Chemistry & Biology, Vol. 12, 657–664, June, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.chembiol.2005.04.009

Structure-Based Engineering of *E. coli* Galactokinase as a First Step toward In Vivo Glycorandomization

Jie Yang, Xun Fu, Jianchun Liao, Lesley Liu, and Jon S. Thorson*
Laboratory for Biosynthetic Chemistry
Pharmaceutical Sciences Division
School of Pharmacy
University of Wisconsin-Madison
777 Highland Avenue
Madison, Wisconsin 53705

Summary

In vitro glycorandomization is a rapid chemoenzymatic strategy to diversify complex natural product scaffolds. The glycorandomization sugar activation pathway is dependent upon the efficient construction of diverse sugar-1-phosphate libraries. In the context of the previously evolved GalK Y371H "gatekeeper" mutation, the active site M173L mutation described herein presents a kinase with remarkably broadened substrate range to include 28 diverse natural and unnatural sugars. Among these new substrates, 6-azido-6-deoxy-galactose and 6-azido-6-deoxy-glucose present unique chemical probes to assess the utility of an *E. coli* Y371H/M173L-GalK-overproducing strain to generate unnatural sugar-1-phosphates in vivo. Remarkably, the in vivo conversion of both unnatural sugars rival that demonstrated in vitro. This notable in vivo success stands as the first step toward constructing short sugar-activation pathways in vivo and, ultimately, in vivo natural-product glycorandomization.

Introduction

Glycosylated natural products remain a key source of pharmaceutical leads, and, given the essential roles of their sugar appendages, the development of efficient methods to rapidly diversify sugars attached to complex natural products is anticipated to enhance drug discovery efforts [1–5]. Glycorandomization (Figure 1A), a process centered upon the inherent promiscuity of secondary metabolite-associated glycosyltransferases, is one of the latest promising developments toward this important goal [6, 7]. Critical to the success of glycorandomization has been our ability to engineer and/or evolve two additional promiscuous enzymes—anomeric kinases [8–11] and nucleotidylyltransferases [12–16]. Taken together with the many elegant methods to synthesize monosaccharide libraries [17–19] and the intrinsic substrate flexibility of many secondary metabolite-associated glycosyltransferases, this two-enzyme short activation pathway allows one to rapidly diversify the sugars attached to complex natural products [6, 7, 20–22]. The glycorandomization process is further enhanced via a final diversification step that relies upon the use of downstream chemoselective ligation [6, 7]. As exemplary proof of the concept, recent in vitro glycopeptide glycorandomization experiments led to >50 vancomycin derivatives, some of which displayed enhanced and distinct biological properties from the parent natural product [6, 7, 23].

Cumulatively, the successful demonstration of in vitro glycorandomization, the illustration that functional NDP-sugar pathways can be reassembled in prokaryotes [24, 25], and the fact that natural and "unnatural" endogenous sugars are processed in vivo by both prokaryotes and eukaryotes [26–30] present the foundation from which to approach in vivo glycorandomization. For example, the expression of a tandem promiscuous sugar-1-kinase (GalK) and nucleotidylyltransferase (E_p)—essentially an unnatural NDP-sugar factory—in a natural aglycon-producing host (e.g., the erythromycin-producing *Saccharopolyspora*) should present the prospect of generating a glycorandomized library in situ, the glycorandomized metabolite output of which is controlled by monosaccharides being fed to the strain (Figure 1B). Alternatively, expression of the tandem two gene "NDP-sugar factory" genes in a nonproducing host (e.g., *S. lividans* or *E. coli*), which expresses a given glycosyltransferase (or glycosyltransferase library), should also provide a vehicle to accomplish glycorandomization via feeding the host with appropriate aglycon acceptors and unnatural sugar donors (Figure 1C). The key to either in vivo scenario is the ability of unnatural sugars to enter the host and serve as efficient substrates of the first enzyme of the short activation pathway (the flexible anomeric kinase). Toward demonstrating this first key step, a kinase able to process sugars bearing unique mass signatures and/or reactive handles would, in addition to further enhancing library diversification, greatly simplify the final analysis of in vivo sugar substrate access and turnover.

To date, we have applied directed enzyme evolution and relied upon a high throughput galactokinase (GalK) assay (DNS assay) for the screening of diverse *E. coli* GalK variants generated via error-prone PCR to elucidate anomeric kinases with broadened substrate specificity [9]. From this approach, one particular GalK mutant (Y371H) demonstrated remarkably widened substrate flexibility toward C-2, C-3, C-5, and C-6 substitutions of D-galactose. Yet, the mutant retained a stringent requirement for the axial C-4 galactose architecture. Based upon a homology model with the recently solved *L. lactis* GalK crystal structure, two highly conserved active-site residues (D37 and Y223 in *E. coli* GalK) are responsible for hydrogen bonding with this C-4 axial hydroxyl group of the substrate [31]. Saturation mutagenesis at these two critical positions in the *E. coli* enzyme failed to provide mutants with enhanced C-4 sugar flexibility [10], although a parallel study revealed the *L. lactis* wild-type GalK and Y385H orthologs (the equivalent to Y371H in the *E. coli* enzyme) to surprisingly display weak activity toward the C-4 epimer, glucose [11].

