Pergamon

Tetrahedron Letters 43 (2002) 3297–3300

TETRAHEDRON LETTERS

Synthesis of mucin-type glycopeptide (β hCG 130–145) by on-resin fragment condensation of the glycopeptide segments carrying unmasked oligosaccharides

Tsuyoshi Ichiyanagi,[a,b,*] Maki Takatani,[a,b] Kimitoshi Sakamoto,[a,b] Yuko Nakahara,[a,b] Yukishige Ito,[a,b] Hironobu Hojo[b,c] and Yoshiaki Nakahara[a,b,c,*]

[a]*RIKEN (The Institute of Physical and Chemical Research), Hirosawa 2-1, Wako-shi, Saitama 351-0198, Japan*
[b]*CREST, Japan Science and Technology Corporation (JST), Tokyo, Japan*
[c]*Department of Applied Biochemistry, Tokai University, Kitakaname 1117, Hiratsuka-shi, Kanagawa 259-1292, Japan*

Received 6 February 2002; revised 11 March 2002; accepted 15 March 2002

Abstract—Use of amino acid pentafluorophenyl esters was found effective for chemoselective *N*-acylation in the peptide elongation. On-resin fragment condensation between **9** and **13** using DCC-PfpOH as an activator afforded coupling product **1** in good yield.

The carbohydrate moieties of glycopeptides have several important functions. For example, they provide protection against proteolysis, influence transport and uptake of proteins, determine human blood groups, and regulate recruitment of leukocytes to the sites of inflammation.[1] To enhance the understanding of the biological functions of glycoproteins, extensive efforts have recently been devoted to the chemical synthesis of glycopeptides.[2]

In the synthesis of glycopeptide, the use of a protective group is crucial in order to obtain a successful outcome.[3–6] Current syntheses of glycopeptides have been preferably performed on solid-phase according to the Fmoc (9-fluorenylmethoxycarbonyl) protocol,[7,8] where *O*-acyl and *O*-benzyl groups have been extensively used for the protection of carbohydrate hydroxyl groups.[9] When *O*-acyl groups are employed, the ultimate deprotection requires basic conditions, which may give grave concern in terms of the potential risks of base-catalyzed β-elimination of serine/threonine-linked oligosaccharide and epimerization of amino acid residues.[10–13] On the other hand, benzyl ether groups are cleavable by neutral hydrogenolysis[14–17] or by a combination of a strong acid and a soft nucleophile.[18] The catalytic hydrogenation conditions, however, are not useful in deprotection of the compounds containing tyrosine, cysteine, or methionine residues owing to liability to overreduction of the tyrosine aromatic ring to the corresponding cyclohexanone,[19] or desulfurization of cysteine or methionine.[17] We have recently established synthesis of a glycopeptide, the extracellular Ig domain I of Emmprin, where detachment from the resin and concomitant deprotection of a benzylated chitobiosyl asparagine-containing glycopeptide were successfully performed by hydrogen fluoride treatment.[20] It is to be noted that use of weak acid-labile protective groups such as silyl, methoxybenzyl or acetal have also been attempted in the Fmoc glycopeptide syntheses.[21]

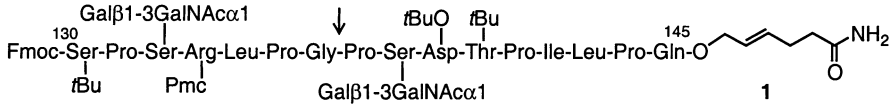


Figure 1. The structure of β-subunit (130–145) of hCG carrying unit at Ser.[132,138] The arrow indicates the site of segment coupling.

Keywords: glycopeptides; condensations; pentafluorophenyl esters.

