Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 14 (2004) 2527–2531

Structure of pentasaccharide of glycopeptide from TIME-EA4, N-glycoprotein in silkworm diapause eggs

Suthasinee Pitchayawasin,[a] Minoru Isobe,[a,*,†] Naoki Tani[a] and Hidenori Kai[b]

[a]*Laboratory of Organic Chemistry, Bioagricultural Sciences, Nagoya University, Chikusa, Nagoya 464-8601, Japan*
[b]*Laboratory of Insect Biochemistry and Biotechnology, Department of Biochemistry and Biotechnology, Faculty of Agriculture, Tottori University, Koyama, Tottori 680-8553, Japan*

Received 17 January 2004; revised 24 February 2004; accepted 28 February 2004

Abstract—The TIME-EA4, from silkworm diapause eggs of pure strain C108, *Bombyx mori*, has glycosylated chain as tetrasaccharide ($Man_2GlcNAc_2$) attaching to the Asn(22) of T3 peptide from tryptic digests. On the other hand, from Showa silkworm strain we additionally observed a pentasaccharide ($Man_3GlcNAc_2$) on T3 at the same linkage site. The linkage pattern of the 5-sugar chain was studied through Smith degradation combined with LC–MS and MS/MS analyses. These advanced methods led us to conclude that the pentasaccharide was branching as Man 1 → 3(Man 1 → 6)Man 1 → 4GlcNAc 1 → 4GlcNAc.

A glycoprotein, TIME-EA4, was isolated from silkworm diapause eggs, *Bombyx mori*, pure strain C108, as a key to terminate embryonic diapause.[1] It was known to play an important role on protein conformational change through the interaction with PIN peptides.[2] The amino acid sequence of TIME-EA4 was found to be close to Superoxide dismutase except TIME-EA4 has the glycosylated chain.[3] In the biosynthesis of *N*-glycosidic linkage in glycoprotein, the primary peptide sequence requirements for oligosaccharyltransferase-catalyzed glycosylation include a minimum Asn-Xaa-Ser/Thr tripeptide recognition motif where Xaa can be any of the 20 natural amino acids except proline.[4,5] In case of TIME-EA4, the glycosylated site has recently been determined by enzymatic digestion with trypsin and the resulting peptides were analyzed with LC-Q-TOF-MS. Among the 14 tryptic peptides, the T3 peptide from Gly(21) to Lys(32) was suggested to be *N*-glycan having tetrasaccharide chain linking at Asn(22).[6]

On the other hand, we found that T3 from another Showa silkworm strains have two kinds of sugar chains (2-3Man+2GlcNAc); thus, tetra- and penta-saccharides. However the linkage position and pattern of these saccharides on T3 were not exactly identified yet because of only trace amount being available from the source glycoprotein. The general methods to determine the oligosaccharide linkage pattern were reported by Takahashi et al.[7] They digested the glycoproteins by glycopeptidase to release free oligosaccharides, which were further modified into pyridylamino derivatives. The structures of oligosaccharides were analyzed by two-dimensional mapping technique (comparison with authentic samples) with co-chromatography on an ODS and an amide–silica column.[8] This method was outstanding but not applicable to the current work due to the trace amount.

In this paper, we tried to establish an alternative, more convenient and direct analytical method; thus, we employed the Smith degradation[9–12] and the tryptic hydrolysates (without component separation) to determine the glycan structure. In a reversed phase LC system, the hydrophobicity of these peptides is useful to let them elute after the non-hydrophobic mass and reasonably before the normal hydrophobic peptides.

We found these glycopeptides were readily separable on an ODS (0.3 mm × 150 mm) column with nano-LC system as demonstrated on ribonuclease B as an example.[13]

Keywords: TIME-EA4; Silkworm diapause eggs; Glycoprotein; Smith degradation.
* Corresponding author. Fax: +81-52-789-4111; e-mail: isobem@agr.nagoya-u.ac.jp
† Bioorganic chemistry of TIME EA4 2; for Part 1, see Ref. 4.

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.02.093

The tryptic peptide mixtures were directly injected to the nano-LC system connected on-line to MS without any flow splitting. LC-ESI-Q-TOF-MS and MS/MS is useful to determine the glycosylated chain linkage pattern of pentasaccharide on TIME-EA4, which is shown in Figure 1.

