Glycoconjugate Journal 21, 197–203, 2004

Synthesis and conformational characterization of the epidermal growth factor-like domain of blood coagulation factor IX carrying xylosyl-glucose

Mayuko Kitamura,[1] Hironobu Hojo,[1] Yoshiaki Nakahara,[1] Takeshi Ishimizu[2] and Sumihiro Hase[2]

[1] *Department of Applied Biochemistry, Institute of Glycotechnology, Tokai University, 1117 Kitakaname, Hiratsuka, Kanagawa, 259-1292, Japan,* [2] *Department of Chemistry, Graduate School of Science, Osaka University, 1-1 Machikaneyamacho, Toyonaka, Osaka, 560-0043, Japan*

Solid-phase synthesis of glycopeptide generally requires the protection of both peptide side chains and hydroxyl groups of the carbohydrate portion. However, if the mild coupling conditions are used, the protection of the carbohydrate portion can be omitted. In this paper, we demonstrated it by the synthesis of Fmoc-serine carrying unmasked xylosyl glucose followed by the solid-phase synthesis of epidermal growth factor (EGF)-like domain of factor IX (45-87) using the unit. The product was well characterized by enzymatic digestion, amino acid analysis and mass spectrometry. The secondary structure of the product as well as glucosylated and non-glycosylated EGF-like domain was characterized by circular dichroism (CD) spectroscopy.
Published in 2004.

Keywords: **unmasked carbohydrate, solid-phase synthesis, EGF-like domain, glycopeptide, CD spectrum**

Introduction

EGF-like domain is one of the most commonly found modules in multidomain proteins. In some of these domains, two kinds of unusual *O*-linked structure has been found, which consists of a fucose or glucose attached to the hydroxyl group of the serine or threonine residues [1]. Recent work demonstrates that the former modification plays a significant role in altering the receptor-ligand interactions. For instance, Fringe, which had been known to modify Notch function, is found to be an *O*-fucose specific β1, 3-GlcNAc transferase [2,3]. By adding GlcNAc to fucose on the EGF repeats of Notch, Fringe enables Notch to interact with its ligands, such as Delta. The latter modification was found in the EGF-like domain of blood coagulation factor VII and IX [4]. The structure of the modification was found to be Xyl-α1, 3-Xyl-α1, 3-Glc-β1-*O*-Ser. In contrast to the *O*-fucose modification, the role of *O*-glucose remains largely unknown, except that the mutation of the carbohydrate-linked seine residue to alanine in factor VII is known to result in a decrease in factor VII clotting activity [5].

To whom correspondence should be addressed: Hironobu Hojo, Department of Applied Biochemistry, Institute of Glycotechnology, Tokai University, 1117 Kitakaname, Hiratsuka, Kanagawa, 259-1292, Japan. E-mail: hojo@keyaki.cc.u-tokai.ac.jp

The solid-phase assembly of glycopeptide requires the protection of amino acid side chains as well as hydroxyl groups of carbohydrate portion to avoid side reactions. The latter protection is generally achieved by acetyl or benzyl groups. However, the removal of these carbohydrate protecting groups might cause damages to the peptide chain. In case of the removal of acetyl groups by NaOMe in methanol (MeOH), there is a potential risk of the β-elimination of *O*-linked carbohydrates and epimerization of amino acid residues due to the basicity of NaOMe. The deprotection of benzyl groups by mild acid treatment is relatively harmless to peptide chains [6,7], but their hydrogenolytic cleavage sometimes reduces the aromatic ring of Tyr [8]. If the *O*-acylation during the solid-phase synthesis can be suppressed, the protection of the carbohydrate moiety can be omitted. Generally, it is considered that *N*-acylation to the terminal amino group is much faster than *O*-acylation. Thus, under a certain condition, the solid-phase synthesis using unmasked glycosylated amino acid unit would be realized. In fact, several reports appear on the use of Fmoc-amino acids carrying unprotected carbohydrates for the solid-phase synthesis of glycopeptide up to 20 amino acid residues [9–16]. These methods use mild coupling methods, such as pentafluorophenyl or benzotriazolyl esters to avoid the *O*-acylation. We previously demonstrated the use of Fmoc-Ser carrying unmasked Glc for

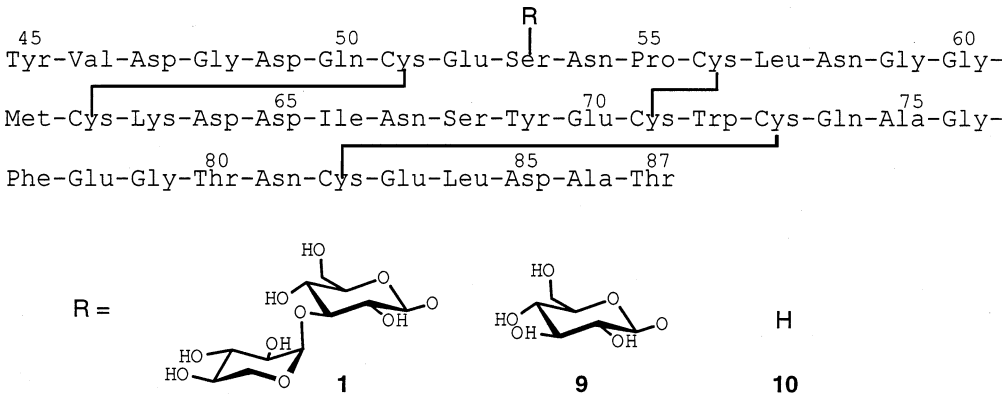

Figure 1. Amino acid sequence of the EGF-like domain of the bovine blood coagulation factor IX (45-87).

the synthesis of longer polypeptide, the EGF-like domain of bovine blood coagulation factor IX comprising 43 amino acid residues [17]. In this report, we examined the applicability of Fmoc-Ser carrying unmasked xylosyl glucose for the synthesis of the same EGF-like domain **1** (Figure 1).

