Pergamon

SCIENCE DIRECT®

Tetrahedron Letters 44 (2003) 2961–2964

TETRAHEDRON LETTERS

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

Hironobu Hojo,[a,*] Eiichiro Haginoya,[a] Yoshiyuki Matsumoto,[a] Yoshiaki Nakahara,[a] Kazuki Nabeshima,[b] Bryan P. Toole[c] and Yasushi Watanabe[d]

[a]*Department of Applied Biochemistry, Tokai University, Kanagawa, 259-1292, Japan*
[b]*Department of Pathology, Fukuoka University Hospital and School of Medicine, Fukuoka 814-0180, Japan*
[c]*Department of Anatomy and Cellular Biology, Tufts University School of Medicine, Medford, MA 02111, USA*
[d]*National Food Research Institute, Protein Laboratory, Ibaraki 305-8642, Japan*

Received 17 January 2003; revised 7 February 2003; accepted 7 February 2003

Abstract—Peptide thioester carrying *N*-linked core pentasaccharide was prepared by the Fmoc solid-phase method with a combination of the benzyl-protection strategy at the carbohydrate portion. The obtained peptide thioester was successfully used for the synthesis of the first Ig domain of emmprin composed of 61 amino acid residues.

Glycosylation is a common post-translational modification of proteins. Although the importance of glycosylation is well recognized in a wide range of biological processes,[1] the mechanism of how carbohydrates on proteins exert their function is poorly understood at the molecular level. This is mainly due to the microheterogeneity at the carbohydrate portions, which makes the isolation of structurally homogeneous glycoproteins difficult. An expression system using E coli., which lacks glycosylation machinery, is also inadequate to prepare glycoproteins.

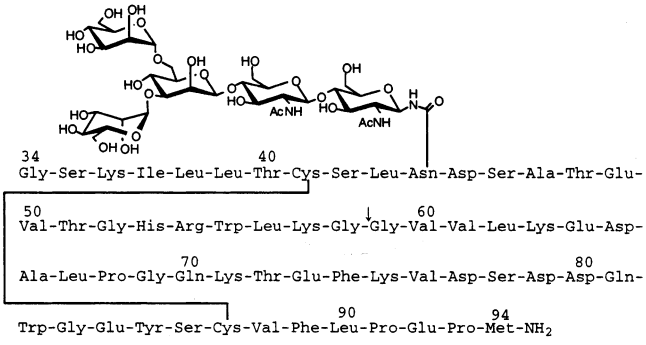

Figure 1. The structure of extracellular Ig domain of emmprin carrying pentasaccharide unit **1**. The arrow indicates the site of segment coupling.

Due to the development of the thioester method[2] as well as the chemical ligation method,[3] chemical synthesis of protein has become practical. In these methods, partially-protected or non-protected peptide thioesters, prepared by the solid-phase method, are consecutively joined to obtain polypeptides. Thus, if the glycosylated peptide thioester can be prepared, these methods can be directly applied to the glycoprotein synthesis. Previously, peptide thioesters have been prepared by the Boc method, because the thioester bond is sensitive to piperidine used for the Fmoc method. In contrast, most of the glycopeptide synthesis has been performed by the Fmoc strategy due to the sensitivity of glycosidic linkages to strong acid such as HF. Recently, this drawback was overcome by development of the method to prepare peptide thioesters by the Fmoc method.[4] In fact, Shin et al. synthesized a peptide thioester carrying a GalNAc residue by the Fmoc method and prepared diptericin by the chemical ligation.[5] However, to analyze the role of carbohydrates on proteins, more complicated oligosaccharide structures have to be introduced. In the case of *N*-linked glycoprotein, at least a pentasaccharide structure, $Man_3GlcNAc_2$, is retained in common.

