Chemistry & Biology, Vol. 9, 287–295, March, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S1074-5521(02)00114-X

Engineered Urdamycin Glycosyltransferases Are Broadened and Altered in Substrate Specificity

Dirk Hoffmeister,[1] Barrie Wilkinson,[2,6] Graham Foster,[2] Philip J. Sidebottom,[3] Koji Ichinose,[4] and Andreas Bechthold[1,5]

[1]Department of Pharmaceutical Biology
Albert-Ludwigs-Universität Freiburg
Stefan-Meier-Strasse 19
79104 Freiburg
Germany

[2]Biotransformations and Strategic Technologies
[3]Analytical Sciences
GlaxoSmithKline Medicines Research Centre
Gunnels Wood Road
Stevenage
Hertfordshire SG1 2NY
United Kingdom

[4]Graduate School of Pharmaceutical Sciences
The University of Tokyo
Hongo
Bunkyo-ku
113-0033 Tokyo
Japan

Summary

Combinatorial biosynthesis is a promising technique used to provide modified natural products for drug development. To enzymatically bridge the gap between what is possible in aglycon biosynthesis and sugar derivatization, glycosyltransferases are the tools of choice. To overcome limitations set by their intrinsic specificities, we have genetically engineered the protein regions governing nucleotide sugar and acceptor substrate specificities of two urdamycin deoxysugar glycosyltransferases, UrdGT1b and UrdGT1c. Targeted amino acid exchanges reduced the number of amino acids potentially dictating substrate specificity to ten. Subsequently, a gene library was created such that only codons of these ten amino acids from both parental genes were independently combined. Library members displayed parental and/or a novel specificity, with the latter being responsible for the biosynthesis of urdamycin P that carries a branched saccharide side chain hitherto unknown for urdamycins.

Introduction

Facing cancer or multiresistant bacterial pathogenes, the search for new compounds has become a major issue in the pharmaceutical and medical sciences. The Streptomycetes, especially, with their diverse secondary products, among them many glycosylated compounds, have attracted researchers' attention. Since glycosylation often provides an essential factor for pharmacologcial activity [1, 2], glycosides promise to be a rewarding group for the development of new bioactive compounds. Attempts to investigate and influence the Streptomycete secondary metabolism led to combinatorial biosynthesis approaches to create unnatural natural products with enhanced structural diversity [3]. For example, much effort was made in engineering polyketide formation, and it has already provided new candidate structures for further pharmaceutical development [4–11]. However, tailoring reactions, especially glycosylation of the polyketide backbone with modified sugars, were looked into to a much smaller extent. Since glycosyltransferases (GTs) often display high specificity toward one or both substrates, in vitro evolution should be an attractive strategy to provide GTs with a broader substrate range.

Technically, three major obstacles have hampered engineering or the in vitro evolution of natural product GTs. First, in the majority of cases, the nucleotide-activated, highly modified sugars are not available. Second, low transformation efficiency or a lack of conjugational competence of suitable host strains, primarily *Streptomyces* species or closely related actinomycetes, strongly limit the size of shuffled gene/enzyme libraries to be expressed and screened for effects on antibiotic biosynthesis. Third, potentially positive clones cannot easily be identified by a rapid prescreening, for example, based on a simple change in color or pH, as was done for screening hydantoinases or carotenoids [12–13], or based on a halo around a bacterial colony when investigating proteases [14]. To circumvent these limitations, we made use of two highly homologous GTs, UrdGT1b and UrdGT1c, both involved in urdamycin biosynthesis (Figure 1) and sharing a 91% identical amino acid sequence, nevertheless transferring different activated deoxysugars to different acceptors. Previous work revealed that one particular protein region, amino acids (aa) 52–82, with an 18 amino acid difference between both parents (Figure 2A), is responsible for both nucleotide sugar and acceptor substrate specificity of these GTs [15]. Thus, we could well confine the region to be engineered. Both parent genes/enzymes originate from the urdamycin producer *Streptomyces fradiae* Tü2717 and belong to GT family 1, according to the CAZy classification [16, 17]. UrdGT1c strictly accepts dNDP-L-rhodinose and transfers it via an α-(1-3)-glycosidic bond to the C-3 hydroxyl of D-olivose, while UrdGT1b transfers D-olivose to the C-4 hydroxyl group of L-rhodinose, forming a β-(1-4)-glycosidic bond [18]. The present work demonstrates that engineering natural product GTs can fuse, alter, or disconnect their donor sugar-, acceptor substrate-, and acceptor position specificities to obtain natural products with unnatural glycosylation patterns.

Results

Reduction of the Number of Amino Acids Significant for UrdGT1b and UrdGT1c Specificity to Ten

A fully independent combination of 18 amino acids would have required a library of 2^{18} members, too many

[5]Correspondence: andreas.bechthold@uni-freiburg.de
[6]Present address: Biotica Technology, 181A Huntington Road, Cambridge CB3 0DJ, United Kingdom.

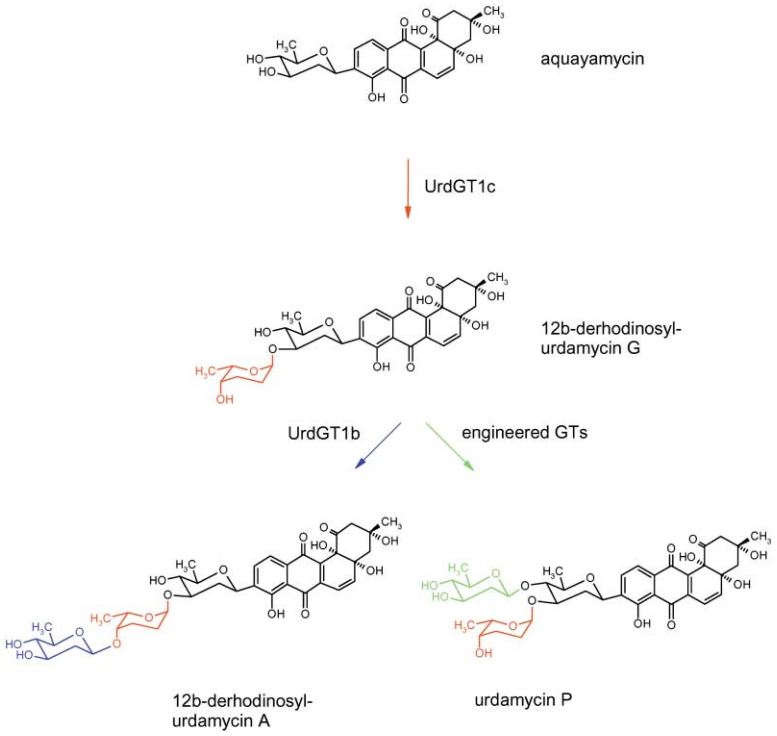


Figure 1. Late Glycosylation Steps during Urdamycin Biosynthesis

Steps catalyzed by UrdGT1b/UrdGT1b-like enzymes and UrdGT1c/UrdGT1c-like enzymes are indicated in blue and red, respectively. The novel mode of glycosylation is shown in green.

