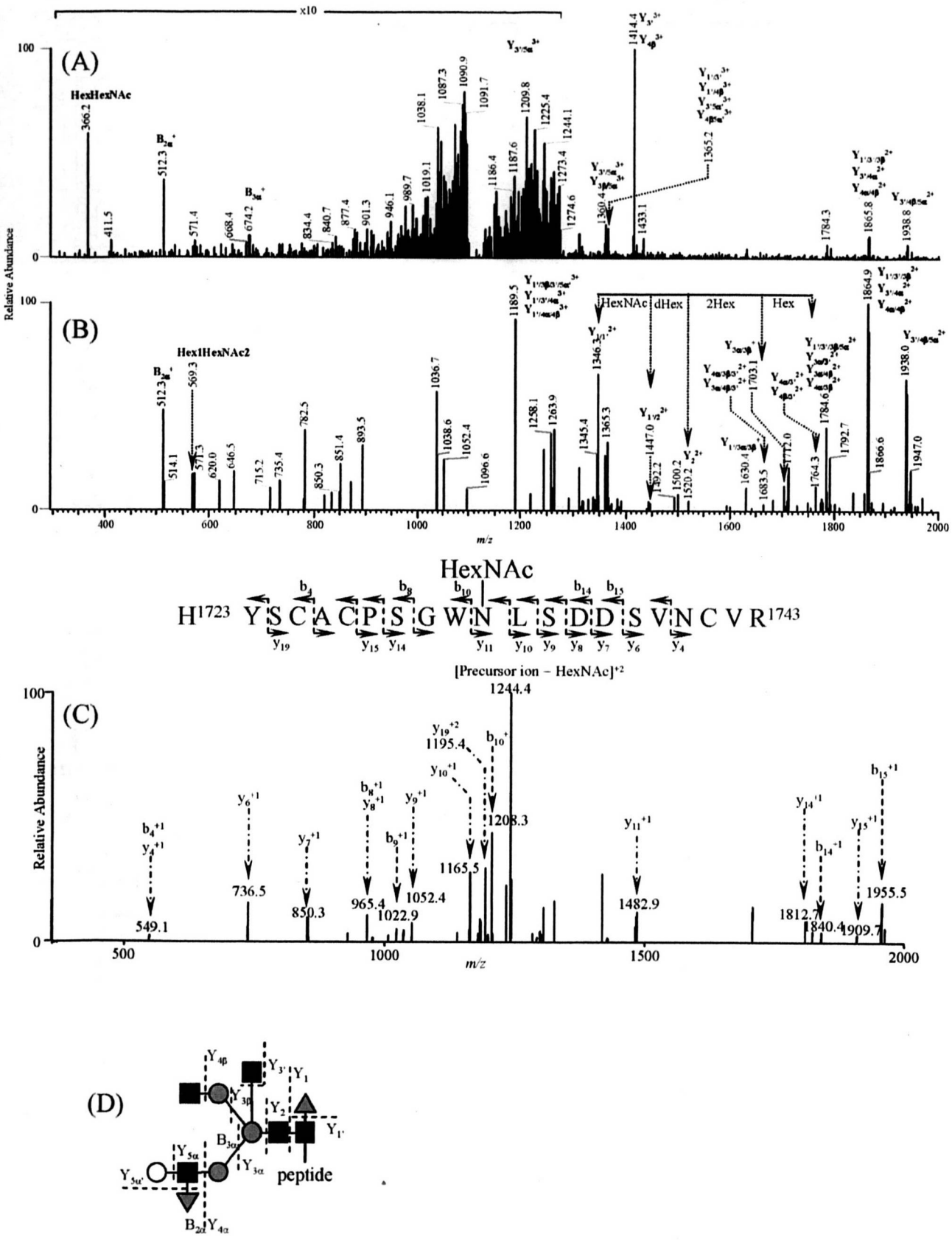


**Figure 5.** Identification of glycopeptide 15. (A) MS/MS spectrum acquired from the molecular ion $[M + 4H]^{4+}$ ($m/z$ 1111.5) of glycopeptide 15 in Figure 2A. (B) MS/MS/MS spectrum acquired from the most intense ion ($m/z$ 1414.4) in the MS/MS. (C) MS/MS/MS/MS spectrum acquired from the product ion ($m/z$ 1346.3) in the MS/MS/MS of glycopeptide 15, and amino acid sequence deduced from the results of database search analysis. (D) Deduced oligosaccharide structure.