Smith degradation includes two-step reactions; thus, sodium metaperiodate oxidation and sodium borohydride reduction. The terminal mannose (hexose) after Smith degradation is cleaved at the C2–C3 and C3–C4 bonds at the same time and the fragmentation mass loss is 134 Da instead of intact mannose being 162 Da.[13] The mass difference 28 Da is corresponding to $C_1H_0O_1$. On the other hand, an internal mannose, which has substitution at the position 2, is cleaved only at the C3–C4 and leaves the mass loss of 164 Da. The internal mannose, which has substitution at the 3 position, has no cleavage and leaves the same value of 162 Da. Among the high mannose core structures of natural glycoproteins, only the cases of substitution at the position 2 or 3 have been reported so far.[14] In order to determine the sugar chain structure of TIME-EA4 from Showa, it was first treated under Smith degradation condition, and secondly hydrolyzed by trypsin. Any reverse operation (first Trypsin and second Smith degradation) never provided good results. The tetra- and penta-saccharides on T3 peptide were easily found at retention time about 16 min, and only the latter (T3-pentasaccharide) being shown in Figure 2a. The doubly charged ions were selected as the respective precursor for the MS/MS, which are illustrated in Figure 3.

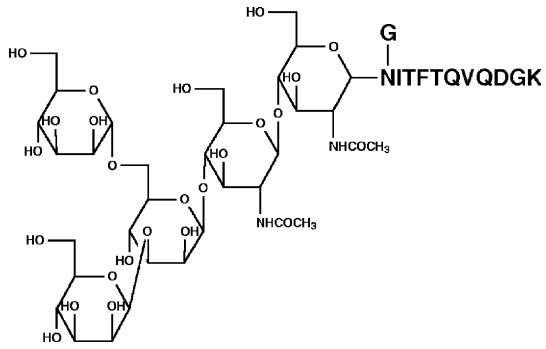

Figure 1. Structure of pentasaccharide [MW 2198.97] attached to the consensus sequence-containing peptide, T3 fragment from N-glycoprotein (TIME-EA4) found in the silkworm diapause eggs.

As has already been established in our laboratory on C108 silkworm strain,[6] TIME-EA4 was digested by trypsin to provide a mixture of 14 peptides (T1–T14). The T3 was found to be a glycopeptide since it appeared at m/z 1019.51 (M+H=2038.02), which is plus 730.25 Da (equivalent to tetrasaccharide) higher than the value m/z 1307.77 M+H of the peptide, Gly-Asn(22)-Ile-Thr-Phe-Thr-Gln-Val-Gln-Asp-Gly-Lys(32). In case of Showa silkworm in Figure 2a, we have also found the tetrasaccharide peak at 15.83 min and newly m/z 1100.49 (M+H=2199.98) at 15.77 min, which is equivalent to pentasaccharide (plus 892.21 Da). Both of these glycopeptides T3-tetra- and T3-penta-saccharides were proven to attach at the Asn(22) under the same method as we have reported.[6]

The peptide mixtures from tryptic TIME-EA4 after the Smith degradation operation were injected to LC-ESI-MS and the glycopeptides again appeared at retention time around 16 min as illustrated in the ESI-MS spec-

