccharide chain was generated (Hoffmeister et al., 2001, 2002). This was indeed one of the first examples describing the generation of a GT with novel aglycon specificity. Further successful examples of GT engineering have been described, most of them by the group of Prof. Dr. J. Thorson. This group worked recently on the GT OleD involved in oleandomycin glycosylation. One mutant of OleD showed improved sugar donor specificity and another mutant showed 200- to 300-fold improved activity toward novobiocic acid (Williams and Thorson, 2008; Williams et al., 2007, 2008).

Model systems used in this study are GTs from different strains involved in the biosynthesis of landomycins, an angucycline type of natural product. Landomycin A containing a long oligosaccharide chain with four D-olivose and two L-rhodinose moieties is produced by *Streptomyces cyanogenus* S136, and landomycin E with two D-olivose and one L-rhodinose moieties is produced by *S. globisporus* 1912 (Figure 1). Detailed studies on the biosynthesis of the landomycin A hexasaccharide side chain revealed that four glycosyltransferases are responsible for its formation (Luzhetskyy et al., 2005). A special highlight of this study is the GT LanGT1 (D-olivosyltransferase) catalyzing the attachment of the second and the fifth sugars during landomycin A biosynthesis (iterative acting GT). In *S. globisporus*, LndGT1 catalyzes the attachment of the second sugar during landomycin E biosynthesis (Ostash et al., 2004). Recently, a *lanGT1* knockout mutant (*S. cyanogenus* S136 Δ *lan*GT1) was created by introducing a 900 bp deletion into the gene (A. Erb, C.K., A.L., and A.B., unpublished data). Whereas *S. cyanogenus* S136 produced landomycins A (containing six sugars), landomycin B (containing five sugars), and landomycin D (containing two sugars) (Figures 1 and 2A), *S. cyanogenus* S136 Δ *lan*GT1 produced landomycin I and 11-deoxy-landomycin (both containing one sugar in the side chain) (Figures 1 and 2B). Further products were landomycin L and 11-deoxy-landomycin L (Figure 1). When *lanGT1* was expressed in *S. cyanogenus*, S136 Δ *lan*GT1 wild-type production could be restored. Most of the products were landomycin A and landomycin B. In addition, 11-deoxy-landomycin K (Figure 1),

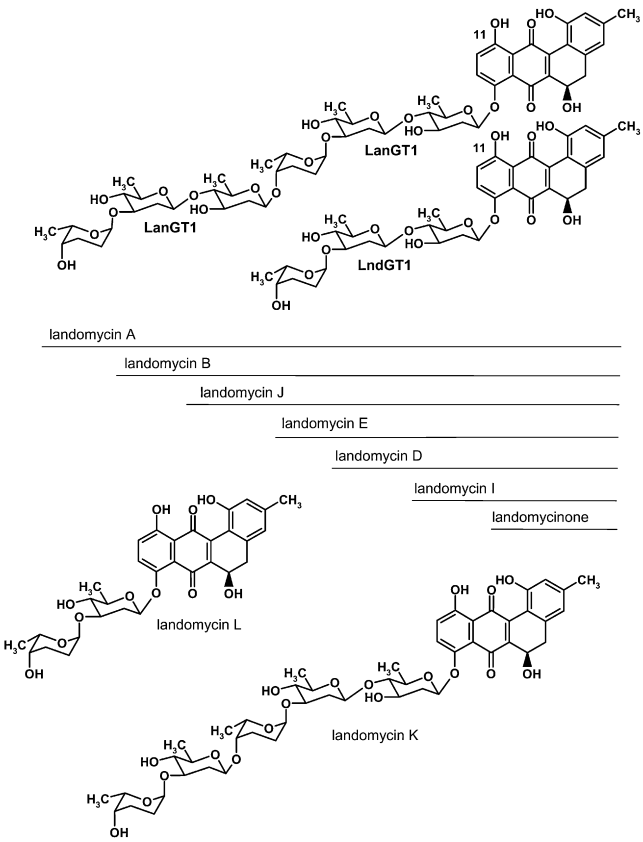


Figure 1. Structures of Various Landomycin Derivatives
LanGT1 is involved in landomycin A biosynthesis by attaching the second and fifth sugars, and LndGT1 is involved in landomycin E biosynthesis by attaching the second sugar.

11-deoxy-landomycin B, 11-deoxy-landomycin A, landomycin J (containing four sugars), landomycin I, and landomycin K were produced (Figures 1 and 2C). When *lndGT1* was expressed in the mutant landomycin J, 11-deoxy-landomycin J, landomycin K, and 11-deoxy-landomycin K (Figure 1) were the major products. Additional compounds generated were landomycin D, 11-deoxy-landomycin D, and 11-deoxy-landomycin I (Figures 1 and 2D). Neither landomycin A nor landomycin B were detectable. The same result was obtained when *lndGT1* was expressed either on a replicative or on an integrative vector. These data showed that LndGT1, which shares 74.8% amino acid identity with LanGT1, does not have the ability to act iteratively (A. Erb, C.K., A.L., and A.B., unpublished data). In this current work, studies were conducted to introduce new aglycon specificity into LndGT1.

RESULTS

S. globisporus Containing *lanGT3* Accumulates a Landomycin Derivative Consisting of Four Sugars in the Side Chain

After expression of *lanGT3* in *S. globisporus*, landomycin J containing four sugars (D-olivose-D-olivose-L-rhodinose-D-olivose) was detected as the major product by HPLC-UV and HPLC-MS. No landomycin A and no landomycin B could be detected. These data clearly prove that LndGT1 is not able to catalyze the attachment of the second sugar, and thus it does not belong to the iteratively acting GTs like LanGT1.

