Glycoconjugate Journal 18, 291–299, 2001


# Engineering neoglycoproteins with multiple *O*-glycans using repetitive pentapeptide glycosylation units

Atsuko Yoneda, Masahiro Asada, Saori Yamamoto, Junko Oki, Yuko Oda, Keiko Ota, Yoko Ogi, Sachi Fujishima and Toru Imamura*

*Gene Discovery Research Center, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan*

**Controlled protein remodeling with *O*-linked glycans has been limited by our incomplete understanding of the process of glycosylation. Here we describe a secretable fibroblast growth factor (FGF) with multiple mucin-type *O*-glycans produced by introducing a minimum pentapeptide glycosylation unit in a decarepeat format at its N- or C-terminus. Expressed in Chinese hamster ovary cells, chemical and biochemical analyses of the resultant proteins (Nm10-FGF and Cm10-FGF, respectively) demonstrated that all *O*-glycosylation units were glycosylated and the dominant structure was sialylated Gal[$\beta$1–3]GalNAc. This indicates that minimum *O*-glycosylation unit in multirepeat format serves as a remarkably efficient acceptor in CHO cells. The Nm10-FGF and Cm10-FGF proteins maintained the mitogenic activity to vascular endothelial cells. In addition, intact Cm10-FGF and its desialylated form interacted with several lectins in the same way as mucin-type glycoproteins. The intact Cm10-FGF with multiple sialylated *O*-glycans exhibited a longer lifetime in circulating blood, whereas the Cm10-FGF with desialylated *O*-glycans exhibited a shorter lifetime than the deglycosylated form of Cm10-FGF. Our approach would thus appear to be highly effective for engineering neoglycoproteins, the characteristics of which are determined by their multiple mucin-type *O*-glycans.**

**Keywords: *O*-glycosylation, multiple, fibroblast growth factor, neoglycoprotein**

## Introduction

Within animal systems, the majority of secreted proteins are glycosylated by one or more of *O*-linked, *N*-linked or glycosaminoglycan-type sugar chains, which provide the core proteins with a variety of advantageous properties. *O*-glycosylation at multiple sites is especially important because it functions to increase the stability [1,2] of the resultant glycoproteins, it provides functional multivalency [3], and it enables sorting of proteins for subsequent release [4,5].

Numerous studies have been performed to identify a sequence motif for *O*-glycosylation. Although the studies have revealed non-random distributions of the individual amino acids at sites encompassing up to four residues at either side of the glycosylated amino acids [reviewed in ref. 6], and have enabled prediction of mucin-type *O*-glycosylation sites in given sequences [7], no specific consensus sequence for *O*-glycosylations have been identified. It was also revealed that some GalNAcTs need prior adjacent *O*-glycosylation to become active [8,9]. Therefore, engineering proteins by controlled modifications with *O*-linked glycans has not been successful. Recently, it was reported that fusion of an *O*-glycosylation domain of known proteins enabled production of neo *O*-glycoproteins [10,11]. However, numbers and modification sites of *O*-glycans were not identified. Along with the absence of a specific motif for *O*-glycosylation, dependence of the *O*-glycan structure on the type and condition of the cells expressing the glycoproteins makes it difficult to control *O*-glycosylation [3,12]. Therefore, establishing mammalian expression systems for glycoproteins with a desired number of *O*-glycans at predetermined sites has been very difficult.

Using Chinese hamster ovary (CHO) cells, that are frequently used for production of various glycoproteins, we have recently established an expression system for glycoproteins modified with a single *O*-glycan, although the precise structure of the sugar chain remained unclear [13]. We utilized a hexapeptide Ala-Ala-Thr-Pro-Ala-Pro (AATPAP) for glycosylation domain, because the Thr residue in this sequence had been found to be an efficient acceptor for *O*-glycosylation in *in vitro* transferase assay [6,14].

*To whom correspondence should be addressed: Toru Imamura, Ph. D., Gene Discovery Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan. Tel.: +81-298-61-6504; Fax: +81-298-61-6149; E-mail: imamura-toru@aist.go.jp

To extend the previous study, our aims in the present study were 1) to investigate whether the minimum pentapeptide motif for a single mucin-type *O*-glycosylation [Ala-Thr-Pro-Ala-Pro (ATPAP)] that was identified by *in vitro* transferase assay of peptide *N*-acetylgalactosaminyl transferase 1 [14] functions efficiently in a multirepeat format *in vivo*; 2) to introduce multiple *O*-glycans into a protein using these repeats, thereby improving its biological stability by mimicking the characteristics of mucin-type *O*-glycans that occur in natural glycoproteins; and 3) to investigate the sugar chain structure.

## Experimental procedures

### Mice and Cell lines

Male ICR mice (4 weeks) were obtained from Japan Clea (Tokyo, Japan) and received humane care. Chinese hamster ovary cells (CHO-K1) (American Type Culture Collection, Manassas, VA) were cultured in Ham F-12 medium supplemented with 5% fetal bovine serum. Dihydrofolate reductase-deficient CHO (CHO-dhfr(−)) cells (American Type Culture Collection) were cultured in alpha-modified minimum essential medium supplemented with 10% fetal bovine serum. ldlD-CHO mutant cells were a kind gift from Dr. Monty Krieger (Massachusetts Institute of Technology, Cambridge, MA) and were cultured in F-12 medium supplemented with 10% fetal bovine serum. Human umbilical vein endothelial cells (HUVECs) were isolated and cultured as previously described [15] in collagen type I-coated plastic dishes.

### Construction of cDNA

A cDNA encoding secretable fibroblast growth factor (secFGF), comprised of a signal peptide of mouse fibroblast growth factor (FGF)-6 (1–40: MSRGAGRVQGTL-QALVFLGVLVGMVVPSPAGARAQGTLLD) plus fibroblast growth factor (FGF)-1 (22–155), was constructed in pBluescript (KS+) vector (Stratagene, La Jolla, CA) [16]. The oligonucleotides encoding a repeated ATPAP sequence, which was inserted at the N- or C-terminus of the FGF-1 portion of secFGF, were amplified by polymerase chain reaction as previously described [17]. The resultant products were subcloned into pBluescript II (KS+) vector. After verifying the cDNA sequence, the inserted cDNA was excised and subcloned into the mammalian expression vectors pMEXneo [13] and pCAD [18] for the gene amplification.

