

# Isoform-specific O-Glycosylation of Osteopontin and Bone Sialoprotein by Polypeptide N-Acetylgalactosaminyltransferase-1<sup>\*[5]</sup>

Received for publication, June 18, 2009, and in revised form, October 8, 2009. Published, JBC Papers in Press, October 30, 2009, DOI 10.1074/jbc.M109.035436

Hazuki E. Miwa<sup>†1</sup>, Thomas A. Gerken<sup>§</sup>, Oliver Jamison<sup>§</sup>, and Lawrence A. Tabak<sup>‡2</sup>

From the <sup>‡</sup>Section on Biological Chemistry, NIDDK, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892 and the <sup>§</sup>Departments of Pediatrics and Biochemistry, W. A. Bernbaum Center for Cystic Fibrosis Research, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106

Mucin-type O-glycan biosynthesis is regulated by the family of UDP-GalNAc polypeptide:N-acetylgalactosaminyltransferases (ppGalNAcTs) that catalyzes the first step in the pathway by transferring GalNAc to Ser or Thr residues in a protein from the sugar donor UDP-GalNAc. Because not all Ser/Thr residues are glycosylated, rules must exist that signal which hydroxyamino acids acquire sugar. To date, no universal consensus signal has emerged. Therefore, strategies to deduce the subset of proteins that will be glycosylated by distinct ppGalNAcTs must be developed. Mucin-type O-glycoproteins are present abundantly in bone, where we found multiple ppGalNAcT isoforms, including ppGalNAcT-1, to be highly expressed. Thus, we compared glycoproteins expressed in wild-type and *Galnt1*-null mice to identify bone-associated proteins that were glycosylated in a ppGalNAcT-1-dependent manner. A reduction in the apparent molecular masses of two SIBLINGs (small integrin binding ligand N-linked glycoproteins), osteopontin (OPN) and bone sialoprotein (BSP) in the *Galnt1*-null mice relative to those of the wild-type was observed. Several synthetic peptides derived from OPN and BSP sequences were designed to include either known or predicted (*in silico*) glycosylation sites. *In vitro* glycosylation assays of these peptides with recombinant ppGalNAcT-1, ppGalNAcT-2, or ppGalNAcT-3 demonstrated that both SIBLINGs contained Thr/Ser residues that were preferentially glycosylated by ppGalNAcT-1. In addition, lysates prepared from wild-type, but not those from *Galnt1*-null derived osteoblasts, could glycosylate these peptides efficiently, suggesting that OPN and BSP contain sites that are specific for ppGalNAcT-1. Our study presents a novel and systematic approach for identification of isoform-specific substrates of the ppGalNAcT family and suggests ppGalNAcT-1 to be indispensable for O-glycosylation at specific sites of the bone glycoproteins OPN and BSP.

Mucin-type O-glycosylation is a ubiquitous post-translational modification of secreted and membrane-bound proteins.

\* This work was supported, in whole or in part, by the National Institutes of Health intramural program of NIDDK and Grant NCI-RO1 CA-78834 (to T. A. G.).

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. S1–S3 and Tables S1–S4.

<sup>1</sup> Present address: Dept. of Cell Biology, Albert Einstein College of Medicine of Yeshiva University, Bronx, NY 10461.

<sup>2</sup> To whom correspondence should be addressed. Tel.: 301-496-3571; Fax: 301-402-2185; E-mail: [tabak@mail.nih.gov](mailto:tabak@mail.nih.gov).

The initial step of this pathway involves the formation of a protein-sugar linkage between GalNAc and Thr and Ser residues (GalNAc  $\alpha$ 1-O-Ser/Thr) catalyzed by members of the UDP-GalNAc polypeptide:N-acetylgalactosaminyltransferase (ppGalNAcT)<sup>3</sup> family (EC 2.4.1.41). It has been estimated that Thr and Ser residues comprise roughly 14% of the proteome of the superkingdom of eukaryotes (1). Because all hydroxyamino acids are not decorated with O-glycans, rules must exist to specify which Thr and Ser become modified. However, despite intensive study, consensus sequences for the addition of mucin-type O-glycans remain undefined. In part, this is due to the complexity of the ppGalNAc family.

Twenty distinct ppGalNAcT isoforms have been detected in humans and all but 4 have demonstrable catalytic activity.<sup>4</sup> A corresponding set of 18 distinct ppGalNAcTs is expressed in mice. Extensive *in vitro* studies with synthetic peptides as substrates suggest that most of the ppGalNAcT isoforms display both unique and overlapping substrate specificities with varying kinetic properties (2–7). In mice, each ppGalNAcT isoform is differentially expressed both spatially and temporally during development, growth, and maturation (8, 9). Thus, the pattern and the extent of mucin-type O-glycosylation on individual substrates are determined by the substrate specificities of the individual ppGalNAcT isoforms that are expressed in any given cell.

To date there is no universal approach to defining the isoform-specific O-glycosylation of the ppGalNAcTs. In this study, we utilized *Galnt1*-null mice to investigate the ppGalNAcT-1-specific regulation of mucin-type O-glycosylation of bone glycoproteins because the skeleton is known to contain at least two highly expressed glycoproteins that are decorated with mucin-type O-glycans, osteopontin (OPN) and bone sialoprotein (BSP) (10–15). Multiple glycoproteins, including the two small integrin binding ligand N-linked glycoproteins (SIBLINGs), OPN and BSP, were underglycosylated in *Galnt1*-null mice. Several synthetic peptides derived from mouse OPN and BSP sequences containing potential glycosylation sites and/or ppGalNAcT-1 selective sites were designed based on

<sup>3</sup> The abbreviations used are: ppGalNAcT, UDP-GalNAc polypeptide: $\alpha$ -N-acetylgalactosaminyltransferase; GalNAc,  $\alpha$ -N-acetylgalactosamine; OPN, osteopontin; BSP, bone sialoprotein; T1, ppGalNAcT-1; SIBLING, small integrin binding ligand N-linked glycoproteins; HRP, horseradish peroxidase; MOPS, 4-morpholinepropanesulfonic acid.

<sup>4</sup> J. Raman and L. A. Tabak, unpublished observations.

the approach/algorithm developed by Gerken *et al.* (7, 16) and Julenius *et al.* (17). These peptides were then glycosylated *in vitro* to demonstrate that both SIBLINGs contained glycosylation sites that required ppGalNAcT-1 activity for efficient glycosylation. Through systematic analysis of the *O*-glycome of mice genetically engineered to ablate expression of one or more ppGalNAcTs, it should be possible to deconvolute the contributions of each individual isoform of this glycosyltransferase family.

## EXPERIMENTAL PROCEDURES

**Animals**—Generation of *Galnt1*-null mice was described previously (18). The bone RNA and protein extracts were prepared from *Galnt1*-null mice (−/−) backcrossed onto C57BL/6NHsd for at least 6 generations and strain-matched control wild-type (+/+) mice C57BL/6NHsd. For bone marrow cell cultures, wild-type and *Galnt1*-null (+/+ and −/−) mice born from *Galnt1*-heterozygous (+/−) parents in C57BL/6NHsd (Harlan, Indianapolis, IN) backgrounds were used. Genotypes of the mice were checked as previously described (18), except that we replaced primer P2: TTCCAGGACAGCCAGGGCTACACAGAG-3′ with WT3′: 5′-ACTTGGAGCCACTTGTCACAGG-3′, to obtain a wild-type-specific product of 145 bp. All animals were maintained in accordance with the institutional guidelines for the Care and Use of Laboratory Animals. The animal protocol was approved by the Animal Care and Use Committee of the National Institutes of Health, NIDDK, Bethesda, MD.

**Antibodies and Reagents**—Rabbit polyclonal antisera against BSP (LF-84) was kindly provided by Dr. Larry W. Fisher (National Institutes of Health, NIDCR, Bethesda, MD) (19). Rabbit polyclonal anti-bone sialoprotein II antibody (catalog AB1854) was purchased from Millipore (Billerica, MA). The mouse monoclonal antibody against mouse OPN (clone 2A1) was purchased from Santa Cruz (Santa Cruz, CA). Anti-rabbit IgG-HRP and anti-mouse IgG-HRP secondary antibodies were purchased from GE Healthcare. Chemical reagents were obtained from Sigma. UDP-[1-<sup>14</sup>C]GalNAc was purchased from PerkinElmer Life Sciences.

**Quantitative Real Time PCR**—Quantitative real time PCR analysis was conducted by the TaqMan method on an MxPro30000 system (Stratagene, La Jolla, CA). RNA was isolated from calvaria and tibia (free of soft tissues and marrow) pooled from two male animals per sample using TRIzol reagent (Invitrogen) and digested with DNase I (Qiagen, Valencia, CA) to remove genomic DNA. The cDNA was synthesized by using random hexamers and SuperScript III (Invitrogen). For each PCR, cDNA equivalent to 25 ng of RNA was used to determine the expression level of each ppGalNAcT with the TaqMan probe and primers (supplemental Table S1) and Universal PCR master mix (Applied Biosystems, Foster City, CA). Real time PCR was performed in duplicate per each gene expression assay on two to four independently isolated RNA preparations from bone tissues. The expression level of each ppGalNAcT was expressed as the fold-amount relative to the reference gene glyceraldehyde-3-phosphate dehydrogenase as described previously (9).

**Preparation of Bone Extracts**—Bone extracts were prepared from male mouse calvaria and tibia bones independently by an

established protocol (20). Briefly, the respective bones were pulverized in liquid nitrogen with a mortar and pestle. The fine bone powder was suspended in 20 to 30 volumes of 4 M guanidine HCl, 50 mM Tris-HCl (pH 7.4), 1 mM phenylmethylsulfonyl fluoride, 5 mM benzamide HCl, and 0.1 M 6-aminohexanoic acid and stirred at 4 °C for about 72 h. The bone powder was rinsed in the same buffer, and then the buffer was replaced with the above buffer containing 0.5 M EDTA and stirred at 4 °C for 72 h. Final extracts were centrifuged at 5000 × *g* for 20 min and concentrated by Centriplus-20 ( $M_r < 10,000$ ) (Millipore). The crude extracts were dialyzed against 50 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 0.1 mM phenylmethylsulfonyl fluoride and further concentrated to ~1 mg/ml. After the final protein concentration was determined with the Pierce BCA protein assay kit, the extracts were stored at −80 °C. Calvaria and tibia from 10 mice yielded ~0.5 and 1 mg of total protein, respectively.

**SDS-PAGE and Western Blot Analysis**—The bone extracts were separated on 4–12% NuPAGE (Invitrogen) in MOPS buffer (Invitrogen) and stained with GelCode Blue (Coomassie Brilliant Blue) (Pierce) followed by 0.1% Stains-all solution (21). For Western blot analysis, proteins were transferred onto polyvinylidene difluoride membranes, blocked with 5% nonfat dry milk for 1 h, followed by incubation with primary antibodies 2A1 (1/1000) for 1 h, or with anti-bone sialoprotein II (1/1000) or LF-84 (1/1000) overnight. Anti-bone sialoprotein II antibody from Millipore (data not shown) and LF-84 gave comparable staining patterns; Western blots obtained by using LF-84 are shown. Membranes were then incubated with either anti-mouse IgG-HRP or anti-rabbit IgG-HRP secondary antibodies (1/5000) for 1 h. Bands were visualized using either SuperSignal West Pico (Pierce) or ECL Plus (GE Healthcare).