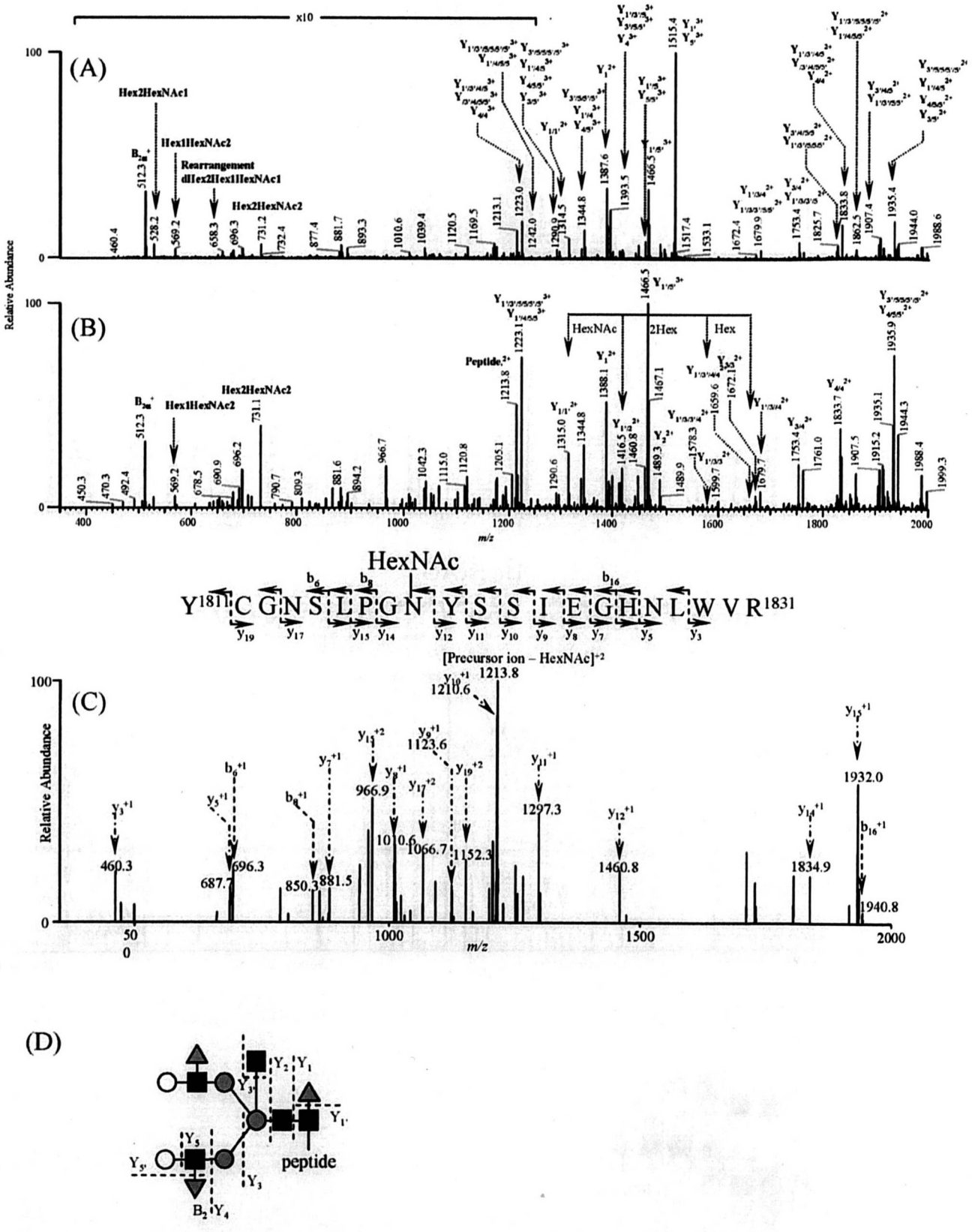


**Figure 6.** Identification of glycopeptide 17. (A) MS/MS spectrum acquired from the molecular ion $[M + 3H]^{3+}$ (*m/z* 1564.0) of glycopeptide 17 in Figure 2A. (B) MS/MS/MS spectrum acquired from the most intense ion (*m/z* 1515.4) in the MS/MS. (C) MS/MS/MS/MS spectrum acquired from the product ion (*m/z* 1315.0) in the MS/MS/MS of glycopeptide 17, and amino acid sequence deduced from the results of database search analysis. (D) Deduced oligosaccharide structure.

**Table 1.** Summary of Oligosaccharide Structures and Amino Acid Sequences of Le^x-Glycopeptides Detected by LC−MS^n and Database Search Analysis

