Chemistry & Biology
Article

Chimeric Glycosyltransferases for the Generation of Hybrid Glycopeptides

Andrew W. Truman,[1,*] Marcio V.B. Dias,[2] Shu Wu,[1] Tom L. Blundell,[2] Fanglu Huang,[1] and Jonathan B. Spencer[1,3]
[1]University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, England, UK
[2]Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, England, UK
[3]Died April 6, 2008
*Correspondence: awt26@cam.ac.uk
DOI 10.1016/j.chembiol.2009.04.013

SUMMARY

Glycodiversification, an invaluable tool for generating biochemical diversity, can be catalyzed by glycosyltransferases, which attach activated sugar "donors" onto "acceptor" molecules. However, many glycosyltransferases can tolerate only minor modifications to their native substrates, thus making them unsuitable tools for current glycodiversification strategies. Here we report the production of functional chimeric glycosyltransferases by mixing and matching the N- and C-terminal domains of glycopeptide glycosyltransferases. Using this method we have generated hybrid glycopeptides and have demonstrated that domain swapping can result in a predictable switch of substrate specificity, illustrating that N- and C-terminal domains predominantly dictate acceptor and donor specificity, respectively. The determination of the structure of a chimera in complex with a sugar donor analog shows that almost all sugar-glycosyltransferase binding interactions occur in the C-terminal domain.

INTRODUCTION

The presence of carbohydrate appendages is often crucial for bioactive natural products to function properly (Varki et al., 1999). Glycosyltransferases (GTs) thus represent a significant class of enzyme, catalyzing the attachment of activated sugar "donors" onto "acceptor" molecules, and glycodiversification is an invaluable tool for generating biochemical diversity (Thibodeaux et al., 2007; Salas and Mendez, 2007) The engineering of bioactive natural product biosynthetic pathways represents an attractive option to generate analogs that possess superior medicinal properties. Examples of engineered glycodiversification strategies include gene disruption of the TDP-desosamine pathway in *S. venezuelae* to generate novel analogs of pikromycin (Borisova et al., 1999), using the reversibility of GTs to exchange sugars between natural products (Minami et al., 2005; Zhang et al., 2006), and the introduction of deoxysugar biosynthesis plasmids into natural product producers (Perez et al., 2006; Schell et al., 2008).

These strategies rely on the native GT possessing relaxed substrate specificity toward the sugar donor, the aglycone acceptor, or both, yet many GTs cannot tolerate anything more than minor modifications to their native substrates. Such high specificity makes many GTs unsuited to current glycodiversification strategies, significantly reducing their applicability. Examples include GtfA, Orf1, and Orf10*, GTs involved in the biosynthesis of the glycopeptides chloroeremomycin (Lu et al., 2004), and teicoplanin (Li et al., 2004; Howard-Jones et al., 2007). However, thus far, few unnatural GT-B GTs have been constructed with the aim of altering substrate specificity; notable examples include GTs involved in the biosynthesis of urdamycin (Hoffmeister et al., 2001, 2002) and the directed evolution of a macrolide GT (Williams et al., 2007).

Glycopeptides are clinically indispensable antibiotics that are decorated with a variety of sugars. The sugars that decorate glycopeptides can increase antibacterial activity by enhancing membrane localization (Beauregard et al., 1995), can aid in glycopeptide dimerization (Mackay et al., 1994), and have even been postulated to be antibacterial without the attached peptide (Ge et al., 1999). All glycopeptide GTs that use nucleotide diphospho-sugar (NDP-sugar) donors belong to the GT-B structural superfamily. The GT-B fold is characterized by a bilobal architecture, with two Rossmann-like $\alpha\beta$ domains facing each other. It is believed that the N-terminal domain contains the acceptor site and the C-terminal domain contains the NDP-sugar donor site; catalysis occurs at the interface of these domains (Coutinho et al., 2003). Structurally characterized examples of this structural superfamily include MurG (Hu et al., 2003), DNA β-glucosyltransferase (Vrielink et al., 1994), and OleD (Bolam et al., 2007). An analysis of the CAZy (carbohydrate-active enzymes) database reveals that over 19,000 sequenced GTs are predicted to adopt this fold.

Domain swapping is a common engineering strategy for biological systems that have distinct substrate binding and catalytic or regulatory regions, yet GT domain swapping is relatively unexplored. A few GT domain swapping examples have been reported, including moving the C domains between UDP-glucosyltransferases (Cartwright et al., 2008) and chimeras of 91% identical urdamycin GTs (Hoffmeister et al., 2001, 2002). An early domain swapping study by Mackenzie (1990) identified that the N domain of rat UDP-glucuronyltransferases is responsible for acceptor binding. In our study we report on the generation and characterization of three chimeric glycopeptide GTs and show that N- and C-terminal domains predominantly dictate acceptor and donor specificity, respectively. Activity analyses reveal these chimeras to possess activities comparable to their parents. Using X-ray crystallography we have solved the structure of GtfAH1,

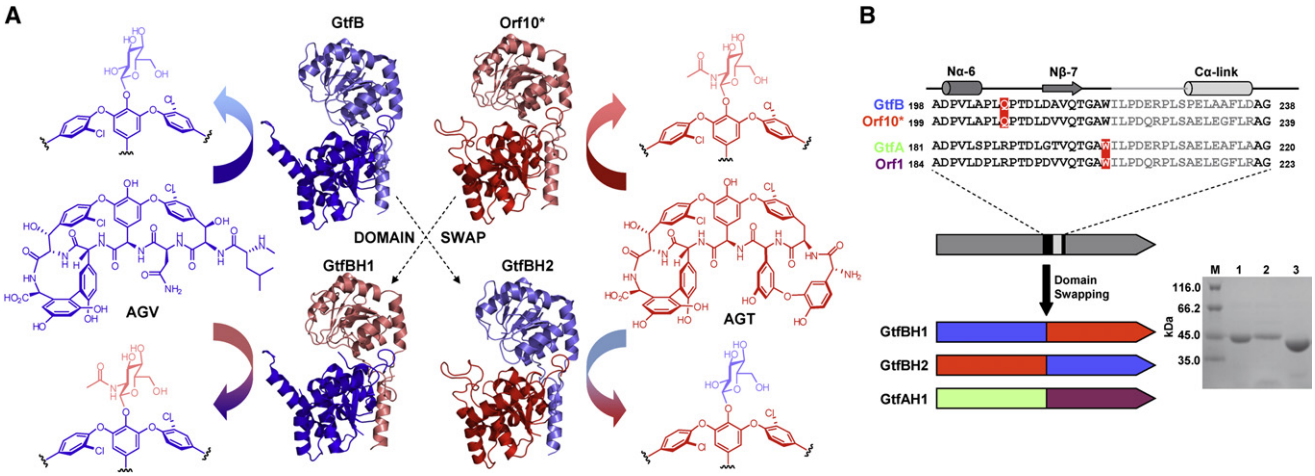

Figure 1. Mixing and Matching GT-B Donor and Acceptor Domains

(A) A schematic of the concept applied to the glycopeptide GTs GtfB and Orf10*. GtfB (Mulichak et al., 2001; PDB 1IIR) and Orf10* structure (predicted using the Phyre fold recognition server [Bennett-Lovsey et al., 2008]). Figures were generated using PyMOL.

