Article

Chemoenzymatic Synthesis of Unnatural Nucleotide Sugars for Enzymatic Bioorthogonal Labeling

Liuqing Wen, Madhusudhan Reddy Gadi, Yuan Zheng, Christopher Gibbons, Shukkoor Muhammed Kondengaden, Jiabin Zhang, and Peng George Wang

ACS Catal., **Just Accepted Manuscript** • Publication Date (Web): 12 Jul 2018

Chemoenzymatic Synthesis of Unnatural Nucleotide Sugars for Enzymatic Bioorthogonal Labeling

Liuqing Wen,* Madhusudhan Reddy Gadi, Yuan Zheng, Christopher Gibbons, Shukkoor Muhammed Kondengaden, Jiabin Zhang, Peng George Wang*

Department of Chemistry, Georgia State University, Atlanta, Georgia, 30303, United States

ABSTRACT: In recent years, the development of enzymatic bioorthogonal labeling strategy offers exciting possibilities in the probing of structure-defined glycan epitopes. This strategy takes advantage of relaxed donor specificity and strict acceptor specificity of glycosyltransferases to label target glycan epitopes with bioorthogonal reactive groups carried by unnatural nucleotide sugars *in vitro*. The subsequent covalent conjugation by bioorthogonal chemical reactions with either fluorescent or affinity tags allows further visualization, quantification, or enrichment of target glycan epitopes. However, the application and development of enzymatic labeling strategy have been hindered due to the limit availability of unnatural nucleotide sugars. Herein, a platform that combines chemical synthesis and enzymatic synthesis for the facile preparation of unnatural nucleotide sugars modified with diverse bioorthogonal reactive groups is described. By this platform, total 25 UDP-GlcNAc and UDP-GalNAc derivatives, including the most well-explored bioorthogonal functional groups, were successfully synthesized. Furthermore, the potential application of these compounds for use in enzymatic bioorthogonal labeling reactions was also evaluated.
KEYWORDS: *unnatural nucleotide sugars; bioorthogonal chemistry; chemoenzymatic labeling; chemoenzymatic synthesis; glycan biomarker; glycan epitopes*

INTRODUCTION

Glycans are involved in a variety of physiological and pathological processes such as protein folding and degradation, immune responses, cell growth and proliferation, cell-cell communications, and cell-pathogen interactions.[1-3] It is well established that aberrant glycosylation in the glycan structure and site occupancy is a hallmark of human cancer and many other diseases.[4-8] Unlike like DNA transcription and protein translation in living cells, glycan assembly is not a template-driven process. Instead, glycans are installed onto proteins or lipids through a stepwise, enzyme-catalyzed process. The size and structural complexity of glycans have posed great challenges in glycan detection. Traditionally, chemical modification and affinity binding by antibody or lectin have been the primary methods used for glycan analysis.[9-12] However, chemical modification of glycans is normally destructive. For example, periodic acid, which is widely used for sialic acid analysis, can cause permanent damage to the treated glycans,[13] preventing downstream analysis. Alternatively, antibody and lectin binding based on non-covalent interaction is a mild and non-destructive method. To date, large numbers of lectins and antibodies have been explored as tools for the detection of glycan epitopes.[14,15] Nevertheless, antibodies and lectins suffer many well-document drawbacks such as low affinity and cross-reactivity.

In recent years, the development bioorthogonal reporter strategy has emerged as a powerful tool for glycan analysis.[16] Typically, bioorthogonal functional groups carried by unnatural monosaccharides analogues are metabolically incorporated into glycans *in vivo*, allowing for further covalent conjugation by bioorthogonal chemical reactions to tag a specific class of glycans (Figure 1).[17,18] The metabolic labeling strategy is based on a covalent modification of target glycans, and therefore, has much better sensitivity than traditional methods such as antibody and lectin binding. The introduction of high-affinity tag such as biotin group in second reaction step makes it possible to globally profile glycan-attached glycoproteins.[19,20] Currently, the most popular two bioorthogonal chemistry groups are azide and alkyne. In addition, many other bioorthogonal groups such as alkene, ketone, nitrile, diazirine have also been explored for use in metabolic labeling.[18] These fundamental studies have made metabolic labeling strategy very popular in glycan analysis. However, higher order glycans that contain disaccharide or trisaccharide motifs cannot be uniquely detected by metabolic labeling method. Moreover, unnatural monosaccharides need to compete with natural substrates while being incorporated into glycans through biosynthetic pathways, resulting in lower labeling efficiencies.

Alternatively, enzymatic bioorthogonal labeling, also known as chemoenzymatic labeling strategy, provides an *in vitro* labeling method to detect structure-defined glycan epitopes.[21] This method takes advantage of the substrate tolerance of recombinant glycosyltransferase to label target epitope with bioorthogonal groups carried by unnatural nucleotide sugars (Figure 1).[22-26] As the glycosyltransferase-mediated reactions and the following bioorthogonal chemical reactions both proceed with high efficiency and specificity, enzymatic bioorthogonal labeling method provides higher sensitivity and selectivity than other analytical methods. More importantly, enzymatic labeling

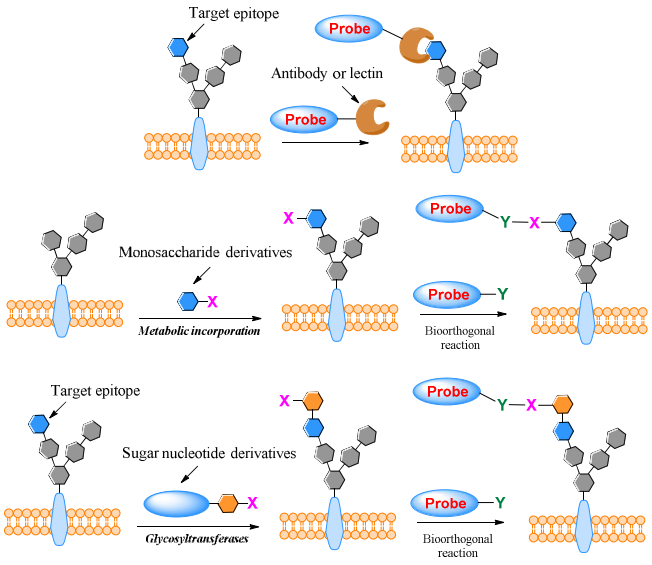


Figure 1. Three common strategies used for glycan analysis: affinity binding strategy, metabolic bioorthogonal labeling strategy, and enzymatic bioorthogonal labeling strategy.

doesn't rely on feeding with metabolic substrates and therefore could be used to label complex samples such as tissue extracts *in vitro*.[27]

To date, the enzymatic bioorthogonal labeling strategy has not proven to be as popular as metabolic labeling for use in glycan analysis. The most important reason is the limited availability of the unnatural nucleotide sugars which functionalized with biorthogonal reactive groups. This has hindered the application and future development of enzymatic bioorthogonal labeling strategy. To address this issue, we describe herein a platform that combines chemical synthesis and enzymatic synthesis to produce unnatural nucleotide sugars modified with bioorthogonal groups. With these unnatural nucleotide sugars in hand, we also evaluated their potentials for use in enzymatic bioorthogonal labeling reactions.