| | Candidate Le^x-glycopeptide | | | | | Peptide | | | | | Deglycosylated peptide prepared by PNGaseF-treatment | | | | Oligosaccharide | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. | Observed m/z value of molecular-related ion[a] | Charge state | Observed m/z value of peptide-related ion[b] | Charge state | Additional sugar residue | Amino acid sequence | Xcorr[c] | Protein identified by database search analysis | GI no.[d] | Theoretical mass[e] | Calculated mass[f] | Observed m/z value of molecular-related ion[a] | Charge state | Xcorr[c] | Theoretical mass[e] | Calculated mass[f] | Deduced Structure | Theoretical mass[g] |
| 1 | 1136.129 | $[M+2H]^{2+}$ | 673.8 | 2 | HexNAc | $E^{[89]}$**GN*ATGHLMGR**$^{[100]}$ | 1.32 | Cubilin precursor | 88909268 | 1141.530 | 1141.530 | 572.280 | 2 | 2.57 | 1142.513 | 1142.545 | | 2281.845 |
| 2 | 1170.147 | $[M+3H]^{3+}$ | 1448.5 | 1 | HexNAc | **N*$^{[77]}$LSY**EAAPDIK$^{[29]}$ | 1.52 | cadherin 16 | 6660900 | 1247.583 | 1243.584 | 623.291 | 2 | 3.43 | 1244.567 | 1244.567 | | 2281.845 |
| 3 | 1152.145 | $[M+3H]^{3+}$ | 770.7 | 2 | HexNAc+dHex | R$^{[31]}$**N*WT**ETEVR$^{[39]}$ | 1.33 | dipeptidase 1 | 6681217 | 1189.584 | 1189.579 | 596.290 | 2 | 2.87 | 1190.568 | 1190.565 | | 2281.845 |
| 4 | 1293.858 | $[M+3H]^{3+}$ | 910.5 | 2 | HexNAc | Y$^{[34]}$**YN*QSA**GGSHTLQR$^{[54]}$ | 1.28 | H-2 class I histocompatibility antigen, K-K alpha precursor | 122155 | 1614.718 | 1614.717 | 808.865 | 2 | 4.19 | 1615.702 | 1615.715 | | 2281.845 |
| 5 | 1225.510 | $[M+3H]^{3+}$ | 1813.7 | 1 | HexNAc | T$^{[1453]}$**WSAFQN*GTDK**R$^{[1470]}$ | 1.32 | low[-]density lipoprotein receptor-related protein 2 (LRP2) | 124487372 | 1409.669 | 1409.673 | 471.227 | 3 | 1.66 | 1410.653 | 1410.659 | | 2281.845 |
| 6 | 1122.804 | $[M+3H]^{3+}$ | 807.7 | 2 | HexNAc | T$^{[1453]}$**WSAFQN*GTDK**R$^{[1470]}$ | 1.30 | low[-]density lipoprotein receptor-related protein 2 (LRP2) | 124487372 | 1409.669 | 1409.666 | 471.227 | 3 | 1.66 | 1410.653 | 1410.659 | | 1973.735 |
| 7 | 1231.510 | $[M+3H]^{3+}$ | 816.2 | 2 | HexNAc | —[h] | — | — | — | — | — | — | — | — | — | — | — | — |
| 8 | 1291.243 | $[M+3H]^{3+}$ | 906.2 | 2 | HexNAc | L$^{[30]}$**HNQLLIN*TTTVEK**$^{[54]}$ | 4.12 | γ-glutamyltransferase 1 (γ-GTP1) | 6679995 | 1606.868 | 1606.871 | 804.937 | 2 | 4.12 | 1607.852 | 1607.859 | | 2281.845 |
| 9 | 1156.466 | $[M+5H]^{5+}$ | 1070.0 | 3 | HexNAc | $^{[29]}$**HHCELQN*CNS**TEFTCNSRPPNRR$^{[70]}$ | — | low[-]density lipoprotein receptor-related protein 2 (LRP2) | 124487372 | 3002.281 | 3002.269 | 601.663 | 5 | 4.77 | 3003.265 | 3003.279 | | 2793.035 |
| 10 | 1100.210 | $[M+4H]^{4+}$ | 1170.0 | 2 | HexNAc | S$^{[90]}$**GQEDHYWLDVEKN*QSAK**$^{[114]}$ | 1.53 | alanyl (membrane) aminopeptidase | 225637487.00 | 2132.976 | 2132.976 | — | — | — | — | — | | 2281.845 |
| 11 | 1125.900 | $[M+5H]^{5+}$ | 915.6 | 3 | HexNAc | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 12 | 1237.506 | $[M+3H]^{3+}$ | 1649.6 | 1 | HexNAc | **N*$^{[345]}$MSS**EFYATQLR$^{[361]}$ | 1.73 | γ-glutamyltransferase 1 (γ-GTP1) | 6679995 | 1445.661 | 1445.660 | 724.340 | 2 | 3.78 | 1446.645 | 1446.665 | | 2281.845 |
| 13 | 1134.803 | $[M+3H]^{3+}$ | 1649.7 | 1 | HexNAc | **N*$^{[345]}$MSS**EFYATQLR$^{[361]}$ | 1.30 | γ-glutamyltransferase 1 (γ-GTP1) | 6679995 | 1445.661 | 1445.664 | 724.340 | 2 | 3.78 | 1446.645 | 1446.665 | | 1973.735 |
| 14 | 1163.044 | $[M+4H]^{4+}$ | 1295.4 | 2 | HexNAc | V$^{[1054]}$**VLVN*TTHK**PFDIHVLHPYR$^{[1081]}$ | 1.79 | low[-]density lipoprotein receptor-related protein 2 (LRP2) | 124487372 | 2584.311 | 2584.312 | 796.118 | 3 | 4.71 | 2585.295 | 2585.332 | | 2281.845 |
| 15 | 1111.453 | $[M+4H]^{4+}$ | 1346.3 | 2 | HexNAc | H$^{[1572]}$**YSCACPSGWN*LSD**DSVNCVR$^{[1594]}$ | 2.87 | low[-]density lipoprotein receptor-related protein 2 (LRP2) | 124487372 | 2486.060 | 2486.059 | — | — | — | — | — | | 1973.735 |
| 16 | 1145.460 | $[M+4H]^{4+}$ | 1157.2 | 2 | HexNAc | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 17 | 1563.318 | $[M+3H]^{3+}$ | 1315.0 | 2 | HexNAc | Y$^{[1911]}$**CGNSLPGN*YSS**IEGHNLWVR$^{[1931]}$ | 2.02 | Cubilin precursor | 88909268 | 2423.097 | 2423.098 | — | — | — | — | — | | 2281.845 |
| 18 | 1481.523 | $[M+2H+Na]^{3+}$ | 925.1 | 2 | HexNAc | V$^{[156]}$**GVTDVFPN*GTQFSR**$^{[171]}$ | — | Meprin A subunit beta precursor (endopeptidase-2) | 2499913 | 1622.805 | 1622.573 | 812.903 | 2 | 3.14 | 1623.789 | 1623.791 | | 2793.035 |
| 19 | 1156.990 | $[M+4H]^{4+}$ | 1282.5 | 2 | HexNAc | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 20 | 1557.382 | $[M+3H]^{3+}$ | 1204.5 | 2 | non-substituted | **N*$^{[1678]}$QSV**VMYSVPQPLGIAIIPSR$^{[1698]}$ | 1.38 | low[-]density lipoprotein receptor-related protein 2 (LRP2) | 124487372 | 2405.289 | 2405.290 | — | — | — | — | — | | 2281.845 |
| 21 | 1311.210 | $[M+H+2Na]^{3+}$ | 924.9 | 2 | HexNAc | V$^{[156]}$**GVTDVFPN*GTQFSR**$^{[171]}$ | — | Meprin A subunit beta precursor (endopeptidase-2) | 2499913 | 1622.805 | 1622.810 | 812.903 | 2 | 3.14 | 1623.789 | 1623.791 | | 2281.845 |
| 22 | 1386.270 | $[M+3Na]^{3+}$ | 948.4 | 2 | HexNAc | V$^{[156]}$**GVTDVFPN*GTQFSR**$^{[171]}$ | — | Meprin A subunit beta precursor (endopeptidase-2) | 2499913 | 1622.805 | 1622.928 | 812.903 | 2 | 3.14 | 1623.789 | 1623.791 | | 2484.925 |

