

Perspective

Natural product glycorandomization[☆]

Jie Yang, Dirk Hoffmeister, Lesley Liu, Xun Fu and Jon S. Thorson*

Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin-Madison, 777 Highland Avenue, Madison, WI 53705 USA

Received 13 October 2003; accepted 19 December 2003

Abstract—Glycorandomization is a chemoenzymatic strategy that overcomes the limitations in natural product derivatization associated with both solely chemistry-based approaches or in vivo engineering. In this article we present the basic strategies for glycorandomization development as a next-generation tool in drug discovery.

© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

A recent estimate suggests that over half of the world's drug leads derive directly from the natural product pool, many of which are glycosylated secondary metabolites (Scheme 1). The functional contribution of carbohydrates to biologically active natural products has been reviewed in a variety of outstanding compilations.^{1–8} Emerging from the exciting work highlighted in these reviews is a clear picture of how sugar ligands can mediate drug targeting, biological activity and pharmacology which suggests that altering glycosylation patterns on secondary metabolites has a high potential for the generation of novel therapeutics.

There are a number of routes for altering the glycosylation of complex metabolites. The first, (i) synthesis/semi-synthesis, relies upon the total synthesis of analogues or synthetic modification of intermediates usually produced via fermentation or hydrolysis of the natural product itself. An advantage of this approach is that the variants produced are only limited by the available chemistries and synthetic expertise. However, a significant disadvantage is the enormous structural complexity of many glycosylated natural products (e.g., **1**). In vivo methods for altering glycosylation include (ii) pathway engineering (often referred to as 'combinatorial biosynthesis', Scheme 2A)^{9–12} and (iii) bioconversion

(Schemes 2B and 2C).^{13–15} A significant advantage of these routes is the ability to access new compounds via fermentation. Yet, many enzymes within sugar ligand biosynthetic pathways (SLBPs) significantly bias the available chemistry and clearly limit the extent of accessible sugar variation in the final product. Moreover, the yields of shunt metabolites are typically <10% that of wild-type product.¹⁶

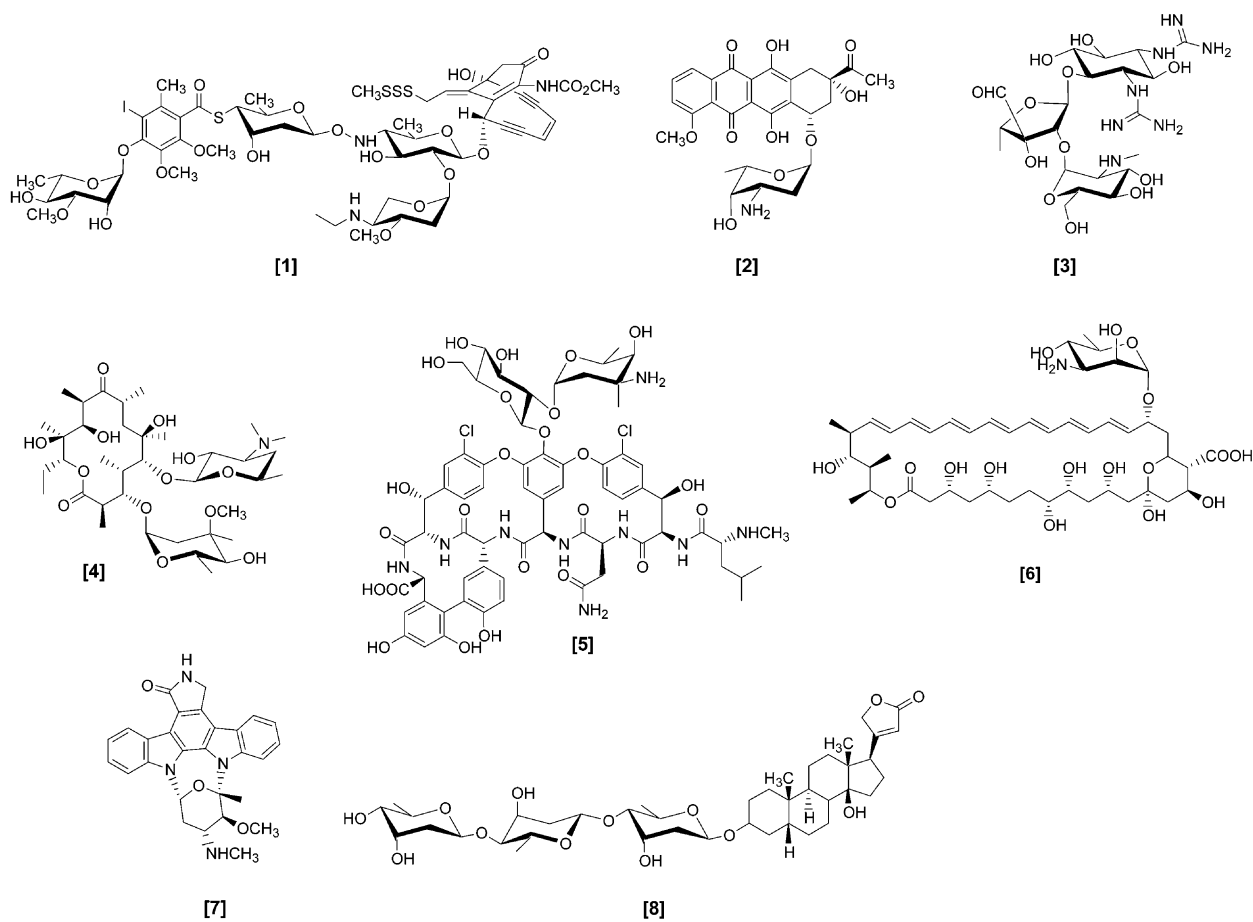
A final route for altering glycosylation is via (iv) biocatalysis and is exemplified by natural product 'glycorandomization'.^{4,8,17–25} In vitro glycorandomization (IVG) takes advantage of combining the limitless flexibility of the chemical synthesis of unique sugar precursors with the inherent or engineered substrate promiscuity of enzymes to activate (sugar kinases and nucleotidyltransferases) and attach (glycosyltransferases) these carbohydrates to various natural product aglycons (Scheme 3). Specifically for IVG, natural and 'unnatural' sugar precursors are chemically synthesized and attached to various aglycons via a single pot, three-enzyme (sugar kinase–nucleotidyltransferase–glycosyltransferase) process. This methodology is advantageous in that it combines the strength of chemical synthesis with the ease of regio- and stereo-specific enzymatic coupling of sugars to extremely complex aglycon structures. Furthermore, this methodology allows for the efficient incorporation of sugars bearing 'reactive handles' (e.g., azides,²⁶ thiols,²⁶ ketones,²⁶ aminoxy²⁷ substituents) which can then be specifically modified, in the context of a very complex natural product architecture, to enhance the diversity of the final glycorandomized natural product library.