### Transfection of DNA and concentration of various forms of secreted FGF-1

CHO-K1 cells and ldlD mutant CHO cells were transfected with pMEXneo containing the cDNA of interest using DOSPER reagent (Roche Diagnostics, Basel, Switzerland) and then selected for resistance to G418 (1 mg/ml). Once the cultures were sub-confluent, the growth medium of CHO-K1 transfectants was replaced by ASF104 serum-free medium (Ajinomoto Co., Tokyo, Japan), after which the medium was changed daily, and the conditioned medium pooled. The various forms of secFGF secreted into the medium were partially purified as previously described [13] and used for later analysis.

ldlD cells are incapable of *O*-glycosylation when grown in culture medium supplemented with 0.5% lipoprotein-deficient serum [19]. The ldlD transfectants were progressively accustomed to F-12 medium supplemented with 0.5% calf serum for at least 72 h, after which the medium was changed daily, and the conditioned medium pooled.

CHO-dhfr(−) cells were transfected with the pCAD containing the cDNA of interest as described above. The transfectants surviving in alpha-minus medium were subjected to gene amplification as previously described [18]. Harvested cells were used for large-scale preparation of proteins.

### Western blotting

Samples were denatured with 2-mercaptoethanol and SDS, and resolved by 10% SDS-polyacrylamide gel electrophoresis (PAGE). The separated proteins were transferred to a nitrocellulose membrane (Schleicher & Schuell, Keene, NH), and after blocking with 5% skim milk solution, the filter was labeled with a monoclonal anti FGF-1 antibody, mAb1 [20] and visualized using horseradish peroxidase-conjugated anti-mouse IgG (Zymed Laboratories Inc., South San Francisco, CA) and ECL (Amersham Pharmacia Biotech Ltd., Arlington Heights, IL).

### Glycosidase digestion

Sialidase digestion was carried out by mixing samples with sialidase (*Arthrobacter ureafaciens*, Nacalai Tesque, Kyoto, Japan) and then incubating them overnight at 37°C. *O*-glycosidase (*Diplococcus pneumoniae*, Roche Diagnostics) and peptide *N*-glycosidase F (PNGase) (Roche Diagnostics) digestion were carried out according to the manufacturer's instructions.

### Nomenclature

Various forms of FGF proteins with *O*-glycosylation cassettes expressed by CHO cells were designated depending on their protein core structures: As *O*-glycosylation cassettes were designated "mucin box", the ATPAP peptide (or AATPAP peptide for single modification) was represented by "m". A name of a variant is C(or N)mx-FGF, where x indicates the number of the "m" peptide, and C or N represents carboxyl- or amino-terminus of FGF polypeptide, respectively, to which the "m" peptide(s) was attached. For example, Cm10-FGF is a form of FGF with decarepeat of a pentapeptide ATPAP attached to its carboxyl terminus and expressed by CHO cells.

### Composition analyses of sugar and amino acid

Cm10-FGF proteins were expressed by CHO-dhfr(−) cells for large scale preparation. As we found that a minor portion of Cm10-FGF was *N*-glycosylated form, and that Cm10-FGF modified solely with *O*-glycans could be separated from Cm10-FGF with *N*-glycans on HiTrap heparin column (Amersham Pharmacia Biotech.) using a linear NaCl gradient, we isolated Cm10-FGF lacking *N*-glycans. Monosaccharide and amino acid compositions of the isolated Cm10-FGF lacking *N*-glycans were determined at the Toray Research Center, Inc. For amino acid analysis, Cm10-FGF was hydrolyzed in 4% thioglycolic acid-6 N HCl for 22 h at 110°C *in vacuo* and applied to amino acid analyzer. Hydrolysis for amino sugar analysis was carried out in 4 N HCl for 6 h at 100°C, for neutral sugar analysis in 2 N trifluoroacetic acid for 6 h at 100°C, and for sialic acid analysis in 0.05 M $H_2SO_4$ for 1 h at 80°C. Monosaccharide compositions were analyzed by fluorometric detection after the formation of borate-monosaccharide complexes using the sugar analysis system (Shimadzu, Tokyo, Japan).

### Mitogenic assay of expressed secFGF-1

Incorporation of [methyl-$^3$H]-thymidine by HUVECs was assayed as previously described [21]. The concentrations of the various forms of FGF-1 were determined by enzyme-linked immunosorbent assay. It was confirmed that this system resolved concentrations of *O*-glycan-modified and unmodified forms of m10-FGF-1 (data not shown).

### Lectin-agarose affinity chromatography

Samples in Tris-buffered saline (TBS) were applied to a 1 mL lectin-agarose column (Seikagaku Corp.) equilibrated with TBS and incubated for 30 min at 4°C. After washing with TBS, bound proteins were eluted with haptenic sugars (0.5 M *N*-acetylglucosamine (GlcNAc) for wheat germ agglutinin (WGA), 0.3 M galactose (Gal) for peanut agglutinin (PNA) and 0.5 M lactose for *Ricinus communis* agglutinin (RCA) 120) in TBS. The FGF content of the fractions were monitored by enzyme-linked immunosorbent assay using a polyclonal anti FGF-1 antibody (R & D Systems Inc., Minneapolis, MN) and horseradish peroxidase-conjugated goat anti rabbit IgG (Zymed Laboratories Inc., South San Francisco, CA).

### Analysis of the concentration of various FGF proteins in mouse blood

Six micrograms of various forms of Cm10-FGF and Cm1-FGF were intravenously administered to mice. Blood samples (8 µl) were then collected from the tail at various times and suspended in 3.8% sodium citrate (15 µl). After centrifugation, the supernatants (15 µl) were boiled in 5 × SDS-PAGE sample buffer (10 µl), and samples were stored in −20°C. The amount of FGF-1 in each sample was semi-quantitatively analyzed by immunoblotting using biotinylated mAb1 and horseradish peroxidase-streptavidin (Amersham Pharmacia Biotech.).