* Corresponding author. Tel.: +81-463-58-1211 (ext. 4173); fax: +81-463-50-2075; e-mail: hojo@keyaki.cc.u-tokai.ac.jp

0040-4039/03/$ - see front matter
doi:10.1016/S0040-4039(03)00394-0

We have been investigating benzyl protection-based solid-phase synthesis of glycopeptide[6] and already succeeded in the synthesis of dodecapeptide CD52 carrying the pentasaccharide unit by the Fmoc strategy.[6b,c] However, this method uses catalytic hydrogenation at the final stage to remove the benzyl groups and hence may not be directly applied to the synthesis of peptide thioesters, which contain a sulfur atom. Recently, we have shown that the low-TfOH,[7] which cleaves protecting groups under the low acidic S_N2 reaction conditions, is suitable to remove the benzyl groups of glycopeptide carrying core-2 *O*-linked tetrasaccharide.[8] We now examined the use of low-TfOH for the deprotection of the benzyl-protected *N*-linked core pentasaccharide, which contains acid-sensitive β-mannoside bond. In this report, syntheses of peptide thioester carrying an *N*-linked pentasaccharide by the Fmoc strategy and the extracellular first Ig domain (34–94) of emmprin **1** by the thioester method are described (Fig. 1). To our knowledge, this is the first synthesis of peptide thioester carrying the pentasaccharide unit.

Emmprin is a glycoprotein located on the surface of cancer cells and plays an essential role in tumor metastasis by stimulating nearby fibroblasts to produce matrix-metalloproteinases (MMPs).[9] The activity derives from the first Ig domain (34–94) which retains the *N*-glycosylation site at Asn[44]. Although the glycosylation is essential to emmprin's activity, the carbohydrate structure, which is essential to the activity, is not known. We have already prepared the Ig domain carrying $GlcNAc_n$ ($n=0$–2) at Asn[44], utilizing a peptide thioester carrying $GlcNAc_n$ unit prepared by the Boc strategy.[10] The biological study using these domains shows that MMP stimulation activity increases as the carbohydrate chain becomes longer from $n=1$ to 2, whereas the Ig domain itself ($n=0$) has no activity (data not shown). To further examine the effect of the structure of carbohydrate on its activity, the Ig domain with a pentasaccharide unit was synthesized in this study.

The sequence of the Ig domain was divided at Gly[58]–Gly[59], and *N*-terminal peptide thioester carrying the pentasaccharide and *C*-terminal peptide amide were prepared by the solid-phase method. Following the procedure of Li et al.,[4b] the *N*-terminal peptide thioester was prepared using a cocktail composed of 1-methylpyrrolidine (MP, 25%), hexamethyleneimine (HMI, 2%), 1-hydroxybenzotriazole (HOBt, 2%) in 1-methyl-2-pyrrolidinone (NMP)-DMSO (1:1) as Fmoc deblocking reagent. Starting from Fmoc-Gly-SCH_2CH_2CONH-resin, Fmoc-amino acids were introduced using *O*-benzotriazol-1-yl-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HBTU), HOBt and *N*,*N*-diisopropylethylamine (DIEA). After the sequence of (45–58) was assembled, a part of the resin was deprotected by TFA and the crude peptide was analyzed by HPLC. Besides the desired thioester, however, about 40% of a defective peptide, which lacks *C*-terminal Lys-Gly was observed. This peptide is most probably derived from the formation of diketopiperazine, and thus, its prevention is essential to improve the yield of the desired thioester.

Sakamoto et al. reported the suppression of diketopiperazine formation by a combination of silyl carbamate and amino acid fluoride.[11] Although this method was used for the synthesis of peptide with *C*-terminal carboxylic acid, the reaction conditions seem to be applicable to the preparation of peptide thioesters. Emmprin (45–58) was resynthesized as shown in Figure 2.[12] Fmoc-Gly-SCH_2CH_2CONH-resin was treated with the Fmoc deblocking reagent and triisopropylsilyloxycarbonyl (Tsoc)-Lys(Z)-OPfp was reacted. The obtained resin was then reacted with Fmoc-Leu-F in the presence of Bu_4NF. This F^- cleaves a part of the Tsoc groups, and the amino groups generated react with Fmoc-Leu-F, which provides F^- required to continue the reaction. As the *N*-deprotection and acylation steps are combined, diketopiperazine formation is suppressed. The chain elongation was continued using Li's method. Analysis of the cleavage mixture of the resin sample of emmprin (45–58) showed that the desired peptide was obtained without detectable diketopiperazine formation.