RESULTS AND DISCUSSION

Design of unnatural nucleotide sugars for glycan analysis by enzymatic bioorthogonal labeling. *N*-Acetylgalactosaminyltransferase, *N*-acetylglucosaminyltransferase, and sialyltransferase are the most popular three classes of enzymes that used in enzymatic labeling strategy. CMP-Neu5Ac analogues can be prepared from ManAc or Neu5Ac analogues by using sialic acid aldolase from *E. coli* K12 and CMP-sialic acid synthetase from *Neisseria meningitis*, both of which have extremely relaxed substrate specificity.[28] However, the synthesis of UDP-GalNAc and UDP-GlcNAc derivatives are more difficult. Now only several azido nucleotide sugars are available (**2**, **3**, **17**, and **18** in Figure 2) for enzymatic labeling. In this work, we design to synthesize 12 UDP-GlcNAc analogues and 13 UDP-GalNAc analogues.

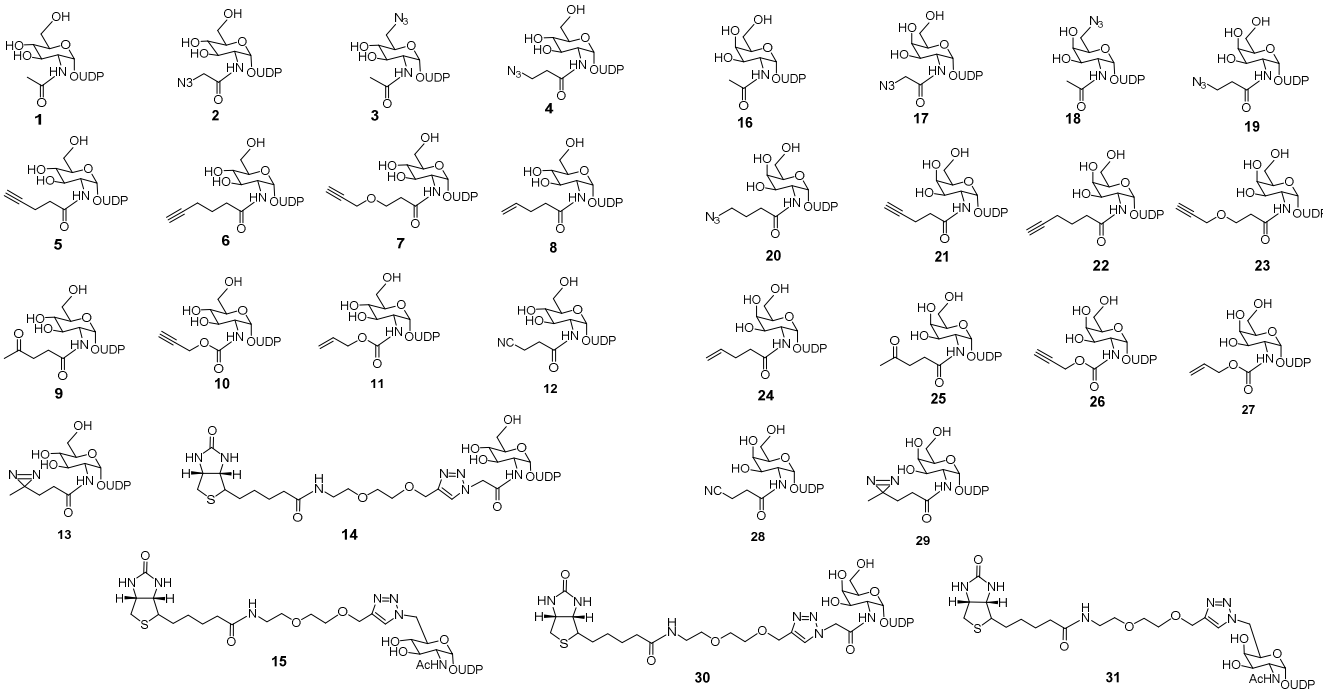
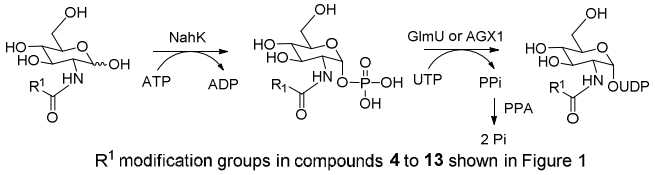


Figure 2. Structures of the designed UDP-GlcNAc and UDP-GalNAc derivatives.

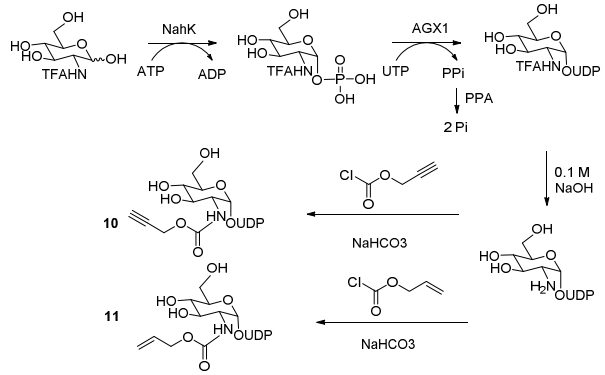
These analogues include most of known bioorthogonal reporter groups that used in metabolic labeling strategy. These reactive groups can be probed by well-established methods such as copper-catalyzed azide-alkyne cycloaddition (azide and alkyne groups),[29] strain promoted alkyne-azide cycloaddition (azide group),[29] Staudinger-Bertozzi ligation (ketone group; **9** and **25**),[30] Pd-catalyzed bioorthogonal elimination reaction (**10**, **11**, **26**, and **27**),[31] Diels-Alder reaction (alkene group; **8**, **11**, **24**, and **27**),[32] Raman reporter strategy (nitrile group; **12** and **18**),[33] photoactivatable crosslinking (diazirine group; **13** and **29**),[34] one-step labeling strategy (biotin group; **14**, **15**, **30** and **31**).[35,36]

Chemoenzymatic synthesis of UDP-GlcNAc derivatives functionalized with bioorthogonal reactive groups. Monosaccharide analogues containing bioorthogonal reactive groups shown in Figure 1 were prepared as described in supporting information. Scheme 1 is a common strategy to prepare UDP-sugars. A prerequisite of this strategy is the availability of a kinase that can phosphorylate monosaccharide to sugar-1-phosphate. NahK (N-Acetylhexosamine 1-Kinase) from *Bifidobacterium* is a kinase that used to phosphorylate GlcNAc and GalNAc and their derivatives.[37] To test whether NahK can recognize our analogues, His-tagged NahK was prepared using *E.coli* expression system and purified by Ni-NTA agarose. Analytical reactions containing monosaccharide analogues, NahK and ATP, were carried in 20 ul reaction system. The reaction was analyzed by both HPLC and TLC. Substrate specificity study showed that all these analogues can be efficiently accepted by NahK (Table S1). Some monosaccharide analogues were found to have higher relative activities than their natural substrates GlcNAc or GalNAc (Table S1). Indeed, all the tested monosaccharides can be clearly converted to corresponding phosphorylated forms after overnight incubation, indicating the excellent feasibility of NahK for large-scale synthesis of the designed compounds.