## Results

### Nm10-FGF and Cm10-FGF are secreted as sialylated glycoproteins from their CHO transfectants

As in our earlier studies [13,16,21], we chose secFGF as our model for glycosylation; in this case to determine whether a decarepeat of pentapeptide Ala-Thr-Pro-Ala-Pro ($[ATPAP]_{10}$) can serve as an acceptor for multiple *O*-glycosylation. secFGF is a chimeric protein composed of a signal peptide and a FGF-1 core peptide (Figure 1). Two cDNAs encoding secFGF proteins with $[ATPAP]_{10}$ at the N- or C-terminus of the FGF-1 region were constructed (Figure 1; Nm10-FGF and Cm10-FGF, respectively) and transfected into CHO-K1 cells. Media conditioned by the respective transfectants were then analyzed by immunoblotting with anti FGF-1 monoclonal antibody, mAb1. Nm10-FGF protein migrated as a single band at 35 kDa on SDS-PAGE (Figure 2A, lane 1). The majority of Cm10-FGF protein migrated at 32.5 kDa, though a small portion migrated at 36 kDa (Figure 2A, lane 2). In all cases, the molecular masses of Nm10-FGF and Cm10-FGF were larger than that of secFGF (18.5 kDa, Figure 2A, lane 5) or those of Nm1-FGF and Cm1-FGF (21 kDa, Figure 2A, lanes 3 and 4, respectively), which harbor a single mucin-type sugar chain attached to a hexapeptide (AATPAP) at the N-terminus or the C-terminus of FGF-1 [13].

To confirm that the secreted Nm10-FGF and Cm10-FGF were glycosylated, they were digested with sialidase or PNGase, resolved by SDS-PAGE and immunoblotted. Electrophoretic mobility of Nm10-FGF was retarded by sialidase (Figure 2B, lanes 1 and 2), but not by PNGase (Figure 2C, lanes 1 and 2), indicating that Nm10-FGF was heavily modified by sialic acids but not by *N*-glycans. The retarded mobility in SDS-PAGE after sialidase digestion is due to removal of negative charges provided by the sialic acids. Electrophoretic mobilities of the Cm10-FGF bands were also retarded by sialidase digestion (Figure 2B, lanes 3 and 4); the minor band at 36 kDa was deemed to be an *N*-glycosylated form of the 32.5-kDa Cm10-FGF as it disappeared after PNGase digestion (Figure 2C, lanes 3 and 4).

### Dominant core structure of *O*-glycans on Nm10-FGF and Cm10-FGF is Gal[$\beta$1-3]GalNAc-(Thr)

To determine the size of the m10-FGF core proteins, their cDNAs were separately transfected into ldlD cells, mutant CHO cells that lack the capacity to synthesize *N*-acetylgalactosamine (GalNAc)-containing sugar chains; the core protein subsequently recovered from the transfectant-conditioned medium migrated at 31 kDa (Figure 2D and 2E, lane 5).

To identify the sugar chain structure, we utilized sialidase and *O*-glycosidase (*O*-glycopeptide endo-D-galactosyl-*N*-acetyl-$\alpha$-galactosamino hydrolase), which specifically removes

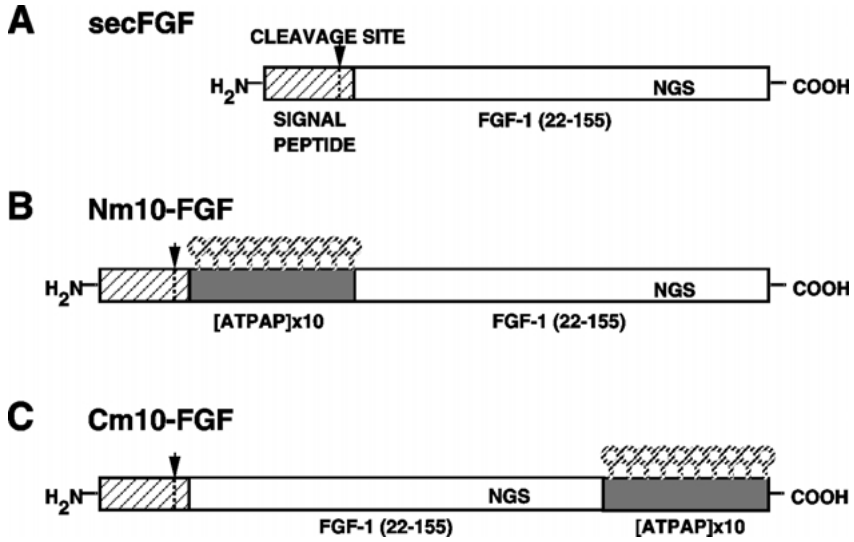


**Figure 1.** Design of neoglycoproteins with multiple *O*-glycans. Schematic diagram of the structures of secFGF, Nm10-FGF and Cm10-FGF. FGF-1 (22–155) is a naturally occurring, N-terminal truncated form of FGF-1 with full mitogenic activity; to this core, a secretion signal peptide was ligated, forming secFGF. A decarepeat of ATPAP was ligated at the N- or C-terminal end of the FGF-1 domain (Nm10- and Cm10-FGF, respectively). Complementary DNAs for these polypeptides, constructed in a eukaryotic expression vector, were transfected into CHO cells as described in the text. The proteins were secreted into the culture medium after modification with *O*-glycans (indicated by ) and cleavage of the signal peptide. NGS indicates the position of an endogenous potential *N*-glycosylation sequence in FGF-1.