Sequence Comparison of LanGT1 and LndGT1 and Construction of Chimeric Genes

The landomycin A biosynthetic gene cluster of *S. cyanogenus* S136 and the landomycin E biosynthetic gene cluster of *S. globisporus* 1912 have been cloned and sequenced. Sequence comparison of LanGT1 and LndGT1 revealed that the enzymes share 74.8% identical amino acids. Most of the differences are located in the N-terminal part between amino acids 50 and 170. Modeling studies of LanGT1 using UrdGT2 as a template demonstrated that these amino acids are located in the N-terminal domain including α helices α4a, α4b, and α5 and β sheet β5 (Figure 3).

To determine which region of LanGT1 would govern iterative functionality, we constructed a set of chimeric GT genes in which differently sized and positioned elements of one parental gene had been replaced by the equivalent of the second parental gene and vice versa. Fragments to be swapped around were named v1a (aa 1–63), v1b (aa 64–82), v2a (aa 83–105), v2b (aa 106–140), v3 (aa 141–208), and c (aa 209–391) (Figure 3). As no convenient restriction sites were available, a megaprimer technology was developed (primers used are listed in Table 1). Ten chimeric genes were constructed representing different combinations of regions of *lanGT1* and *lndGT1* (Figure 4).

Expression of Chimeric Genes in Mutant *S. cyanogenus* S136 Δ *lan*GT1

S. cyanogenus S136 Δ *lan*GT1 was used as an in vivo host to test the functionality of the expressed chimeric genes. Landomycin A and B derivatives indicated iterative functionality and landomycins D, J, and K indicated noniterative functionality. The amount of landomycins without a second D-olivose at the second position of the side chain (landomycins I and L) indicated nonfunctionality, and was a measure for the general "fitness" of the expressed enzyme (Figures 1 and 4).

Expression of the chimeric gene H1 led to the production of landomycins A and B, indicating iterative functionality of the enzyme (Figure 4). Expression of H2 led to the production of compounds with two, three, and four sugars, but the production was reduced, indicating H2 to be noniterative functional with reduced overall activity. These results revealed that the N-terminal part is important for recognition and binding of the aglycon. Amino acids responsible for the iterative functionality of LanGT1 are located in the N-terminal part (aa 1–208). Thus, both enzymes belong to the GT-B superfamily of GTs. The successful generation of H1 and H2, which both were made by domain swapping (generation of GTs containing N- and C-terminal domains from different sources), encouraged us to generate further chimeric genes (see above). We focused the swapping experiments on the N-terminal part of both enzymes. H4b and H5a showed LanGT1-like activity and H4a and H5b showed LndGT1-like activity, indicating that regions v1a and v3 are not important for iterative functionality. Thus, regions v1b, v2a, and v2b are involved in defining LanGT1-like

