Chemistry & Biology, Vol. 11, 1009–1016, July, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.05.009

Deconvoluting the Functions of Polypeptide *N*-α-Acetylgalactosaminyltransferase Family Members by Glycopeptide Substrate Profiling

Matthew R. Pratt,[2] Howard C. Hang,[2] Kelly G. Ten Hagen,[5] Jason Rarick,[6] Thomas A. Gerken,[6] Lawrence A. Tabak,[5] and Carolyn R. Bertozzi[1,2,3,4,*]

[1]Center for New Directions in Organic Synthesis
[2]Department of Chemistry
[3]Department of Molecular and Cell Biology
[4]Howard Hughes Medical Institute
University of California, Berkeley
Berkeley, California 94720
[5]Section of Biological Chemistry
NIDDK
National Institutes of Health
Bethesda, Maryland 20892
[6]Department of Pediatrics and Biochemistry
Case Western Reserve University School of Medicine
Cleveland, Ohio 44016

Summary

The polypeptide *N*-α-acetylgalactosaminyltransferases (ppGalNAcTs) play a key role in mucin-type *O*-linked glycan biosynthesis by installing the intial GalNAc residue on the protein scaffold. The preferred substrates and functions of the >20 isoforms in mammals are not well understood. However, current data suggest that glycosylated mucin domains are created by the successive, often hierarchical, action of several specific ppGalNAcTs. Herein we analyzed the glycopeptide substrate preferences of several ppGalNAcT family members using a library screening approach. A 56-member glycopeptide library designed to reflect a diversity of glycan clustering was assayed for substrate activity with ppGalNAcT isoforms using an azido-ELISA. The data suggest that the ppGalNAcTs can be classified into at least four types, which working together, are able to produce densely glycosylated mucin glycoproteins.

Introduction

Mucin-type *O*-linked glycosylation is initiated by a family of enzymes known as the polypeptide *N*-α-acetylgalactosaminyltransferases (ppGalNAcTs) [1]. These enzymes utilize UDP-GalNAc as the nucleotide donor substrate to modify serine or threonine residues of proteins trafficking through the secretory pathway (Figure 1A). The product of the reaction, α-GalNAc-Ser/Thr, is termed the "Tn-antigen" and is further elaborated to produce complex carbohydrate structures [2]. Mucin-type *O*-linked glycans are often found in clusters defined as mucin domains. Due to their rigid extended structures, induced by glycosylation, these domains play an important role in determining the overall architecture and cell-surface accessibility of the glycoprotein [3, 4]. Mucin-type glycosylation impacts a variety of biological processes, including cell-cell adhesion [5, 6], host-pathogen interactions [7], and intracellular protein trafficking [8–12]. Furthermore, changes in the elaboration of *O*-linked glycans are often correlated with a variety of diseases, such as Wiskott-Aldrich syndrome [13], hematological disorders [14], and cancer [13, 15, 16].

The first ppGalNAcT activity was reported in 1967 by McQuire and Roseman [17]. One of the family members was then purified in 1986 by Elhammer and Kornfeld [18] and subsequently cloned by Tabak [19], Elhammer [20], Clausen [21], and their coworkers. Analysis of the human genome has revealed 24 putative ppGalNAcTs, with orthologs in higher eukaryotes such as *C. elegans* [22], *Drosophila melanogaster* [1, 23, 24], and the unicellular parasite *T. gondii* [25] (9, 14, and 5 isoforms, respectively) that demonstrate >90% sequence homology across species. To date, 21 distinct isoforms of this enzyme family have been biochemically characterized including the 12 mammalian enzymes: ppGalNAcT-1, -T-2, -T-3, -T-4, -T-5, -T-6, -T-7, -T-9, -T-10, -T-11, -T-12, and -T-13 [1]. Each isoform has a discrete distribution in adult tissues as well as spatial and temporal regulation during development. Some isoforms are found in a wide range of tissues and act upon a large repertoire of substrates in vitro (ppGalNAcT-1 and -T-2), while others are more restricted in expression and substrate preference (ppGalNAcT-5, -T-7, -T-10, and -T-11) [1, 26]. Interestingly, isoforms -T-7 and -T-10 have shown a requirement for previous action by another member of the ppGalNAcT family and have thusly been designated glycopolypeptide GalNAc transferases (gppGalNAcTs). It has been speculated that mucin domains are created by the successive action of several ppGalNAcTs on a single polypeptide substrate (Figure 1B) [1].

Despite the important biological functions of the ppGalNAcTs, their individual roles in mammalian physiology remain largely undefined. Mice deficient in ppGalNAcT-1 [27], -T-4 [1], -T-5 [1], or -T-13 [28] have been reported, but no obvious phenotypic abnormalities were identified. However, disruption of the *pgant35A* gene, an ortholog of -T-11, in *Drosophila melanogaster* has been shown to confer lethality [23, 24]. Collectively, these observations argue for a complex balance between redundancy and hierarchy among the members of the ppGalNAcT family. The elucidation of this complex balance is essential for a complete understanding of mucin-type *O*-linked glycosylation.

Although several ppGalNAcTs have been biochemically characterized in vitro, the details of their (glyco)-peptide substrate requirements and preferences have not been addressed in a systematic manner. This information is critical for insight into their preferred biological substrates and participation in mucin domain assembly. Toward this goal, we have undertaken the in vitro profiling of eight members of the ppGalNAcT family (-T-1, -T-2, -T-3, -T-4, -T-5, -T-7, -T-10, and -T-11) with a synthetic glycopeptide library. The library is based on the peptide sequence termed EA2 (PTTDSTTPAPTTK), a