**Glycosidase Treatment of Bone Extracts**—The bone extracts were treated with the enzymatic deglycosylation kit from Prozyme (San Leandro, CA) according to the manufacturer's protocol. Briefly, the bone extracts were first buffer exchanged to water to a final concentration of 1 mg/ml and then reduced/denatured at 99 °C. The denatured bone extracts (10 μg) were digested with various combinations of glycosidases (1.5 milliunits of *N*-glycanase, 1.5 milliunits of sialidase A, 0.6 milliunits of β(1–4)-galactosidase, 12 milliunits of β-*N*-acetylglucosaminidase, and 0.375 milliunits of *O*-glycanase) in 15 μl of reaction buffer for 3 h at 37 °C. Deglycosylation of the bone extract was monitored by GelCode Blue on SDS-PAGE. OPN and BSP deglycosylation was monitored by Western blot analysis as described above.

**Production of Recombinant ppGalNAcTs**—Recombinant ppGalNAcT isoforms were expressed in *Pichia pastoris* as previously described (22). The pKN55–6HmalE-TEV vector was created by cloning the following phosphorylated, annealed oligos into the XhoI/SnaBI site of the pKN55-malE-TEV vector (22): 5′-pTCGAGAAAAGAGAGGCTGAAGCTTACCATCATCATCATCATTAC-3′ and 5′-pGTAATGATGATGATGATGATGGTAAGCTTCAGCCTCTCTTTTC-3′. Residues 42–560 of mouse ppGalNAcT-1 encoding a portion of the stem region and the entire catalytic and lectin domains were cloned as described (22) and inserted between the MluI/AgeI sites of pKN55-N6His-TEV (23). Mouse ppGalNAcT-2 was originally cloned from a mouse spleen cDNA λ library, and a

## ppGalNAcT-1-specific Glycosylation of Bone Glycoproteins

portion of the stem region and the entire catalytic and lectin domains of ppGalNAcT-2 (residues 74–570) were inserted between the MluI/NotI sites of pKN55–6HmaE-TEV. The plasmids were linearized and electroporated into *P. pastoris* strain SMD1168 and selected to create stable transformants as previously described (22). Recombinant soluble mouse ppGalNAcT-1 and ppGalNAcT-2 were purified from the *Pichia* supernatant as described previously (23), except that the His-Trap column-purified transferases were incubated with a half-molar amount of His<sub>6</sub>-tagged TEV protease (23) overnight in 50 mM Tris-HCl (pH 8.0), 25 mM imidazole, 0.2 M NaCl, and 10 mM β-mercaptoethanol (cleavage buffer) to cleave off the tag(s), and then passed through a nickel-nitrilotriacetic acid-agarose (Qiagen) column in cleavage buffer to remove the His<sub>6</sub>-tagged peptide/maltose-binding domains and TEV protease. Glycerol was added to the flow-through fraction to a final concentration of 20%, and these products were used as the source of purified enzymes. The protein concentrations were determined with the Bio-Rad Protein Assay kit (Bio-Rad) according to the manufacturer's protocol. The molar concentrations of ppGalNAcT-1 and ppGalNAcT-2 were determined based on their theoretical molecular masses of 60 and 57 kDa, respectively.

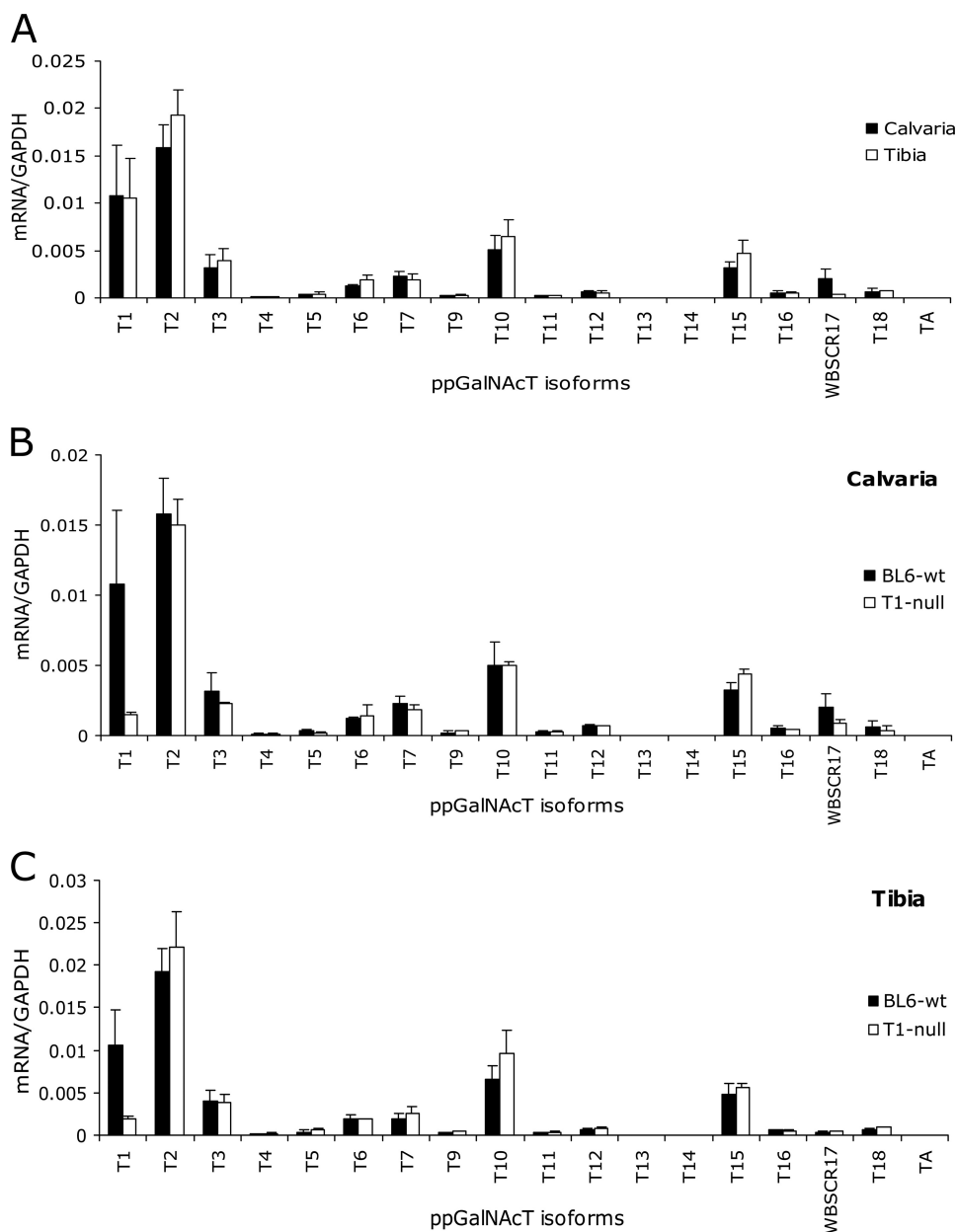
COS-7 cell medium containing secreted recombinant mouse ppGalNAcT-3 was produced as described previously (24). Briefly, COS-7 cells (ATCC, Manassas, VA) were maintained in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen) and transfected with pF1-mT3 (ppGalNAcT-3) or vector without an insert (24). COS-7 cells were plated at  $5 \times 10^4$  cells/cm<sup>2</sup> density on the day before transfection. The cells were then transfected by using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol, and the medium was changed the next day. Two days later, cell media were harvested, cleared, and stored at –80 °C.

**Osteoblast Differentiation**—Bone marrow stromal cells were harvested from the long bones of T1 (+/+ and –/–) male and female littermates by flushing the marrow in minimal essential medium-α (Invitrogen) with 20% fetal bovine serum, 100 units/ml of penicillin, and 100 μg/ml of streptomycin (Invitrogen). The cells were subsequently plated in 12-well plates at  $1.6 \times 10^6$  cells/well in minimal essential medium-α supplemented with 20% fetal bovine serum, 100 units/ml of penicillin, and 100 μg/ml of streptomycin, 50 μg/ml of ascorbic acid, and 10 mM β-glycerophosphate to induce osteoblast differentiation. After 2 days, non-adherent cells were washed off and fresh medium was added. Medium was replaced every 3 days thereafter. After 14 days in culture, cells were harvested or analyzed for various purposes. To extract ppGalNAcT activity, cells were lysed in 50 μl of GALTase lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, and 1× protease inhibitor mixture set III (Calbiochem, Gibbstown, NJ)) per well by scraping followed by sonication at 4 °C. Cell lysates were cleared by centrifugation at  $16,100 \times g$  for 10 min. After the protein concentration in the supernatant was determined with the Pierce BCA protein assay kit, the supernatant was stored at –80 °C. The transcript levels of ppGalNAcT isoforms in RNA isolated with the Nucleospin RNA II kit (Clontech, Mountain View, CA) from bone marrow stromal cell culture were deter-

mined by quantitative real time PCR as described above. The differentiation of bone marrow cells into osteoblast-like cells was monitored by the expression of their alkaline phosphatase activity with the Alkaline Phosphatase kit (Sigma). The formation of mineralization nodules on differentiated osteoblasts in culture was determined by Von Kossa staining with 5.0% silver nitrate solution.

**Designing the Synthetic Peptides Derived from Mouse OPN and BSP**—Because the mucin-type O-glycosylation sites within mouse OPN have been reported, synthetic peptides were designed such that each of the previously determined glycosylation sites was flanked by at least 3 amino acids (25). For mouse BSP, for which O-glycosylation sites have not been mapped, we used both NetOGlyc 3.1 (17) and the enhancement factors obtained by Gerken *et al.* (7, 16) to predict potential glycosylation sites within the protein. For the latter approach, the glycosylation of Thr or Ser residues by ppGalNAcT-1 and ppGalNAcT-2 were predicted for those sites having an enhancement factor product greater than 1.0. Enhancement products were obtained by multiplying together the residue-specific enhancements for positions plus or minus 5 residues from the site of glycosylation. Residues whose enhancements have not been determined at present were taken as one. Synthetic peptides from mouse BSP were then designed in such a way that the Thr or Ser residues, either predicted to be glycosylated by NetOGlyc 3.1 or have an enhancement factor product value larger than 1.0 as described above, were flanked by at least 3 amino acids.