This review briefly highlights the progress in developing the essential components of natural product glyco-

Keywords: Glycosyltransferase; Enzyme catalysis; Galactokinase; Nucleotidyltransferase; Natural products; Antibiotics.

[☆]This contribution was supported in part by the National Institutes of Health (GM58196, CA84374 and AI52218). D.H. is research fellow of the Deutsche Forschungsgemeinschaft (DFG)

*Corresponding author. Fax: +1-608-262-5345; e-mail: jsthorson@pharmacy.wisc.edu



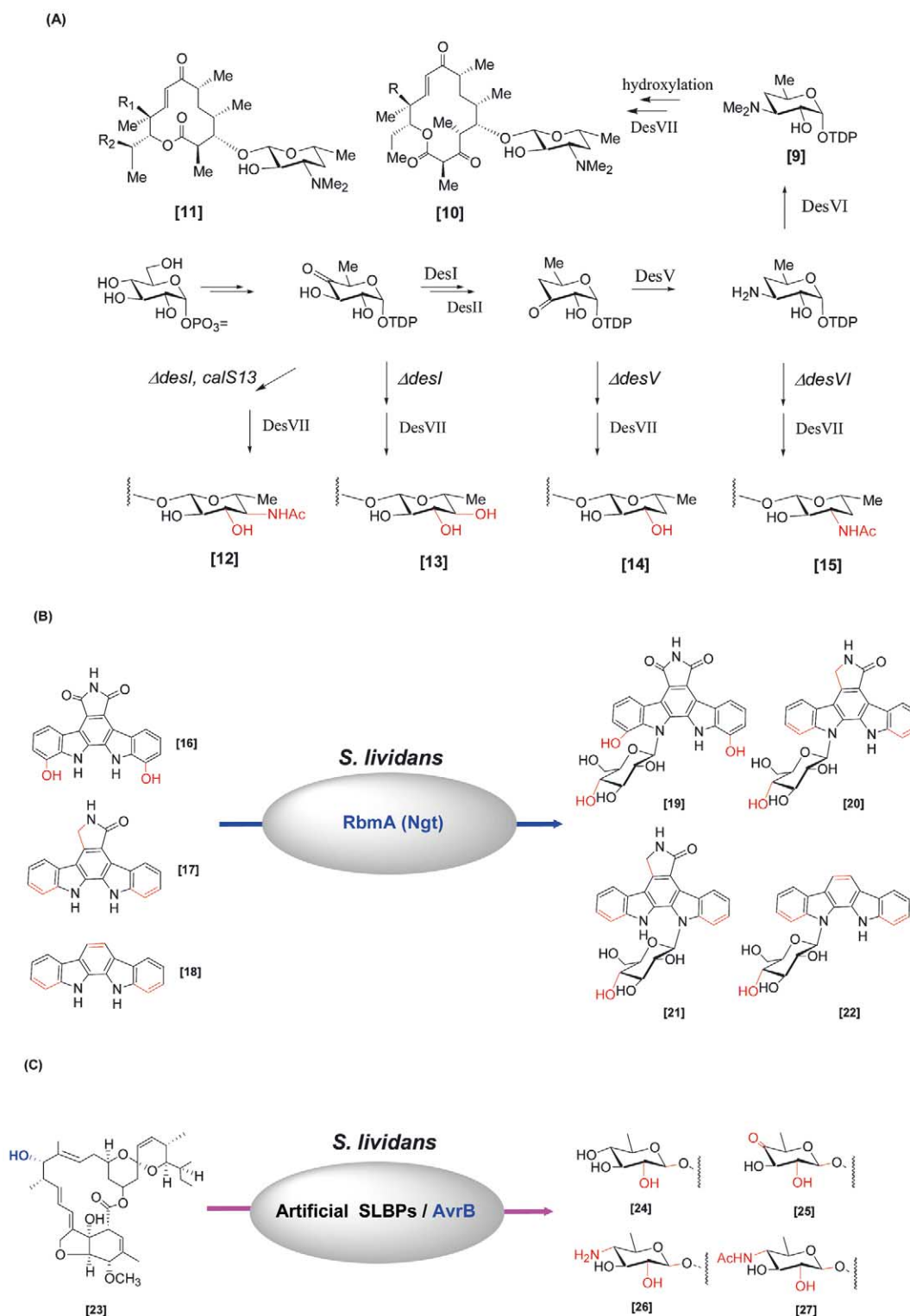
Scheme 1. Representative examples of therapeutically relevant glycosylated secondary metabolites. Natural product sugar ligands have been shown to be critical to DNA-recognition (calicheamicin, **1**) and inhibition of DNA-processing (daunomycin, **2**) RNA-recognition (streptomycin, **3**) and inhibition of translation (erythromycin, **4**), inhibition of cell wall synthesis (vancomycin, **5**) and membrane recognition (amphotericin, **6**), as well as targeting specific proteins (staurosporine, **7**) and/or protein complexes (cardiac glycosides, e.g., digitoxin **8**).

randomization: the generation of promiscuous sugar kinases and nucleotidyltransferases; their successful combination with glycosyltransferases to demonstrate the glycorandomization platform; and finally, the application of downstream chemoselective ligation strategies to further enhance the level of natural product diversity obtained.