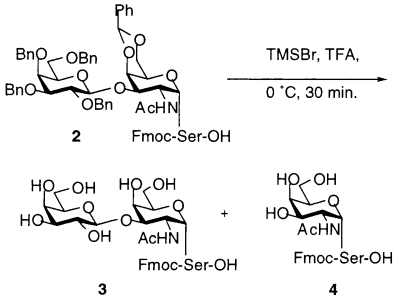
* Corresponding authors. Fax: +81-48-462-4680 (T.I.); +81-463-50-2075 (Y.N.); e-mail: yanagi@postman.riken.go.jp; yonak@keyaki.cc.u-tokai.ac.jp

0040-4039/02/$ - see front matter
PII: S0040-4039(02)00542-7

In the course of our project directed toward the synthesis of larger glycopeptides, we are interested in an alternative approach with minimization of the hydroxyl group protection. Recently, some synthetic methods for glycopeptide synthesis carrying unmasked oligosaccharides research groups were reported.[22] In this paper, we now present a solid-phase synthesis of a mucin-type glycopeptide segment of the β-subunit of human chorionic gonadotropin **1** using a β-D-galactopyranosyl-(1→3)–2 - acetamido - 2 - deoxy - α - D - galactopyranosyl (Thomsen–Friedenreich antigen, T-antigen) serine building block carrying unmasked hydroxyl groups (Fig. 1).

The target compound **1** was retrosynthetically disconnected into two glycopeptide segments, Ser^{130}-Gly^{136} and Pro^{137}-Gln^{145}. Both were synthesized on the allyl linker[9] attached to the Sieber amide resin, and segment coupling was investigated.

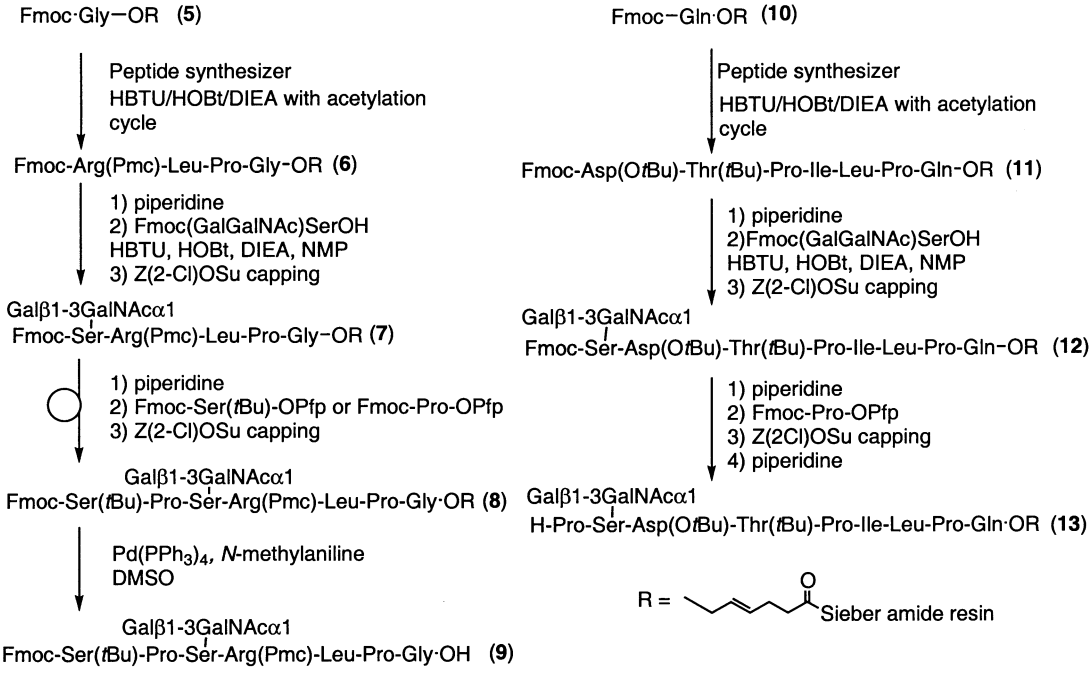
Synthesis of the unprotected-disaccharyl serine **3** is depicted in Scheme 1. Treatment of known glycoserine **2**[14] with bromotrimethylsilane and thioanisole in trifluoroacetic acid at 0°C led to deprotection of the hydroxyl groups to afford glycoserine **3** (65%). The reaction was accompanied in part by degalactosylation to give compound **4** (14%).



Scheme 1. Synthetic route for disaccharide-linked serine unit **3**.