Results and discussion

Fmoc-Ser carrying unmasked xylosyl-glucose **8** was prepared as shown in Figure 2. The same unit without Fmoc group was already prepared by Fukase *et al.* [18]. In their method, xylose was selectively introduced to the 3-position of 1,2,4,6-tetraprotected glucose derivative. However, α and β-isomer formed by this glycosylation were not separated by silicagel column chromatography. In addition, the synthesis of the glucose unit requires multi-step reactions starting from diisopropylideneglucose. In this synthesis, compound **2**, which has two free hydroxyl groups at 2- and 3-positions and which can be easily prepared from glucose [19], was used as a acceptor glucose. The glycosylation of **2** with benzyl-protected xylosyl fluoride using $Cp_2Zr(ClO_4)_2$ as a promoter gave the desired α-xylosylated product **4** in 32% yield. The β-isomer at 3-position was negligible. As expected, the α-xylosylation to 2-position was also occurred but in a reduced ratio (15% yield). The purification of compound **4** was easily achieved by silicagel column chromatography. In addition, purified compound **4** was crystallized from hexane-AcOEt. The product **4** was then benzoylated at 2-position to achieve the subsequent β-glycosylation with serine. The allyl group of the obtained compound **5** was isomerized to propenyl derivative by Ir complex, which was subjected to oximercuration reaction to give disaccharide with free hydroxyl group at 1-position. This compound was fluorinated by diethylaminosulfur trifluoride (DAST) to give **6**, which was then glycosylated with benzyloxycarbonyl (Z)-Ser-OBu^t to give **7**. Benzylidene group and t-Bu ester of compound **7** were removed by 90% trifluoroacetic acid (TFA) and benzoyl group was removed by NaOMe in MeOH. The remaining protecting groups were removed by catalytic hydrogenolysis. Finally, Fmoc group was introduced and the desired Fmoc-serine unit carrying unprotected xylosyl glucose unit **8** was successfully prepared.

The synthesis of the EGF-like domain was carried out as shown in Figure 3. Starting from Fmoc-Thr(Bu^t)-CLEAR Acid resin, the peptide chain was elongated by ABI433A peptide synthesizer using *FastMoc* protocol. After the sequence of EGF-like domain (54-87) was introduced, a part of the resin was taken and compound **8** was introduced manually using its

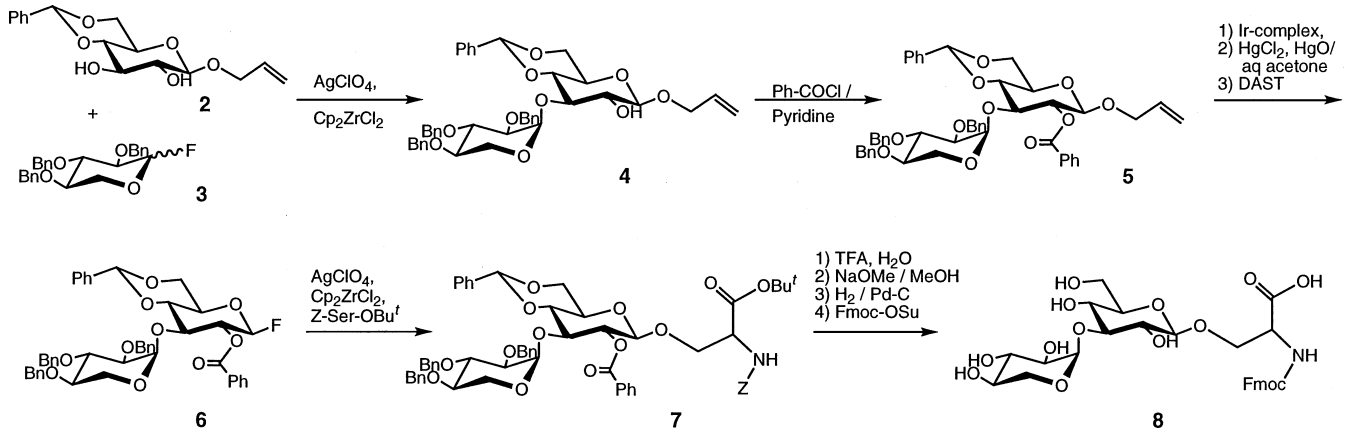

Figure 2. Synthetic procedure for Fmoc-Ser carrying xylosyl glucose.

Fmoc-Thr(Bu^t)-OCH$_2$-CLEAR-Acid-Resin

↓ ABI433A peptide synthesizer, FastMoc protocol

Fmoc-Asn(Trt)-Pro-Cys(Trt)-Leu-Asn(Trt)-Gly-Gly-Met-Cys(Trt)-Lys(Boc)-Asp(OBu^t)-Asp(OBu^t)-Ile-Asn(Trt)-Ser(Bu^t)-Tyr(Bu^t)-Glu(OBu^t)-Cys(Trt)-Trp(Boc)-Cys(Trt)-Gln(Trt)-Ala-Gly-Phe-Glu(OBu^t)-Gly-Thr(Bu^t)-Asn(Trt)-Cys(Trt)-Glu(OBu^t)-Leu-Asp(OBu^t)-Ala-Thr(Bu^t)-OCH$_2$-CLEAR-Acid-Resin

↓ 1) 20% Piperidine/NMP, 2) Fmoc-Ser(Xyl-Glc), DCC, HOBt

Fmoc-Ser(Xyl-Glc)-Asn(Trt)-Pro-Cys(Trt)-Leu-Asn(Trt)-Gly-Gly-Met-Cys(Trt)-Lys(Boc)-Asp(OBu^t)-Asp(OBu^t)-Ile-Asn(Trt)-Ser(Bu^t)-Tyr(Bu^t)-Glu(OBu^t)-Cys(Trt)-Trp(Boc)-Cys(Trt)-Gln(Trt)-Ala-Gly-Phe-Glu(OBu^t)-Gly-Thr(Bu^t)-Asn(Trt)-Cys(Trt)-Glu(OBu^t)-Leu-Asp(OBu^t)-Ala-Thr(Bu^t)-OCH$_2$-CLEAR-Acid-Resin