for a HPLC-based screening procedure. The following GTs were constructed to further constrict the relevant enzyme portion. GT-DH1, GT-DH2, and GT-DH3 are UrdGT1b derivatives but carry residues of UrdGT1c origin within the specificity-conferring region: GT-DH1 carries ^{80}Pro-Ala-Pro^{82}, GT-DH2 additionally carries Val76, and GT-DH3 carries ^{80}Pro-Ala-Pro^{82} and ^{68}Gly-Pro-Asp-Gln^{71} (Figure 2A). The presence of enzymes GT-DH1, GT-DH2, and GT-DH3 in mutant *S. fradiae* AX did not influence glycosylation. Contrastingly, gene expression in *S. fradiae* XTC resulted in the production of mainly 12b-derhodinosyl-urdamycin A (Figure 1) in all cases. These results indicate that the enzymes GT-DH1, GT-DH2, and GT-DH3, respectively, are fully active, like UrdGT1b, and that the amino acid changes did not affect their specificity at all. GT-DH4 is perfectly identical with GT-DH3 at the protein level but carries in its gene a unique AvrII site to facilitate further library engineering. In its enzymatic properties, GT-DH4 is identical to GT-DH3, so it also represents an UrdGT1b-type enzyme. Thus, the key to which glycosyltransfer will be performed should be found among these remaining ten amino acids.

Construction and Expression of a GT Library

To gain further insight into the role of these ten amino acids, a family of *urdGT1b/1c* chimeric GT genes was created by a PCR-based approach. It was carried out such that the codons responsible for these ten amino acids were independently built in, while all other codons remained unchanged, compared to the template gene, *GT-DH4*. For an independent build-in, 11 nucleotide positions must be variable. Each position can be occupied by two nucleotides, thus allowing $2^{11} = 2048$ combinations. The library was prepared in *Escherichia coli*, and chimeric genes were expressed in *S. fradiae* mutants. To assess GT specificity, we made use of two *S. fradiae* mutant strains. The first strain, *S. fradiae* AX [18], produces aquayamycin (Figure 1), which is the acceptor substrate for UrdGT1c. The second strain, *S. fradiae* XTC [15], provides the acceptor substrate for UrdGT1b, 12b-derhodinosyl-urdamycin G (Figure 1).

A total of 400 *S. fradiae* AX and 198 *S. fradiae* XTC transformants were screened for their urdamycin production spectrum by TLC and HPLC. Once an altered pattern was found for one host, the plasmid with the functional GT gene library member was reisolated for DNA sequencing. In addition, it was used to transform the alternative host strain to check whether an additional enzyme activity had been obscured due to the intrinsic metabolite pattern of the first host strain. The screening revealed 40 different chimeric GTs. Basically, they split up into two main categories: GTs with parental enzymatic specificity (single or combined), and those catalyzing a novel glycosylation mode, also single or combined with one or both parental specificity.

GTs with Parental Specificity

Three library members were shown to act like the parental UrdGT1b-type enzyme GT-DH4, i.e., the enzyme specificity had been retained. From ten amino acids in question, Asn52, Asp55, and ^{58}His-Val-Asp-Asp-Met^{62} are conserved in these three GTs. A complete switchover into a strictly UrdGT1c-like enzyme was accomplished five times. Three of the ten amino acids in question were found to be constant; these are Asn52, Lys62, and Leu64. Three GTs displayed both parental activities, depending on the acceptor substrate provided. In the

