Preparation of an *N*-Linked Glycopeptide Containing 6-ThioGlcNAc

Injae Shin,* Sungjin Park, Kwan Soo Kim, Jin Won Cho,† and Dongyeol Lim‡

Department of Chemistry and †Department of Biology, Yonsei University, Seoul 120-749, Korea
‡Department of Chemistry, Sejong University, Seoul 143-747, Korea
Received September 11, 2001

Keywords : Carbohydrates, Glycopeptides, Glycosylations, Chemoselective ligation.

The *O*- and *N*-glycosylated proteins are expressed in eukaryotic cells as heterogeneous mixtures of glycoforms, namely, proteins possessing heterogeneous carbohydrate moieties and thus their purification from natural sources is difficult. As a consequence, the structural effects of carbohydrates on glycoproteins and biological functions of glycoproteins remain elusive. It has been well-documented that carbohydrates of glycoproteins modulate receptor binding and signaling, and influence the intrinsic properties of protein backbones, resulting in the proper folding of proteins, increased thermal stability and resistance to proteases.[1] Therefore, it is imperative to readily access glycoproteins with well-defined oligosaccharide chains to elucidate their biological functions.

Recently, many attempts to introduce carbohydrate moieties into proteins or peptides at a specific position *via* non-native glycosidic linkage in a chemoselective manner have been made.[2,3] In an effort to develop a new methodology to prepare homogeneous glycoproteins, we have investigated the chemoselective ligation of carbohydrates containing a maleimide group to peptides or proteins *via* a stable thioether linkage.[4] As part of our ongoing work, we prepared thiol-containing *N*-acetylglucosaminyl serine (**1**) as a building block for synthesis of *N*-linked glycopeptides bearing 6-ThioGlcNAc, that can be further glycosylated with thiol-reactive carbohydrates at 6-SH site through a disulfide bond as shown in Figure 1. The *N*-linked glycosylation on glycoproteins is catalyzed by oligosaccharyl transferase during co-translational process.[5] Carbohydrate moieties in *N*-glycosylated proteins are covalently attached to an asparagines residue in the consensus sequence of Asn-X-Thr/Ser, where X is any amino acid except proline, and are known to influence the folding of proteins or the stability of the protein backbones.[5]

Synthesis of a protected 6-ThioGlcNAc-Ser monomer (**1**) was efficiently achieved from *N*-acetylglucosamine (**2**) by the reactions delineated in Scheme 1. Peracetylated GlcNAc-N_3 (**3**) derived from **2** according to a known procedure[6] was deacetylated quantitatively under basic conditions. A selective tosylation of a primary alcohol in **4** with 2 equiv of TsCl in pyridine followed by acetylation of the secondary alcohols produced a mono-tosylated azide **5** in 78% yield. Tosylation of **4** with less than 2 equiv of TsCl furnished the desired product in low yield. It was noted that the progress of tosylation of the primary alcohol in **4** should be carefully monitored by TLC to prevent the bis-tosylation of **4**. Substitution of tosyl group by thioacetyl group with potassium thioacetate (KSAc) in the presence of 18-crown-6 at 60 °C provided a thioacetylated azide **6** in 80% yield. It is worthwhile to mention that the substitution reaction of tosyl group in the absence of 18-crown-6 gave a poor yield (< 30 % yield) of **6** as a result of the formation of an unidentified side product. Reduction of **6** and subsequent coupling of the

