

# Expression of Human $\beta$ 3GalT5–1 in Insect Cells as Active Glycoforms for the Efficient Synthesis of Cancer-Associated Globo-Series Glycans

Chih-Chuan Kung,<sup>§</sup> Jennifer M. Lo,<sup>§</sup> Kuo-Shiang Liao, Chung-Yi Wu, Li-Chun Cheng, Cinya Chung, Tsui-Ling Hsu, Che Ma,\* and Chi-Huey Wong\*



Cite This: *J. Am. Chem. Soc.* 2025, 147, 10864–10874



Read Online

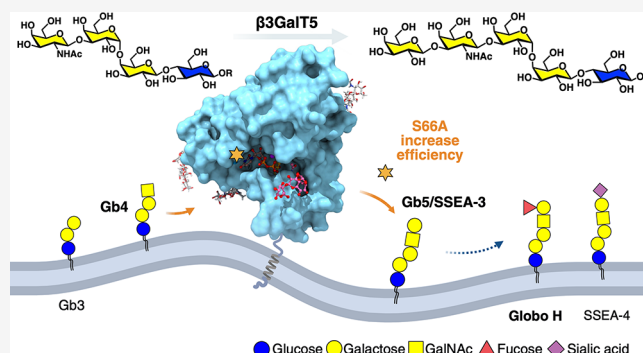
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** The globo-series glycosphingolipids (GSLs) are unique glycolipids exclusively expressed on the cell surface of various types of cancer and have been used as targets for the development of cancer vaccines and therapeutics. A practical enzymatic method has been developed for the synthesis of globo-series glycans, where the conversion of Gb4 to Gb5 (SSEA-3) glycan based on the microbial galactosyltransferase LgtD is relatively inefficient compared to other steps. To improve the efficiency, we explored the two human galactosyltransferase ( $\beta$ 3GalT5) isozymes in cancer cells for this reaction and found that isozyme 1 ( $\beta$ 3GalT5–1) is more active than isozyme 2 ( $\beta$ 3GalT5–2). We then identified a common soluble domain of the two  $\beta$ 3GalT5 isozymes as a candidate and evaluated the activity and substrate specificity of the glycosylated and nonglycosylated glycoforms. The glycoforms expressed in Sf9 cells were selected, and a site-specific alanine scan was performed to identify S66A  $\beta$ 3GalT5 variant with 10-fold more efficiency than LgtD for the synthesis of globo-series glycans. The X-ray structure of  $\beta$ 3GalT5–1 was determined for molecular modeling, and the result together with kinetic data were used to rationalize the improvement in catalysis.



## INTRODUCTION

The globo-series glycosphingolipids (GSLs), including stage-specific embryonic antigen-3 (SSEA-3), SSEA-4, and Globo-H, are often found with irregular expression patterns on the cell surface of various types of cancer<sup>1,2</sup> to promote cancer progression.<sup>3–5</sup> The expression levels of globo-series glycans in clinical samples correlate with tumor metastasis and progression, as well as poor survival.<sup>6–12</sup> These tumor-associated carbohydrate antigens<sup>13</sup> have been considered promising targets for the development of new therapies, especially cancer vaccines.<sup>6–22</sup>

In the biosynthesis of globo-series GSLs, the enzyme  $\beta$ 1,3-galactosyltransferase 5 ( $\beta$ 3GalT5) expressed in cancer cells catalyzes the rate-limiting conversion of Gb4 to Gb5 (SSEA-3), which is subsequently transformed to Globo-H and SSEA-4 by fucosyltransferase 2 (FUT2) and sialyltransferase (ST3Gal1), respectively.  $\beta$ 3GalT5 exists as two distinct isozymes ( $\beta$ 3GalT5–1 and  $\beta$ 3GalT5–2) with 310 amino acids for  $\beta$ 3GalT5–1 and an additional four amino acids in the N-terminal of  $\beta$ 3GalT5–2 (Figure 1). Knockdown of both isozymes simultaneously in MDA-MB-231 breast cancer cells using shRNA was shown to inhibit proliferation and induce cancer cell apoptosis while there was no effect on normal cells.<sup>19</sup> However, the expression of these two isozymes in

normal and cancer cells and their corresponding functions remain unclear. Nevertheless, the cancer vaccine with Globo-H glycan conjugated to keyhole limpet hemocyanin (KLH) and adjuvanted with QS-21 has been advanced to phase 3 clinical trials for the treatment of Globo-H positive patients with early stage triple-negative breast cancer (NCT03562637).<sup>23,24</sup> Another cancer vaccine with Globo-H glycan conjugated to Crim197 and adjuvanted with C34, a glycolipid adjuvant designed to induce class switch, has been advanced to phase 2 trials for the treatment of Globo-H positive patients with tyrosine kinase inhibitor (TKI)-resistant nonsmall cell lung cancer (NCT05442060) and esophageal cancer (NCT05376423).

To enable the late-stage clinical study, we developed a practical enzymatic method coupled with sugar nucleotide regeneration for the synthesis of Globo-H glycan.<sup>25</sup> However,

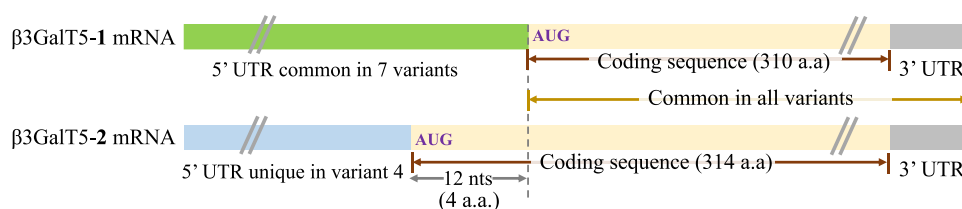
Received: August 26, 2024

Revised: January 13, 2025

Accepted: February 5, 2025

Published: March 25, 2025





**Figure 1.** RNA transcripts of  $\beta 3\text{GalT5-1}$  and  $\beta 3\text{GalT5-2}$ . The yellow boxes represent the full-length translational sequence of  $\beta 3\text{GalT5-1}$  (M1-V310) including the soluble domain (N29–V310), and an additional 4 amino acids in the N-terminal of  $\beta 3\text{GalT5-2}$  (314 amino acids). The green box with 140 nucleotides before the start codon AUG is a common sequence of  $\beta 3\text{GalT5-1}$  in the seven of eight transcripts. The blue box represents the unique sequence in one variant (variant 4) for  $\beta 3\text{GalT5-2}$ .

the microbial enzyme  $\beta 1,3\text{-N-acetylgalactosaminyltransferase}$  (LgtD) used for the synthesis of Gb5 glycan from Gb4 glycan was not efficient. In this study, we explored the synthetic utility of human  $\beta 3\text{GalT5}$  to improve the process. Since  $\beta 3\text{GalT5}$  is a membrane-bound glycoprotein and exists as two isozymes, we first designed specific probes and dicer-substrate siRNA to knockdown specific isozymes in breast cancer cells and found that the expression level and activity of  $\beta 3\text{GalT5-1}$  were higher than those of  $\beta 3\text{GalT5-2}$ . We also investigated the impact of glycosylation on protein folding and activity and found that the soluble domain of  $\beta 3\text{GalT5-1}$  (residues N29–V310) expressed in insect cells is an efficient catalyst for synthesis. In addition, the X-ray structure and substrate specificity of soluble  $\beta 3\text{GalT5-1}$  were determined, and a site-specific alanine scan was performed to identify the S66A  $\beta 3\text{GalT5-1}$  variant with at least 10-fold improvement in activity for the efficient synthesis of globo-series glycans

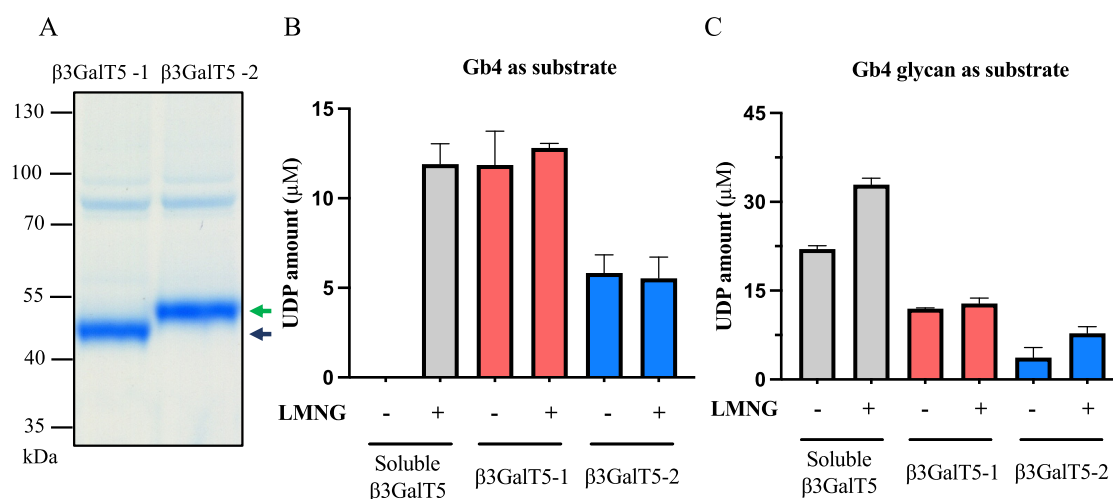
## RESULTS AND DISCUSSION

**Expression of  $\beta 3\text{GalT5-1}$  and  $\beta 3\text{GalT5-2}$  in Normal and Cancer Cells.** Comparing the two isozymes of human  $\beta 3\text{GalT5}$  (Figure 1),  $\beta 3\text{GalT5-1}$  is composed of 310 amino acids, residues 1–7 located in the cytosol, residues 8–28 in the transmembrane domain, and residues 29–310 in the soluble Golgi luminal domain.  $\beta 3\text{GalT5-2}$  contains the same 310-amino acid sequence as  $\beta 3\text{GalT5-1}$  with an additional four amino acids at the N-terminal. At the RNA level (Figure S1), eight variants of human  $\beta 3\text{GalT5}$  exist, seven of which encode the  $\beta 3\text{GalT5-1}$  protein and one variant (variant 4, NM\_033172.3) encoding the  $\beta 3\text{GalT5-2}$  protein. By aligning all transcript variants, we found that except variant 4, the seven variants encoding the  $\beta 3\text{GalT5-1}$  protein have the same 140 nucleotide sequences in the 5' UTR of  $\beta 3\text{GalT5-1}$ . Therefore, we designed specific PCR TaqMan probes to detect the expression of  $\beta 3\text{GalT5-1}$  and  $\beta 3\text{GalT5-2}$  (Figure 1) in different cells including breast, colon, and lung cancer cells. We also employed droplet-digital PCR (ddPCR) to quantify the RNA levels of  $\beta 3\text{GalT5}$  isozymes (Figure S1B) and found that cancer cells contained higher copy numbers than normal cells and the lung cancer cell line CL-5 with greater progression ability than CL-0 exhibited a higher level of  $\beta 3\text{GalT5-1}$  expression. Conversely,  $\beta 3\text{GalT5-2}$  was expressed at extremely low levels in most normal and cancer cells, with slightly higher expression in the breast cancer cell line MDA-MB231 and colon cancer cell lines SW1116 and colo205 (Figure S1C).