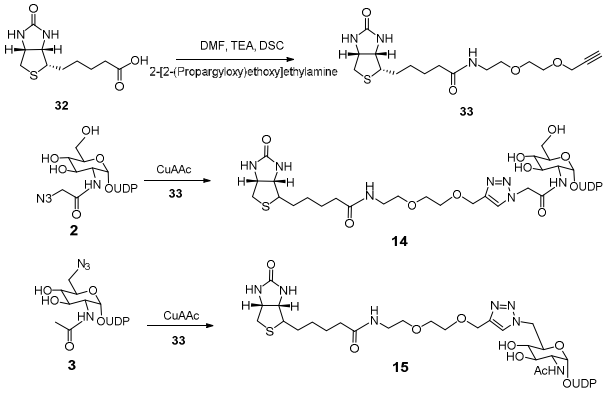
Scheme 1. Chemoenzymatic synthesis of UDP-GlcNAc derivatives **4** to **13**

To prepare UDP-GlcNAc derivatives by scheme 1, we first tested the activity of UDP-*N*-acetylgalactosamine pyrophosphorylase from human (AGX1) and pyrophosphorylase from *Escherichia coli* (GlmU).[38] Analytical reaction was performed in a 20 ul system containing monosaccharide analogues, ATP, UTP, NahK and pyrophosphorylase (AGX1 or GlmU). The reaction was analyzed by both TLC and HPLC. We found AGX1 only tolerate shorter modification group (**4**, **5** and **13**), while GlmU is active to all derivatives (**4** to **13**). We choose GlmU for large-scale synthesis as GlmU has a much higher expression level than AGX1 in *E.coli* system. Preparative-scale synthesis (more than a hundred milligrams) was carried by one-pot multienzyme systems containing monosaccharide analogues (**S4** to **S13**), NahK and GlmU. Inorganic pyrophosphatase PPA was added in reaction to hydrolyze the newly formed inorganic pyrophosphate to improve the overall yields. Products **4**, **8**, **9**, **12** and **13** were purified by silver nitrate precipitation method.[39,40] Silver nitrate precipitation is a newly developed method used for the purification of phosphorylated sugars, by which sugar adenosines (ATP and ADP) can be cleanly precipitated out of aqueous solution while sugar phosphates (with five or more than five carbons) are still in solution.[39,40] This method has widely been used for sugar phosphate purification.[41-43] We noticed that silver nitrate could also cleanly precipitate other phosphorylated nucleotides such as UTP and UDP in aqueous solution, while sugar nucleotides with one or two internal phosphate groups (UDP-sugars, GDP-sugars, and CMP-sugars) were not affected. Therefore, once the above-mentioned reactions finished, excess silver nitrate was added to precipitate ATP, ADP, and UTP. Then, sodium chloride was added to precipitate the redundant silver nitrate. This process can be finished in less than 10 minutes. Then, the final products can be obtained after desalting by P-2 column with high purity. It is worth mentioning **13** is sensitive to light as it contains a diazirine group. The synthetic manipulations require protection from light. Meanwhile, compounds **5**, **6**, **7** and **10** can't be well obtained by silver nitrate precipitation method as they have alkyne group, which can be precipitated by silver ions to a certain extent. Therefore, for the purification of compounds **5**, **6**, **7** and **10**, alkaline phosphatase was added to digest ATP, ADP, and UTP to facilitate further purification by gel filtration. The structures of the isolated products were confirmed by NMR and MS analysis (see supporting information). **4**, **5**, **9** and **12** were obtained in more than 80% isolated yield (Table 1), while the rest derivatives were obtained in relatively lower isolated yields, due to limit efficiency of the second reaction step (a high amount of intermediate phosphorylated sugars was observed on TLC). Prolonging the reaction time or adding more enzyme was not shown to improve the reaction yield.



Scheme 2. Chemoenzymatic synthesis of **10** and **11**

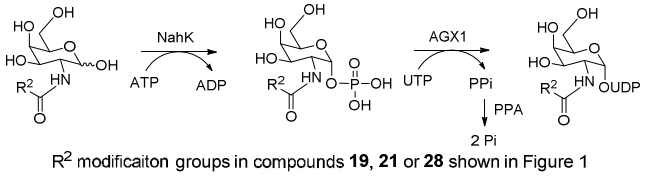
Scheme 2 is an alternative strategy to prepare **10** and **11**. UDP-GlcNTFA was firstly prepared from GlcNTFA by NahK and AGX1. This step gave excellent yield (92 %). Then, the treatment of 0.1 M of sodium hydrate gave a clean conversion of UDP-GlcNTFA to UDP-$GlcNH_2$ in less than two hours. UDP-$GlcNH_2$ was dissolved in water containing sodium bicarbonate, and then propargyl chloroformate, which selectively reacts with amine group, was added slowly to give **10** in quantitative yield. After purification by P-2 column, **10** was obtained in an overall yield of 76% with regards to GlcNTFA. Similarly, incubation of UDP-$GlcNH_2$ with allyl chloroformate gave **11** in total 79% yield. Although more reaction steps are required by Scheme 2 for the synthesis of **10** and **11**, the total isolated yields are still higher than the use of Scheme 1 for the preparation of **10** and **11**.



Scheme 3. Chemoenzymatic synthesis of **14** and **15**

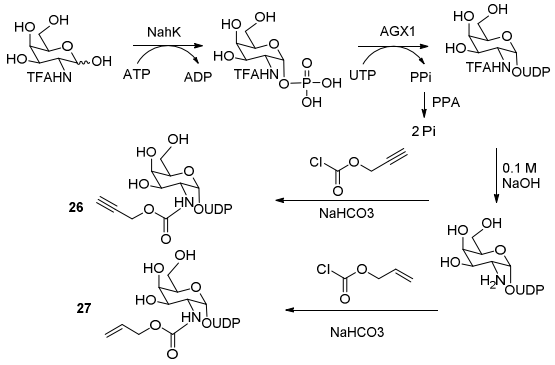
Scheme 3 was used to prepare **14** and **15** from **2** and **3**, respectively. **14** and **15** were conjugated with biotin and can be used for one-step labeling strategy.[36] One-step labeling, whereby a reporter group such as biotin or fluorescence group is directly conjugated with an activated sugar donor for enzymatic labeling, provides not only higher labeling efficiency than the typical two-step labeling but also avoids additional chemical reaction steps needed to introduce a reporter group after the labeling reaction.[35,36] Therefore, it is a more attractive labeling strategy. **2** and **3** were prepared by scheme 1 using NhaK and AGX1 first.[38] Then, **2** and **33** were dissolved in water containing 50% of methanol. Ascorbate, TBTA, and $CusO_4$ were added to initiate the CuAAC ligation. The reaction went to completion in less than 30 minutes with quantitative yield. Product **14** was obtained after P-2 column purification in 82% yield regarding to **2**. Similarly, **15** was prepared from **3** and **33** in 81% yield regarding to **3**. The structures of **14** and **15** were confirmed by NMR analysis (see supporting information).