**In Vitro Glycosylation by Purified Recombinant Enzyme**—A panel of synthetic peptides derived from mouse OPN and BSP sequences (Fig. 5) and the control acceptor substrate EA2 (PTTDSTTPAPTTK) (26) and gp120 (RGPGRAFVTIGKIGNMR) (27) peptides were synthesized by the Facility for Biotechnology Resources, Center for Biologics Evaluation and Research (Bethesda, MD). Either 15 min or 1 h *in vitro* glycosylation assays were conducted as follows at 37 °C in 25 μl of reaction mixture. The reaction mixture contained final concentrations of 10 mM MnCl<sub>2</sub>, 40 mM sodium cacodylate (pH 6.6), 40 mM β-mercaptoethanol, 51 μM UDP-[<sup>14</sup>C]GalNAc (7.8 μCi/mmol), 25 mM Tris-HCl (pH 7.2), 75 mM NaCl, 0.5 mM EDTA, 0.5% Triton X-100, and 1 mM synthetic peptides with the stated enzymes. For reactions with purified enzymes from *P. pastoris*, 0.1 pmol of enzyme (6.0 ng of ppGalNAcT-1 and 5.7 ng of ppGalNAcT-2) was used. For ppGalNAcT-3 activity measurement, 6.25 μl of medium from COS-7 cells transfected with pF1-mT3 was used. ppGalNAcT-3-specific activity was determined by subtracting the activity observed with medium from COS-7 cells transfected with the vector alone. For assays with osteoblast cell lysate, 20 μg of total lysate in 12.5 μl of GALTase lysis buffer was used, and the reaction run for 3 h in the presence of 20 mM CDP-choline in addition to the above reaction mixture to inhibit endogenous non-ppGalNAcT-mediated hydrolysis activity. The reactions were terminated by the addition of 75 μl of 0.1% trifluoroacetic acid, and peptides were purified using Sep-Pak Vac 1cc (50 mg) C18 cartridges (Waters Corp., Milford, MA) as described (28). Briefly, after the column was equilibrated in 0.1% trifluoroacetic acid, peptides were bound in 0.1% trifluoroacetic acid, washed with 3 ml of 5% methanol in



**FIGURE 1. ppGalNAcT expression in bone tissues.** The expression of 18 mouse ppGalNAcT transcripts was detected by quantitative real time PCR using cDNA reverse transcribed from the RNA samples isolated from calvaria and tibia bones of 9-week-old C57BL/6 wild-type and *Galnt1*-null male mice. *A*, comparison of ppGalNAcT transcript expression between calvaria (black) and tibia (white) of wild-type mice. *B* and *C*, comparison of ppGalNAcT transcript expression between wild-type (black) and *Galnt1*-null (white) mice of calvaria (*B*) or tibia (*C*). The wild-type values shown in *B* and *C* are identical to those shown in *A*. Error bars represent standard deviation from average values obtained from two to four independent experiments performed in duplicate.

0.1% trifluoroacetic acid, and then eluted with 0.5 ml of 70% methanol in 0.1% trifluoroacetic acid. The eluted peptides were then counted by scintillation counting to determine the incorporation of GalNAc onto the peptides.

**Determination of Glycosylation Sites by Edman Sequencing**—To determine the sites of glycosylation on each synthetic peptide, *in vitro* glycosylations were carried out at 37 °C for >20 h in 100  $\mu$ l of a reaction mixture containing final concentrations of 10 mM MnCl<sub>2</sub>, 40 mM sodium cacodylate (pH 6.6), 40 mM  $\beta$ -mercaptoethanol, 12.5 mM Tris-HCl, 45 mM NaCl, 0.1% Triton X-100, 0.1 mM EDTA, 3.75 mM imidazole, 1 mM peptide, 1

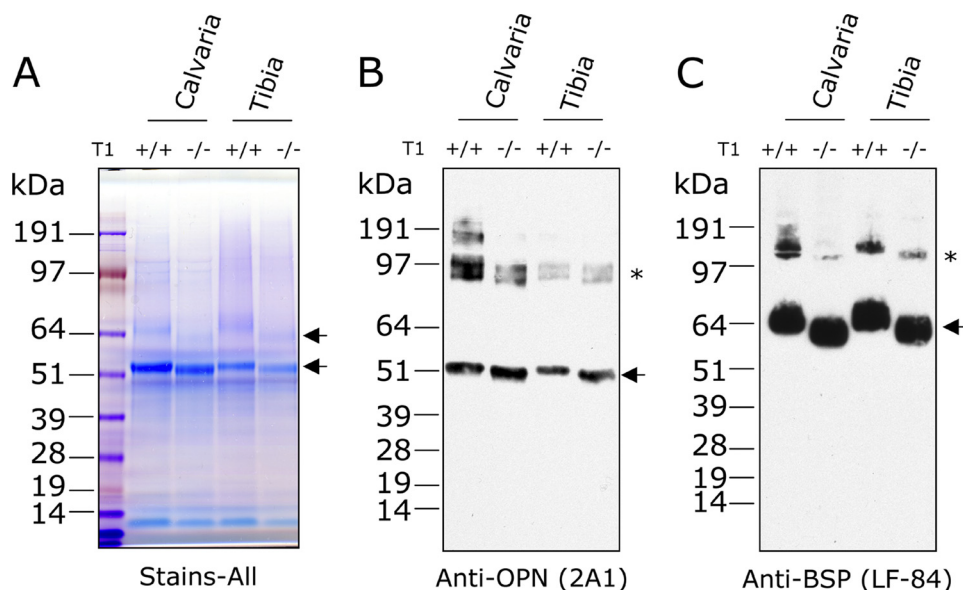
mM UDP-GalNAc, and 200 nM *Pichia*-expressed recombinant mouse ppGalNAcT-1 or ppGalNAcT-2. Prior to sequencing, peptides were chromatographed on columns of Sephadex G-10 (GE Healthcare) in 50 mM acetic acid (pH 4.5) with NH<sub>4</sub>OH to remove the buffer and contaminants (7). GalNAc glycosylation site analysis was performed on a Procise 494 Edman amino acid sequencer (Applied Biosystems). For some peptides, samples were passed through an anion exchange column before being subjected to gel filtration on columns of Sephadex G-10. Integrated peaks for the glycosylated and non-glycosylated Ser and Thr residues were further corrected for lag and preview as described (29, 30).

## RESULTS

**Expression of ppGalNAcT Family Members in Murine Bones**—Bone extracellular matrix is rich in glycoproteins that are produced primarily by bone resident cells, the osteoblasts and osteocytes. To identify ppGalNAcT isoforms that are potentially involved in mucin O-glycosylation of these glycoproteins, we isolated total RNA from two types of bone tissues (*i.e.* calvaria (flat (intramembranous) bone) and tibia (long bone)) of 9-week-old C57BL/6 mice, and performed quantitative real time PCR analysis. We analyzed the two bone tissue types, separately, because they are developed by distinctive mechanisms and may express different sets of ppGalNAcTs. Flat bones are formed by intramembranous ossification, where bone tissues are directly deposited onto condensed mesenchyme, whereas long bones

are developed by endochondral (intracartilaginous) ossification, in which bone replaces the cartilage template. In fact, the expression patterns of ppGalNAcTs in calvaria- and tibia-derived RNA were very similar, except that RNA isolated from calvaria exhibited a slightly higher level of WBSR17 expression (Fig. 1*A*). Overall, in both bone types, moderate to high levels of ppGalNAcT-2 (T2), ppGalNAcT-1 (T1), ppGalNAcT-10 (T10), ppGalNAcT-15 (T15)<sup>4</sup> (formally designated as ppGalNAcT-L2), and ppGalNAcT-3 (T3) were observed. Lower levels of the ppGalNAcT-6 (T6) and ppGalNAcT-7 (T7) transcripts were also detected.

## ppGalNAcT-1-specific Glycosylation of Bone Glycoproteins



**FIGURE 2. OPN and BSP are potential *in vivo* substrates of ppGalNAcT-1.** Total bone extracts prepared from calvaria and tibia of C57BL/6 wild-type (+/+) and *Galnt1*-null (-/-) mice were separated on 4–12% NuPAGE gel and stained with (A) Stains-all, or transferred onto polyvinylidene difluoride membrane for immunoblot analysis against OPN (2A1) (B) and BSP (LF-84) (C). Protein species that migrated differently between the wild-type and *Galnt1*-null mice are marked with arrows. Bands marked with asterisks in B and C are multimers likely formed by transglutamination.

Because ppGalNAcT-1 is highly expressed in bone, we hypothesized that it might be involved in bone protein glycosylation and that a lack of ppGalNAcT-1 activity might result in underglycosylation of several glycoproteins expressed in the skeleton. The generation of *Galnt1*-null mice has been previously reported (18). This mouse lacks *Galnt1* transcript, measurable ppGalNAcT-1 activity, and displays defects in the immune system and blood clotting, but is viable and fertile (18), making it an ideal model to investigate ppGalNAcT-1-mediated glycosylation in adult mice.

As expected, ppGalNAcT-1 expression was essentially absent in *Galnt1*-null mouse calvaria and tibia relative to strain-matched C57BL/6 wild-type (Fig. 1, B and C). Some transcripts detected in *Galnt1*-null mice are likely formed by an occasional splice event. There were no detectable changes in the expression of other tested ppGalNAcT transcripts in bone tissues derived from the null animals.

**Characterization of Bone Mineral Compartment Proteins from Wild-type and *Galnt1*-null Mice**—To identify potential substrates of ppGalNAcT-1 in bone, we extracted proteins in the mineralized matrix from the C57BL/6 wild-type and *Galnt1*-null mice and analyzed them by SDS-PAGE. GelCode blue (Coomassie Brilliant Blue) staining of total bone extracts from calvaria and tibia of wild-type and *Galnt1*-null mice suggested that approximately equivalent extraction efficiency was achieved (Figs. 2A and supplemental S1).

At least four proteins present in the calvaria and tibia extracts migrated differently between wild-type and *Galnt1*-null mice (arrows, Figs. 2A and supplemental S1). A similar difference was observed in bone extracts of *Galnt1*-null and wild-type control littermates (data not shown). Two differentially migrating proteins were stained with Stains-all, which detects acidic proteins effectively (arrows, Fig. 2A). Based on

apparent molecular masses and their acidic nature, it appeared likely that these proteins were OPN and BSP, which are abundantly expressed in bone, highly acidic, and have apparent molecular masses of 50–55 and ~55–65 kDa, respectively. Western blot analysis with antibodies directed against OPN and BSP (Fig. 2, B and C) demonstrated that the apparent molecular mass of each protein expressed in *Galnt1*-null was smaller compared with wild-type, suggesting that they are *in vivo* substrates of ppGalNAcT-1.

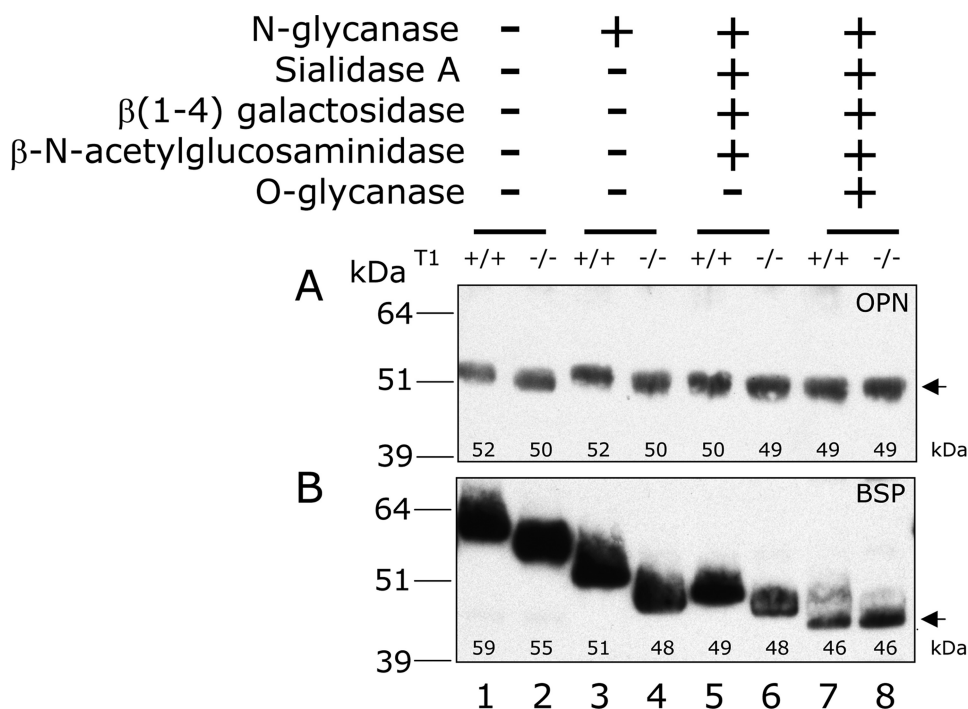
**Characterization of Bone Sialoprotein and Osteopontin Glycosylation**—OPN and BSP are members of the SIBLING family of proteins that are expressed by bone resident cells and play an important role in normal bone development, mineralization, and remodeling (31–33).