**Specific Gene Suppression of  $\beta 3\text{GalT5}$  Isozymes in Breast Cancer Cells.** To investigate the functions of  $\beta 3\text{GalT5-1}$  and  $\beta 3\text{GalT5-2}$  in breast cancer cells, we first designed and screened 11 dicer-substrate siRNAs (dsiRNAs)

based on the common 5' UTR of  $\beta 3\text{GalT5}$ . To specifically target  $\beta 3\text{GalT5-1}$  or  $\beta 3\text{GalT5-2}$  in breast cancer cells, we chose MDA-MB231 as a model because both  $\beta 3\text{GalT5-1}$  and  $\beta 3\text{GalT5-2}$  are expressed in this cell line (Figure S1D). Then, we evaluated the knockdown efficiency of  $\beta 3\text{GalT5-1}$  among 11 dsiRNA in MDA-MB-231 cells. Of all knockdown studies 48 h after electroporation, #4 dsiRNA was the most efficient in inhibiting the expression of  $\beta 3\text{GalT5-1}$  (Figure S2A). In addition, we evaluated the knockdown efficiency at different time points; the percentage of  $\beta 3\text{GalT5-1}$  knockdown within 84 h was about 70% and continued to increase until 144 h (Figure S2B). Next, we analyzed the cell surface levels of SSEA-3, SSEA-4, and Globo-H by flow cytometry at different time points after the knockdown of  $\beta 3\text{GalT5-1}$  and found that the expression of SSEA-3 and Globo-H was significantly reduced in MDA-MB231 cells, while SSEA-4 expression showed little change from 48 to 96 h (Figure S2C), perhaps due to the remaining 20% activity of  $\beta 3\text{GalT5-1}$  and  $\beta 3\text{GalT5-2}$ . We also examined the  $\beta 3\text{GalT5-1}$ -catalyzed glycosylation in MCF-7 cells due to the higher expression of  $\beta 3\text{GalT5-1}$ , SSEA-3, and Globo-H and lower expression of  $\beta 3\text{GalT5-2}$  in this cell line compared to MDA-MB231 (Figures S1C and S2D). We found an increased expression of Gb4 and a decreased expression of Globo-H in MCF-7 cells with  $\beta 3\text{GalT5-1}$  knockdown as measured by mass spectrometry (MS) analysis (Figure S2E), while the surface levels of SSEA-3, SSEA-4, and Globo-H decreased based on flow cytometry analysis (Figure S2D). For the O-glycans, it is known that  $\beta 3\text{GalT5}$  is involved in core 3 and core 4 extension. After the knockdown of  $\beta 3\text{GalT5-1}$ , the core 1-related glycans decreased while the core 2- and 3-related glycans increased, probably due to the disruption of core 3 extension (Figure S2F). This phenomenon is similar to the inhibition of breast cancer development in mice caused by the absence of core 1-derived mucin-type O-glycosylation.<sup>26,27</sup> Additionally, core 3 O-glycan structures are known to decrease in cancer cells.<sup>28,29</sup> However, recent findings demonstrated that, following the knockout of  $\beta 3\text{GalT5}$  in mice, the glycosylation levels of core 1 and 2 increased, while that of core 3 and 4 decreased in colonic MUC2.<sup>30</sup> These contrasting results may be attributed to the differences in cell types between normal and cancerous conditions.

In addition, the percentage of mannose-type N-glycan decreased, while the complexed-type glycans including biantennary, triantennary, and tetra-antennary N-glycans increased (Figure S2G); a similar change was found in previous studies.<sup>31,32</sup> However, the knockdown of  $\beta 3\text{GalT5-2}$ , which was expressed at an extremely low level in MCF-7, showed no effect on SSEA-3/SSEA-4/Globo-H expression as observed in flow cytometry.



**Figure 2.** Purification and activity of  $\beta$ 3GalT5 isozymes from HEK293 cells. A. Two membrane-bound  $\beta$ 3GalT5 isozymes expressed in HEK293 cells were extracted by LMNG and further purified by anti-Flag beads. The blue and green arrows indicated  $\beta$ 3GalT5-1 and  $\beta$ 3GalT5-2, respectively. B, C. Enzymatic activities of membrane-bound  $\beta$ 3GalT5-1 and  $\beta$ 3GalT5-2, and the common soluble domain of both isozymes (soluble  $\beta$ 3GalT5, residues N29–V310 of  $\beta$ 3GalT5-1) using Gb4 and Gb4 glycan as substrates in the presence or absence of 1% LMNG. The results are presented as the mean  $\pm$  standard deviation (SD) of three biological replicates.

**Purification of Membrane-Bound  $\beta$ 3GalT5 Isozymes with Detergent LMNG.** To evaluate the difference in activity of the two isozymes involved in the synthesis of Gb5, two isozymes with a C-terminal Flag tag were expressed in HEK293 cells. To reduce the complexity during purification, organelles, including the Golgi apparatus where  $\beta$ 3GalT5 localized, were specifically extracted by detergents and purified. The cell membrane and organelle fraction were separated by digitonin,<sup>33</sup> and the organelle fractions were further screened with several detergents to optimize the extraction efficiency and the activity of the two isozymes (Figure S3). Of the 20 detergents screened, CHAPS (3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate) and LMNG (lauryl maltose neopentyl glycol) showed better results, and LMNG showed higher extraction efficiency (Figure 2A). We also expressed the soluble domain of  $\beta$ 3GalT5-1 (residues N29–V310, same sequence in  $\beta$ 3GalT5-2) in HEK293 cells and found that it exhibited higher activity compared to membrane-bound  $\beta$ 3GalT5-1 or  $\beta$ 3GalT5-2 using the glycan of Gb4 as a substrate (Figure 2C). However, when the soluble  $\beta$ 3GalT5-1 was incubated with Gb4 glycolipid and UDP-galactose in the absence of LMNG, no activity was observed, and an addition of a small amount of LMNG (1%) was required to have the enzymatic activity, perhaps due to the presence of LMNG to facilitate the orientation of Gb4 glycolipid as an acceptor for  $\beta$ 3GalT5. Interestingly,  $\beta$ 3GalT5-1 showed a 2-fold increase in activity compared to  $\beta$ 3GalT5-2, regardless of whether Gb4 or Gb4 glycan was used as a substrate (Figure 2B and C). Based on these results, we decided to develop soluble  $\beta$ 3GalT5-1 as a catalyst for the synthesis of SSEA-3 glycan.

**Specificity of Soluble  $\beta$ 3GalT5-1 for Acceptor and Donor Substrates.** Knocking down  $\beta$ 3GalT5-1 in MCF-7 cells caused a major effect on the biosynthesis of globo-series glycans and O-glycans and a minor effect on N-glycans. To further explore the substrate specificity of  $\beta$ 3GalT5-1, we investigated the activity of the soluble domain of  $\beta$ 3GalT5-1 expressed in HEK293 toward disaccharides with different linkages and various types of sugars at the nonreducing end. It was observed that the disaccharides with terminal galactose

(Gal), especially in  $\beta$ -1,3 linkage, were excellent substrates and those with terminal N-acetylglucosamine (GlcNAc) or N-acetylgalactosamine (GalNAc) were generally good substrates (Table 1). The glycan parts of both Gb4 and Lc3 were previously reported as native substrates for soluble  $\beta$ 3GalT5.<sup>34</sup> However, the enzyme toward the natural substrate Gb4 exhibited a low activity, which may be attributed to the presence of the ceramide tail that affects the orientation and accessibility of the glycan to  $\beta$ 3GalT5-1 under the assay condition. This problem was mitigated with the addition of detergent LMNG to improve the accessibility of Gb4 glycan to  $\beta$ 3GalT5 (Table 1). In the case of O-linked core structures, both type 3 and type 4 core structures were reported as substrates for human  $\beta$ 3GalT5.<sup>35–37</sup> To investigate the galactosylation preference on core 4 structures, we conducted enzymatic reactions using varying ratios of UDP-Gal to core 4 as the substrate, followed by analysis via LC-MS/MS. The results show that  $\beta$ 3GalT5 prefers the  $\beta$ 1,3-arm of the O-glycan on the core 4 structure, achieving a conversion rate of 65.4% when the molar ratio of core 4 to UDP-Gal is 1:1. However, as the amount of UDP-Gal increases (UDP-Gal: core 4 = 2:1 and 4:1), galactosylation of both the  $\beta$ 1,3- and  $\beta$ 1,6-arms occurred, with conversion rates of 33.52 and 61.22%, respectively, compared to 62.72 and 38.12% when only a single Gal is added to the  $\beta$ 1,3-arm. With an excess of UDP-Gal (UDP-Gal: core 4 = 8:1 and 20:1), more than 90% of the products showed galactosylation on both the  $\beta$ 1,3-arm and  $\beta$ 1,6-arm (Figure S4B and Table S2). In addition, the type 1 LacNAc glycan GlcNAc- $\beta$ 1-3-Gal- $\beta$ 1-3-GlcNAc was also a good acceptor substrate. Regarding donor specificity, soluble  $\beta$ 3GalT5-1 exhibited a preference for UDP-Gal compared to other UDP-sugars using Gb4 glycan as an acceptor (Table 2). We also observed that UDP-Gal-6-aldehyde and UDP-Gal-6-azide were accepted at rates of 78 and 18%, respectively, compared to UDP-Gal (Table 2). The LC-MS results corroborated these findings, showing similar relative activity levels.

