Chemistry & Biology, Vol. 9, 1287–1296, December, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S1074-5521(02)00295-8

Remarkable Structural Similarities between Diverse Glycosyltransferases

Review

Yanan Hu and Suzanne Walker[1]
Department of Chemistry
Princeton University
Princeton, New Jersey 08544

From a functional standpoint, glycosyltransferases (GTases) comprise one the most diverse group of enzymes in existence. Every category of biopolymer (oligosaccharides, proteins, nucleic acids, and lipids) plus numerous natural products are modified by GTases, with remarkably varied effects. Given the structural and functional diversity of the products of glycosyl transfer combined with the often distant evolutionary relationships between glycosyltransferases, it is not surprising that sequence homologies between glycosyltransferases are low. What is surprising is that the majority of glycosyltransferases belong to only two structural superfamilies, implying that nature has come up with only a few solutions to the ubiquitous problem of how to catalyze glycosyl transfer. The conservation of GTase structure suggests that it will be simpler to manipulate glycosyltransferases for various applications than previously envisioned. A new age in glycoconjugate chemistry is beginning.

Introduction

Glycosyltransferases, enzymes that transfer sugars to other molecules, perform critical functions in all living organisms. They store energy in the form of glycogen; synthesize the carbohydrate polymers that support bacterial, fungal, and plant cell membranes; and make the cell surface oligosaccharides that mediate cell-cell recognition events and act as receptors for hormones, bacterial toxins, viruses, and a wide variety of circulating proteins [1–3]. In prokaryotes, glycosyltransferases modulate the activity of many important secondary metabolites, including vancomycin, erythromycin, and daunomycin [4]; in eukaryotes, they regulate the transcription of numerous genes, including those involved in glucose metabolism [5]. The astonishing range of functions in which glycosyltransferases participate is reflected in the diversity of structures they produce. Glycosyltransferases attach sugars to other sugars, to lipids, proteins, nucleic acids, polyketides, and nonribosomally synthesized peptides. Given the structural and functional diversity of the products of glycosyltransfer combined with the divergent evolutionary history of glycosyltransferases, it might be expected that glycosyltransferases themselves would display significant structural diversity. They certainly display significant diversity in terms of sequence [6] (P.M. Coutinho and B. Henrissat, Carbohydrate-Active Enzymes server at http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html). Surprisingly, in the past few years evidence has accumulated that the majority of glycosyltransferases belong to only two different superfamilies, recently named the GT-A and GT-B superfamilies [7, 8]. The GT-B superfamily is particularly remarkable for the diversity of products its members produce. The fact that nature has been able to use the same protein fold to glycosylate so many different types of molecules has numerous implications which will be discussed later in this review.

[1]Correspondence: swalker@princeton.edu

Background

Glycosyltransferases are enzymes that transfer sugars from an activated donor to another molecule. Examples of commonly used glycosyl donors are shown in Figure 1. The sugars are most commonly hexoses, and may have either the D- or L-configuration depending on the glycosyltransferase. The sugars utilized by prokaryotic glycosyltransferases are especially diverse and include a wide range of amino and deoxy sugars as well as more familiar sugars (Figure 2). Leaving groups include mono- and dinucleotides as well as various mono- and diphospholipids. The vast majority of glycosyltransferases utilize donors containing diphosphate leaving groups, with UDP/TDP leaving groups being by far the most common. In fact, of the 7000 known or putative glycosyltransferase sequences listed in the glycosyltranferase database, more than 60% are thought to be sequences of UDP/TDP-glycosyltransferases (Carbohydrate-Active Enzymes server at http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html). The molecules to which GTases transfer sugars, the glycosyl acceptors, include all categories of biopolymers—oligosaccharides, proteins, nucleic acids, and lipids—as well as numerous natural products (Figure 3). Taking into account both the range of structures they produce and the functions they are involved in, glycosyltransferases may be the single most diverse group of enzymes in existence.

Consistent with their functional diversity, glycosyltransferases display a high level of diversity in terms of their primary sequences. Before structural information on different glycosyltransferases became available, it was very difficult to draw any conclusions about glycosyltransferase structure or mechanism from sequence information because homologies were so low. In 1997, Campbell et al. [6] grouped glycosyltransferases into different families based on the identity of the donor sugar, the relative donor/product stereochemistry, and sequence homologies. There are now 62 different families of glycosyltransferases according to this classification scheme (http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html), a number that would lead one to conclude that there are many different glycosyltransferase folds. A large number of different folds have been identified for the glycosidases, enzymes that cleave glycosidic bonds [8, 9], and it would not be unreasonable to surmise that glycosyltransferases, which essentially catalyze the reverse reaction, adopt a similar number of different folds. However, this has not turned out to be the case. Most, if not all, UDP/TDP glycosyltransferases, which comprise by far the largest category of glycosyltransfer-