↓ 1) 20% Piperidine/NMP, 2) Fmoc-amino acid, DCC, HOBt

Fmoc-Tyr(Bu^t)-Val-Asp(OBu^t)-Gly-Asp(OBu^t)-Gln(Trt)-Cys(Trt)-Glu(OBu^t)-Ser(Xyl-Glc)-Asn(Trt)-Pro-Cys(Trt)-Leu-Asn(Trt)-Gly-Gly-Met-Cys(Trt)-Lys(Boc)-Asp(OBu^t)-Asp(OBu^t)-Ile-Asn(Trt)-Ser(Bu^t)-Tyr(Bu^t)-Glu(OBu^t)-Cys(Trt)-Trp(Boc)-Cys(Trt)-Gln(Trt)-Ala-Gly-Phe-Glu(OBu^t)-Gly-Thr(Bu^t)-Asn(Trt)-Cys(Trt)-Glu(OBu^t)-Leu-Asp(OBu^t)-Ala-Thr(Bu^t)-OCH$_2$-CLEAR-Acid-Resin

↓ 1) 20% Piperidine/NMP, 2) Reagent K,
3) Reduced and oxidized glutathione in AcONH$_4$, 4) RPHPLC

EGF-like domain (45-87) **1**

Figure 3. Synthetic procedure for the EGF-like domain (45-87) **1**.

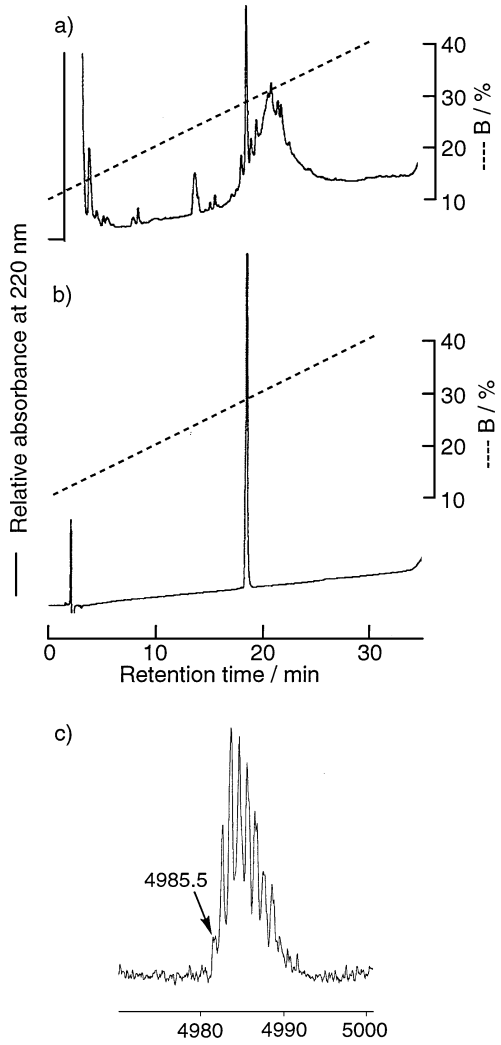

Figure 4. RPHPLC profiles and MALDI-TOF mass spectrum of peptide **1**: (a) crude peptide **1** after oxidation, (b) purified peptide **1**, (c) mass of the purified peptide **1**. Elution conditions in (a) and (b): column, Mightysil RP-18GP (4.6 × 150 mm) at a flow rate of 1 ml min^{-1}; eluent A, distilled water containing 0.1% TFA, B, acetonitrile containing 0.1% TFA.

1-hydroxybenzotriazole (HOBt) active ester at 50°C. After 1 h coupling, the resin became negative to ninhydrin test. The remaining sequence was also introduced by HOBt active ester (10 eq). The protected peptide resin thus obtained was then treated with TFA cocktail (Reagent K [20]) for 2 h at room temperature. As in the case of the EGF-like domain carrying Glc [17], the crude peptide at this stage was strongly adsorbed on HPLC column (data not shown). Thus, the crude peptide was directly subjected to the disulfide bond formation reaction in the presence of oxidized and reduced form of glutathione. After overnight reaction, a sharp peak, which has the desired mass number, was obtained as shown in Figure 4. This peak was purified by preparative RPHPLC and the desired peptide **1** was successfully obtained in 1.6% yield, which was comparable to that of the glucosylated and non-glycosylated domain [17]. The MALDI-TOF mass spectrum and amino acid composition agreed well with the theoretical value of the desired structure. The mode of the disulfide bond in Figure 1 was confirmed by the enzymatic digestion followed by MALDI-TOF mass analysis as described in the experimental section.

CD spectrum measurement

CD spectrum of peptide **1** as well as glucosylated **9** and non-glycosylated EGF-like domain **10** were measured in 10 mM Tris-HCl buffer (pH 7.5) as shown in Figure 5. In contrast to EGF itself, the EGF-like domain has a Ca^{2+} binding site at its *N*-terminal region. Thus, the spectrum was recorded in the presence and the absence of Ca^{2+}. In the Ca^{2+} free state, the spectra are slightly varied among three peptides. This result might show that the conformation of the peptides are less defined in the absence of Ca^{2+} and that the glycan moiety shows some steric effect on peptide structure. In contrast, all peptides show a similar spectrum in the presence of Ca^{2+}. In addition, a shoulder at around 210 nm, which is indicative of β-sheet structure, clearly appeared. The spectrum is similar to that of EGF [21,22]. These results suggest that in the presence of Ca^{2+}, a well defined structure is formed without the need of carbohydrate moiety. This means that the major role of the glycan