**Human  $\beta$ 3GalT5 Expressed in HEK293 or Insect Cells Is More Efficient than the Microbial LgtD in Catalyzing the Synthesis of SSEA-3 Glycan.** In the chemo-enzymatic

Table 1. Substrate Specificity of Soluble  $\beta$ 3GalT5-1 toward Disaccharides, Glycopeptides, and Glycolipids

Substrates	Symbol	Relative activity (%) <sup>a</sup>	Yield (%) <sup>a</sup>
GlcNAc and GalNAc as terminal end			
GlcNAc- $\beta$ 1,3-Gal-OMe		208 ± 4	100
GlcNAc- $\beta$ 1,3-Man		193 ± 1	100
GlcNAc- $\beta$ 1,2-Man		191 ± 8	100
GlcNAc- $\beta$ 1,4-GlcNAc- $\beta$ -OBn		208 ± 6	100
GlcNAc- $\alpha$ 1,4-Gal		0	N.D.
GlcNAc- $\beta$ 1,3-GalNAc		200 ± 2	100
GalNAc- $\beta$ 1,3-Gal		198 ± 6	100
GalNAc- $\beta$ 1,4-Gal-OMe		123 ± 2	100
Galactose and Mannose as terminal end			
Gal- $\beta$ 1,6-GalNAc		19 ± 17	N.D.
Gal- $\beta$ 1,3-GalNAc		52 ± 2	N.D.
Gal- $\beta$ 1,3-GalNAc- $\alpha$ -OMe		71 ± 3	21.6
Gal- $\beta$ 1,3-GalNAc- $\alpha$ -OBn		12 ± 1	28.5
$\beta$ 1-3Galactobiose		48 ± 3	N.D.
Gal- $\beta$ 1-4-GalNAc		8 ± 1	N.D.
Lactose		0	N.D.
Gal- $\beta$ 1-6-Gal		25 ± 1	N.D.
Man- $\alpha$ 1,6-Man		54 ± 6	74.0
Man- $\beta$ 1,6-Man		212 ± 8	100
Glycolipid and its derivatives			
Gb4 glycan		208 ± 10	95.2
Gb4		100	88.5
Gb4 glycan-C5Cl		206 ± 5	100
isoGb4 glycan-C5Cl		206 ± 3	100
Lc3-glycan		208 ± 8	100
O-glycan and its derivatives			
Tn antigen, GalNAc-Serine		0	N.D.
T antigen, Core 1 Gal- $\beta$ 1,3-GalNAc- $\alpha$ -Thr		0	11.9
Core 2, Gal- $\beta$ 1,3[GlcNAc $\beta$ 1,6]-GalNAc- $\alpha$ -Thr		46 ± 2	N.D.
Core 3, GlcNAc- $\beta$ 1,3-GalNAc- $\alpha$ -Thr		206 ± 2	64.3
Core 4, GlcNAc- $\beta$ 1,3-[GlcNAc- $\beta$ 1,6]-GalNAc- $\alpha$ -Thr		187 ± 6	99.9
N-glycan and its derivatives			
Agalacto-SGP		25 ± 17	50.9
Man- $\alpha$ 1-3-(Man- $\alpha$ 1-6)-Man		0	4.9
Type I LacNAc		212 ± 4	100

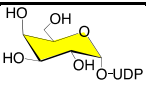
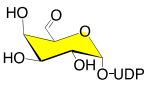
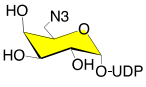
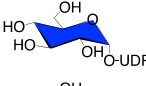
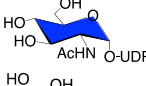
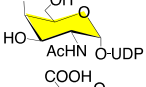
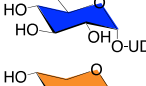
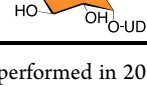
<sup>a</sup>\* Based on natural substrate Gb4 as a reference, all reactions were performed in 20 mM HEPES (pH 7.4) buffer containing 0.1 mM MnCl<sub>2</sub>, 0.1 mM UDP-Gal, 5 mM substrate, and 5  $\mu$ M soluble  $\beta$ 3GalT5-1 expressed from HEK293 cells. isoGb4 glycan: GalNAc- $\beta$ 1,3-Gal- $\alpha$ 1,3-Gal- $\beta$ 1-4-Glc, Lc3 glycan: GlcNAc- $\beta$ 1,3-Gal- $\beta$ 1,4-Glc, and type 1 LacNAc: GlcNAc- $\beta$ 1,3-Gal- $\beta$ 1,3-GlcNAc-O-nitrophenol. For the carbohydrate symbol, yellow circle: galactose (Gal), blue circle: glucose (Glc), green circle: mannose (Man), yellow square: GalNAc, and blue square: GlcNAc. The linker symbol stand for S: serine, T: threonine, K: lysine, V: valine, A: alanine, N: asparagine, Me: methyl group, Bn: benzyl group, and Np: nitrophenyl group. The results are presented as the mean  $\pm$  SD of three biological replicates. \*\* All reactions were carried out using 10  $\mu$ M acceptor substrate, 0.1 mM UDP-Gal, and 0.5  $\mu$ M soluble  $\beta$ 3GalT5-1, incubated at 37 °C for 24 h. The resulting mixture was then diluted 10-fold for LC-MS analysis. N.D.: not determined.

synthesis of globo-series glycans, the bifunctional enzyme LgtD from *Haemophilus influenzae* catalyzed the conversion of Gb4

to Gb5 glycan through  $\beta$ 1,3-galactosylation.<sup>38–40</sup> This galactosylation reaction was relatively slow, and the yield of



Table 2. Donor Specificity of Soluble  $\beta$ 3GalT5-1 toward UDP-sugars and UDP-Galactose Analogs

UDP-Sugar donors	Structure	Relative activity (%) <sup>a</sup>	LC-MS yield (%) <sup>a,b</sup>
UDP-galactose		100	100
6-aldo-UDP-galactose		78 ± 2	53.4
6-azido-UDP-galactose		13 ± 2	30.1
UDP-glucose		1 ± 1	1.7
UDP-GlcNAc		0	1.9
UDP-GalNAc		5 ± 2	N.D.
UDP-glucuronide		0	N.D.
UDP-xylose		2 ± 1	N.D.

<sup>a</sup>\* Based on UDP-Gal as a reference, all reactions were performed in 20 mM HEPES (pH 7.4) containing 0.1 mM MnCl<sub>2</sub>, 0.1 mM UDP-sugar or UDP-Gal analogs, 5 mM Gb4 glycan, and 5  $\mu$ M soluble  $\beta$ 3GalT5-1 expressed from HEK293 cells. The results are presented as the mean  $\pm$  SD of three biological replicates. <sup>b</sup>\*\* All reactions were carried out using 10  $\mu$ M Gb4 glycan as the acceptor substrate, incubated with 0.1 mM UDP-sugar or UDP-Gal analogs, and 0.5  $\mu$ M soluble  $\beta$ 3GalT5-1 at 37 °C for 24 h. The resulting mixture was then diluted 10-fold for LC-MS analysis. N.D.: not determined.

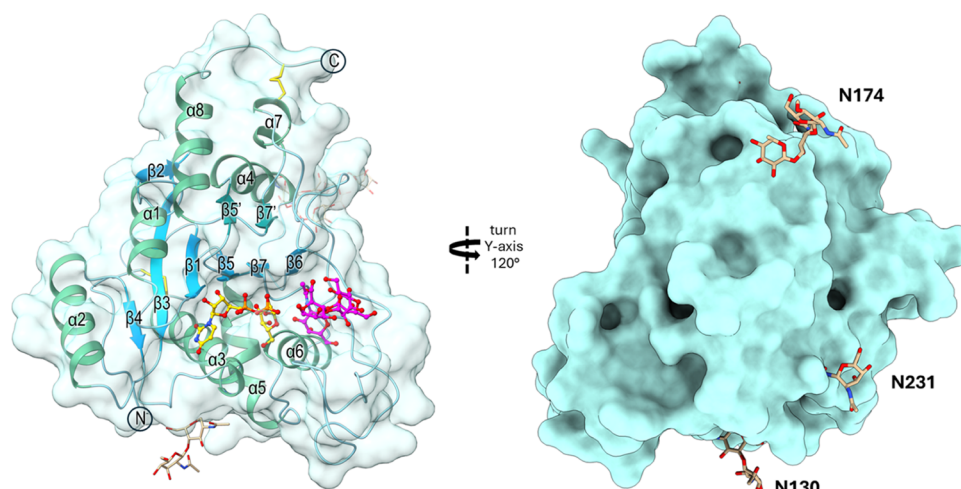
Table 3. Kinetic Parameters for Human Soluble  $\beta$ 3GalT5-1 and LgtD

enzyme <sup>a</sup>	expression host	substrate	$K_m$ (mM)	$k_{cat}$ (min <sup>-1</sup> )	$k_{cat}/K_m$ (min <sup>-1</sup> mM <sup>-1</sup> )
LgtD	<i>E. coli</i>	Allyl-Gb4 glycan	265.4 ± 88.3	0.046 ± 0.003	$1.8 \times 10^{-4} \pm 3.4 \times 10^{-5}$
soluble $\beta$ 3GalT5-1	Insect		0.66 ± 0.4	3.5 ± 0.4	5.3 ± 1.0
	HEK293		0.64 ± 0.3	3.3 ± 0.1	5.2 ± 3.0

<sup>a</sup>The enzymatic reaction of LgtD was performed in MOPS (pH 7.0) buffer with 1 mM MgCl<sub>2</sub>, 0.1 mM UDP-Gal, and a 2-fold serial concentration of allyl-Gb4 (0–500 mM). The soluble  $\beta$ 3GalT5-1 reaction was performed in HEPES buffer with 0.1 mM MnCl<sub>2</sub>, 0.1 mM UDP-Gal, and a 2-fold serial concentration of allyl-Gb4 (0–4 mM). The results were fitted to the Michaelis–Menten enzyme kinetics model using GraphPad Prism10 and are presented as the mean  $\pm$  SD of three biological replicates.