[a] Value obtained by FTICR−MS. [b] Value obtained by IT−MS^n. [c] Value of cross correlation obtained by database search analysis. [d] GenInfo Identifier number. [e] Monoisotopic value. [f] Value calculated by subtraction of the theoretical masses of deduced oligosaccharides from the observed monoisotopic masses of the candidate Le^x-glycopeptide. [g] Monoisotoic value. [h] No data. *, modification with HexNAc; #, modification with HexNAc+dHex. Bold portions of amino acid sequences, consensus sequences of -linked oligosaccharides.

possible modification at Asn with HexNAc (203.1 u) and with dHex + HexNAc (349.1 u), seven glycopeptides were successfully sequenced with a high cross-correlation score (charge +1, Xcorr > 1.5; charge +2, Xcorr > 2.0; charge +3, Xcorr > 2.5; charge +4, Xcorr > 3.0). Figure 4C shows the MS/MS/MS/MS spectrum acquired from glycopeptide 8 (precursor ion: [peptide + HexNAc + 2H]$^{2+}$, *m/z* 906.2). The database search analysis resulted in Leu$^{503}$-Lys$^{516}$ in γ-glutamyl transpeptidase 1 (γ-GTP1) (charge +2, Xcorr: 4.12) (Table 1). The linkage of GlcNAc at Asn$^{510}$ in the *N*-glycosylation consensus sequence, Asn-Thr-Thr, was suggested by the good agreement between the experimental b/y-ion pattern and the predicted pattern.

The MS/MS/MS/MS spectrum acquired from [peptide + HexNAc + 2H]$^{2+}$ (*m/z* 1346.3, glycopeptide 15) is shown in Figure 5C. This peptide was identified as His$^{1723}$-Arg$^{1743}$ in low-density lipoprotein receptor-related protein 2 (LRP2, megalin) (charge +2, Xcorr: 2.82). The b- and y-ion pattern suggested the linkage of GlcNAc at Asn$^{1743}$ in the *N*-glycosylation consensus sequence, Asn-Lys-Ser.

Figure 6C shows the MS/MS/MS/MS spectrum acquired from another expected Le$^x$-conjugated glycopeptide (glycopeptide 17; precursor ion: [peptide + HexNAc + 2H]$^{2+}$, *m/z* 1315.0). Database search analysis revealed that this peptide could be Tyr$^{1011}$-Arg$^{1031}$ in the cubilin precursor (charge +2, Xcorr: 2.02) (Table 1). It was also suggested that the linkage position of GlcNAc was Asn$^{1019}$ in the *N*-glycosylation consensus sequence of Asn$^{1019}$-Tyr-Ser$^{1021}$.

Glycopeptides 2, 10, 12, and 14 were also successfully identified as Asn$^{519}$-Lys$^{529}$ in cadherin 16 (glycosylation site: Asn$^{519}$; charge +1, Xcorr: 1.52), Ser$^{593}$-Lys$^{610}$ in alanyl (membrane) aminopeptidase (glycosylation site: Asn$^{606}$; charge +2, Xcorr: 1.53), Asn$^{343}$-Arg$^{354}$ in γ-GTP1 (glycosylation site: Asn$^{343}$; charge +1, Xcorr: 1.73) and Val$^{3444}$-Lys$^{3463}$ in LRP2 (glycosylation site: Asn$^{3440}$; charge +2, Xcorr: 1.79), respectively. Additionally, we deduced that glycopeptides 1, 3–6, 13, and 20 could be Le$^x$-conjugated glycopeptides from tolerable scores (charges +1 and +2 Xcorr > 1.30) (Table 1). All identified or probable glycopeptides contained consensus sequences of *N*-linked oligosaccharides.

By the present method, three glycoproteins were identified as proteins carrying multiple Le$^x$-conjugated oligosaccharides—namely, γ-GTP1 (glycosylation site: Asn$^{343}$ and Asn$^{510}$; glycopeptides 8, 12 and 13), LRP2 (glycosylation site: Asn$^{1497}$, Asn$^{1676}$, Asn$^{1733}$ and Asn$^{3446}$; glycopeptides 5, 6, 14, 15 and 20), and a cubilin precursor (glycosylation site: Asn$^{1002}$, Asn$^{1819}$, glycopeptides 1 and 17). Only one glycopeptide was sequenced, but it was deduced that cadherin 16, dipeptidase 1, H-2 class I histocompatibility antigen, and K–K alpha precursor (H2–K(k)), and alanyl (membrane) aminopeptidase could be the Le$^x$-conjugated glycoproteins.

The sequences of the Le$^x$-conjugated glycopeptides were confirmed by an additional LC–MS/MS of deglycosylated peptides prepared by PNGaseF-treatment. Because of the deamination of Asn residues by PNGase F treatment, we set the *m/z* values of [peptide + nH + 0.984 u monoisotopic mass]$^{n+}$ ($n$ = 2–5) as precursor ions on the MS/MS. By this conventional method, 8 peptides that were sequenced by our method were identified as shown in Table 1. Moreover, two peptides which could not be sequenced by our method were also identified as His$^{2913}$-Lys$^{2966}$ in LRP2 (glycopeptide 9), and Val$^{510}$-Lys$^{554}$ in the meprin A β subunit precursor (endopeptidase-2, glycopeptides 18, 21 and 22). On the other hand, four glycopeptides that were sequenced by our method were not identified by the conventional method. Using both methods, we failed in the sequencing of four glycopeptides.