To determine whether the shifts in the apparent molecular masses of *Galnt1*-null-derived OPN and BSP were consistent with changes in mucin-type O-glycosylation, we digested bone extracts with a series of glycosidases to sequentially remove both the N-linked and mucin-type O-linked glycans. The changes in the apparent molecular masses were then monitored by Western blot analysis (Fig. 3, A and B).

To characterize the mucin-type O-glycans on each protein, the O-glycans were removed by multiple glycosidases in the presence of N-glycanase. The differences in the apparent molecular masses of OPN and BSP treated with N-glycanase alone or those treated with N-glycanase, sialidase A,  $\beta$ (1–4)-galactosidase, and  $\beta$ -N-acetylglucosaminidase with or without O-glycanase were also compared. In the absence of O-glycanase, the apparent molecular masses of wild-type and *Galnt1*-null-derived BSP and OPN remained different (Fig. 3, A and B, lanes 5 and 6). Following O-glycanase treatment to remove the core-1 moiety, however, the apparent molecular masses of both OPN and BSP became indistinguishable between wild-type and *Galnt1*-null mice (Fig. 3, A and B, lanes 7 and 8). This suggests that both proteins contain ppGalNAcT-1-mediated O-glycans that are further extended to form the core-1 moiety.

N-Glycans were absent in OPN from both wild-type and *Galnt1*-null mice based on its resistance to N-glycanase treatment, even though one potential N-linked glycosylation site is present (Fig. 3A, lanes 1–4). BSP, on the other hand, has four potential N-linked glycosylation sites. The apparent molecular mass of both wild-type and *Galnt1*-null-derived-BSP was reduced by ~7–8 kDa after N-glycanase treatment (Fig. 3B, lanes 1–4).

We estimated that wild-type mouse BSP contained ~5 kDa of O-glycans, whereas BSP from *Galnt1*-null mice O-glycans



**FIGURE 3. Characterization of wild-type and *Galnt1*-null derived OPN and BSP glycosylation.** Bone extracts prepared from C57BL/6 wild-type (+/+) (lanes 1, 3, 5, and 7) and *Galnt1*-null (-/-) (lanes 2, 4, 6, and 8) mice were either incubated without enzymes (lanes 1 and 2), or deglycosylated with *N*-glycanase (lanes 3 and 4), *N*-glycanase, sialidase A,  $\beta(1-4)$ -galactosidase, and  $\beta$ -*N*-acetylglucosaminidase (lanes 5 and 6), or with *N*-glycanase, sialidase A,  $\beta(1-4)$ -galactosidase,  $\beta$ -*N*-acetylglucosaminidase, and *O*-glycanase (lanes 7 and 8). Deglycosylated OPN and BSP were monitored by immunoblotting with antibodies against OPN (A) or BSP (B) as described under "Experimental Procedures." The apparent molecular masses shown in each lane are based on the leading edge and is rounded to a significant figure. Note that some bands with equal molecular masses are apparently, but not significantly different.

were 2 kDa, based on the differences in apparent molecular masses from the *N*-glycanase-treated molecules and those additionally treated with sialidase A,  $\beta(1-4)$ -galactosidase,  $\beta$ -*N*-acetylglucosaminidase, and *O*-glycanase (compare Fig. 3B, lanes 3 and 7, and lanes 4 and 8). This suggests that the synthesis of approximately half of the *O*-glycans on BSP is dependent upon initiation by ppGalNAcT-1. Similarly, the approximate molecular mass of *O*-glycans on wild-type OPN was ~3 kDa, whereas those in OPN from *Galnt1*-null mice were ~1–1.5 kDa (compare Fig. 3A, lanes 3 and 7, and lanes 4 and 8). Taken together, this suggests that ppGalNAcT-1 is a major ppGalNAcT isoform, but not the only one, participating in glycosylation of both OPN and BSP in murine bone.

**Synthetic Peptides from OPN and BSP That Contain Potential Mucin-type O-Glycosylation Sites**—The reduced molecular masses of OPN and BSP in *Galnt1*-null mice indicated that OPN and BSP likely contain amino acid sequence motifs that are highly specific for glycosylation by ppGalNAcT-1. Thus, to identify such motifs within OPN and BSP, we performed *in vitro* glycosylation reactions against a panel of synthetic peptides derived from mouse OPN and BSP sequences with ppGalNAcT-1 as well as ppGalNAcT-2 and ppGalNAcT-3. These three isoforms are expressed well in bone (Fig. 1) and most importantly have each been shown to exhibit both some unique and some overlapping substrate specificities (3, 7). In contrast, although ppGalNAcT-10 and ppGalNAcT-15 are also highly expressed in bone, ppGalNAcT-10 preferably displays catalytic activity toward glycopeptides (34), whereas we

have detected no catalytic activity of ppGalNAcT-15 against an array of peptides available in our laboratory,<sup>5</sup> so these two transferases were not included in this analysis.

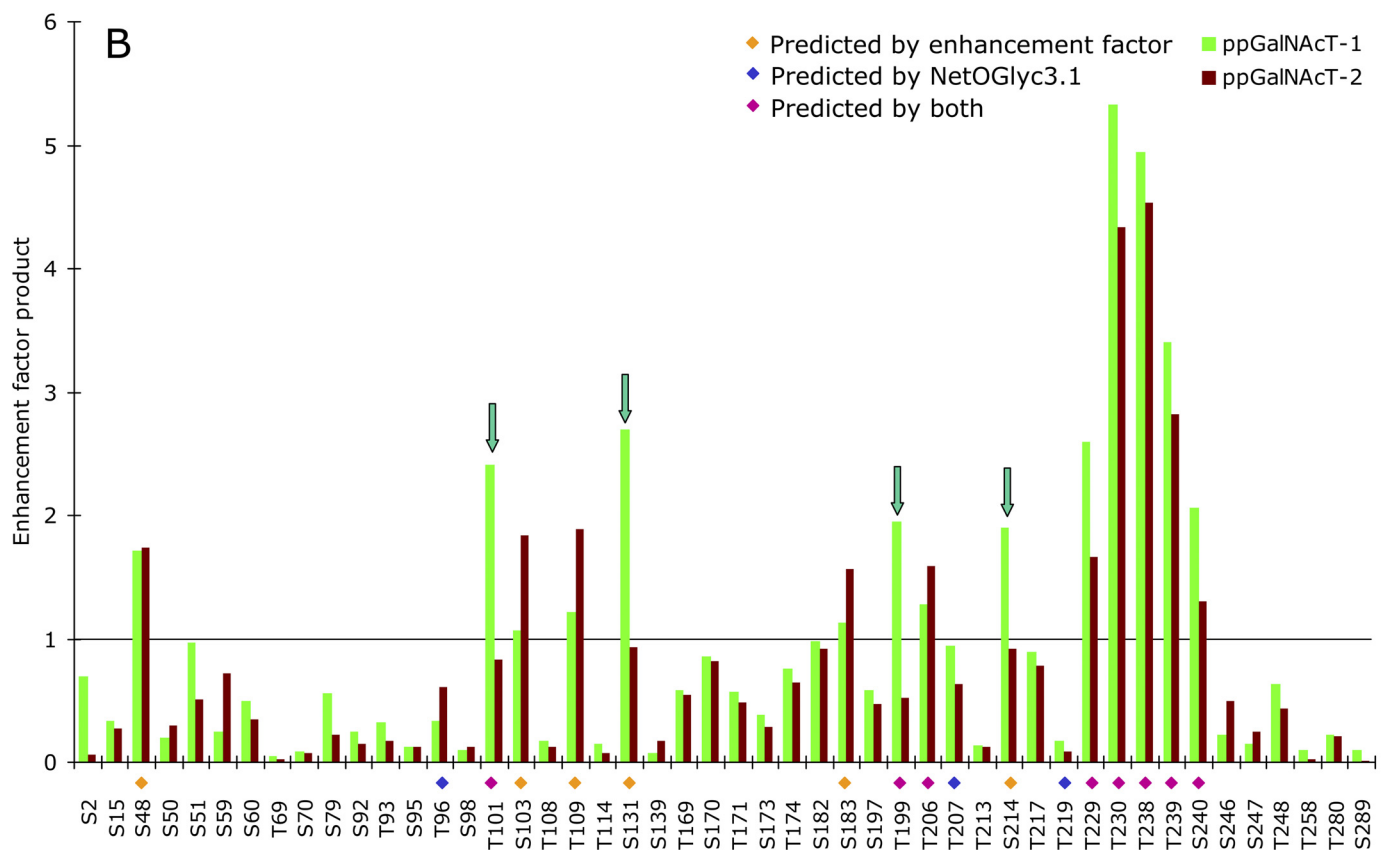
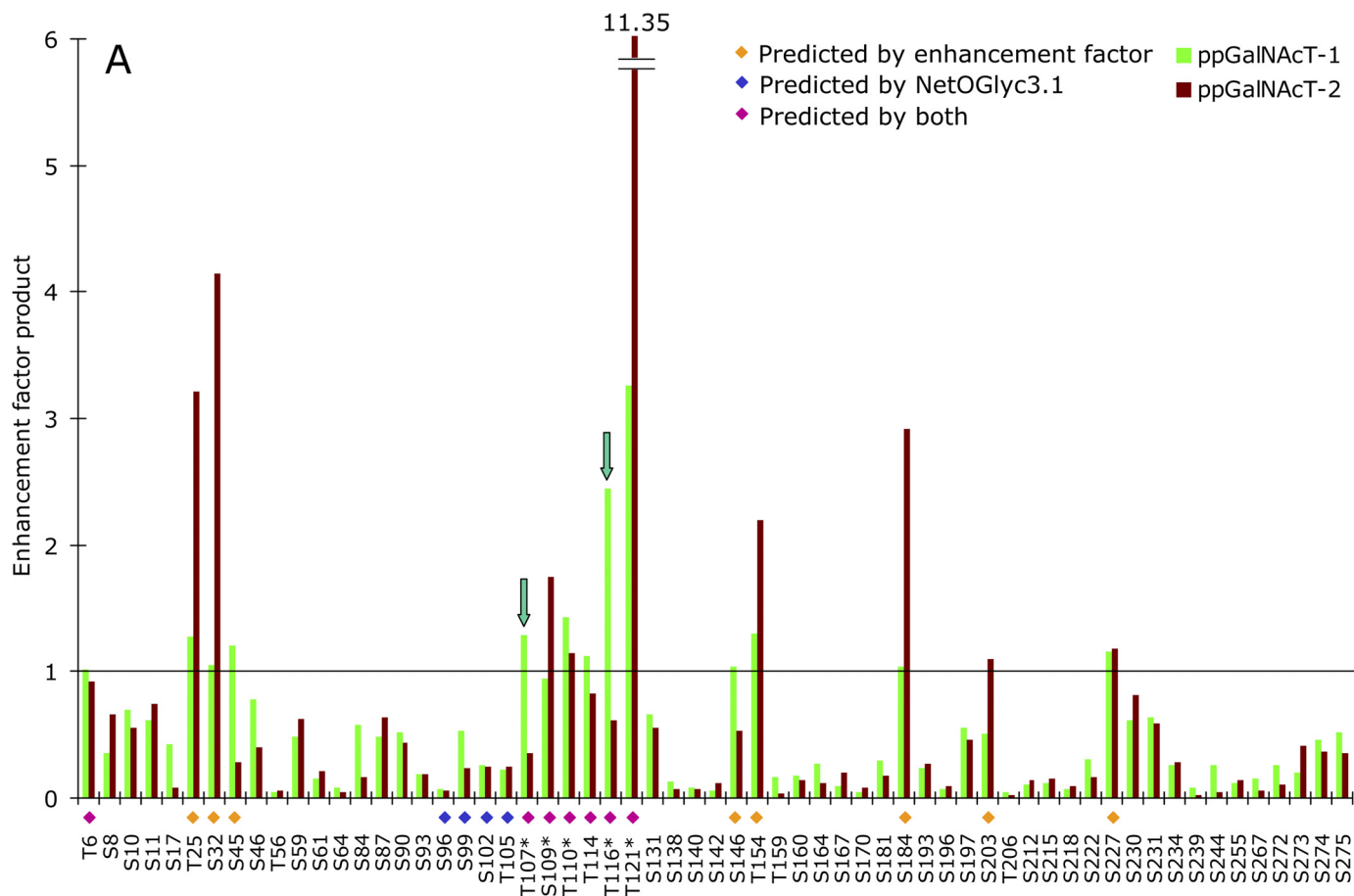
The *O*-glycosylation sites of OPN have been determined in multiple species, including mouse (25), and have been shown to cluster in the region N-terminal to the Arg-Gly-Asp (RGD) integrin binding motif (Fig. 4A, asterisks). Therefore, we employed three synthetic peptides spanning these Thr and Ser residues to determine the glycosylation sites utilized by ppGalNAcT-1 within the OPN sequence (Fig. 5A). Additionally, we also used the preferences reported by Gerken *et al.* (16) to predict potential ppGalNAcT-1 glycosylation sites that would not be glycosylated by ppGalNAcT-2 based on their flanking sequences (16). This approach predicted that Thr-107 and Thr-116 would serve as better substrates for ppGalNAcT-1 compared with ppGalNAcT-2 (arrows in Figs. 4A and 5A). For comparison, NetOGlyc-predicted sites are also shown in Fig. 4A.