The tetra- and heptapeptide (Ser^{133}-Gly^{136} and Pro^{139}-Gln^{145})-linked Sieber amide resin interposing the allyl linker were synthesized by a peptide synthesizer based on standard Fmoc protocol. The side-chain functional groups were blocked with *t*-BuO group for aspartic acid, *t*-Bu group for threonine, and Pmc (2,2,5,7,8-pentamethylchroman-6-sulfonyl) group for arginine. The coupling reactions were performed with HBTU-HOBt-DIPEA in NMP, while the N-terminal Fmoc group was removed with piperidine in NMP. Coupling yield in each step was determined by ninhydrin test and overall yields were over 95%. The glycoserine **3** (1.2 equiv.) was then condensed with *N*-deprotected **6** and **11** (0.25 mmol scale), respectively, by activation with HBTU-HOBt-DIEA (1.1 equiv.). The reaction was performed by using a vortex-mixer for 24 h at room temperature. The unreacted amino group was capped by reaction with *N*-(2-chlorobenzyloxycarbonyloxy)succinimide [Z(2-Cl)-OSu] which was stable to the reagents commonly used for Fmoc/Boc chemistries.[23] On the other hand, use of acetic anhydride as a capping reagent was not appropriate because acetylation of the hydroxyl groups was observed. To our delight, the coupling efficiency proved to be very high (>90%), judging from HPLC analysis after cleavage of peptide **7** or **12** from the resin by treatment with 1% TFA-CH_2Cl_2 (Scheme 2). During the synthesis of **7** and **12** from **6** and **11**, no significant side reactions were observed.

Having prepared the resin-bound glycopeptide **7** carrying unprotected glycan chains, further amino acid



Scheme 2. Synthetic route for glycopeptide segments **9** and **13**.

assembly was next examined with several coupling reagents. Attempted incorporation of proline using DCC-HOBt in NMP or HBTU-HOBt-DIEA in NMP resulted in the formation of small amounts (<10%) of doubly prolinylated products along with the desired glycopeptide **8**. The by-products, presumably derived from *O*-prolinylation, were separated by HPLC and the molecular weights were determined by MALDI-TOF MS. On the other hand, the use of Fmoc amino acid pentafluorophenyl (Pfp) ester as an amino acid source turned out to be effective, obviating the overacylation. Thus, sequential incorporation of proline and serine residues proceeded in a highly satisfactory manner to give the corresponding peptide as the sole product. Fig. 2 shows HPLC profile of the crude glycoheptapeptide **9** which can be readily used for the subsequent fragment condition (vide infra). It is to be noted that the C-terminal glycine residue was liberated as free carboxylic acid after cleavage by treatment of **8** with 10 mol% of $Pd(PPh_3)_4$ and 10 equiv. of *N*-methylaniline in DMSO at room temperature.[24] The compound **9** was obtained in good overall yield (51%) through purification by preparative HPLC (Fig. 3a).

HPLC analysis of the C-terminal glycononapeptide **13** (Pro^{137}-Gln^{145}), similarly prepared on resin and released by 1% TFA-CH_2Cl_2, is demonstrated in Fig. 3b. The peak at 34.44 min corresponds to the glycononapeptide.