Chemoenzymatic synthesis of UDP-GalNAc derivatives modified with bioorthogonal reactive groups. To prepare UDP-GalNAc derivatives by scheme 4, we first tested the activity of AGX1 and GlmU at analytical scales as mentioned above. However, only the compounds with short modifications can be prepared by AGX1 (**19**, **21**, and **28**). The rest either can't be prepared by AGX1 at all or only show trace activity (synthetically useless). Although GlmU has been reported as suitable for the preparation of some UDP-GalNAc derivatives such as **17** and **18**,[38] it is unable to recognize our designed structures in Figure 1. Other pyrophosphorylases from *Pyrococcus furiosus* DSM 3638 and *Bifidobacterium longum*[44,45] were also tested to try to synthesize the rest derivatives. However, no activity was observed. Preparative-scale synthesis (more two hundred milligrams) was carried by one-pot multienzyme systems containing monosaccharide analogues (**S19**, **S21**, or **S28**), NahK and AGX1. Once reaction finished, silver nitrate precipitation was used to purify **19** and **28** as described above. **21** was purified by the treatment of the reaction solution with alkaline phosphatase. After desalting by P-2 column, products were obtained in high yields (Table 1). The structures were confirmed by NMR and MS analysis (see supporting information).

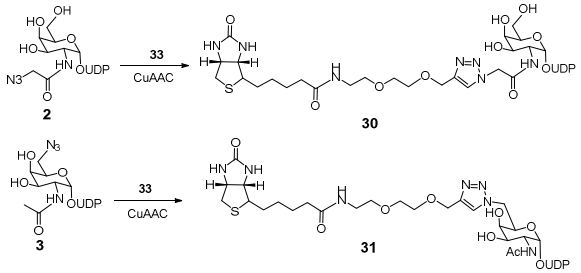


Scheme 4. Chemoenzymatic synthesis of UDP-GalNAc derivatives **19**, **21**, and **28**

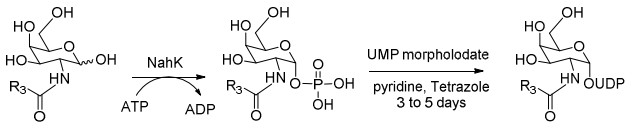
In addition to the preparation of **19**, **21** and **28** by Scheme 4, **26** and **27** were prepared from GalTFA by Scheme 5 as the synthesis of **10** and **11** by Scheme 2. **30** and **31** were prepared from **17** and **18** by CuAAC ligation, respectively (Scheme 6). **17** and **18** were prepared from GlcNAz and 6-N_3-GlcNAc as reported previously.[46] The products **26**, **27, 30** and **31** were obtained in high yields (Table 1) after purification by P-2 column. The structures were confirmed by NMR and MS analysis (see supporting information).



Scheme 5. Chemoenzymatic synthesis of **26** and **27**



Scheme 6. Chemoenzymatic synthesis of **30** and **31**



Scheme 7. Chemoenzymatic synthesis of UDP-GalNAc derivatives **20**, **22**, **23**, **24**, **25** and **29**

For nucleotide sugars that cannot be accessed by enzymatic methods, chemical synthesis is the only means to produce. The most difficult step in the chemical synthesis of nucleotide sugars is to produce sugar phosphate with α or β configuration selectively, which requires complicated protection/deprotection manipulations. Since all the planned monosaccharide analogues can be efficiently phosphorylated by NahK, which produces sugar-α-1-phosphate exclusively, we design Scheme 7 to produce the rest designed sugar nucleotides (**20**, **22**, **23**, **24**, **25**, and **29**). Monosaccharide analogues were incubated with NahK and ATP. Upon reaction completion (no starting material was observed), the silver nitrate precipitation method was employed to purify the phosphorylated sugars (**S30** to **S35**). After desalting by P-2 column, the obtained phosphorylated sugars were confirmed by NMR and MS analysis (see supporting information). In second reaction step, sugar-1-phosphates (**S30** to **S35**) were incubated with UMP morpholodate, which is commercially available, and tetrazole in pyridine solution for 3 to 5 days to produce UDP-sugars.[47] Once the reaction ceased to move forward, the pyridine was removed under vacuum, and the products were dissolved in sodium chloride solution. After purification by P-2 column, products were obtained with yields ranging from 50% to 63% (Table 1). The synthetic manipulations of **29** require protection from light as it has a diazirine group. The products were confirmed by NMR and MS analysis (see supporting information).

Table 1. Chemoenzymatic synthesis of UDP-GlcNAc and UDP-GalNAc derivatives using the strategies shown in schemes 1 to 7

Entry	Starting material[a]	Schemes	Product	Yield[e]
a	**S1**	Scheme 1[b]	**4**	80
b	**S2**	Scheme 1[b]	**5**	84
c	**S3**	Scheme 1[c]	**6**	54
d	**S4**	Scheme 1[c]	**7**	43
e	**S5**	Scheme 1[c]	**8**	47
f	**S6**	Scheme 1[c]	**9**	92
g	**S7**	Scheme 1[c]	**10**	41
h	**S8**	Scheme 1[c]	**11**	39
g1	**S9**	Scheme 2	**10**	76
h1	**S10**	Scheme 2	**11**	79
i	**S11**	Scheme 1[c]	**12**	86
j	**S12**	Scheme 1[b]	**13**	62
k	**2**	Scheme 3	**14**	82
l	**3**	Scheme 3	**15**	81
m	**S15**	Scheme 4[d]	**19**	89
n	**S16**	Scheme 7	**20**	54
o	**S17**	Scheme 4[d]	**21**	77
p	**S18**	Scheme 7	**22**	56
q	**S19**	Scheme 7	**23**	50
r	**S20**	Scheme 7	**24**	63
s	**S21**	Scheme 7	**25**	60
t	**S22**	Scheme 5	**26**	71
u	**S23**	Scheme 5	**27**	78
v	**S24**	Scheme 4[d]	**28**	90
w	**S25**	Scheme 7	**29**	51
x	**17**	Scheme 6	**30**	80
y	**18**	Scheme 6	**31**	86

[a]Compound structures of **2**, **3**, **17**, and **18** were shown in Figure 2, and the rest compound structures were shown in Figure S1. [b]The pyrophosphorylase used in second reaction step can be either AGX1 or GlmU. [c]The pyrophosphorylase used in second reaction step can only be GlmU. [d]The pyrophosphorylase used in second reaction step can only be AGX1. [e]All products were prepared on a preparative (134 mg-306 mg) scale.