Herein, we report a structure-activity model based upon the *L. lactis* active site and its ability to weakly

*Correspondence: jsthorson@pharmacy.wisc.edu

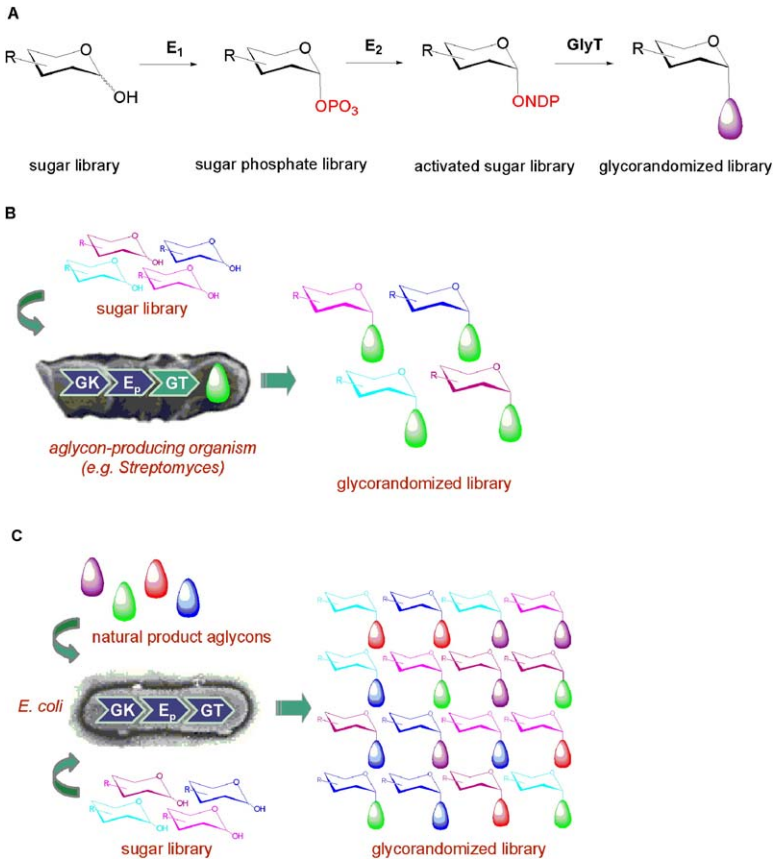

Figure 1. Glycorandomization Overview and Two Potential Scenarios for an In Vivo Approach

(A) In vitro glycorandomization utilizes two enhanced enzymes—E_1 (a general kinase) and E_2 (a general nucleotidylyltransferase)—to generate NDP-sugar substrate libraries to be utilized by a flexible natural product-associated glycosyltransferase (GlyT).

(B) In vivo glycorandomization scenario I: feeding monosaccharides to a natural product-producing host engineered to express the "NDP-sugar factory." In this scenario, both the aglycon and glycosyltransferase are provided by the bacterial host.

(C) In vivo glycorandomization scenario II: feeding monosaccharides and aglycons to a nonproducing host engineered to express the "NDP-sugar factory" and an appropriate glycosyltransferase or glycosyltransferase library.

utilize glucose to serve as a basis for the design of a specific engineered *E. coli* M173L mutant GalK with enhanced C-4 and C-6 promiscuity. Moreover, a combination of the favorable structure-based (M173L) mutation with the beneficial mutation previously discovered via directed evolution (Y371H) provides an enzyme with substrate promiscuity that drastically exceeds an additive enhancement for both C-4 and C-6 substitutions. Most importantly, the additional unnatural sugar substrates accessed by this new double mutant allow for a unique opportunity to assess whether unnatural sugars can enter a bacterial host and serve as efficient substrates of the first enzyme of the glycorandomization pathway (the flexible anomeric kinase). Specifically, feeding of the unnatural substrate 6-azido-6-deoxy-galactose (**2**) or 6-azido-6-deoxy-glucose (**26**) to an *E. coli* host engineered to express M173L-Y371H-GalK followed by the rapid fluorescent labeling of substrates and products via Huisgen 1,3-dipolar cycloaddition revealed the desired unnatural sugar-1-phosphate production to occur in vivo under normal fermentation conditions. This result stands as a critical first step in demonstrating the concept of in vivo glycorandomization.

Results and Discussion

Structural Basis for Engineering GalKs with Expanded Specificity

Prior to any available GalK structural information, the directed evolution of *E. coli* GalK presented a general sugar anomeric kinase with widened flexibility primarily at C-5 and C-6 of the sugar substrate [9]. Interestingly, all C-4-modified derivatives tested in the previous study failed as substrates for the evolved catalyst. In contrast, our latest analysis of the *L. lactis* wild-type GalK revealed moderate in vitro conversion of various C-4-modified analogs, including 4-azido-4-deoxy-D-galactose (**1**), 4-deoxy-D-galactose (**5**), and D-glucose (**7**) [11]. The recent structure elucidation of *L. lactis* GalK [31] potentially allows for a molecular-level assessment of this surprising C-4 specificity distinction between the *E. coli* and *L. lactis* enzymes. With the *L. lactis* structure as a template, an *E. coli*-*L. lactis* GalK homology model revealed one clear difference between the sugar binding pockets. Specifically, L182 in wild-type *L. lactis* GalK is near the C-4 carbon atom of galactose (3.85 Å), and, based upon sequence alignment, this residue is replaced by M173 in *E. coli* GalK. Modeling the C-4 epimer of galactose (D-glucose) within the *L. lactis* active site predicts the Glc-C-4 equatorial hydroxyl to be 3.79 Å from the γ methyl of L182 (Figure 2A). However, the identical model in which L182 has been replaced by methionine (to mimic the *E. coli* GalK active site) revealed the same Glc-C-4 equatorial hydroxyl to be 1.72 Å from M173 sulfur (Figure 2B) and suggests M173 in *E. coli* GalK may limit sugar C-4 specificity to *galacto*-(axial)-configured substrates. Moreover, given the close proximity of the sugar C-6 hydroxyl to M173 in this structural model (2.85 Å) (Figure 2B), we speculate M173 in *E. coli* GalK may also limit C-6 substitutions.