*Correspondence: crb@berkeley.edu

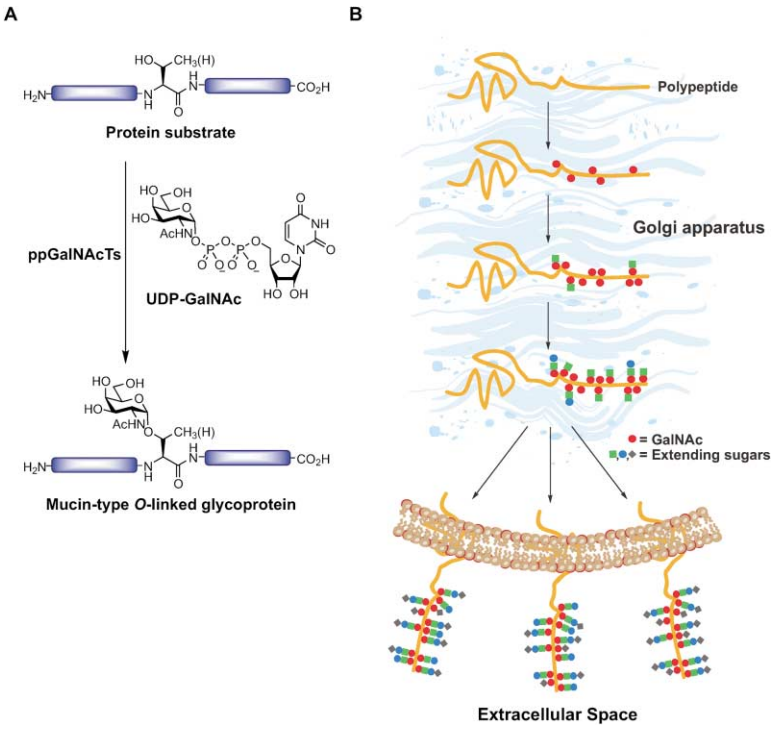

Figure 1. Mucin Biosynthesis

(A) Initiation of mucin-type *O*-linked glycan biosynthesis by ppGalNAcTs.

(B) Biosynthesis of mucin domains by successive action of multiple ppGalNAcTs in the Golgi compartment.

fragment of rat submandibular mucin [29], and contains every possible combination of GalNAc-modified threonine residues up to and including tetra-glycosylated variants (Figure 2). Each peptide in the 56-member library was biotinylated at the C terminus through the introduction of an additional N^{ϵ}-biotin-lysine residue for selective immobilization. Enzymatic glycosylation was quantified by taking advantage of the tolerance of ppGalNAcTs for modified nucleotide sugar donors [30]. An unnatural analog of UDP-GalNAc, UDP-*N*-azidoacetylgalactosamine (UDP-GalNAz), was previously shown to be recognized by the enzymes, allowing for detection of the product by Staudinger ligation with a phosphine-FLAG probe in a high-throughput "azido-ELISA" format (Figure 3A) [30]. An advantage of this method is that the transferred residue, GalNAz, is chemically distinct from the natural GalNAc residues present in the glycopeptide library and can be uniquely detected by virtue of the azido group. The diversity of the library with respect to glycosylation density and pattern was designed to reveal preferences of the ppGalNAcTs for these features and possible hierarchies among the enzymes during the assembly of mucin domains.

Results

The glycopeptide library was prepared by standard solid-phase peptide synthesis (SPPS) using Fmoc-Thr(Ac_3GalNAc)-OH [31, 32] as a building block (Figure 3B). All (glyco)peptides were synthesized from Fmoc-Lys(Biotin)-Wang resin, the biotin moiety allowing for selective immobilization in the ELISA. The peptides were cleaved from resin and deacetylated. The resulting crude library members were analyzed by LCMS for purity and identity (see Supplemental Data). The library members were determined by HPLC analysis to be >85% of the crude product and were therefore used in the azido-ELISA with no further purification. All library members were assayed at the same concentration as determined by UV absorption compared to a pure sample of EA2 peptide.

The azido-ELISA was performed as previously described using recombinant ppGalNAcTs expressed from COS-7 cells as soluble catalytic domains [30, 33]. Transferase reactions were performed under saturating conditions: 500 μM (glyco)peptide, 250 μM UDP-GalNAz for 2 hr. Enzymatic turnover for each library member was normalized to an internal positive control using a known substrate for the enzyme. The parent peptide EA2-K-Biotin (PTTDSTTPAPTTK-K-Biotin) was used as the positive control for ppGalNAcT-1, -T-2, -T-3, -T-4, -T-5, and -T-11. ppGalNAcT-7 and -T-10 are known to require previous glycosylation of their peptide substrates with GalNAc residues and thus required an alternative positive control substrate. For this purpose, we chose the glycopeptide MUC5Ac-3,13-K-Biotin (GTT*PSPVPTTST T*SAP-K-Biotin, T* = α-GalNAc-Thr), which is doubly glycosylated at Thr3 and Thr13 and known to be recognized by the enzymes [34]. The activity from conditioned media derived from mock-transfected COS-7 cells was subtracted as background.

The normalized activities of the ppGalNAcTs with the library members are shown in Figure 4 (structure codes are in Figure 2). The library members are organized from left to right in groups with increasing glycosylation density, and the positive control substrates are at the extreme left and right positions. Several ppGalNAcT isoforms (-T-1, -T-2, and -T-5) demonstrated a strong preference for nonglycosylated or monoglycosylated substrates, acting on the same (glyco)peptide acceptors