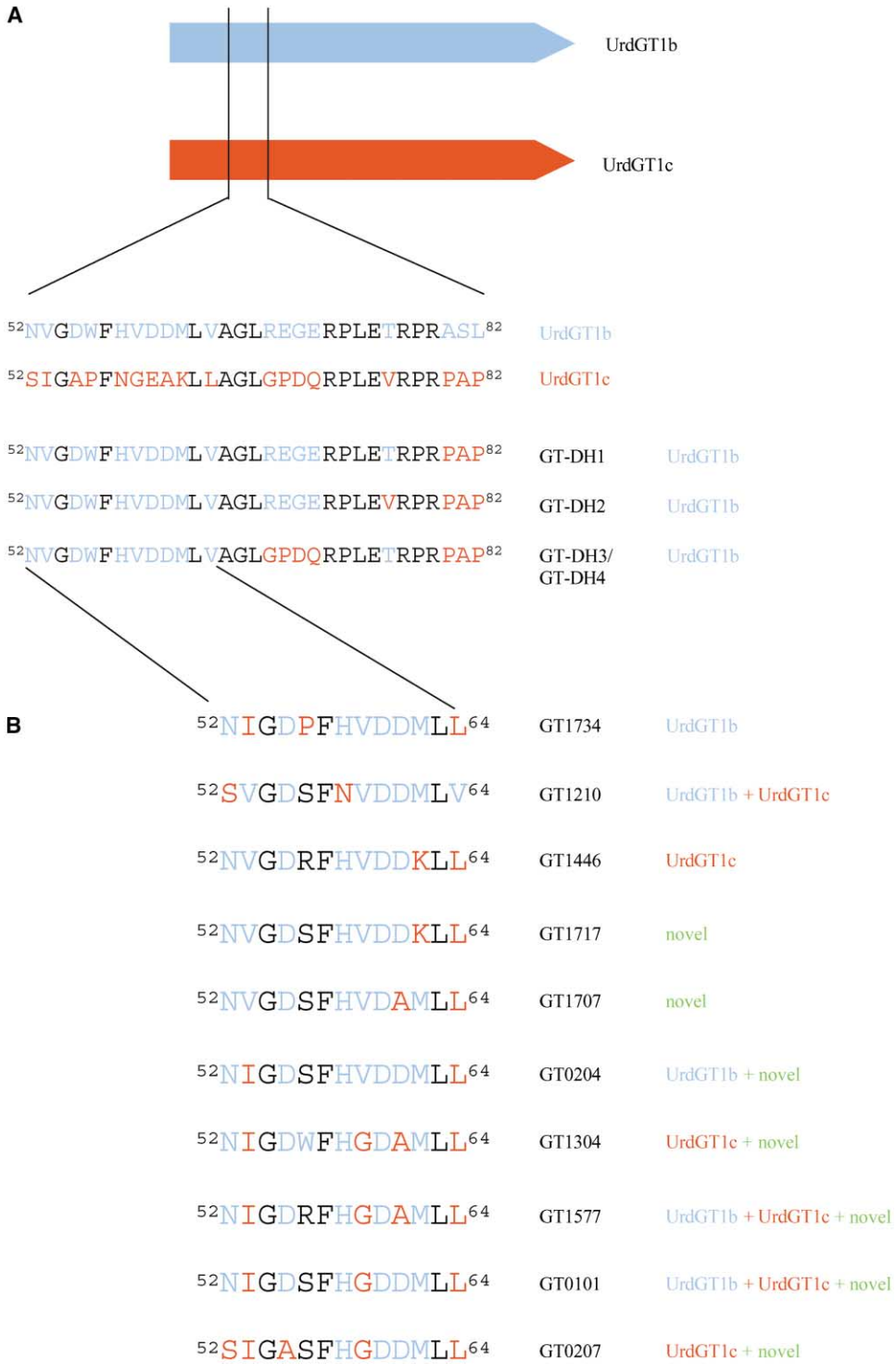


Figure 2. Wild-Type and Engineered Glycosyltransferases

(A) Positioning of the specificity-conferring region (aa 52–82) within the enzymes UrdGT1b (blue arrow) and UrdGT1c (red arrow) and sequence alignment including the parental wild-type enzymes and GT-DH1-4.

(B) Amino acid combinations of selected library members. Only the engineered protein region (aa 52–64) is shown.

presence of aquayamycin (in mutant *S. fradiae* AX), L-rhodinose was added to yield 12b-derhodinosyl-urdamycin G. In the presence of 12b-derhodinosyl-urdamycin G (in mutant *S. fradiae* XTC), D-olivose was appended to yield 12b-derhodinosyl-urdamycin A. Remarkably, two consecutive rounds of enzymatic activity, i.e., a direct conversion from aquayamycin to 12b-derhodinosyl-urdamycin A in *S. fradiae* AX, have never

been observed; although, the product of the first round could have served as an acceptor in the second. In these three dual active GTs, neither activity dominates significantly, and one amino acid, Asp55, is conserved throughout.

GTs with Novel Specificity

Engineering GTs created enzymes that transfer D-olivose within the 12b-derhodinosyl-urdamycin G molecule to the C-4 hydroxy group of D-olivose, forming a β-(1-4)-glycosidic bond. This new glycosylation mode yields urdamycin P (Figures 1 and 3A), carrying a branched saccharide chain that is a feature so far unknown for urdamycins. NMR data of urdamycin P are given in Table 1; additional support for the urdamycin P structure arises from fragmentation experiments (Figure 3C). Surprisingly, with 28 members, the novel active GTs made up the largest group in this screening. It can be subdivided into those with exclusively novel activity (2); those integrating a second activity, UrdGT1b (1) or UrdGT1c (21); or those displaying triple activity (5). GTs 1707 and 1717, both with solely new activity, were found in the screening. It is intriguing that one of them, GT1717, differs only in Ser56 from an UrdGT1c-type enzyme (GT1446), with arginine at this position (Figure 2B). With GT0204, a GT fusing UrdGT1b-like and new specificity was found, and it differs at positions 53 (Val instead of Ile) and 61 (Asp instead of Ala) from GT 1707 (Figure 2B), which lacks GT1b-like activity. With 21 examples, the most abundant GTs were those displaying new activity combined with an UrdGT1c-like mode of enzymatic action. Interestingly, five of them (e.g., GT1304, Figure 2B) accept aquayamycin as a substrate, i.e., the engineered enzyme acts twice on its acceptor substrate to convert it to urdamycin P. These five GTs invariably possess a Gly59 and Leu64, combined with a Met62 four times. Although the target hydroxy group on C-4 of D-olivose is already present in aquayamycin, 16 GTs showed the UrdGT1c-type-, but not the novel activity, when aquayamycin was offered as a substrate. In contrast, novel activity was exerted in these cases with 12b-derhodinosyl-urdamycin G as an acceptor substrate. Among these 16 enzymes, none of the amino acids in question is fully conserved. Nevertheless, the Gly59-Leu64 motif was found 12 times, and a Met62 has been detected 13 times.