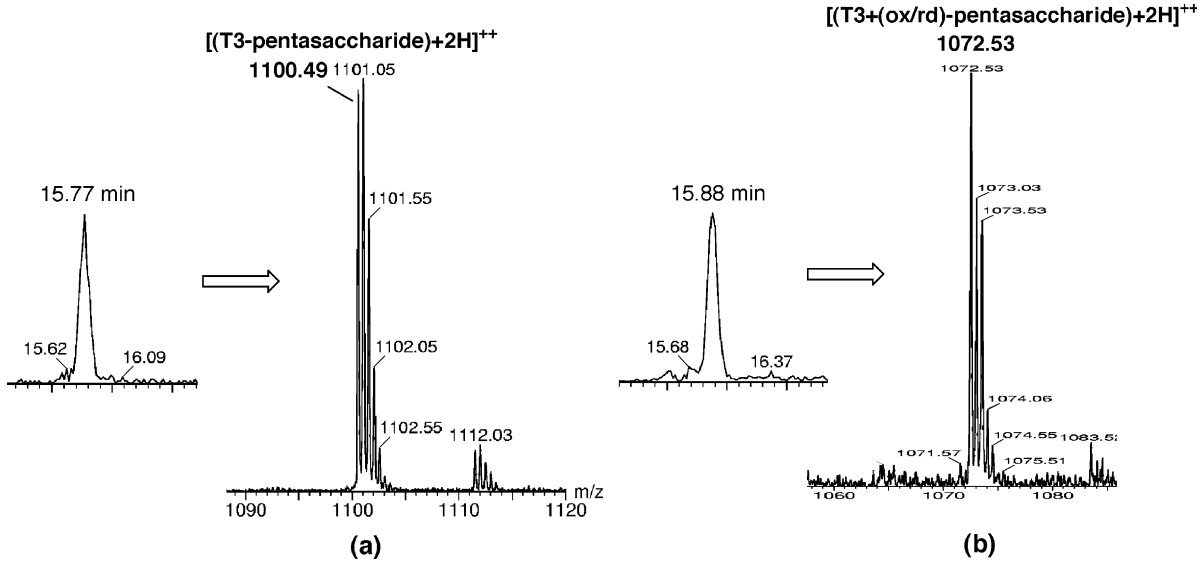

Figure 2. Ion Chromatograms and Mass Spectra of the T3-pentasaccharide obtained from tryptic TIME-EA4 without (**a**) and with (**b**) Smith degradation. Condition: Column, Develosil ODS-HG-5 (0.3 mm i.d. × 150 mm, Nomura, Co. Ltd, Aichi, Japan), Gradient, linear from 0–100% H_2O–CH_3CN, containing 0.025% TFA, flow rate, 5 μL/min. (a) T3-pentasaccharide without Smith degradation: The $[M+2H]^{++}$ ion is observed at m/z 1100.49 (equivalent to M+H 2199.98) with retention time at 15.77 min; (b) T3-pentasaccharide with Smith degradation: The $[M+2H]^{++}$ ion is observed at m/z 1072.53 (equivalent to M+H 2144.06) with retention time at 15.88 min.

