Pergamon

Tetrahedron Letters 42 (2001) 3001–3004

TETRAHEDRON LETTERS

Synthesis of the extracellular Ig domain I of Emmprin carrying a chitobiose unit

Hironobu Hojo,[a,b,*] Jun Watabe,[a] Yoshiaki Nakahara,[a,b,c] Yuko Nakahara,[b,c] Yukishige Ito,[b,c] Kazuki Nabeshima[d] and Bryan P. Toole[e]

[a]*Department of Industrial Chemistry, Tokai University, Kanagawa 259-1292, Japan*
[b]*The Institute of Physical and Chemical Research (RIKEN), Saitama 351-0198, Japan*
[c]*CREST, Japan Science and Technology Corporation (JST), Japan*
[d]*Department of Pathology, Miyazaki Medical College, Miyazaki 889-1692, Japan*
[e]*Department of Anatomy and Cellular Biology, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111, USA*

Received 15 January 2001; revised 19 February 2001; accepted 23 February 2001

Abstract—The extracellular Ig domain I of Emmprin (34-94) carrying a chitobiose unit at Asn[44] was chemically synthesized. Boc-Asn with a benzyl-protected chitobiose unit was synthesized and used for the preparation of peptide thioester with the sequence of Emmprin (34-58) by Boc strategy. C-Terminal peptide amide (59-94) was also prepared by the solid-phase method. These segments were condensed by activation of the thioester group by silver ions to obtain a protected form of Emmprin (34-94)-NH_2. After deprotection and air oxidation, the desired Emmprin (34-94)-NH_2 with chitobiose was successfully obtained.
© 2001 Elsevier Science Ltd. All rights reserved.

Emmprin is a glycoprotein located on the surface of human tumor cells.[1] It belongs to a member of the immunogloblin superfamily and stimulates nearby fibroblasts to produce matrix metalloproteinases. Several experiments have shown that the extracellular immunoglobulin (Ig) domain I (34-94), which possesses an *N*-glycosylation site at Asn[44], plays an essential role in metalloproteinase stimulatory activity (Fig. 1). To analyze the biological function of Emmprin at a molecular level, therefore, the Ig domain with a distinct carbohydrate structure at Asn[44] has to be synthesized. We have previously established synthetic routes for glycosylated amino acids, such as Ser(Neu5Ac-Gal-GalNAc) and Asn(Man_3-GlcNAc-GlcNAc) derivatives, which were used for the syntheses of several glycopeptides by the Fmoc solid-phase method.[2] In these syntheses, construction of the oligosaccharide portions was realized using benzyl-protected sugars, which retain higher reactivity than the generally used acyl-protected sugars. Thus, the deprotection of the carbohydrate portion was carried out at the final stage of the synthesis by catalytic hydrogenation or 1 M trimethylsilyl

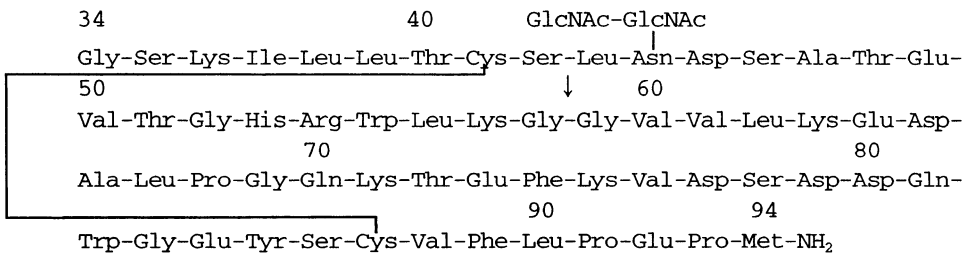

Figure 1. The structure of the first Ig domain (34-94) of Emmprin carrying a chitobiose unit at Asn[44]. The arrow indicates the site of segment coupling.