Evaluation of the use of UDP-GlcNAc and UDP-GalNAc derivatives in enzymatic bioorthogonal labeling reactions. Having these UDP-GlcNAc and UDP-GalNAc derivatives in hand, we further tested their potential applications in enzymatic bioorthogonal labeling reactions.[21] Figure 3 includes several known glycan epitopes that can be detected by enzymatic bioorthogonal labeling strategy. B3GNT6 and GCNT1 were used to analyze Tn antigen and T-antigen with UDP-GlcNAz, respectively.[48] BgtA, CgtA, and Bovine GalT (Y289L) were used to label the Fucα1,2-Gal epitope, Neu5Acα2,3-Gal epitope, and O-GlcNAc-modified proteins with UDP-GalNAz, respectively.[21] Tn antigen and T antigen are famous Mucin-type O-glycans, which have no expression on normal cells but highly expressed on many cancer cells.[49,50] They have shown great potential for application in immunotherapy and clinical diagnosis. Fucα1,2-Gal and Neu5Acα2,3-Gal are complex glycan epitopes that expressed normally on the cell surface. However, both expression levels can significantly change on many cancer cell lines.[51] In addition, they also play important roles in many other biological processes such as cell-cell and cell-pathology interaction.[52,53] O-GlcNAc post modification is an important monosaccharide modification on proteins in multicellular eukaryotes, involving in protein degradation, nutrient sensing, gene expression, and other essential cell processes.[54]

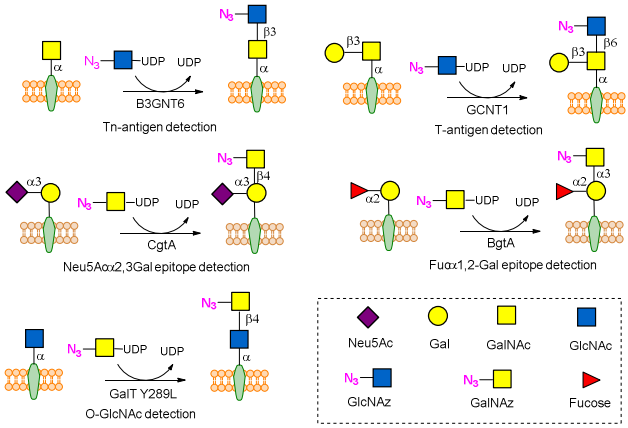


Figure 3. Enzymatic bioorthogonal labeling of Tn antigen, T antigen, Neu5Acα2,3-Gal, and Fuca1,2-Gal, and O-GlcNAc with UDP-GlcNAz and UDP-GalNAz.

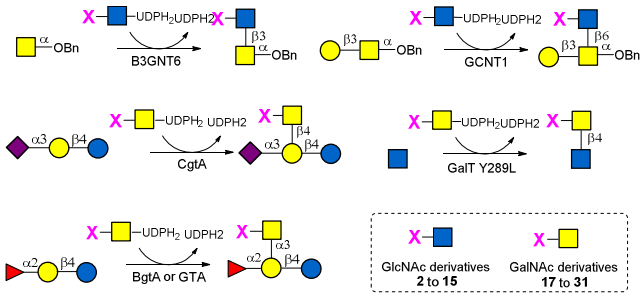


Figure 4. Donor specificity study of B3GNT6, GCNT1, BgtA, GTA, CgtA and Bovine GalT with standard oligosaccharides or monosaccharide.

Therefore, diverse probes (unnatural nucleotide sugars functionalized with bioorthogonal reactive groups) are important tools for the investigation of the detail roles in a variety of physiological and pathological processes.

Human B3GNT6 and GCNT 1 were expressed in baculovirus insect cells. To study the donor specificity of human B3GNT6, GalNAc-O-Bn[55] was incubated with UDP-GlcNAc derivatives (**2** to **15**) in the presence of Mn^{2+}, Tris-HCl, and B3GNT6 (Figure 4). The reaction was analyzed by HPLC. As shown in Table 2, we found some derivatives such as **2**, **4**, and **5** even have better activity than its natural substrate UDP-GlcNAc **1**, indicating their potential use in the diverse labeling of Tn antigen. Similarly, donor specificity of GCNT1 was tested with Galβ1,3-GalNAc-O-Bn[55] (Figure 4) and UDP-GlcNAc derivatives (**2** to **15**). The results indicated that GCNT1 less tolerant of modifications like B3GNT6 (Table 2). Nevertheless, some new substrates such as **3**, **4**, **5**, and **12** can still be accepted by GCNT1.

Table 2. Donor specificity study of B3GNT6 and GCNT1 with UDP-GlcNAc (**1**) and UDP-GlcNAc derivatives (**2** to **15**)

Donor	B3GNT6	GCNT1	Donor	B3GNT6	GCNT1
1	100[a]	100[a]	**9**	2	<1
2	216	25	**10**	17	<1
3	3	26	**11**	8	<1
4	196	22	**12**	27	21
5	155	16	**13**	7	2
6	25	<1	**14**	<1	<1
7	56	<1	**15**	<1	<1
8	<1	5			

[a]the activity with natural substrate UDP-GalNAc 1 was defined as 100. The reaction was carried at 37°C for 30 minutes. See supporting information for experiment details.

BgtA is a bacterial α1,3-*N*-acetylgalactosaminyltransferase from *Helicobacter mustelae*, which specifically recognize Fucα1,2-Gal epitope.[24] We tested the donor substrate specificity of BgtA with Fucα1,2-Gal-β1,4-Glc trisaccharide[24] and UDP-GalNAc derivatives (**17** to **31**). In addition to the known substrates UDP-GalNAc and UDP-GalNAz, only a few derivatives can be accepted by BgtA with relatively lower activity (Table 3). Therefore, we tested another α1,3-*N*-acetylgalactosaminyltransferase GTA from humans,[56] which also recognizes Fucα1,2-Gal epitope specifically. Interestingly, we found GTA has a more relaxed donor specificity than BgtA (Table 3). Many UDP-GalNAc derivatives **17**, **19**, **20**, **21**, **24**, **28** and **29** can be efficiently accepted by GTA, indicating its utility for versatile labeling of Fucα1,2-Gal glycan epitope (further test is undergoing in our lab). More importantly, GTA has much better

Table 3. Donor specificity study of BgtA, GTA, CgtA and GalT with UDP-GalNAc (**16**) and UDP-GalNAc derivatives (**17** to **31**)

Donor	BgtA	GTA	CgtA	GalT
16	100[a]	100[a]	100[a]	100[a]
17	84	120	120	75
18	<1	7	3	<1
19	12	103	38	12
20	7	78	22	8
21	10	53	29	14
22	7	19	21	10
23	<1	4	1	4
24	4	85	8	5
25	6	14	4	15
26	<1	9	86	3
27	<1	11	59	3
28	15	74	43	40
29	2	39	1	2
30	<1	3	2	<1
31	<1	1	1	<1