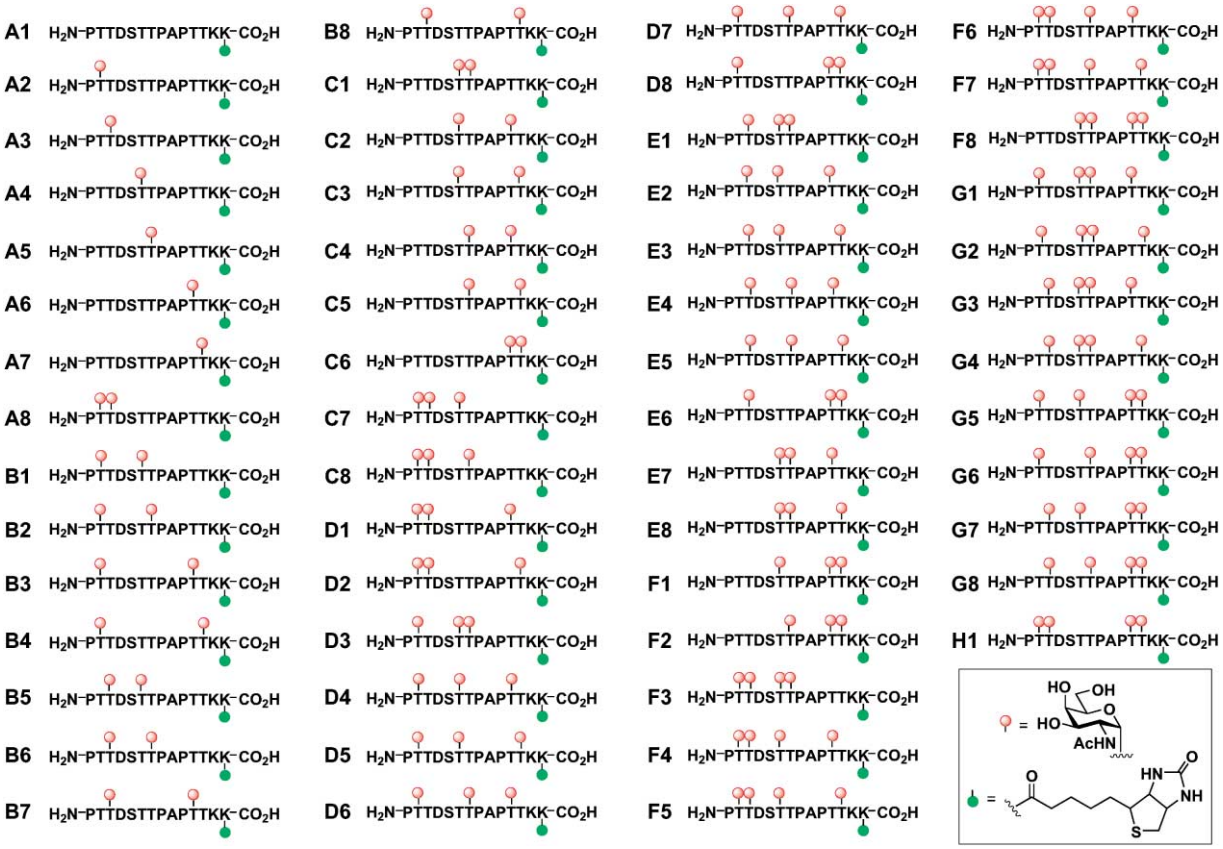

Figure 2. Glycopeptide Library

Biotinylated glycopeptide library derived from EA2 (PTTDSTTPAPTTK), a fragment of rat submandibular mucin. Each library component is named by the code to its left.

although with different optimal sequences. Only ppGalNAcT-1 showed the highest enzymatic turnover with the unglycosylated peptide **A1** (identical to the positive control substrate).

ppGalNAcT-3 and -T-4 prefer mono- and diglycosylated substrates with a small amount of activity observed with triglycosylated acceptors. ppGalNAcT-11 showed discrete specificity for a small number of library members, those with extreme C-terminal glycosylation and no glycosylation toward the peptide's N terminus (compounds **A7** and **C6**, Figure 2). ppGalNAcT-10, a reported gppGalNAcT, was most efficient at modifying tri- and tetraglycosylated substrates and showed little or no activity with monoglycosylated peptides. ppGalNAcT-7 recognized a specific subset of the library: those members that were glycosylated at Thr2 and Thr7 and not at Thr6 (compounds **B2**, **C8**, **D6**, **D7**, **F6**, and **G6**, Figure 2).

In order to determine whether relative activities measured with the azido-ELISA reflect actual difference in substrate kinetics, we selected two pairs of library members and measured K_M and V_{max} values with two enzymes, ppGalNAcT-1 and -T-10. Compounds **A1** and **A3** were compared with ppGalNAcT-1, and MUC5Ac-3,13-K-B and compound **C4** were compared with -T-10 (Supplemental Data and Table 1). Relative V_{max}/K_M values for the substrate pairs were then compared with their relative substrate activities as determined by the azido-ELISA. For both enzymes, kinetic data and azido-ELISA data were in good agreement.

Finally, we determined the sites of glycosylation for two substrates for each enzyme tested. This was accomplished by Edman sequencing of semipurified library members after enzymatic reaction [35–37]. The results of the sequencing are summarized in Figure 5. In the case of ppGalNAcT-1, -T-2, -T-3, -T-4, -T-5, and -T-11, the preferred site of glycosylation was Thr7, regardless of the surrounding glycosylation pattern. Thr7 is located N-terminal to a proline residue and could be in a β-turn motif, consistent with a theorized preferential peptide backbone structure. Another possibility for the partiality toward Thr7 is that the ppGlaNAcTs cannot effectively glycosylate residues at the peptide termini. ppGalNAcT-7 glycosylated Thr6 in both glycopeptides tested. Finally, ppGalNAcT-10 preferred to glycosylate Thr6 in peptide **C4** and Thr3 in the tetraglycopeptide **F8**.