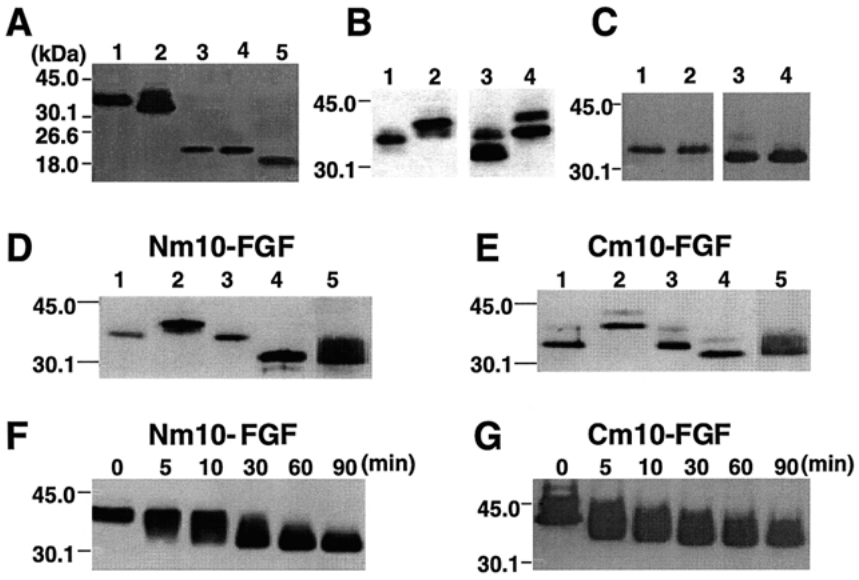


**Figure 2.** Nm10-FGF and Cm10-FGF proteins secreted from respective CHO-K1 transfectants are modified with multiple *O*-glycans. (A) Nm10-FGF and Cm10-FGF proteins in the conditioned supernatants. Supernatants conditioned by CHO-K1 transfectants were analyzed for Nm10-FGF (lane 1), Cm10-FGF (lane 2), Nm1-FGF (lane 3), Cm1-FGF (lane 4) and secFGF (lane 5) by immunoblotting with anti-FGF-1 monoclonal antibody (mAb1). (B) Sialidase treatment. Nm10-FGF (lanes 1 and 2) and Cm10-FGF (lanes 3 and 4) secreted from CHO-K1 transfectants were either not treated (lanes 1 and 3) or were treated with sialidase (lanes 2 and 4) for 18 h at 37°C and analyzed by immunoblotting. (C) PNGase treatment. Secreted Nm10-FGF (lanes 1 and 2) and Cm10-FGF (lanes 3 and 4) were either not treated (lanes 1 and 3) or treated with PNGase (lanes 2 and 4) for 18 h at 37°C. (D and E) *O*-glycosidase treatment. Nm10-FGF (D) and Cm10-FGF (E) were either not treated (lanes 1), treated with sialidase (lane 2), *O*-glycosidase (lane 3), or treated with a mixture of sialidase and *O*-glycosidase (lane 4) for 18 h at 37°C. Proteins from ldlD transfectants were similarly analyzed (lane 5). (F and G) Partial digestion with *O*-glycosidase. Desialylated Nm10-FGF (F) and Cm10-FGF (G) proteins were treated with *O*-glycosidase for the indicated times at 37°C, and aliquots of the digests were analyzed. The bars on the left indicate the positions of molecular weight markers.

the unsialylated Gal[$\beta$1–3]GalNAc structure from Thr/Ser [22]. When intact Nm10-FGF glycoprotein was digested with sialidase plus *O*-glycosidase or with *O*-glycosidase alone, the respective digests migrated at 31 kDa (Figure 2D, lane 4) and 35 kDa (Figure 2D, lane 3), the former being the same as the protein core (Figure 2D, lane 5), and the latter the same as the intact glycoprotein (Figure 2D, lane 1). The Nm10-FGF modification is thus comprised of sialylated Gal[$\beta$1–3]GalNAc structures. Sialidase/*O*-glycosidase digestion of Cm10-FGF yielded primarily the 31-kDa product (Figure 2E, lane 4), while a minor portion was detected at 34 kDa, indicating that Cm10-FGF also contains sialylated Gal[$\beta$1–3]GalNAc structures. These results indicated that the dominant core structure of *O*-glycans on Nm10-FGF and Cm10-FGF is Gal[$\beta$1–3]-GalNAc-(Thr).

### All ATPAP units on Cm10-FGF are modified with *O*-glycan

To investigate whether Nm10-FGF and Cm10-FGF were modified with multiple *O*-glycans, their desialylated forms were digested with *O*-glycosidase for various periods. As digestion progressed, their respective bands shifted stepwise to positions reflecting smaller molecular sizes (Figure 2F and 2G), suggesting that both Nm10-FGF and Cm10-FGF were modified with multiple *O*-glycans.

To analyze the number of *O*-glycans attached, chemical analyses were carried out. As a minor portion of Cm10-FGF was *N*-glycosylated (Figure 2C), Cm10-FGF modified solely with *O*-glycans was purified on HiTrap heparin column using a linear NaCl gradient. The amino acid composition was found to perfectly match that of Cm10-FGF after signal cleavage, as deduced from its cDNA (data not shown). Analysis of the sugars revealed the presence of GalNAc, Gal and *N*-acetylneuraminic acid (Table 1); the content of GalNAc per 1 mole of Cm10-FGF was 10 mole. In contrast, GlcNAc, mannose, fucose, xylose, glucose and *N*-glycolylneuraminic acid were not detected, indicating that Cm10-FGF does not contain *O*-glycans with core 2, 3, 4 or 6 structures [3] or *O*-linked-GlcNAc, mannose, fucose, xylose, and glucose [3]. $\alpha$-*N*-Acetylgalactosaminidase (*Acremonium* sp.) digestion of intact Cm10-FGF and desialylated form did not affect their molecular sizes on SDS-PAGE (data not shown), suggesting that Cm10-FGF did not have sialyl GalNAc-Thr and GalNAc-Thr. All the results indicated that all 10 sites were glycosylated and that the dominant sugar structure was sialylated Gal[$\beta$1–3]GalNAc. Further, the relatively high NeuAc content suggests that some of the Gal[$\beta$1–3]GalNAc chain was modified with two sialic acids.

**Table 1.** Sugar composition of Cm10-FGF *O*-glycans. Monosaccharides composition of the Cm10-FGF *O*-glycans was determined

| | *mole/mole of protein* |
|---|---|
| GalN[a,b] | 10 |
| Gal | 7.9 |
| NeuAc | 14 |
| GlcN[b], Man, Fuc, Xyl, Glc, and *N*-glycolylneuraminic acid | not detected |

[a]Abbreviations used in this table: GalN, galactosamine; Gal, galactose; NeuAc, *N*-acetylneuraminic acid; GlcN, glucosamine; Man, mannose; Fuc, fucose; Xyl, xylose; Glc, glucose.
[b]As the acetyl residues were removed by the acid hydrolysis, the monosugar contents are expressed as aminosugars.