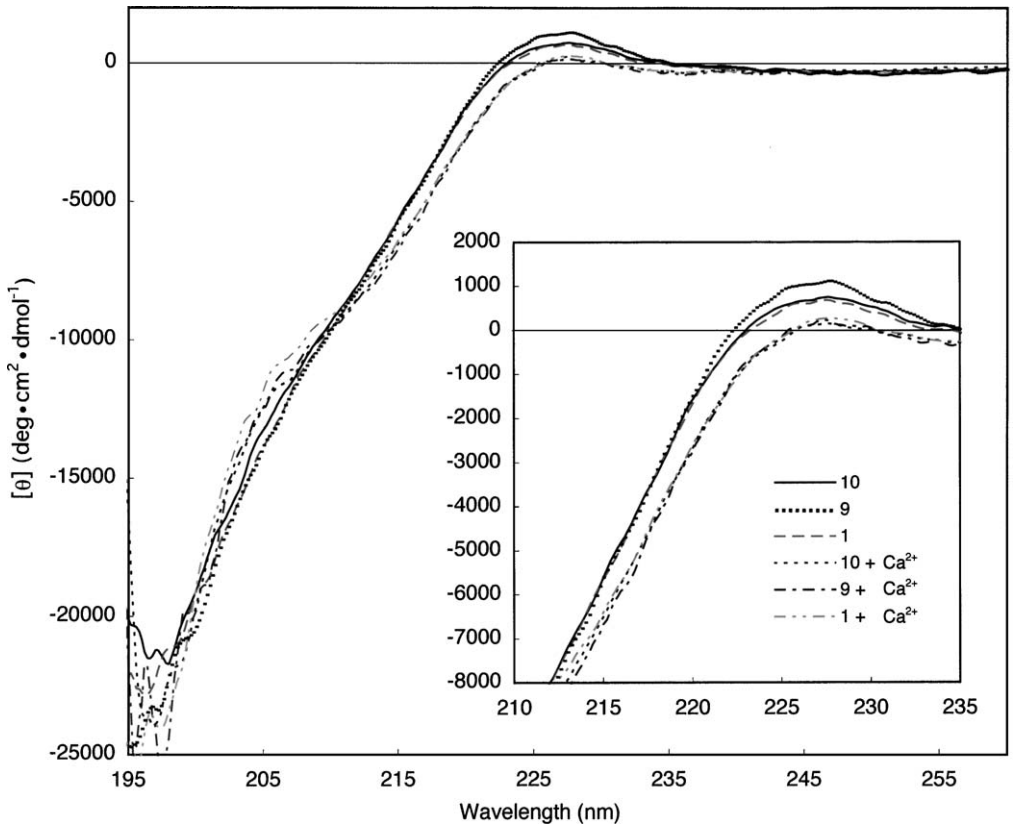

Figure 5. CD spectra of peptide **1**, **9**, **10** in 10 mM Tris-HCl (pH 7.5) in the presence and in the absence of 20 mM $CaCl_2$.

portion in the EGF-like domain is not the formation or the stabilization of the three dimensional structure of the domain. The structure of the non-glycosylated EGF-like domain in the presence of Ca^{2+} was determined by Rao *et al*. [23]. According to the structure, the hydroxyl group of Ser^{53} points outside of the molecule. Thus, it is probable that the carbohydrate moiety can be accommodated without changing the conformation of the peptide chain. Further studies are underway to clarify the role of the glycan moiety.

In conclusion, we prepared Fmoc-Ser carrying unmasked xylosyl glucose and successfully used it for the synthesis of EGF-like domain of factor IX. In the presence of Ca^{2+}, the peptide shows a similar CD spectrum to EGF domain. In addition, the presence of glucose or xylosyl glucose on Ser^{53} has little effect on the conformation of the peptide showing that the carbohydrate points outside of the molecule.

Experimental

Optical rotation values were determined with a DIP-370 polarimeter (Jasco, Tokyo). ^{1}H-NMR spectra were recorded with a AL-400 spectrometer (Jeol, Tokyo) in $CDCl_3$ except compound **8**. MALDI-TOF mass spectra were recorded with a Voyager-DE PRO spectrometer (Applied Biosystems, CA). Amino acid composition was determined with a LaChrom amino acid analyzer (Hitachi, Tokyo) after hydrolysis with 6 M HCl at 150°C for 2 h in an evacuated sealed tube. Fmoc-Thr(Bu^t)-OCH_2-CLEAR-Acid Resin and Fmoc-amino acids were purchased from Peptide Institute Inc. (Minoh, Japan). CD spectra were recorded on J-820 spectroporalimeter (Jasco, Tokyo) with a cell path length of 1 mm. Glucosylated **9** and non-glycosylated EGF-like domain **10** were prepared as described previously [17].

Allyl 2, 3, 4-tri-*O*-benzyl-α-D-xylopyranosyl-(1→3)-4, 6-*O*-benzylidene-β-D-glucopyranoside **4**

To a stirred mixture of Cp_2ZrCl_2 (1.9 g, 6.5 mmol) and $AgClO_4$ (2.7 g, 13 mmol) and dried MS 4A (8 g) in anhydrous CH_2Cl_2 (20 ml) under Ar at −20°C, compound **3** (1.8 g, 4.4 mmol) and **2** (1.3 g, 4.8 mmol) in CH_2Cl_2 (20 ml) were added. The mixture was stirred at −20°C for 1 h and at −5°C for 4 h before the reaction was quenched by aqueous $NaHCO_3$ solution. The mixture was diluted with $CHCl_3$ and filtered through Celite. The filtrate was successively washed with water and brine, dried over Na_2SO_4 and concentrated *in vacuo*. The crude product was chromatographed on silica gel with toluene-EtOAc (4.5:1) to give **4** (1.0 g, 1.4 mmol, 32%). M.p. 109–112°C. $[\alpha]_D$ +15.5° (*c* 0.5 in $CHCl_3$). R_f 0.38 (4:1 toluene-EtOAc). ^{1}H-NMR: δ: 5.94 (m, 1H, $-CH_2CH{=}CH_2$), 5.48 (s, 1H, $PhCH(O)_2$), 5.31 (d, 1H, J = 3.9 Hz, Xyl H-1), 5.25 (dd, 1H, J = 1.3, 10.4 Hz, $-CH_2CH{=}CH_2$), 4.91 (d, 1H,

