Glycobiology vol. 12 no. 11 pp. 763–770, 2002

O-Glycosylation of EGF repeats: identification and initial characterization of a UDP-glucose: protein *O*-glucosyltransferase

Li Shao[3], Yi Luo[3], Daniel J. Moloney[1], and Robert S. Haltiwanger[2,3]

[3]Department of Biochemistry and Cell Biology, Institute for Cell and Developmental Biology, State University of New York at Stony Brook, Stony Brook, NY 11794-5215, USA

Received on April 17, 2002; revised on June 25, 2002; accepted on July 5, 2002

_O_-Glucose is an unusual form of posttranslational modification consisting of glucose directly attached to protein through _O_-linkage. Several serum proteins (factor VII, factor IX, protein Z, and thrombospondin) contain this unique modification on their epidermal growth factor (EGF)–like repeats. Comparison of the glycosylation sites on these proteins revealed a putative consensus sequence for _O_-glucose modification: C^1XSXPC^2, where C^1 and C^2 are the first and second conserved cysteines of the EGF repeat. We identify and characterize an enzymatic activity capable of adding glucose to EGF repeats: UDP-glucose: protein _O_-glucosyltransferase. Using extracts of Chinese hamster ovary cells as the enzyme source, recombinant factor VII EGF repeat as the acceptor, and UDP-[^{3}H]glucose as the donor, we show that the activity is linearly dependent on time, enzyme amount, and substrate concentration. As with most glycosyltransferases, metal ions (such as manganese) are required for activity. Analysis demonstrated that the glucose is added in _O_-linkage to the EGF repeat. Mutation of the serine to alanine in the predicted glycosylation site abrogates glycosylation, as does reduction and alkylation of the EGF repeat, suggesting that the enzyme recognizes not only the consensus sequence but also the 3D structure of the EGF repeat. Detection of _O_-glucosyltransferase activity in extracts of cell lines from insects to humans and a variety of rat tissues suggests that _O_-glucose modification is widespread in biology. These studies lay the foundation for future work on the biological role of the _O_-glucose modification.

Key words: Chinese hamster ovary cells/EGF repeats/glucosyltransferase/glycosyltransferases/*O*-glucose

Introduction

Although *O*-linked carbohydrate modifications on epidermal growth factor (EGF)–like repeats have only recently been described (Harris and Spellman, 1993), they have generated a great deal of excitement. Two kinds of *O*-linked structures have been reported consisting of a fucose (*O*-fucose) or glucose (*O*-glucose) attached to the EGF repeats through the hydroxyl group of serine or threonine residues. EGF repeats are small motifs with about 40 amino acids defined by 6 conserved cysteines forming three disulfide bonds (Campbell and Bork, 1993). They are well known as players in protein–protein interactions, such as receptor–ligand binding. Recent work in a number of laboratories has demonstrated that such receptor–ligand interactions can be affected by the alterations in the *O*-linked carbohydrate modifications on EGF repeats. For instance, Fringe is known to be an *O*-fucose-specific β1,3-N-acetylglucosaminyltransferase capable of altering Notch signal transduction by addition of GlcNAc residues to *O*-fucose modifications on the EGF repeats of Notch (Bruckner *et al.*, 2000; Moloney *et al.*, 2000a). Recently, the presence of an *O*-fucose on the EGF repeat of Cripto was demonstrated to be essential for Cripto to mediate Nodal-dependent signaling (Schiffer *et al.*, 2001; Yan *et al.*, 2002). Coimmunoprecipitation studies suggest that the *O*-fucose is essential for productive interaction between Cripto and Nodal (Yan *et al.*, 2002). *O*-fucose on urinary-type plasminogen activator (uPA) also appears to be required for activation of the uPA receptor (Rabbani *et al.*, 1992). Defucosylated EGF repeat from uPA is able to bind to the uPA receptor, but it does not induce stimulation of signaling. Each of these examples demonstrates how alterations in an *O*-linked carbohydrate modification on an EGF repeat can regulate a signaling event.

The *O*-glucose modification was originally identified on bovine blood coagulation factors VII and IX (Hase *et al.*, 1988). Detailed analysis demonstrated a trisaccharide form of *O*-glucose on bovine factor IX with the structure Xyl-α1,3-Xyl-α1,3-Glc-β1-*O*-Ser (Hase *et al.*, 1990). Similar structures were observed on human factor VII, factor IX, and protein Z (Nishimura *et al.*, 1989). Comparison of the sequences surrounding sites of glycosylation on these proteins led to the proposal of a putative consensus sequence for the *O*-glucose addition (Harris and Spellman, 1993). The proposed consensus sequence for addition of *O*-glucose is C^1-X-S-X-P-C^2, where C^1 and C^2 are the first and second conserved cysteines of the EGF repeat, respectively, and X can be any amino acid (Harris and Spellman, 1993). The actual number of proteins known to be modified with the *O*-glucose modification is limited, but the putative consensus sequence allows database searches (Table I) revealing that numerous secreted and cell surface proteins are predicted to bear *O*-glucose. We recently demonstrated that the putative consensus sequence is capable of accurately predicting whether a protein bears *O*-glucose by showing that the Notch receptor has this modification (Moloney *et al.*, 2000b) (Table I).

[1]Present address: Department of Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA

[2]To whom correspondence should be addressed; E-mail: Robert.Haltiwanger@stonybrook.edu

Table I. *O*-glucose-modified proteins

Proteins reported to be modified with *O*-glucose	Reference
Factor VII	Hase *et al.*, 1988
Factor IX	Hase *et al.*, 1988
Protein Z	Nishimura *et al.*, 1989
Fetal antigen-1/delta-like protein (FA1-DLK)	Krogh *et al.*, 1997
Notch	Moloney *et al.*, 2000b
Thrombospondin	Nishimura *et al.*, 1992
Proteins containing the consensus sequence C-X-S/T-X-P-C	
Secreted	Membranous
Agrin	Delta
Fibrillin	Serrate
Factor X	Jagged
Factor XII	Crumbs
Aggrecan core protein	LDL receptor
Versican core protein	Lin-12
Hepatocyte growth factor activator	Sperm flagellar membrane protein precursor

Proteins that have been identified to contain *O*-glucose are listed along with a partial list of proteins that contain the consensus sequence for *O*-glucose addition and may be modified. Proteins containing a consensus sequence within an EGF module were found by searching the sequence databases, Swiss-Prot and PIP, at the MOTIF Website (http://www.motif.genome.ad.jp) using the query pattern C-x-[ST]-x-P-C.