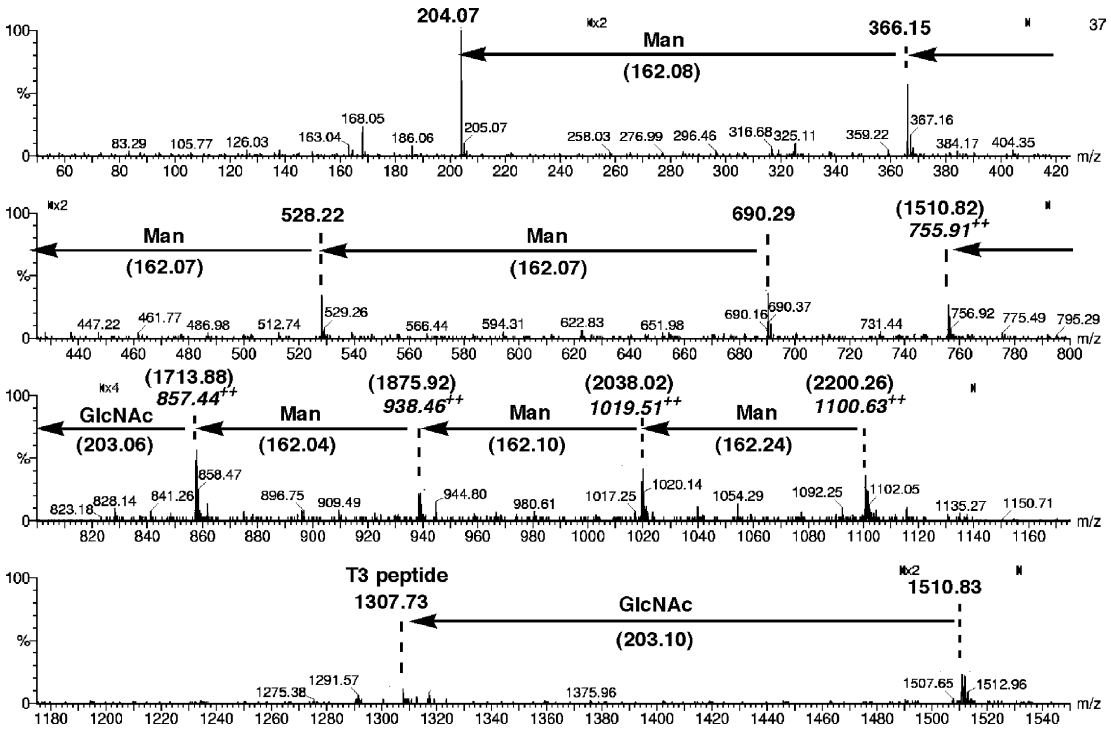

Figure 3. MS/MS spectrum of T3-pentasaccharide from tryptic TIME-EA4 (precursor, $m/z = 1100.63$, cone voltage 40 V, collision energy 35–51 eV).

trum in Figure 2b. The corresponding MS spectrum shows doubly charged ions of glycopeptide, T3-pentasaccharide-derived ion at m/z 1072.53 (M+H=2144.06) with mass difference of 55.92 Da (equivalent of $C_2H_0O_2$) compared with the one without Smith degradation. This suggests that the pentasaccharide includes two terminal mannoses at the nonreducing ends.

The MS/MS spectrum of the T3-pentasaccharide is shown in Figure 3, when m/z 1100.63 $[M+2H]^{++}$ ion was employed as the precursor. The plus 2 charge fragment ions are found as m/z 1100.63, m/z 1019.51, m/z 938.46, m/z 857.44, and m/z 755.91, corresponding to the M+H of m/z 2200.26, m/z 2038.02, m/z 1875.92, m/z 1713.88, and m/z 1510.82, respectively. The mass differences are 162.24, 162.10, 162.04, and 203.06. In addition, the plus 1 charge fragment ions are found at m/z 1510.83 and m/z 1307.73 (T3-peptide), being the mass difference of 203.10. The sugar fragmentation pattern is also followed from the values of m/z 690.29, m/z 528.22, m/z 366.15, and m/z 204.07 (GlcNAc+H), with the mass difference of 162.07, 162.07, and 162.08, corresponding to three mannoses. These data are tabulated in Table 1, concluding that the T3 peptide is linking to $Man_3GlcNAc_2$. The last question is the linkage pattern of the mannoses.

A series operations were processed with the sample of TIME-EA4 firstly by Smith degradation, secondly by trypsin digestion and finally LC–MS/MS measurements from m/z 1072.55 as precursor ion. The spectrum is indicated in Figure 4, wherein the assignments are drawn with arrows and captions. Almost all of the peaks in Figures 3 and 4 are identical, while only the precursor ion m/z 1072.55, and the fragment ions m/z 1005.54 and m/z 500.23 are different so that the mass difference can

Table 1. Assignment of MS/MS spectra of T3-pentasaccharide of tryptic TIME-EA4 without and with Smith degradation

Fragment of T3-pentasaccharide	Without Smith degradation precursor, $m/z = 1100.63$			With Smith degradation precursor, $m/z = 1072.55$		
	M+H obs (cal)	$[M+2H]^{++}$ obs	Mass difference	M+H obs (cal)	$[M+2H]^{++}$ obs	Mass difference
T3+2GlcNAc+3Man	(2200.25)	1100.63		(2144.10)	1072.55	
T3+2GlcNAc+2Man	(2038.02)	1019.51	162.24	(2010.08)	1005.54	134.02
T3+2GlcNAc+Man	(1875.92)	938.46	162.10	(1876.02)	938.51	134.06
T3+2GlcNAc	(1713.88)	857.44	162.04	(1713.96)	857.48	162.06
T3+GlcNAc	1510.83 (1510.82)	755.91	203.06	1510.88 (1510.86)	755.93	203.10
T3	1307.73		203.10	1307.77 (1307.76)	654.38	203.11
GlcNAc+3Man	690.29			634.24		
GlcNAc+2Man	528.22		162.07	500.23		134.01
GlcNAc+Man	366.15		162.07	366.15		134.08
GlcNAc	204.07		162.08	204.07		162.08