Keywords: thioesters; glycopeptides; condensations.
* Corresponding author. Fax: +81-463-50-2075; e-mail: hojo@keyaki.cc.u-tokai.ac.jp

0040-4039/01/$ - see front matter © 2001 Elsevier Science Ltd. All rights reserved.
PII: S0040-4039(01)00342-2

triflate in trifluoroacetic acid (TFA). These treatments are advantageous over the base treatment required for the removal of acyl-protecting groups in that they are free of the danger of β-elimination and racemization. On the other hand, benzyl groups are readily removed by HF. In the case of glycopeptides with acid-stable carbohydrate chains, the Boc strategy would be a better choice, as the peptide and carbohydrate portion can be simultaneously deprotected by HF treatment. However, few reports have appeared on glycopeptide synthesis by the Boc strategy. In addition, all of them are on glycopeptides carrying a monosaccharide.[3] To examine the applicability of the Boc strategy to the synthesis of glycopeptide with a longer carbohydrate chain, a Boc-Asn derivative carrying a benzyl-protected chitobiose was prepared. Based on the result of HF treatment of natural glycoproteins, the chitobiose on Asn seems to be stable to this reagent.[4] This unit was used for the synthesis of a peptide thioester with a sequence of Emmprin (34-58) and then condensed with the Emmprin (59-94) segment to obtain the Ig domain I by the thioester method.[5]

Boc-Asn with a chitobiose unit was prepared as shown in Fig. 2. Starting from a tetra-*O*-acetyl derivative of glucosamine **1**, the fluoride **3** was synthesized. This compound was then coupled with another glucosamine unit to give disaccharide **4**. After removal of the phthaloyl group and acetylation, compound **5** was obtained. Then the azido group of **5** was selectively reduced in the presence of benzyl groups by catalytic hydrogenation using ethylenediamine poisoned Pd/C.[6] The obtained amine was directly used for the coupling with Boc-Asp-OAll by 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (WSCI) without further purification. Removal of the allyl group by $Pd(PPh_3)_4$ and *N*-methylaniline gave the desired product **7**.[7]

The glycosylated Asn **7** was then used for the preparation of Emmprin (34-94) by the thioester method.[5] Emmprin (34-94) was divided at Gly^{58}-Gly^{59} and the solid-phase synthesis of an N-terminal peptide thioester with a chitobiose unit and C-terminal peptide was carried out. The N-terminal segment was synthesized as shown in Fig. 3. Starting from Boc-Gly-SCH_2CH_2CO-