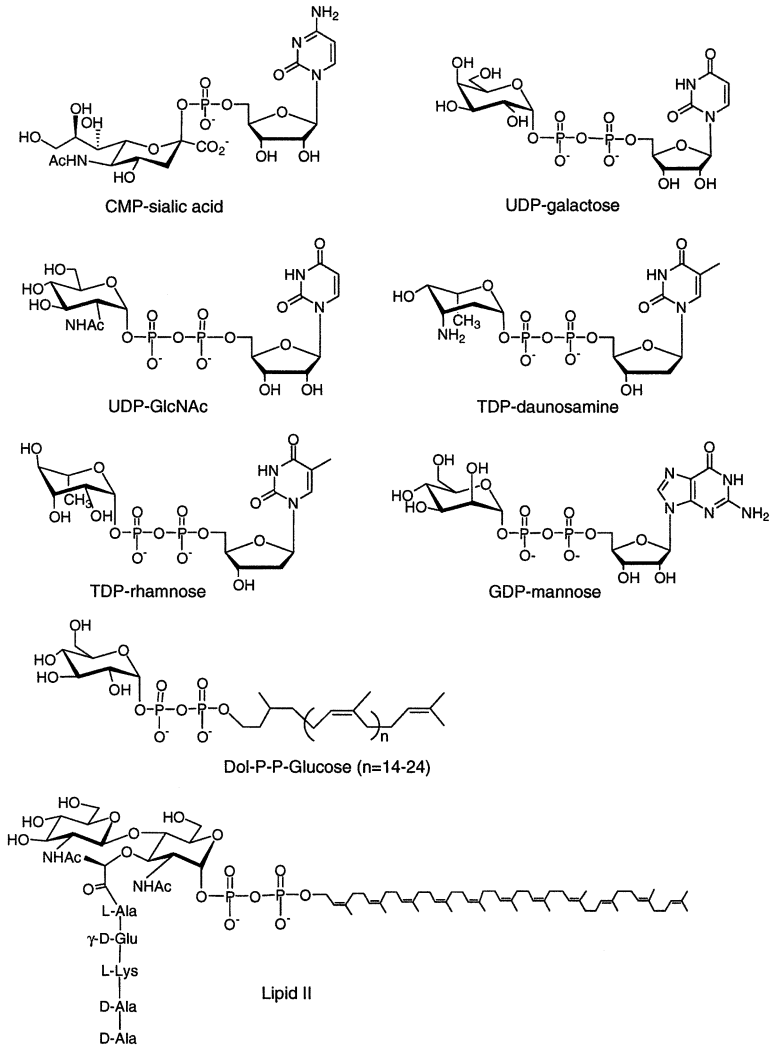

Figure 1. Examples of Glycosyl Donors

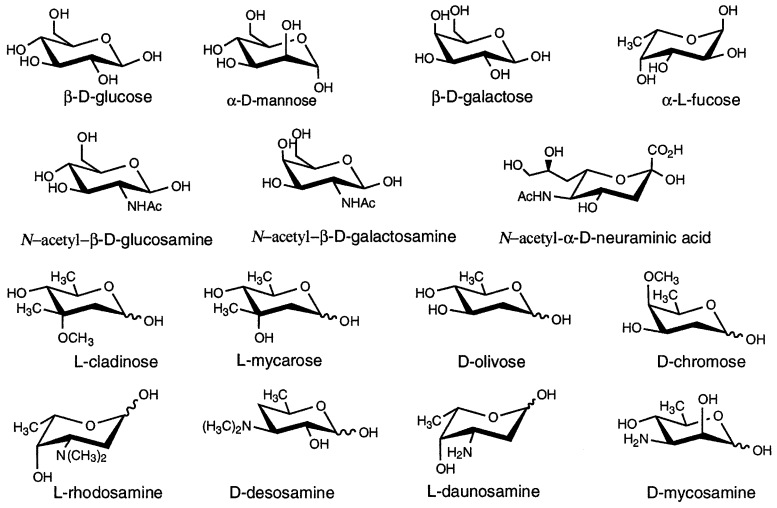

Figure 2. Selected Monosaccharides Transferred by Glycosyltransferases

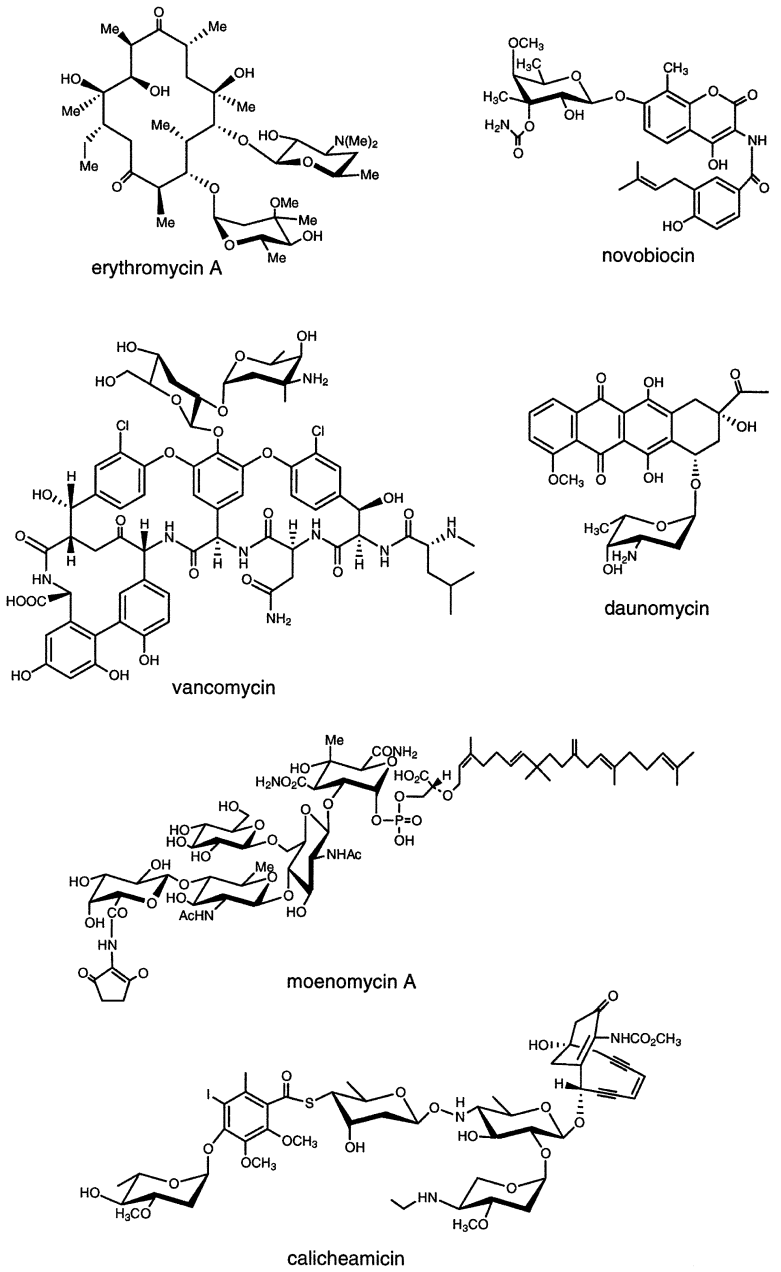

Figure 3. Biologically Active Natural Products Containing Sugar Moieties

ases, fall into only two different structural superfamilies [7, 8]. These superfamilies have different folds, different active sites, and different mechanisms, and they evidently represent two different solutions to the problem of how a protein can catalyze glycosyltransfer.

