Final version published as: Bergfeld, A. K., Pearce, O. M. T., Diaz, S. L., Lawrence, R., Vocadlo, D. J., Choudhury, B., Esko, J. D., & Varki, A. (2012). Metabolism of Vertebrate Amino Sugars with N-Glycolyl Groups: INCORPORATION OF N-GLYCOLYLHEXOSAMINES INTO MAMMALIAN GLYCANS BY FEEDING N-GLYCOLYLGALACTOSAMINE *. Journal of Biological Chemistry, 287(34), 28898–28916. https://doi.org/10.1074/jbc.M112.363499

Metabolism of Vertebrate Amino Sugars with *N*-Glycolyl Groups

*INTRACELLULAR β-O-LINKED N-GLYCOLYLGLUCOSAMINE (GlcNGc), UDP-GlcNGc, AND THE BIOCHEMICAL AND STRUCTURAL RATIONALE FOR THE SUBSTRATE TOLERANCE OF β-O-LINKED β-N-ACETYLGLUCOSAMINIDASE**[S]

Received for publication, March 22, 2012, and in revised form, May 25, 2012 Published, JBC Papers in Press, June 12, 2012, DOI 10.1074/jbc.M112.363721

Matthew S. Macauley[‡1], Jefferson Chan[‡], Wesley F. Zandberg[‡], Yuan He[§2], Garrett E. Whitworth[‡], Keith A. Stubbs[¶3], Scott A. Yuzwa[‖], Andrew J. Bennet[‡], Ajit Varki[**], Gideon J. Davies[§4], and David J. Vocadlo[‡‖5]

*From the Departments of [‡]Chemistry and [‖]Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada, the [§]York Structural Biology Laboratory, Department of Chemistry, University of York, YO10 5DD York, United Kingdom, the [¶]School of Chemistry and Biochemistry, University of Western Australia, Crawley, Western Australia 6009, Australia, and the **Departments of Medicine and Cellular and Molecular Medicine, University of California San Diego, La Jolla, California 92093-0687*

Background: The active site of *O*-GlcNAcase (OGA) is larger than needed to process *O*-GlcNAc from proteins.
Results: GlcNGc is assimilated to form UDP-GlcNGc and to generate *O*-GlcNGc, which can be removed by OGA.
Conclusion: OGA has a conserved active site that enables processing of *O*-GlcNGc.
Significance: Substrate tolerance of OGA occurs to process metabolic variants of *O*-GlcNAc.

The *O*-GlcNAc modification involves the attachment of single *β*-*O*-linked *N*-acetylglucosamine residues to serine and threonine residues of nucleocytoplasmic proteins. Interestingly, previous biochemical and structural studies have shown that *O*-GlcNAcase (OGA), the enzyme that removes *O*-GlcNAc from proteins, has an active site pocket that tolerates various *N*-acyl groups in addition to the *N*-acetyl group of GlcNAc. The remarkable sequence and structural conservation of residues comprising this pocket suggest functional importance. We hypothesized this pocket enables processing of metabolic variants of *O*-GlcNAc that could be formed due to inaccuracy within the metabolic machinery of the hexosamine biosynthetic pathway. In the accompanying paper (Bergfeld, A. K., Pearce, O. M., Diaz, S. L., Pham, T., and Varki, A. (2012) *J. Biol. Chem.* 287, 28865–28881), *N*-glycolylglucosamine (GlcNGc) was shown to be a catabolite of NeuNGc. Here, we show that the hexosamine salvage pathway can convert GlcNGc to UDP-GlcNGc, which is then used to modify proteins with *O*-GlcNGc. The kinetics of incorporation and removal of *O*-GlcNGc in cells occur in a dynamic manner on a time frame similar to that of *O*-GlcNAc. Enzymatic activity of *O*-GlcNAcase (OGA) toward a GlcNGc glycoside reveals OGA can process glycolyl-containing substrates fairly efficiently. A bacterial homolog (*Bt*GH84) of OGA, from a human gut symbiont, also processes *O*-GlcNGc substrates, and the structure of this enzyme bound to a GlcNGc-derived species reveals the molecular basis for tolerance and binding of GlcNGc. Together, these results demonstrate that analogs of GlcNAc, such as GlcNGc, are metabolically viable species and that the conserved active site pocket of OGA likely evolved to enable processing of mis-incorporated analogs of *O*-GlcNAc and thereby prevent their accumulation. Such plasticity in carbohydrate processing enzymes may be a general feature arising from inaccuracy in hexosamine metabolic pathways.

In nature, the biosynthesis of centrally important biomolecules, such as DNA and protein, is prone to mis-incorporation of monomeric units arising from a low level of inaccuracy in the biosynthetic machinery (1, 2). In some cases, such mis-incorporation errors have been proposed to offer evolutionary benefits (3, 4), while in other cases it can be detrimental. As a result of the apparently unavoidable laxness in the accuracy of even the most efficient enzymes, editing mechanisms are critical to ensure that the accumulation of such errors in proteins and nucleic acids is minimized (5–7). Carbohydrates, unlike DNA and protein, are not synthesized using a chemically defined template but instead their biosynthesis is governed by the tissue-specific expression of the enzymes responsible for their biosynthesis (the glycosyltransferase superfamily) as well as the cellular localization of these enzymes and their substrates (8). Glycosyltransferases have been found in many cases to be sur-

* This work was supported in part by National Institutes of Health Grant R01GM32373 (to A. V.). This work was also supported by the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council of Canada (to D. J. V.).

[S] This article contains supplemental Fig. S1, Scheme 1, and Table S1.

The atomic coordinates and structure factors (codes 4aiu and 4ais) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Recipient of a senior scholarship from the Michael Smith Foundation for Health Research and Natural Sciences and Engineering Research Council of Canada.

[2] Recipient of Ph.D. support from the Wild Fund of the University of York.

[3] Recipient of support from the Australian Research Council.

[4] Royal Society/Wolfson Research Merit Award recipient and supported by the Biotechnology and Biological Sciences Research Council.

[5] Canada Research Chair in Chemical Glycobiology, a scholar of the Michael Smith Foundation for Health Research, and an E. W. R. Steacie Memorial Fellow. To whom correspondence should be addressed: 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada. Tel.: 604-291-3530; Fax: 778-782-3765; E-mail: dvocadlo@sfu.ca.

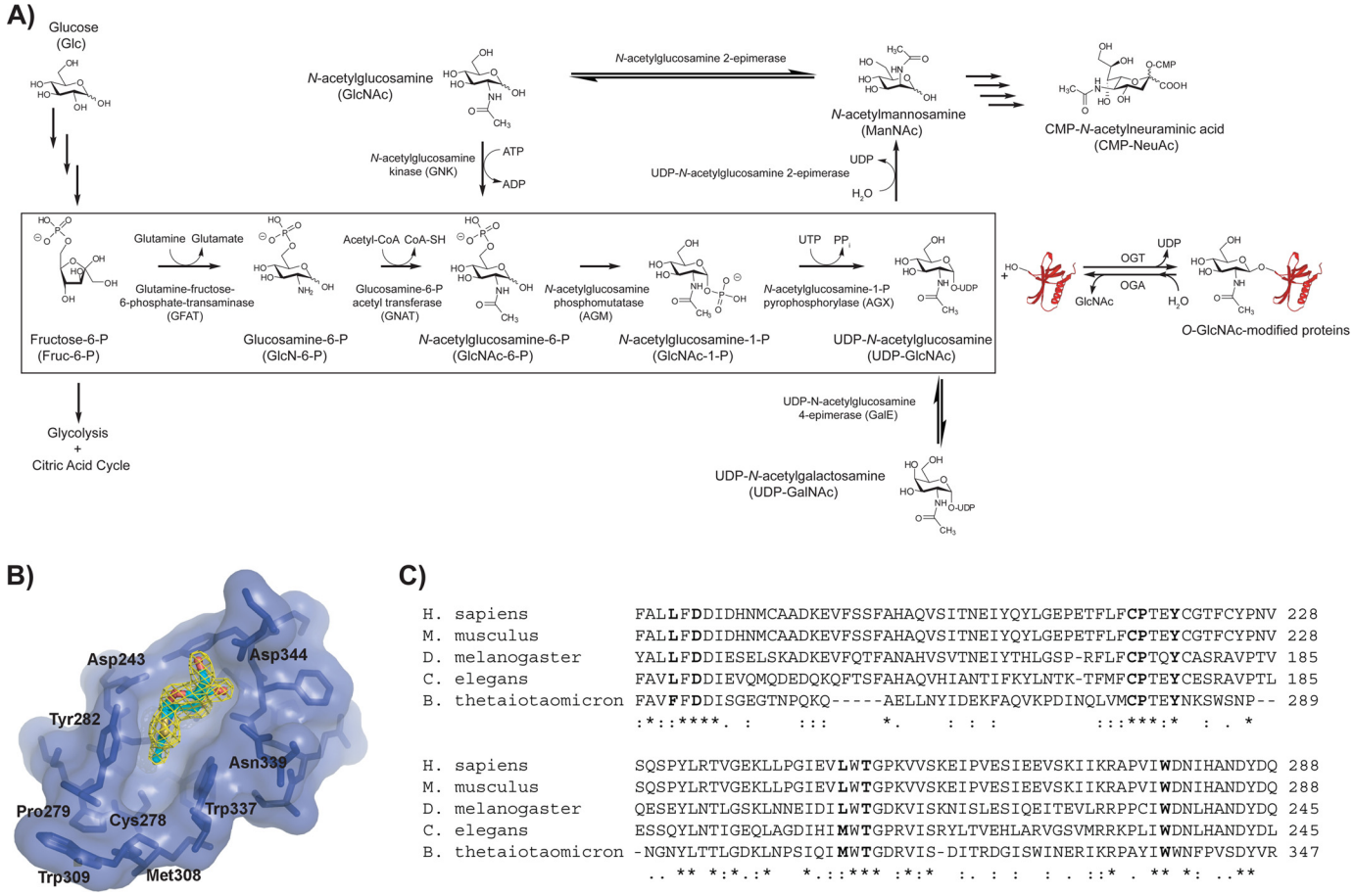


FIGURE 1. **The Hexosamine biosynthetic pathway leads to the formation of UDP-GlcNAc, which is the donor substrate used by OGT to form the *O*-GlcNAc modification, which is in turn processed within the conserved active site pocket found in human OGA and homologs.** *A*, *O*-GlcNAc consists of serine and threonine residues modified with a single β-*O*-linked *N*-acetylglucosamine (GlcNAc) residue. Installation of *O*-GlcNAc is catalyzed by OGT, which uses UDP-GlcNAc as a substrate. Removal of *O*-GlcNAc from modified proteins is catalyzed by OGA. UDP-GlcNAc is made from the hexosamine biosynthetic pathway (*box*). GlcNAc can be salvaged by phosphorylation to generate the HBSP intermediate GlcNAc-6-P. UDP-GlcNAc is also used for the biosynthesis of CMP-Neu5Ac and UDP-GalNAc. *B* and *C*, a conserved active site pocket is present around the 2-acetamido substituent in all GH84 family members. *B*, structure of *Bt*GH84 with NAG-thiazoline bound in its active site (Protein Data Bank code 2CHN). Highlighted are residues that line the active site including the pocket formed beneath the 2′-methyl group of NAG-thiazoline. *C*, sequence alignment of several GH84 family members. Residues in boldface are those that line the active site pocket. Note that Asp-242, Phe-240, Thr-310, and Val-314 are not noted in *B* because they are either out of plane (into the page) or obscured.

prisingly tolerant to variation in the structure of their preferred substrates. Indeed, some organisms exploit this substrate promiscuity by diversifying the repertoire of available substrates (9). These "errors," which arise from the loose substrate specificity of the enzyme, can give rise to further diversification of an already heterogeneous population of glycoforms (10). Such variations may offer evolutionary advantages (11) or, alternatively, remain as an evolutionary relic (12). It is important to point out, however, that cells must be able to degrade the full complement of glycoconjugates to recycle individual components and avoid the deleterious accumulation of these species. This requirement is underscored by the existence of disease states, such as the set of lysosomal storage diseases, in which various glycoconjugates accumulate to toxic levels due to the inability of cells to catabolize certain glycoconjugates (13).

Heterogeneity is well established in forms of glycosylation that exists on the extracellular side of the plasma membrane. However, the modification of proteins within the nucleus and cytoplasm with a single β-*O*-linked *N*-acetylglucosamine (*O*-GlcNAc) has been, to date, described as a having an invariable structure and linkage and lacking elongation to form oligosaccharides. The *O*-GlcNAc modification is found on proteins with a wide range of functions and appears to be essential for life at the single cell level (14). GlcNAc is installed by an enzyme termed *O*-GlcNAc transferase (OGT),[6] which uses UDP-GlcNAc as a substrate (Fig. 1*A*) (15, 16). UDP-GlcNAc is itself a product of the hexosamine biosynthetic pathway (HBSP) (Fig. 1*A*). This pathway involves four enzymes that convert fructose 6-phosphate, an intermediate in glycolysis, to UDP-GlcNAc. Intermediates in the pathway such as glucosamine (GlcN) and GlcNAc can also be incorporated into the pathway by a salvage pathway (the hexosamine salvage pathway (HSP)) involving enzymes that phosphorylate these species to form GlcN-6-P and GlcNAc-6-P, respectively. Another enzyme termed *O*-GlcNAcase (OGA) catalyzes hydrolysis of the glycosidic linkage between GlcNAc and the protein to return pro-

[6] The abbreviations used are: OGT, *O*-GlcNAc transferase; HBSP, hexosamine biosynthetic pathway; GlcNGc, *N*-glycolylglucosamine; OGA, *O*-GlcNAcase; HSP, hexosamine salvage pathway; NeuNGc, *N*-glycolylneuramic acid; ManNGc, *N*-glycolylmannosamine; ddH_2O, double distilled H_2O; Hex, hexosaminidase; rcf, relative centrifugal force; pNP-GlcNGc, 4-nitrophenyl 2-deoxy-2-hydroxyacetamido-β-D-glucopyranoside; GalNGc, *N*-glycolylgalactosamine.

teins to their unmodified state. The actions of OGA give the *O*-GlcNAc modification a dynamic nature since it has been demonstrated that the lifetime of *O*-GlcNAc on a protein is less than the polypeptide backbone of the protein to which it is attached (17).