### Cm10-FGF and Nm10-FGF are endothelial cell mitogens

In the presence of heparin, secFGF is a potent mitogen toward HUVECs [13]. When we tested whether introduction of multiple *O*-glycans affected the biological activity of secFGF by assessing the capacity of Nm10-FGF and Cm10-FGF to induce DNA synthesis in HUVECs, we found that both were mitogenic, but that their potencies were increased by de-*O*-glycosylation (Figure 3A and 3B). The mitogenicity of secFGF, by contrast, was unaffected by glycosidase treatment (Figure 3C), indicating that introduction of multiple *O*-glycans reduced the mitogenic activity of FGF-1 to some extent.

### Cm10-FGF and Nm10-FGF interact with lectins in the same way as mucin-type glycoproteins

To determine whether multiple *O*-glycosylation of Cm10-FGF affects its interaction with sugar recognition molecules, the glycoprotein was treated with PNGase and then examined for its binding to immobilized WGA which recognizes multiple sialic acids [23]. As shown in Table 2, 100% of intact Cm10-FGF was bound to WGA, whereas none of its de-*O*-glycosylated form and secFGF (wild-type FGF-1 expressed by CHO cells), and only 6% of the singly *O*-glycosylated Cm1-FGF were. Intact Cm10-FGF and secFGF were not bound to PNA, which recognizes Gal[$\beta$1–3]-GalNAc, but not sialylated Gal[$\beta$1–3]GalNAc [24], whereas its desialylated form was almost completely bound (Table 2). Desialylated Cm10-FGF, but not its intact form, was also bound to RCA120, which recognizes terminal Gal. Nm10-FGF glycoprotein yielded essentially the same results (data not shown). Thus, m10-FGF glycoproteins interact with lectins just as mucin-type glycoproteins do.

### The lifetimes of intact and desialylated Cm10-FGF glycoproteins in circulating blood differ from that of non-glycosylated FGF

To test whether the addition of the *O*-glycans to Cm10-FGF core protein might alter its biological availability *in vivo*, intact Cm10-FGF or its derivatives were intravenously injected into mice, and the time-dependent changes in their respective blood levels were analyzed. Aliquots of the plasma obtained at selected times were resolved on SDS-PAGE and analyzed for

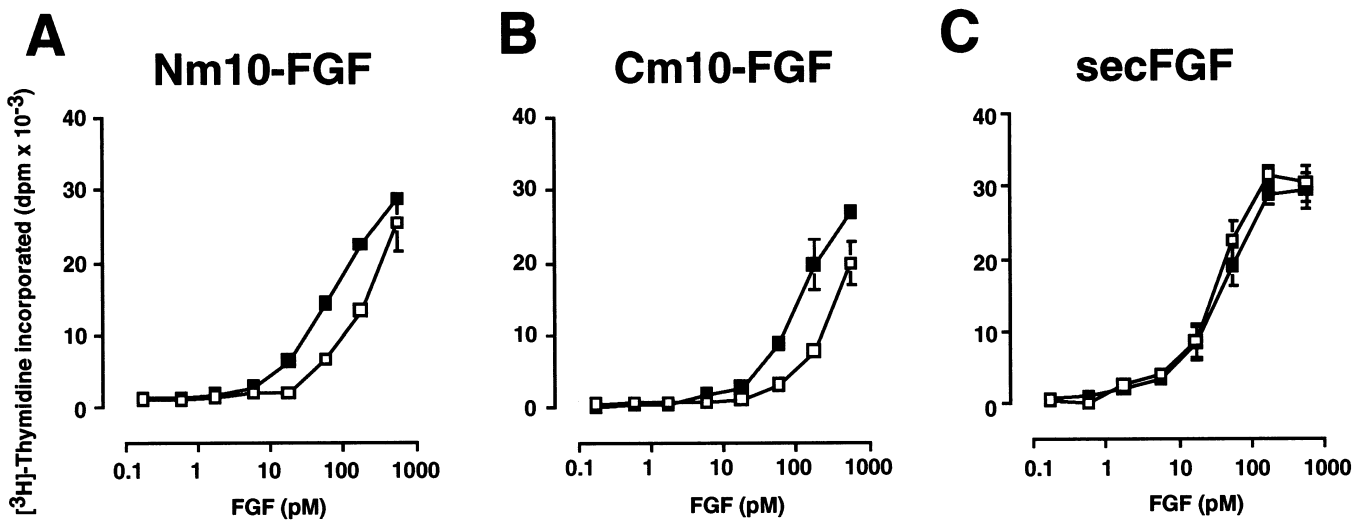


**Figure 3.** Nm10-FGF and Cm10-FGF glycoproteins are mitogenic toward endothelial cells. The effects of intact (open squares) and sialidase/*O*-glycosidase-digested (closed squares) Nm10-FGF (A), Cm10-FGF (B) and secFGF (C) on [$^3$H]-thymidine incorporation into HUVECs in the presence of heparin (5 μg/ml). Values represent means of triplicate samples ± standard deviation. Multiple experiments using different preparations of the sample glycoproteins yielded essentially the same results.

**Table 2.** Cm10-FGF interacts with lectins just as mucin-type glycoproteins do. Samples in Tris buffered saline were applied to a 1 mL lectin-agarose column equilibrated with TBS and incubated for 30 min at 4°C. After washing with TBS, bound proteins were eluted with haptenic sugars (0.5 M *N*-acetylglucosamine for WGA, 0.3 M galactose for PNA and 0.5 M lactose for RCA) in TBS. The FGF content of the fractions were monitored by enzyme-linked immunosorbent assay using a polyclonal anti FGF-1 antibody

| | *WGA*[a] *(%)* | | *PNA (%)* | | *RCA120 (%)* | |
|---|---|---|---|---|---|---|
| | *Unbound* | *Bound* | *Unbound* | *Bound* | *Unbound* | *Bound* |
| Cm10-FGF | | | | | | |
| Intact | 0 | 100 | 91 | 9 | 100 | 0 |
| Desialylated | 96 | 4 | 12 | 88 | 0 | 100 |
| De-*O*-glycosylated | 100 | 0 | n.d. | | n.d. | |
| Cm1-FGF | | | | | | |
| Intact | 94 | 6 | 99 | 1 | n.d. | |
| secFGF[b] | 100 | 0 | 100 | 0 | n.d. | |

[a]Abbreviations used in this table: WGA, wheat germ agglutinin; PNA, peanut agglutinin; RCA, *Ricinus communis* agglutinin; n.d., not determined.
[b]The secFGF is a non-glycosylated, wild-type FGF-1 expressed and secreted by CHO cells.