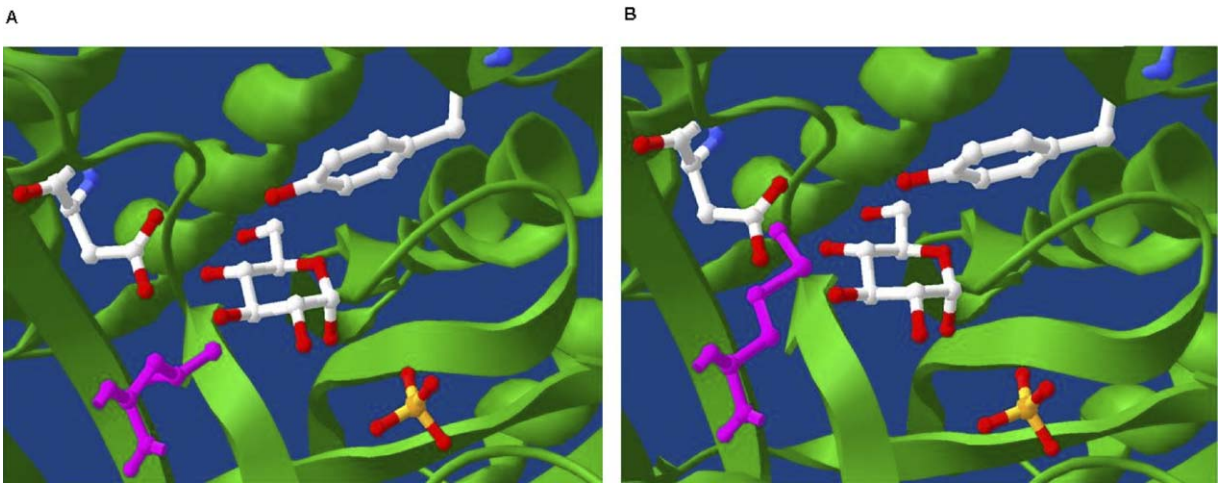

Figure 2. Modeling D-Glucose Bound to GalK Variants
(A) The active site of *L. lactis* GalK. (B) The corresponding homology-model of *E. coli* GalK.

Characterization of Engineered *E. coli* GalK Mutants

A single *E. coli* GalK M173L mutant was generated and screened against a panel 45 potential sugar substrates to test the above hypothesis. As predicted, the *E. coli* M173L-engineered mutant displayed moderate D-glucose (**7**) activity (20% conversion in 3 hr). Moreover, three additional D-*gluco*-configured structures, **8**, **17**, and **18** (Figures 3 and 4), which were not substrates of wild-type *E. coli* GalK (or the evolved *E. coli* Y371H mutant), were also substrates of the new structure-based variant. Although these studies clearly revealed the structure-based M173L mutant to accept a substrate set distinct to that of the previously evolved Y371H mutant, in contrast to the structural model described above, both mutants failed in the presence of substrates presenting even moderate C-6 bulk such as 6-azido-6-deoxy-galactose **2** or 6-azido-6-deoxy-D-glucose **26**. The *E. coli* double M173L-Y371H mutant was examined in an attempt to further generalize the sugar kinase activity. Remarkably, not only did this double mutant retain the activity of both corresponding single mutants, but this prodigy demonstrated a substantial degree of kinase activity toward a variety of new structures (**1–6**, **23–28**). Most of the new substrates share modifications at C-4 and/or C-6 with many of D-*gluco* origin. It is also noteworthy that three among this new substrate set are azidosugars (**1**, **2**, and **26**), thereby setting the stage for rapid analysis of in vivo bioconversion via postbioconversion labeling of substrates and products with a fluorescent tag by using Huisgen 1,3-cycloaddition [6]. Figures 3 and 4 illustrate the complete substrate profiles for wild-type *E. coli* GalK, the *E. coli* GalK mutant M173L, Y371H, and M173L-Y371H.

To better understand the distinct role of the two particular amino acid residues (M173 and Y371) in determining the substrate specificity, we chose the native substrate D-galactose **9**, the unique M173L substrate D-glucose **7**, and the unique Y371H substrate L-altrose **21** for complete comparative steady-state kinetic profiling. In comparison to wild-type *E. coli* GalK, a slight (~2-fold) D-galactose K_m increase was observed in all three variants (M173L, Y371H, and M173L-Y371H) (Table 1). Moreover, a comparison of the D-glucose kinetic parameters for M173L and the L-altrose kinetic values for Y371H to those of the double mutant with D-glucose or L-altrose revealed very little change. Although the determined V_{max} values are consistent with observed 3-hr-conversion yields presented in Figure 4, in contrast to the notable gain of function (in terms of the overall number of new M173L-Y371H substrates) illustrated in Figures 3 and 4, this cursory kinetic analysis predicts the gain of function to be additive at best. In other words, the kinetic analysis suggests the double mutant should likely accept known M173L and Y371H substrates but would not forecast an expansion beyond this dual substrate set. Yet, although it is difficult to explain this remarkable gain of function in the M173L-Y371H variant, the unique ability of this double mutant to accept compounds **1**, **2**, and **26** sets the stage to assess the first step of in vivo glycorandomization as described below.

In Vivo Bioconversion of Unnatural Sugars with an Engineered GalK

In the context of assessing in vivo bioconversion, the specific M173L-Y371H-**1**, **2**, or **26** relationships are advantageous for two reasons. First, as described above, **1**, **2**, and **26** are not substrates for wild-type *E. coli* GalK, and, therefore, the use of a standard *E. coli* host strain (which contains the inherent wild-type *E. coli* GalK) should not interfere with the experimental design. Second, as previously mentioned, **1**, **2**, and **26** each offer a unique functional handle to provide for the rapid installation of a fluorescent label to simplify the chromatographic analysis. In this context, **2** and **26** are equally reactive (>80%) (X.F. and J.S.T., unpublished data) to the required fluorescent labeling via Husigen 1,3-cycloaddition, whereas **1** is poorly reactive (<10%) (X.F. and J.S.T., unpublished data). Thus, for the current in vivo analysis, **1** was excluded. The set selected (**2** and **26**) still offer the opportunity to test a range of sub-