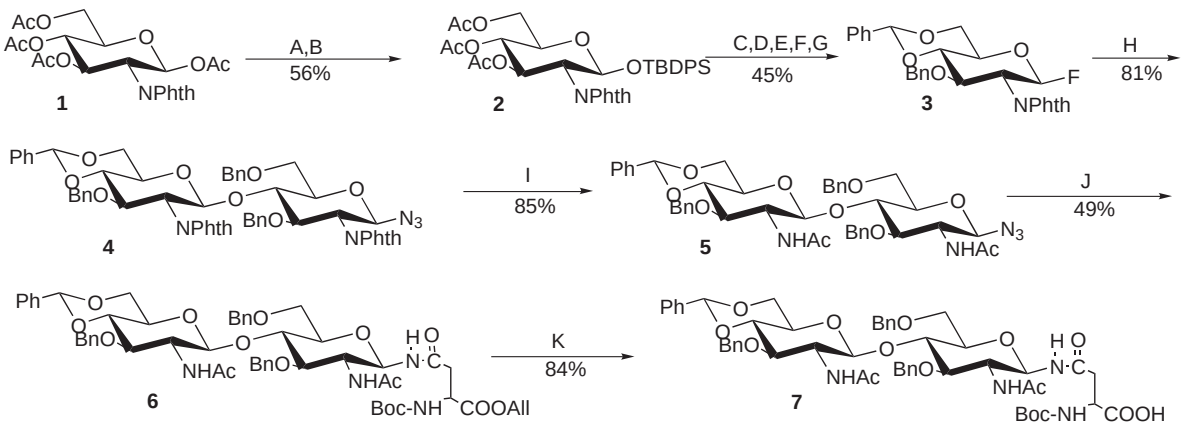

Figure 2. Synthetic route for chitobiose-linked Asn unit **7**. (A) NH_2NH_2–AcOH/DMF. (B) *t*-Butyldiphenylsilyl chloride, imidazole, 4-dimethylaminopyridine/DMF. (C) *t*-BuOK/MeOH, THF. (D) Benzaldehyde dimethylacetal, camphorsulfonic acid/CH_3CN. (E) Benzyl bromide, NaH/DMF. (F) Tetrabutylammonium fluoride, AcOH/THF. (G) Dimethylaminosulfur trifluoride/THF. (H) 3,6-Di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl azide, Cp_2HfCl_2, $AgClO_4/CH_2Cl_2$. (I) 1: Ethylenediamine/BuOH. 2: $(CH_3CO)_2O$/MeOH. (J) 1: H_2, Pd/C(en)/CH_2Cl_2, MeOH. 2: Boc-Asp-OAll, WSCI/CH_2Cl_2. (K) $Pd(PPh_3)_4$, *N*-methylaniline/THF.

Boc-Gly-SCH_2CH_2CO-Leu-OCH_2-Pam-resin

↓ ABI433A Peptide synthesizer system software 2.03; Boc/HOBt/DCC 0.5 mmol protocol without acetylation cycle

Boc-Gly-Ser(Bn)-Lys(ClZ)-Ile-Leu-Leu-Thr(Bn)-Cys(Acm)-Ser(Bn)-Leu-Asn[GlcNAc(Bn)-GlcNAc(Bn_2)]-Asp(OBn)-Ser(Bn)-Ala-Thr(Bn)-Glu(OBn)-Val-Thr(OBn)-Gly-His(Bom)-Arg(Mts)-Trp(Mts)-Leu-Lys(ClZ)-Gly-SCH_2CH_2CO-Leu-OCH_2-Pam-resin (**8**)

↓ 1) anhydrous HF containing 10% 1,4-butanedithiol and 10% anisole, 2) RPHPLC

Gly-Ser-Lys-Ile-Leu-Leu-Thr-Cys(Acm)-Ser-Leu-Asn(R)-Asp-Ser-Ala-Thr-Glu-Val-Thr-Gly-His-Arg-Trp-Leu-Lys-Gly-SCH_2CH_2CO-Leu (**9**, R=GlcNAc-GlcNAc; **10**, R=GlcNAc)

↓ Boc-OSu

Boc-Gly-Ser-Lys(Boc)-Ile-Leu-Leu-Thr-Cys(Acm)-Ser-Leu-Asn(R)-Asp-Ser-Ala-Thr-Glu-Val-Thr-Gly-His-Arg-Trp-Leu-Lys(Boc)-Gly-SCH_2CH_2CO-Leu (**11**, R=GlcNAc-GlcNAc; **12**, R=GlcNAc)

Figure 3. Synthetic route for peptides **11** and **12**.

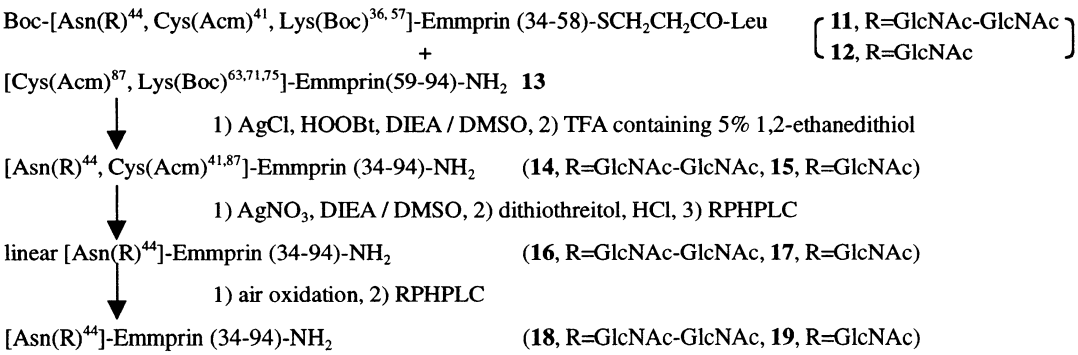

Figure 4. Synthetic route for peptide [Asn(R)[44]]-Emmprin (34-94)-NH_2.