FGF-1 antigenicity. For comparison, FGF-1 simple protein expressed by *E. coli* was analyzed in lieu of secFGF due to its low yield. It was found that intact Cm10-FGF gradually disappeared from the blood over 4–6 h (Figure 4A, panel a); desialylated Cm10-FGF disappeared within 30 min (Figure 4A, panel b); and the de-*O*-glycosylated form disappeared within 2 h (Figure 4A, panel c). Two hours after injection of the intact glycoprotein, 2 nM in plasma remained (Figure 4B). Cm10-FGF glycoprotein also had a longer lifetime in blood than singly glycosylated Cm1-FGF, which in turn had a longer lifetime than the core protein of Cm10-FGF (Figure 4B). FGF-1 expressed by *E. coli* had shorter lifetime than intact Cm10-FGF and Cm1-FGF. These results show that sialylated *O*-glycans, and not elongation of the core polypeptide, stabilize Cm10-FGF in circulating blood. On the other hand, desialylated Cm10-FGF had shorter lifetime in blood than the core polypeptide, presumably due to clearance by hepatic lectins for asialoglycoproteins.

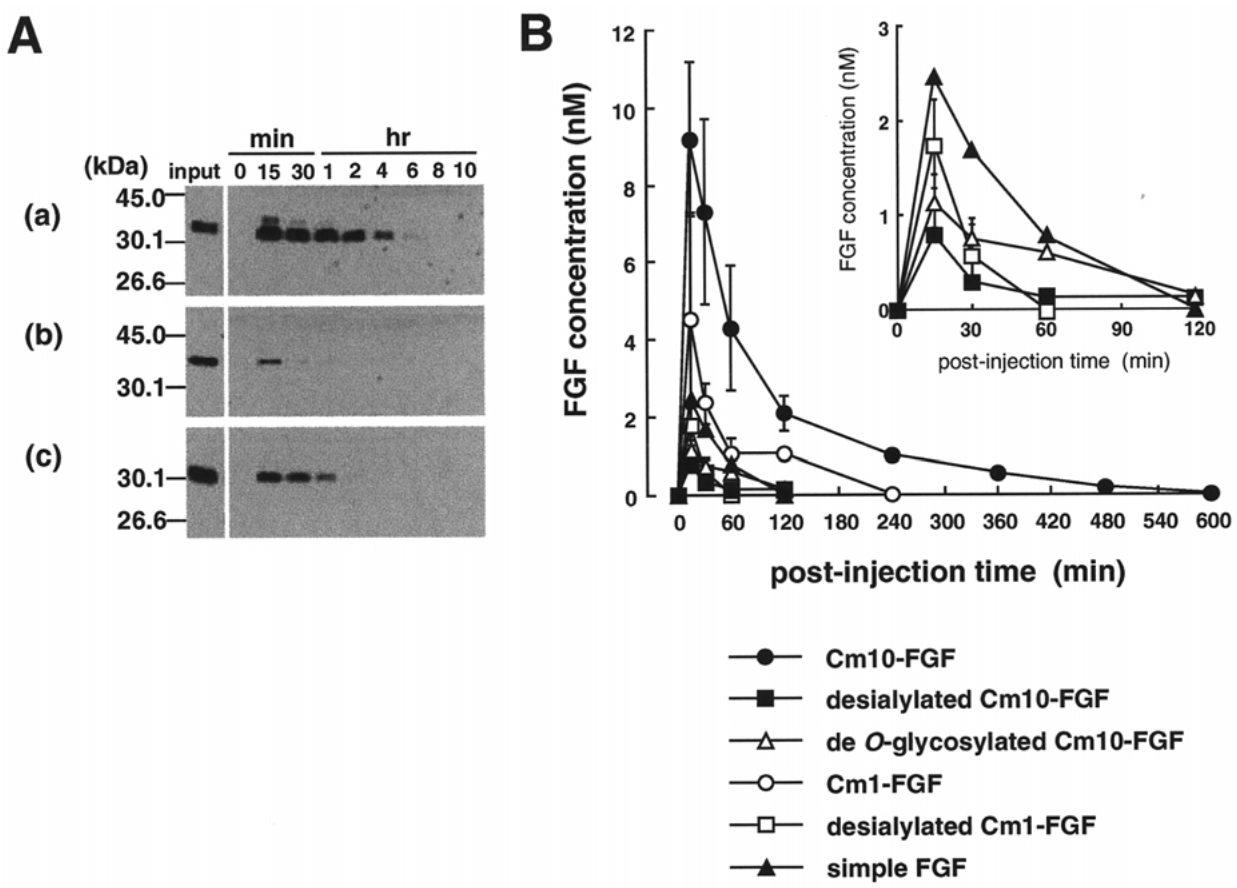


**Figure 4.** The Cm10-FGF glycoprotein has a long lifetime in circulating blood. (A) Clearance of Cm10-FGF glycoprotein from circulating blood. Mice were intravenously injected with intact Cm10-FGF (a) or its desialylated (b) or de-*O*-glycosylated form (c), and samples of circulating blood drawn at the indicated times were analyzed by immunoblotting with biotinylated mAb1. Results of representative immunoblots are shown. The bars on the left indicate the positions of molecular weight markers. (B) Time-dependent decline of Cm10-FGF and Cm1-FGF in circulating blood. Experiments were carried out as in panel A using the non-digested Cm10-FGF and Cm1-FGF, enzyme-digested forms of Cm10-FGF and Cm1-FGF, and non-glycosylated wild-type FGF-1 (simple FGF-1). For each sample, experiments were performed at least three times. The concentrations of the samples in blood (as FGF molarity) were determined by densitometric analyses of the quantitative Western blots and presented as a graph. Inset indicates the magnification of data of desialylated Cm10-FGF, de-*O*-glycosylated Cm10-FGF, desialylated Cm1-FGF and simple FGF. Values represent means of triplicate samples ± standard error.