**Structural Analyses of the Oligosaccharides in the Le$^x$-Conjugated Glycopeptides.** The carbohydrate structures of the Le$^x$-conjugated glycopeptides were deduced from the fragment patterns and molecular masses obtained by the first run using FTICR-MS. The structural assignment of glycopeptide 8 is shown in Figure 4A and B. The carbohydrate composition was estimated to be 3dHex 5Hex 5HexNAc from the molecular mass of the carbohydrate moiety (calculated molecular mass of the glycan moiety: 2281.850). The fragment ions at *m/z* 1681.0 and *m/z* 1425.6 in Figure 4A were assigned to $Y_4^{2+}$ and $Y_{4/4}^{2+}$, which arose from [M + 2H]$^{2+}$ (*m/z* 1681.8) by the dissociation of two molecules of the Lewis-motifs. The presence of $B_2^+$ (*m/z* 512) in both Figure 4A and B also suggested the binding of two Lewis-motifs. The $Y_4^{2+}$ further yielded [Hex + 2HexNAc]$^+$ (*m/z* 569.1), $Y_{1'/3/3}^{2+}$ (*m/z* 1190.3) and $Y_{3/3}^{2+}$ (*m/z* 1263.2) on the MS/MS/MS, which suggested the presence of bisecting GlcNAc. The fucosylation of reducing-end GlcNAc was proven by the detection of $Y_{1/1'}^{2+}$ ([peptide + HexNAc + 2H]$^{2+}$, *m/z* 906.2) and $Y_1^{2+}$ ([peptide + dHex + HexNAc + 2H]$^{2+}$, *m/z* 979.0). Consequently, the glycan of glycopeptide 8 was characterized as a bisected and core-fucosylated oligosaccharide carrying two molecules of Le$^x$-motifs (Figure 4D). The possibility of the deduced structure was confirmed by the good agreement between the experimental mass (2281.850) and the theoretical mass (2281.845) (Table 1).

Figure 5A and B show the assignments of the carbohydrate moiety in the glycopeptide 15. The predominant ion (*m/z* 1414.4) in the MS/MS spectrum was assigned to [M – HexNAc + 3H]$^{3+}$ ($Y_3^{3+}$ or $Y_{4\beta}^{3+}$). This $Y_3^{3+}$ ($Y_{4\beta}^{3+}$) ion yielded the $B_2^+$ (*m/z* 512.3) by MS/MS/MS, suggesting the presence of only one molecule of the Lewis-motif. The presence of $Y_{1/1'}^{2+}$ (*m/z* 1346.3), $Y_{1'/2}^{2+}$ (*m/z* 1447.0) and $Y_2^{2+}$ (*m/z* 1520.2) suggested the fucosylation at the reducing end of GlcNAc. The presence of bisecting GlcNAc was deduced from the detection of the ion [Hex + 2HexNAc]$^+$ (*m/z* 569.3) and $Y_{1'/3\alpha/3\beta}^{2+}$ (*m/z* 1630.4). From these fragments, the oligosaccharide structure was characterized as a bisected and core-fucosylated oligosaccharide carrying one molecule of the Le$^x$-motif (Figure 5D).

The deduced carbohydrate structure of glycopeptide 17 is indicated in Figure 6D. In the MS/MS spectrum, the fragments at *m/z* 1515.4 and *m/z* 1393.5 were assigned to [M – dHex + 3H]$^{3+}$ ($Y_1^{3+}$ or $Y_5^{3+}$) and [M – dHex – Hex – HexNAc + 3H]$^{3+}$ ($Y_4^{3+}$), respectively (Figure 6A). The detection of $B_2^+$ (*m/z* 512.3) in both the MS/MS and MS/MS/MS spectra revealed the binding of two Le$^x$-motifs (Figure 6A and 6B). The presence of bisecting GlcNAc was suggested by the detection of the ion [Hex + 2HexNAc]$^+$ (*m/z* 569.2), $Y_{1/3/3}^{2+}$ (*m/z* 1599.7) and $Y_{3/3}^{2+}$ (*m/z* 1672.1) in the MS/MS/MS spectrum (Figure 6B). The ions $Y_{1/1'}^{2+}$ ([peptide + HexNAc + 2H]$^{2+}$, *m/z* 1315.0) and $Y_1^{2+}$ ([peptide + dHex + HexNAc + 2H]$^{2+}$, *m/z* 1388.1) revealed the fucosylation at the reducing end of GlcNAc. We also found the presence of a distinctive ion of the Lewis y (Le$^y$) motif, (Fucα1–2)Galβ1–4(Fucα1–3)GlcNAc, at *m/z* 658.3 in the MS/MS spectrum. To determine whether *N*-linked oligosaccharides contained the Le$^y$-motif, *N*-linked oligosaccharides were released from mouse kidney proteins and treated with α1–2 fucosidase. Then the glycan profiles of the fucosidase-treated and -untreated oligosaccharides were compared by LC/MS. No change was found in the mass spectrometric glycan profiles between the two samples, but the fragment (*m/z* 658) was still detected in the MS/MS spectra of the enzyme-treated

glycopeptide 17 (data not shown). These results suggest the absence of α1−2 fucose on the glycopeptides. Consequently, we assigned the glycans of glycopeptide 17 to a bisected and core-fucosylated oligosaccharide carrying two Le$^x$-motifs (Figure 6D).

The oligosaccharide structures of other Le$^x$-conjugated glycopeptides were deduced from their B- and Y-type ions as well as the molecular masses obtained by FTICR−MS in the same manner (Table 1 and Figure 7). The most common structure was a bisected and fucosylated complex-type biantennary oligosaccharide carrying two Le$^x$-motifs (glycopeptides 1−5, 8, 10, 12, 14, 17, 20 and 21). A bisected and core-fucosylated complex-type biantennary oligosaccharide carrying one Le$^x$-motif was found in glycopeptides 6, 13, and 15. A bisected and core-fucosylated complex-type triantennary oligosaccharide carrying three Le$^x$-motifs was found in glycopeptides 9 and 18. The oligosaccharide structure of the glycopeptide in glycopeptide 22 was a triantennary carrying two Le$^x$-motifs. All experimental molecular masses of the deduced glycopeptides were identical to their theoretical masses (Table 1).