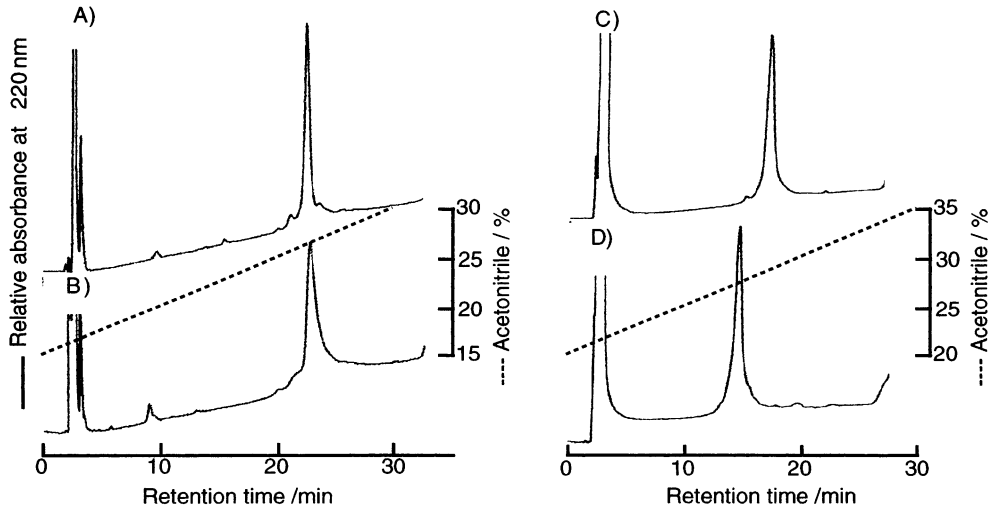

Figure 5. HPLC elution profiles of the peptides; (A) crude peptide **14**, (B) crude peptide **16**, (C) purified peptide **16**, (D) crude peptide **18**. Elution conditions: column, Mightysil RP-4 GP (4.6×150 mm) at a flow rate of 1 ml min^{-1}; eluent, aqueous acetonitrile containing 0.1% TFA.

Leu-OCH_2-Pam-resin, the sequence of Emmprin (34-58) was synthesized by an ABI 433A peptide synthesizer (Applied Biosystems Inc., Foster City, CA) using system software version 2.03. The capping module 'g' of this protocol was omitted, as the TFA treatment after coupling of compound **7** would remove the benzylidene group of the chitobiose moiety, which causes the acetylation of 4,6-hydroxyl groups during the capping cycle. The side chain functional groups of Arg and Trp were protected by the Mts group to achieve the final deprotection using both HF and 1 M triflic acid (TfOH)–TFA. The chitobiose-linked Asn unit **7** was introduced manually using the 1,3-dicyclohexylcarbodiimide (DCC)-1-hydroxybenzotriazole (HOBt) method. After completion of the peptide chain assembly, the protected peptide resin was treated with anhydrous HF or 1 M TfOH–TFA. Both treatments successfully gave the desired peptide **9**. The yield of peptide **9** by HF and 1 M TfOH–TFA treatment was 12 and 14%, respectively, based on the compound **7** used for the preparation of resin **8**. However, the glycosidic linkage between GlcNAc-GlcNAc seems to be not completely stable to these treatments, since GlcNAc-truncated peptide **10** was also obtained in 7.5 and 6% yields. To reduce this side reaction, further optimization of the deprotection conditions is required. Peptide **10** was used for Emmprin (34-94)-NH_2 with one GlcNAc unit.

The amino groups of peptides **9** and **10** were protected with Boc groups using *N*-(*t*-butoxycarbonyloxy)-succinimide (Boc-OSu) to obtain peptides **11** and **12** in yields of 73 and 92%, respectively.[7] The C-terminal segment Gly-Val-Val-Leu-Lys(Boc)-Glu-Asp-Ala-Leu-Pro-Gly-Gln-Lys(Boc)-Thr-Glu-Phe-Lys(Boc)-Val-Asp-Ser-Asp-Asp-Gln-Trp-Gly-Glu-Tyr-Ser-Cys(Acm)-Val-Phe-Leu-Pro-Glu-Pro-Met-NH_2 (**13**) was prepared via a peptide segment synthesized on Fmoc-Rink amide MBHA-resin (Calbiochem-Novabiochem, Switzerland) using *FastMoc* protocol. The yield was 8.9% based on the content of the amino groups on the initial resin.