[a]the activity with natural substrate UDP-GalNAc was defined as 100. The reaction was carried at 37°C for 30 minutes (BgtA, GTA and CgtA tests). For GalT test, the reaction was carried at 20°C for 2 hours. See supporting information for experiment details.

expression level than BgtA in *E.coli* system (~50 mg protein can be obtained from 1 L culture, while only ~5 mg of BgtA can be obtained as the same expression condition). CgtA is a β1,4-*N*-acetylgalactosaminyltransferase from *Campylobacter jejuni*, which recognizes Neu5Acα2,3-Gal epitope.[57] We found it accept UDP-GalNAz as substrate and developed a chemoenzymatic labeling tool to detect cell surface Neu5Acα2,3-Gal glycans.[22] CgtA donor specificity was tested by Neu5Acα2,3-Gal-β1,4-Glc trisaccharide[22] and UDP-GalNAc derivatives **17** to **31**. The results indicated that UDP-GalNAz is the best substrate of CgtA, while other substrates such as **19**, **21**, **26**, **27**, and **28** can also be well recognized by CgtA, also indicating their potentials for versatile labeling of Neu5Acα2,3-Gal glycan epitope. Bovine GalT Y289L is a β1,4-galactosyltransferase, which catalyzes the transfer of galactose from UDP-Gal to GlcNAc.[58] It can also tolerate substitutions at C-2 position such as UDP-GalNAc and UDP-GalNAz for use in O-GlcNAc protein labeling.[26,27] We tested donor specificity of Bovine GalT Y289L using GlcNAc and UDP-GalNAc derivatives **17** to **31**. However, in addition to UDP-GalNAz, only compound **28** can be efficiently recognized by Bovine GalT, indicating its strict donor specificity with UDP-GalNAc derivatives. In addition, we found neither *N*-acetylglucosaminyltransferase nor *N*-acetylgalactosaminyltransferase studied in this work can efficiently recognize probes **14**, **15**, **30**, and **31**, which can be used in one-step labeling strategies. This is because the triazole-linked biotin group is too large to be accepted by these enzymes. Protein engineering to allow for a larger donor may be a potential solution, one that may well be worth pursuing given the benefits of one-step labeling in glycan analysis.

CONCLUSIONS

In summary, we have developed a platform for the efficient preparation of unnatural nucleotide sugars functionalized with diverse bioorthogonal reactive groups by combing of chemical and enzymatic synthesis. A total number of 25 UDP-GlcNAc and UDP-GalNAc derivatives were successfully prepared by this platform. These analogues contain most of the known bioorthogonal reporter groups that were used in metabolic labeling strategy for glycan analysis. In addition, we also evaluated the potential application of these analogues in enzymatic bioorthogonal labeling reactions. Many derivatives could be efficiently accepted by the known glycosyltransferases that have been used in previous enzymatic bioorthogonal labeling systems. These analogues can be used in the versatile labeling of glycan epitopes mentioned above in future studies. We anticipate this work will accelerate the further development and application of enzymatic bioorthogonal labeling strategy for use in glycan analysis.

ASSOCIATED CONTENT

Supporting Information.

The Supporting Information is available free of charge on the at DOI

Experimental details, Supporting Figures, Supporting Table, ^{1}H NMR and ^{13}C NMR spectra,

AUTHOR INFORMATION

Corresponding Authors

*(P.G.W) Email: pwang11@gsu.edu.

*(L.W) Email: lwen2@gsu.edu.

ORCID

Peng George Wang: 0000-0003-3335-6794

Liuqing Wen: 0000-0001-9187-7999

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We thank the National Institutes of Health (U01GM116263) for financial support of this work. We thank Dr. Donald L. Jarvis from University of Wyoming for providing the baculovirus vector used to produce B3GNT6 and GCNT1 for this study. This vector was produced as part of the glycorepository project, which was supported by NIH P41GM103390.

REFERENCES

(1) Varki, A. Sialic Acids in Human Health and Disease. *Trends Mol. Med.* **2008,** *14*, 351-360.

(2) Kato, K.; Ishiwa, A. The Role of Carbohydrates in Infection Strategies of Enteric Pathogens. *Trop Med. Health* **2015,** *43*, 41-52.

(3) Georgi, G.; Bartke, N.; Wiens, F.; Stahl, B. Functional Glycans and Glycoconjugates in Human Milk. *Am. J. Clin. Nutr.* **2013,** *98*, 578S-585S.

(4) Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nat. Rev. Cancer* **2015,** *15*, 540-555.

(5) Magalhaes, A.; Duarte, H. O.; Reis, C. A. Aberrant Glycosylation in Cancer: A Novel Molecular Mechanism Controlling Metastasis. *Cancer Cell* **2017,** *31*, 733-735.

(6) Madsen, C. B.; Petersen, C.; Lavrsen, K.; Harndahl, M.; Buus, S.; Clausen, H.; Pedersen, A. E.; Wandall, H. H. Cancer Associated Aberrant Protein O-Glycosylation Can Modify Antigen Processing and Immune Response. *PLoS One* **2012,** *7*, e50139.

(7) Motteram, J.; Lovegrove, A.; Pirie, E.; Marsh, J.; Devonshire, J.; van de Meene, A.; Hammond-Kosack, K.; Rudd, J. J. Aberrant Protein N-Glycosylation Impacts Upon Infection-Related Growth Transitions of the Haploid Plant-Pathogenic Fungus Mycosphaerella Graminicola. *Mol. Microbiol.* **2011,** *81*, 415-433.

(8) Christiansen, M. N.; Chik, J.; Lee, L.; Anugraham, M.; Abrahams, J. L.; Packer, N. H. Cell Surface Protein Glycosylation in Cancer. *Proteomics* **2014,** *14*, 525-546.

(9) Sharon, N. Lectins: Carbohydrate-Specific Reagents and Biological Recognition Molecules. *J. Biol. Chem.* **2007,** *282*, 2753-2764.

(10) Loureiro, L. R.; Carrascal, M. A.; Barbas, A.; Ramalho, J. S.; Novo, C.; Delannoy, P.; Videira, P. A. Challenges in Antibody Development against Tn and Sialyl-Tn Antigens. *Biomolecules* **2015,** *5*, 1783-1809.

(11) Palmisano, G.; Larsen, M. R.; Packer, N. H.; Thaysen-Andersen, M. Structural Analysis of Glycoprotein Sialylation -Part Ii: Lc-Ms Based Detection. *Rsc Adv.* **2013,** *3*, 22706-22726.

(12) Thaysen-Andersen, M.; Larsen, M. R.; Packer, N. H.; Palmisano, G. Structural Analysis of Glycoprotein Sialylation - Part I: Pre-Lc-Ms Analytical Strategies. *Rsc Adv.* **2013,** *3*, 22683-22705.

(13) Oegema, T. R., Jr.; Cooper, K. M. An Improved Automated Periodate-Resorcinol Method for the Determination of Sialic Acid. *Anal. Biochem.* **1983,** *133*, 233-238.