(B) Fusion sites used for construction of chimeric GTs; 10% SDS-PAGE: M, protein ladder; 1, GtfBH1; 2, GtfBH2; 3, GtfAH1.

a remarkably active and promiscuous chimeric GT, to 1.15 Å resolution in complex with UDP and to 1.30 Å resolution in complex with UDP-2-deoxy-2-fluoro-glucose. This shows that almost all sugar-GT binding interactions occur in the C-terminal domain for glycopeptide GTs. Taken together, these results demonstrate the viability of generating highly active chimeric GTs with novel and predictable activities.

RESULTS

Design and Construction of GtfB Chimeras

The presence of two distinct substrate binding domains makes chimeric GT-B GTs an attractive bioengineering target, and the glycopeptide GTs GtfB and Orf10* were used to assess the feasibility of functional chimeric GTs. These GTs are responsible for transferring glucose (Glc) and *N*-acetylglucosamine (GlcNAc) to 4-OH-PheGly$_4$ of the vancomycin and teicoplanin aglycones (AGV and AGT), respectively (Figure 1A) (Li et al., 2004; Mulichak et al., 2001). Orf10* and GtfB possess significant sequence similarity (70% identity) and perform analogous glycosylations, yet nevertheless act on different donor and acceptor substrates. Success with such a model system would indicate the feasibility of generating functional chimeric GTs by fusing heterologous donor and acceptor domains.

The linker region for GtfB and Orf10* has been approximately defined as amino acids 219–227 (random coil region) and 228–235 (C-terminal linking α helix). Gln206 was chosen as the fusion site as it corresponds to a common PstI restriction site in both genes (Figure 1B). Although this residue precedes the linker region, there is very high similarity between GtfB and Orf10* within this region (90.5% identity from aa 206–226), so it was anticipated that this site would not disrupt the N-terminal domain. Expression constructs were created by digesting gtfB-pET28a and orf10*-pET28a with PstI and MluI then heterologously re-ligating the resulting fragments. GtfBH1 (N-terminal GtfB, C-terminal Orf10*) and GtfBH2 (N-terminal Orf10*, C-terminal GtfB) were then expressed in *E. coli* BL21(DE3) with N-terminal hexahistidine tags and subjected to in vitro analysis of their reactivity and specificity.

Both chimeras were expressed as soluble proteins (Figure 1B), and protein identity was confirmed by liquid chromatography/electrospray ionization tandem mass spectrometry (LC-ESI-MS): GtfBH1 observed mass = 44,493 Da (predicted mass for GtfBH1 after loss of *N*-formyl-methionine = 44,498 Da); GtfBH2 observed mass = 44,577 Da (predicted mass for GtfBH2 after loss of *N*-formyl-methionine = 44,584 Da). The solubility of these proteins indicates that the domain swap does not result in severe protein misfolding, although expression yields appear to be related to the yield of the GT that provides the N terminus (purified protein yields from parallel 1 liter cultures: GtfB = 7.6 mg; Orf10* = 1.6 mg; GtfBH1 = 4.6 mg; GtfBH2 = 1.0 mg).

Analysis of GtfB Chimera Activity and Specificity

GtfBH1 and GtfBH2 were tested for GT activity with AGT, AGV, uridine diphospho-glucose (UDP-Glc), and UDP-GlcNAc (Table 1). We also used competitive assays to probe donor specificity by incubating equimolar amounts of UDP-GlcNAc and UDP-Glc with the appropriate aglycone and GT. The reaction turnover rates of the chimeric GTs indicate that their catalytic efficiencies are comparable to their parent GTs. Furthermore, the relative rates of the GTs with AGV and AGT demonstrate that the N-terminal domain entirely controls acceptor binding. Hence, the acceptor specificities of GtfBH1 and GtfBH2 correspond to the specificities of GtfB and Orf10*, respectively. GtfBH1 and GtfB can glycosylate both AGV and AGT, whereas Orf10* and GtfBH2 can only transfer sugars to AGT. The data also shows that the C-terminal domain predominantly controls sugar specificity, although its control is not as absolute as in the case of the N-terminal domain. GtfBH1 has the C domain of Orf10* but it does not possess the same high selectivity for UDP-GlcNAc as Orf10*. Competitive assays indicate that Orf10* and GtfB both have a >100-fold preference for their preferred sugar, while GtfBH1 only has a 1.48-fold preference for GlcNAc over Glc. In contrast, the sugar specificity of GtfBH2 is comparable to the

Table 1. Activity of Hybrid GtfB-like Proteins in Comparison to Parent GTs

	AGT		AGV		Competitive Assays	
GT	UDP-GlcNAc	UDP-Glc	UDP-GlcNAc	UDP-Glc	GlcNAc/Glc ratio	AGV/AGT ratio
Orf10*	**1.000**	0.575	**-**	-	130:1	-
GtfB	-	0.080	0.039	**0.803**	0.006:1	5.1:1
GtfBH1	0.004	-	**0.141**	0.122	1.48:1	15.1:1
GtfBH2	0.003	**0.578**	-	-	0.014:1	-

Relative activities are normalized to Orf10* activity with AGT and UDP-GlcNAc (4.93 μM per hour per micromole of enzyme). A dash denotes no activity detected by LC-ESI-MS analysis. All enzymes were tested at 20°C with 1 mM acceptor and 2 mM donor and all data are the average of three experiments. Numbers in bold reflect both the maximal activity for the enzyme and the particular set of substrates used.

specificity of GtfB and GtfBH2 can only transfer GlcNAc to AGT at a highly reduced rate. Hence the C-terminal domain appears to be the strongest determinant of sugar specificity but additional regions are also influential. The different specificities of the chimeras may indicate that there is a GtfB N-domain interaction with glucose but no Orf10* N-domain interaction with *N*-acetylglucosamine.

Design and Construction of GtfA Chimeras

The creation of two functional hybrids from Orf10* and GtfB acted as a promising proof of principle for the mixing and matching of GT-B domains to create GTs with novel and predictable activity. This initial success encouraged us to apply domain swapping to a more functionally disparate pair of GTs. A much less glycodiversified site is the benzylic hydroxyl group of β-OH-Tyr$_6$. GtfA and Orf1 transfer *epi*-vancosamine and *N*-acetylglucosamine onto this position on desvancosaminyl vancomycin (DVV) and teicoplanin glucosaminyl-pseudoaglycone **1**, respectively (Figure 2A; Lu et al., 2004; Li et al., 2004). As with GtfB/Orf10*, these two proteins possess significant similarity (66% identity). Crucially, GtfA and Orf1 use very different sugar donors and are very specific toward their native substrates, so no cross-reactivity exists. A consequence of this specificity is that few glycopeptide analogs have been produced by glycodiversification at β-OH-Tyr$_6$. TDP-L-β-4-*epi*-vancosamine is an L-trideoxyhexose which has different substituents at every position around the ring compared to UDP-D-α-4-*N*-acetylglucosamine. In addition, GlcNAc is attached to β-OH-Tyr$_6$ by a β linkage, whereas *epi*-vancosamine is attached via an α linkage. With such differences, a rational site-directed mutagenesis approach is unlikely to switch specificity.

pET28a(+) plasmids containing the chimeric genes were constructed in an analogous fashion to the GtfB chimeras. Trp201 was chosen as the fusion site for the GtfA-Orf1 chimeras. This residue is at the beginning of the unstructured random coil region that links the two substrate-binding domains and corresponds to a common BamHI site in the two genes. One factor that was not assessed for the GtfB chimeras is the effect, if any, of the site of the His$_6$-tag, so constructs were generated for both N- and C- terminally His$_6$-tagged chimeras. GtfAH1 (N-terminal GtfA, C-terminal Orf1) was expressed in *E. coli* BL21(DE3) as a hexahisitidine tagged protein. The GtfAH2 hybrid (N-terminal Orf1, C-terminal GtfA) was also expressed as a soluble protein but its activity could not be investigated due to TDP-*epi*-vancosamine unavailability (GtfAH2 is not functional with glucose-like sugar donors).