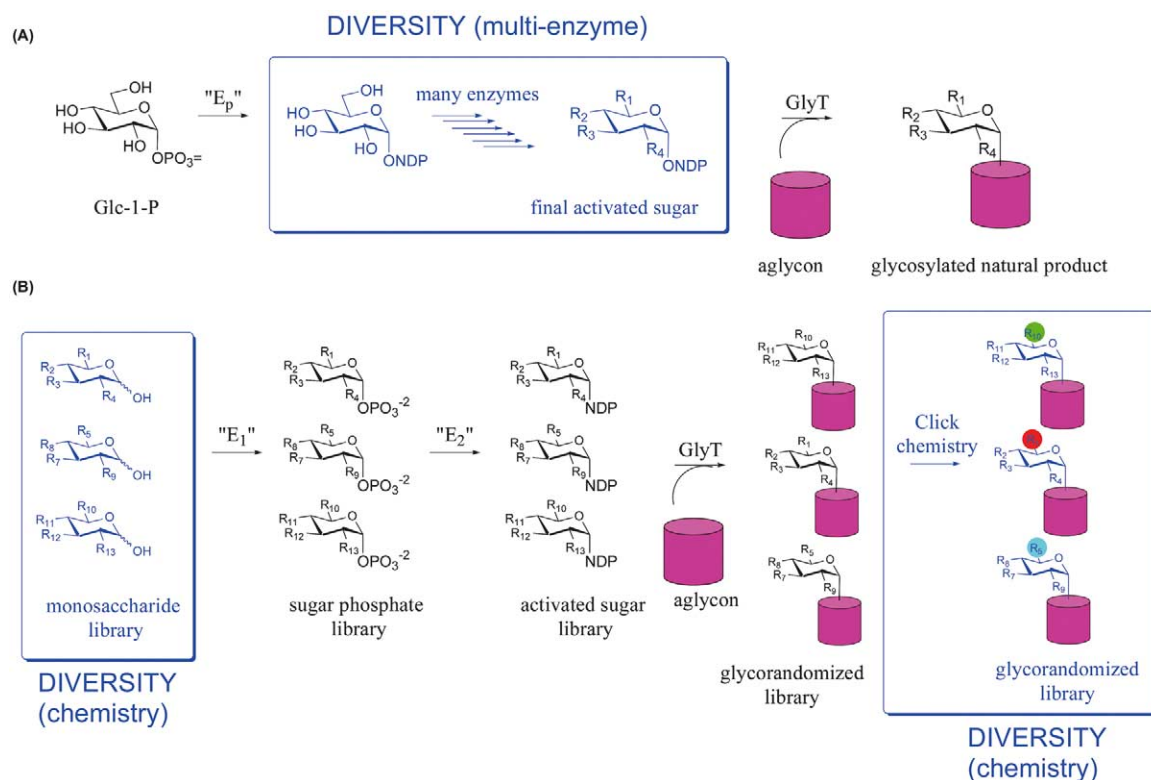
2. GalK, the model anomeric sugar kinase for glyco-randomization

Galactokinases (GalK) catalyze the formation of α -D-galactose-1-phosphate (Gal-1-P) (**29**) from D-galactose (**28**) and ATP (Scheme 4A). Yet, the GalKs characterized to date are known to be specific for one or only a few monosaccharides (Scheme 4B).^{28–31} Moreover, in all C-1 kinases studied thus far, a strict adherence to either D-sugars (GalK and glycogen phosphorylases),^{28–32} or L-sugars (as in fucokinase)³³ was observed. Thus, to apply any of these anomeric kinases toward generating a randomized sugar phosphate library, their monosaccharide substrate promiscuity must first be enhanced.

Prior to the emergence of the first structure for a sugar C-1 kinase (GalK from *Lactococcus lactis*),³⁴ we embarked upon a directed evolution approach to develop a flexible sugar anomeric kinase. As a model system, we selected the well-characterized *Escherichia coli* GalK³⁵ and focused the initial screens toward significant C-5 (e.g., L-sugar variants) and C-6 alterations (e.g., deoxy, amino, uronic acid derivatives). From this approach, one particular GalK mutant (Y371H) displayed a surprising degree of kinase activity toward sugars as diverse as D-talose (**34**), D-galacturonic acid (**36**), L-altrose (**37**), and L-glucose (**38**), all of which failed as wild-type GalK substrates (Scheme 4B).²⁴ Furthermore, this mutant provides enhanced turnover of the small pool of sugars converted by the wild-type enzyme (Scheme 4B). Comparison of this mutation to the recently solved structure of *Lactococcus lactis* GalK, begins to provide a blueprint for further engineering of this vital class of enzyme. Most importantly, our enhanced GalK expedites the production of sugar-1-phosphate substrate libraries for the next enzyme in IVG, Ep.