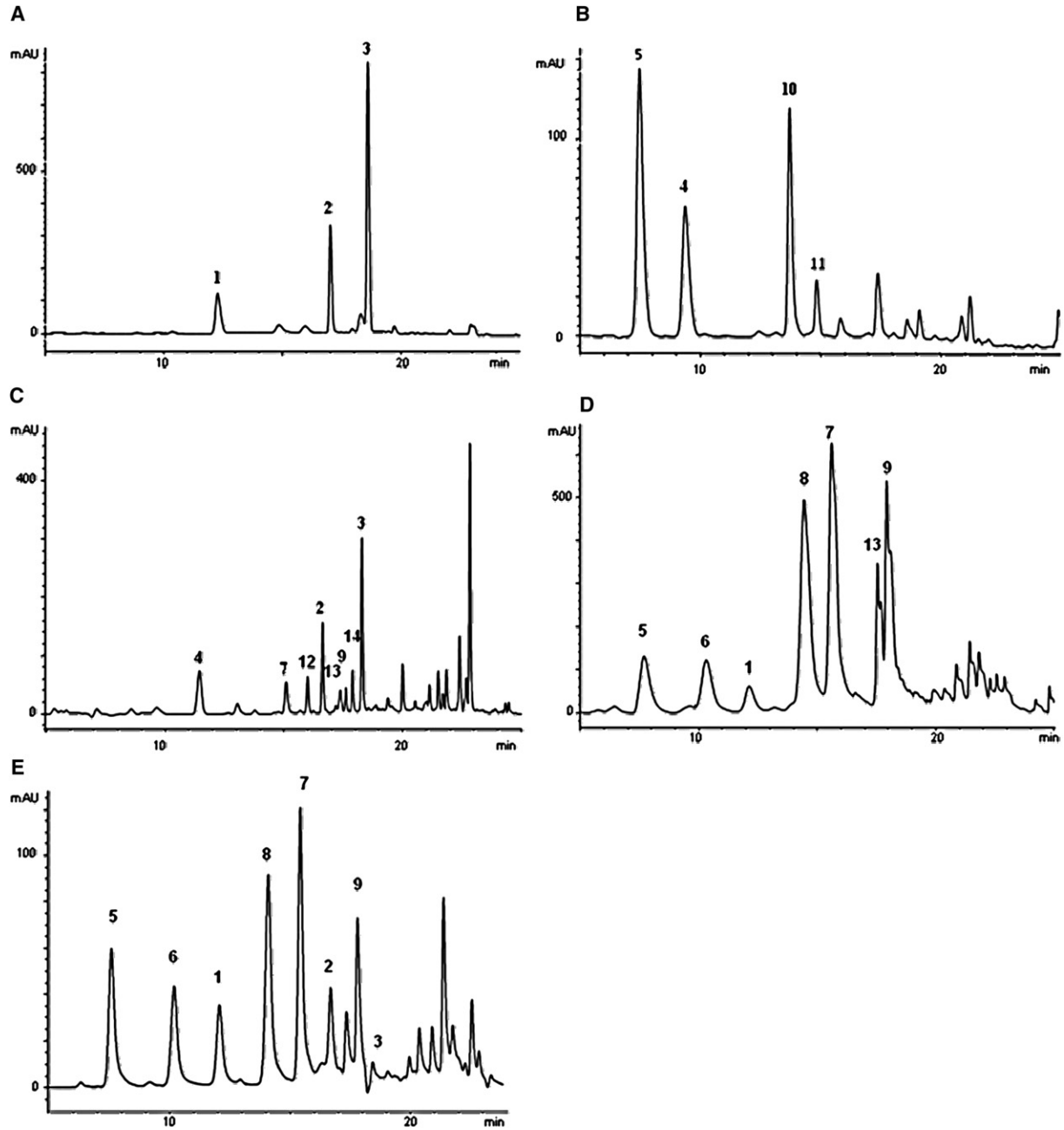


Figure 2. Selected HPLC Chromatograms

(A) *S. cyanogenus* S136.

(B) *S. cyanogenus* Δ lanGT1.

(C) *S. cyanogenus* Δ lanGT1 x lanGT1.

(D) *S. cyanogenus* Δ lanGT1 x lndGT1.

(E) *S. cyanogenus* Δ lanGT1 x lndGT1mut10.

1: landomycin D; **2**: landomycin B; **3**: landomycin A; **4**: landomycin I; **5**: 11-deoxy-landomycin I; **6**: 11-deoxy-landomycin D; **7**: landomycin J; **8**: 11-deoxy-landomycin J; **9**: landomycin K; **10**: 11-deoxy-landomycin L; **11**: landomycin L; **12**: 11-deoxy-landomycin B; **13**: 11-deoxy-landomycin K; **14**: 11-deoxy-landomycin A.

properties. Additional swaps were created that focused on exchanging these regions in the N-terminal part. LanGT1-like activity was strongly reduced in H3a, H3b, and H6b and was not detectable in H6a. These data indicated that iterative functionality of LanGT1 is controlled by up to 20 amino acids which are located at different regions of the protein. We also deduced from these results interactions of the amino acids located in regions v1b, v2a, and v2b.

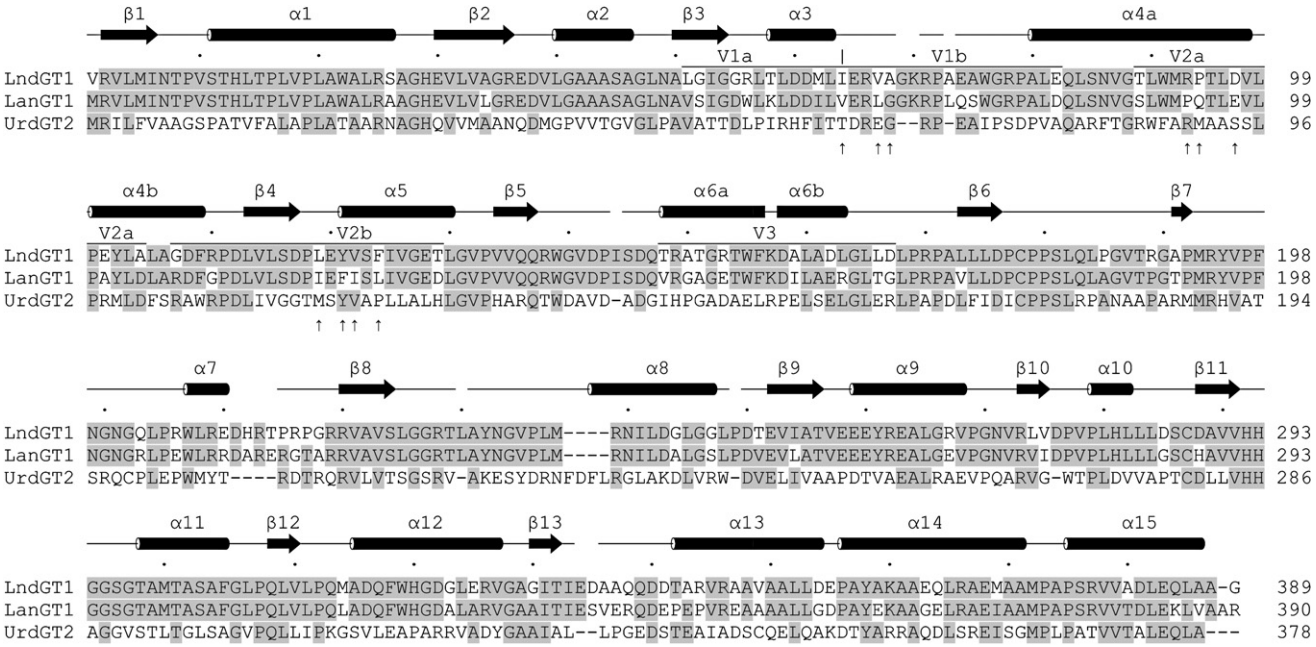


Figure 3. Structure-Based Sequence Alignment of LanGT1, LndGT1, and UrdGT2

The secondary structures are from UrdGT2. Regions v1a, v1b, v2a, and v2b are indicated. Amino acids which are present in at least two of the three enzymes are highlighted. Arrows are pointing to amino acids which have been changed in LndGT1 to introduce iterative functionality (mutant LndGT1-mut10).