Through a series of biochemical and structural studies (18–23), it has been firmly established that OGA harbors an unusually large active site pocket located directly below the 2-acetamido group of the substrate (Fig. 1*B*). The residues that line this pocket are conserved among eukaryotic homologs of human OGA and even among homologs from pathogenic and symbiotic bacteria (Fig. 1*C*). The near complete conservation of these residues suggests functional importance. Yet, within mammals, no derivative of GlcNAc bearing a larger substituent than the 2-acetamido group has been shown to occur naturally. Interestingly, however, the HBSP as well as OGT have been shown to have some plasticity in substrate recognition because they can readily use a substrate with an azido substituent appended off the 2-acetamido group of GlcNAc (24).

We were stimulated to look for a derivative of GlcNAc and found a number of older studies pointing toward the possible existence of one such species. One intriguing study has shown that glucosamine 6-phosphate acetyltransferase (GNAT) can accept glycolyl-CoA as a substrate to form *N*-glycolylglucosamine 6-phosphate (GlcNGc) (Fig. 1*a*). Given that glycolyl-CoA is itself naturally produced in cells from ω-hydroxylated fatty acids (25) makes the production of a GlcNGc species plausible. However, as shown in an accompanying paper (26), metabolic feeding of glycolic acid does not lead to the formation of GlcNGc suggesting that this route is improbable. More compelling is another possible route. Catabolism of *N*-glycolylneuramic acid (Neu5Gc) proceeds to generate *N*-glycolylmannosamine (ManNGc) via an enzyme termed *N*-acylneuraminate pyruvate lyase (27). In the preceding accompanying paper (26), it is clearly shown that ManNGc can be converted to GlcNGc by GlcNAc 2-epimerase and to GlcNGc 6-phosphate. The demonstration that GlcNGc and GlcNGc 6-phosphate are true catabolic intermediates within cells is significant. *In vitro* studies using sheep tissue extracts have shown that GlcNGc can be converted to GlcNGc-6-P and then to GlcNGc-1-P (28), and purified GlcNAc-6-P phosphomutase from pig could convert GlcNGc-6-P to GlcNGc-1-P with only a 3-fold decrease in catalytic efficiency compared with GlcNAc-6-P (29). Given that OGT can modify protein with 2-azidoacetamido-2-deoxy-D-glucopyranose (24), which is similar in size to GlcNGc, it seemed probable that if UDP-GlcNGc could be formed in cells it could be transferred to proteins as *O*-GlcNGc.

With these thoughts in mind, studies were initiated to test two hypotheses. The first hypothesis is that GlcNGc can be salvaged through the HSP and HBSP to generate UDP-GlcNGc within cells. The second is that GlcNGc can enter into the *O*-GlcNAc cycle to form *O*-GlcNGc-modified proteins through the action of OGT and then be removed by the action of OGA. Here, we report multiple lines of evidence supporting both of these hypotheses. The results offer a clear rationale for the presence of the highly conserved active site pocket found in OGA and other GH84 enzymes (for a review of the classification system of glycoside hydrolases see Ref. 30). We further demonstrate the molecular basis for the ability of OGA to process a glycolyl substrate through the crystal structure of a trapped GlcNGc reaction intermediate. We propose that OGA has evolved this ability to compensate for metabolic promiscuity in the HSP and HBSP, as well as in OGT, and that without such a proofreading function, *O*-GlcNGc could accumulate on proteins in an irreversible manner.

EXPERIMENTAL PROCEDURES

General

CTD110.6 was obtained from Covance, and the RL2 and HGAC85 antibodies were obtained from Abcam. Secondary antibodies were obtained from Santa Cruz Biotechnology. Human placental hexosaminidase B (HexB) was obtained from Sigma.

Cell Culture

EMEG$^{+/-}$ and EMEG$^{-/-}$ cells were cultured in standard DMEM (Invitrogen) containing 5% fetal bovine serum (HyClone). Cells were passaged every 2–3 days. The growth rate of EMEG cells was determined as follows: 5000 cells were plated onto 12-well plates and counted 5 days later using a hemocytometer. During these 5 days of growth, the media were replaced after 2 and one-half days to ensure that nutrients and/or supplements were not depleted. All counting was carried out in a minimum of three replicates. For experiments involving HPLC analysis of sugars digested from proteins by an *N*-acetylglucosaminidase from *Bacteroides thetaiotaomicron* (*Bt*GH84), EMEG$^{-/-}$ cells (5×10^6) were plated onto 175-cm flasks and treated with 5 mM GlcNAc, GlcNGc, or PBS as a control. The cells were grown for 48 h, at which time the cells reached confluence. The cells were washed with 5 ml of PBS, 3 ml of trypsin solution was added, and the cells were incubated for 3 min at 37 °C to detach the cells from the plate. The buffer and cells were then transferred to a 15-ml conical tube and pelleted by centrifugation (250 rcf, 8 min), and the supernatant was carefully removed. The cell pellet was stored at −80 °C until further processing.

Immunoblotting

Procedures for Western blotting are the same as those described previously (31). The dilutions of the three anti-*O*-GlcNAc antibodies used were 1:4000, 1:500, and 1:1000, respectively, for CTD110.6, RL2, and HGAC85. To block binding of the CTD110.6 antibody using either GlcNAc or GlcNGc, the antibody was first made up to the appropriate dilution in PBS-T containing 1% BSA. The appropriate amount of sugar was then incubated with this antibody solution for 10 min at room temperature, and this solution was used directly to probe the membranes.

Cloning

The cDNA encoding human GalE was obtained from Origene and amplified by PCR using the following two primers: 5′-AGCAGC**CATATG**ATGGCAGAGAAGGTGCTGG-3′ (NdeI cut site shown in boldface) and 5′-AGCAGC**CTCGAG**TCAGGCTTGCGTGCCAAAGCC-3′ (XhoI cut site shown in boldface). The PCR product was digested with NdeI and XhoI and

ligated into pET28a to produce a construct with an N-terminal His_6 tag. The clone for *Sp*Hex, in pET15 encoding an N-terminal His_6 tag, was obtained as a gift from Dr. Brian Mark (University of Manitoba). The gene encoding human Agx1, in the pTrcHisB expression vector encoding an N-terminal His_6 tag, was obtained as gift from Véronique Piller (Université d'Orléans). Cloning of human OGA and *Bt*GH84 into pET28a to encode fusion proteins with an N-terminal His_6 tag has been described previously (18, 32).

Protein Expression—All vectors were transformed into Tuner™(λDE3) cells (Stratagene) and selected with either 50 μg/ml kanamycin or 100 μg/ml ampicillin with the exception that *Sp*Hex was selected with only 25 μg/ml ampicillin. Cells containing the plasmids were grown until an A_{600} of ~1.0 was reach, after which time protein expression was induced by addition of 0.5 mM isopropyl 1-thio-β-D-galactopyranoside for 3–4 h at 25 °C. Bacteria were pelleted, lysed, and purified by nickel chelate affinity chromatography using a previously reported procedure (32).

Nucleotide Diphosphate Extraction

Nucleotide diphosphate sugars were extracted from cells according to established procedures (33, 34). $EMEG^{-/-}$ cells were plated onto 10-cm plates and grown to 50% confluence and then treated with either 5 mM GlcNAc or 5 mM GlcNGc or PBS as a control. Each growth condition was carried out in triplicate. After 48 h, a time when the cells have reached confluence, the plates of cells were washed once with PBS before the addition of 1 ml of PBS containing 10 mM EDTA. Cells were incubated at 37 °C for 4 min, scraped from the plates using a rubber policeman, transferred into 1.5-ml centrifuge tubes, and placed on ice prior to pelleting by centrifugation (100 rcf, 10 min, 4 °C). Cell pellets were immediately lysed by adding 750 μl of pre-chilled (−20 °C) 75% ethanol and sonicating these samples using a W-375 ultrasonic processor (Heat Systems Ultrasonics, Inc.) using two blasts of 30 s each. Insoluble material was removed from the lysates by centrifugation (21,000 rcf, 15 min, 4 °C), and the supernatants, which contained the sugar nucleotides, were snap-frozen in liquid nitrogen and then lyophilized. To this crude extract was added GDP-Glc (200 pmol) as an internal standard, dissolved in 0.5 ml of ddH_2O, and extracted using Envi-Carb graphite solid-phase extraction cartridges (200 mg) (Supelco) exactly as described previously (35). Sugar nucleotides were eluted three times using 1 ml of 25% CH_3CN in 50 mM triethylammonium acetate, pH 7.0, passed through a 0.22-μm filter (Millipore), lyophilized, and stored in dry form at −20 °C until analysis by capillary electrophoresis.

Capillary Electrophoresis

Lyophilized sugar nucleotides were dissolved in 200 μl of ddH_2O, and an aliquot was diluted 1:4 prior to characterization by capillary electrophoresis by adapting a previously reported method (36). Capillary electrophoresis was performed using a ProteomeLab PA800 (Beckman-Coulter) equipped with fused silica capillaries of 50 μm internal diameter × 44 cm length (to the detector). The running buffer, prepared from components of the highest available purity, was 40 mM $Na_2B_4O_7$ (Sigma), pH 9.5, containing 1.0% (w/v) polyethylene glycol (M_r 20,000) (Fluka) and was filtered through a 0.22-μm filter (Millipore) prior to use. Before samples were introduced, the capillary was conditioned by washing sequentially with 1 N NaOH (2 min, 20 p.s.i.), ddH_2O (3 min, 20 p.s.i.), and running buffer (5 min, 40 p.s.i.). To ensure reproducible migration times, the running buffer was replaced after every second run. After injecting a short (~15 nanoliters) H_2O plug, samples were electrostatically introduced and pre-concentrated at the anode by applying a potential of −10 kV for 10 s according to a field-amplified sample injection technique described previously (37). Electrophoresis was carried out at a constant voltage of 30 kV (which produced a current of about 90 μA), and the capillary was thermostated at 22 °C. Electrophoretograms were derived by measuring the absorbance at 254 (±10) nm at a rate of 4 Hz. Peaks were integrated using 32 Karat 5.0 software (Beckman-Coulter), and all data were normalized to the GDP-Glc internal standard. All NDP-sugar standards, as well as GlcNAc-1-P, were obtained from Sigma, except for UDP-GlcNGc and UDP-GalNGc, which were made as part of this work.

Large Scale Digestion of Lysates with BtGH84

Cell pellets were thawed on ice and resuspended in 1 ml of buffer consisting of 10 mM sodium phosphate and 10 mM sodium chloride, pH 6.5. Solutions were made up very carefully with the highest quality salts and water to avoid small amounts of contaminant that would be concentrated over the following steps. Low amounts of salts and buffer were used to minimize the salt content. The cells were lysed by passing the solution up and down through a 27-gauge needle three times. This suspension was sonicated (two times for 20 s at 15% duty) followed by centrifugation (17,900 rcf, 20 min) to clarify the supernatant. The supernatant was made up to 2.5 ml with the same low salt buffer and desalted using a PD-10 column equilibrated in same low salt buffer as the eluent. One round of desalting did not efficiently desalt the samples; therefore, the 3.5 ml of elution obtained after the first column was further diluted to 5 ml and split into two 2.5-ml aliquots, which were desalted a second time. These eluted proteins were combined (7 ml in total), and fucose and maltose were added as internal standards (1 μM final concentration) to enable a percentage recovery to be calculated. These solutions were treated at 37 °C for 1 h with 200 μl of 1 mg/ml *Bt*GH84 (that had been dialyzed in the same low salt buffer) or 200 μl of buffer as a control. One set of lysates was not incubated at 37 °C but instead was kept on ice to control for any molecules that were not removed by desalting and/or released by endogenous glycosidases during the 1-h 37 °C incubation step. Following glycosidase digestion, the samples were concentrated down to 1 ml on a speed vac (Thermo) heated to 55 °C, and the protein was then precipitated from the reactions by the addition of 7 ml of 95% cold ethanol, and the resulting mixtures were incubated at −20 °C for 30 min. The precipitated protein was pelleted by centrifugation (4000 rpm, 15 min). The supernatant was gently decanted and then lyophilized (overnight). The precipitate was resuspended in 2 ml of methanol by extensive vortexing. Any precipitate that did not dissolve was removed by centrifugation (17,900 rcf, 10 min), and the supernatant was evaporated using a speed vac heated to 55 °C. The residue was taken back up in 150 μl of ddH_2O. These solutions

were vortexed and then centrifuged (17,900 rcf, 5 min) to remove any insoluble debris, and the supernatant was collected for HPLC analysis of GlcNAc or GlcNGc content.

HPLC Analysis of Sugars Liberated by BtGH84 Digestion

Carbohydrates were separated by HPLC anion exchange chromatography using an ICS 3000 HPLC equipped with an ASI 100 automated sample injector (Dionex) and detected using an electrochemical detector (ED_{50}, Dionex) using a gold working electrode and an Ag/AgCl reference electrode. To detect GlcNAc, a CarboPac PA20 column was used, and isocratic elution was carried out using 20 mM NaOH. Preliminary experiments designed to guide optimization of the assay showed that a GlcNGc standard was retained by the PA20 column and eluted very late in the chromatogram. The long retention time of GlcNGc reflects strong interaction with the solid phase of the column (immobilized quaternary ammonium ion), and this interaction is much stronger than GlcNAc because the highly basic mobile phase likely results in partial deprotonation of the acidic group present on the glycolyl moiety. Therefore, for better detection of GlcNGc, a CarboPak PA100 column was used, and isocratic elution was carried out using 100 mM NaOH. In each case, 20 μl of sample was injected; this was followed by 20 min of isocratic elution, followed by a 5-min wash with 500 mM NaOH to clean the column between the samples. After washing, the column was equilibrated for a minimum of 15 min in the same concentration of NaOH used to run the isocratic elution. A constant flow rate of 0.4 ml/min was used with the PA20 column, whereas a flow rate of 0.5 ml/min was used with the PA100 column, as per the manufacturer's instructions. To assess the percent recovery and absolute amount of GlcNAc and GlcNGc present in the samples, standard curves of fucose, maltose, GlcNAc, and GlcNGc were constructed for concentrations ranging from 0.5 to 50 μM. GlcNAc (Bioshop), ManNAc (Jülich), maltose (Sigma), and fucose (Carbosynth) were obtained commercially and carefully prepared in ddH_2O to construct the standard curves.