J = 10.7 Hz, —CH_2-Ph), 4.83 (d, 1H, J = 11.0 Hz, —CH_2-Ph), 4.72 (d, 1H, J = 11.7 Hz, —CH_2-Ph), 4.60 (d, 1H, J = 11.7 Hz, —CH_2-Ph), 4.56 (d, 1H, J = 12.2 Hz, —CH_2-Ph), 4.47 (d, 1H, J = 7.8 Hz, Glc H-1), 4.38 (d, 1H, J = 12.2 Hz, —CH_2-Ph), 4.33 (dd, 1H, J = 4.9, 10.3 Hz, —CH_2CH=CH$_2$), 4.16 (m, 1H, —CH_2CH=CH$_2$), 4.04 (brt, 1H, J = 11.0 Hz, Xyl H-5), 3.95 (brt, 1H, J = 9.1 Hz, Glc H-3), 3.91 (brt, 1H, J = 9.0 Hz, Xyl H-3), 3.81 (brt, 1H, J = 9.3 Hz, Glc H-4), 3.79 (brt, 1H, J = 10.4 Hz, Glc H-6), 3.66 (m, 1H, Glc H-2), 3.59 (dd, 1H, J = 5.7, 11.2 Hz, Xyl H-5), 3.39 (dd, 1H, J = 3.7, 9.8, Xyl H-2), 2.72 (d, 1H, J = 2.7 Hz, Glc OH). HRFABMS: Calcd for $C_{42}H_{46}O_{10}$ $[M+H]^+$: 711.3169. Found: 711.3187.

Allyl 2,3,4-tri-*O*-benzyl-α-D-xylopyranosyl-(1→3)-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranoside **5**

Compound **4** (200 mg, 0.28 mmol) was dissolved in pyridine (3.0 ml) and benzoyl chloride (0.33 ml, 2.8 mmol) was added. After the solution was kept at room temperature overnight, the solvent was removed *in vacuo*. The residue was dissolved in EtOAc, successively washed with water and brine, dried over Na_2SO_4, and concentrated *in vacuo*. The crude product was chromatographed on silica gel with $CHCl_3$-EtOAc-hexane (9:1:2) to give **5** (214 mg, 0.26 mmol, 93%). $[\alpha]_D$ +67.3° (*c* 0.5 in $CHCl_3$). R_f 0.45 (7:1 toluene-EtOAc). ^{1}H-NMR: δ: 5.75 (m, 1H, —CH$_2$CH=CH$_2$), 5.45 (s, 1H, PhCH(O)$_2$), 5.43 (brt, 1H, J = 8.8 Hz, Glc H-2), 5.37 (d, 1H, J = 3.7 Hz, Xyl H-1), 5.22 (brd, 1H, J = 17.3 Hz, —CH$_2$CH=CH_2), 5.11 (brd, 1H, J = 10.5 Hz, —CH$_2$CH=CH_2), 4.83 (d, 1H, J = 10.7 Hz, —CH_2-Ph), 4.75 (d, 1H. J = 8.1 Hz, Glc H-1), 4.74 (d, 1H, J = 10.7 Hz, —CH_2-Ph), 4.52 (d, 1H, J = 12.5 Hz, —CH_2-Ph), 4.28 (brt, 1H, J = 9.5 Hz, Glc H-3), 4.16 (d, 1H, J = 12.0 Hz, —CH_2-Ph), 4.10 (dd, 1H, J = 6.2, 13.3 Hz, —CH_2CH=CH$_2$), 3.95 (brt, 1H, J = 9.3 Hz, Glc H-4), 3.84 (brt, 1H, J = 10.4 Hz, Glc H-6), 3.73 (brt, 1H, J = 9.0 Hz, Xyl H-5), 3.57 (td, 1H, J = 4.9, 9.8, 9.8 Hz, Glc H-5), 3.36 (brt, 1H, J = 10.5 Hz, Xyl H-3), 3.17 (dd, 1H, J = 5.1, 10.3 Hz, Xyl H-4). HRFABMS: Calcd for $C_{49}H_{51}O_{11}$ $[M+H]^+$: 815.3431. Found: 815.3530.

2,3,4-Tri-*O*-benzyl-α-D-xylopyranosyl-(1→3)-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl fluoride **6**