To accumulate the tools necessary to examine the biological function of *O*-linked modifications on EGF repeats, we identify the enzymes responsible for their addition. At least three different enzymes are necessary of the synthesis of the *O*-glucose trisaccharide (Figure 1). Two of them have been previously identified: the UDP-D-xylose: β-D-glucoside α-1,3-D-xylosyltransferase (Omichi *et al.*, 1997) and the UDP-D-xylose: α-D-xyloside α1, 3 xylosyltransferase (Minamida *et al.*, 1996), responsible for adding the α1,3-Xyl onto *O*-glucose and the α1,3-Xyl onto Xyl, respectively. In this article we report the initial identification and characterization of the UDP-glucose: protein *O*-glucosyltransferase activity in extracts of Chinese hamster ovary (CHO) cells, the enzyme responsible for addition of the glucose to protein. The fact that this enzyme adds a unique modification to a number of biologically interesting proteins (Table I) suggests that it may play a role in their biology.

Results

Development of an in vitro *assay for the UDP-glucose: protein O-glucosyltransferase*

To develop the assay, we utilized UDP-[^{3}H]glucose as the donor substrate and a bacterially expressed recombinant factor VII EGF repeat known to contain an *O*-glucose consensus site (Harris and Spellman, 1993) as the acceptor substrate. Extracts of CHO cells were used as the enzyme source because we knew that CHO cells synthesize proteins bearing the *O*-glucose modification (Moloney *et al.*, 2000b). After incubation of the CHO cell lysate with factor VII EGF repeat and UDP-[^{3}H]glucose, the products were separated from unincorporated radiolabel using a C18 cartridge. The *O*-glucosyltransferase activity showed linear dependency on the amount of lysate (Figure 2a), factor VII EGF repeat (Figure 2b), and UDP-glucose (Figure 2c).

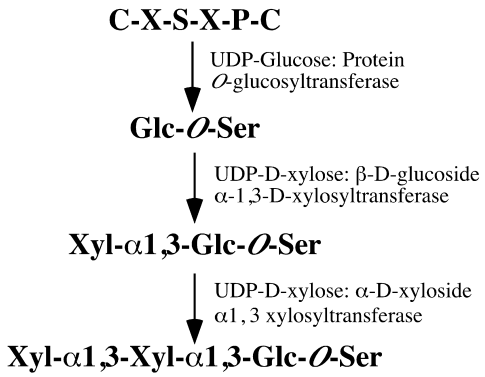

Fig. 1. *O*-glucose glycosylation pathway. Addition of *O*-glucose to the putative consensus site on EGF repeats by UDP-glucose: protein *O*-glucosyltransferase (as described), as well as elongation of *O*-glucose by *O*-glucose α1,3-xylosyltransferase (Omichi *et al.*, 1997) and xylose α1,3-xylosyltransferase are shown (Minamida *et al.*, 1996).

Previous studies on the protein *O*-fucosyltransferase (*O*-FucT-1) demonstrated that the EGF repeat needed to be properly folded for the enzyme to recognize the consensus site (Wang and Spellman, 1998). To examine whether the *O*-glucosyltransferase also requires folded EGF repeats, the factor VII EGF repeat was denatured in urea and reduced and alkylated with dithiothreitol (DTT) and iodoacetamide. Control factor VII EGF repeat was treated with urea but not with DTT or iodoacetamide. Both were repurified by reverse-phase high-performance liquid chromatography (HPLC) (Figure 3a), where the reduced and alkylated form migrated somewhat later, consistent with the carboxyamidomethylation. Both the denatured and native EGF repeats were analyzed as substrates in the *O*-glucosyltransferase assay, and only the native EGF repeat was functional (Figure 3b). The reduced and alkylated EGF repeat did not serve as a substrate for the enzyme, even at high concentrations. These results demonstrate that the enzyme requires a properly folded EGF repeat for recognition.

Human factor VII is modified with *O*-glucose at serine-52 (Hase *et al.*, 1988). To examine whether the predicted serine on the factor VII EGF repeat was being modified, an alanine was substituted for serine-52 (S52A: C^1-A-S-S-P-C^2→C^1-A-A-S-P-C^2). Because the adjacent amino acid is also a serine, it was also mutated to alanine (S53A: C^1-A-S-S-P-C^2→C^1-A-S-A-P-C^2) to examine the possibility of its modification. Both the S52A and S53A mutants were expressed in *Escherichia coli* and purified. Because mutations in EGF repeats can result in misfolding (Wang and Spellman, 1998) and misfolded EGF repeats are not substrates (Figure 3b), we needed an independent measure of whether the S52A and S53A mutants were properly folded. We took advantage of the fact that *O*-FucT-1 also requires properly folded EGF repeats and that the factor VII EGF repeat contains an *O*-fucose glycosylation

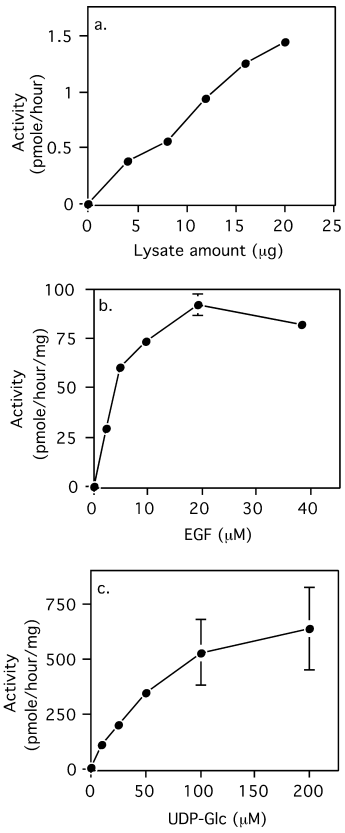

Fig. 2. *O*-glucosyltransferase activity is linearly dependent on the CHO cell lysate amount and the concentration of the substrates. (**a**) CHO cell lysate was varied with constant EGF and UDP-glucose concentration. (**b**) The concentrations of factor VII EGF were varied with constant CHO cell lysate and UDP-glucose concentration. (**c**) UDP-glucose was varied with constant EGF concentration and CHO cell lysate amount.