Comparative Digests of Cell Lysates with SpHex and BtGH84

*Sp*Hex and *Bt*GH84 were prepared as 1 mg/ml stocks in a buffer consisting of 50 mM sodium phosphate and 100 mM sodium chloride at pH 6.5. Cell lysates from $EMEG^{-/-}$ cells cultured for 48 h with either GlcNAc or GlcGc were treated with either *Sp*Hex or *Bt*GH84 at a final concentration of 0.1 mg/ml and incubated at 37 °C. After various times, an aliquot of the reaction was removed and quenched by addition of SDS-PAGE loading buffer and heated at 95 °C for 5 min. Samples were analyzed by Western blot using the CTD110.6 anti-*O*-GlcNAc antibody as described above. Densitometric analysis was performed using ImageQuant software.

Enzyme Kinetics

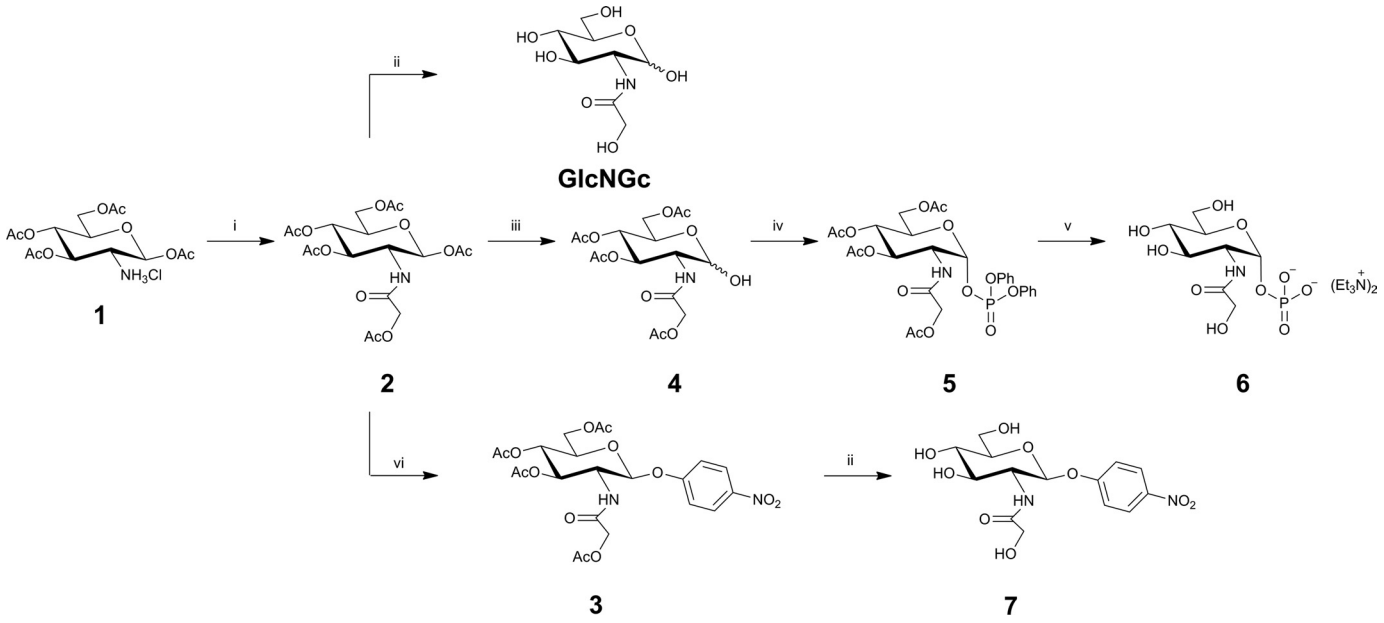
The rates of OGA- and HexB-catalyzed hydrolysis of pNP-GlcNAc and pNP-GlcNGc (**7**) were measured using a continuous UV/visible assay as described previously (38). Briefly, 10 mM stocks of the two substrates were made up in either PBS or low pH citrate buffer (50 mM sodium citrate, 100 mM sodium chloride, pH 4.5) for monitoring the OGA- and HexB-catalyzed reactions, respectively. Reactions consisted of a total of 100 μl with various concentrations of substrate and 100 nM enzyme in 96-well plates. Reactions were initiated by the addition of enzyme and monitored continuously at either 400 nM for OGA-catalyzed reactions or 340 nM for HexB-catalyzed reactions on a 96-well plate reader. Initial rates of reaction were obtained between the 2nd and 5th min of the reaction. Rates were converted into absolute amount of product released by comparing the absorbance values to a standard curve of *p*-nitrophenol made up in the same buffer conditions.

Chemical Syntheses-4-Nitrophenyl 2-deoxy-2-hydroxyacetamido-β-D-glucopyranoside (pNP-GlcNGc) 7 (Scheme 1)

1,3,4,6-Tetra-O-acetyl-2-acetyloxyacetamido-2-deoxy-glucopyranose ***2***—Triethylamine (2.2 ml, 16 mmol) was added to a solution of the hydrochloride salt of 2-amino-2-deoxy-1,3,4,6-tetra-*O*-acetyl-β-D-glucopyranose **1** (39) (2.0 g, 5.2 mmol) in CH_2Cl_2 (20 ml), at which time the starting material dissolved. The reaction mixture was cooled to 0 °C, and acetoxyacetyl chloride (0.7 ml, 6.5 mmol) was added. The resultant mixture was stirred for ~2 h at room temperature. When the reaction mixture was judged complete by TLC analysis, the mixture was diluted with EtOAc. The resulting organic phase was washed successively with water, 1 M NaOH, and saturated sodium chloride. The organic phase was dried over $MgSO_4$, filtered, and concentrated to yield a white crystalline solid. The material was recrystallized using a mixture of ethyl acetate and hexanes to yield the desired compound **2** as a white solid (2.0 g, 86%). For ^{1}H NMR (600 MHz, $CDCl_3$), δ = 6.25 (1H, d, $J_{NH\text{-}H2}$ = 9.2 Hz, NH), 5.71 (1H, dd, $J_{H1\text{-}H2}$ = 8.8 Hz, H-1), 5.19–5.11 (2H, m, H-3/4), 4.53 (1H, A part of ABq, $J_{H\text{-}H}$ = 15.4 Hz, CH_2), 4.40 (1H, A part of ABq, CH_2), 4.32 (1H, ddd, $J_{H2\text{-}H3}$ = 9.1 Hz, H-2), 4.25 (1H, dd, $J_{H5\text{-}H6}$ = 4.3 Hz, $J_{H6\text{-}H6}$ = 12.3 Hz, H-6), 4.12 (1H, dd, $J_{H5\text{-}H6}$ = 2.2 Hz, H-6′), 3.81 (1H, ddd, $J_{H4\text{-}H5}$ = 9.2 Hz, H-5), 2.15 (3H, s, OAc), 2.10 (3H, s, OAc), 2.08 (3H, s, OAc), 2.02 (3H, s, OAc), and 2.01 (3H, s, OAc). For ^{13}C NMR (150 MHz, $CDCl_3$), δ = 171.3, 170.6, 169.7, 169.6, 169.2, 167.5 (C = O), 92.4 (C-1), 73.0, 72.1, 67.6 (C-3, C-4, C-5), 62.6, 61.6 (C-6, CH_2), 52.7 (C-2), 20.8, 20.7, 20.5 (OAc). For mass spectrometry, FAB calculated for $C_{18}H_{26}NO_{12}$ was 448.1455, and found was 448.1445. The melting point = 168–169 °C. For $C_{18}H_{25}NO_{12}$, that calculated was C 48.32, H 5.63 and found was C 48.48, H 5.59.

2-Deoxy-2-hydroxyacetamido-D-glucopyranose (GlcNGc)—Following a published synthetic strategy (40), a solution of 1,3,4,6-tetra-*O*-acetyl-2-acetyloxyacetamido-2-deoxy-β-D-glucopyranose **2** (1.5 g, 3.35 mmol) in MeOH (50 ml) was treated with 0.5 M NaOMe (5 ml). The solution was stirred overnight at room temperature. Amberlite resin (H^+ form) was added to neutralize the reaction and was filtered. The filtrate was concentrated under reduced pressure to afford the desired product as a white solid (711 mg, 90%) and m.p. = 179.0–180 °C. Spectral properties of the desired product (GlcNGc) matched those described previously (40).

4-Nitrophenyl 2-deoxy-2-hydroxyacetamido-β-D-glucopyranoside ***7***—A solution of 1,3,4,6-tetra-*O*-acetyl-2-acetyloxyacetamido-2-deoxy-D-glucopyranose **2** (250 mg, 0.56 mmol) and 4-nitrophenol (117 mg, 0.84 mmol) in anhydrous CH_2Cl_2 (10



SCHEME 1. *(i)* $AcOCH_2C(O)Cl$, Et_3N, CH_2Cl_2; *(ii)* NaOMe, MeOH; *(iii)* $BnNH_2$, THF; *(iv)* $(PhO)_2P(O)Cl$, DMAP, CH_2Cl_2; *(v) a*; PtO_2, THF; *b*, MeOH:H_2O:Et_3N (5:2:1); *(vi)* 4-nitrophenol, $BF_3{\cdot}OEt_2$, CH_2Cl_2.

ml) was cooled to 0 °C in an ice bath. The solution was treated with boron trifluoride etherate (177 μl, 1.4 mmol) and stirred at room temperature for 3 h. The reaction mixture was then washed successively with saturated $NaHCO_3$ (20 ml), water (20 ml), and brine (20 ml). The organic fraction was dried (Na_2SO_4) and concentrated under reduced pressure to give a syrup. This material was purified via flash chromatography (EtOAc/hexanes, 1:1, v/v) to afford the intermediate **3** as an off-white solid. Deacetylation of **3** under Zemplén conditions afforded the desired compound **7** as a light-yellow solid (92 mg, 46% over two steps). For ^{1}H NMR (600 MHz, D_2O), δ = 8.19 – 8.18 (m, 2H, Ar-H), 7.15–7.11 (m, 2H, Ar-H), 5.33 (dd, J = 8.4, 1.2, 1H, H-1), 4.10 – 3.98 (m, 3H), 3.91–3.89 (m, 1H), 3.78 – 3.70 (m, 2H), 3.68 – 3.61 (m, 1H), 3.54 (t, J = 13.9, 1H), and 1.21 (bs, 1H, NH). For ^{13}C NMR (151 MHz, D_2O), δ = 175.4 (C = O), 161.2, 142.2, 125.6, 116.0 (Ar), 97.9 (C-1), 75.8 (C-5), 72.7 (C-3), 69.1 (C-4), 60.4 (CH_2O_2), 59.9 (C-6), and 54.5 (C-2). The melting point = 188.5–189 °C. For $C_{14}H_{18}N_2O_8$, that calculated was C 49.12, H 5.30 and found was C 49.29, H 5.19.

Bis(triethylammonium) 2-Hydroxyacetamido-2-deoxy-α-D-glucopyranosyl Phosphate 6

*3,4,6-Tri-O-acetyl-2-acetyloxyacetamido-2-deoxy-D-glucopyranose **4***—To acetate **2** (1.0 g, 2.2 mmol) dissolved in THF (10 ml) was added benzylamine (0.4 ml, 3.5 mmol), and the mixture was kept (2 h). The mixture was then diluted with CH_2Cl_2 (30 ml) and washed sequentially with cold water, 0.5 M HCl, and saturated sodium bicarbonate solution, then dried ($MgSO_4$), and concentrated affording a colorless oil (800 mg). This material, presumably the hemiacetal **4**, was used in the next step without any further purification.

*Diphenyl (3,4,6-Tri-O-acetyl-2-acetyloxyacetamido-2-deoxy-α-D-glucopyranosyl) Phosphate **5***—A solution of the hemiacetal **4** (0.56 g, 1.4 mmol) and 4-dimethylaminopyridine (DMAP, 1.00 g, 8.2 mmol) in CH_2Cl_2 (30 ml) was cooled to −10 °C. Diphenyl chlorophosphate (0.9 ml, 4.3 mmol) was added dropwise and the solution stirred (4 h), gradually warming to 10 °C. The solution was then diluted with CH_2Cl_2 (30 ml) and washed sequentially with cold water, 0.5 M HCl, and saturated sodium bicarbonate solution, then dried ($MgSO_4$), and concentrated. Flash chromatography (EtOAc/hexanes, 1:3) of the residue gave the α-phosphate **5** as a colorless oil (730 mg, 83%). For ^{1}H NMR (500 MHz, $CDCl_3$), δ = 7.38 –7.32 (4H, m, Ar), 7.24 –7.18 (6H, m, Ar), 6.40 (1H, d, $J_{NH\text{-}H2}$ = 8.4 Hz, NH), 6.04 (1H, dd, $J_{H1\text{-}H2}$ = 3.2 Hz, $J_{H1\text{-}P}$ = 6.0 Hz, H-1), 5.29 (1H, dd, $J_{H3\text{-}H4}$ = 9.7 Hz, $J_{H4\text{-}H5}$ = 9.7 Hz, H-4), 5.22 (1H, dd, $J_{H2\text{-}H3}$ = 9.7 Hz, H-3), 4.50 (1H, A part of ABq, $J_{H\text{-}H}$ = 15.5 Hz, CH_2), 4.39 (1H, m, H-2), 4.18 – 4.14 (2H, m, CH_2, H-6), 4.06 (1H, ddd, $J_{H5\text{-}H6}$ = 2.0 Hz, $J_{H5\text{-}H6}$ = 6.0 Hz, H-5), 3.89 (1H, dd, $J_{H6\text{-}H6'}$ = 12.6 Hz, H-6′), 2.06 (3H, s, OAc), 2.04 (3H, s, OAc), 2.03 (3H, s, OAc), and 1.99 (3H, s, OAc). For ^{13}C NMR (125 MHz, $CDCl_3$), δ = 171.7, 170.5, 169.5, 168.9, 167.6 (C = O), 150.1, 129.9, 125.8, 119.8 (Ar), 96.6 ($J_{C1\text{-}P}$ = 6.7 Hz, C-1), 70.1, 69.5, 66.7 (C-3, C-4, C-5), 62.3, 61.0 (C-6, CH_2), 52.3 ($J_{C2\text{-}P}$ = 8.0 Hz, C-2), 20.5, 20.4, and 20.3 (OAc). For mass spectrometry, FAB calculated for $C_{28}H_{33}NO_{14}P$ was 638.1624 and found was 638.1639. For $C_{28}H_{32}NO_{14}P$, that calculated was C 52.75, H 5.06 and found was C 52.68, H 5.13.