(14) Dan, X. L.; Liu, W. L.; Ng, T. B. Development and Applications of Lectins as Biological Tools in Biomedical Research. *Med. Res. Rev.* **2016,** *36*, 221-247.

(15) Sterner, E.; Flanagan, N.; Gildersleeve, J. C. Perspectives on Anti-Glycan Antibodies Gleaned from Development of a Community Resource Database. *ACS Chem. Biol.* **2016,** *11*, 1773-1783.

(16) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chem. Int. Ed.* **2009,** *48*, 6974-6998.
(17) Palaniappan, K. K.; Bertozzi, C. R. Chemical Glycoproteomics. *Chem. Rev.* **2016,** *116*, 14277-14306.
(18) Cheng, B.; Xie, R.; Dong, L.; Chen, X. Metabolic Remodeling of Cell-Surface Sialic Acids: Principles, Applications, and Recent Advances. *ChemBioChem* **2016,** *17*, 11-27.
(19) Xiao, H.; Tang, G. X.; Wu, R. Site-Specific Quantification of Surface N-Glycoproteins in Statin-Treated Liver Cells. *Anal. Chem.* **2016,** *88*, 3324-3332.
(20) Smeekens, J. M.; Chen, W.; Wu, R. Mass Spectrometric Analysis of the Cell Surface N-Glycoproteome by Combining Metabolic Labeling and Click Chemistry. *J. Am. Soc. Mass. Spectrom.* **2015,** *26*, 604-614.
(21) Lopez Aguilar, A.; Briard, J. G.; Yang, L.; Ovryn, B.; Macauley, M. S.; Wu, P. Tools for Studying Glycans: Recent Advances in Chemoenzymatic Glycan Labeling. *ACS Chem. Biol.* **2017,** *12*, 611-621.
(22) Wen, L.; Zheng, Y.; Jiang, K.; Zhang, M.; Kondengaden, S. M.; Li, S.; Huang, K.; Li, J.; Song, J.; Wang, P. G. Two-Step Chemoenzymatic Detection of N-Acetylneuraminic Acid-Alpha(2-3)-Galactose Glycans. *J. Am. Chem. Soc.* **2016,** *138*, 11473-11476.
(23) Li, Q.; Li, Z.; Duan, X.; Yi, W. A Tandem Enzymatic Approach for Detecting and Imaging Tumor-Associated Thomsen-Friedenreich Antigen Disaccharide. *J. Am. Chem. Soc.* **2014,** *136*, 12536-12539.
(24) Chaubard, J. L.; Krishnamurthy, C.; Yi, W.; Smith, D. F.; Hsieh-Wilson, L. C. Chemoenzymatic Probes for Detecting and Imaging Fucose-Alpha(1-2)-Galactose Glycan Biomarkers. *J. Am. Chem. Soc.* **2012,** *134*, 4489-4492.
(25) Zheng, T.; Jiang, H.; Gros, M.; del Amo, D. S.; Sundaram, S.; Lauvau, G.; Marlow, F.; Liu, Y.; Stanley, P.; Wu, P. Tracking N-Acetyllactosamine on Cell-Surface Glycans in Vivo. *Angew. Chem. Int. Ed.* **2011,** *50*, 4113-4118.
(26) Khidekel, N.; Arndt, S.; Lamarre-Vincent, N.; Lippert, A.; Poulin-Kerstien, K. G.; Ramakrishnan, B.; Qasba, P. K.; Hsieh-Wilson, L. C. A Chemoenzymatic Approach toward the Rapid and Sensitive Detection of O-Glcnac Posttranslational Modifications. *J. Am. Chem. Soc.* **2003,** *125*, 16162-16163.
(27) Aguilar, A. L.; Hou, X.; Wen, L.; Wang, P. G.; Wu, P. A Chemoenzymatic Histology Method for O-Glcnac Detection. *ChemBiochem* **2017,** *18*, 2416-2421.
(28) Yu, H.; Yu, H.; Karpel, R.; Chen, X. Chemoenzymatic Synthesis of Cmp-Sialic Acid Derivatives by a One-Pot Two-Enzyme System: Comparison of Substrate Flexibility of Three Microbial Cmp-Sialic Acid Synthetases. *Bioorg. Med. Chem.* **2004,** *12*, 6427-6435.
(29) Nwe, K.; Brechbiel, M. W. Growing Applications of "Click Chemistry" for Bioconjugation in Contemporary Biomedical Research. *Cancer Biother. Radiopharm* **2009,** *24*, 289-302.
(30) Saxon, E.; Bertozzi, C. R. Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **2000,** *287*, 2007-2010.
(31) Wang, J.; Cheng, B.; Li, J.; Zhang, Z. Y.; Hong, W. Y.; Chen, X.; Chen, P. R. Chemical Remodeling of Cell-Surface Sialic Acids through a Palladium-Triggered Bioorthogonal Elimination Reaction. *Angew. Chem. Int. Ed.* **2015,** *54*, 5364-5368.
(32) Niederwieser, A.; Spate, A. K.; Nguyen, L. D.; Jungst, C.; Reutter, W.; Wittmann, V. Two-Color Glycan Labeling of Live Cells by a Combination of Diels-Alder and Click Chemistry. *Angew. Chem. Int. Ed.* **2013,** *52*, 4265-4268.
(33) Lin, L.; Tian, X.; Hong, S.; Dai, P.; You, Q.; Wang, R.; Feng, L.; Xie, C.; Tian, Z. Q.; Chen, X. A Bioorthogonal Raman Reporter Strategy for Sers Detection of Glycans on Live Cells. *Angew. Chem. Int. Ed.* **2013,** *52*, 7266-7271.
(34) Pham, N. D.; Parker, R. B.; Kohler, J. J. Photocrosslinking Approaches to Interactome Mapping. *Curr. Opin. Chem. Biol.* **2013,** *17*, 90-101.
(35) Sun, T.; Yu, S. H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G. J. One-Step Selective Exoenzymatic Labeling (Seel) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016,** *138*, 11575-11582.
(36) Wen, L.; Liu, D.; Zheng, Y.; Huang, K.; Cao, X.; Song, J.; Wang, P. G. A One-Step Chemoenzymatic Labeling Strategy for Probing Sialylated Thomsen-Friedenreich Antigen. *ACS Cent. Sci.* **2018,** *4*, 451-457.
(37) Li, Y.; Yu, H.; Chen, Y.; Lau, K.; Cai, L.; Cao, H.; Tiwari, V. K.; Qu, J.; Thon, V.; Wang, P. G.; Chen, X. Substrate Promiscuity of N-Acetylhexosamine 1-Kinases. *Molecules* **2011,** *16*, 6396-6407.