Contribution of Various Amino Acids to LanGT1-like Activity

The contribution of one or more of the 20 different amino acids in regions v1b, v2a, and v2b to the iterative functionality of LanGT1 was evaluated by carrying out a set of site-directed mutagenesis studies. Regions were divided into subregions as shown in Figure 5. To assess the influence of the amino acids, we introduced mutations into different chimeric genes/enzymes in such a way that amino acids originating from LndGT1 were replaced by amino acids originating from LanGT1. Hybrid H3a was used to investigate the influence of region v1b, and thus mutants H3aA, H3aB, and H3aC were generated. H3b and H6b were used for region v2b, and thus H3bH, H3bI, H3bJ, H6bH, H6bI, and H6bJ were generated. And hybrid H3bI was used for region v2a, and thus mutants H3bID, H3bIE, H3bIF, and H3bIG were generated (Figure 5).

H3aA, H3aB, and H3aC showed improved LanGT1 activity in comparison to H3a, indicating that not one single amino acid in this region is essential for LanGT1 activity. As after expression of H3aB, landomycins I and L were produced in significant amounts, we believe that the B region is also involved in stabilizing the structure of LanGT1.

LanGT1-like activity was comparable using H6b and H6bH, indicating that amino acids R(107) and G(110) in subregion H of v2b are not involved in the iterative functionality of LanGT1. LanGT1 activity was nearly restored in H6bI and H6bJ, indicating that subregions I and J of v2b are involved in the iterative activity of LanGT1. The exchange of these regions in mutant H3b could not restore LanGT1-like activity. H6b and H3b differ in amino acid composition in v2a and v3. This result might indicate that in H3bI and H3bJ, essential protein-protein interactions between v2a and v2b are affected. Interestingly, LanGT1-like activity was even strongly reduced in H3bJ.

To gain more information about the influence of region v2a, we decided to use H3bI as a template for the generation of further mutants. The results obtained from mutants H3bID, H3bIE, H3bIF, and H3bIG show that regions D, F, and G are less important for iterative functionality than region E. Mutant H3bIE restored landomycin A and B production in *S. cyanogenus* S136 *Δ lan*GT1 nearly to that in LanGT1.

In addition, we generated two additional enzymes containing mutations in the HHGGSGT motif (HHVG and HHCG). This motif is considered to be important for sugar binding (Mulichak et al., 2004). Both enzymes were not enzymatically active.

After analysis of all the available information, we determined ten amino acids which are most essential for the iterative functionality of LanGT1. Three amino acids (V64, L67, G68) were chosen from region v1b, three (P93, Q94, E97) from v2a, and four (I119, F121, I122, L124) from v2b (Figure 5). These amino acids were introduced into LndGT1 at the corresponding positions, resulting in LndGT1-mut10. Expression of the LndGT1-mut10 gene in *S. cyanogenus* ΔLanGT1 led to the production of landomycin B and small amounts of landomycin A (Figure 2E).

DISCUSSION

GT engineering has become an important research field in obtaining novel GTs with altered substrate specificity. Successful engineering requires the existence of at least two GTs with a high number of identical amino acids, the crystal structure of a GT, or the possibility of using a direct evolution approach. Examples in engineering GTs of the superfamily GT-B are rare because just a few GT partners with very similar sequences are known, just a few crystallographic analyses have been performed (structural information is now available for 17 distinct