*Bis(triethylammonium) 2-Hydroxyacetamido-2-deoxy-α-D-glucopyranosyl Phosphate **6***—PtO_2 (40 mg) was added to a solution of the α-phosphate **5** (110 mg) in THF (5 ml), and the mixture was stirred under an atmosphere of hydrogen gas (RT, 48 h). When the reaction was judged complete by TLC analysis, the mixture was filtered through Celite and concentrated *in vacuo*. The crude material was then resuspended in MeOH/H_2O/triethylamine (5:2:1, 5 ml) and stirred for 16 h at 4 °C and then concentrated *in vacuo* to yield the phosphate **6** as a colorless oil with no further purification required (85 mg, 98%). For ^{1}H NMR (500 MHz, D_2O), δ = 5.39 (1H, dd, $J_{H1\text{-}H2}$ = 3.3 Hz,

$J_{H1\text{-}P}$ = 7.6 Hz, H-1), 4.15 (1H, A part of ABq, $J_{H\text{-}H}$ = 16.5 Hz, CH_2), 4.11 (1H, A part of ABq, CH_2), 3.99 (1H, ddd, $J_{H2\text{-}P}$ = 2.1 Hz, $J_{H2\text{-}H3}$ = 10.5 Hz, H-2), 3.94 (1H, ddd, $J_{H5\text{-}H6}$ = 2.2 Hz, $J_{H5\text{-}H6}$ = 5.1 Hz, $J_{H4\text{-}H5}$ = 10.1 Hz, H-5), 3.89–3.84 (2H, m, H-3, H-6), 3.78 (1H, dd, $J_{H6\text{-}H6'}$ = 12.4 Hz, H-6′), 3.51 (1H, dd, $J_{H3\text{-}H4}$ = 9.2 Hz, H-4), 3.20 (12H, q, $J_{CH2\text{-}CH3}$ = 7.4 Hz, CH_2CH_3), and 1.28 (18H, t, CH_2CH_3). For ^{13}C NMR (125 MHz, D_2O), δ = 176.3 (C = O), 93.1 ($J_{C1\text{-}P}$ = 5.5 Hz, C-1), 73.3, 71.9, 70.8 (C-3, C-4, C-5), 62.0, 61.6 (C-6, CH_2), 54.6 ($J_{C2\text{-}P}$ = 7.7 Hz, C-2), 47.6 (CH_2CH_3), and 9.2 (CH_2CH_3). For mass spectrometry, FAB calculated for $C_8H_{15}NNa_2O_{10}P$ was 362.0229 and found was 362.0224.

2-Deoxy-2-hydroxyacetamido-D-mannopyranose (ManNGc) (supplemental Scheme 1)

Following a published synthetic strategy (41), a solution of mannosamine hydrochloride **8** (1.46 g, 6.77 mmol) and $NaHCO_3$ (3.5 g, 42 mmol) in H_2O (60 ml) was cooled in an ice-bath and treated with acetoxyacetyl chloride (3.9 ml, 36 mmol). After stirring for 1 h, the solution was neutralized with aqueous HCl. The reaction mixture was concentrated under reduced pressure, and the crude residue was purified via flash chromatography (EtOAc/*i*PrOH/H_2O, 27:8:4, v/v) to afford the intermediate **9** as a white solid. This material was dissolved in autoclaved H_2O (20 ml), treated with amano lipase A from *Aspergillus niger* (500 mg, activity ≥12,000 units/g) and incubated at 37 °C for 48 h. The solution was filtered through an Amicon filter and lyophilized to afford the title product as a white solid (952 mg, 59%). Spectral properties of the desired product matched those described previously (41).

Crystallography

To obtain the complex structure of B*t*GH84-glycolic acid, apo-B*t*GH84 crystals, grown as described previously (18), were soaked in a drop containing 0.1 M NH_4OAc, 15% (w/v) PEG3350, 20% glycerol, and 0.4 M glycolic acid at pH 6.0 overnight. The crystals were subsequently flash-frozen in liquid nitrogen prior to data collection. X-ray data of *Bt*GH84-glycolic acid were collected from a single crystal to 2.0 Å on beamline ID14-1 at the European Synchrotron Radiation Facility. All data were integrated using MOSFLM (42), followed by data scaling and merging with SCALA from the CCP4 suite of program CCP4 suite (43). The structure was solved using molecular replacement method using the apo structure (Protein Data Bank code 2CHO) as an initial searching model. Manual corrections to the model were made with COOT (44), and refinement cycles were performed with REFMAC (45).

The D242N-glycolyl oxazoline complex was obtained by soaking of D242N crystals (46) in a drop containing crystallization mother liquor supplemented with a minute amount of pNP-GlcGc **7** for 5 min. Soaked crystals were then quickly transferred to a second drop comprising the same solution with glycerol concentration raised to 20% (w/v) before the crystals were flash-frozen in liquid nitrogen. Data for D242N-glycolyl oxazoline were collected at the European Synchrotron Radiation Facility (Grenoble) to 2.25 Å resolution on beamline ID29. Data processing and structure solution were performed essentially as described for the *Bt*GH84-glycolic acid structure, above. Details of data and final structure quality are shown in supplemental Table 1. Coordinates have been deposited with the Protein Data Bank under the codes 4aiu and 4ais.

RESULTS AND DISCUSSION

To begin our search for a viable metabolically-derived GlcNAc analog, we felt that a cell line in which the biosynthesis of UDP-GlcNAc is compromised would be useful. Because UDP-GlcNAc is generally abundant in cells, lower levels of UDP-GlcNAc would facilitate the observation of an analog of GlcNAc and enhance the effects of feeding exogenous GlcNAc analogs that would compete with GlcNAc derived through the HBSP. It is known, however, that the biosynthesis of UDP-GlcNAc is necessary for life; deletion of the gene encoding glucosamine-6-phosphate acetyltransferase (*EMEG32*) results in embryonic lethality in mice (47). Despite the fatal consequence of deleting the *EMEG32* gene at the organismal level, mouse embryonic fibroblasts from *EMEG32*-deficient embryos have been generated and readily propagated (47). Although heterozygous ($EMEG32^{+/-}$) cells have normal growth rates and show levels of UDP-GlcNAc comparable with those of WT cells, the growth rate of $EMEG32^{-/-}$ cells is dramatically impaired, and their cellular UDP-GlcNAc levels are ~1/50th the concentration of those found in $EMEG32^{+/-}$ and WT cells. It is likely that $EMEG32^{-/-}$ cells derive this UDP-GlcNAc from scavenging free GlcNAc and degrading GlcNAc-containing glycans present in the serum used during cell culture. Consistent with this view, we find that the growth rate of $EMEG^{-/-}$ cells is affected by serum concentration to a much larger extent than are $EMEG^{+/-}$ cells. Boehmelt *et al.* (47) demonstrated that the slow growth rate and cellular levels of UDP-GlcNAc of $EMEG32^{-/-}$ cells could both be rescued by addition of GlcNAc to the media. Thus, we felt that $EMEG32^{-/-}$ mouse embryonic fibroblasts would be an excellent tool for investigating the effects of exogenously added GlcNAc, or analogs thereof, on cells as well as evaluating the tolerance of the HBSP to such compounds.

Rescue of the Growth Rate of $EMEG^{-/-}$ Cells by GlcNGc and ManNGc—Because OGA shows a highly conserved active site pocket that most likely accommodates structural variants of *O*-GlcNAc, we thought it would be interesting to see if cells treated with metabolically viable analogs of GlcNAc could incorporate these in place of *O*-GlcNAc. Earlier suggestions regarding the existence of GlcNGc (25), and an accompanying paper (26), which provides compelling evidence that both GlcNGc and GlcNGc-6-phosphate are catabolic intermediates derived from Neu5Gc, provide a clear rationale to evaluate whether GlcNGc could be incorporated as *O*-GlcNGc. Using $EMEG32^{-/-}$ cells, we observe that the growth rate of these cells approached that of heterozygous cells when 10 mM GlcNAc was added to the culture media (Fig. 2*A*). Heterozygous cells showed no enhancement in growth rate when GlcNAc was added to the culture media (Fig. 2*B*); results that are consistent with previous studies using this cell line (47). In parallel, we treated cells with varying concentrations of GlcNGc. In $EMEG32^{-/-}$ cells, GlcNGc rescued the growth rate in a dose-dependent manner up to a concentration of 10 mM, such that

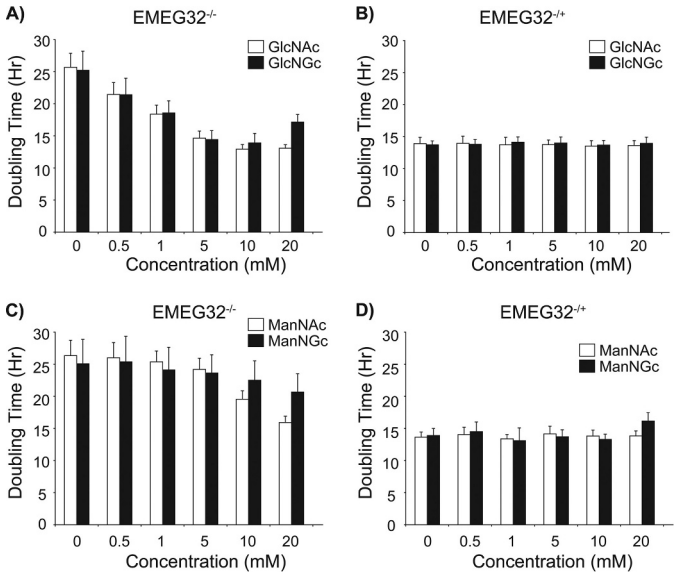


FIGURE 2. **GlcNGc and ManNGc rescue the growth rate of *EMEG32*$^{-/-}$ mouse embryonic fibroblasts.** Approximately 5000 *EMEG32*$^{-/-}$ (*A* and C) *EMEG32*$^{+/-}$ (*B* and *D*) cells were plated onto 12-well plates and at the end of 5 days of growth; the cells were counted on a hemocytometer to obtain the doubling time. Cells were cultured with the indicated concentration of GlcNAc or GlcNGc (*A* and *B*) and ManNAc or ManNGc (*C* and *D*) through the experiment. Each experimental condition was repeated in at least triplicate, and the errors represent the standard deviation between these replicates.

the growth rate essentially matched that observed when using GlcNAc (Fig. 2*A*). Above 10 mM GlcNGc, cells began to grow more slowly, and this appears to be specific to GlcNGc treatment in *EMEG32*-deficient cells, because we do not observe the same effect when using GlcNGc in *EMEG32*$^{+/-}$ cells (Fig. 2*B*). These results suggested to us that it might be that GlcNGc is being converted to UDP-GlcNGc, which can compensate for decreased UDP-GlcNAc and thereby rescue growth. Following the idea that GlcNGc might be derived from Neu5Gc, we treated *EMEG32*$^{-/-}$ cells with various concentrations of ManNAc and found that high concentrations of ManNAc also rescued their growth rate (Fig. 2*C*). The high concentration of ManNAc required to observe this effect is in keeping with earlier reports showing that ManNAc is not as readily taken up into the cell as is GlcNAc (48) and that uptake of ManNAc is nonsaturable up to a concentration of 20 mM (49). When using ManNGc, we found a minor increase in growth rate at 20 mM; however, the difference did not reach statistical significance ($p = 0.09$) (Fig. 2*C*). Furthermore, we found that 20 mM ManNGc had a slightly deleterious effect on the growth rate of *EMEG32*$^{+/-}$ cells (Fig. 2*D*). This latter observation suggested to us that the effect of ManNGc on the growth rate of *EMEG32*$^{-/-}$ cells might be complicated by two countervailing effects. The first is that epimerization of ManNGc to GlcNGc might result in increased UDP-GlcNGc levels, and this could rescue the growth rate by substituting for UDP-GlcNAc. The second is that high levels of ManNGc or some metabolic intermediate derived from ManNGc might have a deleterious effect on growth rate. An additional complication is that because ManNAc appears to only gain access to the cell through passive diffusion (49), the additional hydroxyl group of the glycolyl unit of ManNGc might compromise the ability of this molecule to passively diffuse into the cell. With these complications in mind, the growth rate data for *EMEG32*$^{-/-}$ cells treated with ManNGc does suggest that it can be epimerized to GlcNGc, which together with the growth rate rescue observed using ManNAc suggests that GlcNAc 2-epimerase is operative in these cells. The data further suggest that the HBSP is likely capable of converting GlcNGc to UDP-GlcNGc.