## Discussion

Several glycan-epitopes, including Lewis antigens, HNK-1, and polysialic acid, have been widely shown to be involved in the physiological functions of glycoproteins and certain diseases. Some oligosaccharide-related antigens are being used as diagnostic markers of tumors in a clinical stage.[42,43] However, only a few proteins are known to carry the glycan-epitopes. To understand the physiological roles of the glycan-epitopes and to develop more effective diagnostic markers, we need methods that allow for the identification of target proteins carrying the glycan motif of interest. Glycan-epitopes are often detected by two-dimensional (2D)-electrophoresis in combination with lectin or immuno-blotting. The stained spots are subjected to in-gel tryptic digestion followed by protein identification by MS/MS and database search analysis. There are still problems in this procedure with the verification of the glycan structure in the identified protein. In addition, the procedure cannot be employed on hydrophobic membrane proteins having a high molecular weight.

In the present study, all proteins in the mouse kidney were digested into peptides, and the fucosylated glycopeptides were enriched by lectin-affinity chromatography. The resulting fucosylated glycopeptides were subjected to two different runs of LC−MS$^n$. In the first run, the elution positions of Le$^x$-conjugated glycopeptides in the tryptic peptide map were located based on the presence of Le$^x$-motif-distinctive ions. We picked out the product ion spectra of expected Le$^x$-conjugated glycopeptides from the elution positions and carefully assigned the peptide + HexNAc, peptide + (dHex)HexNAc, and peptide fragment. Then the fucosylated glycopeptides were subjected to a second run in which the peptide-related ions were set as precursor ions. We successfully identified γ-GTP1, LRP2, and the cubilin precursor as Le$^x$-conjugated glycoproteins by sequencing of 2−5 glycopeptides. Although only one glycopeptide was sequenced, cadherin 16, dipeptidase I, H2−K(k) and alanyl (membrane) aminopeptidase were characterized as Le$^x$-conjugated glycoproteins based on the good agreement between the experimental and theoretical masses of glycopeptides and their fragment patterns. Some of these were membrane proteins with high molecular masses over 400 kDa, the identification of which might have been difficult by 2D-electrophoresis with Western blotting.

Carbohydrate structures of the identified glycopeptides were deduced from the accurate molecular masses as well as fragment patterns obtained by the first run. We confirmed that all glycopeptides contained a bisected and core-fucosylated oligosaccharide carrying one or two molecules of Le$^x$-motifs at the *N*-linked oligosaccharide consensus sequence. Our model tissue was a mouse kidney in which we had previously confirmed the presence of Lewis x [Galβ1−4(Fucα1−3)GlcNAc] and/or y [(Fucα1−2)Galβ1−4(Fucα1−3)GlcNAc] motifs as well as the absence of Lewis a [Galβ1−3(Fucα1−4)GlcNAc] or b [(Fucα1−2)Galβ1−3(Fucα1−4)GlcNAc] motifs.[39] In this study, the Le$^y$-distinctive ions, (2dHex + Hex + HexNAc)$^+$ (*m*/*z* 658), were found in all MS/MS spectra of Lewis-conjugated peptides. However, treatment of α 1−2 fucosidase led to no change in the mass spectrometric glycan profile, suggesting the absence of the Le$^y$-motif. Recently, several groups have reported the internal migration of fucose residues in the ESI−CID of underived or derived carbohydrates.[44−48] Fucose residues are transferred between branches in liberated *N*-linked oligosaccharides by the ESI−CID.[49] Our finding suggests that the rearrangement of fucose residues also occurs by the ESI−CID of glycopeptides. This phenomenon makes it difficult to deduce the oligosaccharide structure from only the fragmentation pattern. A simultaneous use of lectins and/or antibodies would be crucial for the identification of the desired glycoproteins.

γ-Glutamyl transpeptidase 1 is associated with glutathione salvage, metabolism of endogenous mediators such as leukotrienes and prostaglandins. The attachment of Le$^x$-conjugated oligosaccharide to mouse γ-GTP 1 has already been demonstrated by Yamashita et al.[51] They determined the carbohydrate structures by the purification of γ-GTP 1 and the sequential exoglycosidase digestion in combination with methylation analysis. The oligosaccharide structures deduced from the MS/MS and MS/MS/MS spectra were in good agreement with those they reported. Furthermore, we revealed the heterogeneity of glycosylation on Asn$^{343}$.

Dipeptidase 1 is a glycosylphosphatidylinositol-anchored membrane glycoprotein. This protein is highly expressed in the kidney and small intestine and plays an important role in the degradation of cysteinyl-glycine, a glutathione produced by the removal of the glutamyl group from γ-glutamyl cysteinyl-glycine by γ-GTP.[52] The present study is the first report on the oligosaccharide structures of a mouse renal dipeptidase.

Cubilin, which is highly expressed in the renal proximal tubules, is a 460 kDa membrane glycoprotein consisting of 27 CUB (complement components C1r/C1s, Uegf, and bone morphogenic protein-1) domains. Cubilin is an endocytic receptor for intrinsic factor vitamin B12, albumin, apolipoprotein A-I, receptor-associated protein, immune globulin light chain and high-density lipoprotein.[50] These factors bind to cubilin through their CUB domains. The Le$^x$-conjugated oligosaccharides we found were all located on Asn$^{1802}$ and Asn$^{1819}$ in the CUB12 domain (Figure 8).