Discussion

The data presented here demonstrate that the ppGalNAcTs have both redundant and hierarchical relationships with respect to mucin domain assembly. The broadly expressed isoforms ppGalNAcT-1 and -T-2 show clear overlap in their substrate preference. This

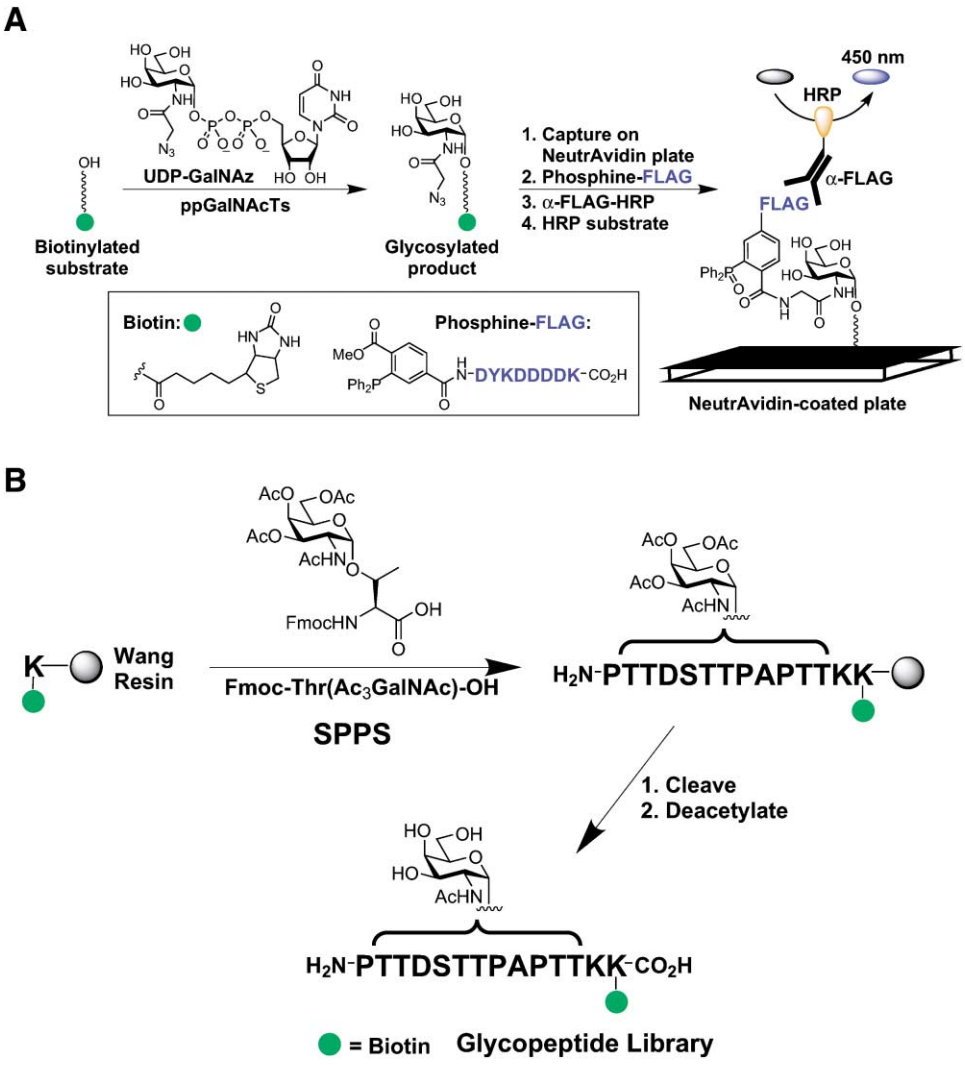

Figure 3. Azido-ELISA and Peptide Synthesis

(A) Schematic diagram of the azido-ELISA.

(B) Fmoc-Thr(Ac_3GalNAc)-OH building block used in solid-phase peptide synthesis (SPPS) to generate the glycopeptide library.

redundancy is consistent with the normal healthy phenotype of mice deficient in -T-1 but in full possession of -T-2 [27]. The similar substrate preference of ppGalNAcT-5 to -T-1 and -T-2, combined with its highly regulated expression during development [26], implies a role in the regulation of glycosylation density on important developmental proteins. Together, ppGalNAcT-1, -T-2, and -T-5 represent the first level of mucin-type *O*-linked glycosylation and could be termed early transferases, i.e., transferases that prefer non- and monoglycosylated peptide substrates. This proposal is supported by the fact that neighboring residue glycosylation inhibits the activity of ppGalNAcT-1 and -T-2 [38]. Their proposed role in mucin assembly is depicted in Figure 6. The apparent redundancy of these transferases may reflect more than a compensatory function. Different levels of expression of these early transferases may be necessary to initiate *O*-linked glycosylation along the entire length of a mucin glycoprotein.

ppGalNAcT-3 and -T-4 represent the next level of transferase in the hierarchical elaboration of mucin domains. These transferases prefer diglycosylated peptides and show lesser but significant activity with triglycosylated substrates. Again, there appears to be redundancy between these two family members, although their preferred substrates are not all identical. This level of transferase may follow the early transferases, thereby increasing the density of glycosylation on the polypeptide backbone (Figure 6). We term these the intermediate ppGalNAcTs.

Of the two glycopeptide transferases analyzed, ppGalNAcT-10 showed partiality for heavily glycosylated substrates of the tri- and even tetraglycosylated variety. Transferases such as -T-10 may complete mucin domain assembly, providing the highly glycosylated polypeptides found in vivo (Figure 6). We term these enzymes the late transferases. Although -T-10 is the only transferase assayed in this study to show this type