Analysis of UDP-GlcNAc Levels—To test if exogenously added GlcNGc can be converted to UDP-GlcNGc within *EMEG32*$^{-/-}$ cells, we used a capillary electrophoresis assay to analyze levels of NDP-sugars from cells. When *EMEG32*$^{-/-}$ cells were incubated with 5 mM GlcNAc overnight, the levels of UDP-GlcNAc dramatically increased as expected (Fig. 3*A, GlcNAc*). On incubation with 5 mM GlcNGc, two new peaks appeared in the electrophoretogram (Fig. 3*A, GlcNGc, peaks 2* and *5*). One peak corresponds to a species eluting slightly before UDP-GlcNAc (Fig. 3*A*, *peak 2*). We confirmed that this was a distinct species by mixing samples derived from cells treated with a mixture of GlcNAc and GlcNGc (Fig. 3*A, mixture*). A second new species gave rise to a peak (Fig. 3*A*, *peak 5*) eluting between UDP-Glc and UDP-GalNAc. We hypothesized that the two new peaks were UDP-GlcNGc and, because UDP-GalNAc is made directly from UDP-GlcNAc, UDP-GalNGc. In the absence of standards, however, the identity of these two new peaks could not be unambiguously assigned. Accordingly, synthetic standards of UDP-GlcNGc **12** and UDP-GalNGc **14** were prepared enzymatically from chemically synthesized GlcNGc-1-P **6** (Scheme 2). We used two recombinant enzymes for this biosynthesis as follows: Agx1, a UDP-HexNAc pyrophosphorylase that catalyzes the transfer of UMP from UTP to GlcNAc-1-P, and GalE, a UDP-HexNAc 4-epimerase that catalyzes the interconversion of UDP-GlcNAc and UDP-GalNAc. As shown in Fig. 3*B*, Agx1 yielded UDP-GlcNAc **11** when incubated with GlcNAc-1-P and UTP (Scheme 2). When GalE was added to the resulting reaction mixture, a second peak could be observed in the electrophoretogram that corresponded to UDP-GalNAc **13**, which indicates that these enzymes are active. When GlcNGc-1-P was used as a substrate for Agx1, a peak was observed that eluted from the capillary slightly before UDP-GlcNAc, and importantly, this peak has the same electrophoretic mobility as the first new peak (peak 2) from the cells treated with GlcNGc. When GalE was added to this reaction mixture (Scheme 2), another species eluted from the capillary at the same time as the second new peak (peak 5) observed in the sample obtained from cells treated with GlcNGc. Carrying out a reaction with GlcNAc-1-P and GlcNGc-1-P at equimolar concentrations in the presence of Agx1 and GalE resulted in the production of the expected four product peaks (Fig. 3*B, peaks 2, 3, 5,* and *6*). It should be noted that the two enzymes we use here, Agx1 and GalE, are recombinant human enzymes. These observations therefore support the ability of these enzymes to process substrates having the glycolyl moiety in human cells and other eukaryotes, such as mice, in which these enzymes are highly conserved at the amino acid level. Although a quantitative kinetic analysis was not performed, it should be noted that in Fig. 3*B* both UDP-GlcNAc **11** and UDP-GlcNGc **12** were produced quantitatively. Also, the absolute amount of UDP-GlcNGc produced in cells treated with 5 mM GlcNGc is similar

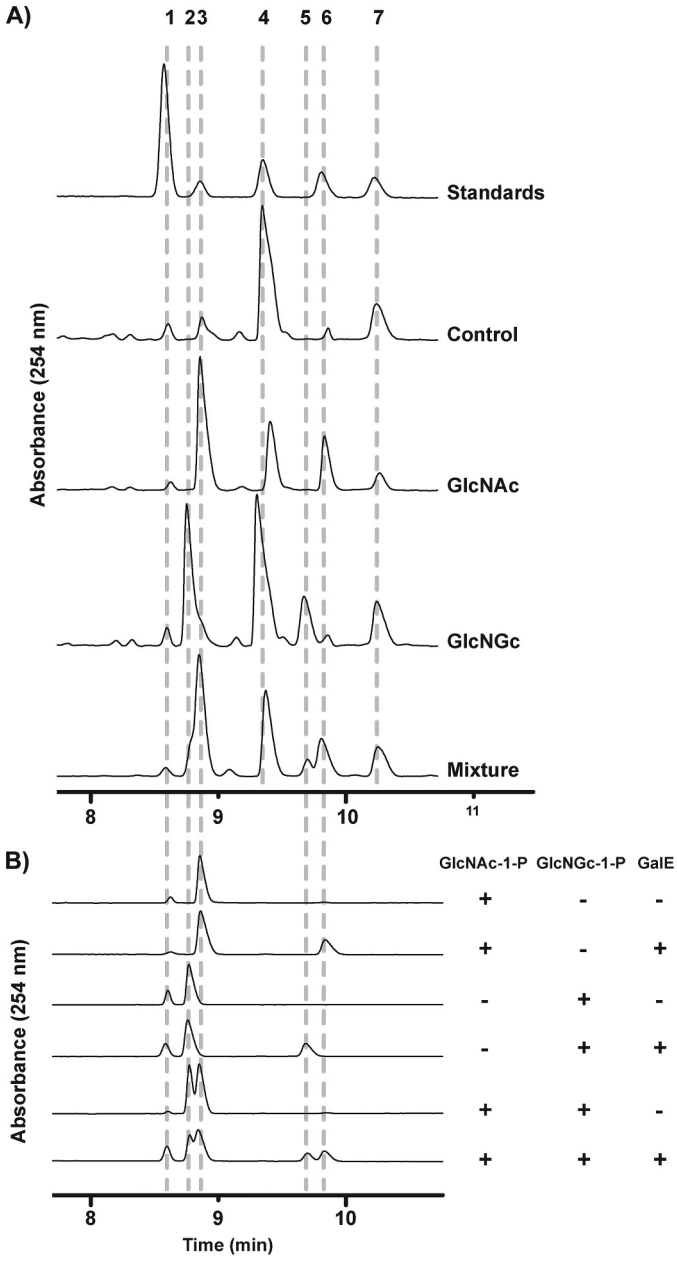
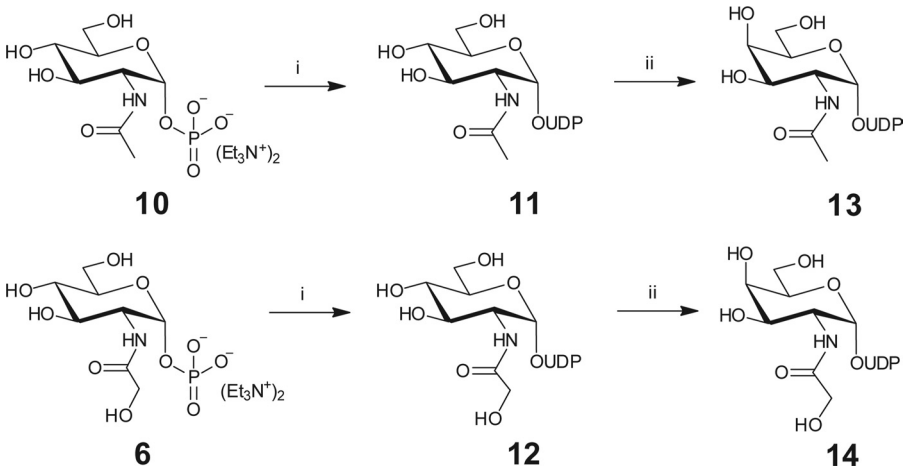


FIGURE 3. **UDP-GlcNGc and UDP-GalNGc can be biosynthesized from exogenously supplied GlcNGc in *EMEG32*$^{-/-}$ mouse embryonic fibroblasts.** *A, EMEG32*$^{-/-}$ cells were treated for 24 h with either 5 mM GlcNAc or GlcNGc. Cells were then harvested, and nucleotide mono-, di-, and triphosphate-containing compounds were isolated from cell extracts and analyzed by capillary electrophoresis. Commercially available standards (*upper trace*) established the electrophoretic mobilities of the following: *1*) GDP-Glc (8.6 min; used as an internal standard); *3*) UDP-GlcNAc **11** (8.9 min); *4*) UDP-Glc (9.3 min); *6*) UDP-GalNAc **13** (9.85 min); and *7*) UDP-Gal (10.2 min). Peaks labeled *2* (8.8 min) and *5* (9.75 min) are unique to cells treated with GlcNGc and mixing samples from GlcNAc- and GlcNGc-treated cells (mixture) strongly suggesting that peak *2* is not the peak arising from UDP-GlcNAc. *B,* to establish the identity of peaks *2* and *5*, GlcNAc-1-P and GlcNGc-1-P were converted to UDP-GlcNAc and UDP-GlcNGc, respectively, by the enzyme Agx1. Additionally, UDP-GlcNAc and UDP-GlcNGc **12** were converted to UDP-GalNAc **13** and UDP-GalNGc **14**, respectively, by the epimerase GalE.

to the amount of UDP-GlcNAc formed when cells were treated with 5 mM GlcNAc. These data indicate that the enzymatic machinery of the HBSP in mammals is capable of efficiently converting GlcNGc to UDP-GlcNGc **12** as well as UDP-GalNGc **14**. An accompanying paper (50) provides complementary evidence for the metabolic promiscuity of the HBSP and GalE by showing exogenous GalNGc is converted to both UDP-GalNGc and UDP-GlcNGc in cultured mammalian cells.

Immunoblot Evidence for Modification of Proteins with O-GlcNGc—Having established that UDP-GlcNGc **12** can be biosynthesized within cells, it follows that it may be used as a substrate by OGT. OGT has been previously shown to use uridine diphosphate *N*-azidoacetylglucosamine UDP-GlcNAz as a substrate (24), which suggests that it might be able to use UDP-GlcNGc **12** with reasonable efficiency because the azide and hydroxyl substituents of the compounds are similar in size. To test this possibility, we treated *EMEG32*$^{-/-}$ cells for 24 h with 5 mM GlcNAc or GlcNGc, after which *O*-GlcNAc levels were analyzed by Western blot using the CTD110.6 anti-*O*-GlcNAc antibody (Fig. 4*A*). As measured this way, we find that *O*-GlcNAc levels increased substantially in cells treated with 5 mM GlcNAc compared with control cells, which is in line with the large increase in UDP-GlcNAc levels observed when these cells were treated with 5 mM GlcNAc (Fig. 4*A*). In contrast, the immunoreactivity of lysates derived from cells treated with 5 mM GlcNGc was greatly diminished compared with control cells despite equivalent loading of all samples. Initially, this result was surprising to us and prompted the investigation of two other anti-*O*-GlcNAc antibodies (RL2 and HGAC85). We therefore tested these antibodies and observed a similar phenomenon (Fig. 4, *B* and *C*). Interestingly, *EMEG32*$^{-/-}$ cells treated with 20 mM of either ManNAc or ManNGc showed a similar trend except that the differences in levels of immunoreactivity were not as dramatic (Fig. 4*D*). Two possibilities could account for the observed decreases in immunoreactivity. The first possibility is that these antibodies have decreased binding affinity to possible *O*-GlcNGc-modified proteins. The second possibility is that some GlcNGc metabolite could act as an inhibitor of OGT. To test the first possibility, the CTD110.6 anti-*O*-GlcNAc antibody was analyzed for its ability to bind GlcNAc and GlcNGc by preincubating it with different concentrations of GlcNAc or GlcNGc prior to using it in a Western blot. We found that at least 100-fold more GlcNGc is required to block antibody binding (Fig. 4*E*). Thus, the CTD110.6 antibody has a preference for the 2-acetamido group of GlcNAc over the 2-*N*-glycolyl group of GlcNGc, and the decrease in immunoreactivity observed in lysates from *EMEG32*$^{-/-}$ cells treated with GlcNGc or ManNGc could represent the accumulation of β-*O*-linked GlcNGc on nucleocytoplasmic proteins.

Evidence for Modification of Proteins with O-GlcNGc by HPLC Analysis—To obtain more direct evidence for the presence of *O*-GlcNGc-modified proteins, we monitored GlcNAc or GlcNGc enzymatically liberated from proteins by anion-exchange HPLC. GlcNAc or GlcNGc was released from proteins using a bacterial homolog of OGA from *B. thetaiodomicron* (*Bt*GH84). This enzyme was used because of the following: (i) it can efficiently remove *O*-GlcNAc from modified proteins (18); (ii) it can be expressed recombinantly at high levels, which allows the use of large quantities to drive reactions toward completion (18), and (iii) it hydrolyzes GlcNGc glycosides only slightly less efficiently than GlcNAc (see below). *EMEG32*$^{-/-}$ cells (5×10^7) were treated with either 5 mM GlcNAc or



SCHEME 2. *(i)* UTP, Agx1; *(ii)* GalE.

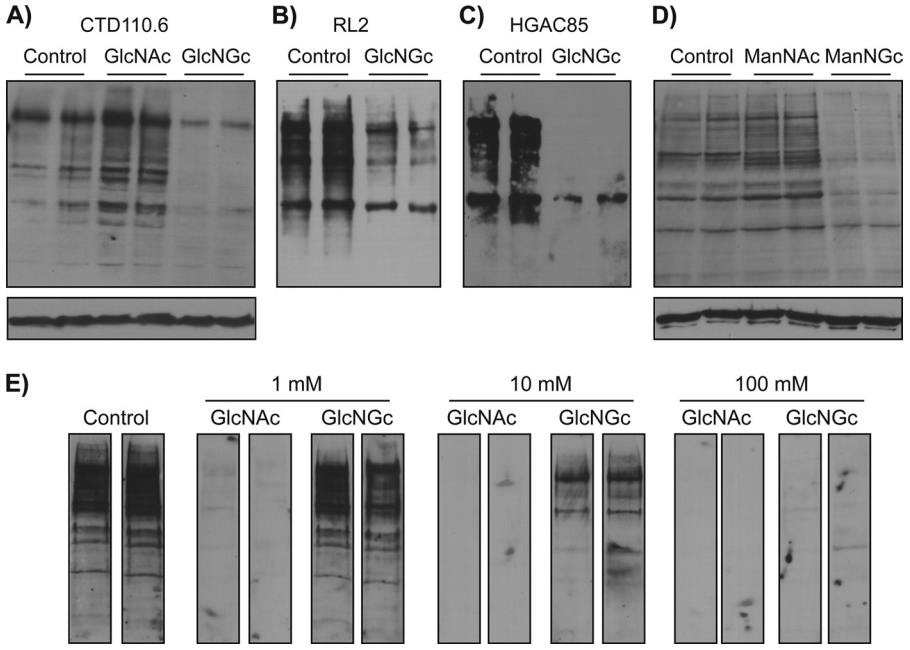


FIGURE 4. **Effect of GlcNGc and ManNGc on *O*-GlcNAc levels in *EMEG32*$^{-/-}$ mouse embryonic fibroblasts.** Cells were treated for 24 h with either 5 mM of GlcNAc or GlcNGc (*A–C*) and 20 mM ManNAc or ManNGc (*D*). *O*-GlcNAc levels were assessed by Western blot analyses using three different *O*-GlcNAc antibodies. In all cases, cells treated with GlcNGc or ManNGc show lower immunoreactivity using these antibodies (*A* and *D*, CTD110.6; *B,* RL2; *C*, HGAC85). The *upper panels* represents Western blots with the specified anti-*O*-GlcNAc antibody, and the *lower panels* are Western blots obtained using an anti-β-actin to demonstrate equal loading of samples. *E,* CTD110.6 anti-*O*-GlcNAc antibody binds GlcNGc with lower affinity than GlcNAc. The antibody was incubated with the specified concentration of either GlcNAc or GlcNGc for 10 min prior to as well as during probing of the blots containing cell lysates from *EMEG32*$^{-/-}$ cells treated with GlcNAc.