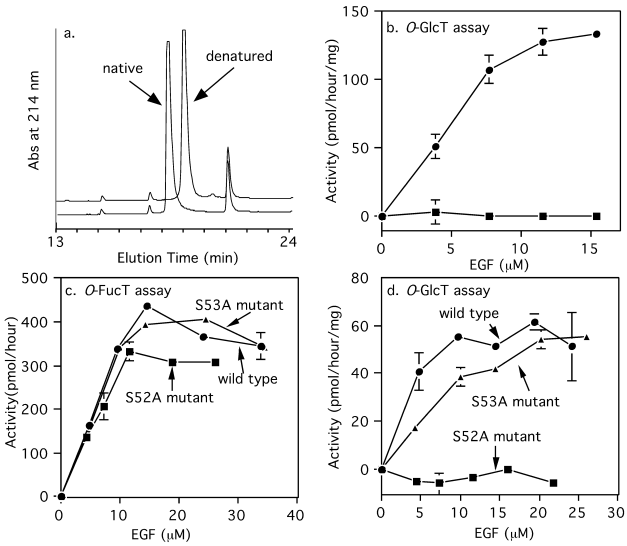

Fig. 3. Factor VII EGF repeat must be properly folded and contain the proper consensus sequence to be modified by *O*-glucosyltransferase. (**a**) After reduction and alkylation, factor VII EGF repeat was purified by reverse-phase HPLC. Elution times of the native and reduced and alkylated (denatured) EGF repeats are shown. (**b**) *O*-glucosyltransferase assays were performed using both control and denatured factor VII EGF repeat as in Figure 2b (circles: native EGF repeat; squares: reduced and alkylated EGF repeat). (**c**) The predicted *O*-glucose consensus site in factor VII EGF repeat was mutated (S52A and S53A) and analyzed for its ability to serve as a substrate for *O*-FucT-1 compared to wild-type factor VII EGF repeat (circles: wild-type EGF repeat; squares: S52A mutant; triangles: S53A mutant). (**d**) Mutated factor VII EGF repeat was analyzed in an *O*-glucosyltransferase assay as in Figure 2b (circles: wild-type EGF repeat; squares: S52A mutant; triangles: S53A mutant).

site (Harris and Spellman, 1993). Both the S52A and S53A mutants were essentially equivalent to wild type as a substrate for *O*-FucT-1 (Figure 3c), indicating that the mutant EGF repeats were properly folded. Analysis of the same mutant in the *O*-glucosyltransferase assay showed no activity for the S52A mutant, whereas the S53A mutant had nearly equivalent activity to wild type (Figure 3d). These results demonstrate that serine-52 is essential for addition of *O*-glucose to factor VII EGF repeat.

Product characterization

A series of chromatographic analyses were carried out to demonstrate that the product generated in the assays consists of glucose in *O*-linkage to the EGF repeat. To demonstrate that the glucose is covalently associated with the EGF repeat, the product was analyzed by reverse-phase HPLC (Figure 4a). Furthermore, to demonstrate that the glucose was attached through an *O*-linkage, alkali-induced β-elimination was performed. The released sugar product from the β-elimination migrated as a monosaccharide on gel filtration chromatography (Figure 4b). Analysis of the monosaccharide by high-performance anion-exchange chromatography (HPAEC) revealed it to be glucitol (Figure 4c), the expected product from β-elimination of *O*-glucose.

Optimization of O*-glucosyltransferase assay*

The *O*-glucosyltransferase assay was optimized for temperature, pH, and time. Product formation increased linearly with time, and, not surprisingly, the enzyme worked best at 37°C (Figure 5a, b). Different pH values of the reaction buffer affected the enzyme activity as well (Figure 5c), with an optimal range between pH 6 and 7.5. As with most glycosyltransferases, the *O*-glucosyltransferase requires metal ions for activity. In the absence of cations or in the presence of ethylenediamine tetra-acetic acid (EDTA), the *O*-glucosyltransferase had very little activity. Among those cations tested, manganese enhanced the activity the greatest; magnesium, calcium, and cobalt also supported some enzyme activity, whereas zinc did not (Figure 5d). Optimal activity was achieved at a concentration of 10 mM manganese (Figure 5e).

Distribution of O*-glucosyltransferase activity*

To address how widespread the *O*-glucose modification is in biology, we assayed extracts of cells from a variety of species and several tissues from rat for *O*-glucosyltransferase activity. Activity was detected in extracts of cells from all metazoans tested, although it was not detected in yeast (Figure 6a). The activities between different cell lines varied dramatically. Lec1 (CHO), Hela (human cervical carcinoma), SF-9 (fall armyworm), and S2 (Schneider) contained moderate levels of activity, whereas HD11 cells (chicken macrophage) and NIH 3T3 (mouse embryo fibroblast) showed relatively high amounts of activity. These data suggest that this enzyme

