Dynamic Glycosylation of Nuclear and Cytosolic Proteins

CLONING AND CHARACTERIZATION OF A UNIQUE O-GlcNAc TRANSFERASE WITH MULTIPLE TETRATRICOPEPTIDE REPEATS

Lisa K. Kreppel‡, Melissa A. Blomberg, and Gerald W. Hart§

From the Department of Biochemistry and Molecular Genetics Schools of Medicine/Dentistry, University of Alabama at Birmingham Station, Birmingham, Alabama 35294 and the Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

O-Linked N-acetylglucosamine (O-GlcNAc) glycosylation is a dynamic modification of eukaryotic nuclear and cytosolic proteins analogous to protein phosphorylation. We have cloned and characterized a novel gene for an O-GlcNAc transferase (OGT) that shares no sequence homology or structural similarities with other glycosyltransferases. The OGT gene is highly conserved (up to 80% identity) in all eukaryotes examined. Unlike previously described glycosyltransferases, OGT is localized to the cytosol and nucleus. The OGT protein contains multiple tandem repeats of the tetratricopeptide repeat motif. The presence of tetratricopeptide repeats, which can mediate protein-protein interactions, suggests that OGT may be regulated by protein interactions that are independent of the enzyme’s catalytic site. The OGT is also modified by tyrosine phosphorylation, indicating that tyrosine kinase signal transduction cascades may play a role in modulating OGT activity.

Unlike other forms of protein glycosylation, serine (threonine)-O-linked N-acetylglucosamine (O-GlcNAc) is found primarily in the cytoplasm and nucleus, and is not modified or elongated to more complex structures (1, 2). Since it was first described in 1984 (3), the O-GlcNAc modification (termed O-GlcNAcylation) has been found on a myriad of eukaryotic nuclear and cytosolic proteins, including RNA polymerase II and its transcription factors, nuclear pore proteins, viral proteins, cytoskeletal proteins, tumor suppressor proteins, and oncoproteins (reviewed in Refs. 1 and 2).

Direct evidence is rapidly accumulating in support of the role of O-GlcNAcylation as a regulatory modification. O-GlcNAc appears to be as abundant as phosphorylation, and several of the known sites of attachment are similar to those used by proline-directed kinases (4, 5). O-GlcNAcylation is a dynamic modification exhibiting properties more like phosphorylation than typical O- and N-linked glycosylation (1, 2). The turnover rate of the O-GlcNAc moiety on cytokeratins (6) and the small heat shock protein aB-crystallin (7) is much higher than the turnover rate of the protein. O-GlcNAcylation also has been shown to regulate a number of cellular functions. For example, recent studies have shown the following. 1) O-GlcNAcylation modulates the DNA binding activity of the p53 tumor suppressor (8). 2) O-GlcNAcylation of p67 regulates protein synthesis by controlling the phosphorylation state of the elongation initiation factor 2 (eIF-2α) (9, 10). 3) O-GlcNAcylation of the head domain of neurofilament-H appears to regulate neurofilament assembly (11). 4) The O-GlcNAc and phosphate modifications of the RNA polymerase II COOH-terminal domain are reciprocal and are likely to regulate transcription (12, 13). 5) O-GlcNAc has a reciprocal relationship with phosphorylation at the site on the c-Myc protein, which has been implicated in modulating its oncogenic activity (14).

Consistent with the dynamic nature of O-GlcNAcylation, both a UDP-N-acetylglucosamine:peptide N-acetylglucosaminyl-transferase (O-GlcNAc transferase) (15), specific for the attachment of O-GlcNAc to proteins, and a soluble N-acetyl-β-D-glucosaminidase with a neutral pH optima (O-GlcNacase) (16), specific for the removal of O-GlcNAc from proteins have been purified and characterized. These two enzymes appear to regulate the attachment and removal of O-GlcNAc in much the same way that kinases and phosphatases regulate protein phosphorylation. Taken together these observations suggest that O-GlcNAc may play a role in modulating either the phosphorylation state or the assembly and disassembly of multimeric protein complexes in several key cellular systems, including transcription, nuclear transport, and cytoskeletal organization.

The O-GlcNAc transferase (OGT) (EC 2.4.1.1) purified from rat liver cytosol appears to be a heterotrimer composed of two catalytic 110-kDa (p110) subunits and one 78-kDa (p78) subunit (17). Here we describe the cloning and characterization of the gene encoding the catalytic p110 subunit. The gene is highly conserved throughout evolution, consistent with the ubiquitous nature of the O-GlcNAc modification. We also find that, like many other regulatory proteins, OGT contains several tandem repeats of the tetratricopeptide repeat (TPR) motif (reviewed in Refs. 18 and 19), suggesting that OGT can interact with other proteins via the TPR domain, to form a regulatory complex. Examination of the posttranslational modifications of OGT shows that the enzyme is modified by both O-GlcNAc and tyrosine phosphorylation. The subcellular localization of the cloned gene is consistent with O-GlcNAcylation as a nuclear and cytosolic modification.