GlcNGc for 48 h, and desalted cell lysates were incubated with active *Bt*GH84 (WT *Bt*GH84) or inactive mutant *Bt*GH84 (*Bt*GH84-D242A). Despite two rounds of desalting, we find sample lysates from cells treated with GlcNAc contained a small amount of GlcNAc even without adding *Bt*GH84 (Fig. 5*A*, *Enzyme* ($t = 0$)). When the lysates were incubated with the inactive mutant *Bt*GH84-D242A, the size of the peak corresponding to GlcNAc remained unchanged (Fig. 5*A*, + *Mut Enzyme*). However, digestion of lysates with WT *Bt*GH84 greatly increased the size of the peak corresponding to GlcNAc (Fig. 5*A*, + *WT Enzyme*). The results obtained for the lysates derived from cells treated with GlcNGc appear initially to be more challenging to interpret (Fig. 5*B*); therefore, we carried out several additional controls (Fig. 5*C*) to clarify the results. The interpretation was initially complicated by the fact that in all the assays, no peak had the same retention time as the GlcNGc standard. There was, however, a peak in all the samples treated with GlcNGc that eluted ~0.4 min earlier than the GlcNGc standard does, and significantly, this peak was larger when analyzing lysates from cells incubated with GlcNGc and then digested with WT *Bt*GH84 (Fig. 5*B*). To investigate whether this peak corresponded to GlcNGc, and whether it simply eluted earlier because of other factors, such as the composition of the sample, we carried out an additional experiment (Fig. 5*C*). As mentioned above, analysis of lysates from GlcNAc-treated cells, which were never exposed to exogenously added GlcNGc, revealed no peaks having an elution time corresponding to either the GlcNGc standard or the peak

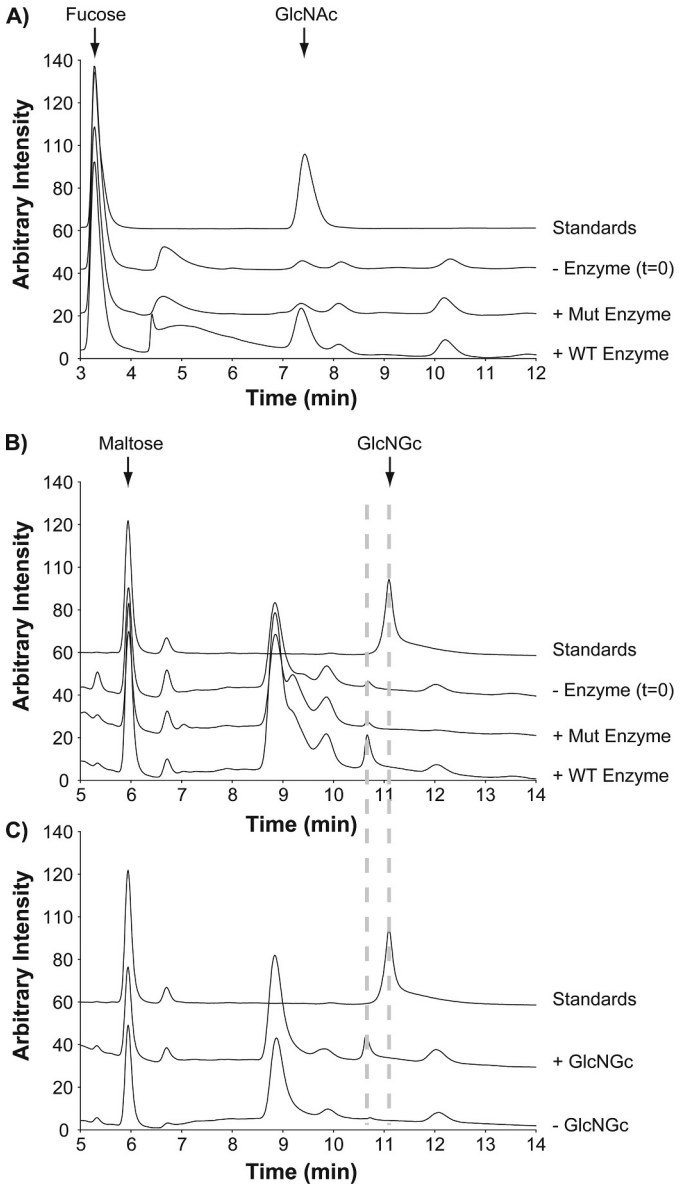


FIGURE 5. **Direct support for the attachment of GlcNGc to proteins via a β-*O*-linkage.** *EMEG32*$^{-/-}$ cells were treated with either 5 mM GlcNAc (*A*) or 5 mM GlcNGc (*B*) for 24 h. Cell lysates were incubated with 1 mg/ml *Bt*GH84 for 60 min, and protein was removed by ethanol precipitation. The samples were then analyzed by anion-exchange HPLC. For cells treated with GlcNAc, samples were analyzed on a PA20 (strong binding) with a mobile phase consisting of 20 mM NaOH and an internal standard of 20 μM fucose. For cells treated with GlcNGc, samples were analyzed on a PA100 (weaker binding) with a mobile phase consisting of 100 mM NaOH and an internal standard of 20 μM maltose. The internal standards were added into the samples prior to *Bt*GH84 digestion to assess the percent recovery through handling steps. The retention times of GlcNAc (7.4 min) and GlcNGc (11.1 min) standards were used to determine the peak corresponding to these species from the reactions. *C*, variation in the retention time of GlcNGc observed when assaying standards compared with samples derived from cell lysates was resolved by adding a standard to lysates. Samples derived from cells treated with GlcNAc, but without added BtGH84, were assayed both with and without added 1 μM GlcNGc.

observed in the digested GlcNGc-treated cell lysates. When a small amount (2 μM) of GlcNGc was added to this cell lysate, a new peak appeared that eluted 0.4 min earlier than did the GlcNGc standard, and significantly, this new peak aligned with the peak hypothesized to be GlcNGc in the lysates obtained from GlcNGc-treated cells. Accordingly, this peak arises from GlcNGc and its retention time differs from that of the standard due to the composition of the samples; the standard is in water, and the experimental sample contains other small molecules that can affect electrophoretic mobility.

Dynamic Addition and Removal of GlcNGc—The results of the HPLC assay shown above indicate that *EMEG32*$^{-/-}$ cells treated with GlcNGc have terminal β-*O*-linked GlcNGc within their glycoconjugates. It is possible that some of this GlcNGc is contained within soluble glycoconjugates other than *O*-GlcNGc present on nucleocytoplasmic proteins. Yet, as many other forms of glycoconjugates are found on proteins stably associated with membranes and on secreted proteins, many of these would have been removed by washing the cells and by the centrifugation step following cell lysis. One hallmark of the *O*-GlcNAc modification of nucleocytoplasmic proteins is its dynamic nature (14). To assess the time-dependent addition and removal of GlcNGc from proteins in *EMEG32*$^{-/-}$ cells, we cultured cells for 48 h in the presence of either 5 mM GlcNAc or GlcNGc. After this time, the cells were washed, and the media were changed to media containing the other compound. Cells were then harvested at various times after switching the media to analyze *O*-GlcNAc levels by Western blot. When cells were initially grown in GlcNGc and then switched to GlcNAc, we found a time-dependent increase in *O*-GlcNAc immunoreactivity up to a maximum at 48 h (Fig. 6*A*). This increase in immunoreactivity represents the removal of GlcNGc from proteins and subsequent addition of GlcNAc. Based on these data, we estimate the half-life for this process to be 24 h. Interestingly, to our knowledge only two published studies have attempted to quantify the dynamics of *O*-GlcNAc; one study estimated a half-life for *O*-GlcNAc of 48 h on a cytokeratin (51), and another study estimated a half-life of 12 h on an α-crystallin (17). Although many factors likely contribute to the residency time of *O*-GlcNAc (or *O*-GlcNGc) on proteins, the global rates that are estimated from these data are in line with the rates determined previously on individual proteins. When we first cultured cells in media containing GlcNAc and then switched to media containing GlcNGc, *O*-GlcNAc immunoreactivity showed a time-dependent decrease over 48 h (Fig. 6*B*). In this experiment, the decrease in signal represents the combined processes of removal of GlcNAc from proteins and subsequent addition of GlcNGc. These results support the view that the GlcNGc can be added to and removed from proteins in a dynamic manner consistent with that observed for *O*-GlcNAc.

O-GlcNGc in Cells Having an Intact HBSP—The evidence we present thus far for *O*-GlcNGc modification of nucleocytosolic proteins in cells treated with GlcNGc comes from EMEG32$^{-/-}$ cells, which lack a functional HBSP and cannot therefore independently biosynthesize GlcNAc. To investigate if exogenously added GlcNGc is able to compete with the regular flux through the HBSP in cells containing a functional HBSP, *EMEG32*$^{+/-}$ cells were used because they have normal UDP-GlcNAc levels (47); consistent with this, their growth rate is not dependent on exogenously added GlcNAc (Fig. 2*B*). We speculated that if GlcNGc can compete with GlcNAc endogenously produced via the HBSP, then exogenously added GlcNGc should result in decreased *O*-GlcNAc immunoreactivity. To test this proposal, we treated *EMEG32*$^{+/-}$ cells for 48 h with either 5 mM GlcNAc

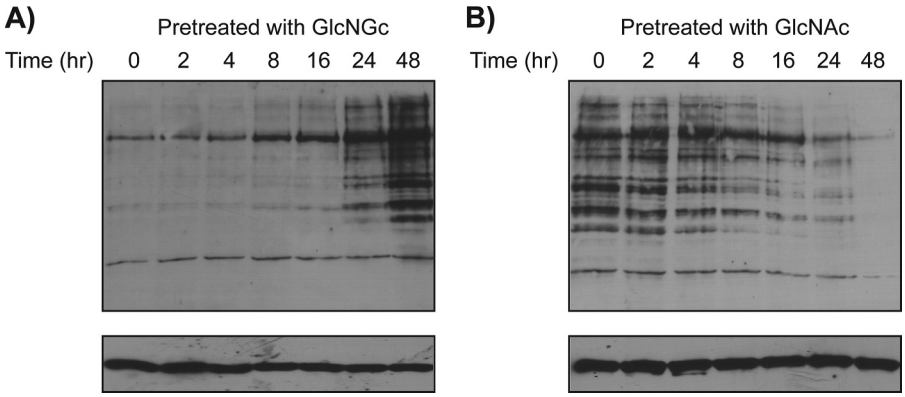


FIGURE 6. **Time course for addition of and removal of GlcNGc from *EMEG32*$^{-/-}$ mouse embryonic fibroblasts.** Cells were grown for 48 h in the presence of 5 mM GlcNGc (*A*) or GlcNAc (*B*), and then the media were replaced with media containing 5 mM of the other compound. At various times, the cells were harvested, and *O*-GlcNAc levels were assessed by Western blot. The *upper panels* represent Western blots obtained using the CTD110.6 anti-*O*-GlcNAc antibody, and the *lower panels* are Western blots obtained using an anti-β-actin, which demonstrates equivalent loading of samples.

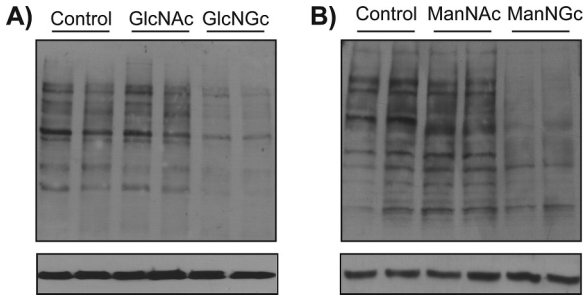


FIGURE 7. **Effect of GlcNGc and ManNGc on *O*-GlcNAc levels in *EMEG32*$^{+/-}$ mouse embryonic fibroblasts.** Cells were treated for 48 h with either 5 mM GlcNAc or GlcNGc (*A*) and either 20 mM ManNAc or ManNGc (*B*). The *upper panels* represent Western blots obtained using the CTD110.6 anti-*O*-GlcNAc antibody, and the *lower panels* are Western blots obtained using an anti-β-actin to demonstrate equivalent loading of samples.

or GlcNGc, and CTD110.6-positive immunoreactivity present in lysates was assessed by Western blot (Fig. 7*A*). Cells treated with GlcNAc showed only a slight increase in *O*-GlcNAc immunoreactivity, whereas cells treated with GlcNGc showed a decrease that is consistent with the studies presented above and can be accounted for by poor recognition of GlcNGc by the antibody. Treatment of *EMEG32*$^{+/-}$ cells with either 20 mM ManNAc or ManNGc produced a similar outcome (Fig. 7*B*). Collectively, these data indicate that GlcNGc can feed into the HBSP via the HSP and be incorporated as *O*-GlcNGc in the presence of an intact HBSP.