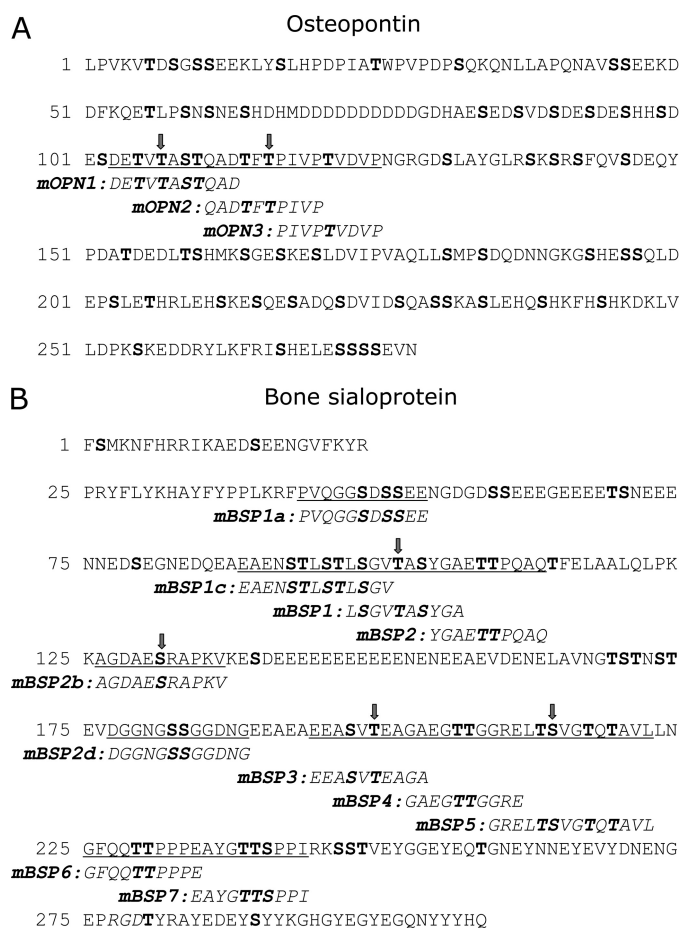
The *O*-glycosylation sites of mouse BSP have not been determined. Some information is available for *O*-glycosylation sites in human BSP (11, 35), but many glycosylation sites are not conserved between the human and mouse sequences (supplemental Fig. S2). Thus, to systematically identify potential acceptor residues of ppGalNAcT-1 in mouse BSP, the approach developed by Gerken *et al.* (16) and the NetOGlyc 3.1 program (17) were used to predict the potential *O*-glycosylation sites within the BSP sequence (Fig. 4B). The former predicted that Thr-101, Ser-131, Thr-199, and Ser-214 would show preference for ppGalNAcT-1 over ppGalNAcT-2 (arrows in Figs. 4B and 5B). Peptides containing these potential glycosylation sites plus 3 to 4 flanking amino acid residues were used for *in vitro* glycosylation assays (Fig. 5B).

**Specific Activity of ppGalNAcT-1, ppGalNAcT-2, and ppGalNAcT-3 against OPN and BSP Peptides**—We conducted *in vitro* glycosylation assays of the peptides in Fig. 5 using purified recombinant ppGalNAcT-1 and ppGalNAcT-2 expressed in *P. pastoris*. As shown in Fig. 6, A and B, and supplemental Table S2, ppGalNAcT-1 could glycosylate a wider range of peptides derived from OPN and BSP compared with ppGalNAcT-2. In particular, we detected the activity of ppGalNAcT-1 on mOPN1, mOPN2, mBSP1c, mBSP1, mBSP2, mBSP3, and mBSP5 peptides, but no activity was observed for ppGalNAcT-2 with these peptides under the same condi-

<sup>5</sup> H. E. Miwa, J. Raman, and L. A. Tabak, unpublished observations.

*ppGalNAcT-1-specific Glycosylation of Bone Glycoproteins*





**FIGURE 5. Sequence and synthetic peptides derived from mouse OPN and BSP to identify ppGalNAcT-1-dependent glycosylation sites.** *A*, synthetic peptides were derived from mouse OPN based on the glycosylation sites previously described in Christensen *et al.* (25). *B*, synthetic peptides were designed from a mouse BSP sequence containing Thr and Ser residues predicted to be glycosylated as depicted in Fig. 4*B*. All of the Thr and Ser residues are shown in bold. Letters with arrows are potential ppGalNAcT-1-selective glycosylation sites.

tions. In contrast, the mOPN3 peptide was preferentially O-glycosylated by ppGalNAcT-2, whereas O-glycosylation by ppGalNAcT-1 was negligible. Glycosylation of mBSP6 and mBSP7 could be efficiently carried out by both ppGalNAcT-1 and ppGalNAcT-2 under these conditions.

Because many of these peptides contain multiple Thr and Ser residues, we identified the sites of glycosylation mediated by ppGalNAcT-1 and ppGalNAcT-2 by Edman sequencing. To obtain sufficient glycopeptides for the analysis, the reaction was allowed to take place overnight with excess enzymes and equal concentrations of peptide and donor substrate UDP-GalNAc. As shown in Fig. 7, the majority of the peptides were glycosylated at only one site. One exception was mBSP7, in which significant GalNAc incorporation was observed at 2 Thr residues by both enzymes. Interestingly, however, the two enzymes displayed differential preferences toward one or the other Thr residues (Fig. 7).

We also investigated whether ppGalNAcT-3 could glycosylate any of the peptides described above by *in vitro* assays using a secreted form of recombinant ppGalNAcT-3 expressed in COS-7 cells. As shown in Fig. 6*C* and supplemental Table S3, recombinant ppGalNAcT-3 glycosylated a subset of peptides (*i.e.* mBSP1, mBSP2, mBSP6, and mBSP7) that were also glycosylated by ppGalNAcT-1. Some of these were poor substrates for ppGalNAcT-2 (Fig. 6*B*). Therefore, ppGalNAcT-3 might be better able than ppGalNAcT-2 to compensate for the absence of ppGalNAcT-1 activity.