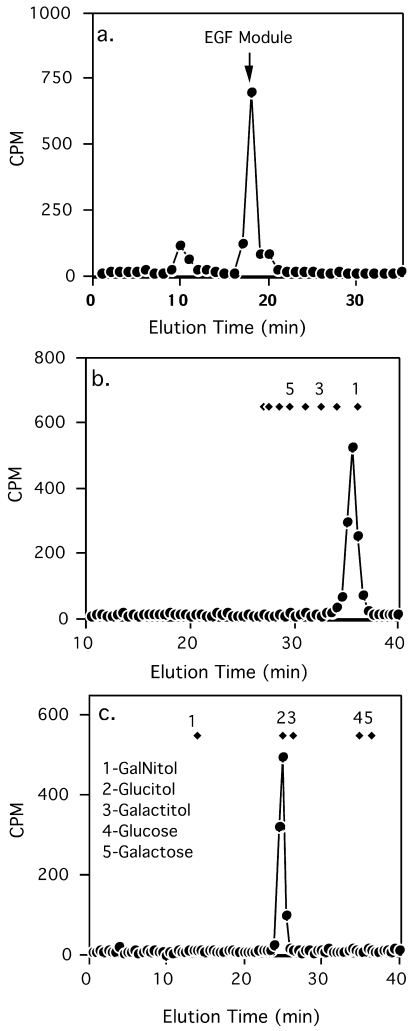

Fig. 4. Product analysis demonstrates that the glucose is attached to the EGF repeat through an *O*-linkage. (**a**) EGF repeat was labeled *in vitro* and purified on reverse-phase HPLC. The arrow indicates the elution position of factor VII EGF repeat. (**b**) The purified EGF repeat with radiolabeled glucose was subjected to β-elimination (see *Materials and methods*), and the products were analyzed by gel filtration chromatography on a calibrated Superdex column. Diamonds represent the elution position of glucose standards. (**c**) The monosaccharide from panel **b** was identified as glucitol by HPAEC analysis. The elution positions of authentic standards are shown.

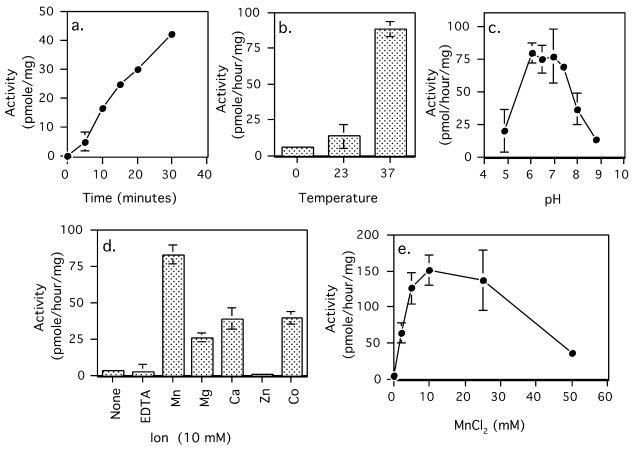

Fig. 5. Characterization of time, temperature, pH, and metal requirements of the *O*-glucosyltransferase activity. Enzyme activity was measured as described in *Materials and methods* as a function of (**a**) time, (**b**) temperature, (**c**) pH, (**d**) metal ions, and (**e**) with increasing concentration of $MnCl_2$.

activity is found in both vertebrate and invertebrate species and that the invertebrate enzymes can recognize the glycosylation site in the EGF repeat from a vertebrate. Analysis of the rat tissue extracts (heart, liver, lung, leg muscle, kidney, spleen, and brain) suggested that this enzyme is also widely expressed in tissues (Figure 6b), although the specific activity varies. The highest activity was found in the extract of spleen and brain, and lowest in heart and leg muscle.

Discussion

In this article, we develop an assay for and initially characterize the UDP-glucose: protein *O*-glucosyltransferase. Like *O*-FucT-1 (Wang *et al.*, 2001; Wang and Spellman, 1998), the *O*-glucosyltransferase requires a properly folded EGF repeat as an acceptor substrate. The activity requires metal ions, and optimal activity is between pH 6 and 7.5. The enzyme activity could be detected in a wide variety of tissues and metazoan species, indicating that this unusual modification is widespread in biology. Identification of this enzyme is an essential step toward a full understanding of the biological significance of the *O*-glucose modification.

Although the biological effect of modifying EGF repeats with *O*-glucose is not yet known, recent work on the *O*-fucose modifications of the Notch receptor (Moloney *et al.*, 2000a), uPA (Rabbani *et al.*, 1992), and Cripto (Schiffer *et al.*, 2001; Yan *et al.*, 2002) demonstrates that carbohydrate modifications of EGF repeats can regulate receptor–ligand interactions. If a simple *O*-fucose monosaccharide can alter the interaction between receptor and ligand, the more complex *O*-glucose modifications on the EGF repeats may have similar effects. Several observations suggest such a role for *O*-glucose modifications. For instance, mutation of serine-52 (the *O*-glucose site) to alanine in human blood coagulation factor VII results in a decrease in factor VII clotting activity by an unknown mechanism (Bjoern *et al.*, 1991). The conservation across species of *O*-glucose consensus sites on many of the proteins either known or predicted to be modified (Table I) strongly suggests that *O*-glucose plays an essential role in the function of these proteins. This is most apparent in the Notch receptor. Although Notch activity is regulated by Fringe through alteration of the *O*-fucose residues, Notch contains more conserved *O*-glucose sites than *O*-fucose sites (Moloney *et al.*, 2000b). Comparison of all Notch sequences currently in the database reveals 5 evolutionarily conserved *O*-fucose sites (EGF modules 3, 20, 24, 26, 31) and 12 evolutionarily conserved *O*-glucose sites (EGF modules 4, 10, 12, 13, 14, 16, 19, 20, 21, 25, 27, 33).

The Notch receptor protein is extremely important for its role in numerous stages of development (Artavanis-Tsakonas *et al.*, 1999). In *Drosophila*, Notch is one of several neurogenic genes that mediate cell fate decisions during development. Targeted disruption of the Notch1 gene in mice causes embryonic lethality (Swiatek *et al.*, 1994). In humans, deregulation of Notch signaling results in serious diseases, such as T cell

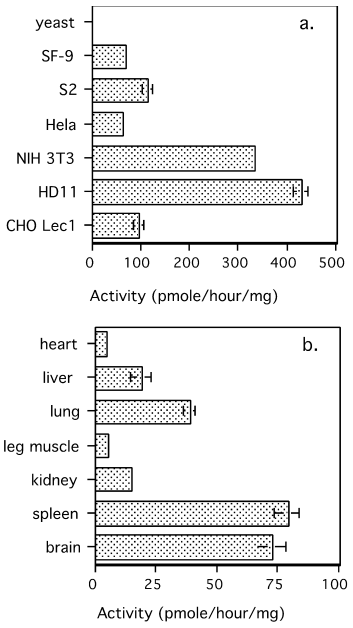

Fig. 6. *O*-glucosyltransferase is expressed in a wide variety of metazoans and mammalian tissues. (**a**) *O*-glucosyltransferase activity was assayed in lysates of a variety of different cultured cells. CHO, Chinese hamster ovary epithelial; HD11, chicken macrophage; NIH 3T3, mouse embryo fibroblast; Hela, human cervical carcinoma; S2: Schneidar cell; SF-9, fall armyworm. (**b**) *O*-glucosyltransferase activity was assayed in lysates of normal adult rat tissues.