The GT-A Superfamily

The better studied UDP/TDP-glycosyltransferase superfamily is the so-called GT-A superfamily. Most of the Leloir pathway GTases that reside in the Golgi apparatus and the endoplasmic reticulum belong to this family, as do many prokaryotic GTases. This superfamily has been reviewed recently [7, 8, 10] and will not be discussed at length here. For purposes of comparison, however, the following features are noted. First, enzymes in the GT-A superfamily employ a DXD motif (or variant thereof) to bind a divalent metal ion (most commonly Mn^{2+}). The metal ion, which is essential for catalysis, helps anchor the pyrophosphoryl group of the UDP-sugar donor in the enzyme active site. The diversity of the acceptors used by GT-A superfamily members is relatively low in that they are almost exclusively other sugars. An ability to manipulate the ER and Golgi enzymes would enable new approaches to probe the myriad roles of cell-surface oligosaccharides [11]. However, the molecular basis for donor and acceptor selectivity is not clear even though ten crystal structures have been reported recently [12–20]. It is possible that spatial and temporal localization in the ER and Golgi influences the selectivity of these glycosyltransferases. In addition, the features

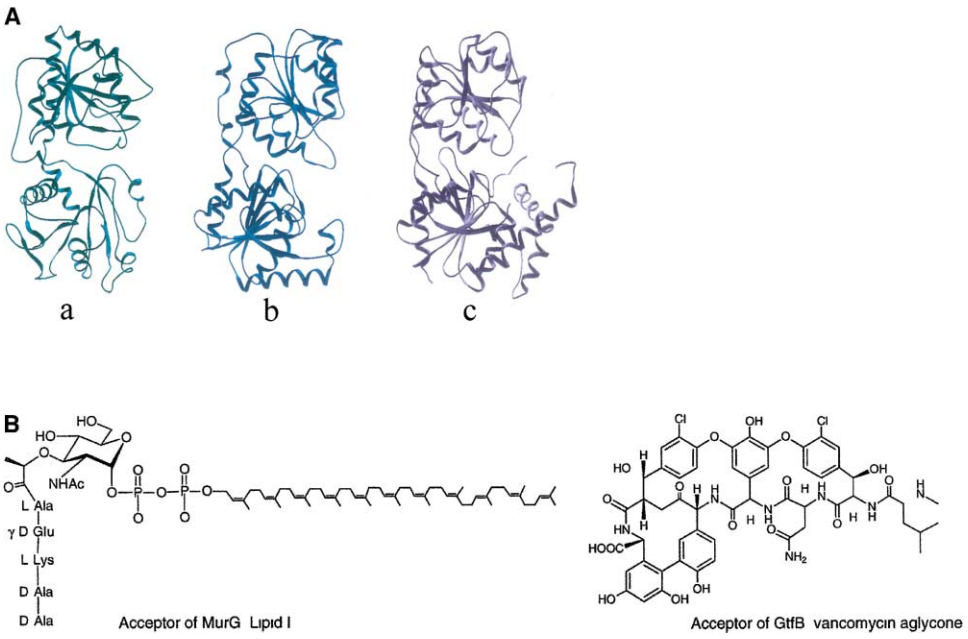

Figure 4. Structures of BGT, *E. coli* MurG, and GtfB with the Acceptors for MurG and GtfB

(A) Structure of three GT-B superfamily members, BGT (a), *E. coli* MurG (b), and GtfB (c). The C domains contain the primary donor binding site. The N domains of MurG and GtfB are proposed to contain the primary acceptor binding site. The figure was produced with Swiss-PdbViewer [61] and rendered by POV-Ray (downloaded from www.povray.org).

(B) The acceptors for MurG and GtfB.

responsible for discrimination can be subtle and thus obscured unless enzymes with substantial structural similarities are compared. For example, by comparing two enzymes with almost identical primary sequences, Patenaude et al. have shown that the ability of these enzymes to distinguish between UDP-GalNAc and UDP-Gal is due primarily to a single amino acid residue [20]. While there are still many unanswered questions regarding this superfamily, it has received much attention recently, and so our primary focus will be on the GT-B superfamily.

The GT-B Superfamily

The GT-B superfamily of glycosyltransferases is remarkably diverse. It includes most of the prokaryotic enzymes that glycosylate secondary metabolites to produce such biologically active natural products as erythromycin, daunomycin, vancomycin, and novobiocin [4]. It also includes prokaryotic enzymes involved in primary metabolic pathways such as cell wall biosynthesis [21]. It includes some of the Leloir pathway enzymes, such as the one that attaches UDP-galactose to ceramide in the biosynthesis of galactosylceramide [22]. It includes a very large number of putative glucuronosyltranferases, enzymes that glycosylate potentially toxic lipophilic compounds (such as bilirubin) for clearance from the body [23]. It includes at least 30% of all the glycosyltransferases found in *C. elegans* and a huge number of insect and plant glycosyltransferases. Finally, it appears to include the enzyme that may well be the single most important glycosyltransferase yet identified: O-GlcNAc transferase, or OGT [24], a GTase that posttranslationally modifies a wide variety of nuclear and cytoplasmic proteins and influences gene transcription in eukaryotes [5, 25, 26].

While there are crystal structures of ten different GT-A superfamily members, there are only three published crystal structures of GT-B superfamily members [21, 27–29]. (Glycogen phosphorylase is structurally related to the GT-B GTases and has been grouped with them [8] but is excluded from this discussion, because is it not a nucleotide-sugar transferase and does not provide obvious clues to the origins of donor and acceptor selectivity for NDP-sugar transferases [30, 31].) Nevertheless, this number is sufficient to reveal the striking structural similarities between family members. The structure of the first GT-B superfamily member was reported in 1994 by Vrielink et al. and belonged to a phage enzyme, T7 phage β-glucosyltransferase (BGT), which attaches glucose to hydroxymethyl cytosines on duplex DNA [27]. This glycosyltransferase structure was reported more than five years before another nucleotide-glycosyltransferase structure was reported, but its relevance to understanding other nucleotide-glycosyltransferases was impossible to assess at the time because BGT shares no useful sequence homologies with any other glycosyltransferases. In addition, the acceptor for BGT, which is a base in duplex DNA, is atypical.

BGT's prominence as the first member of the GT-B superfamily of glycosyltransferases became apparent when a second crystal structure of a family member was reported in 2000 [21]. This structure was of *E. coli* MurG, a glycosyltransferase that catalyzes the transfer of GlcNAc from UDP to the C4 position of a lipid-linked N-acetyl muramyl peptide acceptor to make a disaccharide that is the minimal subunit of the bacterial cell wall.

```
MurG                  261 RSGALTVSEIAAAGLPALFVPFQHKDRQQYWNALPLEKAGAAKI
GtfB                  308 HGGAGTTHVAARAGAPQILLPQMA---DQPYYAGRVAELGVGVA
GnT 1A                372 HAGSHGVYESICNGVPMVMMPLF---GDQMDNAKRMETKGAGVT
Ceramide 1-β-GalT     358 HGGLNSIFETMYHGVPVVGIPVF---GDHYDTMTRVQAKGMGIL
macrolide GT          321 HAGAGGSQEGLATATPMIAVPQA---ADQFGNADMLQGLGVART
daunosamin GT         324 HGGAGTWATAALHGVPQLALAWQ---WDDVFRAGQLEKLGAGIF
zeaxanthin GluT       324 HGGMNTYLDAINYRTPLLALPLA---FDQPGVASRIVYHGIGKR
oleandomycin GT       306 HAGAGGSQEGLATATPMIAVPQA---VDQFGNADMLQGLGVARK
flavonol GluT         341 HCGWNSILESISSCVPLICRPFF---GDQKLNSRMVQDSWKIGV
rhamnosylT            361 HAGFSSVIEALVNDCQVVMLPQK---GDQILNAKLVSGDMEAGV
baumycin GT           327 HGGSGTFMTALAHATPQLIVPDM--MWDAMEKAHGLARSGAGGY
MGD                   357 KAGPGTIAESLIRSLPIILNDYI--PGQEKGNVPYVVENGAGVP
Consensus                 H_G__T__E____G P____P_____DQ__NA_____G_G__
```