Scheme 2. (A) An example of the application of genetic engineering toward changing macrolide glycosylation. Disruption (as indicated by ‘Δ’) of genes leading to the biosynthesis of dTDP-desosamine (**9**), a precursor to pikromycin (**10**)/methymycin (**11**) and related macrolides in *S. venezuelae*, leads to macrolides bearing novel sugars. In addition, introduction of biosynthetic genes from other pathways (*desI*, *calS13* which incorporates a sugar 4-aminotransferase from the calicheamicin pathway in *M. echinospora*) can lead to further diversity in glycosylation (structural alterations from the natural metabolites highlighted in red). (B) An example of the application of bioconversion toward altering indolocarbazole glycosylation. In this example, the *N*-glycosyltransferase (*rbmA*) gene from rebeccamycin biosynthesis in *S. aerocolonigenes* was expressed in the heterologous host *S. lividans*. Unnatural aglycons were then fed to the RbmA-*S. lividans* recombinant strain leading to novel indolocarbazole analogues (structural alterations from rebeccamycin highlighted in red). (C) An example of the application of artificial biosynthetic sugar pathways toward the generation of novel avermectin derivatives. In this example, combinations of the TDP-D-desosamine (pikromycin/methymycin, *S. venezuelae*) and TDP-L-oleandrose (avermectin, *S. avermitilis*) biosynthetic genes were assembled in a non-producing host *S. lividans* engineered to express the avermectin glycosyltransferase gene, *avrB*. Upon feeding this host the avermectin aglycon, novel D-sugar substituted avermectins were produced (alterations from avermectin highlighted in red).



Scheme 3. (A) The standard route for secondary metabolite glycosylation in vivo. Generation of the fully functionalized activated NDP-sugars can require up to ten distinct transformations by unique enzymes prior to the culminating glycosyltransferase (GlyT)-catalyzed attachment to the aglycon. (B) Schematic for natural product in vitro glycorandomization illustrating the great potential for enhanced diversity via the simplistic bypass of the many specific SLBP enzymes. In IVG, the diversity stems from the upstream chemical synthesis of novel monosaccharides and the downstream chemical modification, via chemical ligation strategies. The core of IVG is dependent upon three key enzymes (E_1 , kinase; E_2 , nucleotidyltransferase; and GlyT, glycosyltransferase) to accomplish the difficult sugar-activating and attachment chemistry in a stereo- and regiospecific manner.

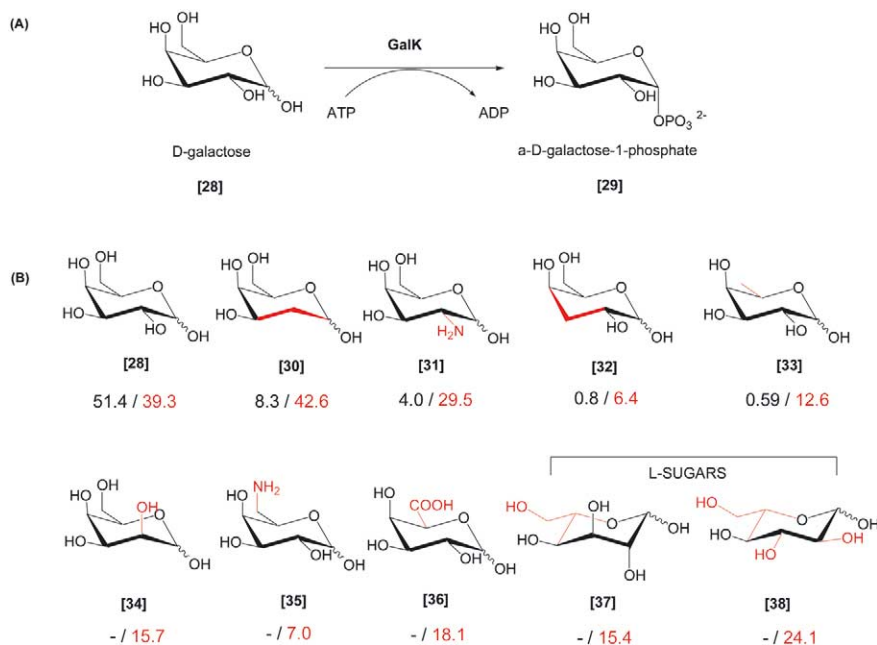
3. E_p , A model nucleotidyltransferase for glycorandomization

Out of the vast number of available nucleotidyltransferases, structure-based engineering work began with the *rmlA*-encoded α -D-glucopyranosyl phosphate thymidyltransferase (E_p) from *Salmonella enterica* LT2.³⁶ E_p catalyzes the conversion of α -D-glucopyranosyl phosphate (Glc-1-P) (**40**) and dTTP (**39**) to dTDP- α -D-glucose (dTDP-Glc) (**41**) and pyrophosphate (PPi) (Scheme 5A), via a single sequential displacement mechanism,¹⁹ and we found E_p to be unique among nucleotidyltransferases in that it displays unusual promiscuity toward both its nucleotide triphosphate (dTTP and UTP) and the sugar phosphate substrates (Scheme 5B).^{17,18,21} Yet limitations, as a function of sterics, ring conformation, and/or electrostatics prohibits the use of E_p in a true combinatorial sense. Thus, a structure-based engineering program, in conjunction with our continued sugar-1-phosphate syntheses program (which has been more recently mediated by the above mentioned enhanced GalK), was pursued to elucidate E_p variants capable of utilizing an expanded sugar-1-

phosphate set.^{19,20,25} This cumulative effort has led to > 50 functional substrates translating to > 100 novel NDP-sugars, some of which are illustrated in Scheme 5B, to serve as substrates for glycosyltransferases as the final enzymatic step of constructing glycorandomized libraries.