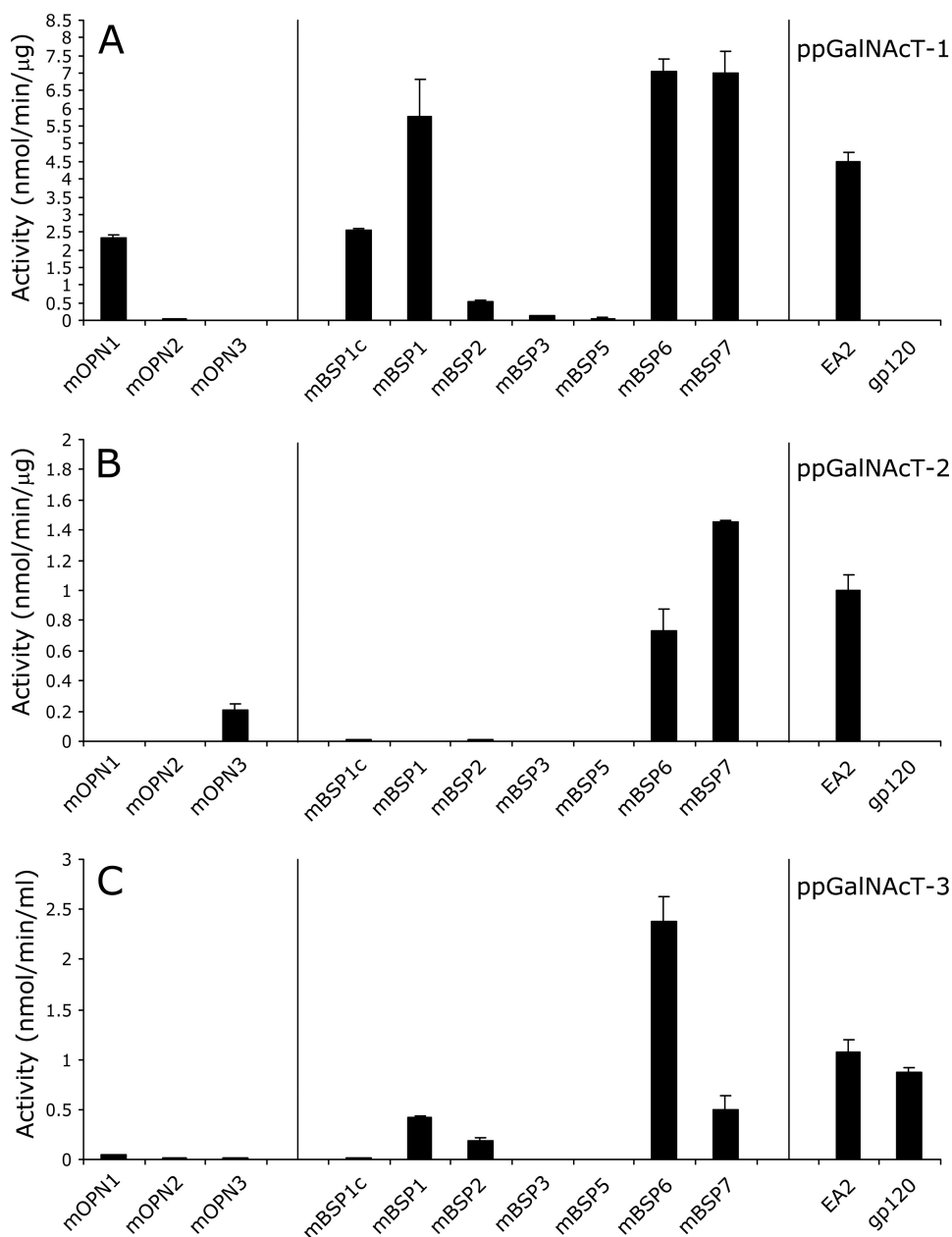
*In Vitro Glycosylation of OPN and BSP Peptides by Extracts of Cultured Osteoblasts*—In bone, osteoblasts are the major cell types that express OPN and BSP. We thus compared the GalNAc transfer activity against the selected peptides using lysates prepared from either wild-type or *Galnt1*-null osteoblasts. Osteoblast-like cells were prepared by growing bone marrow stromal cells from wild-type and *Galnt1*-null mice in osteogenic medium for 14 days. The osteoblast-like phenotype of these cells was confirmed by their expression of alkaline phosphatase activity and formation of mineralized nodules (supplemental Fig. S3*A*). The transcripts of ppGalNAcT isoforms expressed in these cells were determined by quantitative real time PCR analysis to confirm the expression of ppGalNAcT-1, ppGalNAcT-2, and ppGalNAcT-3 as well as other ppGalNAcTs in wild-type osteoblasts (supplemental Fig. S3*B*). As expected, expression of ppGalNAcT-1 was diminished in *Galnt1*-null osteoblasts, whereas the expression levels of other isoforms were similar to those of wild-type. Osteoblast lysates containing these multiple ppGalNAcTs were then used to glycosylate a panel of synthetic peptides (Fig. 8). The average activities detected from independent preparation of osteoblast lysates prepared from wild-type and *Galnt1*-null osteoblasts are shown in Fig. 8*A*. The activity detected in the lysate derived from the *Galnt1*-null osteoblasts was normalized to that of the respective littermate controls as shown in Fig. 8*B*.

The ppGalNAcT activities in lysates were determined by using EA2 and gp120 peptides, which are non-isoform selective and ppGalNAcT-3/T-6 selective substrates (5), respectively. Although the activity in osteoblast lysates derived from the *Galnt1*-null osteoblasts exhibited reduced ppGalNAcT activity against the EA2 control peptide compared with osteoblast lysates prepared from wild-type mice, a considerable level of ppGalNAcT activity remained because of expression of other members of the ppGalNAcT family. As a control, we verified that the ppGalNAcT activity against the gp120 peptide (a poor ppGalNAcT-1 substrate) in both lysates was similar, as expected. Therefore, the observed reduction of ppGalNAcT activity against the EA2 peptide in osteoblast lysates derived from null mice can be attributed to a lack of ppGalNAcT-1 activity. With these lysates, we found that the mBSP6 and mBSP7 peptides (which were efficiently glycosylated *in vitro* by ppGalNAcT-1, ppGalNAcT-2, and ppGalNAcT-3) were glyco-

**FIGURE 4. Prediction of sites of glycosylation by ppGalNAcT-1 and ppGalNAcT-2 in mouse OPN and BSP.** Predicted glycosylation sites of OPN (*A*) and BSP (*B*). Enhancement factor products of ppGalNAcT-1 and ppGalNAcT-2 obtained by the approach of Gerken *et al.* (16) are shown in green and red, respectively. Sites that are found to be glycosylated in mouse OPN from MC3T3-E1 osteoblastic cells are marked with asterisks (25). Sites predicted to be glycosylated by the enhancement factor products, NetOGlyc 3.1 program, or both are marked orange, blue, or purple, respectively, above the residue numbers. Arrows designate potential ppGalNAcT-1-selective glycosylation sites.



## ppGalNAcT-1-specific Glycosylation of Bone Glycoproteins



**FIGURE 6. Substrate preference displayed by ppGalNAcT-1, T-2, and T-3 against OPN and BSP peptides.** The acceptor substrate specificities of ppGalNAcT-1 (A), ppGalNAcT-2 (B), and ppGalNAcT-3 (C) against OPN and BSP peptides and control peptides EA2 (for T1, T2, and T3) and gp120 (for T3) were determined under the conditions described under "Experimental Procedures." No glycosylation was observed in mBSP1a, mBSP2b, mBSP2d, and mBSP4 peptides by any of the three transferases and thus is not depicted. Error bars represent S.D. from average values obtained from two independent experiments performed in duplicate.

sylated by both lysates, although the activity was reduced in the lysates derived from *Galnt1*-null osteoblasts. The mOPN3 peptide, which showed an *in vitro* preference for ppGalNAcT-2 (Fig. 6B), was glycosylated equally by both lysates, confirming that it is a poor substrate of ppGalNAcT-1. In contrast, glycosylation of the mOPN1, mBSP1c, and mBSP1 peptides by osteoblast lysates derived from *Galnt1*-null osteoblasts were reduced by >90% compared with wild-type (Fig. 8B), consistent with their preferred *in vitro* usage by ppGalNAcT-1 (Fig. 6A). Glycosylation of mBSP3 and mBSP5 (which were weakly glycosylated *in vitro* by ppGalNAcT-1) was negligible in osteoblast lysates derived from *Galnt1*-null osteoblasts, even though some

weak activity was observed in lysates prepared from the wild-type mice. Together, these results provide strong evidence that both OPN and BSP contain Thr/Ser residues for which glycosylation is highly dependent on the presence of ppGalNAcT-1 activity.

## DISCUSSION

The full extent of the *O*-glycome remains undefined. The methods to unambiguously map sites of *O*-glycan attachment are highly specialized and relatively insensitive (7, 36, 37). Thus, the few experimentally mapped *O*-glycosylation sites were most recently estimated to be 236 among 50 human proteins (36). The limited data base of verified *O*-glycosylation sites has hampered efforts to develop predictive algorithms. Despite intensive effort, no consensus sequence has emerged although recent work employing different statistical learning methods have reported accuracies and sensitivities as high as 90 and 80%, respectively (38–40). However, none of these approaches were designed to deconvolute the respective contributions of the different family members of the ppGalNAcTs. Indeed, many mucin-type *O*-glycosylated proteins have multiple *O*-glycosylation sites, and thus may be glycosylated by multiple members of the ppGalNAcT family. Evidence from recent studies of mutant organisms and human disease suggests that individual ppGalNAcT isoforms may have unique biological functions (18, 41–44). Thus, we have developed an approach to identify isoform-specific substrates

to better understand how mucin-type *O*-glycosylation is controlled in an isoform-specific manner.

Our strategy was to make use of a mouse model that was genetically engineered to lack a specific ppGalNAcT to identify substrates that appeared to change in size, presumably due to a loss of *O*-glycans. By SDS-PAGE/Western blot analysis, we observed significant molecular mass reductions in two major bone glycoproteins, OPN and BSP, in *Galnt1*-null mice. Subsequent glycosidase treatment confirmed that the *O*-glycosylation status of these glycoproteins was indeed modified in these mice. In common with many mucin-type glycoproteins, both OPN and BSP have multiple *O*-glycosylation sites. Although

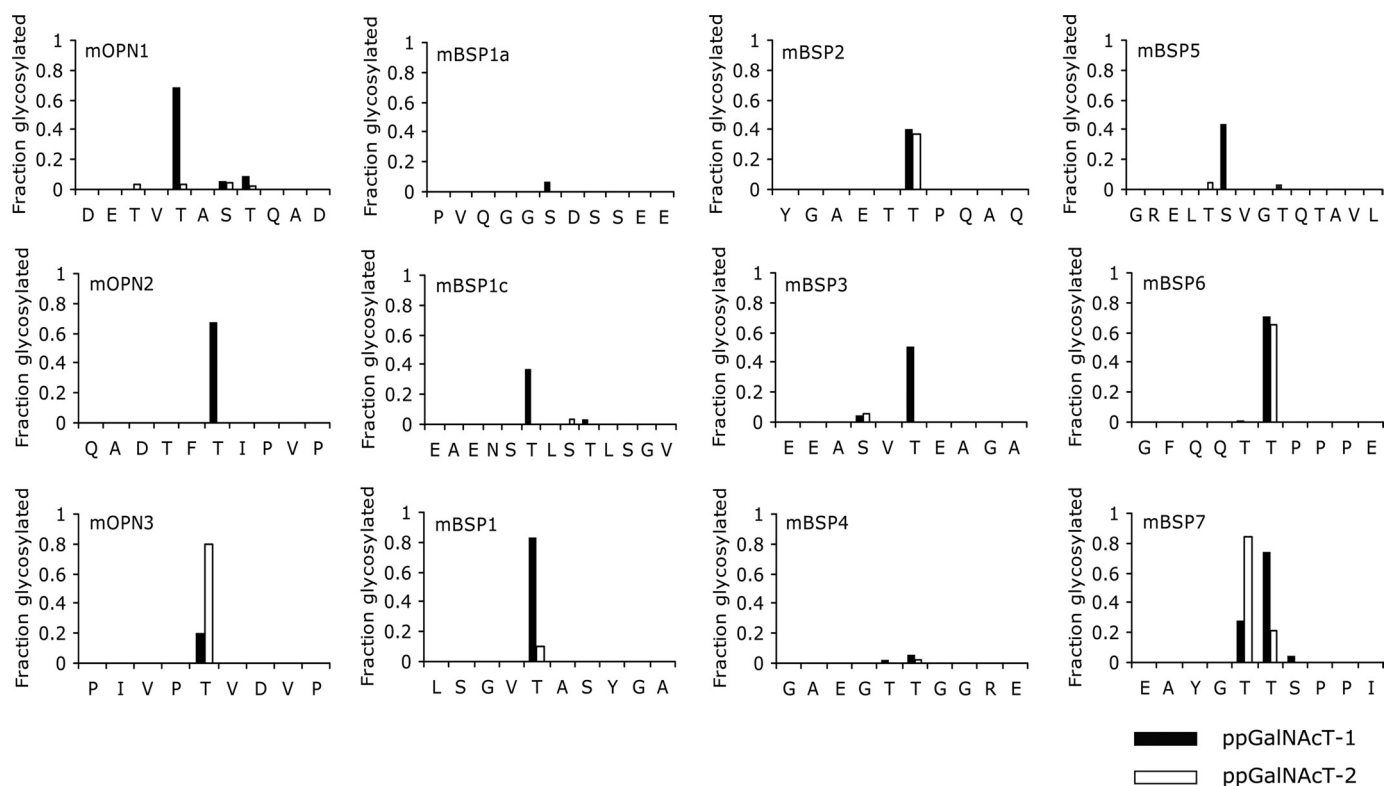


FIGURE 7. Sites of glycosylation in OPN and BSP peptides determined by Edman sequencing. The fraction of peptides glycosylated under the conditions described under "Experimental Procedures" on each site is shown. Sites glycosylated by ppGalNAcT-1 are shown in black, ppGalNAcT-2 in white. No glycosylation was observed in the mBSP2b and mBSP2d peptides and thus is not depicted.

the lack of ppGalNAcT-1 activity resulted in a reduction in the level of *O*-glycans on both OPN and BSP, the presence of *O*-glycans on either protein was not completely abolished. Thus, the question arises whether reduced *O*-glycosylation in *Galnt1*-null mice results from an overall reduction of glycans at all *O*-glycosylation sites, or by a lack of glycans on selective sites.