‡ This work was supported by National Institutes of Health Grant HD13563 (to G. W. H.), by the Juvenile Diabetes Foundation and Filly 50 Foods Inc., and by a gift from Oxford Glycosystems Inc. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: UAB, Dept. of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35294. E-mail: gwhart@bmg.bhs.ua.edu.

© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.

Printed in U.S.A.

This paper is available on line at http://www-jbc.stanford.edu/jbc/

9308
**EXPERIMENTAL PROCEDURES**

**Preparation of Tryptic Peptides**—The O-GlcNAc transferase was purified and concentrated on a Q-Sepharose column as described previously (17). The purified protein was separated by SDS-PAGE (20). The material corresponding to the 110-kDa subunit was visualized, excised, and subjected to in-gel protease digestion with trypsin (Boehringer Mannheim sequencing grade or Worthington soyphyenylalanyl chloromethyl ketone-treated trypsin) by the method of Rosenfeld et al. (21). The resulting peptides were separated by reverse phase high performance liquid chromatography (RP-HPLC). In addition, a second sample was electroblotted to Immobilon-P™ (Millipore) after SDS-PAGE separation and the protein stained with Amido Black 10B (Sigma) as per manufacturer’s recommendations. The band corresponding to the 110-kDa subunit was excised and submitted to the Harvard Microchem protein sequencing facility (Boston, MA) for both NH2-terminal sequencing and internal sequence analysis.

**RP-HPLC**—The tryptic peptides were purified by several rounds of RP-HPLC. A Vydac 5-μm C18 column (4.6 × 250 mm) or a Rainin Microsorb-MV 5-μm C18 column (4.6 × 250 mm) was used for first and second-dimension RP-HPLC. A Vydac narrow-bore C18 column (2.1 × 150 mm) was used for third dimension-RP-HPLC. The peptides were bound to the column in either 0.05% triton x-100 or 0.1% phosphoric acid containing 100 mM sodium perchlorate at pH 2 or pH 7, (22), and eluted with a linear gradient from 0 to 60% acetonitrile.

**Peptide Sequencing**—The RP-HPLC-purified peptides were sequenced by gas phase automated Edman degradation on either a Protein Instruments (Tarzana, CA) model PI 2090E microsequencing system, or an Applied Biosystems Inc. (Foster City, CA) model 470A gas phase sequencer.

**Rabbit Antiserum**—Polyclonal antibodies to the OGT protein were generated as follows. His-tagged protein was expressed in *Escherichia coli* using the pTrExHis vector, and the protein was purified as described below. The purified protein was separated by SDSPAGE, visualized, and excised from the gel as described below. The gel slices were homogenized and used directly as immunogens by Hazelton Research Products Inc. (San Diego, CA) to produce polyclonal antisera in two rabbit sera designated AL-24 and AL-25. Immunoglobulin G (IgG) was purified by passing the rabbit antisera over a protein A-Sepharose column (Pharmacia) as per manufacturer’s recommendations.

**Western Blot Analysis**—Crude protein extracts were prepared either from the dissected tissue of 3–6-month-old Harlan Sprague Dawley rats or from transfected HEK293 cells by following the first two steps of the purification protocol as described previously (17). The proteins were separated by SDS-PAGE and transferred to polyvinylidene difluoride membrane (23). Purified rabbit polyclonal IgG AL-24 (1:5000), or monoclonal anti-phosphotyrosine (Sigma, 1:2000) was used as a primary antibody with anti-rabbit or anti-mouse IgG coupled to horseradish peroxidase (Amersham) as the secondary antibody (1:20,000 dilution). Detection of the horseradish peroxidase activity was enhanced chemiluminescence (ECL) and fluorography as described by the manufacturer. In addition, a second sample was electrophoretically transferred to a nitrocellulose membrane, and probed with the PCR-22b fragment (see above) as described (25). Two rat liver cDNA λ-ZAPIII clones were subjected to double-stranded DNA sequencing using deoxyadenosine 5′-[35S]thiophosphate and Sequenase II (U. S. Biochemical Corp.), as described by the manufacturer. Additional sequence information of the cDNA clones was obtained by automated DNA sequencing on an Applied Biosystems model 373A automated DNA sequencer.