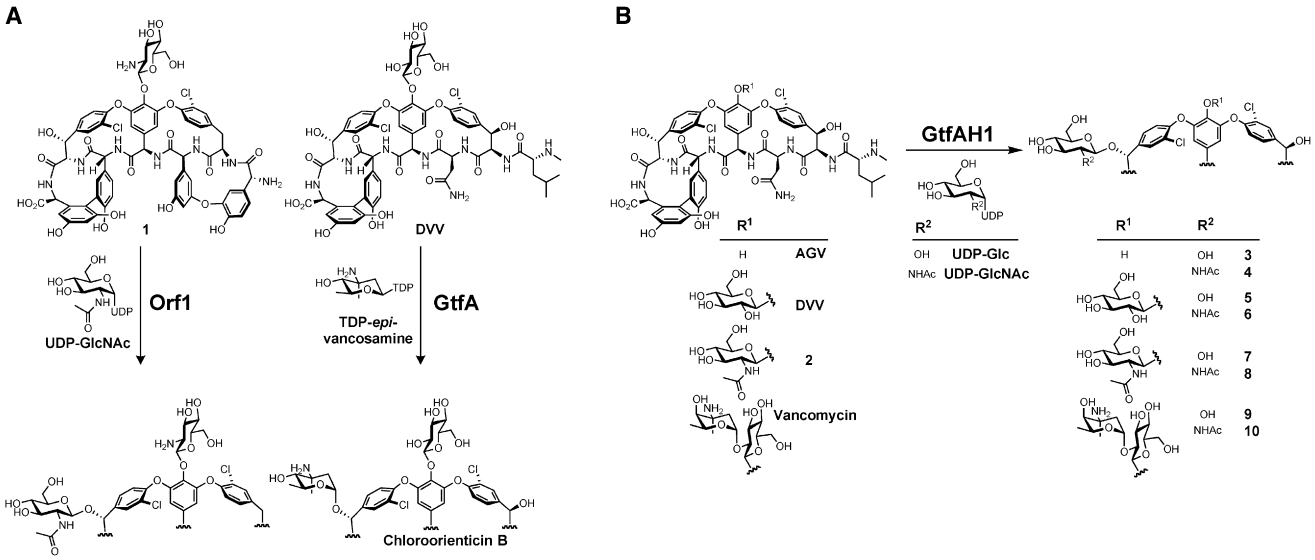

Figure 2. GtfA-Orf1 Domain Swapping

(A) Native activity of GTs Orf1 and GtfA.

(B) Hybrid vancomycin derivatives produced by GtfAH1, a chimera generated by a fusion of the N-terminal domain of GtfA with the C-terminal domain of Orf1.

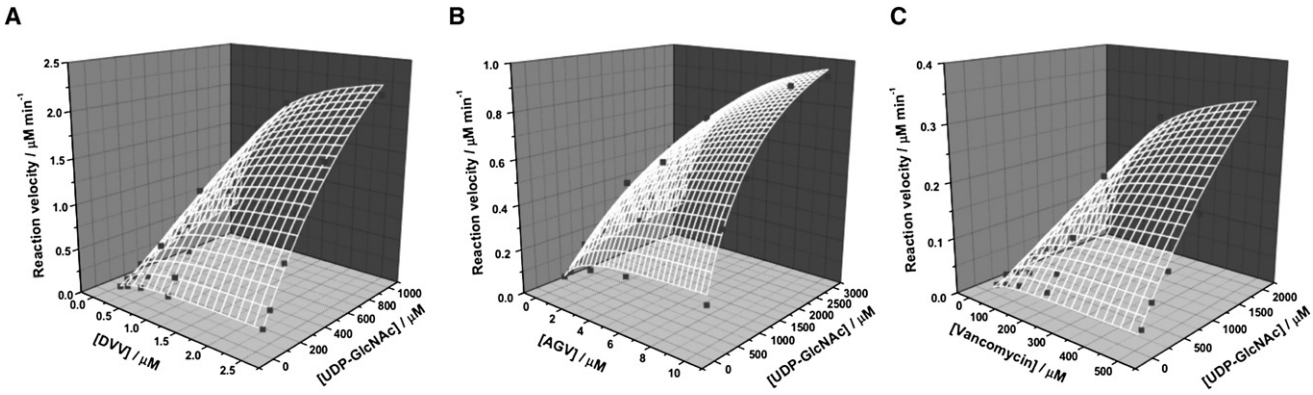

Figure 3. Three-Dimensional Plots of Initial Reaction Velocities of GtfAH1 with Acceptor Glycopeptides and UDP-GlcNAc

All points represent the average of three experiments. The white surface mesh shows the nonlinear curve fitting parameters fitted to this data using Origin 7.

(A) DVV as acceptor.

(B) AGV as acceptor.

(C) Vancomycin as acceptor.

Functional Analysis of GtfAH1

GtfAH1 was assayed with AGV, DVV, AGT, vancomycin, and *N*-acetylglucosaminyl-AGV **2** as acceptor molecules and UDP-GlcNAc/Glc as sugar donors. Remarkably, activity was observed with all substrate combinations (Figure 2B). The position of the GtfAH1 His_6 tag did not affect enzymatic activity. The ability of GtfAH1 to glycosylate all these glycopeptide acceptors (see Figure S1 available online for examples of its activity) is highly unexpected, given the very strict substrate specificity of its parent GTs. Wild-type GtfA possesses only trace activity with AGV or vancomycin (Lu et al., 2004), thus implying strict N-domain specificity, while we detected no activity when Orf1 was tested with UDP-glucose and AGT or **1**. GtfAH1 donor specificity between UDP-Glc and UDP-GlcNAc was assessed using competitive assays, which showed that GtfAH1 only exhibits a 2.2-fold preference for *N*-acetylglucosamine over glucose, regardless of the acceptor molecule (Figure S2). Therefore, domain swapping appears to have significantly loosened substrate specificity.

Kinetic Analysis of GtfAH1

The broad specificity of GtfAH1 clearly demonstrates the potential of hybrid GTs to generate previously difficult to access natural product derivatives. To probe the biosynthetic value of GtfAH1 we determined its two-substrate kinetic parameters with UDP-GlcNAc and DVV/AGV/vancomycin. A two-substrate kinetic analysis generates two binding constants per substrate (K_{iA} and K_{mA}), where K_{mA} is the Michaelis constant for substrate A. The true meaning of K_{iA} depends on whether the substrates bind in an ordered or random manner (Cornish-Bowden, 2004). If substrate A binds first in an ordered mechanism, K_{iA} is the disassociation constant for the enzyme A complex. To determine a full set of kinetic parameters, initial rates were required for a range of both substrate concentrations.