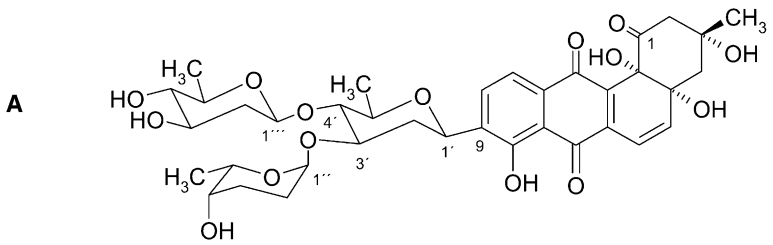
Triple-Specific GTs

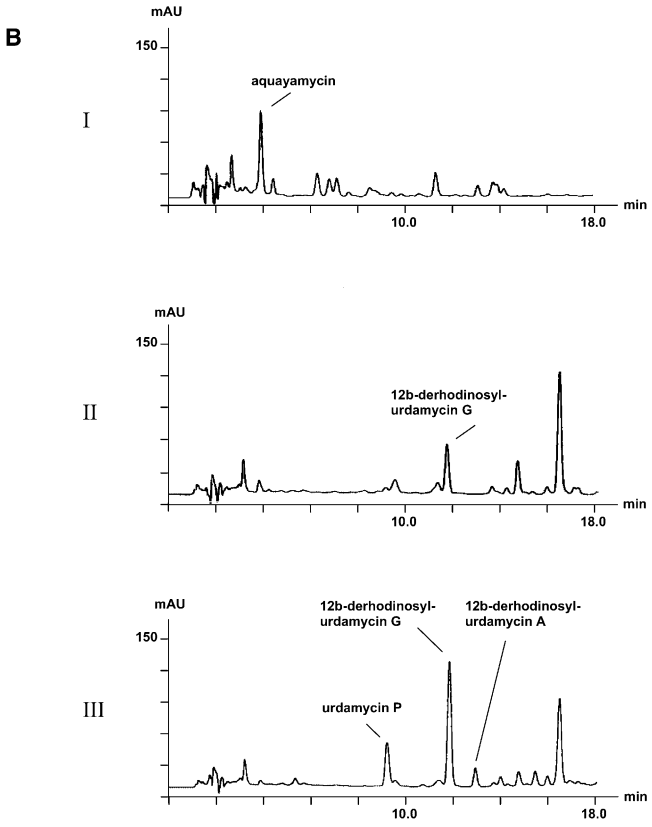
Five enzymes (GTs 1438, 1511, 1577, 1815, and 0101) were shown to act in a highly unspecific manner. In the presence of aquayamycin (in mutant *S. fradiae* AX), GTs 1438, 1511, 1577, and 1815 were acting like UrdGT1c producing 12b-derhodinosyl-urdamycin G. They showed new and UrdGT1b-like specificity, using 12b-derhodinosyl-urdamycin G (in mutant *S. fradiae* XTC) as an acceptor, which resulted in parallel 12b-derhodinosylurdamycin A and urdamycin P production. The fifth enzyme (GT0101, Figure 2B, Table 2) was able to convert aquayamycin to urdamycin P in *S. fradiae* AX, and, in *S. fradiae* XTC, it converted 12b-derhodinosyl-urdamycin G to urdamycin P and 12b-derhodinosyl-urdamycin A simultaneously. In summary, this enzyme is most flexible and exhibited all possible enzyme activities. As for the sequence, GT0101 (which includes UrdGT1b activity) differs only in Asn52 and Asp55 from GT0207 (Ser52 and Ala55), which is void of UrdGT1b-like activity. Apparently, one or both amino acids integrate an UrdGT1b-type mode of action into the enzyme. Also, GT1577 possesses Trp56, while GT1304 with Arg56 lacks UrdGT1b-like activity (Figure 2B, Table 2). Estimated by HPLC peak areas, urdamycin P production dominated 12b-derhodinosyl-urdamycin A biosynthesis in all five cases. However, with a ratio of 1:0.6 GT0101 was found to balance best between both glycosylation products (HPLC-chromatogram shown in Figure 3B).

Discussion

Glycosyltransferases with a new enzyme specificity were generated by genetic engineering. The novel glycosylation mode resulted in a branched saccharide originating from hydroxyl groups at C-3 and C-4 of the carrier sugar, a structural feature that also appears, for example, in saccharomicins [19] or in calicheamicins as a branch via sugar atoms C-2 and C-4 [20]. An urdamycin derivative carrying merely a diolivosyl side chain has not been found in the screening. Thus, we conclude that the sugar side chain of 12b-derhodinosyl-urdamycin G is a prerequisite for the novel glycosylation mode and that attachment of L-rhodinose to C-3 of the C-glycosidically linked D-olivose occurs prior to the novel glycosylation.

Crystallographic and computational investigations identified GTs as two-domain enzymes [21]. It is assumed that the active site resides in a cleft between these two domains. In the absence of solved crystal structures for the GTs described here, we speculate that the recombined amino acids reside in perhaps a substrate- or product-lodging loop or cavity close to the substrate binding sites inside the protein. Such loops have been reported from several solved crystal structures of GTs, for example, SpsA from *Bacillus subtilis* [22], GnTI from rabbit [23], and others. These loops appear flexible and could play a role in preventing the nucleotide sugar from being hydrolyzed or, more likely, in substrate binding and in product release [21]. Once the glycosyltransfer is accomplished (indicated by the nucleotide diphosphate being cleaved off the sugar), the loop is altered structurally. Compared with other GTs of streptomycete secondary metabolism, e.g., those of the pathways for erythromycin [24, 25], daunorubicin [26, 27], mithramycin [28, 29], oleandomycin biosynthesis [30], and many more, the protein region investigated during this study shows little conservation in length and sequence, presumably reflecting the individual steric interactions for the particular glycoside molecule. Due to steric constraints, the ten amino acids in question, and their positional equivalents in other GTs, could decide if a product fits in the loop and if it is being released from the catalytic center to exit the enzyme. Recently, the crystal structure of GtfB, an UDP-glucosyltransferase involved in chloroeremomycin biosynthesis, has been published [31]. A shallow pocket located within the N-terminal protein domain facing the interdomain





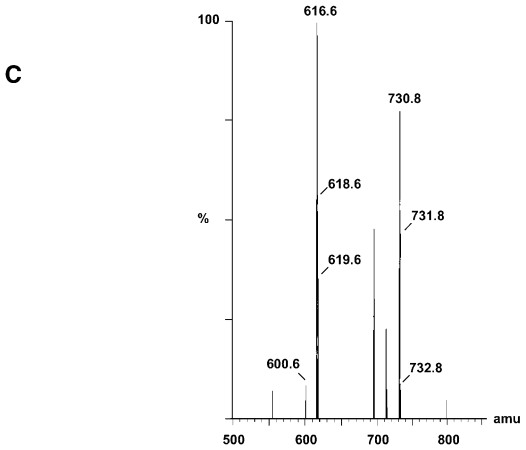


Figure 3. Chemical Analysis of Urdamycin P

(A and B) (A) Molecular structure and (B) HPLC chromatograms of untransformed host strains, *S. fradiae* mutants AX (I) and XTC (II) as controls, and *S. fradiae* XTC expressing GT0101 (III).

(C) This panel documents the fragmentation pattern of urdamycin P after APCI-MS (negative mode). The m/z value range from 500–850 atomic mass units (amu) is shown. Signal intensities are given as relative abundance. The 616.6 amu signal correlates to an urdamycin P fragment lacking L-rhodinose, and the 600.6 amu signal correlates to a fragment lacking D-olivose, whereas the 730.8 amu signal corresponds to the unfragmented molecule. Major artificial follow-up products of urdamycin P and fragments generated during MS analysis are also indicated by their respective m/z value.