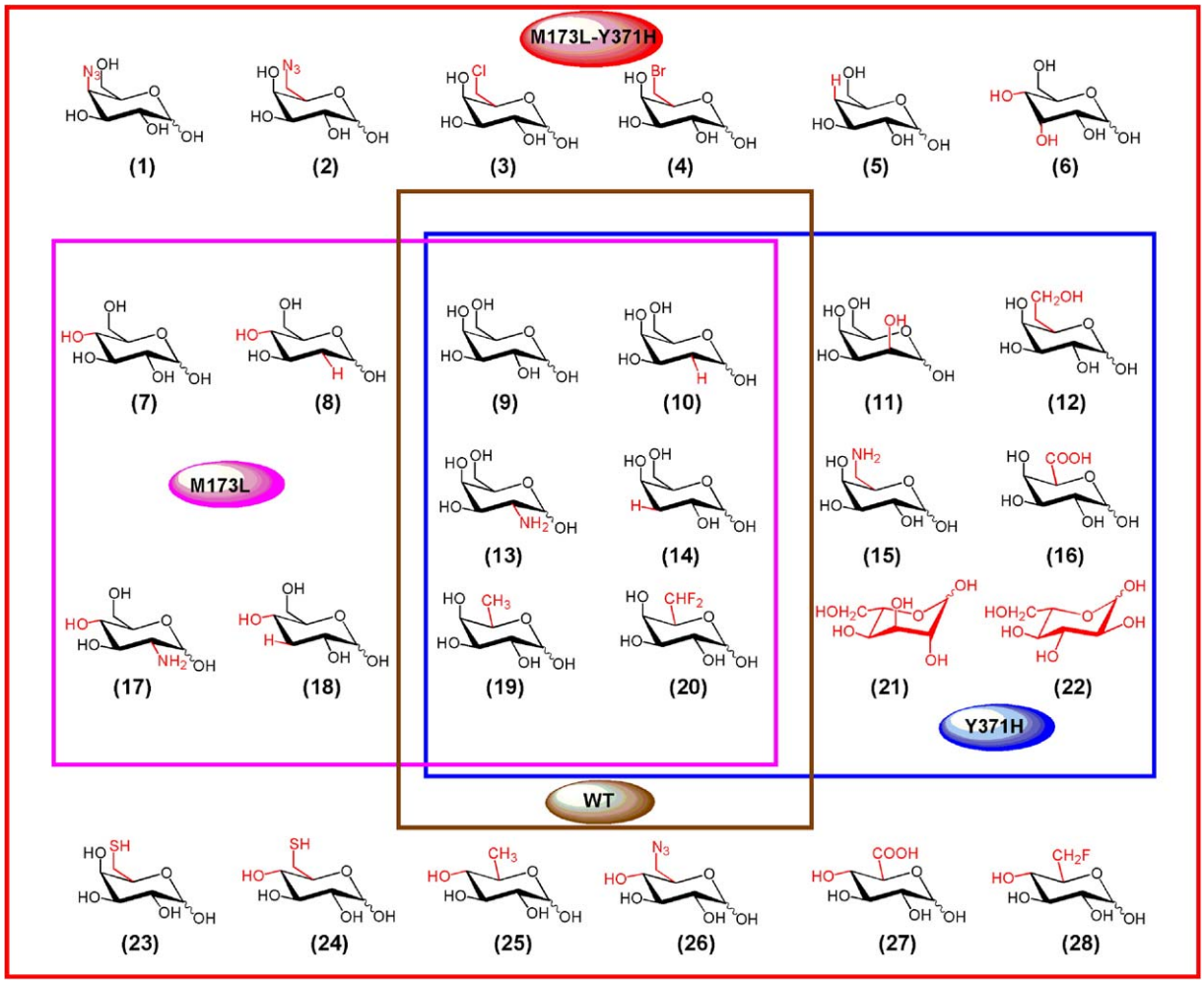

Figure 3. "Natural" and "Unnatural" Substrates of Wild-Type GalK and GalK Mutants

The positions deviating from the natural GalK substrate D-galactose (**9**) are highlighted in red.

strates with distinctly unique in vitro profiles. Specifically, **2** is known to have 44% ± 2% conversion in vitro in 3 hr, whereas **26** shows 14% ± 3% conversion under the same conditions.

The unnatural sugars (**2** or **26**, 4 mM final concentration) were fed to an *E. coli* host (40 ml culture), which overexpressed the promiscuous GalK, to assess the Y371H-M173L GalK-catalyzed in vivo production of unnatural sugar-1-phosphates (Figure 5A). After a designated time, the extracts were analyzed via the specific attachment of a fluorescent tag (**30**) with 1,3-dipolar cycloaddition [6]. Two controls were processed in parallel. The first utilized a strain containing an empty expression vector (pET-15b—the vector used for overexpression of the GalK mutants), whereas the second employed a wild-type GalK overexpression strain. The crude products from each bioconversion were isolated, labeled via 1,3-dipolar cycloaddition, and analyzed by fluorescence HPLC and LC-MS. As illustrated in Figure 5B, 69% ± 5% conversion of intracellular 6-azido-6-deoxy-D-galactose (**2**) was observed, a slight improvement over the in vitro yield (44% ± 2% conversion). In a similar manner, 15% ± 4% conversion of intracellular 6-azido-6-deoxy-D-glucose (**26**) was observed, consistent with the in vitro yield (14% ± 3% conversion). Although the transport of "unnatural" sugars into this typical cloning *E. coli* host is less than optimal (e.g., typically ranging from 5%–25%) (J.Y. and J.S.T., unpublished data), the above results are notable for a variety of reasons. First, the consistency between the in vivo and in vitro yields suggests other potential sugars competing for the mutant GalK do not greatly influence the production of the desired product. Should competition become a problem with future "unnatural" sugars, this system is easily amenable to minimal media growth. Second, the reflected yields also suggest the "unnatural" sugars are not being metabolized via alternative "dead-end" metabolic pathways, some of which could be genetically excised if problematic to future development. Most important, the above result illustrates that unnatural sugars able to enter the heterologous *E. coli* host are clearly accessible to the engineered, heterologously-expressed, promiscuous sugar kinase. In the context of strains capable of enhanced uptake of a