Kinetic parameters were obtained by nonlinear curve fitting using Origin 7 (Figure 3 and Table 2). A comparison with the parameters reported by the Walsh group for GtfA (Lu et al., 2004) and Orf1 (Howard-Jones et al., 2007) shows that GtfAH1 is a highly active chimera that has a surprisingly high affinity

Table 2. Two-Substrate Kinetic Parameters for GtfAH1 with UDP-GlcNAc and Glycopeptide Acceptor Substrates

		Acceptor Parameters		UDP-GlcNAc Parameters	
Glycopeptide Substrate		GtfAH1	GtfA[a]	GtfAH1	Orf1[b]
DVV	k_{cat}/min^{-1}	6.20 ± 0.77	2.3 ± 0.5	6.22 ± 0.76	7.6 ± 0.1
	$K_m/\mu M$	0.51 ± 0.15	107	1398 ± 284	32 ± 6
	$K_i/\mu M$	0.30 ± 0.06	ND	827 ± 229	ND
AGV	k_{cat}/min^{-1}	1.64 ± 0.14	<0.05	1.64 ± 0.14	0.38 ± 0.01[c]
	$K_m/\mu M$	2.84 ± 0.53	ND	1166 ± 216	807 ± 66[c]
	$K_i/\mu M$	0.84 ± 0.45	ND	347 ± 190	ND
Vancomycin	k_{cat}/min^{-1}	1.80 ± 0.05	<0.05[d]	1.80 ± 0.05	NA
	$K_m/\mu M$	352 ± 49	ND	6960 ± 2451	NA
	$K_i/\mu M$	30.7 ± 9.5	ND	609 ± 246	NA

Errors quoted represent the standard error of curve fitting. ND, not determined.

[a] Parameters obtained from Lu et al. (2004).

[b] Parameters obtained from Howard-Jones et al. (2007).

[c] Data refer to Orf1 donor parameters with AGT (Howard-Jones et al., 2007).

[d] k_{cat} is quoted for GtfA activity with epivancomycin (Lu et al., 2004).

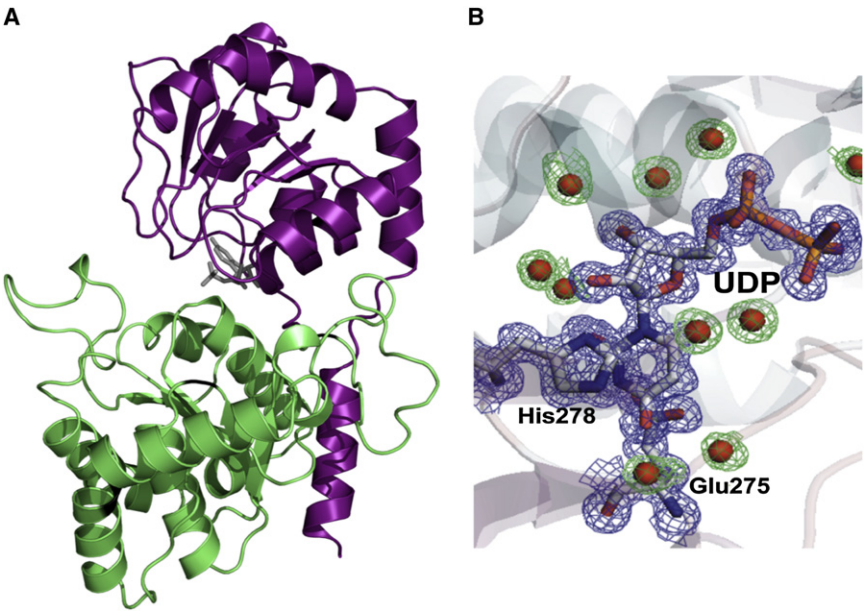

Figure 4. Structural Analysis of GtfAH1

(A) 1.15 Å resolution structure of GtfAH1 containing bound UDP (gray) with GtfA N-terminal domain colored green and Orf1 C-terminal domain colored purple.

(B) 2Fo − 2Fc electron density map of the uracil ring of UDP sandwiched between His278 and the Glu275-Arg11 salt bridge. The electron density for UDP, His278, Glu275, and water molecules present in the UDP binding site is illustrated.

toward DVV and AGV when compared to the DVV:GtfA and Orf1:**1** K_m values (0.51 μM and 2.84 μM versus 107 μM and 85 μM, respectively). Additionally, the K_m of vancomycin with GtfAH1 (352 μM) is comparable to the parent GT K_m values listed above, while all k_{cat} values are of a similar magnitude to the wild-type activities of GtfA and Orf1.

Although substrate specificity has been broadened, there is still a definite trend regarding the relative activity of GtfAH1 with DVV, AGV, and vancomycin. The K_m and K_i increase in the order DVV < AGV < vancomycin, whereas k_{cat} decreases in the order DVV > AGV ≈ vancomycin. This evidently reflects an N domain that is optimized to bind DVV and is least capable in accommodating the additional steric bulk of vancomycin. Inhibition studies are usually used to determine whether a mechanism proceeds via a random or an ordered mechanism. However, a study by Yang et al. (2005) indicated that, for GTs, a ratio of K_{mA}/K_{mB} greater than ten strongly implies an ordered mechanism. A K_m(UDP-GlcNAc)/K_m(DVV) ratio of 2741 indicates that DVV binds first as part of an ordered mechanism in GtfAH1.

Isothermal titration calorimetry was used to verify the low-micromolar affinity between GtfAH1 and DVV. This provides a dissociation constant (K_D) that should be a similar magnitude to K_m and K_i. C- and N-terminally His_6-tagged GtfAH1 were analyzed, along with C-terminally His_6-tagged GtfA (Figure S3). This analysis reinforced the data obtained by kinetic analysis as the K_D for DVV with GtfAH1 was determined to be 2.26 μM. Along with the kinetic data, this tight binding is the clearest indication yet that GT domain swapping can generate biosynthetically relevant enzymes. Interestingly, the K_D of DVV with GtfA is over twenty times stronger than the kinetically determined K_m (4.85 μM compared to 107 μM; Lu et al., 2004). The basis for this difference has not been elucidated, although may be a simple difference in protein preparation and/or its subsequent stability. Alternatively, it may be the difference in GtfA-DVV affinity depending on whether TDP-*epi*-vancosamine is bound to the enzyme or not.

Structural Analysis of GtfAH1

The activity and promiscuity of GtfAH1 demonstrates that domain swapping can not only provide novel GT activity, but can also broaden substrate tolerance while improving catalytic efficiency. To analyze the structural determinants that control specificity and activity we determined the crystal structure of C-terminally His_6-tagged GtfAH1 to 1.15 Å resolution, the highest achieved for a GT-B fold GT (Figure 4A). Statistics for data collection and refinement are listed in Table 3. GtfAH1 possesses characteristic GT-B bilobal architecture, with two facing Rossmann-like αβ domains and contains UDP in the C-domain nucleotide binding site, either acquired prior to purification or by hydrolysis of co-crystallized UDP-GlcNAc. The uracil ring is sandwiched between an Arg11-Glu275 ion pair and potential pi stacking with His278 (Figure 4B). Comparison with the open and closed structural conformations of GtfA (Mulichak et al., 2003) suggests that loop 256–270 swings to allow for a hydrophobic interaction between the side chain of Trp258 and loops 57–60 (VRAG) and 229–231 (GSG), indicating that the GtfAH1:UDP complex exists in the closed conformation, as would be expected. A structural superposition of GtfA and GtfAH1 shows that Trp258 occupies an almost identical position to Trp260 in GtfA (Figure S4A).