Gb5 glycan with sugar nucleotide regeneration ranged from 63 to 74%, which was less efficient when compared to the other enzymatic steps in the large-scale synthesis of globo-series glycans.<sup>25,41,42</sup> Based on the specificity study of  $\beta$ 3GalT5, we next tried to identify an effective expression system to prepare the soluble domain of human  $\beta$ 3GalT5-1 for the synthesis of SSEA-3 glycan. We initially expressed the enzyme in *Escherichia coli*. However, the enzyme was misfolded and could not be efficiently purified using the affinity tag (Figure S6A). We then coexpressed the soluble  $\beta$ 3GalT5-1 with various molecular chaperones to promote folding in *E. coli*.<sup>43,44</sup> Unfortunately, the yield was exceedingly low in all attempts except with the trigger factor (Figure S6B), but it did not have the proper enzymatic activity, probably due to the absence of post-translational glycosylation. We then expressed a fusion protein of  $\beta$ 3GalT5-1 with the trigger factor in *E. coli* to facilitate the disulfide bond formation (Figure S6C). This approach yielded a substantial increase in the production of active  $\beta$ 3GalT5-1, reaching 20–25 mg/L with activity comparable to that of LgtD (Table S3). Upon cleavage of

the trigger factor using thrombin, the activity of  $\beta$ 3GalT5-1 was doubled as compared to LgtD but remained ten times less efficient than human  $\beta$ 3GalT5-1 expressed in insect cells (Figure S6D). The soluble domain of  $\beta$ 3GalT5-1 expressed in HEK293 cells contains complex-type glycoforms with three N-glycosites. To assess the effect of glycosylation on the enzymatic activity, the soluble domain of  $\beta$ 3GalT5-1 was also expressed in insect cells (Sf9 cells) to produce the enzyme with mainly paucimannose glycoforms and in the presence of kifunensine to generate the high mannose-type glycoforms, which were further treated with Endo-H to obtain mono-GlcNAc  $\beta$ 3GalT5-1. The deglycosylated  $\beta$ 3GalT5-1 was also prepared after PNGase F digestion; however, the activity was significantly reduced to less than 50% activity compared to the untreated, a result similar to that from *E. coli*, while the high mannose (Man) and the mono-GlcNAc glycoforms had 2-fold higher activity than the paucimannose structure (Figure S6E). We also expressed the soluble form of  $\beta$ 3GalT5-1 in yeast (*Pichia pastoris*); however, the purified proteins had no function, probably due to improper glycosylation and



**Figure 3.** Surface representation of soluble  $\beta 3\text{GalT5-1}$  with overall structure and *N*-glycosylation. Surface representation of the enzyme  $\beta 3\text{GalT5-1}$ , highlighting its substrate binding cleft and *N*-glycosylation sites. The protein surface is shown in pale turquoise, while the  $\alpha$ -helix is shown in mint, the  $\beta$ -sheet is in dark cyan, and the two disulfide bonds (Cys52-Cys146 and Cys276-Cys307) are in yellow. The divalent ion manganese appears as a purple sphere. The donor substrate UDP-Gal is represented by gold sticks, and the acceptor substrate Gb4 glycan is represented by pink sticks. The three glycosites (N130, N174, and N231) are highlighted with brown sticks. The electron density map reveals the following glycosylation patterns: site N130 up to two GlcNAc residues, site N231 up to one GlcNAc residue, and site N174 displays paucimannose glycosylation.

misfolding. To compare the catalytic activity of LgtD and human  $\beta 3\text{GalT5-1}$  in converting Gb4 to SSEA-3 glycan, the kinetic parameters of soluble  $\beta 3\text{GalT5-1}$  expressed in *E. coli*, HEK293, and insect cells (Sf9 cells) were determined by monitoring the amount of UDP released from UDP-Gal. When allyl-Gb4 glycan was used as the substrate, the  $K_m$  and  $k_{cat}$  values for human  $\beta 3\text{GalT5-1}$  from HEK293 cells were  $0.64 \pm 0.3$  and  $3.3 \pm 0.1$  per minute, respectively, compared to  $265.4 \pm 88.3$  and  $0.046 \pm 0.003$  per minute for LgtD (Table 3). The turnover rate of human  $\beta 3\text{GalT5-1}$  was approximately 30,000-fold higher than that of LgtD.

**Optimization of Soluble  $\beta 3\text{GalT5-1}$  Expressed in Insect Cells for the Chemo-Enzymatic Synthesis of SSEA-3 Glycan.** To efficiently synthesize the globo-series glycans using soluble  $\beta 3\text{GalT5-1}$ , several problems must be resolved, and the most important one is to ensure proper glycosylation of the enzyme to retain the activity. While the expression yield of the enzyme from insect cells was approximately 8–10 mg/L, the activity of soluble  $\beta 3\text{GalT5-1}$  expression from insect cells was similar to that from HEK293 cells for the conversion of Gb4 to SSEA-3 glycan, as confirmed by ultra-performance liquid chromatography (UPLC). However, when we replaced LgtD with soluble  $\beta 3\text{GalT5-1}$  from insect cells and employed a sugar nucleotide regeneration method to synthesize the SSEA-3 glycan, a byproduct with additional galactose on the SSEA-3 glycan was observed, as confirmed by thin-layer chromatography (TLC) and MS. To address this issue, we optimized the amount of soluble  $\beta 3\text{GalT5-1}$  and the molar ratio of Gb4 glycan to UDP-Gal under various conditions. We found that when the molar ratio of UDP-Gal to Gb4 glycan was adjusted to 0.8 and 1, no side product was observed in TLC and MS analyses within 45 h. However, with an excess amount of soluble  $\beta 3\text{GalT5-1}$  and longer reaction time, side product formation increased after 69 h (Table S4). Therefore, we thought that the molar ratio of Gb4 glycan to soluble  $\beta 3\text{GalT5-1}$  should range from 8000 to 10,000, while the ratio of Gal (which then converted to UDP-Gal through sugar nucleotide regeneration) to Gb4 glycan

should range from 0.9 to 1. With this condition, we can minimize the side product formation, and the enzymatic reaction can be completed within a day.

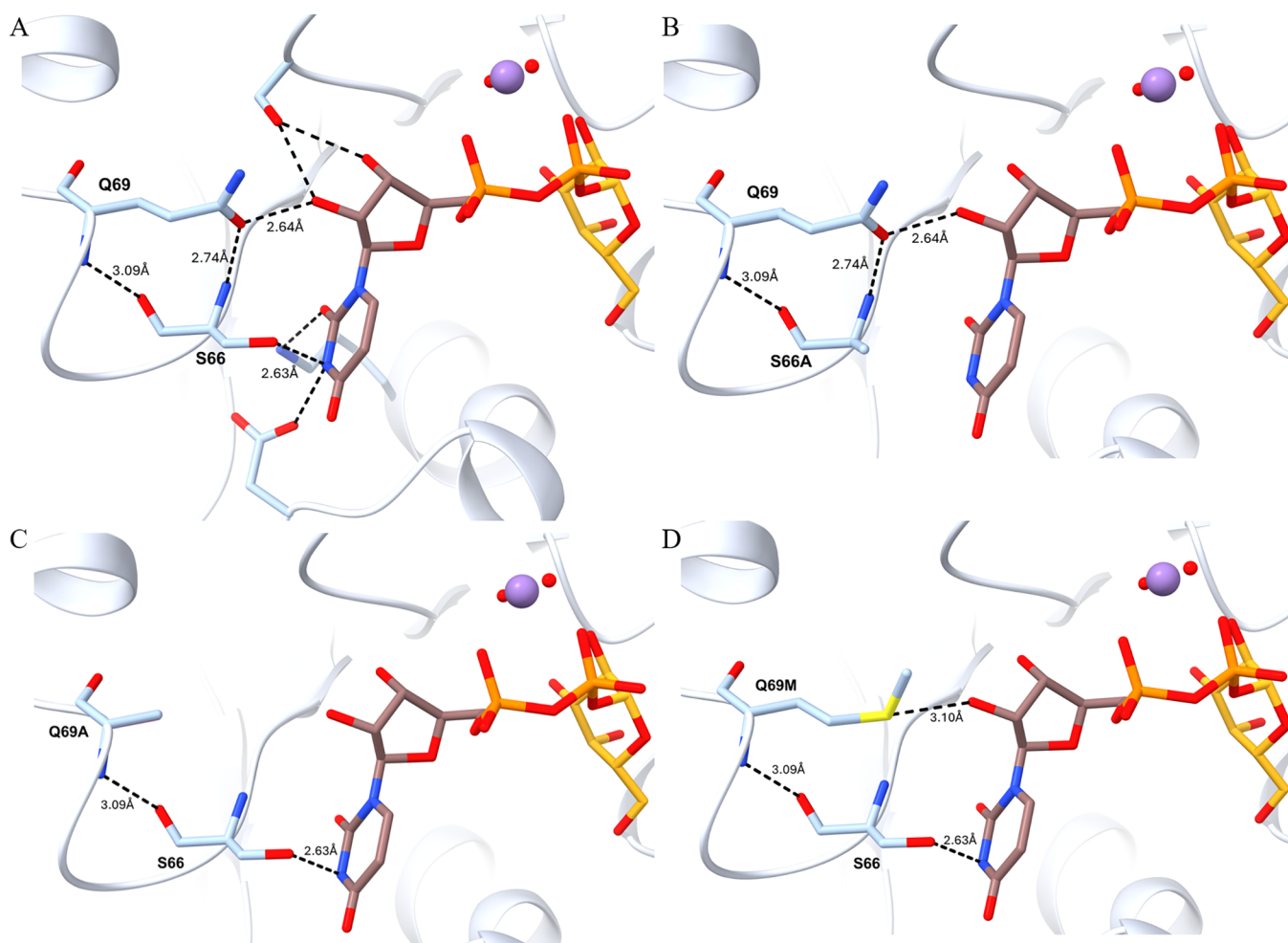
**Structure of Soluble  $\beta 3\text{GalT5-1}$  Expressed in Insect Cells.** The X-ray structure of soluble  $\beta 3\text{GalT5-1}$  from insect cells cocrystallized with UDP-Gal was determined through sulfur single-wavelength anomalous dispersion phasing, with a resolution of 2.20 Å. In addition, the structures of soluble  $\beta 3\text{GalT5-1}$  in complex with either substrates or products through molecular replacement, employing the sulfur derivative structure as a template, were also determined. The soluble enzyme is composed of residues Phe31-Pro308, although only residues Asp41-Pro308 are clearly visible in the electron density maps. The catalytic domain displayed a mixed  $\alpha/\beta$  Rossmann-like fold, typical of the GT-A glycosyltransferases family, including a central seven-stranded  $\beta$ -sheet ( $\beta 1$ – $\beta 7$ ) encircled primarily by  $\alpha$ -helices ( $\alpha 1$ – $\alpha 7$ ), a two-stranded antiparallel  $\beta$ -sheet ( $\beta 5'$  and  $\beta 7'$ ), and an additional 13-residue  $\alpha$ -helix ( $\alpha 8$ ) at the C-terminal. Two disulfide bonds, Cys52-Cys146 and Cys276-Cys307, along with the glycans at all three glycosylation sites (Asn130, Asn174, and Asn231) were identified. Both the donor and acceptor sugar molecules were positioned near the center of the binding cleft (Figure 3). As expected, based on the analysis of *N*-glycans with mass spectrometry, the soluble domain of  $\beta 3\text{GalT5-1}$  expressed from insect cells contains mainly paucimannose glycoforms, whereas the glycoforms from HEK293 cells are mainly complex type (Figure S7).