A suspension of the Ir complex Ir(COD)[PCH$_3$(Ph)$_2$]$_2$PF$_6$ (19 mg, 39 μmol) in tetrahydrofurane (THF) (2 ml) was stirred in an atmosphere of H_2 until a red color disappeared. After the atmosphere was replaced with nitrogen, compound **5** (190 mg, 0.24 mmol) in THF (2 ml) was added. The solution was stirred under N_2 for 15 min and the solvent was removed *in vacuo*. The residue was dissolved in 90% aqueous acetone and stirred with $HgCl_2$ (160 mg, 0.58 mmol) and HgO (17 mg, 78 μmol) for 2.5 h at room temperature. After the solvent was removed *in vacuo*, the residue was dissolved in EtOAc, successively washed with aqueous KI, water, brine, and dried over Na_2SO_4. The solvent was removed *in vacuo*, and the residue was chromatographed on silica gel with toluene-EtOAc (3:1) to give 2,3,4-tri-*O*-benzyl-α-D-xylopyranosyl-(1→3)-2-*O*-benzoyl-4, 6-*O*-benzylidene-β-D-glucopyranose (151 mg, 0.19 mmol, 79%). R_f 0.43 (3:1 toluene-EtOAc). HRFABMS: Calcd for $C_{46}H_{47}O_{11}$ $[M+H]^+$: 775.3118. Found: 775.3132. This compound (145 mg, 0.19 mmol) was dissolved in CH_2Cl_2 (2 ml) and diethylaminosulfur trifluoride (30 μl, 0.23 mmol) was added at 0°C under Ar. After the solution was stirred for 15 min, the reaction was quenched by adding MeOH. The solvent was removed *in vacuo* and the residue was dissolved in EtOAc, washed successively with water, brine and dried over Na_2SO_4. After the solvent was concentrated *in vacuo*, the residue was chromatographed on silica gel with hexane-EtOAc (3:1) to give **6** (100 mg, 0.13 mmol). $[\alpha]_D$ + 71.7° (*c* 0.5 in $CHCl_3$). R_f 0.28 (3:1 hexane-EtOAc). ^{1}H-NMR: δ: 5.53 (dd, 1H, J = 6.1, 54 Hz, Glc H-1), 5.51 (m, 1H, Glc H-2), 5.47 (s, 1H, PhCH(O)$_2$), 5.32 (d, 1H, J = 3.7 Hz, Xyl H-1), 4.85 (d, 1H, J = 10.8 Hz, —CH_2-Ph), 4.77 (d, 1H, J = 10.7 Hz, —CH_2-Ph), 4.55 (d, 1H, J = 12.2 Hz, —CH_2-Ph), 4.46 (d, 1H, J = 11.7 Hz, —CH_2-Ph), 4.44 (d, 1H, J = 12.2 Hz, —CH_2-Ph), 4.41 (dd, 1H, J = 4.9, 10.3 Hz, Glc H-6), 4.28 (d, 1H, J = 11.7 Hz, —CH_2-Ph), 4.27 (brt, 1H, J = 7.3 Hz, Glc H-3), 4.12 (brt, 1H, J = 8.9 Hz, Glc H-4), 3.86 (brt, 1H, J = 10.3 Hz, Glc H-6). HRFABMS: Calcd for $C_{46}H_{46}FO_{10}$ $[M+H]^+$: 777.3075. Found: 777.3066.

N-(Benzyloxycarbonyl)-*O*-[2,3,4-tri-*O*-benzyl-α-D-xylopyranosyl-(1→3)-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl]-L-serine *t*-butyl ester **7**

To a stirred mixture of Cp_2ZrCl_2 (75 mg, 0.26 μmol) and $AgClO_4$ (110 mg, 0.52 μmol) and dried MS 4A (2 g) in anhydrous dichloroethane (10 ml) under Ar at −15°C, compound **6** (100 mg, 0.13 mmol) and Z-Ser-OBu^t (57 mg, 0.19 mmol) in dichloroethane (10 ml) were added. The mixture was stirred at −0°C for 3 h before the reaction was quenched by aqueous $NaHCO_3$ solution. The mixture was diluted with $CHCl_3$ and filtered through Celite. The filtrate was successively washed with water and brine, dried over Na_2SO_4 and concentrated *in vacuo*. The crude product was chromatographed on silica gel with toluene-EtOAc (9:1) to give **7** (66 mg, 63 μmol, 49%). $[\alpha]_D$ +52.0° (*c* 0.5 in $CHCl_3$). R_f 0.27 (9:1 toluene-EtOAc). ^{1}H-NMR: δ 5.44 (s, 1H, PhCH(O)$_2$), 5.36 (d, 1H, J = 3.7 Hz, Xyl H-1), 5.33 (brt, 1H, J = 8.8 Hz, Glc H-2), 5.03 (d, 1H, 12.2 Hz, —CH_2-Ph), 4.94 (d, 1H, J = 12.0 Hz, —CH_2-Ph), 4.81 (d, 1H, J = 10.7 Hz, —CH_2-Ph), 4.72 (d, 1H, J = 11.0 Hz, —CH_2-Ph), 4.62 (d, 1H, J = 8.1 Hz, Glc H-1), 4.52 (d, 1H, J = 12.5 Hz, —CH_2-Ph), 4.15 (d, 1H, J = 12.5 Hz, —CH_2-Ph), 3.90 (brt, 1H, J = 9.3 Hz, Glc H-6), 3.70 (brt, 1H, J = 9.0 Hz, Xyl H-5), 3.51 (m, 1H, Glc H-5), 3.34 (brt, 1H, J = 10.2 Hz, Xyl H-3), 3.17 (dd, J = 4.6, 9.8 Hz, Xyl H-4). MALDI TOF mass: Calcd for $C_{61}H_{65}NNaO_{15}$ $[M+H]^+$: 1074.43. Found: 1074.39.

N-(9-Fluorenylmethoxycarbonyl)-*O*-[α-D-xylopyranosyl-(1 → 3)-β-D-glucopyranosyl]-L-serine **8**

Compound **7** (48 mg, 46 μmol) was dissolved in 90% TFA (0.6 ml) and stirred at room temperature for 30 min. TFA was

removed by N_2 stream and the product was precipitated by adding water. After washing with water 3 times, the product was dried *in vacuo*. The obtained powder was dissolved in dry MeOH (0.4 ml) and 1 M NaOMe in MeOH (0.4 ml) was added. After stirring for 8 h at room temperature, the solution was neutralized by amberlyst and the solvent was removed *in vacuo*. The residual mass was dissolved in dry MeOH (0.8 ml) containing 1 M HCl (60 μl) and hydrogenated in the presence of catalytic amount of 10% Pd/C overnight. The catalyst was filtered off through Celite and the filtrate was evaporated *in vacuo*. The residue was dissolved in 10% aq Na_2CO_3 (0.4 ml) and dropped into Fmoc-OSu (21 mg, 62 μmol) in 1,2-dimethoxyethane (0.4 ml) and the resultant mixture was stirred overnight. The mixture was washed 3 times with ether, acidified by acetic acid and purified by RPHPLC to give **8** (27 mg, 43 μmol, 93%). $[\alpha]_D$ +51.8° (*c* 0.5 in MeOH). R_f 0.47 (*n*-butanol:AcOH:water 4:1:1). ^{1}H-NMR (CD_3OD): δ 5.16 (d, 1H, J = 3.7 Hz, Xyl H-1), 4.44 (m, 1H, Ser αH), 4.32 (d, 1H, J = 7.6 Hz, Glc H-1), 3.81 (dd, 1H, J = 3.3, 10.1 Hz, Ser βH), 3.67 (dd, 1H, J = 5.6, 12.0 Hz, H-6), 3.63 (brt, 1H, J = 8.9 Hz, Xyl H-3), 3.42 (dd, 1H, 3.7, 9.5 Hz, Xyl H-2). MALDI TOF mass: Calcd for $C_{61}H_{65}NNaO_{15}$ $[M+H]^+$: 644.20. Found: 644.40.