4. GtfE, a model glycosyltransferase for glycorandomization

GtfE, the first of two tandem glycosyltransferases in vancomycin biosynthesis (Scheme 6A),^{37,38} served as an early model to demonstrate the glycorandomization of nonribosomal peptide-based natural products.^{22,23} Specifically, 33 natural and 'unnatural' NDP-sugars were individually tested as substrates for GtfE (Scheme 6B).^{22,23} As illustrated in Scheme 6B, 31 from this set were accepted as substrates (> 25% conversion). Prior to this work, the number of total monoglycosylated vancomycin analogues totaled 11 members constructed via multi-step chemical synthesis.³⁹ In addition, the 'unnatural' substrate **44** was previously shown to function



Scheme 4. (A) Reaction catalyzed by galactokinase (GalK). (B) ‘Natural’ and ‘unnatural’ substrates of wild-type GalK (k_{cat}/K_m illustrated in black) and evolved GalK Y371H (k_{cat}/K_m illustrated in red). The positions deviating from the natural GalK substrate D-galactose are also highlighted in red.

with the second glycosyltransferase GtfD,²² which suggests further potential to build upon this library by the enzymatic addition of vancosamine or, possibly, other sugars. More importantly, glycoconjugates bearing reactive ‘handles’ such as compounds **45–49** from this small library present the opportunity to test chemoselective modification, or chemoglycorandomization.²³

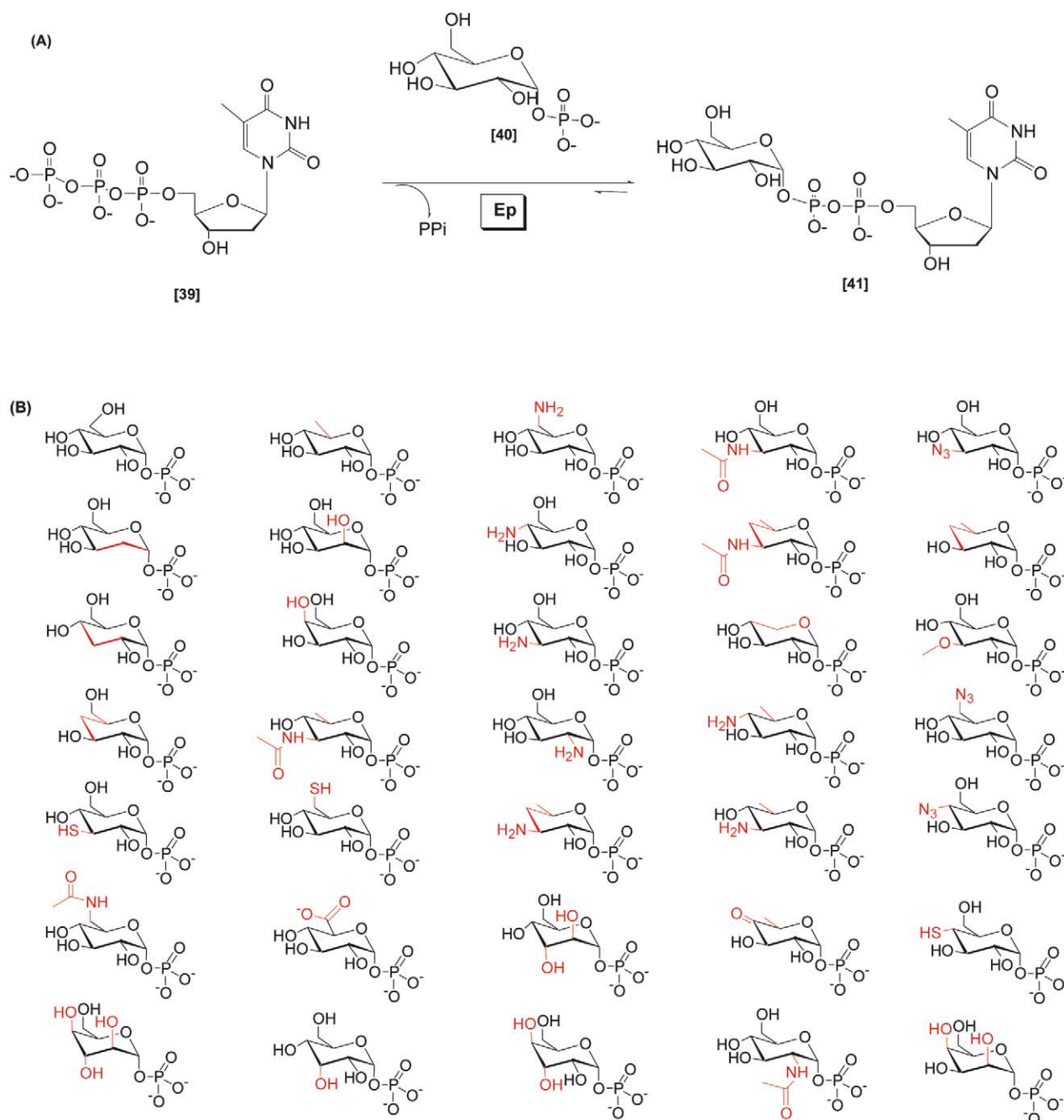
5. Chemoglycorandomization

Chemoselective ligation reactions among two mutually and uniquely reactive functional groups in an aqueous environment have been used extensively in recent years for the modification of biopolymers and offer advantages similar to those of enzymatic reactions (efficiency, regio- and stereospecificity), with the advantage of a much broader range of coupling partners.^{40,41} One such reaction is the Huisgen 1,3-dipolar cycloaddition of azides and acetylenes to give 1,2,3-triazoles (Scheme 6C).^{41,42} To test the feasibility of 1,3-dipolar addition toward the chemoselective diversification of vancomycin, the monoglycosylated variant **45** was incubated with a variety of commercially available and modified alkynes followed by analysis of reaction progress via HPLC and the outcome of representative chemoselective ligations is illustrated in Scheme 6C. From this small pilot demonstration, the library of known monoglycosylated variants was increased to a total of 50 members bearing uniquely diverse functionality, further illustrating the significant potential of this approach toward natural product diversification. Antibacterial screens of this library

revealed the first monoglycosylated vancomycins which rival the biological activity of parent natural product and clearly support the significance of sugar substitution as a means to enhance desired properties.^{39,43,44}