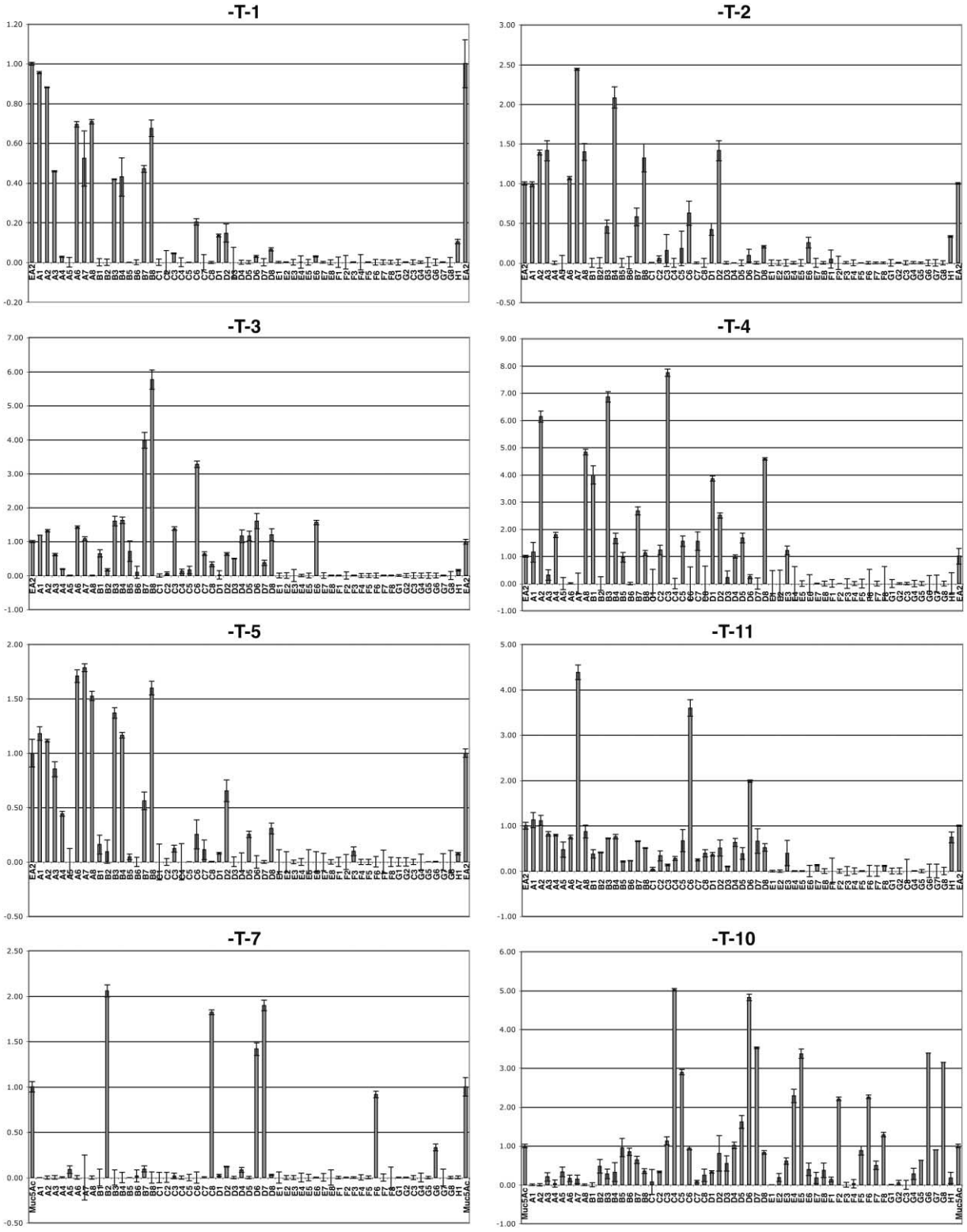

Figure 4. Normalized Activities of the ppGalNAcTs with Library Members

The library members are organized from left to right in groups with increasing glycosylation density, and the positive control substrates are at the extreme left and right positions.

Table 1. Kinetic Data and Comparison of Kinetic Activity with Activity in the Azido-ELISA

		K_M (μM)	V_{max} (μM/min)	V_{max}/K_M (min^{-1})	Relative V_{max}/K_M	ELISA Activity
T-1	PTTDSTTPAPTTK-K-B (**A1**)	121.25 ± 15.17	0.29 ± 0.02	0.0024	1	1
	PTT*DSTTPAPTTK-K-B (**A3**, T* = α-GalNAc-Thr)	108.55 ± 22.2	0.082 ± 0.007	0.00076	0.32	0.46
T-10	GTT*PSPVPTTSTT*SAP-K-B (Muc5Ac-3,13-K-Biotin)	147.94 ± 41.5	0.14 ± 0.02	0.00095	1	1
	PTTDSTT*PAPT*TK-K-B (**C4**)	55.12 ± 12.1	0.60 ± 0.04	0.011	11.6	5

of substrate preference, there may be other family members among the >20 isoforms that contribute to the final stages of mucin domain assembly.

The outliers among the enzymes tested herein are ppGalNAcT-7 and -T-11. ppGalNAcT-7 shows activity on a small subset of library members with a common glycosylation pattern. Analysis of its reaction products demonstrated that ppGalNAcT-7 modifies Thr6 within the context of two consensus GalNAc residues at Thr2 and Thr7. This information may be useful in identifying the biological substrates of -T-7. Combined with its limited expression pattern [26], these data also suggest a discrete biological role for -T-7 in vivo. Gene knockouts have not yet been reported for this enzyme and will be of significant interest. It should be noted that the physiologically relevant substrates for -T-7 may have sequences distinct from the EA2 peptide on which the library was based. Alternative libraries may reveal additional specificities, a subject of future investigation.

ppGalNAcT-11 displays activity with a very limited number of the glycopeptide substrates, much like T-7 but without an obvious consensus. It is possible that -T-11 has a specific substrate sequence requirement that is not satisfied by our library and has no redundancy among the other ppGalNAcTs. Notably, disruption of the *D. melanogaster* -T-11 ortholog *pgant35A* confers lethality [23, 24], consistent with a unique role for this enzyme in vivo.

Significance

Substrate profiling using a glycopeptide library and azido-ELISA revealed both redundancies and hierarchical relationships among the ppGalNAcTs that assemble mucin domains. The data presented here classify the transferases into at least four different types. ppGalNAcT-1, -T-2, and -T-5 are members of the first type, which we have termed early transferases. They are responsible for the beginning stages of mucin domain assembly in which the polypeptide backbone is