## Discussion

We have successfully produced a secretable FGF modified by multiple mucin-type *O*-glycans by adding a decarepeat of the pentapeptide ATPAP to its N- or C-terminus and expressing it in CHO cells. Chemical and biochemical analyses of the resultant glycoproteins (Nm10-FGF and Cm10-FGF) revealed that all of the potential *O*-glycosylation sites were glycosylated; that the dominant sugar structure was sialylated Gal[$\beta$1–3]GalNAc; that the glycoproteins retained their capacity to stimulate proliferation of vascular endothelial cells; that intact Cm10-FGF and its desialylated form interact with lectins, just as mucin-type glycoproteins do; and that intact Cm10-FGF and its desialylated form respectively exhibited longer and shorter lifetimes in circulating blood than the deglycosylated form. Thus, introduction of ATPAP in a repetitive format was highly effective for engineering neoglycoproteins, the characteristics of which were determined by their mucin-type *O*-glycans.

It is noteworthy that in the present study, all of the potential *O*-glycosylation sites were glycosylated while a peptide ATPAP showed an apparent decrease from AATPAP in the GalNAc transfer in the *in vitro* study [14]. To our knowledge this is the first observation that minimum *O*-glycosylation unit in multirepeat format serves as a remarkably efficient acceptor in CHO cells. We expect to be able to vary the number of *O*-glycans, depending on the number of ATPAP repeats in the primary structure of the core protein; we have already produced Cm3-FGF, which contains ATPAP in a tri-repeat format and appears to be modified with three *O*-glycans (Asada and Yoneda, unpublished results).

It is of interest that introduction of multiple sialylated *O*-glycans provided the intact neoglycoprotein with lectin interactivity and a longer lifetime in circulating blood, whereas the desialylated form had a shorter lifetime. On the other hand, we observed no significant difference among the intact

m10-FGF glycoproteins and the FGF simple protein in their susceptibility to several proteases (data not shown). This suggests that the longer lifetime of intact Cm10-FGF is not due to its resistance to proteases in circulating blood, but rather due to its altered interactivity with the host system. The shorter lifetime of desialylated Cm10-FGF is likely to be the result of its clearance by asialoglycoprotein receptors in liver [25].

By utilizing the present method and varying the combination of glycosidase treatments, it should be possible to produce neoglycoproteins modified with multiple sialic acids, with Gal or with GalNAc as the terminal sugars. Introduction of such multiple glycans into proteins could enable targeting of neoglycoproteins to specific tissues by taking advantage of the selectivities of the endogenous lectins.

A portion of Cm10-FGF was *N*-glycosylated, probably at a potential *N*-glycosylation site in FGF-1 core protein (Figure 1), whereas Nm10-FGF was not (Figure 2C). Similar observation was made in our previous study in which we made various secFGF mutants, each with an *N*-glycosylation cassette inserted at random site [16]. When these mutants were expressed by mammalian cells, the extents of their *N*-glycosylation were different. Taken together, these results indicate that the location of potential *N*-glycosylation sequence in the primary structure of a protein affects the efficiency of *N*-glycosylation.

It has been reported that introduction of *O*-glycans into proteins or peptide enhance the stability and antigenicity of the core protein or peptide. For instance, the catalytic domain of human milk cholesterol esterase acquired acid stability and resistance to proteolysis by fusion with a portion of a mucin tandem repeat sequence [11]; the use of synthetic multiple antigenic *O*-linked glycopeptides with the Tn structure (GalNAcα-Ser/Thr) in a therapeutic immunization protocol increased survival of tumor-bearing mice [26]; and short synthetic glycopeptides were able to induce anticancer antibody responses [27]. Our prolongation of the lifetime of neoglycoproteins in circulating blood is yet another example of the promising uses for *O*-glycosylation engineering. The *O*-glycosylation engineering opens new horizons in stabilizing and/or *in vivo* targeting of pharmaproteins. This method could be also applied to CHO cells engineered to express sialyl Lewis X on *O*-glycans as described previously [28,29].

## Acknowledgments

We thank Dr. Monty Krieger (Massachusetts Institute of Technology, Cambridge, MA) for ldlD cells and Dr. Treasa Creavin for critical reading of the manuscript. The text was edited by Dr. William Goldman at MST Editing Company. A.Y. was supported by postdoctoral fellowships of NEDO and NIBH. This work was supported by MITI-AIST Complexed Carbohydrate project.

## References

1 Kingsley DM, Kozarsky KF, Hobbie L, Krieger M, Reversible defects in O-linked glycosylation and LDL receptor expression in a UDP-Gal/UDP-GalNAc 4-epimerase deficient mutant, *Cell* **44**, 749–59 (1986).

2 Remaley AT, Ugorski M, Wu N, Litzky L, Burger SR, Moore JS, Fukuda M, Spitalnik SL, Expression of human glycophorin A in wild type and glycosylation-deficient Chinese hamster ovary cells. Role of *N*- and *O*-linked glycosylation in cell surface expression, *J Biol Chem* **266**, 24176–83 (1991).

3 Van den Steen PE, Rudd PM, Wormald MR, Dwek RA, Opdenakker G, *O*-linked glycosylation in focus, *Trends Glycosci Glycotech* **12**, 35–49 (2000).

4 Alfalah M, Jacob R, Preuss U, Zimmer K-P, Naim H, Naim HY, *O*-linked glycans mediate apical sorting of human intestinal sucrase-isomaltase through association with lipid rafts, *Curr Biol* **9**, 593–6 (1999).

5 Magrané J, Casaroli-Marano RP, Reina M, Gåfvels M, Vilaró S, The role of *O*-linked sugars in determining the very low density lipoprotein receptor stability or release from the cell, *FEBS Lett* **451**, 56–62 (1999).

6 Elhammer ÅP, Kezdy FJ, Kurosaka A, The acceptor specificity of UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases, *Glycoconjugate J* **16**, 171–80 (1999).