Table 1. Primer Used to Generate Chimeric Genes

Construct	Primer	Sequence
*lan*GT1	LanNF	5′-CGAGGTCGACGGTATCGATAAGCTTC AATTGGCAGGAGACGC-3′
	LanCR	5′-GCGGTCGGTCTAGACGGTCTCAGCGG GCCGCGACC-3′
*lnd*GT1	LndNF	5′-TGATTACGCCAATTGCGCGCAATTAAC CCTCACTAAAGGGA-3′
	LndCR	5′-TAATACGACTCACTATAGGGCGAATC TAGAGGTACCGGG-3′
H1	LanNR	5′-GGGCGCGGGGTCCGGTGGTCCTCCCG GAGCCACTCGGGCAGTCG-3′
	LndCF	5′-CGACTGCCCGAGTGGCTCCGGGAGGA CCACCGGACCCCGCGCCC-3′
H2	LndNR	5′-GTTCCGCGCTCCCGGGCATCACGCCGC AGCCATCGGGGCAGCTG-3′
	LanCF	5′-CAGCTGCCCCGATGGCTGCGGCGTGAT GCCCGGGAGCGCGGAAC-3′
H3a	LndGT1 RH3a	5′-GGGAGCCGACGTTGGAGAGCTGCTCC AGGGCCGGCCGGCCCCA-3′
	LanGT1 FH3a	5′-TGGGGCCGGCCGGCCCTGGAGCAGCT CTCCAACGTCGGCTCCC-3′
H3b	LanGT1 RH3b	5′-GCGTTCCGACGTTGGAGAGCTGGTCC AGGGCCGGCCTGCCCCA-3′
	LndGT1 FH3b	5′-TGGGGCAGGCCGGCCCTGGACCAGCT CTCCAACGTCGGAACGC-3′
H4a	LndGT1 RH4a	5′-CTGGTCCGAGATGGGGTCGACGCCCC AGCGCTGCTGGACGACG-3′
	LanGT1 FH4a	5′-CGTCGTCCAGCAGCGCTGGGGCGTCG ACCCCATCTCGGACCAG-3′
H4b	LanGT1 RH4b	5′-CTGGTCCGAGATGGGGTCGACACCCC AGCGCTGCTGCACCACG-3′
	LndGT1 FH4b	5′-CGTGGTGCAGCAGCGCTGGGGTGTCG ACCCCATCTCGGACCAG-3′
H5a	LndGT1 RH5a	5′-GCTTGCCGCCGAGCCGCTCGACCAGCA TGTCGTCCAGGGTGAGC-3′
	LanGT1 FH5a	5′-GCTCACCCTGGACGACATGCTGGTCGA GCGGCTCGGCGGCAAGC-3′
H5b	LanGT1 RH5b	5′-GCTTGCCCGCCACCCGCTCGATCAGTA TGTCGTCCAGCTTGAGC-3′
	LndGT1 FH5b	5′-GCTCAAGCTGGACGACATACTGATCG AGCGGGTGGCGGGCAAGC-3′
H6a	LndGT1 RH6a	5′-GGTCGGGTCCGAAGTCGCGGGCGAGC GCGAGGTACTCCGGCAGCACG-3′
	LanGT1 FH6a	5′-CGTGCTGCCGGAGTACCTCGCGCTCGC CCGCGACTTCGGACCCGACC-3′
H6b	LanGT1 RH6b	5′-GGTCGGGACGGAAGTCGCCGGCCAGG TCCAGGTACGCGGGCAGTACC-3′
	LndGT1 FH6b	5′-GGTACTGCCCGCGTACCTGGACCTGGC CGGCGACTTCCGTCCCGACC-3′
H3bH	HIIIbAF	5′-CTCGCGCTCGCCCGCGACTTCGGACCC GACCTCG-3′
	HIIIbAR	5′-CGAGGTCGGGTCCGAAGTCGCGGGCG AGCGCGAG-3′
H3bI/H4bI	HVIbBF	5′-CCTGTCGGATCCGATCGAGTTCATCTC GCTGATCGTCGGGGAGAC-3′
	HVIbBR	5′-GTCTCCCCGACGATCAGCGAGATGAAC TCGATCGGATCCGACAGG-3′
H3bJ/H4bJ	HVIbCF	5′-CTTCATCGTCGGGGAGGATCTCGGCGT CCCCGTC-3′
	HVIbCR	5′-GACGGGGACGCCGAGATCCTCCCCGA CGATGAAG-3′
H6bH	HVIbAF	5′-GGACCTGGCCCGCGACTTCGGACCCG ACCTCG-3′
	HVIbAR	5′-CGAGGTCGGGTCCGAAGTCGCGGGCC AGGTCC-3′
H3bID	v2aAF	5′-CTCTCCAACGTCGGATCCCTCTGGATG CGCCCGAC-3′
	v2aAR	5′-GTCGGGCGCATCCAGAGGGATCCGAC GTTGGAGAG-3′
H3bIE	v2aBF	5′-GAACGCTCTGGATGCCCCAGACGCTG GAGGTGCTGCCGGAG-3′
	v2aBR	5′-CTCCGGCAGCACCTCCAGCGTCTGGGG CATCCAGAGCGTTC-3′
H3bIF	v2aCF	5′-GGACGTGCTGCCGGCGTACCTGGACCT CGCCGGCGAC-3′
	v2aCR	5′-GTCGCCGGCGAGGTCCAGGTACGCCG GCAGCACGTCC-3′
H3bIG	HIIIbGF	5′-GCCGGAGTACCTCGACCTCGCCGGCGAC-3′
	HIIIbGR	5′-GTCGCCGGCGAGGTCGAGGTACTCCG GC-3′
H3aA	HIIIaAF	5′-CGACATGCTGGTCGAGCGGCTCGGCG GCAAGCGCC-3′
	HIIIaAR	5′-GGCGCTTGCCGCCGAGCCGCTCGACCAGCATGTCG-3′
H3aB	HIIIaB	5′-GGCAAGCGCCCCTTGCAGTCCTGGGG CCGG-3′
H3aC	HIIIaCF	5′-CGGCCCTGGACCAGCTCTCCAACGTC-3′
	HIIIaCR	5′-GACGTTGGAGAGCTGGTCCAGGGCCG-3′

GT families), and because the availability of a suitable high-throughput screen to select for the desired activity is often lacking. In any case, successful examples have been described in all three areas (Hoffmeister et al., 2002; Williams and Thorson, 2008; Williams et al., 2007, 2008; Bolam et al., 2007; Hancock et al., 2006).