**Screening the cDNA Libraries**—Two rat liver cDNA λ-ZAPIII libraries (nos. 936513 and 936507, Stratagene) were plated on XL-1 Blue host cells (Stratagene) and screened by hybridization in 50% formamide (25), with the PCR-22b probe (see above) labeled to a specific activity of ~8 × 10^9 dpm/μg using the Ready-To-Go DNA Labeling Kit (Pharmacia). Hybridization was performed at 42 °C, and the filters were washed in 0.2 × SSC at 70 °C. Nine positive clones were isolated, ranging in size from 1–3.2 kb (designated H1–H9). The longest H1 (Fig. 1A), contained a putative start site, 31 base pairs upstream of untranslated DNA, and an open reading frame encoding 958 residues; however, no in-frame stop codon was present. A rat hippocampus λ-ZAPII library kindly provided by Anthony Lanahan (Johns Hopkins School of Medicine, Baltimore, MD) was screened as above, except that a gel-purified SacI restriction enzyme fragment representing the 5′-end of the partial clone was used as probe (Fig. 1A); all other conditions were identical. Five clones were isolated (designated LTP1–LTP5), ranging in size from 2.5 to 4.6 kb. All of these clones overlapped the H1 clone by 0.3–2.2 kb, and all of them contained a poly(A) tail (data not shown). Positive clones from each library were screened. Positive clones were subjected to in vivo excision protocol according to the manufacturers’ directions, to recover a pBlueScript plasmid containing the cDNA clone of interest for further analysis. A full-length cDNA clone was constructed by ligation of the 5′-end of the H1 clone to the 3′-end of LTP3 at the unique Xhol site (Fig. 1A).

**DNA and RNA Blot Analysis**—DNA and RNA analysis was performed on a prepared Zoo-Blot (Clontech) as per manufacturer’s directions with a PCR-22b probe labeled as above. In addition, total genomic DNA was isolated from HEK293 cells (25), digested to completion with EcoRI, resolved on a 0.8% agarose gel, and transferred to a Nytran (Schleicher & Schuell) filter according to manufacturer’s directions. Hybridization was performed at 42 °C, and the filters were washed in 0.2 × SSC at 70 °C. Nine positive clones were isolated, ranging in size from 1–3.2 kb (designated H1–H9). The longest H1 (Fig. 1A), contained a putative start site, 31 base pairs upstream of untranslated DNA, and an open reading frame encoding 958 residues; however, no in-frame stop codon was present. A rat hippocampus λ-ZAPII library kindly provided by Anthony Lanahan (Johns Hopkins School of Medicine, Baltimore, MD) was screened as above, except that a gel-purified SacI restriction enzyme fragment representing the 3′-end of the partial clone was used as probe (Fig. 1A); all other conditions were identical. Five clones were isolated (designated LTP1–LTP5), ranging in size from 2.5 to 4.6 kb. All of these clones overlapped the H1 clone by 0.3–2.2 kb, and all of them contained a poly(A) tail (data not shown). Positive clones from each library were screened. Positive clones were subjected to in vivo excision protocol according to the manufacturers’ directions, to recover a pBlueScript plasmid containing the cDNA clone of interest for further analysis. A full-length cDNA clone was constructed by ligation of the 5′-end of the H1 clone to the 3′-end of LTP3 at the unique Xhol site (Fig. 1A).

**DNA and RNA Blot Analysis—Genomic**—DNA and RNA analysis was performed on a prepared Zoo-Blot (Clontech) as per manufacturer’s directions with a PCR-22b probe labeled as above. In addition, total genomic DNA was isolated from HEK293 cells (25), digested to completion with EcoRI, resolved on a 0.8% agarose gel, and transferred to a Nytran (Schleicher & Schuell) filter according to manufacturer’s directions. Hybridization was performed at 42 °C, and the filters were washed in 0.2 × SSC at 70 °C. Nine positive clones were isolated, ranging in size from 1–3.2 kb (designated H1–H9). The longest H1 (Fig. 1A), contained a putative start site, 31 base pairs upstream of untranslated DNA, and an open reading frame encoding 958 residues; however, no in-frame stop codon was present. A rat hippocampus λ-ZAPII library kindly provided by Anthony Lanahan (Johns Hopkins School of Medicine, Baltimore, MD) was screened as above, except that a gel-purified SacI restriction enzyme fragment representing the 3′-end of the partial clone was used as probe (Fig. 1A); all other conditions were identical. Five clones were isolated (designated LTP1–LTP5), ranging in size from 2.5 to 4.6 kb. All of these clones overlapped the H1 clone by 0.3–2.2 kb, and all of them contained a poly(A) tail (data not shown). Positive clones from each library were screened. Positive clones were subjected to in vivo excision protocol according to the manufacturers’ directions, to recover a pBlueScript plasmid containing the cDNA clone of interest for further analysis. A full-length cDNA clone was constructed by ligation of the 5′-end of the H1 clone to the 3′-end of LTP3 at the unique Xhol site (Fig. 1A).