Figure 5. Sequence Alignments of Selected GT-B Superfamily Members Showing the Conserved Motif that Corresponds to the α/β/α Fold

The glycosyltransferases represent a range of prokaryotic and eukaryotic glycosyltransferases that play roles in both primary and secondary metabolic processes. In *E. coli* MurG, the consensus E residue (highlighted in bold) plays a role in binding to the hydroxyls on the ribose sugar of the UDP group. The conserved threonine located at the amino terminus of the first α helix helps anchor the α phosphate (along with backbone amides from adjacent residues), and the polar residues (consensus DQ), located in the loop immediately preceding the second α helix, anchor the hexose sugar.

As shown in Figure 4, the overall topologies of BGT and MurG are virtually identical even though the proteins are only about 10% homologous on the sequence level. Both structures reveal a two-domain structure in which each domain adopts an α/β open sheet motif similar to a classic Rossmann fold. Neither BGT nor MurG was found to contain a bound metal ion, and kinetic studies have shown that neither BGT nor MurG requires metal ions for activity. Nevertheless, the rates of both enzymes are accelerated by certain cations for reasons that are not understood [32]. Structural and kinetic analyses of other enzymes in the GT-B superfamily similarly indicate that divalent cations are not essential for activity [29, 33]. In this respect, the GT-B superfamily is fundamentally different from the GT-A superfamily.

The third crystal structure of a GT-B family glycosyltransferase was reported last year by Walsh, Garavito, and coworkers [29]. This structure is of GtfB, a UDP-glucosyltranferase that attaches glucose to the A4 phenol of the chloroeremomycin (or vancomycin) aglycone in the biosynthesis of the glycopeptide antibiotic chloroereomomycin. As shown in Figure 4, GtfB is also topologically almost identical to MurG and to BGT. The overall structural resemblance is remarkable, given that the acceptors for these enzymes are completely unrelated.

Genomic Analysis of GT-B Superfamily Members

Sequence analysis of GT-B family members combined with structural information can provide useful information rapidly about what features of these enzymes are important for binding and catalysis. MurG has been particularly useful as a paradigm for understanding the GT-B class of glycosyltransferases because it plays an essential role in a metabolic pathway that is highly conserved in bacteria, i.e., peptidoglycan synthesis. All bacteria that make peptidoglycan contain MurG homologs that catalyze virtually the same reaction [21]. However, the sequences of MurG homologs from different organisms can vary considerably depending on the evolutionary relationship between them. For example, the homology between between *E. coli* and *E. faecalis* MurG is only 30% even though the structures are predicted to be almost identical. Thus, by comparing sequences from many different organisms it is possible to identify the regions that are most critical for catalysis because these regions are invariant. The invariant residues in MurG homologs are confined to only five regions in the protein, all of which are located at or near the cleft between the Rossmann domains [21]. The longest conserved motif is located in the donor binding domain and consists of a pattern of prolines and glycines with a few other invariant residues interspersed in the sequence (Figure 5). The crystal structure of *E. coli* MurG revealed that this sequence motif encodes an α/β/α subdomain. A similar sequence motif can be identified in most other GT-B superfamily members and has, in fact, become a kind of signature that allows one to identify members of the superfamily quickly [21, 24, 29, 33]. This sequence motif is so highly conserved because it encodes a folding unit that is intimately involved in binding the glycosyl donors. A cocomplex of *E. coli* MurG with UDP-GlcNAc bound reveals the function of the α/β/α folding unit (Y. Hu et al, submitted). This structure reveals that the first α helix in the subdomain makes key contacts to the furanose on the nucleotide, while the second α helix and the loop preceding it contact the pyranose. The negatively charged oxygen of the α phosphate on the glycosyl donor is anchored at the amino terminus of the first α helix, where it is apparently stabilized by the positively charged helix dipole.

It is worth commenting that the type of fold observed in the glycosyl-donor binding domain of GT-B family members, the Rossmann fold, was first characterized for proteins that bind diphosphate-containing cofactors such as NAD(H) [34]. The negatively charged pyrophosphate portions of these cofactors can bind to Rossmann domains without any requirement for positively charged side chains or metal ions because they exploit the stabilization provided by helix dipole effects [35]. Given that glycosyltransferases utilize diphosphate-containing substrates, it is not surprising in retrospect to find that nature has coopted the Rossmann motif for use by the GT-B family of glycosyltransferases. Thus, this motif represents one solution to how to bind NDP-sugar donors. Another solution which is utilized by the GT-A superfamily of glycosyltransferases involves metal-ion coordination of the negatively charged pyrophosphoryl group [7, 8].

Whereas analysis of MurG sequences has been particularly useful for identifying elements of GT-B superfamily members that play essential roles in binding and catalysis, analysis of GtfB-related sequences has been useful for complementary reasons. GtfB belongs to a family of glycosyltransferases that attach sugars to different sites on a very similar group of peptide antibiotics (Figure 6). The sequences of the enzymes within this family are highly homologous even though they utilize different acceptors. The high degree of homology strongly sug-

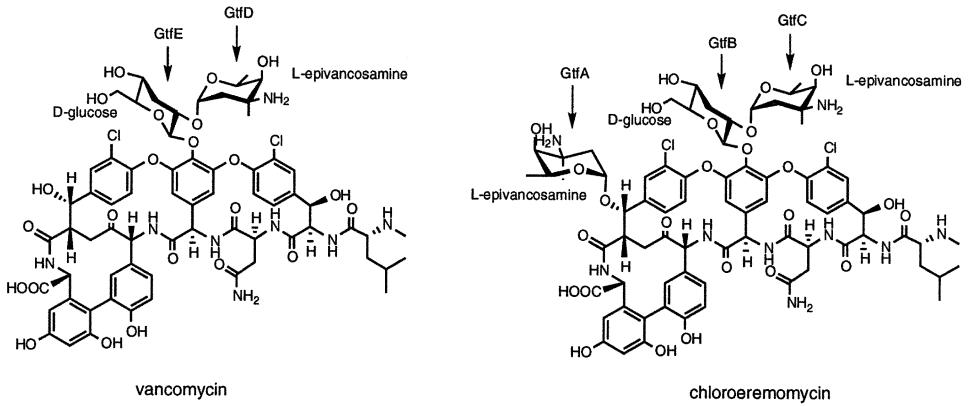

Figure 6. Glycosylation Pattern of Vancomycin Group Antibiotics

The enzymes responsible for the glycosyl transfer are shown above the indicated sugar. Adapted from [29].

gests that all of these enzymes originated recently from the same parent enzyme. The differences among these enzymes are presumed to reflect structural adaptations to accommodate different acceptors. Therefore, it is possible to identify which regions one might vary to alter selectivity by comparing sequence alignments for the family. The peptide segment that varies most with changes in the acceptor structure follows strand Nβ5 in the N-terminal domain, and this region was proposed to be the acceptor binding site for each of the Gtf enzymes [29].