leukemias (Ellisen *et al.*, 1991), a cerebral arteriopathy (Joutel *et al.*, 1996), spondylocostal dysostosis (Bulman *et al.*, 2000), and Alagille syndrome (Joutel and Tournier-Lasserve, 1998; Li *et al.*, 1997). Notch receptor becomes activated by binding to its cell surface ligands, Delta or Serrate/Jagged (Mumm and Kopan, 2000). Notch, Delta, and Serrate/Jagged all have numerous EGF repeats with multiple sites for *O*-glucose and *O*-fucose modifications (Moloney *et al.*, 2000b; Panin *et al.*, 2002). EGF repeats 11 and 12 of Notch are necessary and sufficient for interaction with ligands (Rebay *et al.*, 1991). Interestingly, an evolutionarily conserved *O*-glucose site exists on EGF repeat 12, suggesting that the *O*-glucose modification at this site may play a crucial role in Notch–ligand interactions. Experiments to examine this hypothesis are currently under way.

Glucose is not often found in oligosaccharide modifications of proteins in mammalian systems. Best studied are the three α-linked glucoses attached to mannose on newly synthesized N-linked oligosaccharides. This is a transient modification that plays a role in the correct folding of glycoproteins. The glucoses are removed step by step as an indicator of proper folding (Hebert *et al.*, 1995). Other glucose modifications are found on glycogenin and collagen. Glycogenin, one of the two subunits of glycogen synthase complex, is an enzyme that catalyzes its own autoglucosylation, adding a α-linked glucose to itself at tyrosine 194 (Alonso *et al.*, 1995). The glucosylated glycogenin is the critical protein primer required for *de novo* glycogen synthesis. An α-linked glucose is added on to the 2′-hydroxyl of galactose during posttranslational modification of procollagen (Smith *et al.*, 1983). The function of the disaccharide in collagen is not clear but probably is involved in directing the correct molecular assembly of fibrils.

Only two forms of β-linked glucose modifications are known to exist on mammalian glycoproteins, including *O*-glucose and the disaccharide Glcβ1,3Fuc-*O*-Ser/Thr. Several years ago we demonstrated the presence of the Glcβ1,3Fuc disaccharide *O*-linked to a number of proteins in CHO cells (Moloney *et al.*, 1997) and identified a β1,3-glucosyltransferase activity capable of adding the glucose to fucose (Moloney and Haltiwanger, 1999). We had originally believed these disaccharides to exist on EGF repeats, because *O*-fucose was only known on EGF repeats at that time. A recent elegant study by Hofsteenge and co-workers (2001) using mass spectral analysis of thrombospondin-1 revealed the presence of two unusual carbohydrate modifications: C-linked mannose on several tryptophans and the *O*-linked disaccharide Glc-Fuc on several serine/threonine residues. This work demonstrated that Glc-Fuc disaccharides are not on EGF repeats, but instead are found on a different type of protein domain called a thrombospondin-type 1 repeat (TSR) (Adams and Tucker, 2000). Neither the enzyme responsible for addition of *O*-fucose to the TSR nor the biological function of this modification is yet known.

The anomeric linkage of *O*-glucose was established in a single study (Hase *et al.*, 1990) where nuclear magnetic resonance (NMR) was utilized to analyze the structure of the *O*-glucose modification on bovine factor IX. No other data on the anomeric linkage of *O*-glucose exists. In particular, no successful enzymatic cleavage of *O*-glucose from protein has been reported. We have been unsuccessful in several attempts to remove *O*-glucose from factor VII EGF repeat or glycopeptides derived from glucosylated EGF repeat using either α- or β-glucosidases, suggesting that this linkage is uniquely resistant to enzymatic digestion. Although the *O*-glucose on factor VII EGF repeat is likely to be β-linked, definitive demonstration of the linkage awaits production of sufficient product to pursue NMR analysis.

Unusual carbohydrate modifications are often related to a specific biological event (Varki, 1993). For example, mannose 6-phosphate targets lysosomal enzymes to lysosomes (Kornfeld, 1990), sialyl Lewis x helps recruit leukocytes to sites of inflammation (Zak *et al.*, 2000), and polysialic acid modulates neural cell adhesion molecule interactions during neuronal development (Acheson *et al.*, 1991). The recent work demonstrating that *O*-fucose modifications on the EGF repeats of Notch, uPA, and Cripto play a role in signaling events provide other clear examples of a rare form of glycosylation performing a specific biological function. The *O*-glucose modifications of EGF repeats are another unusual form of glycosylation that we predict will have an equally specific role in the biology of the proteins it modifies. We have initiated attempts to purify the *O*-glucosyltransferase by using a combination of conventional and affinity chromatography. Purifying and cloning this enzyme will allow us to develop the tools necessary to study this unique modification.

Materials and methods

Materials

UDP-β-D-Glucose was purchased from Sigma. UDP-[6-^{3}H]glucose (60 Ci/mmol) was purchased from American Radiolabeled Chemicals (St. Louis, MO). Human factor VII EGF

was both kindly provided by Dr. Yang Wang (Genentech, San Francisco, CA) and synthesized as described later. The CHO Lec1 cell line was developed by Pamela Stanley (Stanley and Siminovitch, 1977). Lec1, Hela, SF-9, S2, and NIH3T3 cells were from American Type Culture Collection (Rockville, MD). HD11 cells (Beug *et al.*, 1979) were kindly provided by Dr. Michael Hayman (Department of Microbiology, SUNY Stony Brook). Alditol sugar standards were prepared by reductions of the corresponding sugar with sodium borohydride as described elsewhere (Haltiwanger *et al.*, 1990). C18 cartridges were obtained from Waters Associated (Milford, MA). Reverse-phase C18 HPLC columns (4.6 × 250 mm; Dyanamax) were purchased from Rainin (Woburn, MA). Protease inhibitor cocktails I and II were prepared as described elsewhere (Holt and Hart, 1986). All other reagents were of the highest quality available.