**Expression of the Ogt cDNA**—The coding region of the GTF was assembled from two cDNA clone fragments at a unique Xhol site. The 5′- untranslated region was removed and replaced with a linker restoring the start codon, and a BamHI site was added 5′ to the start (Fig. 1B). The coding region was then subcloned as a BamHI/HindIII fragment into the polylinker of the pTrExHis Xpress vector (Invitrogen). Destination plasmid pLK51, for expression in *E. coli*, or into pGWL (a kind gift of Mike Lee) designated plK61, for transient expression in mammalian cells. pLK51 was transformed into the XL-1 Blue (Stratagene) strain of *E. coli* and grown to midlog phase, and protein expression was induced by the addition of 1 mM final concentration of isopropyl-1-thio-β-D-galactopyranoside. Cells were harvested 6 h after induction and protein was purified on Hi-Trap Chelating Columns (Pharmacia) under urea denaturing conditions as per manufacturer’s instructions. HepK293...
cells were grown in six-well plates for protein expression assays, or on
glass coverslips for immunolocalization in Dulbecco’s modified Eagle’s
medium, 10% fetal bovine serum until 50% confluence. The cells were
then transfected with pLK61 by calcium phosphate-mediated transfec-
tion (27) and grown for an additional 24–48 h to allow protein
expression.

Immunofluorescence—Transfected HEK293 cells (see above) or CHO
cells, grown on glass coverslips in Dulbecco’s modified Eagle’s medium/
Ham’s F-12 (1:1) supplemented with 10% fetal bovine serum until 50%
confluence, were washed twice in serum-free medium and fixed in 4%
formaldehyde for 30 min. The cells were then washed four times in
phosphate-buffered saline, pH 7.5 (PBS), and permeabilized in 0.5%
Triton X-100 in PBS for 5 min. Cells were washed in PBS and blocked
with goat serum and 3% bovine serum albumin in PBS (1:3) for 15 min
at 37 °C. The cells were then incubated in primary antibody for 30 min
at 37 °C. Excess primary antibody was washed away with four 10 min
incubations in PBS, and the cells were incubated in secondary antibody,
at room temperature for 30 min in the dark. The cells were rinsed as
for the primary antibody and mounted onto slides in 0.1% paraphe-
nylene diamine in 90% glycerol. Secondary antibody alone gave no sig-
nal, and no signal was observed in non-permeabilized cells. Primary
antibodies, AL-25 or preimmune, were used at a 1:500 dilution. The
secondary antibodies, FITC-conjugated goat anti-rabbit IgG (Jackson
ImmunoResearch Laboratories) were diluted 1:200; DAPI stain was
used at a final concentration of 0.1 μg/ml. All antibodies were diluted in
3% bovine serum albumin in PBS.

RESULTS

The Gene Encoding p110 Is Evolutionarily Conserved—Four-
teen unique peptide sequences were obtained from protease
digests of the p110 subunit of OGT purified from rat liver
(underlined in Fig. 1A). Polymerase chain reaction amplification
and standard cDNA library screening techniques were
combined to clone the gene encoding the p110 subunit (see
“Experimental Procedures”). Two overlapping clones were iso-
lated and the open reading frame was reconstructed at a con-
venient restriction site (Fig. 1B). The full-length cDNA
contains a single open reading frame encoding a protein of 1037
residues, designated p110OGT.

Computer searches of the standard GenBank™ data bases
using the BLAST algorithm (28) revealed 61% sequence iden-
tity between p110OGT and a hypothetical 1194 residue protein
encoded at locus K04G7.3 of
Caenorhabditis elegans
(accession
number U21320). A matrix plot of the predicted
C. elegans
protein and p110OGT is shown in Fig. 2A. The homology ex-
tends through the entire clone with regions as long as 350
residues sharing
.80% identity. A recently cloned gene in
Arabidopsis, SPINDLY, involved in gibberellin signal trans-
duction (29) also shares extensive homology with p110OGT throughout the entire coding region. Thus, both K04G7.3 and
SPINDLY are likely to encode homologues of p110OGT.
In addition, searches of the dbEST data base of expressed sequence
tags (30) revealed homology between p110OGT and the concep-
tional translation products of expressed sequence tags from hu-
man (R7594; 93% identity over 414 residues),
Schistosoma mansoni
(T14553; 65% identity over 442 residues), and rice
(D24403; 67% identity over 326 residues).