Sequence analysis of *E. coli* MurG homologs shows that the loop between Nβ5 and Nα5 is invariant across all MurG homologs, indicating its importance for the function of this enzyme. The Nβ5-Nα5 loop in MurG was proposed to be the acceptor binding site [21], like the corresponding region of GtfB. The fact that a loop identified as invariant for a sequence-diverse group of enzymes that use the same acceptor corresponds to a loop identified as varying for a sequence-homologous group of enzymes that use slightly different acceptors raises the confidence level that focus on this region is one way to alter selectivity.

Figure 7 shows three of the five invariant regions of MurG juxtaposed with the corresponding regions of GtfB, and the structural correspondence is notable. The α/β/α motif, which constitutes a large part of the donor binding site, is almost superimposable except that the loop between Cβ5 and Cα5 is longer in MurG than in GtfB. A sulfate in the GtfB structure is located exactly where the α-phosphate of the glycosyl donor is located in MurG. Furthermore, several of the amino acids that anchor the GlcNAc moiety superimpose on counterparts located at the exact same position in GtfB. These comparisons and the high conservation of the α/β/α motif suggest that the donor sugars are held in a similar manner in many members of the superfamily, regardless of the nature of the acceptor. Since nature has clearly used the same two-domain architecture throughout evolution to attach hexoses to an enormous range of different acceptors, we have concluded that it should be possible to alter the acceptor selectivity by varying the region between Nβ5 and Nα5 while maintaining the overall architecture and preserving certain other key features or residues. It should also be possible to tune the donor selectivity by making judicious amino acid changes in the α/β/α domain.

Before concluding this discussion of what regions of the GT-B glycosyltransferases play key roles in influencing selectivity, it is important to note work reported recently by Bechthold and coworkers on two glycosyltransferases involved in urdamycin biosynthesis [36, 37]. These enzymes share 91% sequence identity but have different acceptor and donor selectivities. Sequence differences between these enzymes are largely located between amino acids 50–80 in the N-terminal region of the protein. The authors have shown that it is possible to alter selectivity for both the donor and the acceptor by varying this region. Based on existing structures of enzymes within the GT-B superfamily, this region is predicted to correspond to the crossover region between Nβ3 and Nβ4 (Figure 7C). The same region in the Gtf family members is also variable depending on the acceptor. Thus, this region in the GT-B superfamily can also have a profound influence on selectivity, although the structural basis for the influence on donor selectivity is not clear. It is possible that the structure of the acceptor can affect donor selectivity. In this regard, it should be noted that a study of two oleandomycin glycosyltransferases indicates that the acceptors for these enzymes bind first in a compulsory ordered mechanism [38, 39]. Acceptor binding thus facilitates donor binding and could influence donor selectivity.

Additional structural information on GTases within the GT-B superfamily would be useful for learning more about how to alter substrate selectivity. While structures of three GT-B superfamily members currently exist, and complexes with intact UDP-glycosyl donor and UDP product have been obtained [27, 28, 32] (Y. Hu et al, submitted), it would be helpful to have structures containing bound glycosyl acceptors. Such structures might shed light on how the crossover region in the N domain influences selectivity. It would also be useful to have additional donor-sugar complexes where the hexoses differ significantly from glucose or GlcNAc. Ternary complexes would be especially valuable for providing additional insight into the mechanism. Finally, the value of sequence comparisons is inestimable and can help