Activity of Human OGA toward pNP-GlcNGc **7**—Biochemical studies of human OGA and the structures of OGA homologs (18–23) have revealed the active site contains a pocket that might enable processing *O*-GlcNGc-modified proteins and other GlcNGc glycosides. To address this question, we prepared (Scheme 1) pNP-GlcNGc, **7** and tested it as a substrate for both human OGA and HexB. Consistent with the cell-based studies, we find that OGA is able to catalyze cleavage of this substrate (Fig. 8*A*) with a catalytic efficiency ($V_{max}/[E]\cdot K_m$) toward pNP-GlcNGc that is 48-fold lower compared with pNP-GlcNAc (Fig. 8*A* and Table 1). However, the drop in catalytic efficiency of human HexB toward pNP-GlcNGc as compared with pNP-GlcNAc was much greater (550-fold) (Fig. 8*B* and Table 1). Therefore, it appears that the large active site pocket of OGA revealed previously (18–20, 52) makes it more tolerant toward glycolyl-containing substrates as compared with the lysosomal HexB, the structure of which reveals there is no pocket beneath the 2-acetamido group (53, 54). This highly conserved pocket (Fig. 1, *B* and *C*) found in *N*-acetylglucosaminidases from GH84 is highly conserved and is likely present to ensure that any *O*-GlcNGc, or other adventitiously incorporated GlcNAc analogs, formed within cells can be removed and not accumulate on proteins.

Molecular Basis for the Ability of OGA to Process of GlcNGc Substrate—In an effort to provide structural insight into GlcNGc binding to OGA, we solved the structure of *Bt*GH84 in complex with glycolic acid. *Bt*GH84 is a good model of human OGA because active site residues are entirely conserved between these two enzymes. The rationale for using glycolic acid was that the initial "apo" *Bt*GH84 structure had shown a molecule of acetate binding in the *N*-acetyl binding "pocket" and glycerol (used as a cryo-protectant) binding in place of the O4-C4-O5-C5-C6-O6 atoms of *N*-acetylglucosamine; the two fragments mimicked well the position of the natural substrate (18). When glycolate was soaked into the apo-crystal, we observed it bound in a position previously occupied by an acetate molecule, suggesting this is representative of the binding of the *N*-glycolyl moiety of GlcNGc (together also with the binding of glycerol, Fig. 9, *A* and *B* (supplemental Table 1)). Interestingly, the hydroxyl group of glycolic acid displayed different binding modes in the two molecules found within the unit cell of the crystal. In molecule A, it forms OH-π interactions with Tyr-282 (3.4 Å from ring center) and makes one H-bond with the thiol functionality of Cys-278. In molecule B, this hydroxyl group is within hydrogen bond distance to carboxyl group of Asp-242 and also makes OH-π interactions with Trp-337 (2.6 Å from ring center). The different binding modes observed for this hydroxyl group implies high mobility of the molecule within the crystal, which is also reflected by its relatively higher *B*-factor values. To obtain more direct structural evidence for a catalytically relevant GlcNGc species, we aimed to obtain a *Bt*GH84 bound GlcNGc structure. To achieve this aim, we built upon our previous observation of a catalytically competent oxazoline intermediate (46), and thus, by using the D242N mutant of *Bt*GH84 in combination with the substrate pNP-GlcNGc **7**, we were able to trap a glycolyl-oxazoline reaction intermediate bound within the active site. Unambiguous electron density

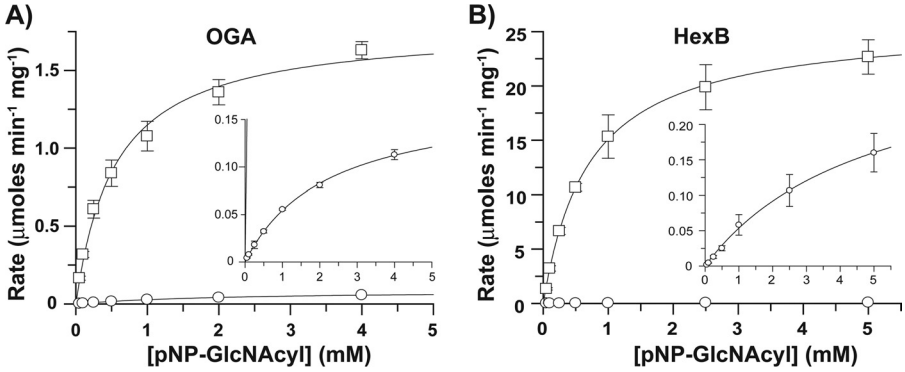


FIGURE 8. **Activity of human OGA and HexB toward pNP-GlcNGc (7).** The catalytic efficiency of human OGA (*A*) and HexB (*B*) was assessed by obtaining full Michaelis-Menten plots against pNP-GlcNAc (*squares*) and pNP-GlcNGc (*circles*). *Insets* show that rates with the glycolyl derivative were slower but did show Michaelis-Menten behavior.

TABLE 1
Enzymatic parameters describing the hydrolysis of pNP-GlcNAc and pNP-GlcNGc 7 by human OGA and HexB

Enzyme	Substrate	$V_{max}/[E]_0$	K_m	$V_{max}/[E]_0K_m$	$(V_{max}/[E]_0K_m)_{Ac}/(V_{max}/[E]_0K_m)_{Gc}$
		$\mu mol\ min^{-1}\ mg^{-1}$	mM	$\mu mol\ min^{-1}\ mg^{-1} mM^{-1}$	
OGA	pNP-GlcNAc	1.8	0.55	3.3	48
	pNP-GlcNGc	0.15	2.2	0.068	
HexB	pNP-GlcNAc	25	0.71	36	550
	pNP-GlcNGc	0.32	4.9	0.065	

revealed the glycolyl oxazoline derivative (Fig. 9*C*) bound in a $^4C_1$ chair conformation that is essentially identical to that of previously described D242N-oxazoline (Cα root mean square deviation = 0.23 Å), with almost invariant active site interactions (Fig. 9*D*). In brief, the C6 hydroxyl (carbohydrate numbering) donates one H-bond to Asp-344. The C4 hydroxyl is H-bonded to Asp-344 and Asn-372. The C3 hydroxyl receives one H-bond from the main chain carbonyl of Gly-135. The oxazoline ring nitrogen is H-bonded to Asp-242, although the oxazoline oxygen receives one H-bond from Asn-339. One notable feature of the new structure lies in the glycolyl hydroxyl group, which is closely sandwiched by Trp-337 and Tyr-282, to form an H-bond with the thiol of Cys-278 at the bottom of the active site. These two structures therefore support the hypothesis that *Bt*GH84 is able to bind and process metabolically derived *O*-GlcNGc and also offer a structural rationale for this binding.

Conclusions—These collective observations support the view that *O*-GlcNGc modification of nucleocytoplasmic proteins can occur from metabolites such as GlcNGc and GlcNGc 6-phosphate that are generated by catabolism of Neu5Gc as shown in an accompanying paper (26). Less likely, based on data from another accompanying paper (50) is that GlcNGc could arise by *de novo* synthesis of GlcNGc via the HBSP. We also find that GlcNGc introduced to cells having a functional HBSP can compete with the normal flux through the HBSP and become covalently attached to nucleocytosolic proteins as *O*-GlcNGc. Interestingly, *O*-GlcNGc can turn over within a similar time frame as *O*-GlcNAc, indicating that the enzymes involved have a fair level of tolerance for this monosaccharide and for *O*-GlcNGc. Enzyme kinetic and structural studies reveal the quantitative and structural basis for the tolerance of OGA toward *O*-GlcNGc, supporting the cell-based observations. Given the ubiquitous nature of the HBSP, OGA, and OGT in tissues and their conservation among mammals, it appears likely that GlcNGc could be found in a wide range of tissues and mammals.

The consequence of these observations and the biological significance of *O*-GlcNGc and how it might interact with *O*-GlcNAc are difficult to predict. Clearly, GlcNGc can substitute for GlcNAc as measured by its ability to rescue the growth rate of *EMEG32*$^{-/-}$ cells (Fig. 2*A*). One interesting possible scenario is that cells under certain metabolic pressures might produce various short chain fatty acids that could be incorporated to form analogs of GlcNAc, and this may either confer some adaptive benefit or be maladaptive. It is interesting in this regard that short chain *N*-acyl variants have been found on the lysine residues of histones (55). These alternative modifications are likely derived from cellular pools of propionyl-CoA and butyryl-CoA as acetyltransferases can use these substrates with reasonable efficiency (for review see Refs. 56, 57). Whether such variants have biologically significant roles, or are the consequence of the imperfection of enzymatic processes in their ability to distinguish small structural variants of substrates, remains a topic of interest. In this regard, it should be noted that although we find good support for *O*-GlcNGc being able to occur, there is no evidence to support that it is an abundant or functionally distinct species. In this regard, however, the accompanying paper (26) provides clear evidence for GlcNGc and GlcNGc 6-phosphate as true catabolic intermediates derived from Neu5Gc, which is abundant in nonhuman vertebrates. Indeed, even if GlcNGc levels are significantly lower than GlcNAc levels and the efficiency of processing of GlcNGc is lower than that of GlcNAc, principles of enzyme kinetics dictate that inevitably *O*-GlcNGc will form at some level. Given that we show here that UDP-GlcNGc and UDP-GalNGc are formed within cells from GlcNGc, it seems very possible that other glycan structures could incorporate *N*-glycolyl-hexosamines. Consistent with this view, in an accompanying paper

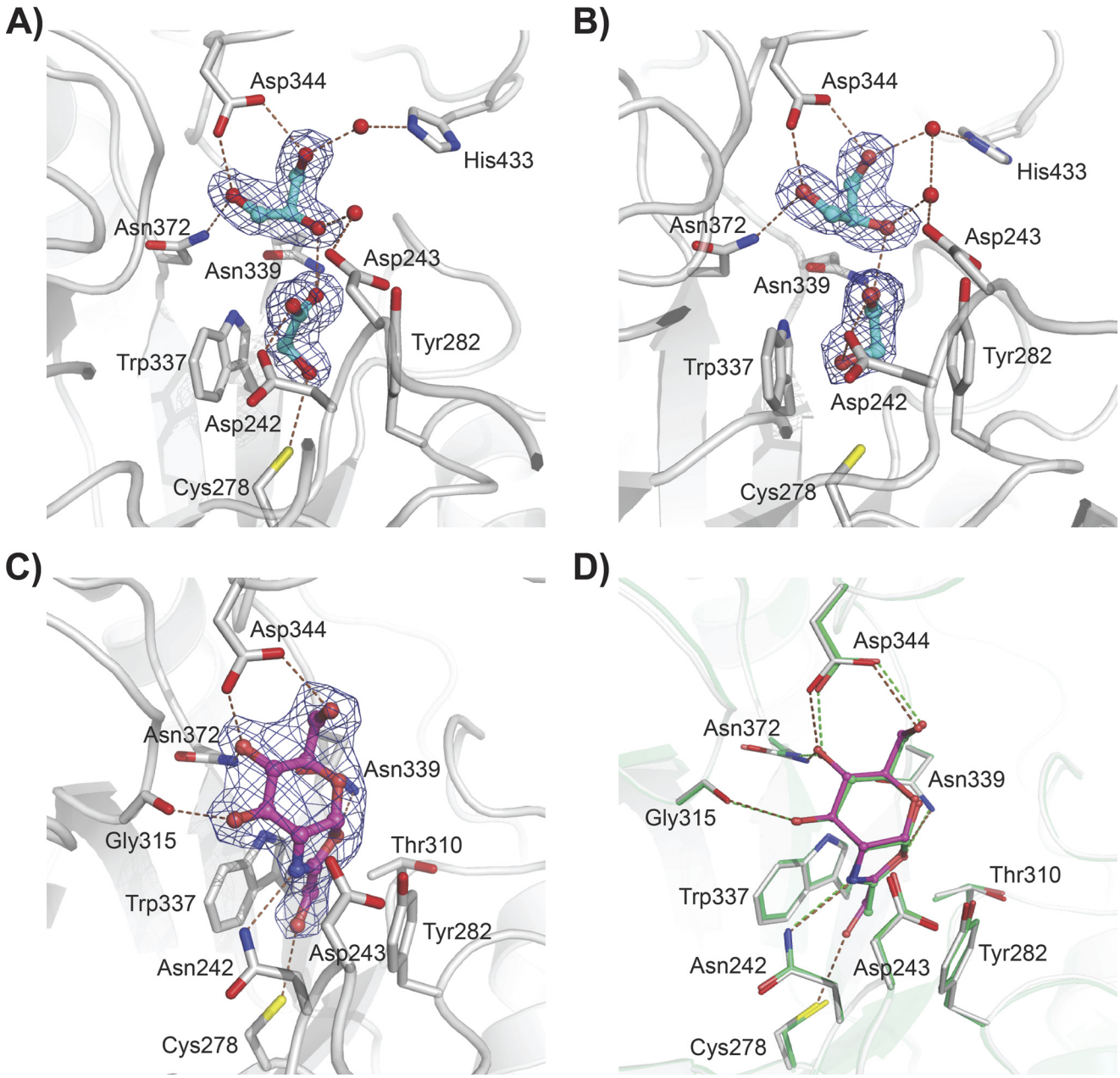


FIGURE 9. **Three-dimensional structures of glycolic acid and glycolyl oxazoline bound within the *N*-acyl binding pocket of *Bt*GH84 support the catalytic competence of OGA toward *O*-GlcNGc.** *A* and *B*, binding modes of glycerol and glycolic acid (*cyan*) to two *Bt*GH84 molecules in the crystal structure to mimic interactions with the substrate. Electron densities are maximum likelihood weighted $2F_{obs} - F_{calc}$ countered at 1σ. Hydrogen bonds, within 3.2 Å, to enzyme active site residues are shown as *dashed lines*. *C*, the glycolyl-oxazoline intermediate (*magneta*) bound to the D242N variant of *Bt*GH84. *D*, overlay of D242N-glycolyl oxazoline with D242N-oxazoline (*green*), highlighting invariant interactions in both structures. These figures were drawn with PyMOL.