These trans-domain interactions are fully analogous to those that occur in the wild-type GtfA:TDP complex; the ability of GtfAH1 to adopt this conformation illustrates that such structural subtleties can be retained in a chimera. However, the presence of these trans-domain interactions may complicate the generation of future chimeras, especially between more distantly related GTs. To investigate the regularity of a trans-domain Arg-Glu salt bridge to cap the pyrimidine ring, a sequence alignment was performed across 136 GT-B bacterial natural product GTs (Figure S5). This shows that the salt bridge is confined to the glycopeptide GTs, the putative rifampin GT Rgt (NCBI accession number AAK84835), and the mycinamicin GT MycD (Anzai et al., 2003). The majority of GTs appear to use a C-terminal aromatic residue that aromatically caps the pyrimidine ring (61% of sequences analyzed). This mechanism is observed in the structures of OleD (Bolam et al., 2007), *Vv*GT1 (Offen et al., 2006), and MurG (Hu et al., 2003). The trans-domain tryptophan interaction also appears to be confined to the glycopeptide GTs (data not shown).

The structures of the N-terminal domains of GtfA and GtfAH1 are almost identical (aa 1–200; backbone rmsd = 0.355 Å), while

Table 3. Crystallographic Data Collection and Refinement Statistics

	GtfAH1:UDP	GtfAH1:UDP-2F-Glc
Data collection statistics		
Space group	$C222_1$	$C222_1$
Cell parameters (Å)	a = 111.1, b = 129.7, c = 67.1	a = 111.5, b = 129.8, c = 67.4
Wavelength (Å)	0.977	0.98
Resolution range (Å) (outer shell)	42.80–1.15 (1.21–1.15)	29.92–1.30 (1.37–1.30)
No. of total reflections (outer shell)	115,4832 (130,919)	838,486 (121,662)
No. of unique reflections (outer shell)	169,485 (23,415)	119,259 (17,259)
Multiplicity (outer shell)	6.8 (5.6)	7.0 (7.0)
R_{merge} (%) (outer shell)	7.7 (64.7)	7.4 (56.9)
Completeness (%) (outer shell)	99.2 (94.7)	99.9 (100)
Refinement statistics		
Resolution range (Å)	42.80–1.15	29.25–1.30
R_{cryst} (%)[a]	17.1	16.4
R_{free} (%)[b]	18.5	18.4
No. of reflections	160,958	113,256
No. of protein residues	383	392
No. of water molecules	608	630
No. of phosphates	4	2
Ligand molecules	1 UDP	1 UDP-2F-Glc
Average B_{factor} (Å^2)		
Protein (main, side, and whole chain)	9.4, 10.9, 10.1	10.6, 17.6, 14.5
Phosphate	29.1	40.2
UDP or UDP-2F-Glc	7.9	14.8
Water	26.4	28.8
PDB code	3H4T	3H4I

[a] $R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.
[b] $R_{free} = R_{factor}$ for 5% of the data that were not included during crystallographic refinement.

their C-terminal domains are highly similar despite significant differences in sequences and sugar specificities (aa 218–389; sequence identity = 65%; backbone rmsd = 0.703 Å). The most significant structural difference between the two proteins, and a possible rationale for increased acceptor promiscuity, is the $_{318}$NVVE$_{321}$ loop that is present in GtfA but not in GtfAH1. This loop has been proposed to act in a trans-domain manner to assist in DVV binding (Mulichak et al., 2003), although the structure of GtfAH1 indicates the contrary. Without this loop, the turn between Cβ5 and Cα5 (GtfAH1 aa 312–315) is now within H-bonding distance of DVV. In particular, Lys313 is well positioned to H bond with the C terminus of DVV (Figure S4B). Such additional interactions may explain this chimera's enhanced affinity toward DVV.

GtfAH1 Complex with UDP-2-deoxy-2-fluoro-glucose

Owing to the weak natural hydrolytic activity of GtfAH1, we were unable to obtain the crystal structure of a co-complex with UDP-GlcNAc. Sugar recognition in glycopeptide GTs is poorly understood, especially as the sugars utilized can differ considerably, and a structure with an intact NDP-sugar would structurally rationalize the observed switch in sugar specificities seen with the GT chimeras.

We therefore co-crystallized GtfAH1 with UDP-2-deoxy-2-fluoro-glucose (UDP-2F-Glc; Gordon et al., 2006). The highly electronegative 2-fluoro group withdraws electron density from the anomeric carbon where glycosyltransfer occurs. This effect is sufficient to render the molecule unreactive toward glycosyltransfer or hydrolysis, thus making it an excellent GT inhibitor. A structure of GtfAH1 in complex with UDP-2F-Glc at a 1.30 Å resolution was obtained (see Figure S6 for a numbered stereo view), where the electron density for UDP-2F-Glc is clearly visible in a position expected for UDP-GlcNAc (Figure 5A). A comparison with the GtfAH1:UDP complex shows that there is very little structural reorganization after the sugar is transferred to an acceptor molecule (backbone rmsd = 0.097 Å). Despite the surprisingly low K_i and K_m for DVV with GtfAH1, it was not possible to obtain a co-complex including DVV, possibly due to unfavorable crystal packing when a glycopeptide is bound.

As expected, there is a network of H bonds between the sugar and the C-terminal domain (Figure 5B). These include H bonds between the side chain of Asp315 with O3 and O4 of UDP-2F-Glc (2.51 Å and 3.68 Å, respectively) and between the Gln316 side chain with O3 and/or F2 (Gln316 may function as an H bond acceptor/donor with the 2-NAc moiety of GlcNAc). The array of interactions that this DQ motif makes suggests it has a key role in glucose binding. Additionally, the backbone amide of Ala294 H bonds with O4.

Asp13, a putative base for catalysis, is 3.54 Å from the likely position of the DVV hydroxyl, based on a structural superposition

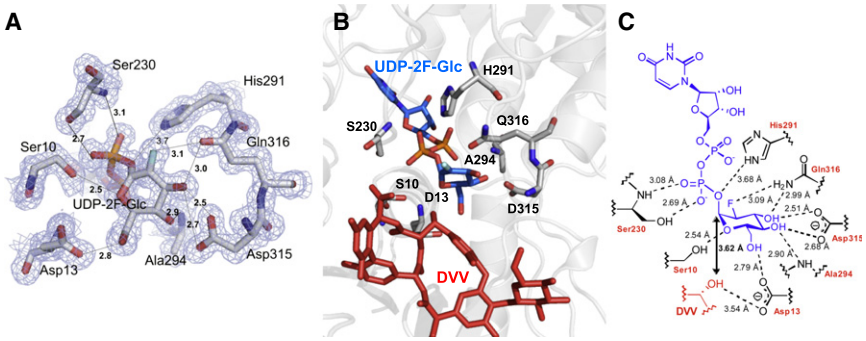

Figure 5. Analysis of UDP-2F-Glc Binding Residues in the GtfAH1:UDP-2F-Glc Complex

(A) 2Fo − 2Fc electron density of UDP-2F-Glc and interacting GtfAH1 residues at 1.30 Å resolution.

(B) Significant active site residues and fitted DVV structure (fitted from N-terminal structural superposition with GtfA:DVV complex; PDB code 1PN3).

(C) Schematic of GtfAH1 active site H-bonding network.

with the GtfA:DVV structure, and is also within H-bonding distance (2.79 Å) of O6 of UDP-2F-Glc. The distance from the DVV nucleophilic hydroxyl to the anomeric carbon is 3.62 Å, so it is likely that a small degree of structural reorganization is required to adopt a reactive conformation. Ser10, another possible active site base (Mulichak et al., 2003), forms an H bond with the sugar ring oxygen (2.54 Å). His291 is 3.68 Å away from the leaving group oxygen on UDP, so it could possibly stabilize the charge generated. Ser230 (side chain and main chain) and Gly12 both H bond with the UDP phosphate groups. With the exception of residues 10–13, which are highly conserved across the GT-B superfamily, no N-terminal residues interact with the sugar. This supports the hypothesis that the C-terminal domain dictates sugar binding affinity, at least for glycopeptide GTs.