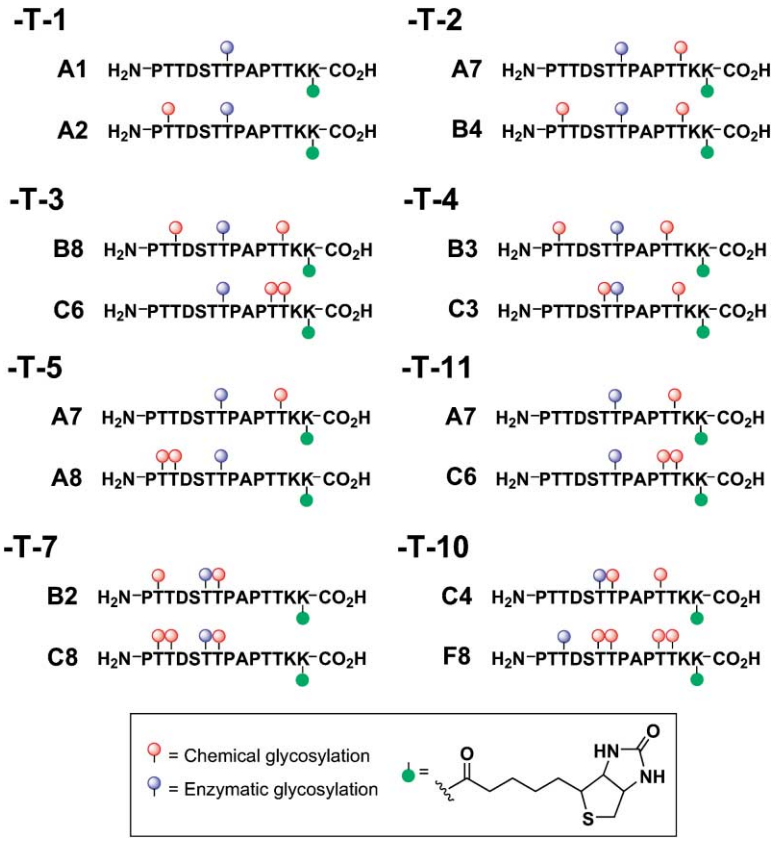

Figure 5. Results of Glycosylation Site Mapping by Edman Degradation

Blue symbols indicate the determined sites of enzymatic glycosylation. Red symbols represent site of glycosylation present in the glycopeptide substrates.

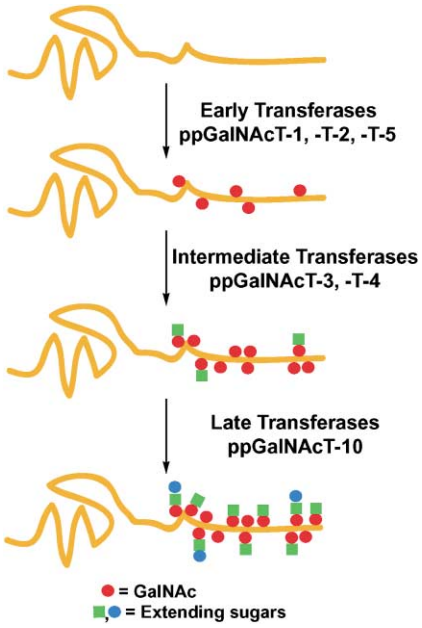

Figure 6. The Three Types of ppGalNAcTs and Their Hierarchy during Mucin Domain Assembly

first glycosylated. The second type of ppGalNAcTs, represented by ppGalNAcT-3 and -T-4, are defined as intermediate transferases. These enzymes increase the density of *O*-linked glycans within the mucin domain. ppGalNAcT-10 was the only family member tested that we define as a late transferase. This enzyme completes mucin domain assembly, affording the densely glycosylated mucin glycoproteins found in vivo. The final type of transferases are those with discrete acceptor substrate requirements, represented by ppGalNAcT-7 and -T-11. These enzymes may modify a small subset of biological targets and play discrete roles in vivo with little redundant activity from the other family members. In summary, the family of ppGalNAcTs comprises a redundant and hierarchical system that ensures proper and full *O*-linked glycosylation patterns necessary for mucin function.

Experimental Procedures

Materials and Methods

N^{α}-Fmoc-amino acids, Fmoc-Lys(Biotin) Wang resin, HOBt and DIC were purchased from Advanced Chemtech. All other chemical reagents were obtained from Sigma or Aldrich and used without further purification. Liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS) was performed on a Hewlett-Packard 1100 mass spectrometer.

(Glyco)peptide Library Synthesis

Fmoc-Thr(Ac_3GalNAc)-OH was synthesized as previously described [31, 32]. The synthesis of all peptides was performed on an Advanced Chemtech APEX-396 96-well automated synthesizer using Fmoc-Lys(Biotin)-Wang resin (0.01 mmol scale). N^{α}-Fmoc-protected amino acids were coupled using DCC/HOBt in NMP. A 5-fold excess of amino acid was used in two sequential coupling steps for each residue. The library was cleaved and deprotected by treatment with 1 ml of 2.5% water, 2.5% triisopropylsilane, and 95% TFA at room temperature for 5 hr. The crude glycopeptides were concentrated and deacetylated with 10% aq. $N_2H_4 \cdot H_2O$ at rt for 20 min. Following concentration, the purity and identity of all library members was ascertained by LC-ESI-MS (see Supplemental Data). The final concentration of the glycopeptides was normalized to 5 mM by comparison to the absorbance of a purified EA2 standard.

ppGalNAcT Reactions for Azido-ELISA

Reactions contained 500 μM biotinylated (glyco)peptide, 250 μM UDP-GalNAz, 10 mM $MnCl_2$, 40 mM sodium cacodylate, 40 mM β-mercaptoethanol, 0.1% Trition X-100 (pH 6.5), and 5 μl of conditioned media from COS-7 cells transiently transfected with plasmids encoding the (g)ppGalNAcT genes [33]. The total reaction volume was 20 μl. All reactions were performed in duplicate. Reactions were allowed to proceed at 37°C for 2 hr and were terminated by addition of 10 μl of 0.1 M EDTA. The azido-ELISA for ppGalNAcTs was performed as previously described [30].