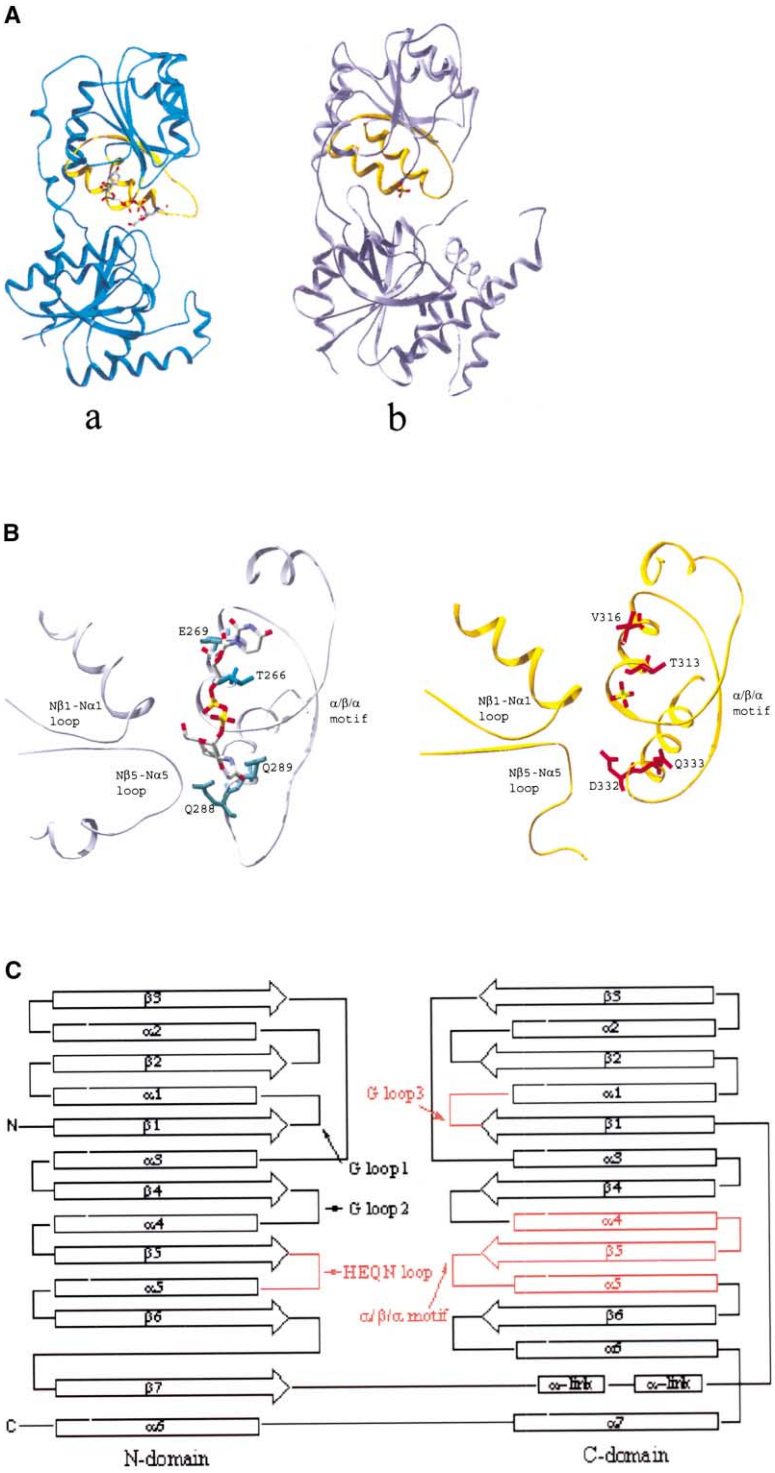

Figure 7. Juxtaposition of MurG:UDP-GlcNAc Cocomplex and GtfB

(A) Full structures with α/β/α motif shown in gold.

(B) Selected invariant regions in MurG (Gloop 1, Nβ1-Nα1 loop; HEQN loop, Nβ5-Nα5 loop; and α/β/α motif) juxtaposed with corresponding regions from GtfB. The loop between Cβ5 and Cα5 in MurG (part of the α/β/α motif) is characteristically longer than in other GT-B family members, including GtfB. The Nβ5-Nα5 loop of GtfB, the proposed acceptor binding site, contains a long proline-rich polypeptide, and the electron density is not continuous to Nα5 so only a portion of this region is shown. The figure was produced with Swiss-PdbViewer [61] and rendered by POV-Ray (www.povray.org).

(C) Topology diagram of MurG. The key regions in the GT-B superfamily are highlighted in red.

focus attention on which enzymes are useful targets for structure analysis. The examples of the Gtf and urdamycin families emphasize the value of comparing enzymes that have different selectivities but very similar sequences for understanding donor and acceptor selectivity. Sequence comparisons of divergent enzymes, on the other hand, can provide more insight into key structural features of the superfamily as well as important catalytic residues.

Glycosyltransferases in Synthesis: Prospects

Many biologically active natural products contain sugars, and the sugar components play crucial roles in determining the biological activity of the compounds. In some natural products, for example, it has been shown that changes to the carbohydrates affect the activity dramatically. The sugars on erythromycin are required for biological activity [40], while the sugars on vancomycin can be modified to overcome vancomycin resistance

[41–43]. Thus, it is known that attaching different or unnatural sugars to biologically active molecules can alter both the spectrum of activity and the potency when compared with the parent natural product [44].

It would be useful to be able to attach a wider range of sugars to naturally produced aglycones such as erythromycin and vancomycin. Unfortunately, this has not been straightforward to accomplish. Until recently, the best way to attach unnatural sugars to natural products was by using chemical glycosylation methods. These methods have serious limitations for glycosylating complex systems. Yields are often low and, except for 1,2-trans glycosidic linkages where neighboring group participation can be used, the stereochemical control is often poor. Furthermore, protecting group schemes have to be worked out to mask reactive groups and protect and solubilize natural product aglycones in the organic solvents compatible with chemical glycosylation reactions. Thus, although dramatic progress is being made in some areas of oligosaccharide synthesis such as automated oligosaccharide synthesis, the substitution of one sugar for another on a complex natural product remains difficult [45–47].

Enzymatic glycosylation is emerging as the solution of choice to the problem of how to glycosylate natural product aglycones and their derivatives to introduce diversity in the carbohydrate portions of the molecules. Enzymes react in water to give specific stereochemical and regiochemical outcomes, obviating the need for protecting groups on either the acceptor or the donor. What was thought to be a major limitation of enzymes—namely, their specificity for a particular substrate—appears not to be as significant a barrier as imagined.

It has been recognized for quite some time that many enzymes in the GT-A superfamily do not have stringent selectivity and can thus be used to make a wide range of unnatural oligosaccharides. Progress in this area has been reviewed recently by, inter alia, Koeller and Wong [48–50]. Furthermore, with crystallographic information available it has become possible to make structure-based mutations that alter or broaden donor selectivity [51]. Glycosidases can also be engineered to utilize glycosyl fluorides as donors in glycosyltransfer reactions, and there is every expectation that these "glycosynthases" can be further engineered for novel selectivities using structure-based methods or directed evolution [52, 53]. Nevertheless, the majority of the GT-A superfamily enzymes and the glycosynthases are likely biased toward sugar-sugar couplings, which may mean that their greatest utility is in making oligosaccharides. Given how many biologically active oligosaccharides with therapeutic potential there are, this is not a drawback. However, if the goal is to make glycoconjugates where the acceptors are natural product aglycones or secondary metabolites, then the GT-B family offers the greatest opportunity to achieve this enzymatically.

The GT-B superfamily appears to be nature's preferred solution to how to glycosylate a wide range of structurally unrelated aglycones. Although these aglycones are sometimes other sugars, as in the case of MurG homologs, they can also be peptides of both ribosomal and nonribosomal origin, nucleic acids, polyketides, assorted lipids, terpenes, etc. Studies examining the donor and acceptor selectivity of GT-B superfamily members are at a relatively early stage compared to studies on GT-A superfamily members, but it is already clear that there is enormous potential to use GT-B glycosyltransferases to generate diverse sets of glycoconjugates. Furthermore, the structural data on glycosyltransferases in the GT-B superfamily suggests that it will be possible to engineer/evolve glycosyltransferases to have new and/or loosened acceptor selectivity by manipulating defined regions of the enzymes. It has also been suggested that the two-domain architecture of GT-B family glycosyltransferases implies that these enzymes may be modular, making domain swapping a potentially useful strategy for changing acceptor selectivity [54].