Low-density lipoprotein receptor-related protein 2, a high molecular weight membrane protein (520 kDa), is an endocytic receptor for several ligands, vitamin-binding proteins, apolipoproteins, hormones and enzymes. Cubilin and LRP2 are coexpressed in the renal proximal tubules and are associated with tubular protein reabsorption, vitamin metabolism and calcium homeostasis. Low-density lipoprotein receptor-related protein 2 consists of four ligand-binding sites containing cysteine-rich complement-type repeats and epidermal growth

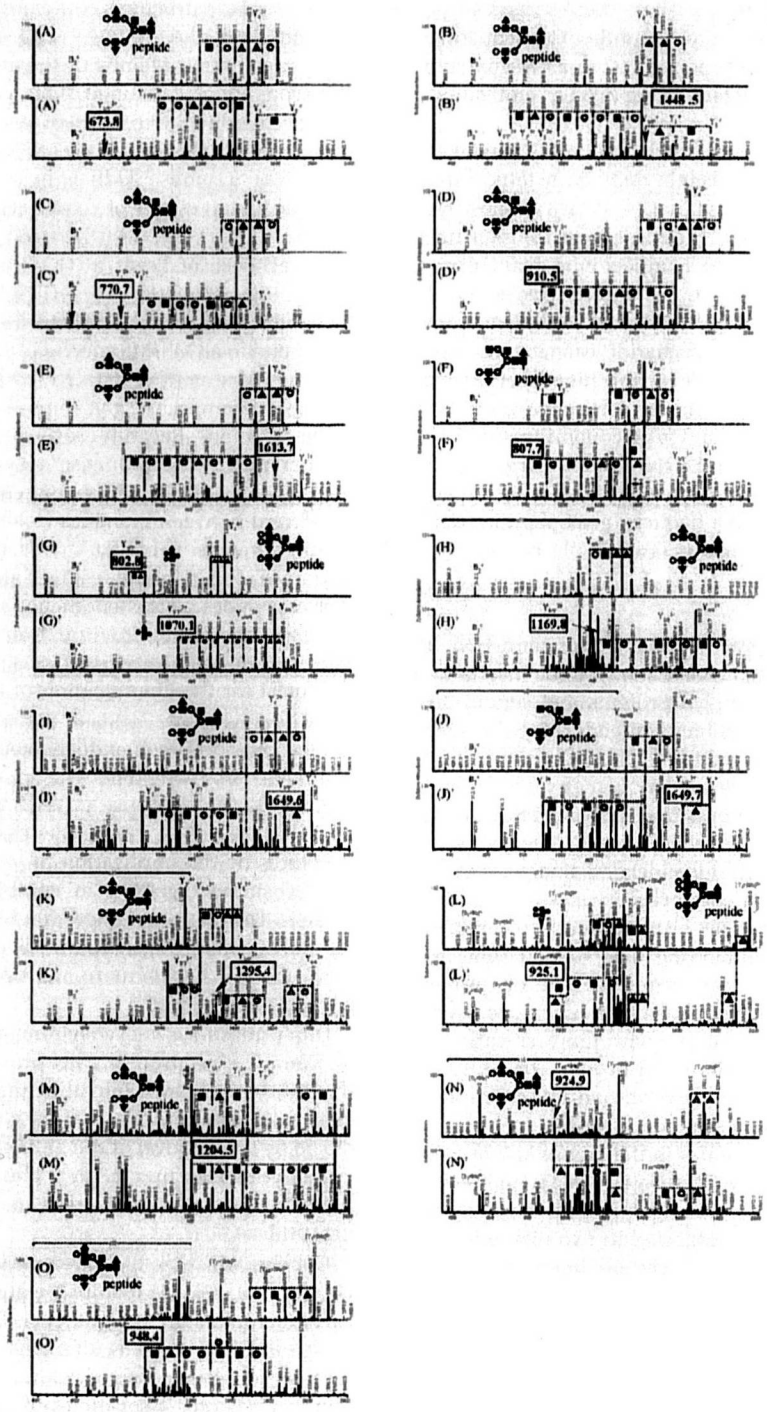


**Figure 7.** MS/MS and the MS/MS/MS spectra of glycopeptides 1−6, 9, 10, 12−14, 18 and 20−22, and deduced oligosaccharide structures. Boxed values are peptide-related ions. (A) MS/MS spectrum of the molecular ion (*m/z* 1136.4) of glycopeptide 1. (A′) MS/MS/MS spectrum of the predominant ion (*m/z* 1448.2) of (A). (B) MS/MS spectrum of glycopeptide 2 (*m/z* 1170.8). (B′) MS/MS/MS spectrum of the ion (*m/z* 1499.3) in (B). (C) MS/MS spectrum of glycopeptide 3(*m/z* 1152.5). (C′) MS/MS/MS spectrum of the ion (*m/z* 1472.5) in (C). (D) MS/MS spectrum of glycopeptide 4 (*m/z* 1294.2). (D′) MS/MS/MS spectrum of the ion (*m/z* 1685.0) in (D). (E) MS/MS spectrum of glycopeptide 5 (*m/z* 1226.0). (E′) MS/MS/MS spectrum of the ion (*m/z* 1582.6) in (E). (F) MS/MS spectrum of glycopeptide 6(*m/z* 1123.4). (F′) MS/MS/MS spectrum of the ion (*m/z* 1428.4) in (F). (G) MS/MS spectrum of glycopeptide 9 (*m/z* 1156.9). (G′) MS/MS/MS spectrum of the ion (*m/z* 1318.5) in (G). (H) MS/MS spectrum of glycopeptide 10 (*m/z* 1100.8). (H′) MS/MS/MS spectrum of the ion (*m/z* 1297.1) in (H). (I) MS/MS spectrum of glycopeptide 12 (*m/z* 1238.1). (I′) MS/MS/MS spectrum of the ion (*m/z* 1673.6) in (I). (J) MS/MS spectrum of glycopeptide 13 (*m/z* 1135.5). (J′) MS/MS/MS spectrum of the ion (*m/z* 1600.3) in (J). (K) MS/MS spectrum of glycopeptide 14 (*m/z* 1163.5). (K′) MS/MS/MS spectrum of the predominant ion (*m/z* 1380.8) in (K). (L) MS/MS spectrum of glycopeptide 18 (*m/z* 1482.6). (L′) MS/MS/MS spectrum of the predominant ion (*m/z* 1433.9) in (L). (M) MS/MS spectrum of glycopeptide 20 (*m/z* 1558.1). (M′) MS/MS/MS spectrum of the predominant ion (*m/z* 1509.8) in (M). (N) MS/MS spectrum of glycopeptide 21 (*m/z* 1279.2). (N′) MS/MS/MS spectrum of the predominant ion (*m/z* 1263.5) in (N). (O) MS/MS spectrum of glycopeptide 22 (*m/z* 1387.5). (O′) MS/MS/MS spectrum of the predominant ion (*m/z* 1825.5) in (O). White circle, galactose; gray circle, mannose; black square, *N*-acetylglucosamine; gray triangle, fucose.