6. Future directions

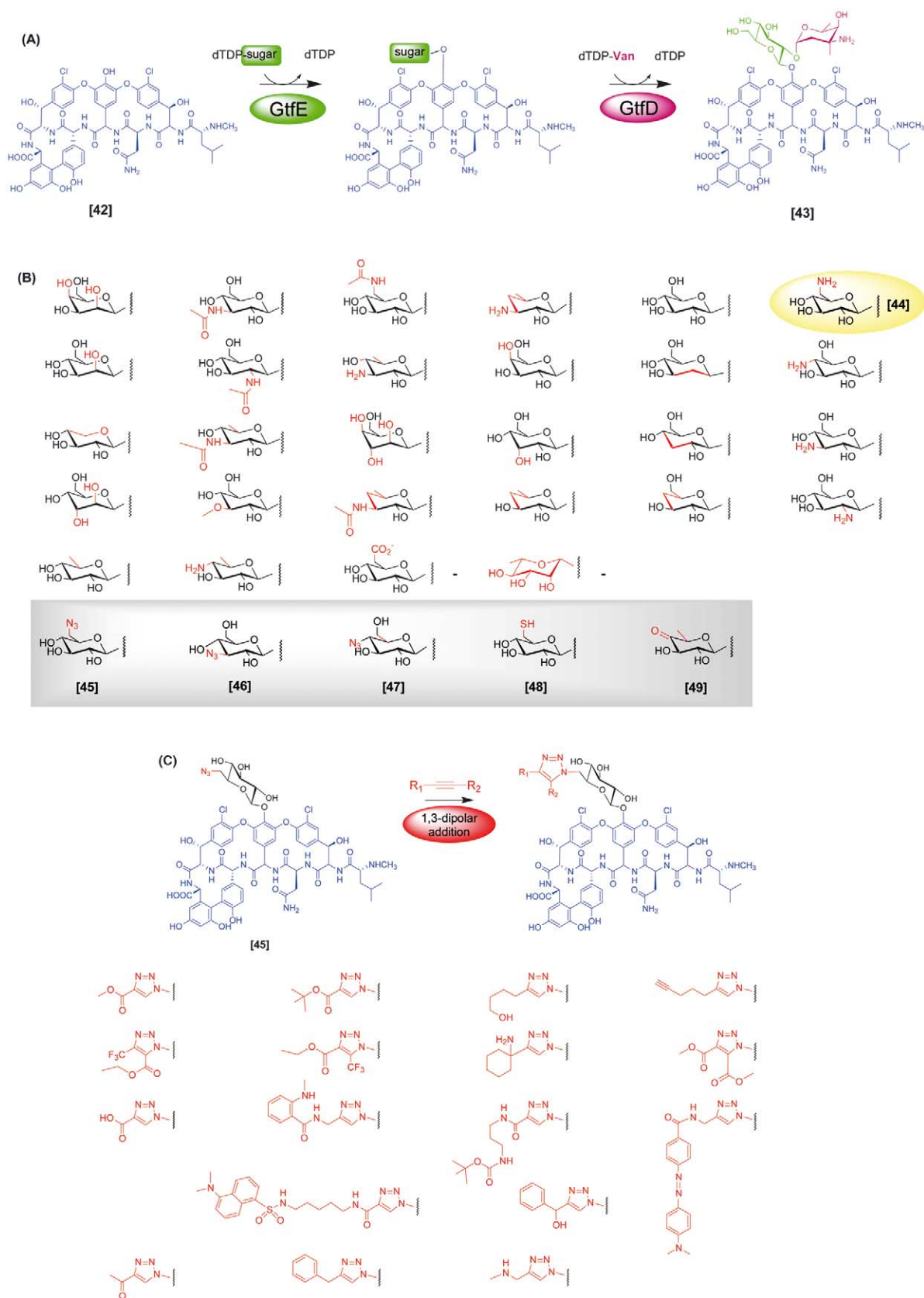
Given glycosyltransferases for the biosynthesis of many vital therapeutics are available and known to be promiscuous,^{2,4,45} our current results suggest IVG holds significant promise for future drug development. In its present form, a primary limitation of IVG stems from its reliance upon an in vitro process which potentially contributes to two key issues (1) scaling and (2) generalizing the process. First, scalability of the process is hampered by expensive substrates/cofactors albeit, spectacular solutions by Wong and others for regenerating these reagents are available.^{46,47} Second, the application of IVG to most, or all, classes of glycosylated natural products is heavily dependent upon the expression of appropriate glycosyltransferases and establishing in vitro conditions for an active enzyme which, in some cases, can be severely dictated by the solubility of the aglycon acceptor. We expect to circumvent these limitations by converting IVG to an in vivo process and propose the advent of kinase-enhanced IVG potentially opens the door to explore at least two in vivo glycorandomization scenarios (Scheme 7). First, the expression of a tandem promiscuous sugar-1-kinase (GalK) and nucleotidyltransferase (E_p)—essentially an unnatural NDP-sugar factory—in a natural