The condensation of the segments was carried out following the procedure of Kawakami et al. as shown in Fig. 4.[5b] For example, the preparation of peptide **18** was briefly described. Peptides **11** and **13** were dissolved in DMSO containing 3,4-dihydro-4-oxo-3-hydroxy-1,2,3-benzotriazine (HOOBt) and *N*,*N*-diisopropylethylamine (DIEA). Then, AgCl was added to initiate the coupling reaction. As shown in Fig. 5, the reaction was almost completed within 4 h without significant side

reactions. The product was successively treated with TFA and $AgNO_3$ in the presence of DIEA to remove all protecting groups. The crude peptide was purified by HPLC to obtain the linear [Asn(GlcNAc-GlcNAc)[44]]-Emmprin (34-94)-NH_2 **16** in a yield of 77%. To form a disulfide bond between Cys^{41} and Cys^{87}, the peptide was dissolved in DMSO and the solution was diluted by 1% $AcONH_4$ (pH 7.8) to 20% DMSO solution. After the reaction mixture was gently stirred in the presence of air overnight, the final product **18** was obtained as shown in Fig. 5D. The yield of peptide **18** was 50% from **16**. In the same manner, peptide **19** was obtained in 27% yield based on peptide **13** used for segment coupling. In conclusion, the Boc-Asn derivative with a benzyl-protected chitobiose moiety was prepared and successfully applied to the synthesis of the Ig domain I of Emmprin. The collagenase stimulation activity of the peptides **18** and **19** is now being studied.