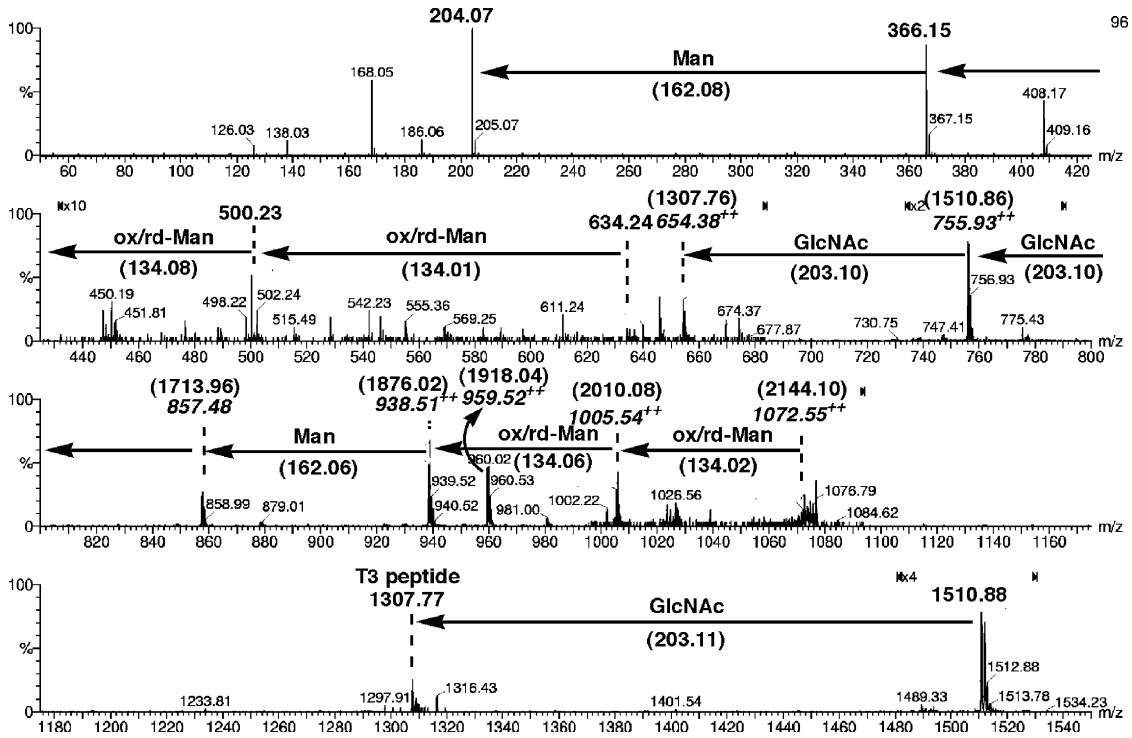

Figure 4. MS/MS spectrum of T3-(ox/rd)pentasaccharide from tryptic TIME-EA4 (precursor, $m/z = 1072.55$, cone voltage 40 V, collision energy 35–51 eV).

be comparable as follows: thus, the series of fragment ions [m/z 1100.63, m/z 1019.51, m/z 938.46] in Figure 3 correspond to [m/z 1072.55, m/z 1005.54, m/z 938.51] in Figure 4, respectively, leaving the difference mass of 134.02 and 134.06. This suggests the two terminal mannoses, in fact, lost 28 Da from the parent mass, respectively. The reason why no change of the peaks m/z 938 and 857 before and after the Smith degradation depends on the linkage position of the internal mannose should be at the 3 position, but not at either 2 or 4 position. Incidentally, the fragment doubly charged ion at m/z 959.52 is assigned to be a loss of glycerin (92.05 Da) from m/z 1005.54. Thus, the structure of T3-pentasaccharide must have the structure as (**a**) in Figure 5.

The structures of T3-pentasaccharide are proposed as the four possible linkage patterns as shown in Figure 5; structure **a**, internal Man has substitution at the 3 and 6 positions (branched chain), structure **b**, internal Man has substitution at the 2 and 6 positions (branched chain), structure **c**, 2 internal Man's have substitution at 2 position (linear chain) and structure **d**, one internal Man has substitution at the 2 position and the other internal Man has substitution at the 3 position (linear chain).