Owing to the number of important interactions it makes, it appears possible that the DQ motif could be an important determinant for selecting glucose-like donor sugars. This observation is supported by structures of UDP-Glc and UDP-2F-Glc with the plant GTs UGT71G1 (Shao et al., 2005) and *Vv*GT1 (Offen et al., 2006): both use the DQ motif to bind glucose in an analogous manner to GtfAH1. A structural alignment with *Vv*GT1 (Figure S7) reveals remarkably high structural similarity in the sugar-binding region despite massive evolutionary divergence (22% overall sequence identity).

To investigate whether this motif is a key feature for selecting glucose-like sugars, 136 bacterial natural product GTs (see above) were analyzed. Of these, 25 contain a DQ or EQ motif in the correct region (Figure S8). Despite the structural information detailed above, many of these GTs do not transfer glucose-like sugars, indicating that sugar selectivity must be determined by additional interactions in these cases. For example, the residues that H bond 2F-Glc in GtfAH1 are conserved throughout all the glycopeptide GTs, even those that use vancosamine-like donors (Figure S9). A structural superposition of GtfD, GtfA, and GtfAH1 shows that most of these residues are structurally conserved. The exception is D315 and Q316, which are disrupted by the $_{318}$NVVE$_{321}$ insertion in GtfA. However, the DQ motif is undisturbed in GtfD (Mulichak et al., 2004), so almost completely superimposes with the analogous GtfAH1 residues (data not shown). Elsewhere, there are a few regions that may facilitate the discrimination between glucose and vancosamine-like substrates. The six residues following Ser230 are highly variable between glycopeptide GT subgroups (those that use glucose-like donors and those that utilize vancosamine-like donors), including a two amino acid insertion in the vancosaminyltransferases. The three residues before Asp315 also appear to significantly differ between groups.

The existence of less obvious interactions is supported by directed evolution studies from Thorson and colleagues (Williams et al., 2007, 2008). These indicate that substrate specificity and selectivity, as well as catalytic activity, can be influenced by residues remote from the expected substrate binding sites. Site-directed mutagenesis and further structural and activity analyses will be required to fully dissect the structural basis for precise sugar selectivity.

DISCUSSION

This study demonstrates the viability of constructing chimeric GTs to generate novel natural product derivatives and the method should have wide applicability, given the ubiquity of GT-B fold GTs. The generality of GT domain swapping is supported by recent studies by the groups of Bowles (Brazier-Hicks et al., 2007; Cartwright et al., 2008), Kim (Park et al., 2009), and Bechthold (Krauth et al., 2009), who have reported interesting examples of functional chimeric GTs (plant flavanoid, bacterial aminoglycoside-glycopeptide, and bacterial angucycline GTs, respectively). All show that the N-terminal domain is responsible for acceptor binding specificity.

Our study expands the method by creating highly functional chimeras with novel substrate specificity; those previous studies, along with a study on UDP-glucuronyltransferases (Mackenzie, 1990), did not attempt to modify sugar specificity and therefore the effect of the C-terminal domain on sugar selectivity could not be established. Chimeras of 91% identical urdamycin GTs possess novel activities (Hoffmeister et al., 2001), but this activity was difficult to predict as the C domains of these GTs are almost identical despite using different sugars. This is very unusual for GT-B fold GTs and may mark these as atypical GTs.

Our functional and structural analysis of chimera selectivity clearly identifies acceptor and donor binding sites for the glycopeptide GTs, although sugar binding appears to be appreciably affected by the N-terminal domain. Current evidence on GT-B trans-domain binding of the sugar residue is conflicting. The chimeric urdamycin GTs revealed N-terminal domain residues that control sugar binding (Hoffmeister et al., 2001), and the sugar specificity of OleD, a macrolide GT, was changed by mutating residues from the analogous N-domain region in a site-directed evolution study (Williams et al., 2008). However, in the co-complexes of sugar donors with MurG (Hu et al., 2003), UGT71G1 (Shao et al., 2005), *Vv*GT1 (Offen et al., 2006), and GtfAH1 almost

all sugar-binding residues are in the C domain. This variation in the exact mode of sugar binding within the GT-B family should be expected and may account for the differences observed in GtfBH1 and GtfBH2 specificities, along with the relaxed GtfAH1 UDP-GlcNAc/Glc selectivity.

Nevertheless, these results are clearly encouraging and provide an indication of the potential that chimeric GTs possess: rational protein engineering to create GTs with novel and predictable activities. It will be intriguing to establish the scope of this method with GTs involved in the biosynthesis of other classes of clinically efficacious natural products. As with domain swapping in proteins such as polyketide synthases, it is essential that appropriate fusion sites are chosen to maximize the likelihood of a functional chimera (Kellenberger et al., 2008). The ever expanding GT-B superfamily structural data set will aid in this process along with the screening method of Park et al. (2009) for GT activity. Our structural and sequence analysis implies that trans-domain interactions will have to be considered when constructing chimeric GTs.

Hybrid GTs should prove particularly useful at derivatizing natural products whose native GTs possess strict substrate specificity, perhaps in combination with the C-terminal domain of GTs, which have proven sugar donor promiscuity. Additionally, this strategy should be amenable to in vivo engineering of glycosylation pathways; it will be useful to establish whether the activities and specificities observed in vitro are mirrored in vivo. Furthermore, the GtfAH1:UDP-2F-Glc crystal structure should allow for a more surgical approach to modifying the sugar specificity of glycopeptide GTs, as well as for GTs that are predicted to possess analogous sugar-binding regions.

SIGNIFICANCE

Our work demonstrates that the sugar donor specificity of GTs can be rationally controlled by fully swapping substrate binding domains. This method has been applied to generate vancomycin analogs that would be difficult to obtain using existing glycodiversification techniques. In conjunction with work performed by the groups of Bechthold (Krauth et al., 2009), Bowles (Cartwright et al., 2008), and Kim (Park et al., 2009), a picture is emerging that domain swapping is a viable strategy for generating highly active unnatural GTs. The kinetic characterization of GtfAH1 and the activity analysis of GtfBH1 and GtfBH2 show that domain swapping does not severely impair catalytic activity and can actually improve activity over wild-type GTs. In addition, the crystal structure of GtfAH1 with and without a bound UDP-2F-Glc provides detailed information regarding the generation of a successful chimera and is an unprecedented insight into the way in which glycopeptide GTs bind sugars. Since the vast majority of GTs involved in natural product biosynthesis belong to the same GT-B structural superfamily, GT domain swapping may be a general strategy for directing the biosynthesis of clinically relevant natural products.

EXPERIMENTAL PROCEDURES

Chemicals

Teicoplanin was purchased from Advanced Separation Technologies Ltd., DNA primers were purchased from MWG Biotech, acetonitrile (HPLC grade) was purchased from Fisher Scientific, and all other chemicals were purchased from Sigma-Aldrich and were of analytical grade. AGT was synthesized by the acid-catalyzed degradation of teicoplanin using published procedures (Boger et al., 2000). AGV and DVV were produced by the acid-catalyzed degradation of vancomycin (Nagarajan et al., 1989). Vancomycin *N*-acetylglucosaminyl pseudoaglycone (**2**) was enzymatically generated using GtfB as previously described (Truman et al., 2008).