**Identification of S66A  $\beta 3\text{GalT5-1}$  with Enhanced Catalytic Efficiency through Site-Specific Alanine Scan for the Synthesis of Gb5 Glycan from Gb4 Glycan.** To improve the catalytic efficiency of soluble  $\beta 3\text{GalT5-1}$  in chemo-enzymatic synthesis, we introduced site-directed mutagenesis of residues that interact with UDP-Gal and Gb4 glycan, followed by the evaluation of their activity through an enzymatic assay. Remarkably, the activities of T64A, S66A, and Q69A variants of  $\beta 3\text{GalT5-1}$ , which interacted with UDP-Gal, were significantly increased. Particularly intriguing

Table 4. Kinetic Parameters of Soluble  $\beta$ 3GalT5-1 Mutants

$\beta$ 3GalT5-1 <sup>a</sup>	substrate	$K_m$ (mM)	$k_{cat}$ (min <sup>-1</sup> )	$k_{cat}/K_m$ (min <sup>-1</sup> mM <sup>-1</sup> )
WT	Allyl-Gb4	1.40 $\pm$ 0.2	2.69 $\pm$ 0.2	1.92 $\pm$ 1.0
	UDP-Gal	0.14 $\pm$ 0.04	7.51 $\pm$ 0.7	52.18 $\pm$ 17.6
S66A	Allyl-Gb4	1.20 $\pm$ 0.2	4.66 $\pm$ 0.3	3.89 $\pm$ 1.9
	UDP-Gal	0.35 $\pm$ 0.1	16.55 $\pm$ 2.5	46.86 $\pm$ 24.5
Q69A	Allyl-Gb4	1.21 $\pm$ 0.3	0.72 $\pm$ 0.07	0.60 $\pm$ 0.2
	UDP-Gal	0.30 $\pm$ 0.05	2.98 $\pm$ 0.2	9.88 $\pm$ 3.8
Q69M	Allyl-Gb4	0.42 $\pm$ 0.3	0.64 $\pm$ 0.1	1.52 $\pm$ 0.5
	UDP-Gal	0.21 $\pm$ 0.05	3.51 $\pm$ 0.3	16.93 $\pm$ 6.2
S66A/Q69A	Allyl-Gb4	1.11 $\pm$ 0.3	0.47 $\pm$ 0.05	0.42 $\pm$ 0.2
	UDP-Gal	0.32 $\pm$ 0.09	1.64 $\pm$ 0.2	9.35 $\pm$ 2.7
S66A/Q69M	Allyl-Gb4	1.65 $\pm$ 0.4	1.49 $\pm$ 0.1	0.90 $\pm$ 0.4
	UDP-Gal	0.27 $\pm$ 0.06	4.64 $\pm$ 0.4	17.05 $\pm$ 6.8

<sup>a</sup>All reactions were conducted in a buffer containing 20 mM HEPES (pH 7.4) and 0.1 mM MnCl<sub>2</sub>, with 0.5  $\mu$ M soluble  $\beta$ 3GalT5-1 mutants expressed in HEK293 cells. To measure the kinetic parameters of allyl-Gb4 or UDP-Gal, reactions were set up with a fixed concentration of 10 mM UDP-sugar or 2 mM Gb4 glycan and a 2-fold serial dilution of allyl-Gb4 glycan (0–4 mM) or UDP-Gal (0–1 mM). The resulting data were analyzed using GraphPad Prism10 to fit the Michaelis–Menten enzyme kinetics model. The results are presented as the mean  $\pm$  SD of three biological replicates.

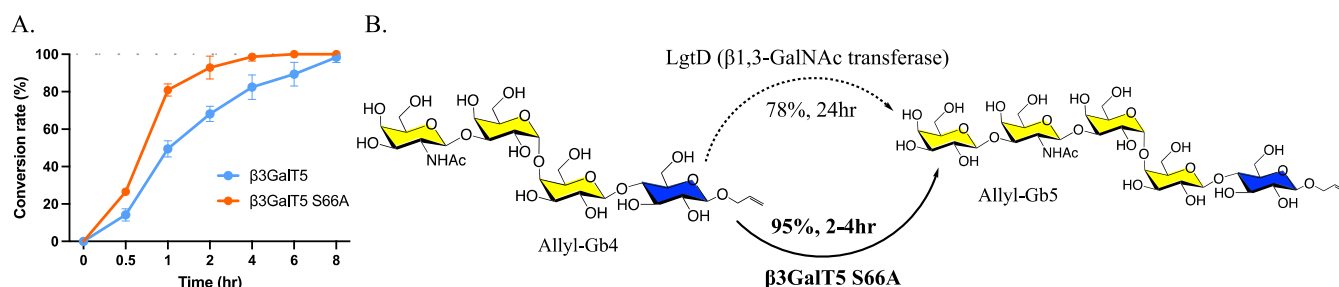


**Figure 4.** Mutations of S66A and Q69A/M in soluble  $\beta$ 3GalT5-1 affect interactions with UDP. A. Residues S66 and Q69 in wild-type  $\beta$ 3GalT5-1 interact (dashed lines) with the UDP moiety only. Side chains of residues involved are shown as sticks. UDP is shown as brown sticks, Gal as gold sticks, divalent ion Mn<sup>2+</sup> as a purple sphere, and water molecules as red spheres. B. Modeling structure of S66A mutant. C. Modeling structure of Q69A mutant. D. Modeling structure of Q69M. All modeled mutant structures are based on the wild-type  $\beta$ 3GalT5-1 structure.

were the findings that S66A and Q69A exhibited approximately 3-fold and 1.5-fold increases in activity compared to the wild-type enzyme, whereas the activity of other mutants decreased significantly (Figure S8). Furthermore, we conducted muta-

tions of the S66 and Q69 positions with alternative amino acids. When the S66 position was substituted with valine, threonine, and glycine, a slight reduction in activity was observed compared to alanine but remained higher than that of





**Figure 5.** Optimized procedure for the enzymatic synthesis of SSEA-3 glycan. A.  $\beta 3\text{GalT5}$ –1 and  $\beta 3\text{GalT5}$ –1-S66A catalyzed the synthesis of 2-AB-Gb5 glycan at various time points. The results are presented as the mean  $\pm$  SD of three biological replicates. B. Synthesis of allyl-SSEA-3 glycan using  $\beta 3\text{GalT5}$ –1-S66A and LgtD.

the wild type. When the Q69 position was replaced with methionine, its activity matched that of alanine and exhibited an approximately 2-fold increase in activity compared to the wild type. In comparing the kinetic parameters of these single mutants (S66A, Q69A, and Q69M) and double mutants (S66A/Q69A or S66A/Q69M) using UDP-Gal as donor and allyl-Gb4 glycan as acceptor, we observed a significant reduction in the enzymatic activities of  $\beta 3\text{GalT5}$ –1 with S66A/Q69A or S66A/Q69M double mutations compared to the wild type. Among the single mutants, S66A displayed the highest efficiency, with  $K_m$  values for UDP-Gal and allyl-Gb4 glycan of 0.35 and 1.20 mM,  $k_{\text{cat}}$  values at 16.55 and 4.66 per minute, and  $k_{\text{cat}}/K_m$  values at 46.86 and 3.89  $\text{min}^{-1} \text{mM}^{-1}$ , respectively. Notably, S66A exhibited a decreased binding affinity for UDP-Gal and Gb4 glycan, but it showed an increased turnover rate and efficiency for Gb4 glycan when compared to the wild-type  $\beta 3\text{GalT5}$ –1. The specificity for UDP-Gal remained unchanged compared to that of wild type (Table 4).