The amino-terminal portion of p110OGT shares homology with a diverse group of proteins all containing a common motif
designated the TPR motif (18, 19), while the carboxyl terminus
shares no significant homology to any known protein in the
data bases. Thus, the p110OGT appears to consist of two dis-
tinct domains: the amino-terminal 463 residues containing
11.5 tandem repeats of the TPR motif (Fig. 1C), and the car-

FIG. 1. Restriction map and deduced amino acid sequence of the p110OGT clone. Panel A, deduced amino acid sequence of the p110OGT
clone. The 14 peptide sequences obtained from purified rat liver p110 are underlined. The region covered by the probe PCR22b is indicated by
brackets, and the predicted receptor tyrosine kinase phosphorylation site is outlined. Panel B, restriction map of the two overlapping clones
isolated. The coding region is represented by the black box, and the start and stop codons are indicated. The two probes used in these studies are
represented by gray stippled boxes. Panel C, alignment of the 11.5 TPR motifs in the amino-terminal half of p110OGT. The TPR consensus motif
is indicated at the bottom, and the conserved residues are printed in bold type. Residues not a part of the TPR consensus sequence but conserved
between the repeats are underlined.
OGT activity is immunoprecipitated from rat liver extracts by an antibody against recombinant p110\textsuperscript{OGT} expressed in E. coli—Polyclonal rabbit antibody (designated AL-25) was prepared against purified, recombinant p110\textsuperscript{OGT} overproduced in E. coli (see "Experimental Procedures"). AL-25 immunoglobulin G (IgG) is highly specific for the OGT and shows no cross-reactivity to other proteins present in partially purified preparation of rat liver OGT designated the Q-Sepharose pool (17) (Fig. 3A, compare lanes 1 and 2). Preimmune IgG shows no activity (data not shown). On Western blots, AL-25 antibody recognizes both the p110 and the p78 subunits of the rat liver OGT (lane 2), suggesting that p110 and p78 are related at the polypeptide level. Similar results were obtained with antibodies raised against synthetic peptides derived from the p110 subunit sequence (data not shown).

Both the p110 and p78 subunits of the native OGT are immunoprecipitated from the Q-Sepharose pool using AL-25, while no protein is precipitated using preimmune IgG (lanes 3 and 4). The pellets and supernatants from the immunoprecipitation were assayed for OGT enzyme activity (17). OGT enzymatic activity is precipitated from the Q-Sepharose pool using AL-25, while no activity is precipitated by buffer alone or preimmune IgG (Fig. 3B). These studies demonstrate that the cloned cDNA indeed represents the p110 subunit of the rat OGT.

Levels of OGT RNA, Protein, and Activity Vary in Different Tissues—Northern blot analysis (Fig. 4A) indicates that there are four transcripts ~of 8.0, 6.0, 4.2, and 1.7 kb, present in all rat tissues examined thus far. The 6.0-kb transcript is closest in size to the cloned cDNA (5.7 kb). The larger 8.0-kb transcript may be an alternate splicing product containing additional exons, as is seen in C. elegans, which has two distinct cDNAs representing alternative splicing events at the 5-prime end of

**Fig. 2. OGT is highly conserved.** Panel A, dot blot comparison of the predicted open reading frame K04G7.3 from C. elegans to the deduced amino acid sequence of the p110 cDNA clone. Panel B, Southern blot analysis of rat genomic DNA digested with BamHI (B), EcoRI (E), HindIII (H), PstI (P), SacI (S), and probed with PCR22b (Fig. 1). Molecular size standards are indicated to the left. Panel C, Southern blot analysis of genomic DNA from rat (Rt), mouse (M), dog (D), cow (C), rabbit (R), and human (H) probed with PCR22b (Fig. 1). Molecular size standards are indicated to the left.
the message. The longer message produces a 130-kDa protein as predicted by Wilson et al., while the smaller message produces a 112-kDa protein. The pattern of protein expression was examined by Western blot analysis of 30% ammonium sulfate cytosolic pellets (30% pellet) from rat tissues using the AL-25 antibody. The positions of the p110 and p78 subunits of OGT as purified from rat liver are indicated by the arrows. Panel C, the extracts from panel B were assayed for OGT activity. Abbreviations for all panels are as follows: B, brain; H, heart; K, kidney; L, liver; L, lung; M, muscle; T, thymus; U/O, uterus and ovary; S, spleen.

OGT Is Modified by Tyrosine Phosphorylation and O-GlcNAcylation—Western blot analysis of the Q-Sepharose pool using an anti-phosphotyrosine antibody shows that the p110 and p78 subunits are immunoreactive. This reactivity is blocked by the addition of 10 mM phosphotyrosine (Fig. 5A, compare lanes 1 and 2), but not 10 mM tyrosine (data not shown). Similar experiments using antibodies against phosphoserine and phosphothreonine showed no immunoreactivity (data not shown). Examination of the p110Ogt amino acid sequence indicates that there is only one well conserved receptor-protein-tyrosine kinase phosphorylation site, Tyr979 (32, 33) (outlined in Fig. 1A).