Then, Fmoc-Asn($Man_3GlcNAc_2Bn_{12}$) **2**,[13] was introduced using *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU) and DIEA. The coupling yield was about 60% judging from the HPLC analysis of the cleavage product of the resin sample. It is to be noted that the coupling reaction did not proceed by the use of the 1,3-dicyclohexylcarbodiimide (DCC)–HOBt method. The remaining amino group was protected using *N*-(2-chlorobenzyloxycarbonyloxy)succinimide (Z(2-Cl)-OSu), and the chain elongation was continued using Fmoc-amino acid Pfp ester. After the complete assembly of the sequence (34–58), the whole resin was treated with reagent K[14] at

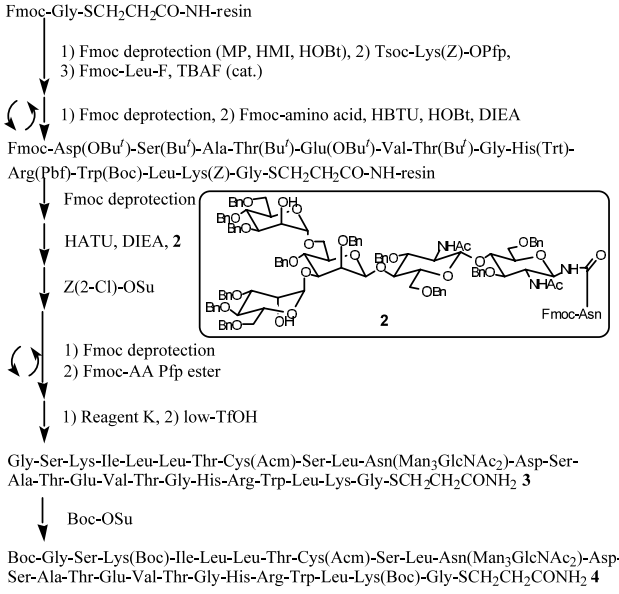

Figure 2. Synthetic procedure of the *N*-terminal peptide thioester carrying pentasaccharide unit **4**.

rt for 1 h. The crude peptide was further treated by the low-TfOH[7] at −10°C for 2 h to remove all benzyl groups of the saccharide portion. As shown in Figure 3, the desired peptide thioester carrying pentasaccharide unit **3** was successfully obtained without significant decomposition of the glycosidic linkages. The purified yield of peptide **3** was 1.8% based on the amino groups on the initial resin. The amino groups of the peptide **3** were protected by the Boc groups and peptide **4** was obtained. *C*-terminal segment **5** was prepared according to the previous procedure.[10]

The condensation reaction was carried out as shown in Figure 4.[15] Peptide **4** and **5** were dissolved in DMSO containing 3-hydroxy-3,4-dihydro-4-oxo-1,2,3-benzotriazine (HOOBt) and DIEA, and AgCl was added. The reaction was almost complete without significant side reactions within 12 h (Fig. 5a). Then the product was treated with TFA and purified by HPLC to obtain peptide **6**. This peptide was further treated with $AgNO_3$ and DIEA to remove Acm groups. After HPLC purification, the peptide was air-oxidized in 0.1 M $AcONH_4$ (pH 8) containing 6 M guanidine hydrochloride as shown in Figure 5c. The Ig domain carrying pentasaccharide **1** was successfully obtained after HPLC purification. The analytical data of the synthesized **1** were in good agreement with those of the desired structure.