(38) Guan, W.; Cai, L.; Wang, P. G. Highly Efficient Synthesis of Udp-Galnac/Glcnac Analogues with Promiscuous Recombinant Human Udp-Galnac Pyrophosphorylase Agx1. *Chemistry* **2010,** *16*, 13343-13345.
(39) Wen, L.; Huang, K.; Wei, M.; Meisner, J.; Liu, Y.; Garner, K.; Zang, L.; Wang, X.; Li, X.; Fang, J.; Zhang, H.; Wang, P. G. Facile Enzymatic Synthesis of Ketoses. *A Angew. Chem. Int. Ed.* **2015,** *54*, 12654-12658.
(40) Wen, L. Q.; Huang, K.; Liu, Y. P.; Wang, P. G. Facile Enzymatic Synthesis of Phosphorylated Ketopentoses. *ACS Catal.* **2016,** *6*, 1649-1654.
(41) Wen, L.; Zang, L.; Huang, K.; Li, S.; Wang, R.; Wang, P. G. Efficient Enzymatic Synthesis of L-Rhamnulose and L-Fuculose. *Bioorg. Med. Chem. Lett.* **2016,** *26*, 969-972.
(42) Wen, L.; Huang, K.; Zheng, Y.; Liu, Y.; Zhu, H.; Wang, P. G. A Two-Step Strategy for the Preparation of 6-Deoxy-L-Sorbose. *Bioorg. Med. Chem. Lett.* **2016**. 26, 4358-4361
(43) Wen, L. Q.; Huang, K.; Zheng, Y.; Fang, J. Q.; Kondengaden, S. M.; Wang, P. G. Two-Step Enzymatic Synthesis of 6-Deoxy-L-Psicose. *Tetrahedron Lett.* **2016,** *57*, 3819-3822.
(44) Mizanur, R. M.; Zea, C. J.; Pohl, N. L. Unusually Broad Substrate Tolerance of a Heat-Stable Archaeal Sugar Nucleotidyltransferase for the Synthesis of Sugar Nucleotides. *J. Am. Chem. Soc.* **2004,** *126*, 15993-15998.
(45) Muthana, M. M.; Qu, J. Y.; Li, Y. H.; Zhang, L.; Yu, H.; Ding, L.; Malekan, H.; Chen, X. Efficient One-Pot Multienzyme Synthesis of Udp-Sugars Using a Promiscuous Udp-Sugar Pyrophosphorylase from Bifidobacterium Longum (Blusp). *Chem. Commun.* **2012,** *48*, 2728-2730.
(46) Chen, Y.; Thon, V.; Li, Y. H.; Yu, H.; Ding, L.; Lau, K.; Qu, J. Y.; Hie, L.; Chen, X. One-Pot Three-Enzyme Synthesis of Udp-Glcnac Derivatives. *Chem. Commun.* **2011,** *47*, 10815-10817.
(47) Schultz, V. L.; Zhang, X.; Linkens, K.; Rimel, J.; Green, D. E.; DeAngelis, P. L.; Linhardt, R. J. Chemoenzymatic Synthesis of 4-Fluoro-N-Acetylhexosamine Uridine Diphosphate Donors: Chain Terminators in Glycosaminoglycan Synthesis. *J. Org. Chem.* **2017,** *82*, 2243-2248.
(48) Wu, Z. L.; Huang, X.; Burton, A. J.; Swift, K. A. Probing Sialoglycans on Fetal Bovine Fetuin with Azido-Sugars Using Glycosyltransferases. *Glycobiology* **2016,** *26*, 329-334.
(49) Guzman-Aranguez, A.; Argueso, P. Structure and Biological Roles of Mucin-Type O-Glycans at the Ocular Surface. *Ocular Surface* **2010,** *8*, 8-17.
(50) Bergstrom, K. S. B.; Xia, L. J. Mucin-Type O-Glycans and Their Roles in Intestinal Homeostasis. *Glycobiology* **2013,** *23*, 1026-1037.
(51) Cheever, M. A.; Allison, J. P.; Ferris, A. S.; Finn, O. J.; Hastings, B. M.; Hecht, T. T.; Mellman, I.; Prindiville, S. A.; Viner, J. L.; Weiner, L. M.; Matrisian, L. M. The Prioritization of Cancer Antigens: A National Cancer Institute Pilot Project for the Acceleration of Translational Research. *Clin. Cancer Res.* **2009,** *15*, 5323-5337.
(52) Danishefsky, S. J.; Shue, Y. K.; Chang, M. N.; Wong, C. H. Development of Globo-H Cancer Vaccine. *Acc. Chem. Res.* **2015,** *48*, 643-652.
(53) You, Y.; Cao, Y.; Guo, S.; Xu, J.; Li, Z.; Wang, J.; Xue, C. Purification and Identification of A 2–3 Linked Sialoglycoprotein and A 2–6 Linked Sialoglycoprotein in Edible Bird's Nest. *Eur. Food Res. Technol.* **2015,** *240*, 389-397.
(54) Li, J.; Wang, J.; Wen, L.; Zhu, H.; Li, S.; Huang, K.; Jiang, K.; Li, X.; Ma, C.; Qu, J.; Parameswaran, A.; Song, J.; Zhao, W.; Wang, P. G. An Oga-Resistant Probe Allows Specific Visualization and Accurate Identification of O-Glcnac-Modified Proteins in Cells. *ACS Chem. Biol.* **2016,** *11*, 3002-3006.
(55) Gao, Y.; Aryal, R. P.; Ju, T.; Cummings, R. D.; Gahlay, G.; Jarvis, D. L.; Matta, K. L.; Vlahakis, J. Z.; Szarek, W. A.; Brockhausen, I. Acceptor Specificities and Selective Inhibition of Recombinant Human Gal- and Glcnac-Transferases That Synthesize Core Structures 1, 2, 3 and 4 of O-Glycans. *Biochim. Biophys. Acta.* **2013,** *1830*, 4274-4281.
(56) Alfaro, J. A.; Zheng, R. B.; Persson, M.; Letts, J. A.; Polakowski, R.; Bai, Y.; Borisova, S. N.; Seto, N. O.; Lowary, T. L.; Palcic, M. M.; Evans, S. V. Abo(H) Blood Group a and B Glycosyltransferases Recognize Substrate Via Specific Conformational Changes. *J. Biol. Chem.* **2008,** *283*, 10097-10108.
(57) Gilbert, M.; Karwaski, M. F.; Bernatchez, S.; Young, N. M.; Taboada, E.; Michniewicz, J.; Cunningham, A. M.; Wakarchuk, W. W. The Genetic Bases for the Variation in the Lipo-Oligosaccharide of the Mucosal Pathogen, Campylobacter Jejuni. Biosynthesis of Sialylated Ganglioside Mimics in the Core Oligosaccharide. *J. Biol. Chem.* **2002,** *277*, 327-337.

(58) Boeggeman, E. E.; Balaji, P. V.; Sethi, N.; Masibay, A. S.; Qasba, P. K. Expression of Deletion Constructs of Bovine Beta-1,4-Galactosyltransferase in Escherichia Coli: Importance of Cys134 for Its Activity. *Protein Eng.* **1993,** *6*, 779-785.