To determine if OGT was itself modified by O-GlcNAC, highly purified OGT was probed with galactosyltransferase (see “Experimental Procedures”). Galactosyltransferase is a specific probe for terminal GlcNAc residues (24, 34) that is commonly used to detect O-GlcNAC by covalently labeling the GlcNAC with UDP-[3H]galactose. Both the p110 and the p78 subunits are labeled with [3H]galactose (Fig. 5A, lane 3), indicating that they are modified by GlcNAC. The labeled p110 band was excised from the gel and subjected to alkaline β-elimination. The label was released by this treatment indicating that the sugar was an O-linked glycan (data not shown). The released sugars were then analyzed using HPAE-PAD chromatography (see “Experimental Procedures”). The radioactivity was seen to migrate with the disaccharide alditol of Galβ1,3GalNAcitol (a), Galβ1,3GlcNAcitol (b), Galβ1,4GalNAcitol (c), and Galβ1,4GlcNAcitol (d).

**Fig. 4. Tissue distribution of OGT message, protein, and activity.** Panel A, total RNA (25 μg) was probed with the PCR22b DNA fragment. **Panel B**, 30% ammonium sulfate pellets of tissue cytosol were desalted and electrophoresed on a 7.5% polyacrylamide gel, and immunoblotted with AL-25. The positions of the p110 and p78 subunits of OGT as purified from rat liver are indicated by the arrows. Panel C, the extracts from panel B were assayed for OGT activity. Abbreviations for all panels are as follows: B, brain; H, heart; K, kidney; L, liver; L, lung; M, muscle; T, thymus; U/O, uterus and ovary; S, spleen.

**Fig. 5. Posttranslational modifications of OGT.** Panel A, partially purified OGT was electrophoresed on a 7.5% polyacrylamide gel and immunoblotted with α-phosphotyrosine antibody alone (lane 1), or α-phosphotyrosine antibody in the presence of 10 mM free phosphotyrosine (lane 2). Highly purified OGT was labeled with [3H]galactose using galactosyltransferase as a probe for terminal GlcNAc residues, electrophoresed on a 7.5% polyacrylamide gel, and visualized by fluorography (lane 3). The positions of the p110 and p78 OGT subunits are indicated by the arrows. The galactosyltransferase is also modified by terminal GlcNAc residues, and will label itself. Autogalactosylated galactosyltransferase is indicated by the arrowhead. Panel B, the elution profile of the alkaline β-elimination product of the [3H]galactose labeled p110 separated by HPAE-PAD. The elution positions of common disaccharide alditols are indicated by arrowheads: Galβ1,3GalNAcitol (a), Galβ1,3GlcNAcitol (b), Galβ1,4GalNAcitol (c), and Galβ1,4GlcNAcitol (d).
The OGT cloned in the present study displays several features that are unique and provides clues with respect to the general functional significance of the O-GlcNAc modification. 1) The high evolutionary conservation of the enzyme suggests that it has a fundamental cellular function. 2) The enzyme’s nuclear and cytosolic localization is consistent with its action on a myriad of proteins in both compartments. 3) The tyrosine phosphorylation of the enzyme implies that it may be regulated by one or more of the receptor tyrosine kinases, linking O-GlcNAcylation to signal transduction cascades (37, 38). 4) The presence of multiple TPR repeats suggests a two-binding site model for the regulation of the O-GlcNAcylation of proteins (Fig. 8).

The Ogt gene described above represents a novel glycosyltransferase that has no structural or sequence similarities to any previously described glycosyltransferase (39). It has been highly conserved among higher eukaryotes such as rats, nematodes, and plants. This level of conservation indicates strong evolutionary pressure and suggests that Ogt encodes a protein with an essential cellular function. However, aside from the common TPR motif domain, p110<sup>OGT</sup> does not share any significant homology at the primary sequence level with any protein in the Saccharomyces cerevisiae data bases. OGT also shares no sequence homology with the α-toxin from Clostridium novyi that catalyzes the incorporation of O-GlcNAc into the Rho family of proteins in a manner very analogous to OGT (40). However, the α-toxin does share some sequence homology with an uncharacterized open reading frame from yeast (accession no. Z73530). Thus, a protein unrelated to p110<sup>OGT</sup> at the amino acid level could perform a similar enzymatic function in yeast or other eukaryotes.

Southern blot analysis and sequence comparisons indicate that Ogt is not a member of a closely related gene family. However, O-GlcNAc is found on a diverse group of proteins at a multitude of glycosylation sites. Thus, it seems unlikely that one enzyme could be responsible for the specific addition of O-GlcNAc to all these proteins. We cannot rule out the possibility that a family of OGT proteins exists, which, like the Golgi glycosyltransferases, share significant sequence homology, only structural similarity (39, 41). Alternatively, there may exist a mechanism for the regulation of a single OGT enzyme, which could confer both temporal and substrate specificity in response to cellular signals.

While there is only one Ogt gene, there are multiple transcripts and proteins related to p110<sup>OGT</sup>, some of which are tissue-specific. These related proteins likely arise from one gene by a combination of alternative RNA splicing and specific proteolysis. The presence of the p110<sup>OGT</sup> subunit in nearly every tissue examined leads us to postulate that this form of the enzyme provides the majority of the basal cellular OGT activity. However, the levels of activity in the various tissues do not always correlate with the p110 protein and RNA levels, indicating that additional factors regulate the activity of OGT. These additional factors may be limiting when the p110<sup>OGT</sup> is overexpressed in mammalian cells. Thus we do not see a proportional increase in OGT activity with protein expression.