Plasmid Construction

The generation of pET28a(+)-based plasmids encoding GtfB and Orf10* with N-terminal hexahistidine tags have been described elsewhere (Mulichak et al., 2001; Li et al., 2004). Plasmids encoding GtfA and Or1 with N- and C-terminal hexahistidine tags were constructed as follows. Briefly, the genes were amplified from cosmid DNA (van Wageningen et al., 1998; Li et al., 2004) using the following primers (restriction sites introduced are shown in **bold**): gtfAN-for, 5′- GGATGC**CATATG**CGCGTGTTGATTACG-3′; gtfAN-rev, 5′- GTTCCATAGG**CTCGAG**GTGGTTCAGGC-3′; gtfAC-for, 5′-GGATGCG**CCATGG**GCGTGTTGATTACGGG-3′; gtfAC_rev, 5′- CATAGGCTCGGTGT**CTCGAG**GGCGGGAAC-3′; orf1N-for, 5′- GGATGTG**CATATG**CGCGTGCTGTTTTC GTC-3′; orf1N-rev, 5′- GGCG**GAATTC**ACGCGGGAACCGACGATC-3′; orf1C-for, 5′- GGATGTGA**CCATGG**GCGTGCTGTTTTCGTC-3′; orf1C-rev, 5′- GGCG C**AAGCTT**CGCGGGAACCGACGATCTC-3′.

PCR reactions were carried out using KOD Hot Start DNA polymerase (Novagen) according to the manufacturer's instructions. After digestion (New England Biolabs), the products were ligated (T4 Ligase; Stratagene) into the corresponding sites in a pET28(a)+ expression vector (Novagen). *gtfB* and *orf10** plasmids were doubly digested using MluI and PstI to generate linear fragments of 1508 bp/5044 bp (Orf10*) and 1505 bp/5046 bp (GtfB). The heterologous re-ligation of the small fragments with the large fragments resulted in the generation of pET28(a)+ constructs harboring the hybrid GT genes (1508 bp + 5046 bp = N-terminal Orf10* and C-terminal GtfB; 1505 bp + 5044 bp = N-terminal GtfB and C-terminal Orf10*). GtfA-Orf1 hybrid constructs were generated by digesting the Orf1 and GtfA plasmids using MluI and BamHI. The heterologous re-ligation of the small and large fragments generated pET28-based plasmids encoding hybrids with N- or C-terminal hexahistidine tags. The Orf1 constructs were only partially digested with BamHI owing to a second BamHI restriction site at nucleotide 1067 in the gene. The identity of all plasmids was confirmed by sequencing.

General Enzyme Expression Procedure

E. coli BL21(DE3) competent cells (Novagen) containing chimeric GT plasmid were grown in LB medium (1 liter) containing kanamycin (50 μg/mL) with shaking at 37°C until the OD_{600} reached approximately 0.6. Isopropyl-β-D-thiogalactopyranoside was then added to a final concentration of 0.2 mM and cell growth was continued with shaking at 16°C for 16 hr. The cells were then harvested by centrifugation and the resulting cell pellet was suspended in 30 ml binding buffer (10 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, and 10% glycerol [pH 7.9]) and lysed by sonication (Vibra-Cell Sonicator; Sonics and Materials Inc.). The resulting cell lysate was clarified by centrifugation and hexahistidine-tagged protein was purified on a His•Bind resin column (Novagen) at 4°C following the manufacturer's instructions.

After elution, the protein solution was desalted using an Amicon Ultra centrifugal filter (Millipore) and the buffer was exchanged to 50 mM Tris-HCl and 100 mM KCl (pH 8.4) for GtfBH1 and GtfBH2 and to 50 mM Tris-HCl and 100 mM KCl (pH 7.5) for GtfAH1. Where necessary, the enzyme was further purified by gel filtration on an ÄKTA Explorer FPLC system with a HiLoad 16/60 Superdex 200 Prep Grade column. The isocratic mobile phase contained 100 mM KCl and 20 mM Tris-HCl (pH 8.4).

Protein identity was confirmed by LC-ESI-MS analysis: GtfBH1 observed mass = 44496 Da (predicted mass after loss of *N*-formyl-methionine = 44498 Da); GtfBH2 observed mass = 44581 Da (predicted mass after loss of *N*-formyl-methionine = 44584 Da); GtfAH1-N observed mass = 42600 Da (predicted mass after loss of *N*-formyl-methionine = 42601 Da); GtfAH1-C observed mass = 41858 Da (predicted mass after loss of *N*-formyl-methionine = 41859 Da).

Glycopeptide LC-ESI-MS Data Acquisition

Spectra were obtained using a Hewlett-Packard HPLC 1100 series instrument coupled to a Finnigan MAT LCQ ion trap mass spectrometer fitted with

a positive mode ESI source. Samples were injected onto a Phenomenex Luna C18(2) column (250 mm × 2.0 mm, 5 μm), eluting with a linear gradient of 0 to 60% acetonitrile containing 0.1% trifluroacetic acid (TFA) in water (0.1% TFA) over 30 min with a flow-rate of 0.3 mL/min.

Chimeric GT Assays

A typical glycosylation reaction involved incubating the GT (5 μM) with glycopeptide acceptor (50 μM–1 mM) and NDP-sugar (0.5–2 mM) in buffer (75 mM Tris-HCl [pH 7.4] or 75 mM Tricine [pH 9]), 8 mM $MgCl_2$, 2.5 mM TCEP, 1 mg/mL BSA, and 10% (v/v) DMSO at 20°C–37°C. Assays were quenched with an equivalent amount acetonitrile containing 0.1% TFA and directly subjected to LC-ESI-MS analysis. Kinetic parameters were generated by curve-fitting using Origin 7 (OriginLab). See Supplemental Methods for additional experimental details on kinetic analyses and competitive assays.

Crystallization and Structural Determination

The protein was crystallized using vapor diffusion by the hanging drop technique using a protein concentration of 10 mg/mL. Equal parts of crystallization solution and protein solution were mixed forming drops of 2 μL. The crystals were obtained in 0.1 M Tris-HCl (pH 8.5) and 1.2–2.0 M ammonium sulfate. To facilitate data collection at 100 K, these crystals were cryoprotected using the well solution in the presence of 30% v/v glycerol. Co-crystallization with UDP-GlcNAc and UDP-2F-Glc was performed using concentrations of five to ten higher than the molar concentration of the protein. X-ray data for GtfAH1 crystallized in complex with UDP were collected to 1.15 Å resolution at the European Synchrotron Radiation Facility. The structure was solved by molecular replacement using AMoRe (CCP4, 1994) with GtfA from *A. orientalis* as a search probe (Mulichak et al., 2003). Refinement led to a final R_{factor} and R_{free} of 17.1% and 18.5%, respectively. X-ray data for GtfAH1 crystallized in complex with UDP-2F-Glc were collected to 1.30 Å resolution. See Table 3 for a summary of the statistics of data collection and refinement.

ACCESSION NUMBERS

Atomic coordinates for GtfAH1:UDP and GtfAH1:UDP-2F-Glc have been deposited in the Protein Data Bank (www.rcsb.org) under the accession numbers 3H4T and 3H4I, respectively.

SUPPLEMENTAL DATA

Supplemental Data contain nine figures and Supplemental Methods and can be found with this article online at http://www.cell.com/chemistry-biology/supplemental/S1074-5521(09)00178-1.

ACKNOWLEDGMENTS

This work was supported by the Biotechnology and Biological Sciences Research Council (A.W.T., F.H., and J.B.S.), the Cambridge Overseas Trust (S.W.), and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (M.V.B.D.). We thank Peter Leadlay (University of Cambridge) for helpful discussions and advice, and Stephen Withers (University of British Columbia) for the kind donation of UDP-2F-Glc.