Solid-phase synthesis of the EGF-like domain carrying xlyosyl-glucose **1**

Starting from Fmoc-Thr(Bu^t)-OCH_2-CLEAR-Acid Resin (content of Thr: 0.2 mmol/g), peptide chain was elongated by ABI433A peptide synthesizer (Applied Biosystems, Foster City, CA) using *FastMoc* protocol. After the peptide chain corresponding to the sequence of bovine EGF-like domain (54-87), Fmoc-Asn(Trt)-Pro-Cys(Trt)-Leu-Asn(Trt)-Gly-Gly-Met-Cys(Trt)-Lys(Boc)-Asp(OBu^t)-Asp(OBu^t)-Ile-Asn(Trt)-Ser(Bu^t)-Tyr(Bu^t)-Glu(OBu^t)-Cys(Trt)-Trp(Boc)-Cys(Trt)-Gln(Trt)-Ala-Gly-Phe-Glu(OBu^t)-Gly-Thr(Bu^t)-Asn(Trt)-Cys(Trt)-Glu(OBu^t)-Leu-Asp(OBu^t)-Ala-Thr(Bu^t)-OCH_2-CLEAR-Acid Resin was obtained, a part of this resin (ca 11 μmol) was taken and the Fmoc group was removed by 20% piperidine-1-methyl-2-pyrrolidinone (NMP) for 5 and 15 min. Fmoc-Ser(Xyl-Glc) **8** (12 mg, 19 μmol) was preactivated with 1M 1,3-dicyclohexylcarbodiimide (DCC)/NMP (25 μl) and 1M HOBt/NMP (25 μl) for 30 min at room temperature and reacted with the resin at 50°C for 1 h. Ninhydrin test indicated the complete introduction of the glycosylated serine residue. The remaining sequence was introduced manually using Fmoc-amino acid benzotriazolyl ester (100 μmol). After the complete assembly of the peptide sequence, Fmoc group was removed and the resin was dried *in vacuo*. A part of the resin (86 mg out of 136 mg obtained) was treated with Reagent K (1.6 ml) for 2 h at room temperature. TFA was removed by nitrogen stream and the ether was added to form a precipitate, which was washed twice with ether and dried *in vacuo*. The peptide was extracted by 6M guanidine hydrochloride (2 ml) and filtered through a membrane filter. The solution was dropped into 0.1 M ammonium acetate (90 ml, pH 8.0) containing oxidized (6 mg) and reduced form of glutathione (30 mg) and stirred 2 days at room temperature. The solution was acidified by acetic acid and loaded on RPHPLC column. The main fraction was collected and lyophilized to give the desired peptide **1**. The yield was 120 nmol (1.6% based on the Thr content on the starting resin) MADLI-TOF mass Found: m/z 4985.5 $(M+H)^+$. Calcd: m/z 4985.8 $(M+H)^+$; Amino acid analysis: $Asp_{8.73}Thr_{1.89}Ser_{1.72}Glu_{5.84}Pro_{0.96}Gly_5Ala_{2.00}Cys_{nd}Val_{0.90}Met_{0.94}Ile_{0.94}Leu_{2.01}Tyr_{1.91}Phe_{0.98}Lys_{0.85}$.

Determination of the disulfide bond pairings

Peptide **1** (30 μg) was dissolved in 0.1 M pyridine acatate (pH 6.5, 30 μl) and thermolysin (10 μg) was added. The solution was kept at 45°C for 6 h. The peptide fragments were separated by RPHPLC (column: Mightysil RP-18GP, 4.6 × 150 mm) using a linear gradient of acetonitrile containing 0.1% TFA from 2% to 15% over 26 min. A peak eluted at 24 min had a mass number of 899.2, which corresponds to the calculated mass value ($[M+H]^+ = 899.3$) of the disulfide bonded fragment WCQ(72-74)-TNCE(80-83). To confirm the disulfide bond between 51-62, peptide **1** (30 μg) was dissolved in 70% aq formic acid (30 μl) and BrCN (30 μg) was added. After the solution was kept overnight in the dark, the solvent was removed *in vacuo*. The residual mass was dissolved in distilled water and lyophilized. The obtained powder was dissolved in 0.1 M NH_4HCO_3 (20 μl) and digested with *Staphylococcus aureus* protease V8 (3 μg) at 37°C for 6 h. The fragments were separated by the same HPLC column using a linear gradient of acetonitrile containing 0.1% TFA from 10% to 30% over 20 min. The peak eluted at 10.8 min had a mass number of 2012.1, which corresponds to the calculated mass value ($[M+H]^+ = 2011.8$) of the disulfide bonded form of YVDGDQCE(45-52) – CKDDINSYE(62-70).

CD spectrum measurement

EGF-like domains were dissolved in 10 mM Tris-HCl (pH 7.5, 0.4 ml) at the concentration of ca 0.4 mg/ml and the spectra were recorded between 190 and 260 nm. The mean residue ellipticity (deg cm^2/dmol) was determined assuming a mean residue weight of 109. The effect of Ca^{2+} was measured by adding 2 M $CaCl_2$ (4 μl) to the above solution.

Acknowledgment

We are indebted to Dr. Yuko Nakahara (Institute of Glycotechnology) for high resolution mass analysis. This work was supported by Grant-in Aid for Scientific Research (16580093) from the Ministry of Education, Culture, Sports, Sciences and Technology of Japan. We also thank Tokai University for a grant-aid for high-technology research.

References

1 Harris RJ, Spellman MW, *O*-linked fucose and other post-translational modifications unique to EGF modules, *Glycobiology* **3**, 219–24 (1993).