Available technology is insufficiently sensitive to perform direct analysis of purified protein glycoproteins derived from a reasonable number of null animals. To gain insight into this question, we examined ppGalNAcT-1-specific glycosylation *in vitro* by using synthetic peptides based on mouse OPN and BSP sequences that were predicted to contain sites of *O*-glycan attachment with predictive algorithms using NetOGlyc3.1 and amino acid sequence preferences for ppGalNAcT-1 and ppGalNAcT-2, separately determined by the work of Gerken *et al.* (7, 16). Transcripts of ppGalNAcT-1 and ppGalNAcT-2 were both highly expressed in bone, and ppGalNAcT-2 showed particularly high expression in osteoblasts. Because it was shown that these two isoforms exhibit similar, but also some unique, preferences for amino acids flanking the acceptor hydroxyl amino acid, we anticipated that this approach would allow us to look for sequence motifs within the two SIBLINGs that might be selectively glycosylated by ppGalNAcT-1, but not by ppGalNAcT-2. Although the amino acid sequence preference for ppGalNAcT-3 is not yet understood, we included it in our study, as it is highly expressed in bone and has been shown to display some overlapping substrate specificity with ppGalNAcT-1 (3).

Our study demonstrated that ppGalNAcT-1 showed preferences over ppGalNAcT-2 as well as ppGalNAcT-3 on several of

the SIBLING peptides. The selective usage of these peptides by ppGalNAcT-1 was further demonstrated in osteoblast lysates prepared from wild-type and *Galnt1*-null mice. OPN and BSP are highly expressed in osteoblasts, in which we have detected the expression of various ppGalNAcT transcripts. Although lysates from wild-type osteoblasts glycosylated those peptides that showed high selectivity for ppGalNAcT-1, the GalNAc transfer activity of *Galnt1*-null lysates against these peptides was reduced to less than 10% of the wild-type levels. Therefore, the *O*-glycosylation of these sites was poorly compensated by other isoforms and thus in *Galnt1*-null mice.

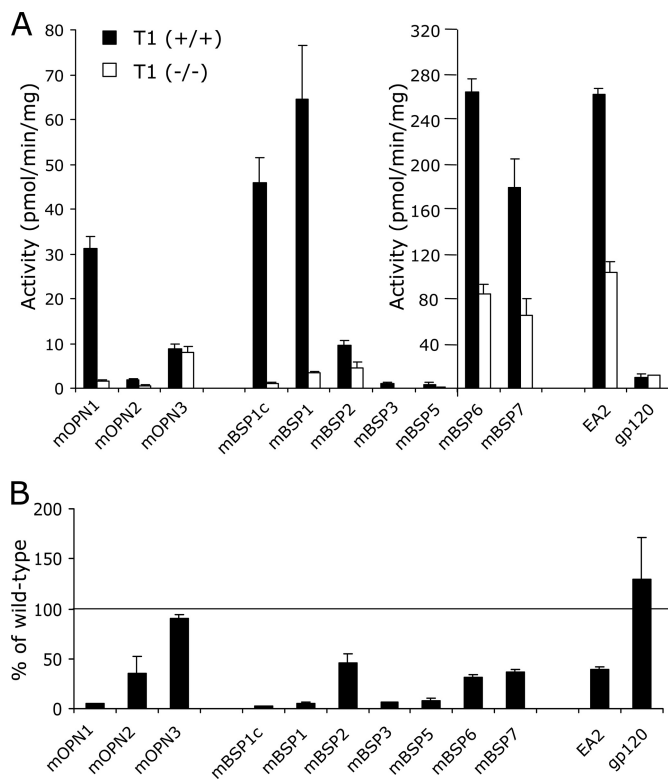
Mucin-type *O*-glycosylation appears to occur in a hierarchical order, where some transferases such as ppGalNAcT-7 and ppGalNAcT-10 require pre-existing *O*-glycosylated sites near their acceptor residues for *O*-glycosylation (6, 45, 46). Recent studies have revealed that the substrate specificity of transferases such as ppGalNAcT-2 are also modulated by pre-existing *O*-glycosylated sites (47, 48). Therefore, in addition to sites directly glycosylated by ppGalNAcT-1, it is possible that glycosylation that occurs subsequent to ppGalNAcT-1-mediated glycosylation may also be affected in *Galnt1*-null mice.

Glycosylation sites of OPN have been determined from a wide range of species and sources (13–15), including those of mouse OPN isolated from the conditioned medium of a murine osteoblastic cell line (MC3T3-E1) and transformed fibroblasts (25). Based on these studies, the mucin-type *O*-glycosylation of OPN is exclusively found in a phosphorylation-free region located N-terminal to the RGD integrin-binding motif (Fig. 4A). Five glycosylation sites have been identified in mouse, three of which are also conserved in human OPN (13). In the

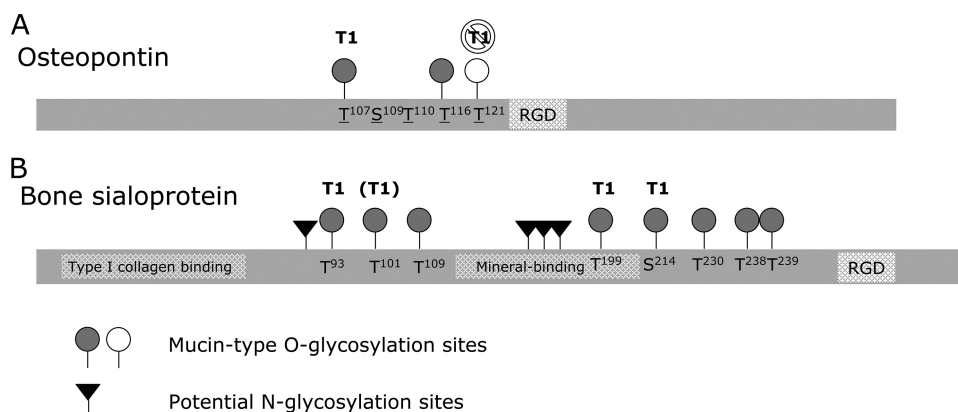
## ppGalNAcT-1-specific Glycosylation of Bone Glycoproteins

present study, we examined all five sites as potential acceptor substrates of ppGalNAcT-1, ppGalNAcT-2, and ppGalNAcT-3 *in vitro*. We found that Thr-107 and Thr-116 were glycosylated

by ppGalNAcT-1, and Thr-121 by ppGalNAcT-2 (Fig. 9A), whereas the activity of ppGalNAcT-3 against the peptides containing these residues was very low. Preferential glycosylation of Thr-107 and Thr-116 by ppGalNAcT-1 was in agreement with our prediction, as was the glycosylation of Thr-121 by ppGalNAcT-2 (Fig. 4A and supplemental Table S4). Thus, we show this approach to be useful for screening of potential isoform-selective sites. It is interesting to note that Christensen *et al.* (25) have shown that Thr-107, a ppGalNAcT-1-selective site, was only half-glycosylated in osteoblast-derived OPN, whereas all of the other four sites were fully occupied (25). Therefore, there might be a regulatory mechanism that allows this site to be differentially glycosylated based on the level of ppGalNAcT-1 activity present in individual cells. Last, glycosylation of Ser-109 and Thr-110 by these two enzymes was negligible, although they are found to be glycosylated *in vivo* (25). Therefore, it is likely that either these two residues are glycosylated by other ppGalNAcTs or these sites may require a longer peptide sequence or prior glycosylation of other nearby residues to be glycosylated *in vitro*.



**FIGURE 8. ppGalNAcT activity in osteoblasts derived from wild-type and *Galnt1*-null bone marrow cells.** Bone marrow cells were isolated from 9-week-old wild-type (+/+) and *Galnt1*-null (-/-) littermates and grown in the osteogenic medium for 14 days to allow them to differentiate into osteoblasts. A, ppGalNAcT activity in osteoblast lysates from each genotype (+/+, black; -/-, white) was determined against various peptides under the conditions described under "Experimental Procedures." Note that two scales are used to cover the wide range of activity. B, the specific activity detected in osteoblast lysates derived from *Galnt1*-null mice was normalized to that of the wild-type control. Error bars represent S.D. from average values obtained from 2 pairs of independently prepared wild-type and *Galnt1*-null osteoblasts. Each assay was performed in duplicate.



**FIGURE 9. Schematic representation of OPN and BSP domains and glycosylation sites.** Functional domains and mucin-type *O*-glycosylation sites within OPN (A) and BSP (B) are depicted. Sites that can be glycosylated at least by ppGalNAcT-1 are depicted as gray circles. Those that are highly specific for or preferred for ppGalNAcT-1 are labeled T1, or (T1), respectively. A site in OPN that can be glycosylated at least by ppGalNAcT-2, but poorly by ppGalNAcT-1, is shown as a white circle. Glycosylation sites identified in mouse OPN are underlined (25). Potential *N*-linked glycosylation sites in mouse BSP are determined based on the presence of N-X(T/S) motifs (black triangles). Type I collagen and mineral binding domains within mouse BSP are based on the studies by Tye *et al.* (49) and Tye *et al.* (51), respectively. The Arg-Gly-Asp (RGD) integrin binding motif is also shown in each protein.

Specific mucin-type *O*-glycosylation sites within mouse BSP have not been previously described. We thus utilized both the amino acid residue preferences described by the approach/algorithm of Gerken *et al.* (7, 16) and Julenius *et al.* (17) to screen for potential *O*-glycosylation sites within mouse BSP. Using a panel of synthetic peptides with recombinant ppGalNAcTs, we identified at least 8 glycosylation sites that can be glycosylated efficiently *in vitro* (Fig. 9B). Among these residues, at least three of them, Thr-93, Thr-199, and Ser-214, were strongly preferred by ppGalNAcT-1 in addition to Thr-101 that showed a preference for ppGalNAcT-1 over ppGalNAcT-2. Importantly, Thr-101, Thr-199, and Ser-214 were predicted to be preferred substrates for ppGalNAcT-1, but not for ppGalNAcT-2 (Fig. 4B and supplemental Table S4), once again demonstrating the effectiveness of using this approach to look for isoform-specific substrates. The presence of multiple ppGalNAcT-1-preferred sites in BSP is also in good agreement with the apparent molecular mass reduction observed in BSP from *Galnt1*-null mice, in which multiple mucin-type *O*-glycans present on wild-type BSP appear to be missing.