References

1. Nabeshima, K.; Kataoka, H.; Koono, M.; Toole, B. P. In *Collagenases*; Hoeffler, W., Ed.; R.G. Landes Company: Texas, 1999; pp. 91–113.
2. (a) Nakahara, Y.; Nakahara, Y.; Ito, Y.; Ogawa, T. *Carbohydr. Res.* **1998**, *309*, 287–296; (b) Guo, Z.-W.; Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Bioorg. Med. Chem.* **1997**, *5*, 1917–1924.
3. (a) Lavielle, S.; Ling, N. C.; Guillemin, R. C. *Carbohydr. Res.* **1981**, *89*, 221–228; (b) Mizuno, M.; Haneda, K.; Iguchi, R.; Muramoto, I.; Kawakami, T.; Aimoto, S.; Yamamoto, K.; Inazu, T. *J. Am. Chem. Soc.* **1999**, *121*, 284–290; (c) Teshima, T.; Urabe, T.; Yamamoto, T.; Isaka, S.; Kumagaye, K.; Nakajima, K.; Shiba, T. In *Peptide Science 1998*; Kondo, M., Ed.; Protein Research Foundation: Osaka, 1999; pp. 157–160.
4. Mort, A. J.; Lamport, D. T. A. *Anal. Biochem.* **1977**, *82*, 289–309.
5. (a) Hojo, H.; Yoshimura, S.; Go, M.; Aimoto, S. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 330–336; (b) Kawakami, T.; Yoshimura, S.; Aimoto, S. *Tetrahedron Lett.* **1998**, *39*, 7901–7904; (c) Aimoto, S. *Biopolymers (Peptide Science)* **1999**, *51*, 247–265.
6. Sajiki, H.; Hattori, K.; Hirota, K. *J. Org. Chem.* **1998**, *63*, 7990–7992.
7. Selected physical data are given below.
5: Calcd for $C_{44}H_{49}N_5O_{10}$: C, 65.41; H, 6.11; N, 8.67. Found: C, 65.20; H, 6.11; N, 8.59. ^{1}H-NMR: 6.13 (d, 1H, *J* 7.9 Hz, NHAc), 5.53 (s, 1H, PhCH(O)$_2$), 4.88 (d, 1H, *J* 7.0 Hz, H-1a), 4.87 (d, 1H, *J* 12.0 Hz, PhCH$_2$), 4.78 (d, 1H, *J* 8.6 Hz, NHAc), 4.70 (d, 1H, *J* 11.9 Hz, PhCH$_2$), 4.65 (d, 2H, *J* 11.2 Hz, PhCH$_2$), 4.63 (d, 1H, *J* 12.2 Hz, PhCH$_2$), 4.47 (d, 1H, *J* 11.7 Hz, PhCH$_2$), 4.41 (d, 1H, *J* 8.3 Hz, H-1b), 4.22 (dd, 1H, *J* 5.0, 10.6 Hz, H-6b), 3.97 (brt, 1H, *J* 6.0 Hz, H-3b), 3.29–3.25 (m, 1H, H-5b), 1.94 (s, 3H, Ac), 1.80 (s, 3H, Ac). **7**: Calcd for $C_{53}H_{64}N_4O_{15}\cdot H_2O$: C, 62.71; H, 6.45; N, 5.52. Found: C, 62.43; H, 6.44; N, 5.66. ^{1}H NMR: 5.52 (s, 1H, PhCH(O)$_2$), 4.88–4.79 (m, 3H, PhCH$_2$×2, H-1a), 4.50–4.43 (m, 3H, PhCH$_2$, H-1b, Asn β-H), 1.89 (s, 3H, Ac), 1.84 (s, 3H, Ac), 1.43 (s, 9H, *t*-Bu). **11**: MALDI-TOF MS: found: *m*/*z* 3665.2 $(M+H)^+$, calcd for $(M+H)^+$: 3664.8. Amino acid analysis: $Asp_{2.01}Thr_{2.83}Ser_{2.68}Glu_{1.04}Gly_3Ala_{1.03}Val_{1.25}1/2Cys_{0.14}Ile_{0.86}Leu_{4.98}Lys_{1.96}His_{1.02}Arg_{1.00}$. **12**: MALDI-TOF MS: found: *m*/*z* 3461.8 $(M+H)^+$, calcd for $(M+H)^+$: 3461.8. Amino acid analysis: $Asp_{1.92}Thr_{2.69}Ser_{2.59}Glu_{1.00}Gly_3Ala_{1.01}Val_{1.27}1/2Cys_{0.14}Ile_{0.89}Leu_{4.98}Lys_{1.94}His_{0.97}Arg_{0.99}$. **13**: MALDI-TOF MS: found: *m*/*z* 4453.6 $(M+Na)^+$ (average), calcd for $(M+Na)^+$: 4452.0 (average). Amino acid analysis: $Asp_{4.05}Thr_{0.94}Ser_{1.75}Glu_{6.21}Pro_{3.03}Gly_3Ala_{1.02}Val_{2.92}1/2Cys_{0.06}Met_{0.82}Leu_{3.01}Tyr_{0.88}Phe_{1.92}Lys_{2.97}$. **18**: MALDI-TOF MS: found: *m*/*z* 7131.9 $(M+H)^+$ (average), calcd for $(M+H)^+$: 7132.0 (average). Amino acid analysis: $Asp_{5.99}Thr_{3.74}Ser_{4.47}Glu_{7.33}Pro_{3.13}Gly_6Ala_{2.10}Val_{4.51}1/2Cys_{0.52}Met_{0.84}Ile_{0.90}Leu_{7.52}Tyr_{1.07}Phe_{1.92}Trp_{0.74}Lys_{4.84}His_{1.02}\ Arg_{0.98}$. **19**: MALDI-TOF MS: found: *m*/*z* 6929.0 $(M+H)^+$ (average), calcd for $(M+H)^+$: 6928.8. Amino acid analysis: $Asp_{6.01}Thr_{3.79}Ser_{4.49}Glu_{7.39}Pro_{3.25}Gly_6Ala_{2.07}Val_{4.37}1/2Cys_{0.75}Met_{0.84}Ile_{0.92}Leu_{7.58}Tyr_{1.09}Phe_{1.94}Trp_{0.79}Lys_{4.88}His_{1.03}Arg_{1.04}$.