The fragments from MS/MS spectra of glycopeptide, T3-pentasaccharide from tryptic TIME-EA4, without and with Smith degradation, are summarized in Table 1. The inner Man, which connect to GlcNAc was not modified after Smith degradation, so this Man and 2 GlcNAc's act as the core structure in the pentasaccharide. The linkage pattern of sugar chains on TIME-EA4 can be distinguished by Smith degradation due to the changes of mass values from MS/MS spectra of the glycopeptide (without and with Smith degradation).

This can be doubly confirmed among all four possible structures, which are commonly found in nature as part of the high mannose type N-linked glycoprotein. Each of the structure can be differentiated from two major points of the mass differences; thus, (i) the inner mannose shows either 162 Da (substituent at the 3 position) or 164 Da (substituent either at the 2 or 4 position), (ii) the number(s) of terminal mannose(s) is (are) countable from the numbers of 134 mass loss. We described herein a detailed analysis of glycosylated chain on TIME-EA4 by Smith degradation method combined with LC-ESI-MS and MS/MS. This method was successful to determine the 5 sugar chain structure on this glycoprotein to have the branched chain. Although no data have been shown in this paper because of the space, the T3-tetrasaccharide of TIME-EA4 has the structure without the mannose linking with $1 \rightarrow 6$ (Fig. 1). The details will be reported elsewhere with additional data from C108 strain. These molecular varieties might be affecting the time control during the diapause termination, and the details are to be reported in due course. The current methodology is of powerful tool for the trace analysis, but not all mighty since the information is simply obtained by the numbers of ion masses. The disadvantage would be covered by other methods, so that the big analytical application would expand to the glycan structure of the glycoprotein obtainable in only trace amount from natural sources.