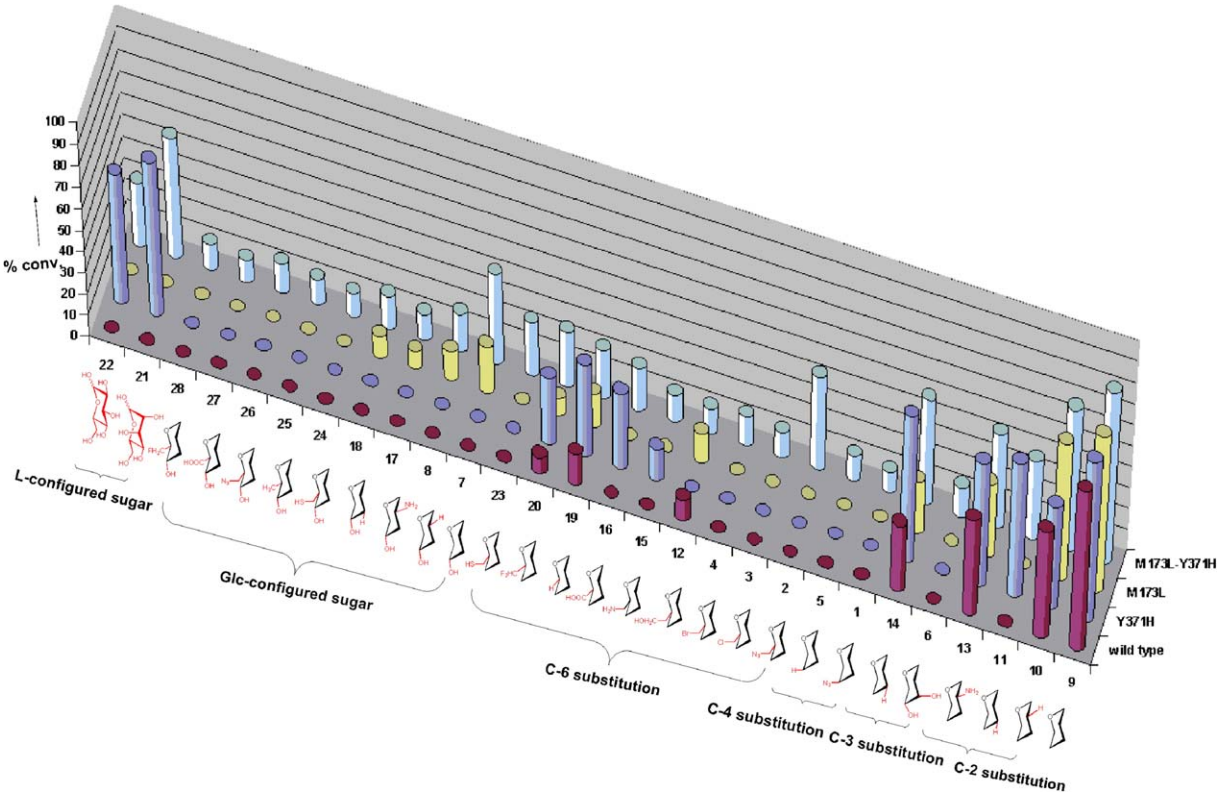

Figure 4. Percent Conversion of Sugar Substrates by Wild-Type and Mutant GalKs

Standard reaction conditions for all enzymatic reactions: [sugar], 8 mM; [ATP], 14 mM; [$MgCl_2$], 3.5 mM; [enzyme], 15.0 μM; reaction time, 180 min. The overall average error is ± 3%.

broad range of carbohydrates [32–34], this prototype system is particularly promising.

Significance

The work described herein impacts upon two distinct areas—enzyme engineering and diversification of complext natural product libraries with therapeutic potential. With respect to enzyme engineering, we report the application of structure-based engineering to provide the most promiscuous anomeric sugar kinase to date—capable of accepting nearly 30 diverse monosaccharide substrates—that will provide a foundation for the further rational engineering of this vital class of enzyme. In the context of natural product library generation, such flexible sugar kinases are a vital component of in vitro glycorandomization—a rapid chemoenzymatic method to alter the glycosylation patterns of complex natural products. However, the general application of in vitro glycorandomization approach is limited by the two primary issues. First, expensive substrates/cofactors/enzymes hamper the overall scale of the process. Second, the in vitro application is dependent upon establishing in vitro conditions for an active multienzyme process that can be severely limited by a variety of factors. A complimentary in vivo process is expected to overcome some, or all, of these limitations. Our current result illustrates the facile entry of unnatural sugars and their subsequent utilization by an engineered GalK-*E. coli* host. We anticipate the current strain to provide the ability to generate novel sugar-1-phosphate libraries in vivo, whereas overexpression of the next enzyme in our short activation pathway (an engineered sugar-

Table 1. Kinetic Data of the GalK Variants with D-Gal, D-Glc, and L-Alt

	D-Gal (**9**)		D-Glc (**7**)		L-Alt (**21**)	
	K_m mM	V_{max} mM · min^{-1}	K_m mM	V_{max} mM · min^{-1}	K_m mM	V_{max} mM · min^{-1}
WT GalK	2.09 (± 0.48)	1.54 (± 0.54)	—	—	—	—
M173L	5.87 (± 1.18)	1.95 (± 0.59)	2.57 (± 0.54)	0.02 (± 0.01)	—	—
Y371H	5.62 (± 0.86)	2.18 (± 0.62)	—	—	6.04 (± 1.26)	0.16 (± 0.04)
M173L-Y371H	4.59 (± 1.08)	2.26 (± 0.38)	4.02 (± 0.97)	0.02 (± 0.01)	6.28 (± 1.19)	0.07 (± 0.03)