The expression of tissue-specific forms of OGT is one mech-
anism by which the substrate specificity of the enzyme could be regulated. Posttranslational modifications of OGT present another mechanism for modulating activity or specificity. We have shown that OGT is modified by O-GlcNAcylation. In addition, we have shown that OGT is modified by tyrosine phosphorylation. A single receptor protein-tyrosine kinase phosphorylation consensus site is present in the putative catalytic domain, where it could act as a regulatory modification modulating the activity of OGT in response to signal transduction cascades (38).

The presence of TPR motifs in p110OGT is interesting, as TPRs have been found in a large number of proteins of diverse function and are believed to play a role in modulating a variety of cellular processes, including cell cycle (42–44), transcription regulation (45–47), and protein transport (48). Direct evidence for TPR-mediated protein-protein interactions regulating cellular functions is seen for the yeast transcription factor Cyc8. Cyc8 contains 10 TPR domains that are directly involved in recruitment of the Cyc8-Tup1 co-repressor regulating transcription from a distinct set of genes (47). Other examples are the yeast Cdc proteins: Cdc16p, Cdc23p, and Cdc27p, which directly interact with each other via their TPRs in a sequence-dependent manner during mitosis. Mutational analysis of the TPR domains of the Cdc proteins shows that a given TPR modulates a specific protein interaction (43). The ability of TPRs to regulate cellular processes via protein-protein interactions suggests that the TPR motifs of p110OGT could mediate specific protein interactions with accessory proteins, thereby modulating the activity or specificity of OGT. The tyrosine phosphorylation and O-GlcNAcylation modifications of OGT would provide an additional level of regulation.

A model for OGT regulation, combining the TPR accessory proteins and the posttranslational modifications, is presented in Fig. 8. In this model OGT has a basal level of activity for a narrow range of substrates. Binding of TPR accessory proteins would allow O-GlcNAcylation of additional specific substrates. The basal activity of OGT is up-regulated by changes in the phosphorylation or O-GlcNAcylation state of the protein. This model is not without precedence, as both the activity and specificity of RNA polymerase II are regulated by a large array of transcription factors as well as posttranslational modifications. Although RNA polymerase II does not contain TPR motifs, many of the transcription factors required for activation bind directly to the protein (reviewed in Refs. 49–51). In addition, both O-GlcNAcylation (12, 13, 52) and phosphorylation (52) of the COOH-terminal domain of RNA polymerase II have been documented. Additional study of OGT will allow further elucidation of the mechanisms regulating O-GlcNAcylation and will facilitate the direct evaluation of O-GlcNAc’s functions in cellular metabolism.

Acknowledgments—We thank Bill Kelly for work on the characterization of the C. elegans transcripts, Anthony Lanahan from the Department of Neuroscience at Johns Hopkins Medical School generously providing the rat hippocampus library, Betty Jean Earles for peptide sequencing and synthesis, and all the members of the Hart laboratory for helpful discussions.
REFERENCES