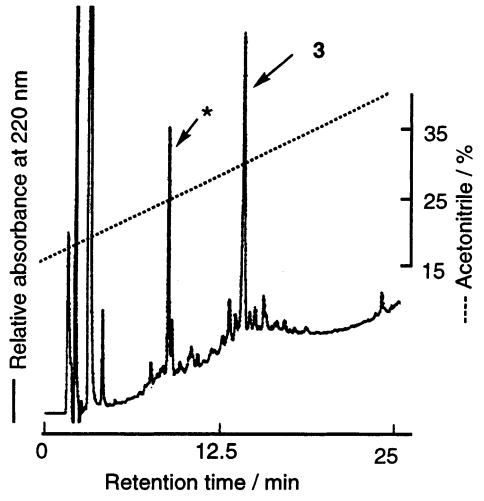

Figure 3. HPLC profile of the crude peptide thioester **3**. The peptide indicated by asterisk derived from the incomplete introduction of pantasaccharide-Asn unit and subsequent Z(2-Cl)-capping. Elution conditions: column, Mightysil RP-18 GP (4.6×150 mm) at a flow rate of 1 ml min^{-1}; eluent, aqueous acetonitrile containing 0.1% TFA.

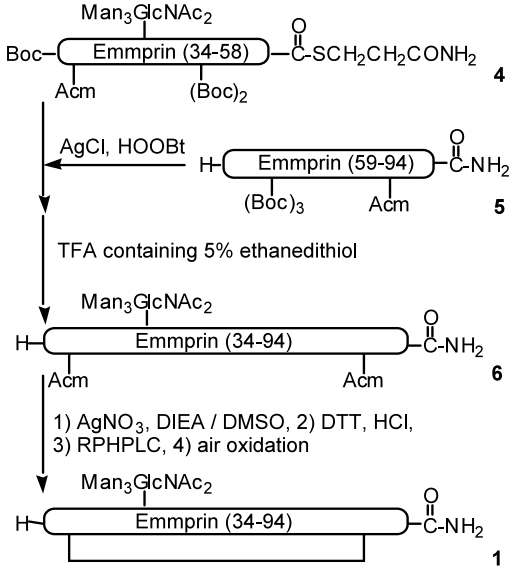

Figure 4. Synthetic procedure for the Ig domain (34–94) **1**.

In conclusion, we have succeeded in the synthesis of a peptide thioester containing the *N*-linked core pentasaccharide by the modified Li's Fmoc method combining benzyl protection strategy. This peptide thioester was successfully used for the synthesis of the Ig domain of emmprin. This method would be applicable for the synthesis of peptide thioesters carrying the *N*-linked core pentasaccharide and thus would be useful for the preparation of larger *N*-glycopeptides or *N*-glycoproteins. The MMP-stimulatory activity measurement as well as conformational analysis of this domain are now under investigation.

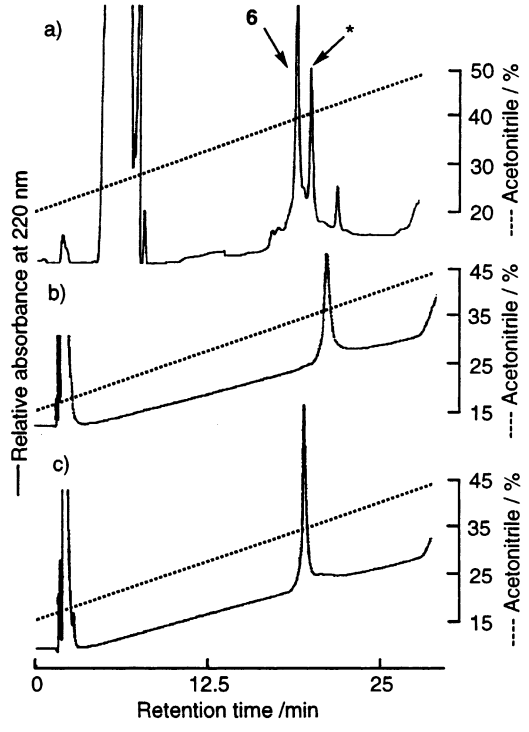

Figure 5. HPLC profile of the crude peptides: a) TFA treated reaction mixture of the segment coupling of **4** and **5**; b) crude peptide obtained after removal of Acm groups from peptide **6**; c) crude peptide **1**. The asterisked peak in panel a) derived from the unreacted peptide **5**. Elution conditions of a): column, Cosmosil 5C18AR (10×250 mm) at a flow rate of 2.5 ml min^{-1} and at 60°C; b) and c): same as Figure 3.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan [(B) No. 13460051] and in part by CREST program of Japan Science and Technology Corporation. The authors also acknowledge Tokai University for a grant-aid for high-technology research.