Received: March 8, 2009
Revised: April 23, 2009
Accepted: April 28, 2009
Published: June 25, 2009

REFERENCES

Anzai, Y., Saito, N., Tanaka, M., Kinoshita, K., Koyama, Y., and Kato, F. (2003). Organization of the biosynthetic gene cluster for the polyketide macrolide mycinamicin in *Micromonospora griseorubida*. FEMS Microbiol. Lett. *218*, 135–141.

Beauregard, D.A., Williams, D.H., Gwynn, M.N., and Knowles, D.J.C. (1995). Dimerization and membraneanchors in extracellular targeting of vancomycin group antibiotics. Antimicrob. Agents Chemother. *39*, 781–785.

Bennett-Lovsey, R.M., Herbert, A.D., Sternberg, M.J., and Kelley, L.A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. Proteins *70*, 611–625.

Boger, D.L., Weng, J.-H., Miyazaki, S., McAtee, J.J., Castle, S.L., Kim, S.H., Mori, Y., Rogel, O., Strittmatter, H., and Jin, Q. (2000). Thermal atropisomerism of teicoplanin aglycon derivatives: preparation of the P,P,P and M,P,P atropisomers of the teicoplanin aglycon via selective equilibration of the DE ring system. J. Am. Chem. Soc. *122*, 10047–10055.

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.

Borisova, S.A., Zhao, L., Sherman, D.H., and Liu, H.-w. (1999). Biosynthesis of desosamine: construction of a new macrolide carrying a genetically designed sugar moiety. Org. Lett. *1*, 133–136.

Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.K., Bowles, D.J., Davies, G.J., and Edwards, R. (2007). Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. Proc. Natl. Acad. Sci. USA *104*, 20238–20243.

Cartwright, A.M., Lim, E.K., Kleanthous, C., and Bowles, D.J. (2008). A kinetic analysis of regiospecific glucosylation by two glycosyltransferases of *Arabidopsis thaliana*: domain swapping to introduce new activities. J. Biol. Chem. *283*, 15724–15731.

Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Cornish-Bowden, A. (2004). Fundamentals of Enzyme Kinetics, Third Edition (London: Portland Press).

Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. J. Mol. Biol. *328*, 307–317.

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.

Gordon, R.D., Sivarajah, P., Satkunarajah, M., Ma, D., Tarling, C.A., Vizitiu, D., Withers, S.G., and Rini, J.M. (2006). X-ray crystal structures of rabbit *N*-acetylglucosaminyltransferase I (GnT I) in complex with donor substrate analogues. J. Mol. Biol. *360*, 67–79.

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.

Howard-Jones, A.R., Kruger, R.G., Lu, W., Tao, J., Leimkuhler, C., Kahne, D., and Walsh, C.T. (2007). Kinetic analysis of teicoplanin glycosyltransferases and acyltransferase reveal ordered tailoring of aglycone scaffold to reconstitute mature teicoplanin. J. Am. Chem. Soc. *129*, 10082–10083.

Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003). Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. Proc. Natl. Acad. Sci. USA *100*, 845–849.

Kellenberger, L., Galloway, I.S., Sauter, G., Bohm, G., Hanefeld, U., Cortes, J., Staunton, J., and Leadlay, P.F. (2008). A polylinker approach to reductive loop swaps in modular polyketide synthases. ChemBioChem *9*, 2740–2749.

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

Li, T.-L., Huang, F., Haydock, S.F., Mironenko, T., Leadlay, P.F., and Spencer, J.B. (2004). Biosynthetic gene cluster of the glycopeptide antibiotic

teicoplanin. Characterization of two glycosyltransferases and the key acyltransferase. Chem. Biol. *11*, 107–119.

Lu, W., Oberthür, M., Leimkuhler, C., Tao, J., Kahne, D., and Walsh, C.T. (2004). Characterization of a regiospecific epivancosaminyl transferase GtfA and enzymatic reconstitution of the antibiotic chloroeremomycin. Proc. Natl. Acad. Sci. USA *101*, 4390–4395.

Mackay, J.P., Gerhard, U., Beauregard, D.A., Maplestone, R.A., and Williams, D.H. (1994). Dissection of the contributions toward dimerization of glycopeptide antibiotics. J. Am. Chem. Soc. *116*, 4573–4580.

Mackenzie, P.I. (1990). Expression of chimeric cDNAs in cell culture defines a region of UDP glucuronosyltransferase involved in substrate selection. J. Biol. Chem. *265*, 3432–3435.

Minami, A., Kakinuma, K., and Eguchi, T. (2005). Aglycon switch approach toward unnatural glycosides from natural glycoside with glycosyltransferase VinC. Tetrahedron Lett. *46*, 6187–6190.

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.

Nagarajan, R., Berry, D.M., and Schabel, A.A. (1989). The structural relationships of A82846B and its hydrolysis products with chloroorienticins A, B and C. J. Antibiot. (Tokyo) *42*, 1438–1440.

Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B.G., Tarling, C.A., Ford, C.M., Bowles, D.J., and Davies, G.J. (2006). Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. EMBO J. *25*, 1396–1405.

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

Perez, M., Lombo, F., Baig, I., Brana, A.F., Rohr, J., Salas, J.A., and Mendez, C. (2006). Combinatorial biosynthesis of antitumor deoxysugar pathways in *Streptomyces griseus*: reconstitution of unnatural natural gene clusters for the biosynthesis of four 2,6-D-dideoxyhexoses. Appl. Environ. Microbiol. *72*, 6644–6652.

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

Schell, U., Haydock, S.F., Kaja, A.L., Carletti, I., Lill, R.E., Read, E., Sheehan, L.S., Low, L., Fernandez, M.J., Grolle, F., et al. (2008). Engineered biosynthesis of hybrid macrolide polyketides containing D-angolosamine and D-mycaminose moieties. Org. Biomol. Chem. *6*, 3315–3327.

Shao, H., He, X., Achnine, L., Blount, J.W., Dixon, R.A., and Wang, X. (2005). Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. Plant Cell *17*, 3141–3154.

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

Truman, A.W., Fan, Q., Rottgen, M., Stegmann, E., Leadlay, P.F., and Spencer, J.B. (2008). The role of Cep15 in the biosynthesis of chloroeremomycin: reactivation of an ancestral catalytic function. Chem. Biol. *15*, 476–484.

van Wageningen, A.M., Kirkpatrick, P.N., Williams, D.H., Harris, B.R., Kershaw, J.K., Lennard, N.J., Jones, M., Jones, S.J., and Solenberg, P.J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chem. Biol. *5*, 155–162.

Varki, A., Cummings, R., Esko, J., Freeze, H., and Marth, J. (1999). Essentials of Glycobiology (New York: Cold Spring Harbor Laboratory Press).

Vrielink, A., Rüger, W., Driessen, H.P., and Freemont, P.S. (1994). Crystal structure of the DNA modifying enzyme β-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. EMBO J. *13*, 3413–3422.

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.

Yang, M., Brazier, M., Edwards, R., and Davis, B.G. (2005). High-throughput mass-spectrometry monitoring for multisubstrate enzymes: determining the kinetic parameters and catalytic activities of glycosyltransferases. ChemBioChem *6*, 346–357.

Zhang, C., Griffith, B.R., Fu, Q., Albermann, C., Fu, X., Lee, I.-K., Li, L., and Thorson, J.S. (2006). Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. Science *313*, 1291–1294.