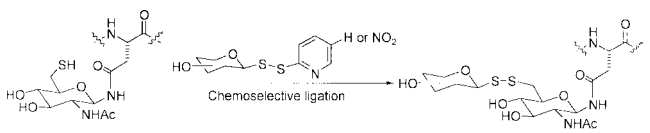
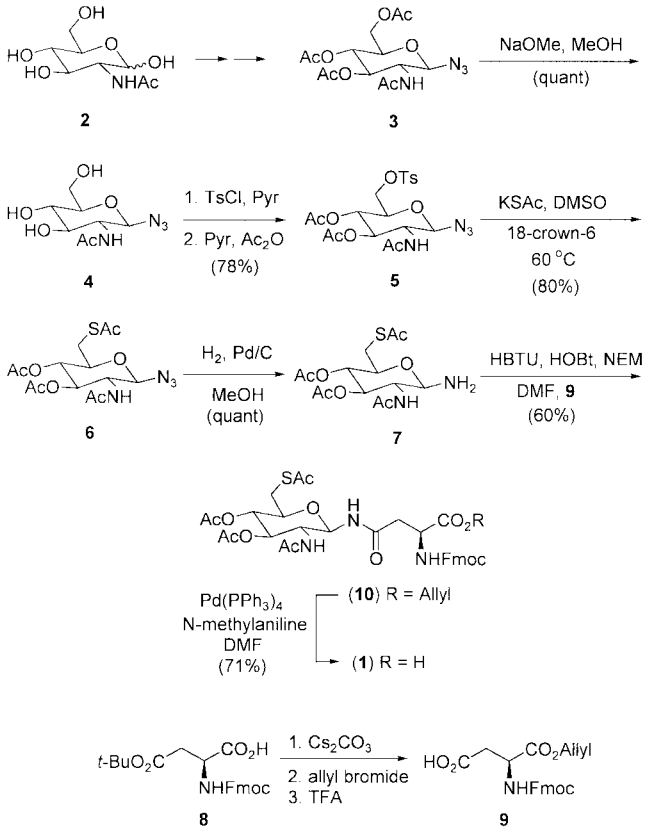
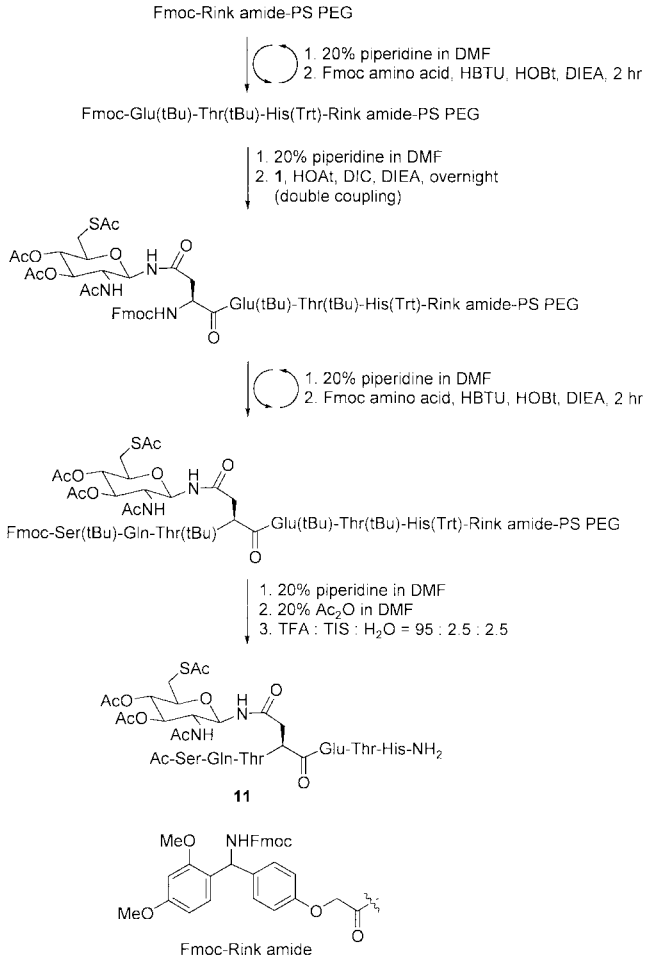


Figure 1



Scheme 1

*To whom correspondence should be addressed. Tel: +82-2-2123-2631; Fax: +82-2-364-7050; E-mail: injae@yonsei.ac.kr



Scheme 2

resultant amine **7** to **9** obtained from Fmoc-Asp(tBu)-OH (**8**) in three steps[7] by *O*-benzotriazole-*N,N,N'N'*-tetramethyluronium-hexafluorophosphate (HBTU), *N*-hydroxybenzotriazole (HOBt) and *N*-ethylmorpholine (NEM) afforded a glycosylated serine **10** in 60% yield. Finally, the allyl group in **10** was removed using $Pd(PPh_3)_4$ in the presence of *N*-methylaniline as a hydrogen donor to give the desired monomer **1** in 71% yield.[8]

Next, we prepared a glycopeptide **11** possessing 6-ThioGlcNAc. The glycopeptide **11** was synthesized on PS-PEG (polystyrene-polyethylene glycol) Rink amide resin on a 0.5 mmol scale using Fmoc-amino acids (Scheme 2).[9] Stepwise peptide assembly was performed by a manual peptide synthesis using HOBt/HBTU-mediated couplings, except for the coupling of **1**, which was carried out by double coupling using the more powerful coupling reagents of DIC/HOAt.[10] After completion of chain assembly, Fmoc group of *N*-terminus was removed and the exposed NH_2 was acetylated. Finally, glycopeptide **11** was cleaved from the solid support by treatment with TFA : triisopropylsilane : H_2O (95 : 2.5 : 2.5) and characterized by ESI MS.[11]

In summary, we developed an efficient synthesis of a thiol-containing *N*-acetylglucosaminyl serine monomer and synthesized the glycopeptide mimetic using a prepared monomer. Further glycosylation of a glycopeptide with thiol-reactive carbohydrates after deacetylation is in progress.