Preparation of cell lysates and rat tissues

All mammalian cells were grown in 100-mm dishes in a humidified 37°C incubator with 5% CO_2. Lec1 cells were grown in αMEM medium (Mediatech, Herndon, VA) supplemented with 10% fetal bovine serum (HyClone, Logan, UT). Hela and NIH3T3 cells were grown in Dulbecco's modified Eagle's medium (Gibco, Life Technologies, Rockville, MD) supplemented with 10% fetal bovine serum. SF-9 cells were grown at room temperature in Grace's insect medium (Gibco, Life Technologies) supplemented with 10% fetal bovine serum. S2 cells were also grown at room temperature in Shields and Sang M3 insect medium (Sigma) with 10% fetal bovine serum. Cells were collected by scraping and washed three times with Tris-buffered saline (10 mM Tris–HCl, pH 7.5, 0.15 M NaCl). Yeast cell pellets were generously provided by Rolf Sternglanz (Department of Biochemistry and Cell Biology, SUNY Stony Brook). Cell pellets were lysed on ice in Tris-buffered saline with 1% (w/v) Nonidet P-40 and protease inhibitor cocktails I and II (yeast cell pellets were lysed by vortexing with glass beads) (Sutton *et al.*, 1991). Cell debris was pelleted by centrifugation (14,000 rpm, 10 min). The supernatants were aliquoted, frozen in liquid nitrogen, and stored at –80°C. Extracts of rat tissues were prepared as described (Moloney and Haltiwanger, 1999).

Production of recombinant factor VII EGF1 repeat and O*-glucose site mutants in* E. coli

DNA sequences encoding human factor VII EGF1 domain were cloned into a pET20b (+) vector (Novagen, Madison, WI) using BamHI (Invitrogen, Carlsbad, CA) and XhoI (Invitrogen) restriction sites with primers containing BamHI and XhoI restriction sites on the 5′- and 3′-ends, respectively (5′-CCGAAGGATCCGGCAAGT GATGGTGACCAG- 3′, and 5′-CTGCCCTCGAGCCCGTCATCCTTGTG-3′). A plasmid encoding EGF1 from human factor VII (generously provided by Yang Wang; Wang *et al.*, 1996) was used as template. Polymerase chain reaction was carried out for 30 cycles with the following conditions: denaturing at 95°C for 0.5 min; annealing at 65°C for 1 min; elongating at 72°C for 1 min. Serine-52 was changed to Ala (S52A) using the Strategene Quik Change Site-Directed Mutagenesis Protocol, with the primers 5′-GGTGACCAGTGT GCCGCGAGTCCATGCC-AG-3′ and 5′-CTGGCATGGACTCGCGGCACACTGGTCA CC-3′. Serine-53 was changed to Ala (S53A) using the primers 5′-GACCAGTGTGCCTCGGCTCCATGCCAGAATGGT-3′ and 5′-ACCATTCTGGCATGGAGCCGAGGCACACTGG-TC-3′. All constructs were sequenced prior to further study.

To produce protein, transformed bacteria BL21 (DE3) were grown to an absorbency of 0.6 at 600 nm in the presence of 100 mg/L ampicillin, 37°C, and induced with 1 mM isopropyl-β-thiogalactopyranoside overnight at 25°C. Cells were harvested by centrifugation, and periplasmic shock was performed according to the Qiagen manual. The shockate was dialyzed extensively against dialysis buffer (50 mM Tris–HCl pH 8.0, 10 mM imidazole) at 4°C, and purified using Qiagen Ni-NTA resin. The resin was washed with the dialysis buffer, and the protein was eluted with 50 mM Tris–HCl, pH 8.0, 100 mM imidazole.

To identify properly folded EGF repeat, a portion of the purified EGF repeat was radiolabeled using baculovirus overexpressed *O*-FucT-1 and GDP-[^{3}H]fucose (as described in Wang *et al.*, 2001; Wang and Spellman, 1998). Reverse-phase HPLC was performed to separate the properly folded EGF repeats from misfolded variants as described later, and the properly folded form was identified by the presence of [^{3}H]fucose. The remainder of the EGF repeat was then purified in the same manner on a preparative scale. The final concentration of factor VII EGF1 was determined by a BCA assay (Pierce). The factor VII EGF1 repeats containing the S52A and S53A mutants were expressed and purified in the same manner.

O*-Glucosyltransferase assay*

The reaction mixture (50 µl final volume) contained 50 mM HEPES, pH 7.0, 10 mM $MnCl_2$, 4.8 µM recombinant human factor VII EGF-1 repeat, 0.2 µM UDP-[^{3}H]glucose, 2.8 µM UDP-glucose, and 10–20 µg cellular protein. The reaction was incubated at 37°C for 20 min and stopped by addition of 450 µl 50 mM EDTA, pH 8. The EGF repeat was separated from unincorporated radioactivity on C18 cartridges using a vacuum manifold. The sample was loaded onto a C18 cartridge (100 mg). The cartridge was washed with 8 ml water, and the EGF repeat was eluted with 1.5 ml 80% methanol. Incorporation of [^{3}H]glucose into the EGF repeat was determined by scintillation counting of the eluate. Reactions without EGF repeat were used as background control.

Reduction and alkylation of EGF repeat

Two aliquots of recombinant factor VII EGF repeat (4.8 nmole each) were dried in a Speed Vac. Both samples were dissolved in 25 µl 0.4 M NH_4HCO_3 containing 8 M urea (final pH between 7.5 and 8.5). Water (10 µl) was added to one sample as control. The other sample was reduced by adding 5 µl 45 mM DTT and incubating at 50°C for 15 min. After cooling to room temperature, the sample was alkylated by adding 5 µl 200 mM iodoacetamide and incubating at room temperature for another 15 min. Both the control EGF repeat and the reduced and alkylated EGF repeat were purified on reverse-phase HPLC as described.