A very small number of GTs have been shown to act twice during the biosynthesis of a natural compound. Examples are

Name	Chimeric and wild-type GTs	Landomycins produced indicating:		
		LndGT1-like activity [%] (mono-functional)	LanGT1-like activity [%] (bi-functional)	No activity [%]
LanGT1		51.9	37.8	10.3
LndGT1		95.3	<0.1	4.7
H1		54.5	35.1	10.4
H2		72.8	<0.1	27.2
H3a		95.2	0.4	4.4
H3b		79.1	5.8	15.1
H4a		57.7	<0.1	42.3
H4b		47.5	36.3	16.2
H5a		66.1	33.9	<0.1
H5b		58.5	<0.1	41.5
H6a		87.1	<0.1	12.9
H6b		26.9	0.6	72.5

Figure 4. Functionality of Chimeric GTs

S. cyanogenus S136 Δ *lan*GT1 was used as in vivo host to test the functionality of the expressed chimeric genes. The amount of landomycin A and B derivatives indicates LanGT1-like activity, the amount of landomycins D, J, and K indicates LndGT1-like activity, and the amount of landomycins I and L indicates no activity.

LanGT1 and LanGT4 from *S. cyanogenus* S136 and AveBI from *S. avermitilis* (Zhang et al., 2006).

As LanGT1 catalyzes key steps during the biosynthesis of the hexasaccharide antibiotic landomycin A (the attachment of the second and fifth sugars), we focused our research on this interesting enzyme. LndGT1, a GT from a different strain involved in landomycin E biosynthesis attaching the second sugar (dTDP-D-olivose) during the biosynthesis, was shown not to have the ability to act twice in this study. As the amino acid sequences of LanGT1 and LndGT1 were rather similar (74.8% identical amino acids), we decided that hybrids of these enzymes could be useful to find sequence regions responsible for the difference in substrate specificity.

An in vivo test system had to be developed due to the lack of availability of the donor substrate of LanGT1, dTDP-D-olivose. Thus, all hybrids and mutants had to be tested in vivo in the *lanGT1* deletion mutant, a time- and effort-consuming process.

As GTs seem to have independent structural domains, each having distinct acceptor and donor substrate binding functions, we started our work with the generation of chimeric enzymes H1 and H2 consisting of the N- and C-terminal domains of both GTs. Both are examples of functionally active chimeric enzymes, indicating that possible amino acid contacts in the interface between the domains in LanGT1 and LndGT1 are conserved. Although the generation of chimeric enzymes has been described (e.g., Fischbach et al., 2007), this is one of the first reports on domain swapping using GTs involved in antibiotic biosynthesis. Another example comes from the group of Prof. B.G. Kim, which published a chimeric library with N-terminal fragments of KanF involved in kanamycin biosynthesis and C-terminal fragments of GtfE involved in vancomycin biosynthesis (Park et al., 2009).

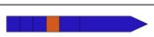
In our study, further swapping experiments isolated a 77 amino acid segment containing 20 amino acids which are different between LanGT1 and LndGT1. This segment was divided into three regions, v1b, v2a, and v2b.

The X-ray structures of the antibiotic GTs GtfA (Mulichak et al., 2003), GtfB (Mulichak et al., 2001), GtfD (Mulichak et al., 2004), UrdGT2 (Mittler et al., 2007), OleD (Bolam et al., 2007), OleI (Bolam et al., 2007), and CalG3 (Zhang et al., 2008) were recently solved. LanGT1 and LndGT1 were the closest related to UrdGT2, which is a C-C GT involved in the biosynthesis of urdamycin A (Dürr et al., 2004), an angucycline similar to landomycin A. Because LndGT1 and LanGT1 share up to 28% identical and 48% similar residues with UrdGT2, the structure of UrdGT2 allows for structural interpretation of the functional effect of the mutations discovered in this study.

Region v1b is present in a loop region between helices α3 and α4a (nomenclature taken from UrdGT2) that is hypervariable in other GT-B-fold GTs, and it has been stated several times that this loop constitutes a part of the acceptor-binding site (Ostash et al., 2004; Mittler et al., 2007). Mutations in β3 and α3 have been performed in the GTs UrdGT1b and UrdGT1c involved in urdamycin A biosynthesis, leading to altered substrate specificity (Hoffmeister et al., 2001, 2002). And recently, Thorson and coworkers determined P67 in this loop in the oleandomycin GT OleD as an important amino acid. A triple mutant containing inter alia P67T showed increased rates of glycoside formation for different acceptor substrates in comparison to the wild-type enzyme (Williams et al., 2008).

Region v2a is part of α helices (α4a and α4b in UrdGT2) which are also variable in other GTs. In contrast to UrdGT2, LangGT1 and LndGT2 contain a proline residue at position 93 and 94, respectively, that will probably introduce a bend into helix α4a at this position. This may result in better accessibility of region v2a, which would explain its impact on substrate specificity. Interestingly, exchange of amino acids in this region, as performed in H3bIE, resulted in improved LanGT1-like activity.

Region v2b consists of a β sheet (β4) and the α helix α5 with some variability between different GTs. It is located at the bottom of the active site of GTs. Mutations at this position may directly change the substrate-enzyme interaction and/or distort the geometry of the active site, which could explain