Glycosylation Site Mapping

Standard conditions for (g)ppGalNAcT reactions were as follows. The reaction mixture contained the following components in a final volume of 100 μl: 10 mM $MnCl_2$, 40 mM sodium cacodylate, 40 mM β-mercaptoethanol, 0.1% Trition X-100 (pH 6.5), 25 μl of conditioned media containing recombinant transferase, 500 μM peptide, and 250 μM UDP-GalNAc. The reactions were incubated at 37°C for 16 hr and were terminated by the addition of 800 μl of 20 mM sodium borate, 1 mM EDTA (pH 9.1). The reaction products were passed through AG 1-X8 resin (2 ml), eluted with 4 ml of water, and finally desalted on a C-18 SepPak cartridge. The resulting solution of glycopeptide was then concentrated under vacuum. The reaction products were confirmed by ESI-MS. Pulsed liquid phase Edman degradation amino acid sequencing was performed on an Applied Biosystems Procise 494 protein sequencer (Applied Biosystems, Foster City, CA), and site-specific glycosylation was determined as described previously [35–37]. Improved separation of the first phenylthiohydantoin (PTH)-Thr-*O*-GalNAc diastereomer and the PTH-Ser peak was achieved by reducing the PTH column temperature to 45°C from a normal temperature of 55°C. Corrections for the overlap of the second eluting PTH-Thr-*O*-GalNAc diastereomer with PTH-Thr were made as described [35–37]. Quantification of the extent of glycosylation was performed after eliminating long-range cycle preview and lag for each PTH derivative by a simple base line subtraction approach [35]. Percent glycosylation was determined from the relative picomoles of glycosylated and nonglycosylated Ser or Thr after correcting for the experimentally determined recovery of the individual PTH species [37].

Supplemental Data

LC-ESI-MS traces and mass spectra for all library members and the kinetic analysis of selected peptides are available as Supplemental Data at http://www.chembiol.com/cgi/content/full/11/7/1009/DC1.

Acknowledgments

This research was supported by a grant to C.R.B. from the National Institutes of Health (GM66047). T.A.G. and J.R. were supported by a grant from National Institutes of Health, National Cancer Institute (CA078834). The Center for New Directions in Organic Synthesis is supported by Bristol-Myers Squibb as supporting member and Novartis as sponsoring member.

Received: March 9, 2004
Revised: May 5, 2004
Accepted: May 6, 2004
Published: July 23, 2004

References

1. Ten Hagen, K.G., Fritz, T.A., and Tabak, L.A. (2003). All in the family: the UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases. Glycobiology *13*, 1–16.
2. Brockhausen, I. (2000). Biosynthesis of the *O*-glycan chains of mucins and mucin-type glycoproteins. In Carbohydrates in Chemistry and Biology, Volume 3, B. Ernst, G.W. Hart, and P. Sinay, eds. (New York: Wiley-VCH), pp. 313–328.