2 Bruckner K, Perez L, Clausen H, Cohen S, Glycosyltransferase activity of Fringe modulates Notch-Delta interactions, *Nature* **406**, 411–5 (2000).

3 Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, Wilson R, Wang Y, Stanley P, Irvine KD, Haltiwanger RS, Vogt TF, Fringe is a glycosyltransferase that modifies Notch, *Nature* **406**, 369–75 (2000).

4 Hase S, Kawabata S, Nishimura H, Takeya H, Sueyoshi T, Miyata T, Iwanaga S, Takao T, Shimonishi Y, Ikenaka T, A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX, *J Biochem (Tokyo)* **104**, 867–8 (1988).

5 Bjoern S, Foster DC, Thim L, Wiberg FC, Christensen M, Komiyama Y, Pedersen AH, Kisiel W, Human plasma and recombinant factor VII, *J Biol Chem* **266**, 11051–7 (1991).

6 Takano Y, Habiro M, Someya M, Hojo H, Nakahara Y, Preparation of core 2 type tetrasaccharide carrying decapeptide by benzyl protection-based solid-phase synthesis strategy, *Tetrahedron Lett* **43**, 8395–9 (2002).

7 Hojo H, Haginoya E, Matsumoto Y, Nakahara Y, Nabeshima K, Toole BP, Watanabe Y, The first synthesis of peptide thioester carrying *N*-linked core pentasaccharide through modified Fmoc thioester preparation: Synthesis of an *N*-glycosylated Ig domain of emmprin, *Tetrahedron Lett* **44**, 2961–4 (2003).

8 Guo ZW, Nakahara Y, Nakahara Y, Ogawa T, Solid-phase synthesis of a glycosylated peptide fragment of the IL-8 receptor containing two vicinal oligosaccharide chains, *Carbohydr Res* **303**, 373–7 (1997).

9 Otvos LJ, Urge L, Hollosi M, Wroblewski K, Graczk G, Fasman GD, Thurin J, Automated solid-phase synthesis of glycopeptides. Incorporation of unprotected mono- and disaccharide units of *N*-glycoprotein antennae into T cell epitopic peptides, *Tetrahedron Lett* **31**, 5889–92 (1990).

10 Reimer KB, Meldal M, Kusumoto S, Fukase K, Bock K, Small-scale solid-phase *O*-glycopeptide synthesis of linear and cyclized hexapeptides from blood-clotting factor IX containing *O*-(α-D-Xyl-1-3-D-Xyl-1-3-β-D-Glc)-L-Ser, *J Chem Soc, Perkin Trans* **1**, 925–32 (1993).

11 Seitz O, Wong C-H, Chemoenzymatic solution- and solid-phase synthesis of *O*-glycopeptides of the mucin domain of MAdCAM-1. A general route to *O*-LacNAc, *O*-sialyl-LacNAc, and *O*-sialyl-Lewis-X peptides, *J Am Chem Soc* **119**, 8766–76 (1997).

12 Mizuno M, Haneda K, Iguchi R, Muramoto I, Kawakami T, Aimoto S, Yamamoto K, Inazu T, Synthesis of a glycopeptide containing oligosaccharides: Chemoenzymatic synthesis of a eel calcitonin analogues having natural N-linked oligosaccharides, *J Am Chem Soc* **121**, 284–90 (1999).

13 Ichiyanagi T, Takatani M, Sakamoto K, Nakahara Y, Ito Y, Hojo H, Nakahara Y, Synthesis of mucin-type glycopeptide (β hCG 130–145) by on-resin fragment condensation of the glycopeptide segments carrying unmasked oligosaccharides, *Tetrahedron Lett* **43**, 3297–300 (2002).

14 Gobbo M, Biondi L, Filira F, Gennaro R, Benincasa M, Scolaro B, Rocchi R, Antimicrobial peptides: Synthesis and antibacterial activity of linear and cyclic drosocin and apidaecin 1b analogues, *J Med Chem* **45**, 4494–504 (2002).

15 Xue J, Guo Z, Efficient synthesis of complex glycopeptides based on unprotected oligosaccharides, *J Org Chem* **68**, 2713–9 (2003).

16 Yamamoto N, Ohmori Y, Sakakibara T, Sasaki K, Juneja LR, Kajihara Y, Solid-phase syntheis of sialylglycopeptides through selective esterification of the silalic acid residues of an Asn-linked complex-type sialyloligsaccharide, *Angew Chem Int Ed* **42**, 2537–40 (2003).

17 Takemura T, Hojo H, Nakahara Y, Ishimizu T, Hase S, Application of Fmoc-amino acid carrying an unmasked carbohydrate to the synthesis of the epidermal growth factor-like domain of bovine blood coagulation factor IX, *Org Biomol Chem* **2**, 133–6 (2004).

18 Fukase K, Hase S, Ikenaka T, Kusumoto S, Synthesis of new serine-linked oligosaccharides in blood-clotting factors VII and IX and protein Z, *Bull Chem Soc Jpn* **65**, 436–45 (1992).

19 Sugawara F, Nakayama H, Ogawa T, Stereoselective synthesis of 1-α-cellobiosyloxy-2(R)- and 2(S)-propanol, *Carbohydr Res* **108**, C5–9 (1982).

20 King DS, Fields CG, Fields GB, A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis, *Int J Peptide Protein Res* **36**, 255–66 (1990).

21 Taylor JM, Mitchell WM, Cohen S, Epidermal growth factor, *J Biol Chem* **247**, 5928–34 (1972).

22 Holladay LA, Savage Jr. CR, Cohen S, Puett D, Conformation and unfolding of epidermal growth factor and derivatives, *Biochemistry* **15**, 2624–33 (1976).

23 Rao Z, Handford P, Mayhew M, Knott V, Brownlee GG, Stuart D, The structure of a Ca^{2+}-binding epidermal growth factor-like domain: Its role in protein-protein interactions, *Cell* **82**, 131–41 (1995).