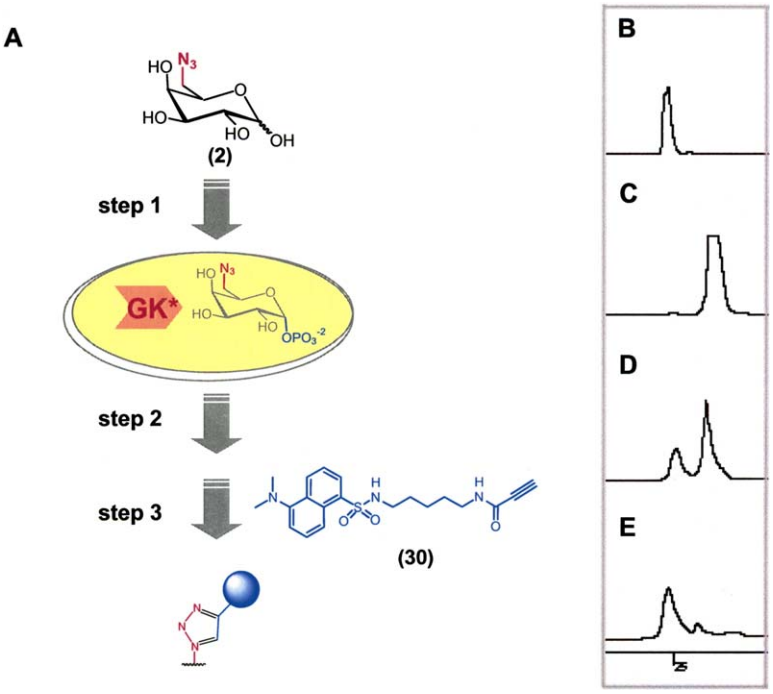

Figure 5. In Vivo GalK-Catalyzed Sugar-1-Phosphate Production

(A) Experimental design overview with 6-azido-6-deoxy-D-galactose (**2**) as an example: step 1, **2** is fed to an *E. coli* host expressing the M173L-Y371H GalK double mutant; step 2, after fermentation, the cell pellet is washed, resuspended, and boiled to delineate in vivo conversion from any potential in vitro conversion during work up; step 3, the reactants and products are subsequently labeled by using 1,3-dicycloaddition with the fluorescent tag **30** and analyzed via HPLC (Phenomenex Luna 5 μ 250 × 4.6 mm C18 reverse phase; 30 mM potassium phosphate [pH 6.0], 5 mM tetrabutylammonium-hydrogensulfate, 2% acetonitrile; 1.0 ml · min^{-1}; λ_{ex} = 335 nm, λ_{em} = 518 nm).

(B–E) HPLC traces. (B) Labeled-**2** standard, calculated for $C_{26}H_{36}N_6O_8S$ 592.2, and found m/z 591.2 $[M-H]^-$. (C) Labeled-**2**-1-phosphate standard, calculated for $C_{26}H_{37}N_6O_{11}PS$ 672.2, and found m/z 671.2 $[M-H]^-$. (D) **2** with Y371H-M173L GalK-*E. coli*. (E) **2** with wild-type GalK-*E. coli*.

1-phosphate nucleotidylyltransferase) will provide correspondingly diverse NDP-sugar libraries in a similar manner. Although these unique sugar-1-phosphate and NDP-sugar libraries will clearly serve as versatile chemical tools for glycobiology, this result more notably stands as evidence supporting the overall feasibility of in vivo glycorandomization and provides the fundamental elements for the eventual in vivo glycorandomization of a variety of clinically important natural products.

Experimental Procedures

Materials

The syntheses of 4-azido-4-deoxy-D-galactose (**1**), 6-ido-6-deoxy-D-galactose (**2**), 6-chloro-6-deoxy-D-galactose (**3**), 6-bromo-6-deoxy-D-galactose (**4**), 4-deoxy-D-galactose (**5**), 6-hydroxymethylene-D-galactose (**12**), 3-deoxy-D-galactose (**14**), 6-amino-6-deoxy-D-galactose (**15**), 6-deoxy-6,6-difluoro-D-galactose (**20**), and 5-dimethylaminonaphthalene-1-(N-[5-propargylamidepentyl])-sulfonamide (**30**) were reported previously [6, 8, 11]. The syntheses of 6-thio-6-deoxy-D-galactose (**23**), 6-thio-6-deoxy-D-glucose (**24**), and 6-azido-6-deoxy-D-glucose (**26**) will be reported elsewhere (J.L., X.F., and J.S.T., unpublished data), whereas other monosaccharide compounds **6–11**, **13**, **16–19**, **21**, **22**, **25**, **27**, and **28** were purchased from Sigma (St. Louis, Missouri), Fisher/Acros Organics (Hanover Park, Illinois), or Fluka (Milwaukee, Wisconsin). *E. coli* strains XL1-Blue and BL21 (DE3) were purchased from Stratagene (La Jolla, California). Expression vector pET15b was purchased from Novagen (Madison, Wisconsin). Enzymes were purchased from Promega (Madison, Wisconsin).

Structure Modeling

The PDB file for the crystal structure of *L. lactis* wild-type GalK was obtained from the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, New Jersey (http://www.rcsb.org/) [31]. The structure modeling was accomplished with Swiss-PdbViewer software (version 3.7). L182 was replaced by methionine in *L. lactis* wild-type GalK-galactose structure, and the best rotamors for M182 substitution were selected by the scoring function of the software. The galactose C-4 axial hydroxyl within the resulting minimized mutant structure was inverted to present the plausible representation of the *E. coli* wild-type GalK active site bound to D-glucose.

Site-Specific Mutagenesis

The GalK M173L single mutant and M173L-Y371H double mutant were generated with the QuikChange II Site-Directed Mutagenesis Kit from either wild-type or Y371H template, respectively (Stratagene). The corresponding mutated plasmids pGalKM173L and pGalKMLYH were constructed by using PfuUltra DNA polymerase for mutagenic primer-directed replication from pGalK [8] or pGalKY371H [9] template, respectively, with a pair of mutagenic primers (5′-GTAACTGCGGGATCCTGGATCAGCTAATTTCCG-3′ and 5′-CGGAAATTAGCTGATCCAGGATCCCGCAGTTAC-3′). Amplification was accomplished under the following conditions: 5 μl of 10× reaction buffer, 40 ng template DNA, 120 ng of each oligonucleotide primer, 1 μl dNTPs mixture (2.5 mM), and 2.5 U of PfuUltra high-fidelity DNA polymerase in a total volume of 50 μl ddH_2O (thermocycler parameters: initial denaturation, 2 min at 95°C; amplification, 12 cycles, 0.5 min at 95°C, 1 min at 55°C, 6.5 min at 68°C; terminal hold, 5 min at 68°C). The amplified plasmids were treated with *Dpn*I to digest the parental DNA template and the mutated prodigy plasmid subsequently used to transform *E. coli* XL1-blue. The desired point mutation was verified by sequencing.