References

1. Varki, A. *Glycobiology* **1993**, *3*, 97–130.
2. (a) Hojo, H.; Aimoto, S. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111–117; (b) Aimoto, S. *Biopolymers (Peptide Science)* **1999**, *51*, 247–265.
3. (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779; (b) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
4. (a) Futaki, S.; Sogawa, K.; Maruyama, J.; Asahara, T.; Niwa, M.; Hojo, H. *Tetrahedron Lett.* **1997**, *38*, 6237–6240; (b) Li, X.; Kawakami, T.; Aimoto, S. *Tetrahedron Lett.* **1998**, *39*, 8669–8672; (c) Alsina, J.; Yokumu, T. S.; Albericio, F.; Barany, G. *J. Org. Chem.* **1999**, *64*, 8761–8769; (d) Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. *J. Am. Chem. Soc.* **1999**, *121*, 11369–11374; (e) Clippingdale, A. B.; Barrow, C. J.; Wade, J. D. *J. Peptide Sci.* **2000**, *6*, 225–234; (f) Swinnen, D.; Hilvert, D. *Org. Lett.* **2000**, *2*, 2439–2442; (g) Ishii, A.; Hojo, H.; Nakahara, Y.; Ito, Y.; Nakahara, Y. *Biosci. Biotechnol. Biochem.* **2002**, *66*, 225–232.
5. Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 11684–11689.
6. (a) Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Carbohydr. Res.* **1996**, *292*, 71–81; (b) Guo, Z.-W.; Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1464–1466; (c) Guo, Z.-W.; Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Bioorg. Med. Chem.* **1997**, *5*, 1917–1924; (d) Nakahara, Y.; Nakahara, Y.; Ito, Y.; Ogawa, T. *Carbohydr. Res.* **1998**, *309*, 287–296.
7. Tam, J. P.; Heath, W. F.; Merrifield, R. B. *J. Am. Chem. Soc.* **1986**, *108*, 5242–5251.
8. Takano, Y.; Habiro, M.; Someya, M.; Hojo, H.; Nakahara, Y. *Tetrahedron Lett.* **2002**, *43*, 8395–8399.
9. Nabeshima, K.; Kataoka, H.; Koono, M.; Toole, B. P. In *Collagenases*; Hoeffler, W., Ed.; R.G. Landes Company: Texas, 1999; pp. 91–113.
10. Hojo, H.; Watabe, J.; Nakahara, Y.; Nakahara, Y.; Ito, Y.; Nabeshima, K.; Toole, B. P. *Tetrahedron Lett.* **2001**, *42*, 3001–3004.
11. Sakamoto, K.; Nakahara, Y.; Ito, Y. *Tetrahedron Lett.* **2002**, *43*, 1515–1518.
12. Fmoc-Rink amide MBHA resin (0.137 g, 0.1 mmol) was treated with the Fmoc deblocking reagent, and was reacted with Fmoc-Gly-SCH_2CH_2COOBt (0.22 mmol). The Fmoc group was removed and Tsoc-Lys(Z)-OPfp (129 mg, 0.2 mmol) in THF was reacted for 15 min. The coupling reaction was repeated. After washing with NMP and dichloromethane (DCM), Fmoc-Leu-F (85 mg, 0.24 mmol) dissolved in DCM was added to the resin. Then, 1 M Bu_4NF in THF (10 μl, 0.01 mmol) was added and the resin was stirred for 1 h. The resin was subjected to the automated synthesis by ABI 433A peptide synthesizer using FastMoc protocol. The Fmoc deprotection protocol was modified so that the premixed reagent was introduced to the reaction vessel without dilution. The deprotection time was 3 and 15 min. After the synthesis of (45–58) was completed, half