(50), we show that GalNGc can be incorporated into various cell surface glycans. Notably, even enzymatic activities that are both ancient and critical for life, including DNA and protein biosynthesis, show less than complete fidelity. In response to such enzymatic promiscuity, adaptive mechanisms have evolved to correct and cope with such errors (5–7). Here, we propose that the active site pocket of OGA and the substrate promiscuity of OGA are highly conserved exactly because of the lack of complete fidelity of the HBSP and OGT. OGA therefore acts to "proofread" the incorporation of variants of GlcNAc. More generally, these observations set precedent for the possibility that other carbohydrate-processing enzymes may show similar plasticity. This possibility is a topic of interest given the observations regarding metabolic promiscuity of glycosylation pathways and glycosyltransferases made in this and the accompanying papers (26, 50).

Acknowledgments—We thank Prof. Tak Mak of the Ontario Cancer Institute for providing EMEG32 cells and Prof. Mario Pinto for providing access to capillary electrophoresis instrumentation. We thank Dr. Véronique Piller of Université d'Orléans for the pTrcHisB expression vector encoding human Agx1 bearing an N-terminal His_6 tag.

REFERENCES

1. Kunkel, T. A. (2004) DNA replication fidelity. *J. Biol. Chem.* **279,** 16895–16898
2. Kurland, C. G. (1992) Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* **26,** 29–50
3. Davies, J., Gilbert, W., and Gorini, L. (1964) Streptomycin, suppression, and the code. *Proc. Natl. Acad. Sci. U.S.A.* **51,** 883–890
4. Tawfik, D. S. (2010) Messy biology and the origins of evolutionary innovations. *Nat. Chem. Biol.* **6,** 692–696
5. Nick McElhinny, S. A., Kumar, D., Clark, A. B., Watt, D. L., Watts, B. E., Lundström, E. B., Johansson, E., Chabes, A., and Kunkel, T. A. (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat. Chem. Biol.* **6,** 774–781
6. Zaher, H. S., and Green, R. (2009) Fidelity at the molecular level. Lessons from protein synthesis. *Cell* **136,** 746–762
7. Goldsby, R. E., Lawrence, N. A., Hays, L. E., Olmsted, E. A., Chen, X., Singh, M., and Preston, B. D. (2001) Defective DNA polymerase-δ proofreading causes cancer susceptibility in mice. *Nat. Med.* **7,** 638–639
8. Hart, G., Cummings, R., Varki, A., Esko, J., and Freeze, H. (2008) *Essentials of Glycobiology*, 2d Ed., Cold Springs Harbor Laboratory Press, Cold Springs Harbor, NY
9. Thibodeaux, C. J., Melançon, C. E., and Liu, H. W. (2007) Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446,** 1008–1016
10. North, S. J., Hitchen, P. G., Haslam, S. M., and Dell, A. (2009) Mass spectrometry in the analysis of *N*-linked and *O*-linked glycans. *Curr. Opin. Struct. Biol.* **19,** 498–506
11. Angata, T., and Varki, A. (2002) Chemical diversity in the sialic acids and related α-keto acids. An evolutionary perspective. *Chem. Rev.* **102,** 439–469
12. Varki, A. (1996) "Unusual" modifications and variations of vertebrate oligosaccharides. Are we missing the flowers for the trees? *Glycobiology* **6,** 707–710
13. Boomkamp, S. D., and Butters, T. D. (2008) Glycosphingolipid disorders of the brain. *Subcell. Biochem.* **49,** 441–467
14. Hart, G. W., Housley, M. P., and Slawson, C. (2007) Cycling of *O*-linked β-*N*-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446,** 1017–1022
15. Kreppel, L. K., Blomberg, M. A., and Hart, G. W. (1997) Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique *O*-GlcNAc transferase with multiple tetratricopeptide repeats. *J. Biol. Chem.* **272,** 9308–9315
16. Lubas, W. A., Frank, D. W., Krause, M., and Hanover, J. A. (1997) *O*-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *J. Biol. Chem.* **272,** 9316–9324
17. Roquemore, E. P., Chevrier, M. R., Cotter, R. J., and Hart, G. W. (1996) Dynamic *O*-GlcNAcylation of the small heat shock protein αB-crystallin. *Biochemistry* **35,** 3578–3586
18. Dennis, R. J., Taylor, E. J., Macauley, M. S., Stubbs, K. A., Turkenburg, J. P., Hart, S. J., Black, G. N., Vocadlo, D. J., and Davies, G. J. (2006) Structure and mechanism of a bacterial β-glucosaminidase having *O*-GlcNAcase activity. *Nat. Struct. Mol. Biol.* **13,** 365–371
19. Macauley, M. S., Whitworth, G. E., Debowski, A. W., Chin, D., and Vocadlo, D. J. (2005) *O*-GlcNAcase uses substrate-assisted catalysis. Kinetic analysis and development of highly selective mechanism-inspired inhibitors. *J. Biol. Chem.* **280,** 25313–25322
20. Rao, F. V., Dorfmueller, H. C., Villa, F., Allwood, M., Eggleston, I. M., and van Aalten, D. M. (2006) Structural insights into the mechanism and inhibition of eukaryotic *O*-GlcNAc hydrolysis. *EMBO J.* **25,** 1569–1578
21. Whitworth, G. E., Macauley, M. S., Stubbs, K. A., Dennis, R. J., Taylor, E. J., Davies, G. J., Greig, I. R., and Vocadlo, D. J. (2007) Analysis of PUGNAc and NAG-thiazoline as transition state analogs for human *O*-GlcNAcase. Mechanistic and structural insights into inhibitor selectivity and transition state poise. *J. Am. Chem. Soc.* **129,** 635–644
22. Martinez-Fleites, C., He, Y., and Davies, G. J. (2010) Structural analyses of enzymes involved in the *O*-GlcNAc modification. *Biochim. Biophys. Acta* **1800,** 122–133
23. Macauley, M. S., and Vocadlo, D. J. (2010) Increasing *O*-GlcNAc levels. An overview of small molecule inhibitors of *O*-GlcNAcase. *Biochim. Biophys. Acta* **1800,** 107–121
24. Vocadlo, D. J., Hang, H. C., Kim, E. J., Hanover, J. A., and Bertozzi, C. R. (2003) A chemical approach for identifying *O*-GlcNAc-modified proteins in cells. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 9116–9121
25. Vamecq, J., Draye, J. P., and Poupaert, J. H. (1990) Studies on the metabolism of glycolyl-CoA. *Biochem. Cell Biol.* **68,** 846–851
26. Bergfeld, A. K., Pearce, O. M., Diaz, S. L., Pham, T., and Varki, A. (2012) *Metabolism of vertebrate amino sugars with N-glycolyl groups. Elucidating the intracellular fate of the non-human sialic acid N-glycolylneuraminic acid. J. Biol. Chem.* **287,** 28865–28881
27. Schauer, R., and Wember, M. (1996) Isolation and characterization of sialate lyase from pig kidney. *Biol. Chem. Hoppe-Seyler* **377,** 293–299
28. Kent, P. W., and Allen, A. (1968) The biosynthesis of intestinal mucins. The effect of salicylate on glycoprotein biosynthesis by sheep colonic and human gastric mucosal tissues *in vitro*. *Biochem. J.* **106,** 645–658
29. Fernandez-Sorensen, A., and Carlson, D. M. (1971) Purification and properties of phosphoacetylglucosamine mutase. *J. Biol. Chem.* **246,** 3485–3493
30. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy). An expert resource for glycogenomics. *Nucleic Acids Res.* **37,** D233–D238
31. Macauley, M. S., He, Y., Gloster, T. M., Stubbs, K. A., Davies, G. J., and Vocadlo, D. J. (2010) Inhibition of *O*-GlcNAcase using a potent and cell-permeable inhibitor does not induce insulin resistance in 3T3-L1 adipocytes. *Chem. Biol.* **17,** 937–948
32. Macauley, M. S., Stubbs, K. A., and Vocadlo, D. J. (2005) *O*-GlcNAcase catalyzes cleavage of thioglycosides without general acid catalysis. *J. Am. Chem. Soc.* **127,** 17202–17203
33. Tomiya, N., Ailor, E., Lawrence, S. M., Betenbaugh, M. J., and Lee, Y. C. (2001) Determination of nucleotides and sugar nucleotides involved in protein glycosylation by high performance anion-exchange chromatography. Sugar nucleotide contents in cultured insect cells and mammalian cells. *Anal. Biochem.* **293,** 129–137
34. Turnock, D. C., and Ferguson, M. A. (2007) Sugar nucleotide pools of *Trypanosoma brucei*, *Trypanosoma cruzi*, and *Leishmania major*. *Eukaryot. Cell* **6,** 1450–1463
35. Räbinä, J., Mäki, M., Savilahti, E. M., Järvinen, N., Penttilä, L., and Renkonen, R. (2001) Analysis of nucleotide sugars from cell lysates by ion-pair solid-phase extraction and reversed-phase high performance liquid chromatography. *Glycoconj. J.* **18,** 799–805
36. Feng, H. T., Wong, N., Wee, S., and Lee, M. M. (2008) Simultaneous determination of 19 intracellular nucleotides and nucleotide sugars in Chinese hamster ovary cells by capillary electrophoresis. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **870,** 131–134
37. Lagarrigue, M., Bossée, A., Bégos, A., Delaunay, N., Varenne, A., Gareil, P., and Bellier, B. (2008) Field-amplified sample stacking for the detection of chemical warfare agent degradation products in low conductivity matrices by capillary electrophoresis-mass spectrometry. *J. Chromatogr. A* **1178,** 239–247
38. Greig, I. R., Macauley, M. S., Williams, I. H., and Vocadlo, D. J. (2009) Probing synergy between two catalytic strategies in the glycoside hydrolase *O*-GlcNAcase using multiple linear free energy relationships. *J. Am. Chem. Soc.* **131,** 13415–13422
39. Cunha, A. C., Pereira, L. O., de Souza, M. C., and Ferreira, V. F. (1999) Use of protecting groups in carbohydrate chemistry. An advanced organic synthesis experiment. *J. Chem. Educ.* **76,** 79–80
40. Sinaÿ, P. (1971) Synthé du 3-*O*-(D-1-carboxyéthyl)-2-désoxy-2-glycoamido-D-glucose (acide *N*-glycolylmuramique). *Carbohydr. Res.* **16,** 113–122
41. Baumann, W., Freidenreich, J., Weisshaar, G., Brossmer, R., and Friebolin, H. (1989) Cleavage and synthesis of sialic acids with aldolase. NMR studies of stereochemistry, kinetics, and mechanisms. *Biol. Chem. Hoppe-Seyler* **370,** 141–149
42. Leslie, A. G. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EACMB Newsletter on*

Protein Crystallography, No. 26, Daresbury Laboratory, Warrington, UK
43. Collaborative Computational Project, No. 4 (1994) The CCP4 suite. Programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
44. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
45. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
46. He, Y., Macauley, M. S., Stubbs, K. A., Vocadlo, D. J., and Davies, G. J. (2010) Visualizing the reaction coordinate of an *O*-GlcNAc hydrolase. *J. Am. Chem. Soc.* **132,** 1807–1809
47. Boehmelt, G., Wakeham, A., Elia, A., Sasaki, T., Plyte, S., Potter, J., Yang, Y., Tsang, E., Ruland, J., Iscove, N. N., Dennis, J. W., and Mak, T. W. (2000) Decreased UDP-GlcNAc levels abrogate proliferation control in EMeg32-deficient cells. *EMBO J.* **19,** 5092–5104
48. Saxon, E., Luchansky, S. J., Hang, H. C., Yu, C., Lee, S. C., and Bertozzi, C. R. (2002) Investigating cellular metabolism of synthetic azidosugars with the Staudinger ligation. *J. Am. Chem. Soc.* **124,** 14893–14902
49. Diaz, S., and Varki, A. (1985) Metabolic labeling of sialic acids in tissue culture cell lines. Methods to identify substituted and modified radioactive neuraminic acids. *Anal. Biochem.* **150,** 32–46
50. Bergfeld, A., Pearce, O., Diaz, S., Lawrence, R., Vocadlo, D., Choudhury, B., Esko, J., and Varki, A. (2012) *Metabolism of Vertebrate Amino Sugars with N-glycolyl Groups. Incorporation of N-glycolylhexosamines into mammalian glycans by feeding N-glycolylgalactosamine. J. Biol. Chem.* **287,** 28898–28916
51. Chou, C. F., Smith, A. J., and Omary, M. B. (1992) Characterization and dynamics of *O*-linked glycosylation of human cytokeratin 8 and 18. *J. Biol. Chem.* **267,** 3901–3906
52. Cetinbas, N., Macauley, M. S., Stubbs, K. A., Drapala, R., and Vocadlo, D. J. (2006) Identification of Asp-174 and Asp-175 as the key catalytic residues of human *O*-GlcNAcase by functional analysis of site-directed mutants. *Biochemistry* **45,** 3835–3844
53. Mark, B. L., Mahuran, D. J., Cherney, M. M., Zhao, D., Knapp, S., and James, M. N. (2003) Crystal structure of human β-hexosaminidase B. Understanding the molecular basis of Sandhoff and Tay-Sachs disease. *J. Mol. Biol.* **327,** 1093–1109
54. Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G., and James, M. N. (2001) Crystallographic evidence for substrate-assisted catalysis in a bacterial β-hexosaminidase. *J. Biol. Chem.* **276,** 10330–10337
55. Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., Falck, J. R., Peng, J., Gu, W., and Zhao, Y. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol. Cell. Proteomics* **6,** 812–819
56. Berndsen, C. E., and Denu, J. M. (2008) Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* **18,** 682–689
57. Smith, B. C., Hallows, W. C., and Denu, J. M. (2008) Mechanisms and molecular probes of sirtuins. *Chem. Biol.* **15,** 1002–1013