To understand the higher activity of the S66A mutant, we performed molecular modeling of the wild-type enzyme in complex with UDP-Gal. It was shown that the S66 backbone carbonyl group interacted with the backbone NH of Q69 through hydrogen bonding. This interaction remains the same even with S66A and/or Q69A/M mutations (Figure 4A). In the S66A mutant, the interaction between S66A and the N3 of uracil in the UDP moiety was abolished (Figure 4B). In addition, the hydrogen bonding between the S66 carbonyl group and the Q69 NH group no longer existed in the Q69A and Q69M mutations (Figure 4C and D). The lack of S66A or Q69A/M side chain interactions with UDP led to decreased binding to UDP-Gal. Consequently, wild-type  $\beta 3\text{GalT5}$  binds to UDP-Gal more strongly ( $K_m = 0.14 \text{ mM}$ ) compared to the S66A ( $K_m = 0.35 \text{ mM}$ ) and Q69A ( $K_m = 0.30 \text{ mM}$ ) mutants. Moreover, a similar  $K_m$  value for UDP-Gal was observed for the Q69M, S66A, and Q69A mutants. Although structural analysis can only directly explain the  $K_m$  value of the kinetic parameters, the loss of the S66A side chain interaction with the UDP moiety might contribute to the higher turnover rate observed for the S66A mutant ( $k_{\text{cat}} = 16.55 \text{ UDP-Gal/min}$ ; Table 4). The weaker binding to UDP potentially contributes to a faster release of UDP and faster reaction as more UDP-Gal is accepted by the enzyme in each catalytic cycle, and the UDP generated is more quickly released from the enzyme. In enzymatic synthesis, the concentrations of substrates usually are higher than their  $K_m$  values; consequently, the enzyme is saturated by substrates in the reaction, and the efficiency is thus determined by the  $k_{\text{cat}}$  value (the turnover rate). This

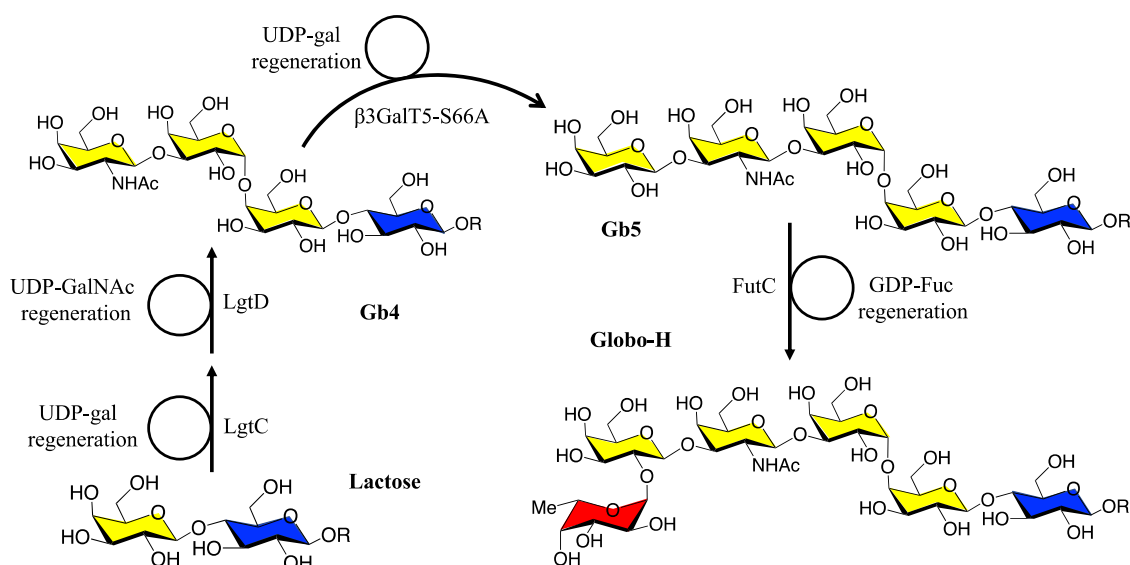
enhanced catalytic efficiency means that the S66A mutants reduced the time needed to synthesize the SSEA-3 glycan.

The substrate specificity of soluble S66A  $\beta 3\text{GalT5}$ –1 closely resembled that of wild-type soluble  $\beta 3\text{GalT5}$ –1 (Figure S5). Notably, the relative activity for disaccharides with Gal at the terminal end was elevated compared to the wild type. This increase in activity could lead to an additional galactosylation in the synthesis of Gb4 glycan. However, this side reaction can be minimized by adjusting the molar ratio of Gal to Gb4 glycan ranging from 0.8 to 1. This ratio is consistent with the conditions used for synthesizing allyl-Gb5 via sugar nucleotide regeneration with the wild-type soluble  $\beta 3\text{GalT5}$ –1. Moreover, by comparing the thermostability of wild type and S66A mutant of  $\beta 3\text{GalT5}$ –1, based on the melting temperature ( $T_m$ ) calculated by Boltzmann and derivative methods, the WT  $\beta 3\text{GalT5}$  expressed from insect cells showed  $T_{m,\text{Boltzmann}}$ : 42.95  $^{\circ}\text{C}$  and  $T_{m,\text{Derivative}}$ : 45.15  $^{\circ}\text{C}$ , whereas the S66A mutant showed  $T_{m,\text{Boltzmann}}$ : 42.42  $^{\circ}\text{C}$  and  $T_{m,\text{Derivative}}$ : 44.72  $^{\circ}\text{C}$  (Figure S9). The melting temperature between the WT and S66A mutant showed similar thermostability.

**Enzymatic Synthesis of Allyl-Gb5 Glycan from Allyl-Gb4 Glycan with Soluble  $\beta 3\text{GalT5}$ –1-S66A.** To verify the relative activity and kinetic parameters of  $\beta 3\text{GalT5}$ –1-S66A, we performed the chemo-enzymatic synthesis of allyl-Gb5 glycan using the S66A mutant expressed from insect cells as a catalyst and 2-aminobenzamide (2-AB) labeled Gb4 glycan as a substrate, and the reaction was monitored by UPLC. Within 2 h, approximately 90% of Gb4 glycan was converted to SSEA-3 glycan, whereas the wild-type  $\beta 3\text{GalT5}$ –1 achieved only a 60% conversion. In 4 h, almost all Gb4 glycan was converted to SSEA-3 glycan by  $\beta 3\text{GalT5}$ –1-S66A, while the reaction time required for wild-type  $\beta 3\text{GalT5}$ –1 was doubled (Figure 5A). In a preparative-scale synthesis starting with 100 mg of allyl-Gb4 glycan and  $\beta 3\text{GalT5}$ –1-S66A along with other enzymes and reagents, nearly all of the allyl-Gb4 glycan was transformed into allyl-SSEA-3 glycan within 2 h as monitored by TLC. The yield of allyl-SSEA-3 glycan after purification through a C-18 column and confirmed with MS and NMR was 95% (Figure 5B). With the improved conversion efficiency from allyl-Gb4 to allyl-SSEA-3 glycan, this method should enable the large-scale one-pot synthesis of SSEA-4 and Globo-H glycans, as demonstrated previously, with lower cost and less reaction time.

**Enzymatic Synthesis of Allyl-Globo-H Glycan from Allyl-Lactose Using Soluble  $\beta 3\text{GalT5}$ –1-S66A Coupled with Sugar Nucleotide Regeneration.** To synthesize the allyl-glycan of Globo-H, the starting material allyl-lactose was mixed with other components following the procedure described previously,<sup>17</sup> except the galactosyltransferase used





**Figure 6.** Enzymatic synthesis of allyl-Globo-H glycan from allyl-lactose with sugar nucleotide regeneration. Gb4 glycan was synthesized through the sequential action of  $\alpha$ -1,4-galactosyltransferase (LgtC) and  $\beta$ -1,4-*N*-acetylgalactosaminyltransferase (LgtD) by transferring UDP-Gal to lactose and UDP-GalNAc to Gb3, achieving a synthetic yield of 94%. Subsequently, Gb5 glycan was synthesized using  $\beta$ -1,3-galactosyltransferase (soluble  $\beta$ GalT5–1 with S66A mutation) with a synthetic yield of 95%. Finally, Globo-H glycan was synthesized by using  $\alpha$ -1,2-fucosyltransferase (FutC), resulting in a synthetic yield of 95%.

for converting allyl-Gb4 to allyl-Gb5 glycan. In brief, 500 mg of allyl-lactose in a final volume of 22.3 mL solution containing 25 mM Tris-HCl (pH 7.5), 15 mM magnesium chloride ( $\text{MgCl}_2$ ), 50 mM Gal, 5 mM adenosine triphosphate (ATP), 5 mM uridine 5'-triphosphate (UTP), and 100 mM phosphoenol-pyruvate (PEP) was mixed with all required enzymes, including 0.2 mM UDP-sugar pyrophosphorylase (AtUSP), 0.4 mM galactose kinase (GalK), 0.1 mM pyruvate kinase (pykF), 1 mM pyrophosphatase (PPA), and 1 mM  $\alpha$ -1,4-galactosyltransferase (LgtC). The reaction was incubated at 30 °C with gentle shaking and monitored with thin-layer chromatography using the solvent system butanol/acetic acid/water (5:3:2 v/v/v) until conversion to allyl-Gb3 glycan was completed. For the conversion of allyl-Gb3 to allyl-Gb4 glycan, the solution was further added to the following reagents and enzymes: 15 mM Tris-HCl (pH 7.5), 50 mM GalNAc, 15 mM  $\text{MgCl}_2$ , 5 mM ATP, 5 mM UTP, 100 mM PEP, 0.1 mM pykF, 1 mM PPA, 0.2 mM *N*-acetylhexosamine 1-kinase (NahK), 0.4 mM *N*-acetylglucosamine-1-phosphate uridylyltransferase (GlmU), and 0.2 mM LgtD until the reaction was finished, and then the following reagents and enzymes were added to the mixture to generate allyl-Gb5 (SSEA-3) glycan: 25 mM Tris-HCl (pH 7.5), 10 mM  $\text{MgCl}_2$ , 10 mM manganese(II) chloride ( $\text{MnCl}_2$ ), 50 mM Gal, 5 mM ATP, 5 mM UTP, 100 mM PEP, 0.2 mM GalK, 0.2 mM AtUSP, 0.1 mM pykF, 0.5 mM PPA, and 0.06 mM S66A  $\beta$ GalT5–1. When allyl-Gb4 glycan was fully converted to allyl-SSEA-3 glycan, the mixture was then passed through 10K MWCO AmiconUltra, washed with water several times, and further purified by C-18 column with a gradient concentration of methanol. To synthesize allyl-Globo-H glycan, 500 mg of purified allyl-SSEA-3 glycan was dissolved in 10.2 mL of the solution containing 0.6 mM pykF, 1.5 mM PPA, 0.4 mM fucokinase/GDP-fucose pyrophosphorylase (FKP), 0.4 mM fucosyltransferase (FUTC), 55 mM fucose, 5 mM ATP, 5 mM GTP, 25 mM Tris-HCl (pH 7.5), 10 mM  $\text{MgCl}_2$ , and 100 mM PEP. The product allyl-Globo-H glycan was then purified by a C-18 column with a gradient of

methanol and further confirmed the structure by NMR and mass spectrometry. The overall yield of allyl-Globo-H glycan from allyl-lactose is 85% (Figure 6).