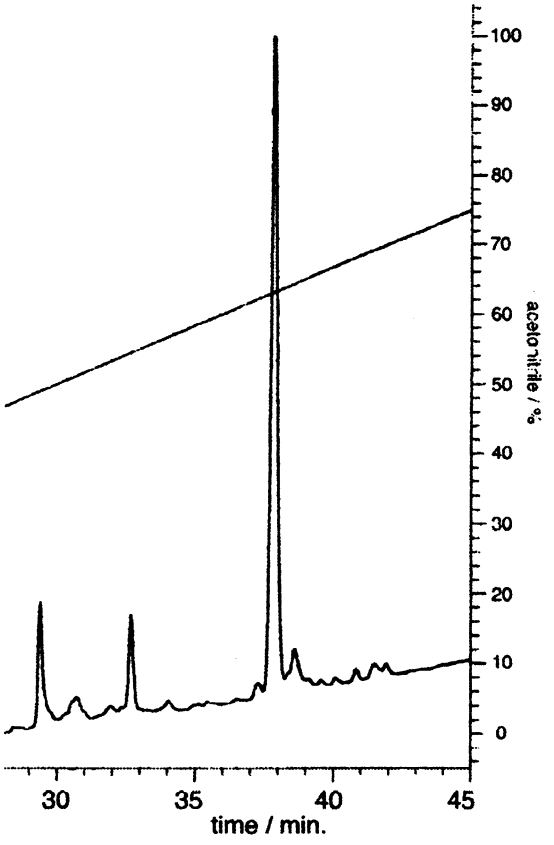


Figure 2. HPLC elution profiles of crude Ser^{130}-Gly^{136}. Elution conditions: column, Mightysil RP-18 GP (3.0×150 mm) at a flow rate of 0.6 ml min^{-1}; eluent, aqueous acetonitrile containing 0.1% TFA.

Coupling reaction of the peptide resin **13** with 2.0 equiv. of peptide segment **9** was performed using DCC-PfpOH (2.0 and 6.0 equiv.) as an activator in NMP for 7 days at room temperature. Formation of the coupling product **1** was monitored by HPLC analysis and the MALDI-TOF MS measurement after cleavage of the sample resin by treatment with 1% TFA-CH_2Cl_2 (Scheme 3). The high efficiency of this coupling reaction was clearly seen by the chromatographic profile of the final reaction mixture cleaved from the resin (Fig. 3c). The peak for C-terminal fragment was no longer observed and a near peak corresponding to the desired product **1** appeared at

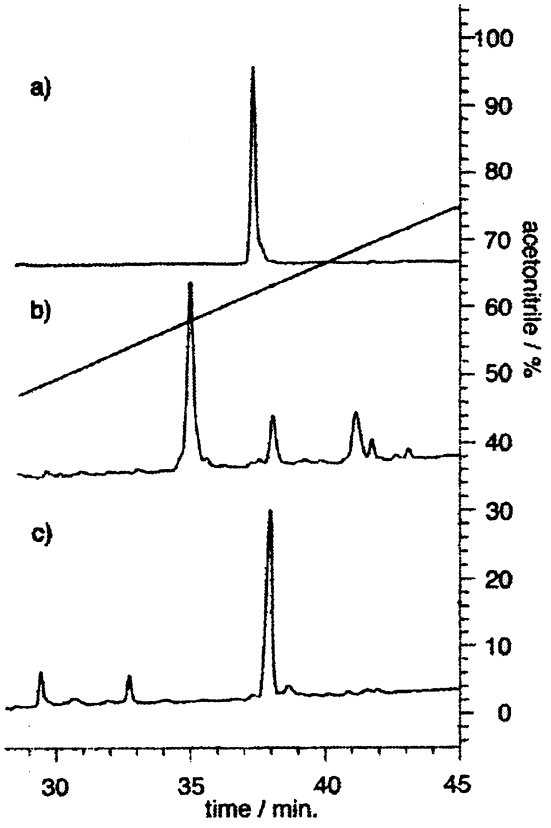
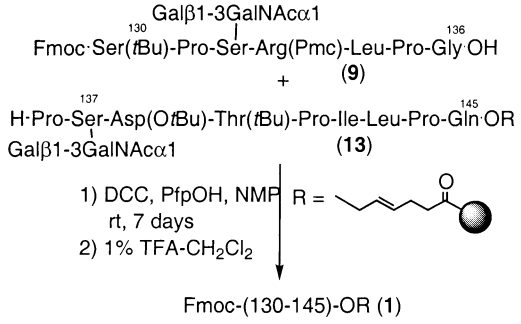


Figure 3. HPLC elution profiles of glycopeptide; (a) pure Ser^{130}-Gly^{136} (**9**), (b) crude Pro^{137}-Gln^{145} (**13**), (c) crude Ser^{130}-Gln^{145} (**1**) after cleavage from resin. Elution conditions: column, Mightysil RP-18 GP (3.0×150 mm) at a flow rate of 0.6 ml min^{-1}; eluent, aqueous acetonitrile containing 0.1% TFA.



Scheme 3. Segment coupling of glycopeptides **9** and **13**.