GT-B glycosyltransferases involved in the biosynthesis of secondary metabolites appear, in general, to have relaxed selectivity, perhaps because there is not much selection pressure for them to make a specific glycoside. They can therefore be used to make many different variations on a particular natural product to probe the role of the sugar moiety. For example, Walsh, Kahne, and coworkers have recently shown that glycosyltransferases that attach sugars to different dalbaheptide scaffolds in the biosynthesis of glycopeptide antibiotics will utilize unnatural acceptors [55]. Furthermore, it is possible to replace hydroxyls in the hexoses utilized by these glycosyltransferases with amines, providing a handle for introducing other functional groups to increase diversity [56]. One drawback to using glycosyltransferases for synthesis is that the glycosyl donors can take considerable effort to make. Thorson and coworkers have recently shown that it is possible to make a range of structurally diverse TDP- or UDP-sugars by exploiting the relaxed specificity of selected hexose-1-phosphate nucleotidylytransferases [40, 57, 58]. Furthermore, Wong and others have developed methods to regenerate nucleotide-sugar donors in situ [59, 60], making it possible to drive reactions to completion and reducing the expense of synthesizing nucleotide-sugar donors. We expect that it will eventually be possible to use combinations of glycosyltransferases and the enzymes that produce glycosyl donors to make libraries of glycoconjugates for screening purposes.

Conclusion

The GT-B superfamily of glycosyltransferases shares a highly conserved two-domain architecture. By comparing structures and sequences of different family members, it is possible to identify the regions that determine selectivity. It should be possible to alter those regions to achieve new selectivities and/or to relax the existing selectivity. Glycosyltransferases with new selectivities will be very useful for the combinatorial synthesis of biologically active glycoconjugates. Reengineered glycosyltransferases also have potential in vivo applications as tools for probing the roles of glycosyltransferases and their products. Finally, reengineered glycosyltransferases may be useful for altering cellular pathways involving glycosyltransferases, providing additional tools to use in the remodeling of cell surfaces [11].

References

1. Gagneux, P., and Varki, A. (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. Glycobiology *9*, 747–755.
2. Monsigny, M., Midoux, P., Mayer, R., and Roche, A.C. (1999). Glycotargeting: influence of the sugar moiety on both the uptake and the intracellular trafficking of nucleic acid carried by glycosylated polymers. Biosci. Rep. *19*, 125–132.
3. Reuter, G., and Gabius, H.J. (1999). Eukaryotic glycosylation: whim of nature or multipurpose tool? Cell. Mol. Life Sci. *55*, 368–422.
4. Weymouth-Wilson, A.C. (1997). The role of carbohydrates in biologically active natural products. Nat. Prod. Rep. *14*, 99–110.
5. Zachara, N.E., and Hart, G.W. (2002). The emerging significance of O-GlcNac in cellular rgulation. Chem. Rev. *102*, 431–438.
6. 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*, 929–939.
7. Unligil, U.M., and Rini, J.M. (2000). Glycosyltransferase structure and mechanism. Curr. Opin. Struct. Biol. *10*, 510–517.
8. Bourne, Y., and Henrissat, B. (2001). Glycoside hydrolases and glycosyltransferases: family and functional modules. Curr. Opin. Struct. Biol. *11*, 593–600.
9. Henrissat, B., and Davies, G.J. (1997). Structural and sequence-based classification of glycoside hydrolases. Curr. Opin. Struct. Biol. *7*, 637–644.
10. Breton, C., Mucha, J., and Jeanneau, C. (2001). Structural and functional features of glycosyltransferases. Biochimie *83*, 713–718.
11. Bertozzi, C.R., and Kiessling, L.L. (2001). Chemical glycobiology. Science *291*, 2357–2364.
12. 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.
13. Gastinel, L.N., Cambillau, C., and Bourne, Y. (1999). Crystal structures of the bovine beta4galactosyltransferase catalytic domain and its complex with uridine diphosphogalactose. EMBO J. *18*, 3546–3557.
14. Rao, M., and Tvaroska, I. (2001). Structure of bovine alpha-1,3-galactosyltransferase and its complexes with UDP and UDPGal inferred from molecular modeling. Proteins *44*, 428–434.
15. Gastinel, L.N., Bignon, C., Misra, A.K., Hindsgaul, O., Shaper, J.H., and Joziasse, D.H. (2001). Bovine alpha1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases. EMBO J. *20*, 638–649.
16. Persson, K., Ly, H.D., Dieckelmann, M., Wakarchuk, W.W., Withers, S.G., and Strynadka, N.C. (2001). Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs. Nat. Struct. Biol. *8*, 166–175.
17. Pedersen, L.C., Darden, T.A., and Negishi, M. (2002). Crystal structure of beta 1,3-glucuronyltransferase I in complex with active donor substrate UDP-GlcUA. J. Biol. Chem. *277*, 21869–21873.
18. Ramakrishnan, B., Balaji, P.V., and Qasba, P.K. (2002). Crystal structure of beta-1,4-galactosyltransferase complex with UDP-Gal reveals an oligosacharide acceptor binding site. J. Mol. Biol. *318*, 491–502.
19. Gibbons, B.J., Roach, P.J., and Hurley, T.D. (2002). Crystal structure of the autocatalytic initiator of glycogen biosyntheis, glycogenin. J. Mol. Biol. *319*, 463–477.
20. Patenaude, S.I., Seto, N.O.L., Borisova, S.N., Szpacenko, A., Marcus, S.L., Palcic, M.M., and Evans, S.V. (2002). The structural basis for specificity in human ABO(H) blood group biosynthesis. Nat. Struct. Biol. *9*, 685–690.
21. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000). The 1.9 A crystal structure of *Escherichia coli* MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. Protein Sci. *9*, 1045–1052.
22. Kapitonov, D., and Yu, R.K. (1999). Conserved domains of glycosyltransferases. Glycobiology *9*, 961–978.
23. King, C.D., Rios, G.R., Green, M.D., and Tephly, T.R. (2000). UDP-glucuronosyltransferases. Curr. Drug Metab. *1*, 143–161.
24. Wrabl, J.O., and Grishin, N.V. (2001). Homology between O-linked GlcNAc transferases and proteins of the glycogen phosphorylase superfamily. J. Mol. Biol. *314*, 365–374.
25. Vosseler, K., Wells, L., and Hart, G.W. (2001). Nucleocytoplasmic O-glycosylation: O-GlcNAc and functional proteomics. Biochimie *83*, 575–581.
26. Havover, J.A. (2001). Glycan-dependent signaling: O-linked N-acetylglucosamine. FASEB J. *15*, 1865–1876.
27. Vrielink, A., Ruger, W., Driessen, H.P., and Freemont, P.S. (1994). Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. EMBO J. *13*, 3413–3422.
28. Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W., and Freemont, P.S. (1999). T4 phage beta-glucosyltransferase: substrate binding and proposed catalytic mechanism. J. Mol. Biol. *292*, 717–730.
29. 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.
30. Artymiuk, P.J., Rice, W.D., Poirrette, A.R., and Willet, P. (1995). β-glucosyltransferase and phophorylase reveal their common theme. Nat. Struct. Biol. *2*, 117–120.
31. Holm, L., and Sander, C. (1995). Evolutionary link between glycogen phosphorylase and a DNA modifying enzyme. EMBO J. *14*, 1287–1293.
32. Morera, S., Lariviere, L., Kurzeck, J., Aschke-Sonnenborn, U., Freemont, P.S., Janin, J., and Ruger, W. (2001). High resolution crystal structures of T4 phage beta-glucosyltransferase: induced fit and effect of substrate and metal binding. J. Mol. Biol. *311*, 569–577.
33. Abdian, P.L., Lellouch, A.C., Gautier, C., Ielpi, L., and Geremia, R.A. (2000). Identification of essential amino acids in the bacterial alpha-mannosyltransferase aceA. J. Biol. Chem. *275*, 40568–40575.
34. Rossmann, M.G., Moras, D., and Olsen, K.W. (1974). Chemical and biological evolution of nucleotide-binding proteins. Nature *250*, 194–195.
35. Hol, W.G., van Duijnen, P.T., and Berendsen, H.J. (1978). The alpha-helix dipole and the properties of proteins. Nature *273*, 443–446.
36. 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.
37. 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.
38. Quiros, L.M., and Salas, J.A. (1995). Biosynthesis of the macrolide oleandomycin by Streptomyces antibioticus. J. Biol. Chem. *270*, 18234–18239.
39. Quiros, L.M., Carbajo, R.J., Brana, A.F., and Salas, J.A. (2000). Glycosylation of macrolide antibiotics: furification and kinetic studies of a macrolide glycosyltransferase from *Streptomyces antibioticus*. J. Biol. Chem. *275*, 11713–11720.
40. He, X.M., and Liu, H.-W. (2002). Formation of unusual sugars: mechanistic studies and biosynthetic applications. Annu. Rev. Biochem. *71*, 701–754.
41. Ge, M., Chen, Z., Onishi, H.R., Kohler, J., Silver, L.L., Kerns, R., Fukuzawa, S., Thompson, C., and Kahne, D. (1999). Vancomycin derivatives that inhibit peptidoglycan biosynthesis without binding D-Ala-D-Ala. Science *284*, 507–511.
42. Kerns, R., Dong, S.D., Fukuzawa, S., Carbeck, J., Kohler, J., Silver, L., and Kahne, D. (2000). The role of hydrophobic substituents in the biological activity of glycopeptide antibiotics. J. Am. Chem. Soc. *122*, 12608–12609.
43. Sun, B., Chen, Z., Eggert, U.S., Shaw, S.J., LaTour, J.V., and Kahne, D. (2001). Hybrid glycopeptide antibiotics. J. Am. Chem. Soc. *123*, 12722–12723.