## CONCLUSIONS

Human galactosyltransferase  $\beta$ 3GalT5 is the key enzyme that catalyzes the galactosylation of Gb4 to the cancer-associated SSEA-3 (Gb5), leading to the subsequent synthesis of the other two globo-series GSLs, SSEA-4 and Globo-H. The globo-series GSLs are often exclusively expressed on the surface of various types of cancer, and their glycans are considered targets for cancer therapies. The development of a practical enzymatic method for the synthesis of globo-series glycans is important, as it will provide sufficient materials for the evaluation of these glycans as targets for clinical development. Previously, we developed an enzymatic method for the large-scale synthesis of Globo-H glycan to enable the manufacture of a cancer vaccine for the phase 3 clinical trial. In this enzymatic process, the microbial galactosyltransferase LgtD was used in the synthesis of SSEA-3 glycan but was significantly less efficient than the human galactosyltransferase  $\beta$ 3GalT5. However, human  $\beta$ 3GalT5 is a glycoprotein with three *N*-glycosites and exists as two isoforms with only a difference in four N-terminal amino acids; it is not clear about the roles of each isoform and glycosylation on the proteins in catalysis and synthesis of globo-series glycolipids. In this study, we developed a precise dsRNA approach to specifically knockdown each of the two isoforms of human  $\beta$ 3GalT5 in MCF-7 cells and identified  $\beta$ 3GalT5–1 as the main isoform responsible for the synthesis of SSEA-3 (Gb5) from Gb4 in cancer cells. We also found that among the various glycosylation patterns, the glycoforms expressed in human (complex type) and insect cells (Sf9) (paucimannose type) are more active than the glycoforms with other types of glycosylation patterns from other species including yeast (such as *Pichia pastoris*). We therefore successfully expressed the common soluble domains of  $\beta$ 3GalT5–1 and  $\beta$ 3GalT5–2

(N29–V310) in insect cells as the most active glycoforms and investigated the substrate specificity. Furthermore, we performed a site-specific alanine scan and mutation in the binding site and identified the S66A mutant of  $\beta$ 3GalT5–1 with a 10-fold increase in efficiency compared to the microbial enzyme LgtD. Molecular modeling of wild-type  $\beta$ 3GalT5–1 and its S66A mutant revealed that the loss of key interactions with UDP-Gal in the mutant decreases binding affinity for UDP-Gal (higher  $K_m$  value) but enhances turnover rate (higher  $k_{cat}$ ) with no effect on binding to the glycan acceptor. With the soluble domain of  $\beta$ 3GalT5–1 readily available from insect cell expression, it can be used with other enzymes for the efficient synthesis of oligosaccharides and facilitate the evaluation of their translational applications.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.4c11723>.

Synthetic and experimental procedures, supporting figures and tables displaying characterization and structural analyses and crystallographic information are included (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Che Ma** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan; [orcid.org/0000-0002-4741-2307](https://orcid.org/0000-0002-4741-2307); Email: [cma@gate.sinica.edu.tw](mailto:cma@gate.sinica.edu.tw)

**Chi-Huey Wong** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan; Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037, United States; [orcid.org/0000-0002-9961-7865](https://orcid.org/0000-0002-9961-7865); Email: [wong@scripps.edu](mailto:wong@scripps.edu)

### Authors

**Chih-Chuan Kung** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan

**Jennifer M. Lo** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan

**Kuo-Shiang Liao** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan

**Chung-Yi Wu** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan

**Li-Chun Cheng** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan

**Cinya Chung** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan

**Tsui-Ling Hsu** – Genomics Research Center, Academia Sinica, Taipei 11529, Taiwan

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/jacs.4c11723>

### Author Contributions

<sup>§</sup>The authors C.-C.K. and J.M.L. contributed equally. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank the Academia Sinica Glycoscience Core Facility, funded by the Academia Sinica Core Facility and Innovative Instrument Project (AS-CFII-112-102), Core Facilities of Translational Medicine of BioTRC (National Biotechnology Research Park, Yi-Ping Huang for technical support, and Dr. Ting-Jen Cheng for providing original plasmid of human  $\beta$ 3GalT5. Mass spectrometry analyses were performed by GRC Mass Core Facility of Genomics Research Center, Academia Sinica. This research was supported by Academia Sinica.

## ■ REFERENCES

- (1) Hakomori, S.-i.; Zhang, Y. Glycosphingolipid antigens and cancer therapy. *Chem. Biol.* **1997**, *4* (2), 97–104.
- (2) Hakomori, S.-i. Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines. *Adv. Exp. Med. Biol.* **2001**, *491*, 369–402.
- (3) Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H., et al., *Essentials of Glycobiology*, 4th ed., Cold Spring Harbor Laboratory Press, 2022.
- (4) Häuselmann, I.; Borsig, L. Altered tumor-cell glycosylation promotes metastasis. *Front. Oncol.* **2014**, *4*, 28.
- (5) Vasconcelos-Dos-Santos, A.; Oliveira, I. A.; Lucena, M. C.; Mantuano, N. R.; Whelan, S. A.; Dias, W. B.; Todeschini, A. R. Biosynthetic Machinery Involved in Aberrant Glycosylation: Promising Targets for Developing of Drugs Against Cancer. *Front. Oncol.* **2015**, *5*, 138.
- (6) Aloia, A.; Petrova, E.; Tomiuk, S.; Bissels, U.; Deas, O.; Saini, M.; Zickgraf, F. M.; Wagner, S.; Spaich, S.; Sutterlin, M.; et al. The sialyl-glycolipid stage-specific embryonic antigen 4 marks a subpopulation of chemotherapy-resistant breast cancer cells with mesenchymal features. *Breast Cancer Res.* **2015**, *17* (1), No. 146.
- (7) Chang, W. W.; Lee, C. H.; Lee, P.; Lin, J.; Hsu, C. W.; Hung, J. T.; Lin, J. J.; Yu, J. C.; Shao, L. E.; Yu, J.; et al. Expression of Globo H and SSEA3 in breast cancer stem cells and the involvement of fucosyl transferases 1 and 2 in Globo H synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (33), 11667–11672.
- (8) Gottschling, S.; Jensen, K.; Warth, A.; Herth, F. J.; Thomas, M.; Schnabel, P. A.; Herpel, E. Stage-specific embryonic antigen-4 is expressed in basaloid lung cancer and associated with poor prognosis. *Eur. Respir. J.* **2013**, *41* (3), 656–663.
- (9) Huang, Y. L.; Hung, J. T.; Cheung, S. K.; Lee, H. Y.; Chu, K. C.; Li, S. T.; Lin, Y. C.; Ren, C. T.; Cheng, T. J.; Hsu, T. L.; et al. Carbohydrate-based vaccines with a glycolipid adjuvant for breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (7), 2517–2522.
- (10) Danishefsky, S. J.; Shue, Y. K.; Chang, M. N.; Wong, C. H. Development of Globo-H cancer vaccine. *Acc. Chem. Res.* **2015**, *48* (3), 643–652.
- (11) Chen, N. Y.; Lin, C. W.; Lai, T. Y.; Wu, C. Y.; Liao, P. C.; Hsu, T. L.; Wong, C. H. Increased expression of SSEA-4 on TKI-resistant non-small cell lung cancer with EGFR-T790M mutation. *Proc. Natl. Acad. Sci. U.S.A.* **2024**, *121* (5), No. e2313397121.
- (12) Lin, C. W.; Wang, Y. J.; Lai, T. Y.; Hsu, T. L.; Han, S. Y.; Wu, H. C.; Shen, C. N.; Dang, V.; Chen, M. W.; Chen, L. B.; Wong, C. H. Homogeneous antibody and CAR-T cells with improved effector functions targeting SSEA-4 glycan on pancreatic cancer. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118* (50), No. e2114774118.
- (13) Hakomori, S.-i. Tumor-Associated Carbohydrate Antigens. *Annu. Rev. Immunol.* **1984**, *2*, 103–126.
- (14) Feng, D.; Shaikh, A. S.; Wang, F. Recent Advance in Tumor-associated Carbohydrate Antigens (TACAs)-based Antitumor Vaccines. *ACS Chem. Biol.* **2016**, *11* (4), 850–863.
- (15) Ingale, S.; Wolfert, M. A.; Gaekwad, J.; Buskas, T.; Boons, G. J. Robust immune responses elicited by a fully synthetic three-component vaccine. *Nat. Chem. Biol.* **2007**, *3* (10), 663–667.