37.83 min. The target glycopeptide **1** was isolated in 30% overall yield and characterized by MALDI-TOF MS and NMR.[25]

In summary, we have achieved Fmoc-based solid-phase synthesis of the glycopeptide fragment (Ser^{130}-Gln^{145}) of the β-subunit of human chorionic gonadotropin by utilizing a glycosyl serine building block carrying unmasked hydroxyl groups. It was shown that amino acid pentafluorophenyl esters could serve as the chemoselective acyl donor. That the on-resin coupling of glycopeptide carrying unmasked oligosaccharide was accomplished in high efficiency may well open the way to chemical synthesis of native glycoproteins.

Acknowledgements

This work was supported by the CREST program of the Japan Science and Technology Corporation and by the Science and Technology Agency of the Japanese Government through Special Researcher's Basic Science Program at RIKEN. We are grateful to Ms. A. Takahashi for technical assistance.

References

1. (a) Varki, A. *Glycobiology* **1993**, *3*, 97–130; (b) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1–27.
2. (a) Kunz, H. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 294–308; (b) Kunz, H.; Brill, W. K.-D. *Trends Glycosci. Glycotechnol.* **1992**, *4*, 71–82.
3. (a) Meldal, M.; Bock, K. *Glycoconjugate J.* **1994**, *11*, 59–63; (b) Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H. *Chem. Rev.* **2000**, *100*, 4495–4538; (c) Garsequell, G.; Valencia, G. *Tetrahedron: Asymmetry* **1997**, *8*, 2839–2876.
4. Paulsen, H. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 823–839.
5. Kihlberg, J.; Elofsson, M. *Curr. Med. Chem.* **1997**, *4*, 79–110.
6. Kihlberg, J. In *Fmoc Solid-Phase Peptide Synthesis: A Practical Approach*; Chan, W. C.; White, P. D., Eds.; Oxford University Press: Oxford, 2000; pp. 195–213.
7. Atherton, E.; Fox, H.; Harkiss, D.; Logan, C. J.; Sheppard, R. C.; Williams, B. J. *J. Chem. Soc., Chem. Commun.* **1978**, 537–539.
8. Chang, C.-D.; Meienhofer, J. *Int. J. Peptide Protein Res.* **1978**, *11*, 246–249.
9. Nakahara, Y.; Ando, S.; Itakura, M.; Kumabe, N.; Hojo, H.; Ito, Y.; Nakahara, Y. *Tetrahedron Lett.* **2000**, *41*, 6489–6493.
10. (a) Hoogerhout, P.; Guis, C. P.; Erkelens, C.; Bloemhoff, W.; Kerling, K. E. T.; van Boom, J. H. *Recl. Trav. Chim. Pays-Bas* **1985**, *104*, 54–59; (b) Andrews, D. M.; Seale, P. W. *Int. J. Pept. Protein Res.* **1993**, *42*, 165–170.
11. Sjölin, M.; Elofsson, M.; Kihlberg, J. *J. Org. Chem.* **1996**, *61*, 560–565.
12. Erbing, B.; Lindberg, B.; Norberg, T. *Acta Chem. Scand.* **1978**, *B32*, 308–310.
13. Paulsen, H.; Schultz, M.; Klamann, J.-D.; Waller, B.; Paal, M. *Liebigs Ann. Chem.* **1985**, 2028–2048.
14. Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Carbohydr. Res.* **1996**, *292*, 71–81.
15. Guo, Z.-W.; Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Bioorg. Med. Chem.* **1997**, *5*, 1917–1924.
16. Guo, Z.-W.; Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1464–1466.