1. Hart, G. W., Kreppel, L. K., Comer, F. C., Arnold, C. S., Snow, D. S., Ye, Z., Chen, X., DellaManna, D., Caine, D. S., Earles, B. J., Akimoto, Y., Cole, R. N., and Hayes, B. K. (1990) Glycobiology 6, 711–716
2. Hart, G. W. (1997) Annu. Rev. Biochem., in press
3. Torres, C., and Hart, G. W. (1984) J. Biol. Chem. 259, 3308–3317
4. Pearson, R. B., and Kemp, B. E. (1991) Methods Enzymol. 200, 62–81
5. Kemp, B. E., and Pearson, R. B. (1991) Methods Enzymol. 200, 121–134
6. Chou, C.-F., Smith, A. J., and Omary, M. B. (1992) J. Biol. Chem. 267, 3901–3906
7. Roquemore, E. P., Cherrier, M. R., Cotter, R. J., and Hart, G. W. (1996) Biochemistry 35, 3578–3586
8. Shaw, P., Freeman, J., Bovey, R., and Iggo, R. (1996) Oncogene 12, 921–930
9. Chakrabarti, D., Wylie, D. E., and Gupta, N. K. (1989) J. Biol. Chem. 264, 4350–4354
10. Datta, B., Ray, M. K., Chakrabarti, D., Wylie, D. E., and Gupta, N. K. (1994) Trends Biochem. Sci. 19, 848–845
11. Dong, L.-Y., Xu, Z.-S., Hart, G. W., and Cleveland, D. W. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 8049–8057
12. Kelly, W. G., Dahmus, M. E., and Hart, G. W. (1993) J. Biol. Chem. 268, 10416–10424
13. Comer, F. I., and Hart, G. W. (1996) FASEB J. 10, A1119
14. Chou, T.-Y., Hart, G. W., and Dang, C. V. (1995) J. Biol. Chem. 270, 18961–18965
15. Haltiwanger, R. S., Holt, G. D., and Hart, G. W. (1990) J. Biol. Chem. 265, 2563–2568
16. Dong, L.-Y., and Hart, G. W. (1994) J. Biol. Chem. 269, 19321–19330
17. Haltiwanger, R. S., Blomberg, M. A., and Hart, G. W. (1992) J. Biol. Chem. 267, 9085–9093
18. Goebel, M., and Yanagida, M. (1991) Trends Biochem. Sci. 16, 173–177
19. Lamb, J. R., Tugendreich, S., and Hieter, P. (1995) Trends Biochem. Sci. 20, 257–259
20. Laemmli, U. K. (1970) Nature 227, 680–685
21. Ralph, J. H., Capdevielle, J., Guillomet, J. C., and Ferrara, P. (1992) Anal. Biochem. 203, 171–179
22. Meck, J. L. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 1632–1636
23. Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4350–4354
24. Roquemore, E. P., Chou, T.-Y., and Hart, G. W. (1994) Methods Enzymol. 230, 443–460
25. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G., and Struhl, K. (1989) Current Protocols in Molecular Biology, John Wiley and Sons, New York
26. Boyle, J. S., and Lew, A. M. (1995) Trends Genet. 11, 8
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor NY
28. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
29. Jacobsen, S. E., Binkowski, K. A., and Olszewski, N. E. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 9292–9296
30. Boguski, M. S., Lowe, T. M., and Tolstoshev, C. M. (1993) Nat. Genet. 4, 332–333
31. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Benfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laistner, N., LaFrette, P., Lightning, J., Lloyd, C., Mortimore, B., O’Callaghan, M., Parsons, J., Perry, C., Rikken, L., Roopra, A., Saunders, D., Shawknree, R., Sims, M., Smalbon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P. (1994) Nature 368, 32–38
32. Patschinsky, T., Hunter, T., Esté, F. S., Cooper, J. A., and Selton, B. M. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 973–977
33. Songyang, Z., and Cantley, L. C. (1995) Trends Biochem. Sci. 20, 441–448
34. Haltiwanger, R. S., Kelly, W. G., Roquemore, E. P., Blomberg, M. A., Dong, L.-Y. D., Kreppel, L. K., Chou, T.-Y., and Hart, G. W. (1992) Biochem. Soc. Trans. 20, 264–269
35. Holt, G. D., and Hart, G. W. (1986) J. Biol. Chem. 261, 8049–8057
36. Kelly, W. G., and Hart, G. W. (1989) Cell 57, 243–251
37. Posada, J., and Cooper, J. A. (1992) Mol. Biol. Cell 3, 583–592
38. Sun, H., and Tenks, N. K. (1994) Trends Biochem. Sci. 19, 480–485
39. Field, M. C., and Wainwright, L. J. (1995) Glycobiology 5, 463–472
40. Selcer, J., Hofmann, F., Rex, G., Wilm, M., Mann, M., Just, I., and Aktories, K. (1996) J. Biol. Chem. 271, 25173–25177
41. Paulson, J. C., and Colley, K. J. (1989) J. Biol. Chem. 264, 17615–17618
42. Hirana, T., Kinoshita, N., Morikawa, K., and Yanagida, M. (1999) Cell 100, 319–326
43. Lamb, J. R., Michaual, W. A., Sikorski, R. S., and Hieter, P. A. (1994) EMBO J. 13, 4321–4329
44. Tugendreich, S., Tomkiel, J., Earnshaw, W., and Hieter, P. (1995) Cell 81, 261–268
45. Schults, J., Marshall-Carlson, L., and Carlson, M. (1990) Mol. Cell. Biol. 10, 4741–4756
46. Rameau, G., Puglia, K., Crowe, A., Sethy, I., and Willis, I. (1994) Mol. Cell. Biol. 14, 822–830
47. Tzamarias, D., and Struhl, K. (1995) Genes Dev. 9, 821–831
48. Haucke, V., Horst, M., Schatz, G., and Lithgow, T. (1996) EMBO J. 15, 1231–1237
49. Li, Y., and Kornberg, R. D. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 2362–2366
50. Koleske, A. J., and Young, R. A. (1995) Trends Biochem. Sci. 20, 113–116
51. Kaiser, K., and Meisterernst, M. (1996) Trends Biochem. Sci. 21, 342–345
52. Dahmus, M. E. (1996) Biochim. Biophys. Acta 1261, 171–182

Cytosolic and Nuclear O-GlcNAc Transferase 9315