7 Hansen JE, Lund O, Tolstrup N, Gooley AA, Williams KL, Brunak S, NetOglyc: Prediction of mucin type *O*-glycosylation sites based on sequence context and surface accessibility, *Glycoconjugate J* **15**, 115–30 (1998).

8 Ten Hagen KG, Tetaert D, Hagen FK, Richet C, Beres TM, Gagnon J, Balys MM, VanWuyckhuyse B, Bedi GS, Degand P, Tabak LA, Characterization of a UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase that displays glycopeptide *N*-acetylgalactosaminyltransferase activity, *J Biol Chem* **274**, 27867–74 (1999).

9 Ten Hagen KG, Bedi GS, Tetaert D, Kingsley PD, Hagen FK, Balys MM, Beres TM, Degand P, Tabak LA, Cloning and characterization of a ninth member of the UDP-GalNAc : polypeptide N-acetylgalactosaminyltransferase family, ppGaN-Tase-T9, *J Biol Chem* **276**, 17395–404 (2001).

10 Furuhashi M, Shikone T, Fares FA, Sugahara T, Hsueh AJ, Boime I, Fusing the carboxy-terminal peptide of the chorionic gonadotropin (CG) β-subunit to the common α-subunit: retention of *O*-linked glycosylation and enhanced *in vivo* bioactivity of chimeric human CG, *Mol Endocrinol* **9**, 54–63 (1995).

11 Loomes KM, Senior HE, West PM, Roberton AM, Functional protective role for mucin glycosylated repetitive domains, *Eur J Biochem* **266**, 105–11 (1999).

12 Brockhausen I, Pathways of *O*-glycan biosynthesis in cancer cells, *Biochim Biophys Acta* **1473**, 67–95 (1999).

13 Asada M, Orikawa N, Yoneda A, Oda Y, Ota K, Imamura T, The AATPAP sequence is a very efficient signal for *O*-glycosylation in CHO cells, *Glycoconjugate J* **16**, 321–6 (1999).

14 Yoshida A, Suzuki M, Ikenaga H, Takeuchi M, Discovery of the shortest sequence motif for high level mucin-type *O*-glycosylation, *J Biol Chem* **272**, 16884–8 (1997).

15 Imamura T, Mitsui Y, Heparan sulfate and heparin as a potentiator or a suppressor of growth of normal and transformed vascular endothelial cells, *Exp Cell Res* **172**, 92–100 (1987).

16 Yoneda A, Asada M, Suzuki M, Imamura T, Introduction of an N-glycosylation cassette into proteins at random sites: expression of neoglycosylated FGF, *BioTechniques* **27**, 576–82 (1999).
17 Asada M, Yoneda A, Oda Y, Imamura T, Construction of a cDNA encoding a repetitive amino acid sequence, *BioTechniques* **29**, 978, 981 (2000).
18 Asada M, Yoneda A, Oda Y, Ota K, Ozawa K, Fukuta K, Omae F, Asanagi M, Orikawa N, Suzuki M, Oka S, Makino T, Imamura T, Characterization of fibroblast growth factor-6 expressed by chinese hamster ovary cells as a glycosylated mitogen for human vascular endothelial cells, *Growth Factors* **16**, 293–303 (1999).
19 Krieger M, Reddy P, Kozarsky K, Kingsley D, Hobbie L, Penman M, Analysis of the synthesis, intracellular sorting, and function of glycoproteins using a mammalian cell mutant with reversible glycosylation defects, *Methods Cell Biol* **32**, 57–84 (1989).
20 Imamura T, Oka S, Tanahashi T, Okita Y, Cell cycle-dependent nuclear localization of exogenously added fibroblast growth factor-1 in BALB/c 3T3 and human vascular endothelial cells, *Exp Cell Res* **215**, 363–72 (1994).
21 Yoneda A, Asada M, Oda Y, Suzuki M, Imamura T, Engineering of an FGF-proteoglycan fusion protein with heparin-independent, mitogenic activity, *Nat Biotechnol* **18**, 641–4 (2000).
22 Brooks M, Savage A, The substrate specificity of the enzyme endo-alpha-*N*-acetyl-D-galactosaminidase from Diplococcus pneumonia, *Glycoconjugate J* **14**, 183–90 (1997).
23 Bhavanandan VP, Katlic AW, The interaction of wheat germ agglutinin with sialoglycoproteins. The role of sialic acid, *J Biol Chem* **254**, 4000–8 (1979).
24 Sueyoshi S, Tsuji T, Osawa T, Carbohydrate-binding specificities of five lectins that bind to *O*-Glycosyl-linked carbohydrate chains. Quantitative analysis by frontal-affinity chromatography, *Carbohydr Res* **178**, 213–24 (1988).
25 Ashwell G, Harford J, Carbohydrate-specific receptors of the liver, *Annu. Rev. Biochem.* **51**, 531–54 (1982).
26 Lo-Man R, Bay S, Vichier-Guerre S, Deriaud E, Cantacuzene D, Leclerc C, A fully synthetic immunogen carrying a carcinoma-associated carbohydrate for active specific immunotherapy, *Cancer Res* **59**, 1520–4 (1999).
27 Vichier-Guerre S, Lo-Man R, Bay S, Deriaud E, Nakada H, Leclerc C, Cantacuzene D, Short synthetic glycopeptides successfully induce antibody responses to carcinoma-associated Tn antigen, *J Pept Res* **55**, 173–80 (2000).
28 Ramachandran V, Nollert MU, Qiu H, Liu WJ, Cummings RD, Zhu C, McEver RP, Tyrosine replacement in P-selectin glycoprotein ligand-1 affects distinct kinetic and mechanical properties of bonds with P- and L-selectin, *Proc Natl Acad Sci U S A* **96**, 13771–6 (1999).
29 Zeng S, Dinter A, Eisenkratzer D, Biselli M, Wandrey C, Berger EG, Pilot scale expression and purification of soluble protein A tagged beta 1,6N-acetylglucosaminyltransferase in CHO cells, *Biochem Biophys Res Commun* **237**, 653–8 (1997).

Received 28 June 2001; revised 3 October 2001; accepted 4 October 2001