Table 1. NMR Data for Urdamycin P in CD_3OD

Position Number	Chemical Shifts	
	δ^1H	$\delta^{13}C$
1	–	n.o
2	2.82 (d, 13)	52.9
	2.67 (dd, 2.5, 13)	
3	–	77.4
4	2.04	44.5
5	6.40 (d, 9.5)	145.7
6	6.87 (d, 9.5)	117.8
7	–	n.o
8	–	n.o
9	–	n.o
10	7.85 (d, 7.5)	133.9
11	7.58 (d, 7.5)	119.5
12	–	n.o.
3 Me	1.24 (s)	29.8
1′	4.86	72.0
2′	2.59 (ddd, 2, 4.5, 12.5)	37.3
	1.29	
3′	3.91 (ddd, 4.5, 9, 11)	73.8
4′	3.37 (t, 7)	82.2
5′	3.53	77.2
6′	1.39 (d, 6)	18.7
1″	4.93 (brd, 3)	93.7
2″	2.00	24.5
	1.35	
3″	2.25	26.2
	1.62 (m)	
4″	3.52	68.1
5″	4.59 (dq, 1.5, 6.5)	66.8
6″	1.14 (d, 6.5)	17.2
1‴	4.74 (dd, 2, 9.5)	100.9
2‴	2.23	40.6
	1.44 (dt 9.5, 12)	
3‴	3.51	71.8
4‴	2.84 (t, 9)	78.3
5‴	3.24 (dq, 9, 6)	73.1
6‴	1.29 (d, 6)	18.3

^1H chemical shifts are referenced to CHD_2OD at 3.31 ppm. For resolved signals, multiplicities and coupling constants (±0.5 Hz) are given in parentheses. The other ^1H chemical shifts are derived from 2D spectra. ^{13}C chemical shifts are derived from a 2D ^1H-^{13}C correlation spectrum (HMQC) and are referenced to CHD_2OD at 48.9 ppm. (n.o. = signal not observed).

Table 2. Amino Acids at the Ten Engineered Positions of All Active GT Library Members

Specificity Type	GT #	Amino Acid Combinations (Positions 52–64)
UrdGT1b	2–1	NV-DP-HVDDM-V
	3–24	NI-DR-HVDDM-V
	1734	NI-DP-HVDDM-L
UrdGT1c	6–13	NI-DS-NGEDK-L
	1106	NI-AP-HGEDK-L
	1423	NI-AR-HVDDK-L
	1446	NV-DR-HVDDK-L
	1701	NV-DS-NGDAK-L
Novel	1707	NV-DS-HVDAM-L
	1717	NV-DS-HVDDK-L
UrdGT1b + UrdGT1c	3–11	NI-DW-HVDAK-L
	1210	SV-DS-NVDDM-V
	1548	NV-DR-NGDAM-L
Novel + UrdGT1b	0204	NI-DS-HVDDM-L
Novel + UrdGT1c*	0207	SI-AS-HGDDM-L
	1304	NI-DW-HGDAM-L
	1322	NV-DS-HGEDM-L
	1565	SI-AW-NGDDK-L
	1814	SI-DW-HGDAM-L
Novel + UrdGT1c	3–17	NV-AS-NVEAM-V
	3–19	NV-DS-NVEAM-L
	6–19	SV-AS-HGDDK-L
	9–1	NI-AS-HGDDM-L
	0820	IV-AP-NGDAM-L
	1111	NV-DW-HGEDK-L
	1128	SI-AS-NVDAM-L
	1141	NV-DW-NGDAM-V
	1413	SV-DP-HGDAK-L
	1416	SV-AS-HGDAM-L
	1574	SV-DP-HGDAM-L
	1703	SI-AS-HGDAM-L
	1709	NV-DS-HGDAM-L
	1720	NV-DP-HGDAM-L
	1806	SF-DP-NGEAM-L
	1812	SV-DP-NGDDM-L
Novel + UrdGT1b + UrdGT1c	0101	NI-DS-HGDDM-L
	1438	SI-DS-NVDDM-L
	1511	NV-DP-NGDDM-L
	1577	NI-DR-HGDAM-L
	1815	NI-AW-HGDDM-V

The category marked by an asterisk comprises those novel + UrdGT1c-type enzymes that also accept aquayamycin as a substrate.

cleft has been proposed as an acceptor substrate binding site. Two leucine residues are central to the pocket, surrounded by six tyrosine or phenylalanine residues plus three alanines and a cysteine. We found that none of these residues that are crucial in GtfB are conserved within UrdGT1b/1c, except Tyr122 (which is Tyr110 in UrdGT1b/1c). The six aromatic residues are part of the GtfB region comprising aa 122–190. The equivalent protein part in the GTs described here, however, comprises only one tyrosine and phenylalanine each. In turn, those amino acid positions detected to be relevant to specificity of UrdGT1b/1c have not been found to interact with at least the acceptor substrate of GtfB. Observation of the nucleotide-bound sugar within the crystals is extremely difficult; therefore, little is known about particular side chain interactions. Thus, it still seems possible that GtfB residues, so far not recognized as being important for sugar substrate acceptance, have their counterpart amino acid within the region investigated during this study. In kinetic studies into a number of pinpoint mutations performed concurrently with the GtfB crystallization work, Asp332 was identified as a potential member of the GtfB catalytic apparatus [31]. Interestingly, the adjacent amino acids are strongly conserved among GtfB and UrdGT1b/1c (^{327}Leu-Pro-Gln-Met-Ala-Asp-Gln^{333} in GtfB, and ^{311}Leu-Pro-Gln-Leu-Ala-Asp-Gln^{317} in UrdGT1b). Therefore, further work on UrdGT1b/1c could include crystallization to obtain a clearer picture about the role of this residue and whether wild-type and chimera involve the same catalytic base.

Possibly due to structural flexibility, certain amino acid combinations favor a particular activity without strictly dictating it. Unlike what is reported for a donor substrate of human blood group GTs [32], we demonstrated that substrate range is not dictated by a single amino acid residue or position. As has been shown during this study, various amino acid combinations result in an identical pattern of enzymatic specificity. However, almost all rhodinose-transferring GTs possess Leu64,

and when additionally integrating novel or novel and UrdGT1b activity, Gly59 and Leu64 were found almost throughout. GTs with UrdGT1b activity (single or in combination) almost always combine Asp60 and Met62. With about 7% active library members, the protein structure subjected to engineering seems to be vulnerable against amino acid substitution, and none of the engineered GTs performed better than the wild-type enzymes, according to HPLC-peak area comparison. Nevertheless, the investigated region appears to be flexible enough to tolerate nonparental amino acids in functional enzymes, i.e., Ser and Arg on position 56 due to the construction procedure, or Phe53 and Ile52, possibly due to a PCR error in GT1806 and GT0820, respectively (Table 2).