Some of the sites in mouse BSP that can be *O*-glycosylated *in vitro* are conserved in human BSP (supplemental Fig. S2), including those reported to be glycosylated in human BSP (11, 35). Although the sequence motifs differ slightly between human and mouse, we found that mouse BSP is likely composed of at least two highly *O*-glycosylated regions, as has been shown in human BSP (Fig. 9B and supplemental S2). The first region is located C-terminal to the type-I collagen binding region (49), a region that has also been shown to interact with MMP-2 (50). This region con-

tains two glycosylation sites (found in mBSP1c and mBSP1) that show a strong preference for ppGalNAcT-1. The second region is located between the hydroxyapatite (51) and integrin binding domains, which contain two motifs (found in mBSP6 and mBSP7) that can be glycosylated efficiently by multiple isoforms. This suggests that the glycosylation within these two regions is controlled in a distinctive manner.

In conclusion, we have identified OPN and BSP as *in vivo* substrates of ppGalNAcT-1. Our study utilized *in vivo* (mouse model), *in silico* (prediction), and *in vitro* (glycosylation assays) approaches to demonstrate the importance of the specific members of the ppGalNAcT family in site-specific regulation of mucin-type O-glycosylation.

*Acknowledgments*—We thank Dr. Jamey D. Marth for help in creating the Galnt1-null mice (18) and Dr. Larry W. Fisher for helpful advice, provision of LF-84, and critical reading of the manuscript. We thank Dr. Timothy A. Fritz for construction of Pichia expression vectors and Dr. Michele Forsythe for cloning of mouse ppGalNAcT-2. We thank Drs. Kelly G. Ten Hagen, Marian F. Young, Jayalakshmi Raman, and Pamela Stanley for helpful suggestions and critical reading of the manuscript.

## REFERENCES

1. Pe'er, I., Felder, C. E., Man, O., Silman, I., Sussman, J. L., and Beckmann, J. S. (2004) *Proteins* **54**, 20–40
2. Hagen, F. K., Ten Hagen, K. G., Beres, T. M., Balys, M. M., VanWuyckhuysse, B. C., and Tabak, L. A. (1997) *J. Biol. Chem.* **272**, 13843–13848
3. Wandall, H. H., Hassan, H., Mirgorodskaya, E., Kristensen, A. K., Roepstorff, P., Bennett, E. P., Nielsen, P. A., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (1997) *J. Biol. Chem.* **272**, 23503–23514
4. Ten Hagen, K. G., Hagen, F. K., Balys, M. M., Beres, T. M., Van Wuyckhuysse, B., and Tabak, L. A. (1998) *J. Biol. Chem.* **273**, 27749–27754
5. Bennett, E. P., Hassan, H., Mandel, U., Hollingsworth, M. A., Akisawa, N., Ikematsu, Y., Merckx, G., van Kessel, A. G., Olofsson, S., and Clausen, H. (1999) *J. Biol. Chem.* **274**, 25362–25370
6. Pratt, M. R., Hang, H. C., Ten Hagen, K. G., Rarick, J., Gerken, T. A., Tabak, L. A., and Bertozzi, C. R. (2004) *Chem. Biol.* **11**, 1009–1016
7. Gerken, T. A., Raman, J., Fritz, T. A., and Jamison, O. (2006) *J. Biol. Chem.* **281**, 32403–32416
8. Kingsley, P. D., Hagen, K. G., Maltby, K. M., Zara, J., and Tabak, L. A. (2000) *Glycobiology* **10**, 1317–1323
9. Young, W. W., Jr., Holcomb, D. R., Ten Hagen, K. G., and Tabak, L. A. (2003) *Glycobiology* **13**, 549–557
10. Midura, R. J., McQuillan, D. J., Benham, K. J., Fisher, L. W., and Hascall, V. C. (1990) *J. Biol. Chem.* **265**, 5285–5291
11. Wuttke, M., Müller, S., Nitsche, D. P., Paulsson, M., Hanisch, F. G., and Maurer, P. (2001) *J. Biol. Chem.* **276**, 36839–36848
12. Midura, R. J., and Hascall, V. C. (1996) *Glycobiology* **6**, 677–681
13. Christensen, B., Nielsen, M. S., Haselmann, K. F., Petersen, T. E., and Sørensen, E. S. (2005) *Biochem. J.* **390**, 285–292
14. Keykhosravani, M., Doherty-Kirby, A., Zhang, C., Brewer, D., Goldberg, H. A., Hunter, G. K., and Lajoie, G. (2005) *Biochemistry* **44**, 6990–7003
15. Sorensen, E. S., Højrup, P., and Petersen, T. E. (1995) *Protein Sci.* **4**, 2040–2049
16. Gerken, T. A., Hagen, K. G., and Jamison, O. (2008) *Glycobiology* **16**, 861–870
17. Julenius, K., Mølgaard, A., Gupta, R., and Brunak, S. (2005) *Glycobiology* **15**, 153–164
18. Tenno, M., Ohtsubo, K., Hagen, F. K., Ditto, D., Zarbock, A., Schaerli, P., von Andrian, U. H., Ley, K., Le, D., Tabak, L. A., and Marth, J. D. (2007) *Mol. Cell. Biol.* **27**, 8783–8796
19. Fisher, L. W., Stubbs, J. T., 3rd, and Young, M. F. (1995) *Acta Orthop. Scand. Suppl.* **266**, 61–65
20. Fisher, L. W., Hawkins, G. R., Tuross, N., and Termine, J. D. (1987) *J. Biol. Chem.* **262**, 9702–9708
21. Fedarko, N. S., Jain, A., Karadag, A., and Fisher, L. W. (2004) *FASEB J.* **18**, 734–736
22. Fritz, T. A., Hurley, J. H., Trinh, L. B., Shiloach, J., and Tabak, L. A. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15307–15312
23. Fritz, T. A., Raman, J., and Tabak, L. A. (2006) *J. Biol. Chem.* **281**, 8613–8619
24. Nehrke, K., Hagen, F. K., and Tabak, L. A. (1998) *Glycobiology* **8**, 367–371
25. Christensen, B., Kazanecki, C. C., Petersen, T. E., Rittling, S. R., Denhardt, D. T., and Sørensen, E. S. (2007) *J. Biol. Chem.* **282**, 19463–19472
26. Albone, E. F., Hagen, F. K., VanWuyckhuysse, B. C., and Tabak, L. A. (1994) *J. Biol. Chem.* **269**, 16845–16852
27. Zara, J., Hagen, F. K., Ten Hagen, K. G., Van Wuyckhuysse, B. C., and Tabak, L. A. (1996) *Biochem. Biophys. Res. Commun.* **228**, 38–44
28. O'Connell, B. C., and Tabak, L. A. (1993) *Anal. Biochem.* **210**, 423–425
29. Gerken, T. A., Owens, C. L., and Pasumarthy, M. (1997) *J. Biol. Chem.* **272**, 9709–9719
30. Gerken, T. A., Owens, C. L., and Pasumarthy, M. (1998) *J. Biol. Chem.* **273**, 26580–26588
31. Malaval, L., Wade-Guéye, N. M., Boudiffa, M., Fei, J., Zirngibl, R., Chen, F., Laroche, N., Roux, J. P., Burt-Pichat, B., Dubouef, F., Boivin, G., Jurdic, P., Lafage-Prost, M. H., Amédée, J., Vico, L., Rossant, J., and Aubin, J. E. (2008) *J. Exp. Med.* **205**, 1145–1153
32. Yoshitake, H., Rittling, S. R., Denhardt, D. T., and Noda, M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8156–8160
33. Ishijima, M., Rittling, S. R., Yamashita, T., Tsuji, K., Kurosawa, H., Nifuji, A., Denhardt, D. T., and Noda, M. (2001) *J. Exp. Med.* **193**, 399–404
34. Perrine, C. L., Ganguli, A., Wu, P., Bertozzi, C. R., Fritz, T. A., Raman, J., Tabak, L. A., and Gerken, T. A. (2009) *J. Biol. Chem.* **284**, 20387–20397
35. Zaia, J., Boynton, R., Heinegård, D., and Barry, F. (2001) *Biochemistry* **40**, 12983–12991
36. Wada, Y., and Tajiri, M. (2008) *Trends Glycosci. Glycotechnol.* **20**, 69–80
37. Seipert, R. R., Dodds, E. D., and Lebrilla, C. B. (2009) *J. Proteome Res.* **8**, 493–501
38. Sasaki, K., Nagamine, N., and Sakakibara, Y. (2009) *IPSI Transactions on Bioinformatics* **2**, 25–35
39. Lu, L., Niu, B., Zhao, J., Liu, L., Lu, W. C., Liu, X. J., Li, Y. X., and Cai, Y. D. (2009) *Peptides* **30**, 359–364
40. Chen, Y. Z., Tang, Y. R., Sheng, Z. Y., and Zhang, Z. (2008) *BMC Bioinformatics* **9**, 101
41. Guda, K., Moinova, H., He, J., Jamison, O., Ravi, L., Natale, L., Lutterbaugh, J., Lawrence, E., Lewis, S., Willson, J. K., Lowe, J. B., Wiesner, G. L., Parmigiani, G., Barnholtz-Sloan, J., Dawson, D. W., Velculescu, V. E., Kinzler, K. W., Papadopoulos, N., Vogelstein, B., Willis, J., Gerken, T. A., and Markowitz, S. D. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106**, 12921–12925
42. Tian, E., and Ten Hagen, K. G. (2007) *J. Biol. Chem.* **282**, 606–614
43. Topaz, O., Shurman, D. L., Bergman, R., Indelman, M., Ratajczak, P., Mizrachi, M., Khamaysi, Z., Behar, D., Petronius, D., Friedman, V., Zelikovic, L., Raimer, S., Metzker, A., Richard, G., and Sprecher, E. (2004) *Nat. Genet.* **36**, 579–581
44. Zhang, L., Zhang, Y., and Hagen, K. G. (2008) *J. Biol. Chem.* **283**, 34076–34086
45. 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) *J. Biol. Chem.* **276**, 17395–17404
46. Ten Hagen, K. G., Tetaert, D., Hagen, F. K., Richet, C., Beres, T. M., Gagnon, J., Balys, M. M., VanWuyckhuysse, B., Bedi, G. S., Degand, P., and Tabak, L. A. (1999) *J. Biol. Chem.* **274**, 27867–27874
47. Raman, J., Fritz, T. A., Gerken, T. A., Jamison, O., Live, D., Liu, M., and Tabak, L. A. (2008) *J. Biol. Chem.* **283**, 22942–22951
48. Wandall, H. H., Irazoqui, F., Tarp, M. A., Bennett, E. P., Mandel, U., Takeuchi, H., Kato, K., Irimura, T., Suryanarayanan, G., Hollingsworth, M. A., and Clausen, H. (2007) *Glycobiology* **17**, 374–387
49. Tye, C. E., Hunter, G. K., and Goldberg, H. A. (2005) *J. Biol. Chem.* **280**, 13487–13492
50. Jain, A., Fisher, L. W., and Fedarko, N. S. (2008) *Biochemistry* **47**, 5986–5995
51. Tye, C. E., Rattray, K. R., Warner, K. J., Gordon, J. A., Sodek, J., Hunter, G. K., and Goldberg, H. A. (2003) *J. Biol. Chem.* **278**, 7949–7955