of the resin was removed. Then, Fmoc-Asn($Man_3GlcNAc_2Bn_{12}$) (259 mg, 0.11 mmol), HATU (42 mg, 0.11 mmol) and DIEA (38 μl, 0.22 mmol) was added and the resin was mixed for 30 min. DIEA (20 μl, 0.12 mmol) was further added and the resin was mixed for 1 h at 50°C. The unreacted amino groups were capped using Z(2-Cl)-OSu (142 mg, 0.5 mmol) in the presence of DIEA (0.17 ml, 1 mmol) for 20 min. The remaining sequence was introduced using Fmoc amino acid Pfp ester (0.25 mmol) and 197 mg of the protected peptide resin was obtained. The whole resin was treated with reagent K (1.9 ml) at room temperature for 1 h. The peptide precipitated with ether was further treated with low TfOH [TfOH: TFA: dimethylsulfide: *m*-cresol (1:5:3:1)] (1.9 ml) at −10°C for 2 h. The product was precipitated by ether and purified by HPLC to give peptide **3** (0.88 μmol, 1.8% yield based on the amino groups on the initial resin). Peptide **3** (1.1 μmol) was dissolved in DMSO (32 μl) containing DIEA (1.7 μl, 9.7 μmol) and Boc-OSu (2.1 mg, 9.8 μmol) was added. After the reaction mixture was stood for 1 h, the product was precipitated by ether and lyophilized from dioxane suspension to give peptide **4** (1.1 μmol). MALDI-TOF mass found: m/z 4039.2. Calcd for $(M+H)^+$ 4039.5 (average). Amino acid analysis: $Asp_{1.92}Thr_{2.72}Ser_{2.70}Glu_{1.06}Gly_{2.84}Ala_1Val_{1.60}Ile_{0.86}Leu_{3.99}Lys_{2.03}His_{0.92}Arg_{0.96}$.
13. Matsuo, I.; Nakahara, Y.; Ito, Y.; Nukada, T.; Nakahara, Y.; Ogawa, T. *Bioorg. Med. Chem.* **1995**, *3*, 1455–1463.
14. King, D. S.; Fields, C. G.; Fields, G. B. *Int. J. Peptide Protein Res.* **1990**, *36*, 255–266.
15. Peptide **4** (1.1 μmol) and **5** (2.2 μmol) were dissolved in DMSO (100 μl) containing HOOBt (7.8 mg, 48 μmol) and DIEA (5.6 μl, 32 μmol). AgCl (0.7 mg, 4.9 μmol) was added and the solution was stirred overnight in the dark. The product was precipitated by ethyl acetate and treated with 10% ethanedithiol-TFA (180 μl) for 10 min. The peptide, precipitated with ether, was extracted with 50% aqueous acetonitrile to remove silver salt. After lyophilization, the peptide was dissolved in DMSO (100 μl) containing DIEA (3.9 μl, 22 μmol) and $AgNO_3$ (1.9 mg, 11 μmol) was added. The solution was stood for 1 h in the dark. Dithiothreitol (17 mg, 110 μmol) and 0.5 M HCl (220 μl) was added and the peptide was purified by HPLC to obtain a reduced form of the Ig domain (559 nmol). The peptide was dissolved in 10 ml of 0.1 M $AcONH_4$ (pH 8) containing 6 M guanidine HCl and stirred for 2 days. The product was purified by HPLC to obtain the desired peptide **1** (353 nmol). MALDI-TOF mass found: m/z 7618.9; calcd for $[M+H]^+$: 7618.4 (average). Amino acid analysis: $Asp_{6.25}Thr_{3.56}Ser_{4.06}Glu_{7.28}Pro_{3.16}Gly_{5.76}Ala_2Val_{4.63}Met_{0.95}Ile_{0.86}Leu_{7.59}Tyr_{1.02}Phe_{1.97}Lys_{5.10}His_{1.02}Arg_{0.97}$.