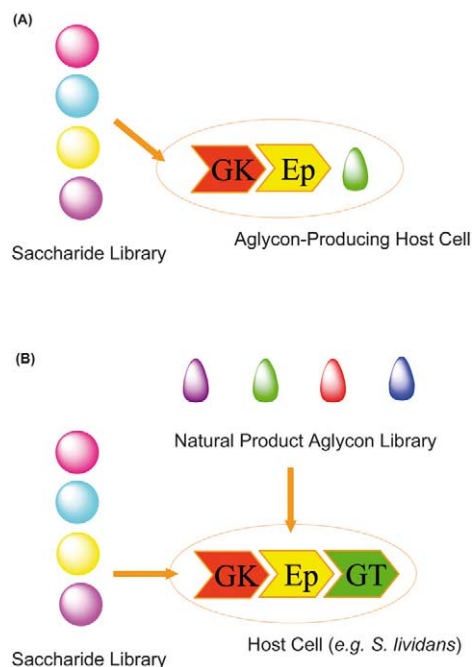
Scheme 5. (A) Reaction catalyzed by 'Ep'. (B) 'Unnatural' substrates of wild-type Ep and structure-based engineered Ep variants. The positions deviating from the natural Ep substrate D-Glc-1-P are highlighted in red.

aglycon-producing host (e.g., the 1-producing *Microspora*) 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 (Scheme 7A). Alternatively, expression of the tandem two gene 'NDP-sugar factory' genes in a non-producing host (e.g., *S. lividans* or possibly even *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 (Scheme 7B). The observations that functional artificial NDP-sugar pathways can be assembled in prokaryotes,^{48,49} in conjunction with the fact that natural and 'unnatural' sugar are processed in vivo by both prokaryotes and eukaryotes,^{50–53} lend support to this idea.



Scheme 6. (A) The final stages of vancomycin biosynthesis catalyzed by glucosyltransferase GtfE (green) and vancosaminyltransferase GtfD (magenta). (B) Variants from IVG of the vancomycin aglycon where (–) indicates no conversion and all others >25% conversion. The changes from the GtfE wild-type substrate (glucose) are highlighted in red. The compound highlighted in the yellow circle represents the substrate previously shown to function with the second glycosyltransferase GtfD.²² Compounds highlighted in the grey box represent the substrates presenting the opportunity to test chemoselective modification. (C) Products from the chemoselective Huisgen 1,3-dipolar cycloaddition of azide **[45]** and acetylenes to give the corresponding 1,2,3-triazoles.



Scheme 7. Two scenarios for natural product in vivo glyco-randomization illustrating the great potential for enhanced diversity. (A) Feeding monosaccharides to a native aglycon-producing host engineered to express the 'NDP-sugar factory'. (B) Feeding monosaccharides and aglycons to a non-producing host engineered to express the 'NDP-sugar factory' and appropriate glycosyltransferases.