Name	Chimeric GTs	Landomycins produced indicating		
		LndGT1-like activity [%] (mono-functional)	LanGT1-like activity [%] (bi-functional)	No activity [%]
H3a		95.2	0.4	4.4
H3aA	I_{64} V_{67} A_{68} → V_{64} L_{67} G_{68}	73.4	26.7	<0.1
H3aB	A_{73} E_{74} A_{75}→ L_{73} Q_{74} S_{75}	77.3	13	9.6
H3aC	D_{82} → E_{82}	85.8	14.3	<0.1
H3b		79.1	5.8	15.1
H3bH	G_{107} R_{110} → R_{107} G_{110}	76.6	1.7	21.7
H3bI	L_{119} Y_{121} V_{122} F_{124} → I_{119} F_{121} I_{122} L_{124}	70.9	6.8	22.3
H3bJ	T_{129} → D_{129}	39.0	0.5	60.5
H6b		26.9	0.6	72.5
H6bH	G_{107} R_{110} → R_{107} G_{110}	20.6	0.2	79.2
H6bI	L_{119} Y_{121} V_{122} F_{124} → I_{119} F_{121} I_{122} L_{124}	67.3	28.5	4.2
H6bJ	T_{129} → D_{129}	59.2	32.3	8.5
H3bI		70.9	6.8	22.3
H3bID	T_{89} → S_{89}	91.8	0.3	7.9
H3bIE	R_{93} P_{94} D_{97} → P_{93} Q_{94} E_{97}	41.9	34.9	23.2
H3bIF	E_{101} → A_{101}	42.5	<0.1	57.5
H3bIG	A_{104} → D_{104}	88.6	<0.1	11.4
LndGT1		95.3	<0.1	4.7
LndGT1-mut10	I_{64} V_{67} A_{68} R_{93} P_{94} D_{97} L_{119} Y_{121} V_{122} F_{124}→ V_{64} L_{67} G_{68} P_{93} Q_{94} E_{97} I_{119} F_{121} I_{122} L_{124}	72.6	17.4	9.7

Figures 5. Functionality of Mutated Chimeric GTs
GTs which were used as hosts for further mutations are shown as arrows. Mutations which were introduced are given below each GT (see Figure 4 for further information).

why H6bI and H6bJ showed improved LanGT1 activity in comparison to H6b.

We observed variability in the overall production amounts by around 30% when comparing three different clones harboring the same chimeric gene. But the overall production of landomycin derivatives was not dependent on the kind of chimeric GT expressed in the mutant. This indicates that LanGT1 and LndGT1 are not limiting factors of landomycin biosynthesis.

LanGT1 is an iteratively working enzyme that does not act like other enzymes which perform one kind of reaction several times. Examples are cellobiohydrolases and peptidoglycan GTs. These enzymes contain one active site inside a tunnel where the product/substrate is being processed to reach the active center several times. In contrast, LanGT1 has to release the product after the first glycosylation step. It is worthwhile to mention that LanGT1 is a highly specific enzyme, only recognizing landomycin I and landomycin J as substrate but no other landomycin with two, three, five, or six sugars. Thus, the mechanism of recognition of LanGT1 is very unusual.

SIGNIFICANCE

GTs are a very important class of enzyme involved in many processes in living cells. Understanding the basis for substrate specificity of these enzymes is very important. This article focuses on two related GTs which differ in substrate specificity. One is capable of iterative glycosylation, the other is not. Amino acids were identified that specifically contribute to iterative action. To our knowledge, this is the first systematic study of GT iterative action, an activity unique to a few, but important, GTs involved in natural product biosynthesis. This work will guide the design of GTs with engineered substrate specificity, an important issue for the development of novel glycosylated natural products.

EXPERIMENTAL PROCEDURES

General

S. cyanogenus S136 was obtained from the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany; DSM 5087); *S. globisporus* 1912 was a kind gift from Prof. Dr. V. Fedorenko (Lviv, Ukraine). *Escherichia coli* XL1-Blue MRF was purchased from Stratagene, *E. coli* ET12567 (pUZ8002) was obtained from Prof. C. Smith (Manchester, UK), pKC1132 was obtained from Eli Lilly, and pUC19 was purchased from New England Biolabs. pMunII was constructed as described earlier (Trefzer et al., 2001). pKC1218E (containing the erythromycin promoter) was kindly provided by Prof. Dr. J. Salas, Universidad de Oviedo (Oviedo, Spain). pSET*erm*E was obtained from Combinature Biopharm (Berlin, Germany). All chemicals were reagent grade and purchased from Promega, New England Biolabs, Roth, or Sigma-Aldrich. Primers were ordered from Eurofins MWG Operon (Ebersberg, Germany). *E. coli* was grown in LB medium as described (Sambrook and Russell, 2001). *S. cyanogenus* was grown in TSB medium (3% tryptic soy broth) and in SG medium (1% soybean meal, 2% glucose, 0.2% $CaCO_3$, 0.234% valine) for production. Apramycin, kanamycin, and carbenicillin were used in concentrations of 50 μg/ml; phosphomycin was used at a concentration of 100 μg/ml. Isolation of plasmid DNA from *E. coli* and restriction, purification, ligation, and site-directed mutagenesis were performed according to the protocols of the manufacturers of the kits and enzymes (Promega, QIAGEN, Macherey-Nagel, and Stratagene).

Chemical Analysis

Landomycins were extracted from liquid cultures with ethyl acetate. TLC was performed on silica gel plates (silica gel 60 F_{254}; Merck) with methylene chloride:methanol (9:1, v/v) as solvent. LC-electrospray ionization (ESI)-MS analysis was performed on an Agilent 1100 series system equipped with a quadrupole detector. The column used on the LC-MS system was a Zorbax Eclipse XDB-C8 main column (150 × 4.6 mm; 5 μm particles). The UV detection wavelength was 254 nm. ESI mass spectrometry was carried out using chamber settings as follows: drying gas flow, 12 ml/min; drying gas temperature, 350°C; nebulizer pressure, 50 psig; capillary voltage, 3000 V. Samples were dissolved in acetonitrile:H_2O (1:1, v/v), and analyzed at a 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, 19–23.10 min 85% B, 23.10–26 min 95% B (solvent A: 99.5% water, 0.5% acetic acid; solvent B: acetonitrile). Analysis was carried out in negative-ionization mode with a mass range set to 300–1150 Da.