Acknowledgment. This work was supported by a grant of the Korea Science and Engineering Foundation (1999-2-12300-005-5).

References

1. (a) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683. (b) Varki, A. *Glycobiology* **1993**, *3*, 97.
2. For reviews (a) Stowell, C. P.; Lee, Y. C. *Adv. Carbohydr. Chem. Biochem.* **1980**, *37*, 225. (b) Lemieux, G. A.; Bertozzi, C. R. *TIBTECH.* **1998**, *16*, 506.
3. (a) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1991**, *32*, 6793. (b) Wong, S. Y. C.; Guile, G. R.; Dwek, R. A.; Arsequell, G. *Biochem. J.* **1994**, *300*, 843. (c) Macindoe, W. M.; van Oijen, A. H.; Boons, G.-J. *Chem. Commun.* **1998**, 847. (d) Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A.; Jones, J. B. *Tetrahedron:Asymmetry* **2000**, *11*, 245. (e) Rodriguez, E. C.; Winans, K. A.; King, D. S.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1997**, *119*, 9905. (f) Marcaurelle, L. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2001**, *123*, 1587. (g) Andersson, L.; Stenhagen, G.; Baltzer, L. *J. Org. Chem.* **1998**, *63*, 1366.
4. (a) Shin, I.; Jung, H.-j.; Cho, J. *Bull. Korean Chem. Soc.* **2000**, *21*, 845. (b) Shin, I.; Jung, H.-j.; Lee, M.-r. *Tetrahedron Lett.* **2001**, *42*, 1325.
5. Imperiali, B.; O'Connor, S. E. *Curr. Opin. Chem. Biol.* **1999**, *3*, 643.
6. Tropper, F. D.; Andersson, F. O.; Braun, S.; Roy, R. *Synthesis* **1992**, 618.
7. (a) Wang, S.-S.; Gisin, B. F.; Winter, D. P.; Makofske, R.; Kulesha, I. D.; Tzougraki, C.; Meienhofer, J. *J. Org. Chem.* **1977**, *42*, 1286. (b) Holm, B.; Linse, S.; Kihlberg, J. *Tetrahedron* **1998**, *54*, 11995.
8. LR FAB MS: calcd for $[M-1]^-$ 700.21, found 700.2. ^1H NMR (DMSO) δ 8.75 (d, J = 6.4 Hz, 1H), 7.98 (d, J = 7.3 Hz, 2H), 7.80 (d, J = 7.0 Hz, 2H), 7.51 (t, J = 7.0 Hz, 2H), 7.42 (t, J = 7.0 Hz, 2H), 5.25 (t, J = 9.0 Hz, 1H), 5.19 (t, J = 9.5 Hz, 1H), 4.8 (t, J = 9.3 Hz, 1H), 4.37-4.31 (m, 4H), 3.82 (t, J = 9.97 Hz, 1H), 3.76 (m, 4H), 3.1-3.0 (m, 2H), 2.62 (s, 3H), 1.96 (s, 3H), 1.87 (s, 3H), 1.70 (s, 3H). ^{13}C NMR (DMSO) δ 194.3, 170.2, 169.7, 169.4, 169.2, 162.2, 155.7, 143.8, 140.6, 127.6, 127.0, 125.2, 120.0, 77.9, 73.3, 73.0, 70.3, 65.6, 52.1, 50.7, 46.6, 37.5, 35.7, 30.6, 30.3, 29.4, 22.5, 20.4, 20.3.
9. Brief procedure for solid phase glycopeptide synthesis: Fmoc amino acid was manually coupled on PS-PEG Rink amide resin on a 0.5 mmol scale, using 3.0 equiv of amino acid and activation with HBTU (3.0 equiv) and HOBt (3.0 equiv) in the presence of DIEA (3.0 equiv). Incorporation of 6-ThioGlcNAc-Ser **1** (2.0 equiv) into the glycopeptide was carried out by double coupling using DIC (3 equiv), HOAt (3 equiv) and DIEA (3 equiv). After removal of the *N*-terminal Fmoc group of a glycosylated peptide with 20% piperidine in DMF, the resin-bound peptide was treated with 20% Ac_2O in DMF. Peptide cleavage/deprotection was achieved under 95% TFA, 2.5% TIS and 2.5% H_2O conditions. The crude peptide was precipitated with ether and then purified by preparative RP-HPLC with a gradient of 5-100% CH_3CN in water (0.1% TFA) over 30 min.
10. Carpino, L. A.; El-Faham, A. *Tetrahedron* **1999**, *55*, 6813.
11. *Selected data* for **11**: (ESI MS): calcd for $C_{75}H_{71}N_{13}O_{22}S$ $[M^+]$ 1201, found 1201.