Reverse-phase HPLC purification

Reverse-phase HPLC was performed on a 250 × 4.6 mm Dynamax C18 column using a 30-min linear gradient from 0% to 80% acetonitrile in 0.1% trifluoracetic acid at 1 ml/min. Peptide was detected at 214 nm. Fractions (1 min) were collected as necessary.

Product analysis by gel filtration chromatography and HPAEC

Radiolabeled product in the eluant from the glucosyltransferase assay was dried in a Speed-Vac. Alkali-induced β-elimination, gel filtration on Superdex column, and HPAEC analysis of the products on a Carbopac MA-1 column (Dionex, Sunnyvale, CA) were all done as described (Moloney *et al.*, 1997, 2000b)

Acknowledgments

The authors would like to thank Annie Li, Malgosia Kokoszka, and Raajit Rampal for critical reading and helpful discussions. We also thank Dr. Michael Hayman and Dr. Sternglanz for providing the HD11 and yeast cells, respectively. This work was supported by National Institutes of Health grant GM 61126.

Abbreviations

EGF, epidermal growth factor-like; CHO, Chinese hamster ovary; DTT, dithiothreitol; EDTA, ethylenediamine tetraacetic acid; HPAEC, high pH anion-exchange chromatography; HPLC, high-performance liquid chromatography; NMR, nuclear magnetic resonance; *O*-FucT-1, *O*-fucosyltransferase; TSR, thrombospondin type 1 repeat; uPA, urinary-type plasminogen activator.

References

Acheson, A., Sunshine, J.L., and Rutishauser, U. (1991) NCAM polysialic acid can regulate both cell–cell and cell-substrate interactions. *J. Cell Biol.*, **114**, 143–153.

Adams, J.C. and Tucker, R.P. (2000) The thrombospondin type 1 repeat (TSR) superfamily: diverse proteins with related roles in neuronal development. *Dev. Dyn.*, **218**, 280–299.

Alonso, M.D., Lomako, J., Lomako, W.M., and Whelan, W.J. (1995) Catalytic activities of glycogenin additional to autocatalytic self-glucosylation. *J. Biol. Chem.*, **270**, 15315–15319.

Artavanis-Tsakonas, S., Rand, M.D., and Lake, R.J. (1999) Notch signaling: cell fate control and signal integration in development. *Science*, **284**, 770–776.

Beug, H., von Kirchbach, A., Doderlein, G., Conscience, J.F., and Graf, T. (1979) Chicken hematopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation. *Cell*, **18**, 375–390.

Bjoern, S., Foster, D.C., Thim, L., Wiberg, F.C., Christensen, M., Komiyama, Y., Pedersen, A.H., and Kisiel, W. (1991) Human plasma and recombinant factor VII: characterization of O-glycosylation at serine residues 52 and 60 and effects of site-directed mutagenesis of serine 52 to alanine. *J. Biol. Chem.*, **266**, 11051–11057.

Bruckner, K., Perez, L., Clausen, H., and Cohen, S. (2000) Glycosyltransferase activity of Fringe modulates Notch–Delta interactions. *Nature*, **406**, 411–415.

Bulman, M.P., Kusumi, K., Frayling, T.M., McKeown, C., Garrett, C., Lander, E.S., Krumlauf, R., Hattersley, A.T., Ellard, S., and Turnpenny, P.D. (2000) Mutations in the human delta homologue, DLL3, cause axial skeletal defects in spondylocostal dysostosis. *Nature Genet.*, **24**, 438–441.

Campbell, I.D. and Bork, P. (1993) Epidermal growth factor–like modules. *Curr. Opin. Struct. Biol.*, **3**, 385–392.

Ellisen, L.W., Bird, J., West, D.C., Soreng, A.L., Reynolds, T.C., Smith, S.D., and Sklar, J. (1991) TAN-1, the human homolog of the *Drosophila* Notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell*, **66**, 649–661.

Haltiwanger, R.S., Holt, G.D., and Hart, G.W. (1990) Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide β-N-acetylglucosaminyltransferase. *J. Biol. Chem.*, **265**, 2563–2568.

Harris, R.J. and Spellman, M.W. (1993) O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology*, **3**, 219–224.

Hase, S., Kawabata, S., Nishimura, H., Takeya, H., Sueyoshi, T., Miyata, T., Iwanaga, S., Takao, T., Shimonishi, Y., and Ikenaka, T. (1988) A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. *J. Biochem. (Tokyo)*, **104**, 867–868.

Hase, S., Nishimura, H., Kawabata, S., Iwanaga, S., and Ikenaka, T. (1990) The structure of $(xylose)_2$glucose-*O*-serine 53 found in the first epidermal growth factor-like domain of bovine blood clotting factor IX. *J. Biol. Chem.*, **265**, 1858–1861.

Hebert, D.N., Foellmer, B., and Helenius, A. (1995) Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell*, **81**, 425–433.

Hofsteenge, J., Huwiler, K.G., Macek, B., Hess, D., Lawler, J., Mosher, D.F., and Peter-Katalinic, J. (2001) C-mannosylation and O-fucosylation of the thrombospondin type 1 module. *J. Biol. Chem.*, **276**, 6485–6498.

Holt, G.D. and Hart, G.W. (1986) The subcellular distribution of terminal N-acetylglucosamine moieties. Localization of a novel protein-saccharide likage, O-linked GlcNAc. *J. Biol. Chem.*, **261**, 8049–8057.

Joutel, A. and Tournier-Lasserve, E. (1998) Notch signalling pathway and human diseases. *Sem. Cell Dev. Biol.*, **9**, 619–625.

Joutel, A., Corpechot, C., Ducros, A., Vahedi, K., Chabriat, H., Mouton, P., Alamowitch, S., Domenga, V., Cecillion, M., Marechal, E., and others. (1996) *Notch3* mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature*, **383**, 707–710.

Kornfeld, S. (1990) Lysosomal enzyme targeting. *Biochem. Soc. Trans*, **18**, 367–374.

Krogh, T.N., Bachmann, E., Teisner, B., Skjodt, K., and Hojrup, P. (1997) Glycosylation analysis and protein structure determination of murine fetal antigen 1 (mFA1)—the circulating gene product of the delta-like protein (dlk), preadipocyte factor 1 (Pref-1) and stromal-cell-derived protein 1 (SCP-1) cDNAs. *Eur. J. Biochem.*, **244**, 334–342.