Construction of Expression Plasmids for *lanGT1*, *lndGT1*, and *lanGT3*

For expression, a PCR fragment containing *lanGT1* was ligated into pSET behind the ermE promoter. The plasmid was transferred to *S. cyanogenus* Δ *lan*GT1 by conjugation. Landomycin production was assayed as described above.

Copies of *lndGT1* were synthesized as PCR fragments containing the ribosome-binding site and MfeI and XbaI sites or two XbaI sites, respectively, and

cloned into the MfeI and XbaI sites of pSET or into the XbaI sites of pKC1218, in both cases behind the ermE promoter.

Plasmid pKC1218lanGT3 was constructed as described (Hoffmeister et al., 2002).

Preparation of Chimeric Genes

For the construction of the chimeric genes, DNA fragments were synthesized by PCR amplification. Reaction conditions in general involved 25 cycles of denaturing at 94°C for 1 min (first cycle for 6 min), annealing at 60°C for 1 min, and extending at 72°C for 2 min using the *Pfu* polymerase. Primers used for each construct are listed in Table 1. These fragments, containing approximately 40 bp of overlapping DNA, were used as megaprimers in the next PCR reactions to create the hybrids. Reaction conditions were used as described, except for extending for 4 min. These hybrids were used as templates for an additional round of PCR with primers covering the 5′and 3′ ends of either *lanGT1* or *lndGT1*. Conditions were identical to those for the creation of hybrids. An MfeI site and an XbaI site were introduced upstream and downstream of the chimeric genes. PCR products were restricted by MfeI and XbaI and ligated into pMunII. Genes were sequenced and cloned into pSET152 behind the *erm*E promoter. Mutations were introduced using the QuikChange site-directed mutagenesis kit from Stratagene (for primers, see Table 1).

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), Be 1389/5-2 grant, to A.B. We thank Gisela Grabellus for excellent technical assistance and Andreas Günther for carrying out LC-MS analyses.

Received: September 24, 2008
Revised: December 1, 2008
Accepted: December 4, 2008
Published: January 29, 2009

REFERENCES

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.

Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., and Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. Glycobiology *16*, 29R–37R.

Dürr, 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 transfer. Angew. Chem. Int. Ed. Engl. *43*, 2962–2965.

Fischbach, M.A., Lai, J.R., Roche, E.D., Walsh, C.T., and Liu, D.R. (2007). Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. Proc. Natl. Acad. Sci. USA *104*, 11951–11956.

Hancock, S.M., Vaughan, M.D., and Withers, S.G. (2006). Engineering of glycosidases and glycosyltransferases. Curr. Opin. Chem. Biol. *10*, 509–519.

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.

Liang, D., and Qiao, J. (2007). Phylogenetic analysis of glycosyltransferases. J. Mol. Evol. *64*, 342–353.

Luzhetskyy, A., Fedoryshyn, M., Dürr, C., Taguchi, T., Novikov, V., and Bechthold, A. (2005). Iteratively acting glycosyltransferases involved in the hexasaccharide biosynthesis of landomycin A. Chem. Biol. *12*, 725–729.

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

Méndez, C., Luzhetskyy, A., Bechthold, A., and Salas, J.A. (2008). Deoxysugars in bioactive natural products: development of novel derivatives by altering the sugar pattern. Curr. Top. Med. Chem. *8*, 710–724.

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 A. J. Mol. Biol. *372*, 67–76.

Mulichak, A.M., Losey, H.C., Walsh, C.T., and Garavito, R.M. (2001). Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. Structure *9*, 547–557.

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.

Newman, D.J., Cragg, G.M., and Snader, K.M. (2003). Natural products as sources of new drugs over the period 1981–2002. J. Nat. Prod. *66*, 1022–1037.

Ostash, B., Rix, U., Rix, L.L., Liu, T., Lombo, F., Luzhetskyy, A., Gromyko, O., Wang, C., Braña, A.F., Méndez, C., et al. (2004). Generation of new landomycins by combinatorial biosynthetic manipulation of the LndGT4 gene of the landomycin E cluster in *S. globisporus*. Chem. Biol. *11*, 547–555.

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., in press. Published online September 29, 2008. 10.1002/bit.22150.

Salas, J.A., and Méndez, C. (2007). Engineering the glycosylation of natural products in actinomycetes. Trends Microbiol. *15*, 219–232.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual, Third Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Trefzer, A., Fischer, C., Stockert, S., Westrich, L., Künzel, 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.

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

Zhang, C., Albermann, C., Fu, X., and Thorson, J.S. (2006). The in vitro characterization of the iterative avermectin glycosyltransferase AveBI reveals reaction reversibility and sugar nucleotide flexibility. J. Am. Chem. Soc. *128*, 16420–16421.

Zhang, C., Bitto, E., Goff, R.D., Singh, S., Bingman, C.A., Griffith, B.R., Albermann, C., Phillips, G.N., Jr., and Thorson, J.S. (2008). Biochemical and structural insights of the early glycosylation steps in calicheamicin biosynthesis. Chem. Biol. *15*, 842–853.