Characterization of GalK Mutants

The GalK Y371H mutant was overexpressed following the procedure previously described for wild-type *E. coli* GalK [8], whereas the overexpression of mutants M173L and M173L-Y371H was accomplished at 16°C as described for *L. lactis* GalK [11]. The mutant enzymes were purified by using metal affinity chromatography on Ni-NTA Spin Columns (QIAGEN, Valencia, California), and fractions containing homogenous protein were collected, concentrated, and quantified by using the Bradford protein assay [35]. The DNS assay was used to assess the substrate specificity of the purified GalK mutants as previously described [8]. A library of 45 different sugars was screened with each mutant (M173L, Y371H, and M173L-Y371H). For each sugar, the DNS assay was used to monitor the reaction progress, and control assays in the absence of enzyme or

sugar were performed in parallel. Standard curves for each sugar were prepared as described [8]. Each reaction contained 15.0 μM enzyme, 8 mM sugar, 14 mM ATP, and 3.5 mM $MgCl_2$ to assess general percent conversion. For total conversion assays, upon final addition of the enzyme, the reactions were incubated at 37°C for 3 hr after which the reactions were quenched with MeOH, centrifuged (10 min, 12,000 rpm), and then the supernatant (diluted 20-fold) submitted for LC-MS and MS/MS analysis. All assays were minimally run in triplicate with the average yields reported.

For monosaccharide kinetic determinations, the sugar concentration was varied over a range of 1–8 mM, under saturating ATP (14 mM), minimally in triplicate. Reaction progress was assessed via the DNS assay wherein a change in absorbance at 575 nm as a function of time was obtained and the initial velocity determined by the slope of the linear phase in the progress curve. The kinetic data was analyzed with Enzyme Kinetics Module software (SSPS, Inc., Chicago, Illinois) as previously described [8].

In Vivo Conversion Analysis

The GalK double mutant pGalKMLYH-*E. coli* was overexpressed at 16°C via induction of a 40 ml of culture at an OD_{600} ~0.7 with IPTG (1 mM). The induced cultures were incubated with shaking (140 rpm) for 1 hr, and 100 mM of **2** or **26** was added to the culture to a final concentration of 4 mM. The cultures were further incubated at 16°C with shaking (140 rpm) for 16 hr. The cells were harvested by centrifugation (15 min, 12,000 rpm), and the recovered cell pellet (380 mg) was washed until DNS analysis of the wash solution revealed no remaining endogenous extracellular-free monosaccharide (2 × 20 ml sodium phosphate buffer) to assess bioconversion. The pellet was frozen, thawed, resuspended in H_2O:MeOH (1:1), heated at 100°C for 15 min, and then sonicated 5 × 45 s on ice. An assay of this crude extract for GalK activity revealed all enzyme had been inactivated, and, thus, any final conversion detected is clearly due to in vivo conversion. Cell debris was collected by centrifugation (15 min, 12,000 rpm) and lyophilized to give the pale white solid (18 mg).

The amount of the fluorescent tag (**30**) needed for 1,3-dipolar cycloaddition was roughly calculated based upon the maximum amount of intracellular azidosugar conversion. A rough determination of extracellular-free sugar concentration was provided by pelleting the cells and assaying the supernatant via our DNS assay to assess the amount of intracellular sugar conversion. This was accomplished both at the beginning of the bioconversion experiment and at the end of the 16 hr incubation, the difference (above background based upon an assay of the same host in the same growth media but in the absence of endogenous unnatural sugar) we attributed to the maximum attainable intracellular free sugar. To a solution containing all the soluble crude product in 160 μl of H_2O:MeOH (1:1) was added 10 μmol **30** and 3.2 μmol of CuI, followed by heating to 50°C for 24 hr [6]. The reaction mixture was subsequently centrifuged to remove CuI, and the supernatant (diluted 2-fold) directly analyzed by HPLC and LC-MS. The yield calculation was based on the integration of the peak area for the substrates and products and corresponds to only intracellular reactants and products. All assays were minimally run in triplicate with the average yield reported.

Acknowledgments

This contribution was supported in part by the National Institutes of Health grants AI52218, CA84374, and GM70637. We are grateful to the University of Wisconsin-Madison School of Pharmacy Analytical Facility for analytical support.

Received: January 31, 2005
Revised: March 28, 2005
Accepted: April 11, 2005
Published: June 24, 2005