3. Strous, G.J., and Dekker, J. (1992). Mucin-type glycoproteins. Crit. Rev. Biochem. Mol. Biol. *27*, 57–92.
4. Carraway, K.L., and Hull, S.R. (1991). Cell surface mucin-type glycoproteins and mucin-like domains. Glycobiology *1*, 131–138.
5. Yeh, J.-C., Hiraoka, N., Petryniak, B., Nakayama, J., Ellies, L.G., Rabuka, D., Hindsgaul, O., Marth, J.D., Lowe, J.B., and Fukuda, M. (2001). Novel sulfated lymphocyte homing receptors and their control by a core1 extension β1,3-*N*-acetylglucosaminyltransferase. Cell *105*, 957–969.
6. Somers, W.S., Tang, J., Shaw, G.D., and Camphausen, R.T. (2000). Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. Cell *103*, 467–479.
7. Hooper, F.L., and Gordon, J.I. (2001). Glycans as legislators of host microbial interactions: spanning the spectrum from symbiosis to pathogenicity. Glycobiology *11*, 1R–10R.
8. Alfalah, M., Jacob, R., Preuss, U., Zimmer, K.-P., Naim, H., and Naim, H.Y. (1999). *O*-linked glycans mediate apical sorting of human intestinal sucrase-isomaltase through association with lipid rafts. Curr. Biol. *9*, 593–596.
9. Alschuler, Y., Kinlough, C.L., Poland, P.A., Bruns, J.B., Apodaca, G., Weisz, O.A., and Houghley, R.P. (2000). Clathrin-mediated endocytosis of MUC1 is modulated by its glycosylation site. Mol. Biol. Cell *11*, 819–831.
10. Breuza, L., Garcia, M., Delgrossi, M.-H., and Le Bivic, A. (2002). Role of the membrane-proximal *O*-glycosylation site in sorting of the human receptor forneurotrophins to the apical membrane of MDCK cells. Exp. Cell Res. *273*, 178–186.
11. Naim, H.Y., Joberty, G., Alfalah, M., and Jacob, R. (1999). Temporal association of the *N*- and *O*-linked glycosylation events and their implication in the polarized sorting of intestinal brush border sucrase-isomaltase, aminopeptidase N, and dipeptidyl peptidase IV. J. Biol. Chem. *274*, 17961–17967.
12. Zheng, X., and Sadler, J.E. (2002). Mucin-like domain of enteropeptidase directs apical targeting in madin-darby canine kidney cells. J. Biol. Chem. *277*, 6858–6863.
13. Tsuboi, S., and Fukuda, M. (2001). Roles of *O*-linked oligosaccharides in immune responses. Bioessays *23*, 46–53.
14. Berger, E.G. (1999). Tn-syndrome. Biochim. Biophys. Acta *1455*, 255–268.
15. Brockhausen, I. (1999). Pathways of *O*-glycan biosynthesis in cancer cells. Biochim. Biophys. Acta *1473*, 67–95.
16. Hollingsworth, M.A., and Swanson, B.J. (2004). Mucins in cancer: protection and control of the cell surface. Nat. Rev. Cancer *4*, 45–60.
17. McGuire, E.J., and Roseman, S. (1967). Enzymatic synthesis of the protein-hexosamine linkage in sheep submaxillary mucin. J. Biol. Chem. *242*, 3745–3747.
18. Elhammer, A., and Kornfeld, S. (1986). Purification and characterization of UDP-*N*-acetylgalactosamine: polypeptide *N*-acetylgalactosaminyltransferase from bovine colostrum and murine lymphoma BW5147 cells. J. Biol. Chem. *261*, 5249–5255.
19. Hagen, F.K., Van Wuyckhuyase, B., and Tabak, L.A. (1993). Purification, cloning, and expression of a bovine UDP-GalNAc: polypeptide *N*-acetyl-galactosaminyltransferase. J. Biol. Chem. *268*, 18960–18965.
20. Homa, F.L., Hollander, T., Lechman, D.J., Tomsen, D.R., and Elhammer, A.P. (1993). Isolation and expression of a cDNA clone encoding a bovine UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltrasferase. J. Biol. Chem. *268*, 12609–12616.
21. Sorensen, T., White, T., Wandall, H.H., Kirstensen, A.K., Roepstorff, P., and Clausen, H. (1995). UDP-*N*-acetyl-α-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase. Identification and separation of two distinct transferase activities. J. Biol. Chem. *270*, 24166–24173.
22. Hagen, F.K., and Nehrke, K. (1998). cDNA cloning and expression of a family of UDP-*N*-acetyl-D-galactosamine:polypepitde *N*-acetylgalactosaminyltransferase sequence homologs from *Caenorhabditis elegans*. J. Biol. Chem. *273*, 8268–8277.
23. Ten Hagen, K.G., and Tran, D.T. (2002). A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is essential for viability in Drosophila melanogaster. J. Biol. Chem. *277*, 22616–22622.
24. Schwientek, T., Bennett, E.P., Flores, C., Thacker, J., Hollmann, M., Reis, C.A., Behrens, J., Mandel, U., Keck, B., Schafer, M.A., et al. (2002). Functional conservation of subfamilies of putative UDP-*N*-acetylgalactosamine:polypeptide *N*-acetylgalactosaminyltransferases in *Drosophila*, *Caenorhabiditis elegans*, and mammals. One subfamily composed of 1(2)35Aa is essential in *Drosophila*. J. Biol. Chem. *277*, 22623–22638.
25. Wojczyk, B.S., Stwora-Wojczyk, M.M., Hagen, F.K., Striepen, B., Hang, H.C., Bertozzi, C.R., Roos, D.S., and Spitalnik, S.L. (2003). cDNA cloning and expression of UDP-*N*-acetyl-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase T1 from *Toxoplasma gondii*. Mol. Biochem. Parasitol. *131*, 93–107.
26. Young, W.W., Jr., Holcomb, D.R., Ten Hagen, K.G., and Tabak, L.A. (2003). Expression of UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase isoforms in murine tissues determined by real-time PCR: a new view of a large family. Glycobiology *13*, 549–557.
27. Westerman, E.L., Ellies, L.G., Hagen, F.K., Marek, J.D., Sutton-Smith, M., Dell, A., Tabak, L.A., and Marth, J.D. (1999). Selective loss of *O*-glycans in mice lacking polypeptide GalNAcT-1. Glycobiology *9*, 1121.
28. Hennet, T., Hagen, F.K., Tabak, L.A., and Marth, J.D. (1995). T-cell-specific deletion of a polypeptide *N*-acetylgalactosaminyl-transferase gene by site-directed recombination. Proc. Natl. Acad. Sci. USA *92*, 12070–12074.
29. Albone, E.F., Hagen, F.K., Van Wuyckhuyase, B.C., and Tabak, L.A. (1994). Molecular cloning of a rat submandibular gland apomucin. J. Biol. Chem. *269*, 16845–16852.
30. Hang, H.C., Yu, C., Pratt, M.R., and Bertozzi, C.R. (2004). Probing glycosyltransferase activities with the staudinger ligation. J. Am. Chem. Soc. *126*, 6–7.
31. Winans, K.A., King, D.S., Rao, V.R., and Bertozzi, C.R. (1999). A chemically synthesized version of the insect antibacterial glycopeptide, diptericin, disrupts bacterial membrane integrity. Biochemistry *38*, 11700–11710.
32. Mitchell, S.A., Pratt, M.R., Hruby, V.J., and Polt, R. (2001). Solid-phase synthesis of *O*-linked glycopeptide analogues of enkephalin. J. Org. Chem. *66*, 2327–2342.
33. Hagen, F.K., Ten Hagen, K.G., Beres, T.M., Balys, M.M., VanWuyckhuyse, B.C., and Tabak, L.A. (1997). cDNA cloning and expression of a novel UDP-*N*-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase. J. Biol. Chem. *272*, 13843–13848.
34. Ten Hagen, K.G., Bedi, G.S., Tetaert, D., Kingsley, P.D., Hagen, F.K., Balys, M.M., Beres, T.M., Degand, P., and Tabak, L.A. (2001). Cloning and characterization of a ninth member of the UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferase faminly, ppGaNTase-T9. J. Biol. Chem. *276*, 17395–17404.
35. Gerken, T.A., Owens, C.L., and Pasumarthy, M. (1997). Determination of the site-specific *O*-glycosylation pattern of the porcine submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:GalNAc transferase peptide binding site. J. Biol. Chem. *272*, 9709–9719.
36. Gerken, T.A., Owens, C.L., and Pasumarthy, M. (1998). Site-specific core 1 *O*-glycosylation pattern of the porcine submaxillary gland mucin tandem repeat. Evidence for the modulation of glycan length by peptide sequence. J. Biol. Chem. *273*, 26580–26588.
37. Gerken, T.A., Gilmore, M., and Zhang, J. (2002). Determination of the site-specific oligosaccharide distribution of the *O*-glycans attached to the porcine submaxillary mucin tandem repeat. Further evidence for the modulation of *O*-glycan side chain structures by peptide sequence. J. Biol. Chem. *277*, 7736–7751.
38. Gerken, T.A., Zhang, J., Levine, J., and Elhammer, A. (2002). Mucin core *O*-glycosylation is modulated by neighboring residue glycosylation status. J. Biol. Chem. *277*, 49850–49862.