Li, L.H., Krantz, I.D., Deng, Y., Genin, A., Banta, A.B., Collins, C.C., Qi, M., Trask, B.J., Kuo, W.L., Cochran, J., and others. (1997) Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nature Genet.*, **16**, 243–251.

Minamida, S., Aoki, K., Natsuka, S., Omichi, K., Fukase, K., Kusumoto, S., and Hase, S. (1996) Detection of UDP-D-xylose: α-D-xyloside α1, 3xylosyltransferase activity in human hepatoma cell line HepG2. *J. Biochem. (Tokyo)*, **120**, 1002–1006.

Moloney, D.J. and Haltiwanger, R.S. (1999) The O-linked fucose glycosylation pathway: Identification and characteriztion of a UDP-glucose: O-fucose β1, 3-glucosyltransferase. *Glycobiology*, **9**, 679–687.

Moloney, D.J., Lin, A.I., and Haltiwanger, R.S. (1997) The O-linked fucose glycosylation pathway: evidence for protein specific elongation of O-linked fucose in Chinese hamster ovary cells. *J. Biol. Chem.*, **272**, 19046–19050.

Moloney, D.J., Panin, V.M., Johnston, S.H., Chen, J., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K.D., Haltiwanger, R.S., and Vogt, T.F. (2000a) Fringe is a glycosyltransferase that modifies Notch. *Nature*, **406**, 369–375.

Moloney, D.J., Shair, L., Lu, F.M., Xia, J., Locke, R., Matta, K.L., and Haltiwanger, R.S. (2000b) Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *J. Biol. Chem.*, **275**, 9604–9611.

Mumm, J.S. and Kopan, R. (2000) Notch signaling: from the outside in. *Dev. Biol.*, **228**, 151–165.

Nishimura, H., Kawabata, S., Kisiel, W., Hase, S., Ikenaka, T., Takao, T., Shimonishi, Y., and Iwanaga, S. (1989) Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl_2-Glc) *O*-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z. *J. Biol. Chem.*, **264**, 20320–20325.

Nishimura, H., Yamashita, S., Zeng, Z., Walz, D.A., and Iwanaga, S. (1992) Evidence for the existence of O-linked sugar chains consisting of glucose and xylose in bovine thrombospondin. *J. Biochem. (Tokyo)*, **111**, 460–464.

Omichi, K., Aoki, K., Minamida, S., and Hase, S. (1997) Presence of UDP-D-xylose: β-D-glucoside α-1, 3-D-xylosyltransferase involved in the biosynthesis of the Xylα1-3Glcβ-Ser structure of glycoproteins in the human hepatoma cell line HepG2. *Eur. J. Biochem.*, **245**, 143–146.

Panin, V.M., Shao, L., Lei, L., Moloney, D.J., Irvine, K.D., and Haltiwanger, R.S. (2002) Notch ligands are substrates for EGF protein O-fucosyltransferase and Fringe. *J. Biol. Chem.*, in press.

Rabbani, S.A., Mazar, A.P., Bernier, S.M., Haq, M., Bolivar, I., Henkin, J., and Goltzman, D. (1992) Structural requirements for the growth factor activity of the amino-terminal domain of urokinase. *J. Biol. Chem.*, **267**, 14151–14156.

Rebay, I., Fleming, R.J., Fehon, R.G., Cherbas, L., Cherbas, P., and Artavanis-Tsakonas, S. (1991) Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor. *Cell*, **67**, 687–699.

Schiffer, S.G., Foley, S., Kaffashan, A., Hronowski, X., Zichittella, A., Yeo, C.Y., Miatkowski, K., Adkins, H.B., Damon, B., Whitman, M., and others. (2001) Fucosylation of Cripto is required for its ability to facilitate Nodal signaling. *J. Biol. Chem.*, **276**, 37769–37778.

Smith, E.L., Hill, R.L., Lehman, I.R., Lefkowitz, R.J., Handler, P., and White, A. (1983) *Principles of biochemistry: mammalian biochemistry*. McGraw Hill, New York.

Stanley, P. and Siminovitch, L. (1977) Complementation between mutants of CHO cells resistant to a variety of plant lectins. *Som. Cell Genet.*, **3**, 391–405.

Sutton, A., Immanuel, D., and Arndt, K.T. (1991) The SIT4 protein phosphatase functions in late G1 for progression into S phase. *Mol. Cell Biol.*, **11**, 2133–2148.

Swiatek, P.J., Lindsell, C.E., del Amo, F.F., Weinmaster, G., and Gridley, T. (1994) Notch1 is essential for postimplantation development in mice. *Genes Dev.*, **8**, 707–719.

Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, **3**, 97–130.

Wang, Y. and Spellman, M.W. (1998) Purification and characterization of a GDP-fucose: polypeptide fucosyltransferase from Chinese hamster ovary cells. *J. Biol. Chem.*, **273**, 8112–8118.

Wang, Y., Lee, G.F., Kelley, R.F., and Spellman, M.W. (1996) Identification of a GDP-L-fucose: polypeptide fucosyltransferase and enzymatic addition of O-linked fucose to EGF domains. *Glycobiology*, **6**, 837–842.

Wang, Y., Shao, L., Shi, S., Harris, R.J., Spellman, M.W., Stanley, P., and Haltiwanger, R.S. (2001) Modification of epidermal growth factor-like repeats with O-fucose: molecular cloning of a novel GDP-fucose: protein O-fucosyltransferase. *J. Biol. Chem.*, **276**, 40338–40345.

Yan, Y.T., Liu, J.J., Luo, Y.E.C., Haltiwanger, R.S., Abate-shen, C., and Shen, M.M. (2002) Bi-functional activity of Cripto as a ligand and co-receptor in the Nodal signaling pathway. *Mol. Cell Biol.*, **22**, 4439–4449.

Zak, I., Lewandowska, E., and Gnyp, W. (2000) Selectin glycoprotein ligands. *Acta Biochim. Pol.*, **47**, 393–412.