17. Nakahara, Y.; Nakahara, Y.; Ito, Y.; Ogawa, T. *Carbohydr. Res.* **1998**, *309*, 287–296.
18. (a) Ando, S.; Nakahara, Y.; Ito, Y.; Ogawa, T.; Nakahara, Y. *Carbohydr. Res.* **2000**, *329*, 773–780; (b) Nakahara, Y.; Nakahara, Y.; Ito, Y.; Ogawa, T. *Tetrahedron Lett.* **1997**, *38*, 7211–7214.
19. Guo, Z.-W.; Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Carbohydr. Res.* **1997**, *303*, 373–377.
20. Hojo, H.; Watabe, J.; Nakahara, Y.; Nakahara, Y.; Ito, Y.; Nabeshima, K.; Toole, B. P. *Tetrahedron Lett.* **2001**, *42*, 3001–3004.
21. (a) Christiansen-Brams, I.; Jansson, A. M.; Meldal, M.; Breddam, K.; Bock, K. *Bioorg. Med. Chem.* **1994**, *2*, 1153–1167; (b) Elofsson, M.; Salvador, L. A.; Kihlberg, J. *Tetrahedron* **1997**, *53*, 369–390; (c) Broddefalk, J.; Bergquist, K.-E.; Kihlberg, J. *Tetrahedron* **1998**, *54*, 12047–12070; (d) Broddefalk, J.; Bäcklund, J.; Almqvist, F.; Johansson, M.; Holmdahl, R.; Kihlberg, J. *J. Am. Chem. Soc.* **1998**, *120*, 7676–7683; (e) Holm, B.; Linse, S.; Kihlberg, J. *Tetrahedron* **1998**, *54*, 11995–12006; (f) Broddefalk, J.; Forsgren, M.; Sethson, I.; Kihlberg, J. *J. Org. Chem.* **1999**, *64*, 8948–8953.
22. (a) Mizuno, M.; Haneda, K.; Iguchi, R.; Muramoto, I.; Kawakami, T.; Aimoto, S.; Yamamoto, K.; Inazu, T. *J. Am. Chem. Soc.* **1999**, *121*, 284–290; (b) Inazu, T.; Ishida, H.; Nagano, R.; Tanaka, K.; Haneda, K. *Pept. Sci.* **1999**, 121–124; (c) Mizuno, M.; Muramoto, I.; Kawakami, T.; Seike, M.; Aimoto, S.; Haneda, K.; Inazu, T. *Tetrahedron Lett.* **1998**, *39*, 55–58; (d) Unverzagt, C. *Tetrahedron Lett.* **1997**, *38*, 5627–5630; (e) Reimer, K. B.; Meldal, M.; Kusumoto, S.; Fukase, K.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1993**, 925–932.
23. Ball, H. L.; Mascagni, P. *Lett. Pept. Sci.* **1995**, *2*, 49–57.
24. Kunz, H.; Unverzagt, C. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 436–437.
25. Selected physical data are given below. **3**: $[\alpha]_D$ +6.79 (*c* 3.73, CF_3CH_2OH), ^1H NMR (400 MHz, DMSO-d_6, 60°C) δ 7.90 (d, $J=7.6$ Hz, 2H; Ar), 7.70–7.73 (m, 2H; Ar), 7.45 (dd, $J=6.8$ Hz, 7.2 Hz, 2H; Ar), 7.36 (dd, $J=6.8$ Hz, 7.6 Hz, 2H; Ar), 4.77 (d, $J=3.6$ Hz, 1H; H-1a), 4.21–4.38 (m, 7H, including at 4.34 ppm, d, $J=7.6$ Hz, H-1b), 4.01 (brs, 1H), 3.30–3.81 (m, 11H), 1.86 (s, 3H; Ac); HRMS (FAB+) Calcd for $C_{32}H_{40}N_2O_{15}$: 693.2507. Found: 693.2515. **9**: ^1H NMR (600 MHz, DMSO-d_6, 60°C) δ 7.88 (d, $J=7.6$ Hz, 2H; Ar), 7.50–7.64 (m, 4H; Ar), 7.28–7.43 (m, 4H; Ar), 0.80–0.92 (m, 6H; Leu). ^{13}C NMR (150 MHz, DMSO-d_6, 60°C) δ 104.4, 98.2. HRMS (FAB+) Calcd for $C_{77}H_{111}N_{11}O_{25}S$: 1622.7552. Found: 1622.7545. **1**: ^1H NMR (600 MHz, DMSO-d_6, 60°C) δ 7.88 (d, $J=7.6$ Hz, 1H; Ar), 7.68–7.72 (m, 2H; Ar), 7.51–7.55 (m, 1H; Ar), 7.39–7.41 (m, 2H; Ar), 7.25–7.33 (m, 3H; Ar), 5.71–5.78 (m, 1H; allyl), 5.48–5.56 (m, 1H; allyl), 0.80–0.88 (m, 12H; Leu), 0.76–0.79 (m, 3H; Ile). MALDI-TOF MS: Calcd for $(M+Na)^+$: 3183.0. Found $(M+Na)^+$: 3182.9.