For numerous applications, rational design or directed evolution approaches have proved to be a promising strategy [33–36]. For natural product derivatization, this class of enzymes provides particularly versatile tools, since they account for specific structural changes with consequences on bioactivity and could close the gap between glycorandomized sugar synthesis [37] and a vast number of polyketide-derived aglycons.

Significance

Genetic engineering of the enzyme regions responsible for selecting sugar and acceptor substrates in urdamycin GTs UrdGT1b and UrdGT1c has provided GTs with combined and/or novel specificity, leading to a new urdamycin derivative with an unusual branched saccharide side chain. This study should stimulate further investigations on GTs as valuable tools for drug derivatization and lead compound generation as well as rational saccharide design. With critical hot spots now being located, saturation mutagenesis could open up new possibilities for engineering new specificities into GTs. From the mechanistic point of view, it is intriguing that different amino acid combinations account for the same substrate spectrum and that minor changes in the amino acid sequence enable or disable an additional catalytic activity. These findings might give an impetus for comparing crystallization or computational studies to unveil further how GTs work in respect to substrate choice and product release.

Experimental Procedures

Growth Conditions, Media, and Genetic Manipulation of *Streptomyces fradiae*

For screening purposes, transformants were fermented in miniature culture vessels, each filled with 7 ml NL 111 V liquid medium [18] supplemented with 25 μg/ml thiostrepton and kept at 250 rpm for 96 hr at 28°C. For urdamycin P production, strain *S. fradiae* XTC + GT0101 was grown as a seed culture for 48 hr in a 500-ml single-baffled erlenmeyer flask filled with 100 ml medium at 180 rpm and 28°C. This culture was then used to inoculate the main culture (1% v/v), which was fermented for 60 hr. All DNA manipulation was carried out in *Escherichia coli* XL-1 blue MRF' (Stratagene) following described procedures [38] and the instructions of the suppliers of enzymes and reagents (Promega, Pharmacia). Before transforming *S. fradiae* mutants AX and XTC, the gene library and the constructs carrying the edited genes were passed through *E. coli* ET12567 [39] to obtain nonmethylated DNA. *E. coli* strains were grown under standard conditions. Vector pUC18 was used for routine cloning purposes. Gene expression constructs were based on vector pUWL201 [40]. *Streptomyces* protoplast preparation, transformation, and regeneration were done as described [41].

Construction of Edited Glycosyltransferase Genes *GT-DH1*, *GT-DH2*, *GT-DH3*, and *GT-DH4*

Gene *GT-DH1* was constructed by PCR using primers M13 reverse and P1 to alter codons 80–82 of *urdGT1b* (Pro-Ala-Pro is encoded instead of Ala-Ser-Leu). P1: 5′-CGCCCG**CTCGAG**ACGCGTCCGC GGCCGGCGCCTGAGTCGATGGGCGGG-3′; the XhoI site is indicated in bold. The template for the construction of *GT-DH1* was pUC18 with the *urdGT1b* cloned into the HindIII and BglII sites. PCR conditions were: 7 min at 97°C; 30 cycles at 95°C for 90 s, at 49°C for 60 s, and at 72°C for 60 s; followed by a terminal hold at 72°C for 10 min. The PCR product was cleaved with XhoI and BglII to replace the equivalent sequence in plasmid pUC-urdGT1b, from which the original sequence had been removed, to generate plasmid pUC-DH1. The altered gene was excised with HindIII/XbaI and ligated into pUWL201 to create plasmid pUWL-DH1. Gene *GT-DH2* and plasmids pUC-DH2/pUWL-DH2 were constructed as described for *GT-DH1*, but using primer P2 (5′-GAGCGCCCG**CTCGAG**GT GCGTCCGCGGCCGG-3′), pUC-DH1 as the PCR template, and restriction enzymes XhoI (bold) and BglII to insert the PCR product into gene *GT-DH1*, lacking the equivalent sequence, resulting in gene *GT-DH2*. PCR conditions were: 7 min at 97°C; 30 cycles at 95°C for 90 s, at 62°C for 60 s, and at 72°C for 60 s; followed by a terminal hold at 72°C for 10 min. Additionally, to the pinpoint mutations of *GT-DH1*, *GT-DH2* carries a codon for valin on position 76, instead of threonine. Gene *GT-DH3* and plasmids pUC-DH3/pUWL-DH3 were constructed as described for *GT-DH1*, but using primers M13 universal and P3 (5′-GGACGCAC**CTCGAG**CGGCCGCTGA TCCGGCCCCAACCCGGCGACCAAC-3′), pUC-DH1 as a PCR template, and enzymes XhoI (bold) and HindIII to insert the 0.25-kb PCR product into gene *GT-DH1*, lacking the equivalent sequence, resulting in gene *GT-DH3*. PCR conditions were the same as those used for *GT-DH2*. Additionally, to the pinpoint mutations of *GT-DH1*, *GT-DH3* carries codons for Gly-Pro-Asp-Gln, instead of Arg-Glu-Gly-Glu, on positions 68–71. Glycosyltransferase gene *GT-DH4* was constructed as described for DH3, but using primer P4 (5′-GTCTCGA**GCGGCCGC**TGATCCGGCCCTAGGCCGGCGACCAAC-3′), pUC-DH3 as a PCR template, and enzymes HindIII and NotI (bold) to insert the PCR product into gene *GT-DH3*, lacking the equivalent sequence, resulting in gene *GT-DH4*. This gene carries silent mutations in codons 66 (Gly) and 67 (Leu) (GGGTTG in *GT-DH3* now replaced by GGCCTA in *GT-DH4*), thus generating a unique AvrII site (underlined).