References

1. Weymouth-Wilson, A.C. (1997). The role of carbohydrates in biologically active natural products. Nat. Prod. Rep. *14*, 99–110.
2. Potier, P. (1999). Nature: a good source of new useful compounds for the health of man, animals and plants... for a long time to come! Actual. Chim. *11*, 9–11.
3. Thorson, J.S., and Vogt, T. (2003). Glycosylated natural products. In Glycosylated Natural Products in Carbohydrate-based Drug Discovery, Volume II, C.-W. Wong, ed. (Germany: Wiley-VCH Weinheim), pp. 685–711.
4. Kren, V., and Martinkova, L. (2001). Glycosides in medicine: "the role of glycosidic residue in biological activity". Curr. Med. Chem. *8*, 1303–1328.
5. Thorson, J.S., Hosted, T.J., Jr., Jiang, J., Biggins, J.B., and Ahlert, J. (2001). Nature's carbohydrate chemists: the enzymatic glycosylation of bioactive bacterial metabolites. Curr. Org. Chem. *5*, 139–167.
6. Fu, X., Albermann, C., Jiang, J., Liao, J., Zhang, C., and Thorson, J.S. (2003). Antibiotic optimization via in vitro glycorandomization. Nat. Biotechnol. *21*, 1467–1469.
7. Fu, X., Albermann, C., Zhang, C., and Thorson, J.S. (2004). Diversifying vancomycin via chemoenzymatic strategies. Org. Lett. *7*, 1513–1515.
8. Yang, J., Fu, X., Jia, Q., Shen, J., Biggins, J.B., Jiang, J., Zhao, J., Schmidt, J.J., Wang, P.G., and Thorson, J.S. (2003). Studies on the substrate specificity of *Escherichia coli* galactokinase. Org. Lett. *5*, 2223–2226.
9. Hoffmeister, D., Yang, J., Liu, L., and Thorson, J.S. (2003). Creation of the first anomeric D/L-sugar kinase by means of directed evolution. Proc. Natl. Acad. Sci. USA *100*, 13184–13189.
10. Hoffmeister, D., and Thorson, J.S. (2004). Mechanistic implications of *Escherichia coli* galactokinase structure-based engineering. ChemBioChem *5*, 989–992.
11. Yang, J., Liu, L., and Thorson, J.S. (2004). Structure-based enhancement of the first anomeric glucokinase. ChemBioChem *5*, 992–996.
12. Barton, W.A., Biggins, J.B., Jiang, J., Thorson, J.S., and Nikolov, D.B. (2002). Expanding pyrimidine diphosphosugar libraries via structure-based nucleotidylyltransferase engineering. Proc. Natl. Acad. Sci. USA *99*, 13397–13402.
13. Barton, W.A., Biggins, J.B., Lesniak, J., Jeffrey, P.D., Jiang, J., Rajashankar, K.R., Thorson, J.S., and Nikolov, D.B. (2001). Structure, mechanism and engineering of a nucleotidylyltransferase as a first step toward glycorandomization. Nat. Struct. Biol. *8*, 545–551.
14. Jiang, J., Albermann, C., and Thorson, J.S. (2003). Application of the nucleotidylyltransferase Ep toward the chemoenzymatic synthesis of dTDP-desosamine analogues. ChemBioChem *4*, 443–446.
15. Jiang, J., Biggins, J.B., and Thorson, J.S. (2001). Expanding the pyrimidine diphosphosugar repertoire: the chemoenzymatic synthesis of amino- and acetamidoglucopyranosyl derivatives. Angew. Chem. Int. Ed. Engl. *40*, 1502–1505.
16. 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.
17. Andreana, P.R., McLellan, J.S., Chen, Y., and Wang, P.G. (2002). Synthesis of 2,6-dideoxysugars via ring-closing olefinic metathesis. Org. Lett. *4*, 3875–3878.
18. Northrup, A.B., and MacMillan, D.W. (2004). Two-step synthesis of carbohydrates by selective aldol reactions. Science *305*, 1752–1755.
19. Northrup, A.B., Mangion, I.K., Hettche, F., and MacMillan, D.W. (2004). Enantioselective organocatalytic direct aldol reactions of alpha-oxyaldehydes: step one in a two-step synthesis of carbohydrates. Angew. Chem. Int. Ed. Engl. *43*, 2152–2154.
20. Albermann, C., Soriano, A., Jiang, J., Vollmer, H., Biggins, J.B., Barton, W.A., Lesniak, J., Nikolov, D.B., and Thorson, J.S. (2003). Substrate specificity of NovM: implications for novobiocin biosynthesis and glycorandomization. Org. Lett. *5*, 933–936.
21. Thorson, J.S., Barton, W.A., Hoffmeister, D., Albermann, C., and Nikolov, D.B. (2003). Structure-based enzyme engineering and its impact on in vitro glycorandomization. ChemBioChem *5*, 16–25.
22. Yang, J., Hoffmeister, D., Liu, L., Fu, X., and Thorson, J.S.

(2004). Natural product glycorandomization. Bioorg. Med. Chem. *12*, 1577–1584.
23. Langenhan, J.M., and Thorson, J.S. (2005). Recent carbohydrate-based chemoselective ligation application. Curr. Org. Syn. *2*, 59–81.
24. Zhang, J., Kowal, P., Chen, X., and Wang, P.G. (2003). Large-scale synthesis of globotriose derivatives through recombinant *E. coli*. Org. Biomol. Chem. *1*, 3048–3053.
25. Zhang, J., Chen, X., Shao, J., Liu, Z., Kowal, P., Lu, Y., and Wang, P.G. (2003). Synthesis of galactose-containing oligosaccharides through superbeads and superbug approaches: substrate recognition along different biosynthetic pathways. Methods Enzymol. *362*, 106–124.
26. Dube, D.H., and Bertozzi, C.R. (2003). Metabolic oligosaccharide engineering as a tool for glycobiology. Curr. Opin. Chem. Biol. *7*, 616–625.
27. Saxon, E., and Bertozzi, C.R. (2000). Cell surface engineering by a modified Staudinger reaction. Science *287*, 2007–2010.
28. Fuster, M.M., Brown, J.R., Wang, L., and Esko, J.D. (2003). A disaccharide precursor of sialyl Lewis X inhibits metastatic potential of tumor cells. Cancer Res. *63*, 2775–2781.
29. Mong, T.K., Lee, L.V., Brown, J.R., Esko, J.D., and Wong, C.H. (2003). Synthesis of N-acetyllactosamine derivatives with variation in the aglycon moiety for the study of inhibition of sialyl Lewis x expression. ChemBioChem *4*, 835–840.
30. Shao, J., Hayashi, T., and Wang, P.G. (2003). Enhanced production of alpha-galactosyl epitopes by metabolically engineered Pichia pastoris. Appl. Environ. Microbiol. *69*, 5238–5242.
31. Thoden, J.B., and Holden, H.M. (2003). Molecular structure of galactokinase. J. Biol. Chem. *278*, 33305–33311.
32. Chatterjee, R., Millard, C.S., Champion, K., Clark, D.P., and Donnelly, M.I. (2001). Mutation of the ptsG gene results in increased production of succinate in fermentation of glucose by *Escherichia coli*. Appl. Environ. Microbiol. *67*, 148–154.
33. Dien, B.S., Nichols, N.N., and Bothast, R.J. (2002). Fermentation of sugar mixtures using *Escherichia coli* catabolite repression mutants engineered for production of L-lactic acid. J. Ind. Microbiol. Biotechnol. *29*, 221–227.
34. Nichols, N.N., Dien, B.S., and Bothast, R.J. (2001). Use of catabolite repression mutants for fermentation of sugar mixtures to ethanol. Appl. Microbiol. Biotechnol. *56*, 120–125.
35. Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254.