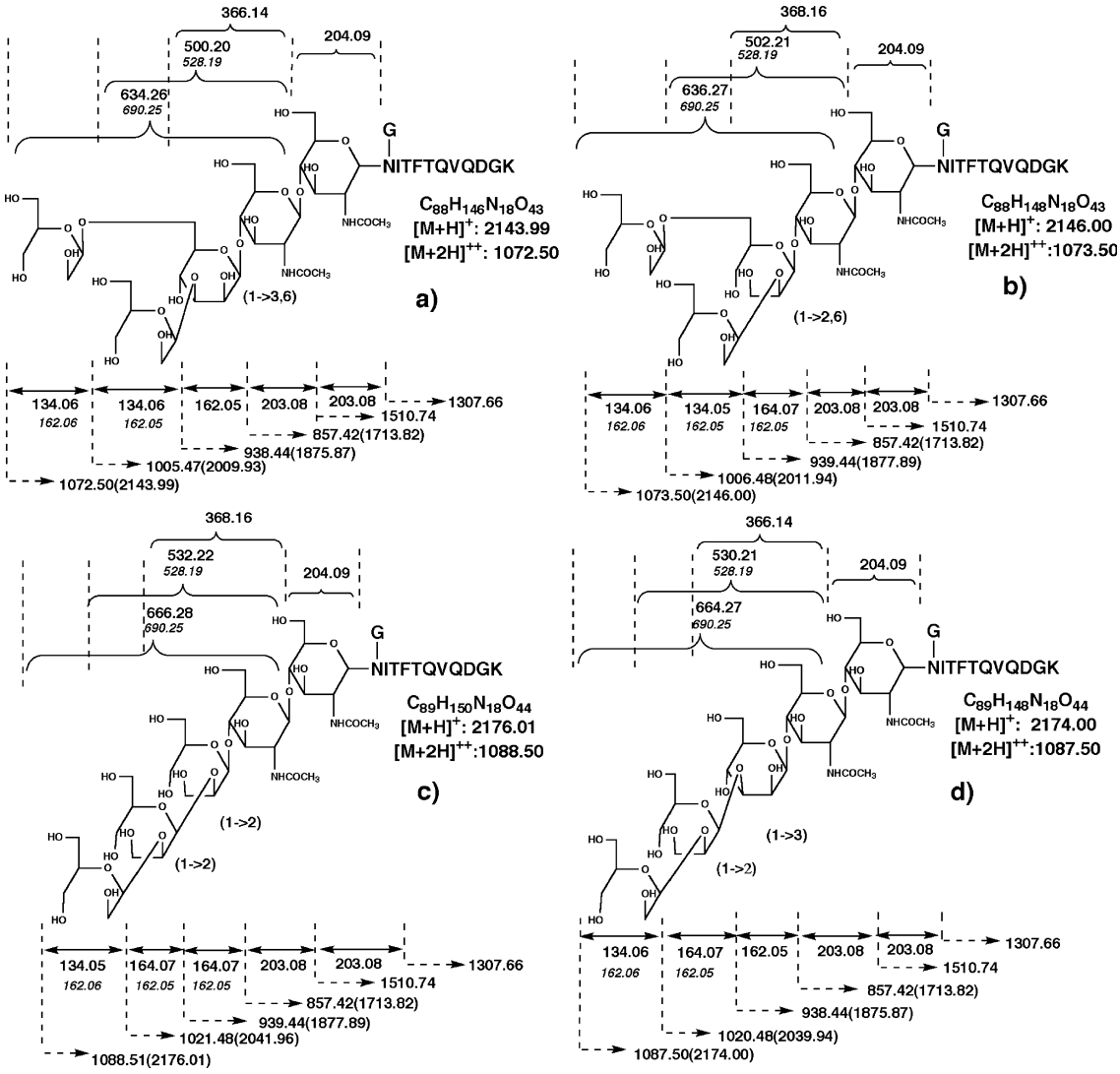

Figure 5. The four possible structures and MS/MS fragmentation of T3-pentasaccharide (T3+2GlcNAc+3Man) from tryptic TIME-EA4 with Smith degradation.

Acknowledgements

The authors are grateful to Jin-nai International Scholarship Program for S.P. during her PhD studies and we also thank to the Ministry of Education, Sciences, Culture and Sports for COE post-docteral researcher scholarship for N.T. We acknowledge Ueda Silkworm Egg Production Co. Inc., Nagano, Japan for cooperation to collect the eggs and the financial support from Ono Pharm. Co. Ltd.

References and notes

1. Tani, N.; Kamada, G.; Ochiai, K.; Isobe, M.; Suwan, S.; Kai, H. *J. Biochem. (Tokyo)* **2001**, *129*, 221–227.
2. Isobe, M.; Suwan, S.; Kai, H.; Katagiri, N.; Ikeda, M. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2851–2854.
3. Kurahashi, T.; Miyazaki, A.; Suwan, S.; Isobe, M. *J. Am. Chem. Soc.* **2001**, *123*, 9268–9278.
4. Wittmann, V. In *Glycosience III, Chemistry and Chemical Biology*; Fraser-Reid, B., Tatsuta, K., Thiem, J., Eds.; Springer: New York, 2001; pp 2257–2258.
5. Metzler, D. E. Sugars, Polysaccharides and Lycoproteins. In *Biochemistry*; Academic: San Deigo, CA, 2001; Vol. 1, pp 180–185.
6. Kurahashi, T.; Miyazaki, A.; Murakami, Y.; Suwan, S.; Franz, T.; Isobe, M.; Tani, N.; Kai, H. *Bioorg. Med. Chem.* **2002**, *10*, 1703–1710.
7. Takahashi, N.; Hitotsuya, H.; Hanzawa, H.; Arata, Y.; Kurihara, Y. *J. Biol. Chem.* **1990**, *265*, 7793–7798.
8. Takahashi, N.; Khoo, K. H.; Suzuki, N.; Johnson, J. R.; Lee, Y. C. *J. Biol. Chem.* **2001**, *276*, 23230–23239.
9. Krusius, T.; Fine, J. *Carbohydr. Res.* **1981**, *90*, 203–214.
10. Angel, A. S.; Nilsson, B. In *Mass Spectrometry, Methods in Enzymology*; McCloskey, J. A., Ed.; Academic: San Deigo, CA, 1990; Vol. 193, pp 587–607.
11. Hase, S.; Kikuchi, N.; Ikenaka, T.; Inoue, K. *J. Biochem.* **1985**, *98*, 863–874.
12. Witte, K.; Sears, P.; Martin, R.; Wong, C. *J. Am. Chem. Soc.* **1997**, *119*, 2114–2118.
13. Pitchayawasin, S.; Isobe, M. *Biosci. Biotechnol. Biochem.*, in press.
14. Dell, A.; Morris, H. R. *Science* **2001**, *291*, 2351–2354.