Construction of a Library of Positionally Shuffled Glycosyltransferase Genes

Primers F: 5′-GATCCGGC**CCTAGG**CCGGCGASCAACWTGKCSTC CMCGTKGAACSRGKCGCCGAYGYTCACCGCGTTGAGC-3′, and R: 5′-AGAGAAGG**CCTAGG**GCCGGATCAGCGGCCGCTCGAGGTGCG TCCGCGGC-3′ (AvrII site in bold) were used for PCR; the template plasmid was pUC-DH4. The primers maintain the AvrII site of *GT-DH4*, overlap there, and are oriented divergently. Thus, the entire template plasmid was amplified by PCR. PCR conditions were: 7 min at 96°C; 35 cycles at 95°C for 60 s, at 56°C for 50 s, and at 72°C for 2 min; followed by a terminal hold at 72°C for 7 min, also using Taq DNA polymerase. The PCR product was cut with AvrII and ligated to constitute proper circular plasmids. Then, they were corestricted with HindIII and XbaI to remove the diversified genes, which subsequently were ligated into pUWL201 and used to transform *S. fradiae* mutants AX and XTC protoplasts.

Metabolite Screening

A total of 1 ml culture broth was extracted with an equal volume of ethyl acetate. The organic layer was evaporated, redissolved in methanol, and subjected to TLC and HPLC analysis. TLC was carried out on silica gel 60 F_{254} plates; the solvent was chloroform/methanol/ethyl acetate (8:1:1, v/v). HPLC was done using a Hewlett-Packard HP1100 series integrated system. Chromatography was over a 3-μ Hypersil HyPurity Elite C18 column eluting at 1.0 ml/min.Solvent A was 10% acetonitrile in H_2O and 10 mM ammonium acetate, and solvent B was 90% acetonitrile in H_2O and 10 mM ammonium ace-

tate. The gradient was 15%–90% B at 0–20 min; 90%–95% B at 20–21 min, held at this relation at 21–23 min; 95%–15% B at 23–24 min, held at 15% B at 24–27 min. The detection wavelength was $\lambda = 254$ nm; UV-VIS spectra were recorded from $\lambda = 190$–600 nm. R_f values and retention times of these substances were determined by comparison to those of authentic samples.

Extraction, Isolation, and Chemical Analysis of Urdamycin P

A total of 4 liters of fermentation broth of *S. fradiae* XTC + GT0101 was extracted twice with an equal volume of ethyl acetate. After evaporating the organic phase, the dry crude extract was dissolved in 8 ml methanol and applied in equal portions to four preparative HPLC runs. The equipment used was a Gilson 306 solvent pump, a Gilson 118 UV-VIS detector, and a 5-μ Kromasil C8 column (250 × 21.2 mm i.d.). Elution was at 15 ml/min with a gradient of 0%–60% B over 75 min, then held at 60% B for an additional 15 min. Solvent A was H_2O (10 mM ammonium acetate), and solvent B was 80% acetonitrile in H_2O (20 mM ammonium acetate). Fractions containing urdamycin P were bulked, diluted with an equal amount of water, and reapplied to the HPLC system/gradient described above. For a final workup, the urdamycin P fraction was adsorbed into a 500-mg Isolute Env+ cartridge (International Sorbent Technology). It was washed with 10 ml H_2O, then eluted with 10 ml methanol. About 4 mg urdamycin P was obtained. NMR spectra were recorded in CD_3OD on a Bruker AMX 500 spectrometer. Urdamycin P has a molecular formula of $C_{37}H_{46}O_{15}$; MW is 730.8, confirmed by atmospheric pressure chemical ionization (APCI) mass spectrometry under described conditions [42]. Production cultures of XTC + GT0101 yielded approximately 1 mg/l urdamycin P and 0.6 mg/l 12b-derhodinosyl-urdamycin A. Background production of 12b-derhodinosyl-urdamycin G, intrinsic to *S. fradiae* XTC, was about 3 mg/l.

Acknowledgments

The authors would like to thank Javier Velasco for excellent support and Nicholas L. Taylor for additional MS analysis. This work was supported by a grant of the European Union (QLRT-1999-00095) to A.B.