44. Dong, S.D., Oberthur, M., Losey, H.C., Anderson, J.W., Eggert, U.S., Peczuh, M.W., Walsh, C.T., and Kahne, D. (2002). The structural basis for induction of VanB resistance. J. Am. Chem. Soc. *124*, 9064–9065.
45. Kahne, D. (1997). Combinatorial approaches to carbohydrates. Curr. Opin. Chem. Biol. *1*, 130–135.
46. Seeberger, P.H., and Haase, W.-C. (2000). Solid-phase oligosaccharide synthesis and combinatorial carbohydrate libraries. Chem. Rev. *100*, 4349–4394.
47. Plante, O.J., Palmacci, E.R., and Seeberger, P.H. (2001). Automated solid-phase synthesis of oligosaccharides. Science *291*, 1523–1527.
48. Koeller, K.M., and Wong, C.-H. (2000). Synthesis of complex carbohydrates and glycoconjugates: enzyme-based and programmable one-pot strategies. Chem. Rev. *100*, 4465–4493.
49. Elhalabi, J.M., and Rice, K.G. (1999). Synthesis and applications for unnatural sugar nucleotides. Curr. Med. Chem. *6*, 93–116.
50. Ohrlein, R. (1999). Glycosyltransferase-catalyzed synthesis of non-natural oligosaccharides. Top. Curr. Chem. *200*, 227–254.
51. Ramakrishnan, B., and Qasba, P.K. (2002). Structure-based design of β1,4-galactosyltransferase I (β4gal-T1) with equally efficient N-acetylgalactosaminyltransferase activity. J. Biol. Chem. *277*, 20833–20839.
52. Ly, H.D., and Withers, S.G. (1999). Mutagenesis of glycosidases. Annu. Rev. Biochem. *68*, 487–522.
53. Mayer, C., Jakeman, D.L., Mah, M., Karjala, G., Gal, L., Warren, R.A.J., and Withers, S.G. (2001). Directed evolution of new glycosynthases from *Agrobacterium* β-glucosidase: a general screen to detect enzymes for oligosaccharide synthesis. Chem. Biol. *8*, 437–443.
54. Khosla, C., and Harbury, P.B. (2001). Modular enzymes. Nature *409*, 247–252.
55. Losey, H.C., Peczuh, M.W., Chen, Z., Eggert, U.S., Dong, S.D., Pelczer, I., Kahne, D., and Walsh, C.T. (2001). Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglcones: novel glycopeptides. Biochemistry *40*, 4745–4755.
56. Losey, H.C., Jiang, J., Biggins, J.B., Oberthür, M., Ye, X.-Y., Dong, S.D., Kahne, D., Thorson, J.S., and Walsh, C.T. (2002). Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides. Chem. Biol. *9*, 1305–1314, this issue.
57. Jiang, J., Biggins, J.B., and Thorson, J.S. (2000). A general enzymatic method for the synthesis of natural and "unnatural" UDP- and TDP-nucleotide sugars. J. Am. Chem. Soc. *122*, 6803–6804.
58. 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 nucleotidylytransferase as a first step toward glycorandomization. Nat. Struct. Biol. *8*, 545–551.
59. Wong, C.-H., Haynie, S.L., and Whiteside, G. (1982). Enzyme-catalyzed synthesis of N-acetyllactosamine within situ regeneration of uridine 5′-diphosphate glucose and uridine 5′-diphosphate galactose. J. Org. Chem. *47*, 5416–5418.
60. Chen, X., Fang, J., Zhang, J., Liu, Z., Shao, J., Kowal, P., Andreana, P., and Wang, P.G. (2001). Sugar nucleotide regeneration beads(superbeads): a versatile tool for the practical synthesis of ologosaccharides. J. Am. Chem. Soc. *123*, 2081–2082.
61. Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis *18*, 2714–2723.

Note Added in Proof

The data referred to throughout as "Y. Hu et al., submitted" are now in press: Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003). Crystal structure of MurG:UPD-GlcNAc complex reveals common strucutral principles of a superfamily of glycosyltransferases. Proc. Natl. Acad. Sci. USA, in press.