- (16) Renaudet, O.; BenMohamed, L.; Dasgupta, G.; Bettahi, I.; Dumy, P. Towards a self-adjuvanting multivalent B and T cell epitope containing synthetic glycolipopeptide cancer vaccine. *ChemMedChem* **2008**, *3* (5), 737–741.
- (17) Buskas, T.; Thompson, P.; Boons, G. J. Immunotherapy for cancer: synthetic carbohydrate-based vaccines. *Chem. Commun.* **2009**, No. 36, 5335–5349.
- (18) Berti, F.; Adamo, R. Recent mechanistic insights on glycoconjugate vaccines and future perspectives. *ACS Chem. Biol.* **2013**, *8* (8), 1653–1663.
- (19) Chuang, P. K.; Hsiao, M.; Hsu, T. L.; Chang, C. F.; Wu, C. Y.; Chen, B. R.; Huang, H. W.; Liao, K. S.; Chen, C. C.; Chen, C. L.; et al. Signaling pathway of globo-series glycosphingolipids and  $\beta$ 1,3-galactosyltransferase V ( $\beta$ 3GalT5) in breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116* (9), 3518–3523.
- (20) Wang, S. W.; Ko, Y. A.; Chen, C. Y.; Liao, K. S.; Chang, Y. H.; Lee, H. Y.; Yu, Y. H.; Lih, Y. H.; Cheng, Y. Y.; Lin, H. H.; et al. Mechanism of Antigen Presentation and Specificity of Antibody Cross-Reactivity Elicited by an Oligosaccharide-Conjugate Cancer Vaccine. *J. Am. Chem. Soc.* **2023**, *145* (17), 9840–9849.
- (21) Cheung, S. K.; Chuang, P. K.; Huang, H. W.; Hwang-Verslues, W. W.; Cho, C. H.; Yang, W. B.; Shen, C. N.; Hsiao, M.; Hsu, T. L.; Chang, C. F.; et al. Stage-specific embryonic antigen-3 (SSEA-3) and  $\beta$ 3GalT5 are cancer specific and significant markers for breast cancer stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113* (4), 960–965.
- (22) Lou, Y. W.; Wang, P. Y.; Yeh, S. C.; Chuang, P. K.; Li, S. T.; Wu, C. Y.; Khoo, K. H.; Hsiao, M.; Hsu, T. L.; Wong, C. H. Stage-specific embryonic antigen-4 as a potential therapeutic target in glioblastoma multiforme and other cancers. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (7), 2482–2487.
- (23) Rugo, H. S.; Cortes, J.; Xu, B. H.; Huang, C. S.; Kim, S. B.; Melisko, M. E.; Nanda, R.; Sharma, P.; Schwab, R.; Hsu, P. A phase 3, randomized, open-label study of the anti-Globo H vaccine adagloxad simolenin/obi-821 in the adjuvant treatment of high-risk, early-stage, Globo H-positive triple-negative breast cancer. *J. Clin. Oncol.* **2022**, *40* (16 suppl), No. TPS611.
- (24) Huang, C. S.; Yu, A. L.; Tseng, L. M.; Chow, L. W. C.; Hou, M. F.; Hurvitz, S. A.; Schwab, R. B.; Murray, J. L.; Chang, H. K.; Chang, H. T.; et al. Globo H-KLH vaccine adagloxad simolenin (OBI-822)/OBI-821 in patients with metastatic breast cancer: phase II randomized, placebo-controlled study. *J. Immunother. Cancer* **2020**, *8* (2), No. e000342.
- (25) Tsai, T. I.; Lee, H. Y.; Chang, S. H.; Wang, C. H.; Tu, Y. C.; Lin, Y. C.; Hwang, D. R.; Wu, C. Y.; Wong, C. H. Effective sugar nucleotide regeneration for the large-scale enzymatic synthesis of Globo H and SSEA4. *J. Am. Chem. Soc.* **2013**, *135* (39), 14831–14839.
- (26) Song, K.; Herzog, B. H.; Fu, J.; Sheng, M.; Bergstrom, K.; McDaniel, J. M.; Kondo, Y.; McGee, S.; Cai, X.; Li, P.; et al. Loss of Core 1-derived O-Glycans Decreases Breast Cancer Development in Mice. *J. Biol. Chem.* **2015**, *290* (33), 20159–20166.
- (27) Burchell, J. M.; Beatson, R.; Graham, R.; Taylor-Papadimitriou, J.; Tajadura-Ortega, V. O-linked mucin-type glycosylation in breast cancer. *Biochem. Soc. Trans.* **2018**, *46* (4), 779–788.
- (28) Brockhausen, I. Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. *EMBO Rep.* **2006**, *7* (6), 599–604.
- (29) Scott, D. A.; Drake, R. R. Glycosylation and its implications in breast cancer. *Expert Rev. Proteomics* **2019**, *16* (8), 665–680.
- (30) Zhang, J. H.; Huang, Y.; Li, H.; Xu, P. F.; Liu, Q. H.; Sun, Y.; Zhang, Z. J.; Wu, T.; Tang, Q.; Jia, Q. Y.; et al. B3galt5 functions as a PXR target gene and regulates obesity and insulin resistance by maintaining intestinal integrity. *Nat. Commun.* **2024**, *15* (1), No. 5919.
- (31) Ščupáková, K.; Adelaja, O. T.; Balluff, B.; Ayyappan, V.; Tressler, C. M.; Jenkinson, N. M.; Claes, B. S.; Bowman, A. P.; Cimino-Mathews, A. M.; White, M. J.; et al. Clinical importance of high-mannose, fucosylated, and complex N-glycans in breast cancer metastasis. *JCI Insight* **2021**, *6* (24), No. e146945.
- (32) Benesova, I.; Nenutil, R.; Urminsky, A.; Lattova, E.; Uhrík, L.; Grell, P.; Kokas, F. Z.; Halamkova, J.; Zdrahal, Z.; Vojtesek, B.; et al. N-glycan profiling of tissue samples to aid breast cancer subtyping. *Sci. Rep.* **2024**, *14* (1), No. 320.
- (33) Holden, P.; Horton, W. A. Crude subcellular fractionation of cultured mammalian cell lines. *BMC Res. Notes* **2009**, *2*, 243.
- (34) Zhou, D. P.; Henion, T. R.; Jungalwala, F. B.; Berger, E. G.; Hennet, T. The  $\beta$ 1,3-galactosyltransferase  $\beta$ 3GalT-V is a stage-specific embryonic antigen-3 (SSEA-3) synthase. *J. Biol. Chem.* **2000**, *275* (30), 22631–22634.
- (35) Isshiki, S.; Togayachi, A.; Kudo, T.; Nishihara, S.; Watanabe, M.; Kubota, T.; Kitajima, M.; Shiraishi, N.; Sasaki, K.; Andoh, T.; et al. Cloning, expression, and characterization of a novel UDP-galactose: $\beta$ -N-acetylglucosamine  $\beta$ 1,3-galactosyltransferase ( $\beta$ 3Gal-T5) responsible for synthesis of type 1 chain in colorectal and pancreatic epithelia and tumor cells derived therefrom. *J. Biol. Chem.* **1999**, *274* (18), 12499–12507.
- (36) Zhou, D.; Berger, E. G.; Hennet, T. Molecular cloning of a human UDP-galactose:GlcNAc $\beta$ 1,3GalNAc  $\beta$ 1, 3 galactosyltransferase gene encoding an O-linked core3-elongation enzyme. *Eur. J. Biochem.* **1999**, *263* (2), 571–576.
- (37) Salvini, R.; Bardoni, A.; Valli, M.; Trinchera, M.  $\beta$ 1,3-Galactosyltransferase  $\beta$ 3Gal-T5 acts on the GlcNAc $\beta$ 1 $\rightarrow$ 3Gal $\beta$ 1 $\rightarrow$ 4GlcNAc $\beta$ 1 $\rightarrow$ R sugar chains of carcinoembryonic antigen and other N-linked glycoproteins and is down-regulated in colon adenocarcinomas. *J. Biol. Chem.* **2001**, *276* (5), 3564–3573.
- (38) Shao, J.; Zhang, J.; Kowal, P.; Wang, P. G. Donor substrate regeneration for efficient synthesis of globotetraose and isoglobotetraose. *Appl. Environ. Microbiol.* **2002**, *68* (11), 5634–5640.
- (39) Shao, J.; Zhang, J.; Kowal, P.; Lu, Y.; Wang, P. G. Overexpression and biochemical characterization of  $\beta$ -1,3-N-acetylglactosaminyltransferase LgtD from *Haemophilus influenzae* strain Rd. *Biochem. Biophys. Res. Commun.* **2002**, *295* (1), 1–8.
- (40) Randrianitsoa, M.; Drouillard, S.; Breton, C.; Samain, E. Synthesis of globopentaose using a novel  $\beta$ 1,3-galactosyltransferase activity of the *Haemophilus influenzae*  $\beta$ 1,3-N-acetylglactosaminyltransferase LgtD. *FEBS Lett.* **2007**, *581* (14), 2652–2656.
- (41) Li, P. J.; Huang, S. Y.; Chiang, P. Y.; Fan, C. Y.; Guo, L. J.; Wu, D. Y.; Angata, T.; Lin, C. C. Chemoenzymatic Synthesis of DSGb5 and Sialylated Globo-series Glycans. *Angew. Chem., Int. Ed.* **2019**, *58* (33), 11273–11278.
- (42) Chen, C. Y.; Lin, Y. W.; Wang, S. W.; Lin, Y. C.; Cheng, Y. Y.; Ren, C. T.; Wong, C. H.; Wu, C. Y. Synthesis of Azido-Globo H Analogs for Immunogenicity Evaluation. *ACS Cent. Sci.* **2022**, *8* (1), 77–85.
- (43) Nishihara, K.; Kanemori, M.; Kitagawa, M.; Yanagi, H.; Yura, T. Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2 in *Escherichia coli*. *Appl. Environ. Microbiol.* **1998**, *64* (5), 1694–1699.
- (44) Nishihara, K.; Kanemori, M.; Yanagi, H.; Yura, T. Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. *Appl. Environ. Microbiol.* **2000**, *66* (3), 884–889.