Received: August 6, 2001
Revised: October 15, 2001
Accepted: October 30, 2001

References

1. Weymouth-Wilson, A.C. (1997). The role of carbohydrates in biologically active natural products. Nat. Prod. Rep. *14*, 99–110.
2. Thorson, J.S., Hosted, T.J., Jiang, J., Biggins, J.B., and Ahlert, J. (2001). Nature's carbohydrate chemists: the enzymatic glycosylation of bioactive bacterial metabolites. Curr. Org. Chem. *5*, 139–167.
3. Reynolds, K.A. (1998). Combinatorial biosynthesis: lesson learned from nature. Proc. Natl. Acad. Sci. USA *95*, 12744–12746.
4. Jacobsen, J.R., Hutchinson, C.R., Cane, D.E., and Khosla, C. (1997). Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase. Science *277*, 367–369.
5. McDaniel, R., Kao, C.M., Hwang, S.J., and Khosla, C. (1997). Engineered intermodular and intramodular polyketide synthase fusions. Chem. Biol. *4*, 667–674.
6. Ruan, X., Pereda, A., Stassi, D.L., Zeidner, D., Summers, R.G., Jackson, M., Shivakumar, A., Kakavas, S., Staver, M.J., Donadio, S., et al. (1997). Acyltransferase domain substitutions in erythromycin polyketide synthase yield novel erythromycin derivatives. J. Bacteriol. *179*, 6416–6425.
7. Marsden, A.F.A., Wilkinson, B., Cortés, J., Dunster, N.J., Staunton, J., and Leadlay, P.F. (1998). Engineering broader specificity into an antibiotic-producing polyketide synthase. Science *279*, 199–202.
8. Katz, L., and McDaniel, R. (1999). Novel macrolides through genetic engineering. Med. Res. Rev. *19*, 543–558.
9. Xue, Q., Ashley, G., Hutchinson, C.R., and Santi, D.V. (1999). A multiplasmid approach to preparing large libraries of polyketides. Proc. Natl. Acad. Sci. USA *96*, 11740–11745.
10. McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., Betlach, M., and Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. Proc. Natl. Acad. Sci. USA *96*, 1846–1851.
11. Tang, L., Fu, H., and McDaniel, R. (2000). Formation of functional heterologous complexes using subunits from the picromycin, erythromycin and oleandomycin polyketide synthases. Chem. Biol. *7*, 77–84.
12. May, O., Nguyen, P.T., and Arnold, F.H. (2000). Inverting enantioselectivity by directed evolution of hydantoinase for improved production of L-methionine. Nat. Biotechnol. *18*, 317–320.
13. Schmidt-Dannert, C., Umeno, D., and Arnold, F.H. (2000). Molecular breeding of carotenoid biosynthetic pathways. Nat. Biotechnol. *18*, 750–753.
14. Ness, J.E., Welch, M., Giver, L., Bueno, M., Cherry, J.R., Borchert, T.V., Stemmer, W.P.C., and Minshull, J. (1999). DNA shuffling of subgenomic sequences of subtilisin. Nat. Biotechnol. *17*, 893–896.
15. 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.
16. Campbell, J.A., Davies, G.J., Bulone, V., and Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem. J. *326*, 929–942.
17. Coutinho, P.M., and Henrissat, B. (1999). Carbohydrate-Active Enzymes Server (http://afmb.cnrs.mrs.fr/~pedro/CAZYdb.html).
18. Trefzer, A., Hoffmeister, D., Künzel, E., Stockert, S., Weitnauer, G., Westrich, L., Rix, U., Fuchser, J., Bindseil, K.U., Rohr, J., et al. (2000). Function of glycosyltransferase genes involved in urdamycin A biosynthesis. Chem. Biol. *7*, 133–142.
19. Singh, M.P., Petersen, P.J., Weiss, W.J., Kong, F., and Greenstein, M. (2000). Saccharomicins, novel heptadecaglycoside antibiotics produced by *Saccharothrix espanensis*: antibacterial and mechanistic activities. Antimicrob. Agents Chemother. *44*, 2154–2159.
20. Zein, N., Sinha, A.M., McGahren, W.J., and Ellestad, G.A. (1988). Calicheamicin gamma 1I: an antitumor antibiotic that cleaves double-stranded DNA sites specifically. Science *240*, 1198–1201.
21. Ünligil, U.M., and Rini, J.M. (2000). Glycosyltransferase structure and mechanism. Curr. Opin. Struct. Biol. *10*, 510–517.
22. Charnock, S.J., and Davies, G.J. (1999). Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. Biochemistry *38*, 6380–6385.
23. Ünligil, U.M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J.M. (2000). X-ray crystal structure of rabbit *N*-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. EMBO J. *19*, 5269–5280.
24. Salah-Bey, K., Doumith, M., Michel, J.M., Haydock, S., Cortés, J., Leadlay, P.F., and Raynal, M.C. (1998). Targeted inactivation for the elucidation of deoxysugar biosynthesis in the erythromycin producer *Saccharopolyspora erythraea*. Mol. Gen. Genet. *257*, 542–553.
25. Summers, R.G., Donadio, S., Staver, M.J., Wendt-Pienkowski, E., Hutchinson, C.R., and Katz, L. (1997). Sequencing and mutagenesis of genes from the erythromycin biosynthetic gene cluster of *Saccharopolyspora erythraea* that are involved in L-mycarose and D-desosamine production. Microbiology *143*, 3251–3262.
26. Otten, S.L., Liu, X., Ferguson, J., and Hutchinson, C.R. (1995). Cloning and characterization of the *Streptomyces peucetius dnrQS* genes encoding a daunosamine biosynthesis enzyme and a glycosyltransferase involved in Daunorubicin biosynthesis. J. Bacteriol. *177*, 6688–6692.
27. Olano, C., Lomovskaya, N., Fonstein, L., Roll, J.T., and Hutchinson, C.R. (1999). A two-plasmid system for the glycosylation of polyketide antibiotics: bioconversion of ϵ-rhodomycinone to rhodomycin D. Chem. Biol. *6*, 845–855.

28. Fernández, E., Weissbach, U., Sánchez Reillo, C., Brana, A.F., Méndez, C., Rohr, J., and Salas, J.A. (1998). Identification of two genes from *Streptomyces argillaceus* encoding glycosyltransferases involved in transfer of a disaccharide during biosynthesis of the antitumor drug mithramycin. J. Bacteriol. *180*, 4929–4937.
29. Blanco, G., Fernández, E., Fernández, M.J., Brana, A.F., Weissbach, U., Künzel, E., Rohr, J., Méndez, C., and Salas, J.A. (2000). Characterization of two glycosyltransferases involved in early glycosylation steps during biosynthesis of the antitumor polyketide mithramycin by *Streptomyces argillaceus*. Mol. Gen. Genet. *262*, 991–1000.
30. Olano, C., Rodriguez, A.M., Michel, J.M., Méndez, C., Raynal, M.C., and Salas, J.A. (1998). Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring. Mol. Gen. Genet. *259*, 299–308.
31. 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.
32. Seto, N.O.L., Compston, C.A., Evans, S.V., Bundle, D.R., Narang, S.A., and Palcic, M.M. (1999). Donor substrate specificity of recombinant human blood group A, B and hybrid A/B glycosyltransferases expressed in *Escherichia coli*. Eur. J. Biochem. *259*, 770–775.
33. Bornscheuer, U.T., and Pohl, M. (2001). Improved biocatalysts by directed evolution and rational protein design. Curr. Opin. Chem. Biol. *5*, 137–143.
34. Liebeton, K., Zonta, A., Schimossek, K., Nardini, M., Lang, D., Dijkstra, B.W., Reetz, M., and Jaeger, K.-E. (2000). Directed evolution of an enantioselective lipase. Chem. Biol. *7*, 709–718.
35. Petrounia, I.P., and Arnold, F.H. (2000). Designed evolution of enzymatic properties. Curr. Opin. Biotechnol. *11*, 325–330.
36. Coco, W.M., Levinson, W.E., Crist, M.J., Hektor, H.J., Darzins, A., Pienkos, P.T., Squires, C.H., and Monticello, D.J. (2001). DNA shuffling method for generating highly recombined genes and evolved enzymes. Nat. Biotechnol. *19*, 354–359.
37. Barton, W.A., Lesniak, J., Biggins, J.B., Jeffrey, P.D., Jiang, J., Rajashankar, K.R., Thorson, J.S., and Nikolov, D.B. (2001). Structure, mechanism and engineering of a nucleotidylyltransferase as a first step toward glycorandomization. Nat. Struct. Biol. *8*, 545–551.
38. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
39. 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.
40. 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.
41. Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M., and Schrempf, H. (1985). Genetic Manipulation of *Streptomyces:* A Laboratory Manual (Norwich, UK: The John Innes Foundation).
42. Wilkinson, B., Foster, G., Rudd, B.A.M., Taylor, N.L., Blackaby, A.P., Sidebottom, P.J., Cooper, D.J., Dawson, M.J., Buss, A.D., Gaisser, S., et al. (2000). Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. Chem. Biol. *7*, 111–117.