References and notes

- Potier, P. *Actual. Chim.* **1999**, *11*, 9.
- Weymouth-Wilson, A. C. *Nat. Prod. Rep.* **1997**, *14*, 99.
- Vogt, T.; Jones, P. *Trends Plant Sci.* **2000**, *5*, 380.
- Thorson, J. S.; Hosted, T. J., Jr.; Jiang, J.; Biggins, J. B.; Ahlert, J. *Curr. Org. Chem.* **2001**, *5*, 139.
- Kren, V.; Martinkova, L. *Curr. Med. Chem.* **2001**, *8*, 1303.
- Jones, P.; Vogt, T. *Planta* **2001**, *213*, 164.
- Keegstra, K.; Raikhel, N. *Curr. Opin. Plant Biol.* **2001**, *4*, 219.
- Thorson, J. S.; Vogt, T. In *Carbohydrate based Drug Discovery*; Wong, C.-W. (Ed.); Glycosylated Natural Products, 2002, Wiley-VCH.
- Zhao, L. S.; Que, N. L. S.; Xue, Y.; Sherman, D. H.; Liu, H.-w. *J. Am. Chem. Soc.* **1998**, *120*, 12159.
- Zhao, L. S.; Sherman, D. H.; Liu, H.-w. *J. Am. Chem. Soc.* **1998**, *120*, 10256.
- Borisova, S. A.; Zhao, L. S.; Sherman, D. H.; Liu, H.-w. *Org. Lett.* **1999**, *1*, 133.
- Zhao, L. S.; Ahlert, J.; Xue, Y. Q.; Thorson, J. S.; Sherman, D. H.; Liu, H.-w. *J. Am. Chem. Soc.* **1999**, *121*, 9881.
- (a) Ohuchi, T.; Ikeda-Araki, T. A.; Watanabe-Sakamoto, A.; Kojiri, K.; Nagashima, M.; Okanishi, M.; Suda, H. *J. Antibiot.* **2000**, *53*, 393. (b) Sanchez, C.; Butovich, I. A.; Brana, A. F.; Rohr, J.; Mendez, C.; Salas, J. A. *Chem. Biol.* **2002**, *9*, 519. (c) Hyun, C.-G.; Bililign, T.; Liao, J.; Thorson, J. S. *ChemBioChem.* **2003**, *1*, 114.
- Albermann, C.; Hyun, C.-G.; Thorson, J. S. Unpublished.
- Liao, J.; Lomovskaya, N.; Fonstein, L.; Wohler, S.-E.; Hutchinson, C. R.; Thorson, J. S. Unpublished.
- He, X. M.; Liu, H.-w. *Annu. Rev. Biochem.* **2002**, *71*, 701.
- Jiang, J.; Biggins, J. B.; Thorson, J. S. *J. Am. Chem. Soc.* **2000**, *122*, 6803.
- Jiang, J.; Biggins, J. B.; Thorson, J. S. *Angew. Chem. Intl. Ed. Engl.* **2001**, *40*, 1502.
- Barton, W. A.; Lesniak, J.; Biggins, J. B.; Jeffrey, P. D.; Jiang, J.; Rajashankar, K. R.; Thorson, J. S.; Nikolov, D. B. *Nat. Struct. Biol.* **2001**, *8*, 545.
- Barton, W. A.; Biggins, J. B.; Jiang, J.; Thorson, J. S.; Nikolov, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1339.
- Jiang, J.; Albermann, C.; Thorson, J. S. *ChemBioChem.* **2003**, *4*, 443.
- Losey, H. C.; Jiang, J.; Biggins, J. B.; Oberthur, M.; Ye, X.-Y.; Dong, S. D.; Thorson, J. S.; Kahne, D.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 1305.
- Fu, X.; Albermann, C.; Jiang, J.; Liao, J.; Zhang, C.; Thorson, J. S. *Nat. Biotechnol.* **2003**, *21*, 1467.
- Hoffmeister, D.; Yang, J.; Liu, L.; Thorson, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13184.
- Thorson, J. S.; Barton, W. A.; Hoffmeister, D.; Albermann, C.; Nikolov, D. B. *ChemBioChem.* **2004**, *5*, 16.
- Albermann, C.; Jiang, J.; Thorson, J. S. Unpublished.
- Langenhan, J. M.; Liao, J.; Thorson, J. S. Unpublished.
- Lavine, J. E.; Cantlay, E.; Roberts, C. T., Jr.; Morse, D. E. *Biochim. Biophys. Acta* **1982**, *717*, 76.
- Dey, P. M. *Eur. J. Biochem.* **1983**, *136*, 155.
- Thomas, P.; Bessell, E. M.; Westwood, J. H. *Biochem. J.* **1974**, *139*, 661.
- Yang, J.; Fu, X.; Jia, Q.; Shen, J.; Biggins, J. B.; Jiang, J.; Zhao, J.; Schmidt, J. J.; Wang, P. G.; Thorson, J. S. *Org. Lett.* **2003**, *5*, 2223.
- Johnson, L. N.; Barford, D.; Biol, J. *Chem.* **1990**, *265*, 2409.
- Park, S. H.; Pastuszak, I.; Drake, R.; Elbein, A. D. *J. Biol. Chem.* **1998**, *273*, 5685.
- Thoden, J. B.; Holden, H. M. *J. Biol. Chem.* **2003**, *278*, 33305.
- Debouck, C.; Riccio, A.; Schumperli, D.; McKenney, K.; Jeffers, J.; Hughes, C.; Rosenberg, M.; Heusterspreute, M.; Brunel, F.; Davison, J. *Nuc. Acids Res.* **1985**, *13*, 1841.
- Lindquist, L.; Kaiser, R.; Reeves, P. R.; Lindberg, A. A. *Eur. J. Biochem.* **1993**, *211*, 763.
- Solenberg, P. J.; Matsushima, P.; Stack, D. R.; Wilkie, S. C.; Thompson, R. C.; Baltz, R. H. *Chem. Biol.* **1997**, *4*, 195.
- Losey, H. C.; Peczu, M. W.; Chen, Z.; Eggert, U. S.; Dong, S. D.; Pelczer, I.; Kahne, D.; Walsh, C. T. *Biochem.* **2001**, *40*, 4745.
- Nicolaou, K. C.; Cho, S. Y.; Hughes, R.; Winssinger, N.; Smethurst, C.; Labischinski, H.; Endermann, R. *Chem. Eur. J.* **2001**, *7*, 3798.
- Hang, H. C.; Bertozzi, C. R. *Acc. Chem. Res.* **2001**, *34*, 727.
- Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004.
- Hlasta, D. J.; Ackerman, J. H. *J. Org. Chem.* **1994**, *59*, 6184.
- Nicolaou, K. C.; Boddy, C. N.; Brase, S.; Winssinger, N. *Angew. Chem. Int. Ed.* **1999**, *38*, 2096.
- Ge, M.; Chen, Z.; Onishi, H. R.; Kohler, J.; Silver, L. L.; Kerns, R.; Fukuzawa, S.; Thompson, C.; Kahne, D. *Science* **1999**, *284*, 507.
- Mendez, C.; Salas, J. A. *Trends Biotechnol.* **2001**, *19*, 449.
- Koeller, K. M.; Wong, C.-H. *Nature* **2001**, *409*, 232.
- Koeller, K. M.; Wong, C. H. *Glycobiol.* **2000**, *10*, 1157.
- Zhang, J.; Kowal, P.; Chen, X.; Wang, P. G. *Org. Biomol. Chem.* **2003**, *1*, 3048.
- Zhang, J.; Chen, X.; Shao, J.; Liu, Z.; Kowal, P.; Lu, Y.; Wang, P. G. *Methods Enzymol.* **2003**, *362*, 106–124.
- Luchansky, S. J.; Hang, H. C.; Saxon, E.; Grunwell, J. R.; Yu, C.; Dube, D. H.; Bertozzi, C. R. *Methods Enzymol.* **2003**, *362*, 249.
- Fuster, M. M.; Brown, J. R.; Wang, L.; Esko, J. D. *Cancer Res.* **2003**, *63*, 2775.
- Mong, T. K.; Lee, L. V.; Brown, J. R.; Esko, J. D.; Wong, C. H. *ChemBiochem* **2003**, *4*, 835.
- Shao, J.; Hayashi, T.; Wang, P. G. *Appl. Environ. Microbiol.* **2003**, *69*, 5238.