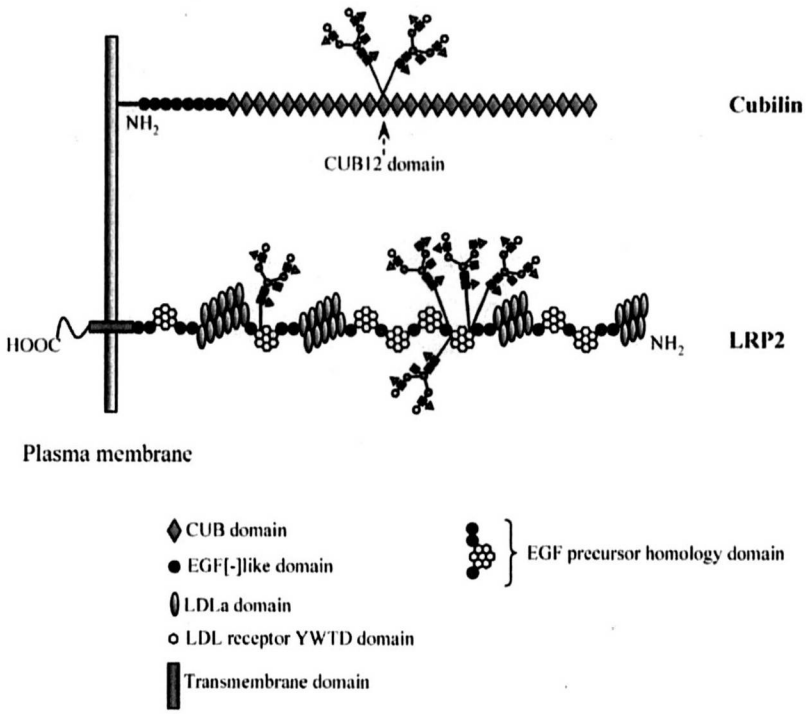


**Figure 8.** Location of Le$^x$-conjugated oligosaccharides on cubilin and LRP2. CUB, C1r/C1s, Uegf, and bone morphogenic protein-1; LDLa, Low-density lipoprotein receptor domain class A.



factor (EGF) precursor homology domains, which are associated with pH-dependent ligand dissociation. A previous study demonstrated the linkage of high-mannose-type and bi- and triantennary complex-type oligosaccharides bearing core fucose and bisecting GlcNAc. But there have been no reports on the presence of Le$^x$-conjugated oligosaccharides in rat kidney LRP2.[53] Interestingly, all of the Le$^x$-conjugated oligosaccharides we found were located in the EGF precursor homology domains (Asn$^{1497}$, Asn$^{1676}$, Asn$^{1733}$ and Asn$^{3448}$). Furthermore, heterogeneity of glycosylation on Asn$^{1497}$ was observed (Figure 8).

The cadherin 16, H2−K(k) protein and alanyl (membrane) aminopeptidase (aminopeptidase N) were also identified as Le$^x$-conjugated glycoproteins. Cadherin 16 is a kidney-specific cadherin that is associated with $Ca^{2+}$-dependent cell−cell adhesion.[54] The H2−K(k) protein, which is a mouse-specific histocompatibility antigen (H-2 antigen), is involved in the presentation of foreign antigens to the immune system. Alanyl (membrane) aminopeptidase (aminopeptidase N) is a transmembrane protein that is expressed predominantly in intestinal mucosa and kidney tissue.[55,56] It was reported that this enzyme is involved in several biological events such as tumorigenesis and immune system.[57] These proteins were unknown to be Le$^x$-conjugated proteins.

Using the present method, we successfully identified 14 Le$^x$-conjugated glycopeptides (12 peptides). Some peptides were found to be glycosylated with different Le$^x$-conjugated oligosaccharides (glycopeptides 5 and 6; glycopeptides 12 and 13). In most cases only a peptide carrying a major oligosaccharide was identified as a Le$^x$-conjugated glycopeptide. Minor Le$^x$-conjugated glycopeptides were not subjected to MS/MS/MS in the first run because they were less intense. Such minor glycopeptides might be identified by an additional run in which the glycopeptides identified in the first run are excluded. In addition, our method tended to fail in the identification of glycopeptides having high molecular mass (>4500 Da), glycopeptides detected as triply charged ions, and glycopeptides containing triantennary oligosaccharides. These glycopeptides yielded a smaller number of peptide-related ions, which was insufficient for further CID, and database search analysis resulted in false-positive proteins, for example, a peptide not containing *N*-glycan consensus sequences. Using a conventional approach that included LC−MS/MS of the PNGase F-treated tryptic digest, we found two additional Le$^x$-expected peptides. One was a high molecular mass peptide (3002 Da) containing a triantennary oligosaccharide, and the other was a peptide detected as sodium adducts in our method. The conventional approach has the advantage of peptide sequencing, but does not allow confirmation of the Lewis-motif in the oligosaccharide. Unlike the classical glycomic approaches that are used for the comprehensive analysis of glycopeptides, our method focused only on Lewis-conjugated glycopeptides. Accordingly, it could be applicable to the identification and screening of glycoproteins carrying target glycan-motifs.

**Acknowledgment.** This study was supported in part by a Grant-in-Aid from the Ministry of Health, Labor, and Welfare, by the Core Research for the Evolutional Science and Technology Program (CREST), and by the Japan Science and Technology Corp (JST).

## References

(1) Dwek, R. A. Glycobiology: Toward Understanding the Function of Sugars. *Chem. Rev.* **1996**, *96*, 683–720.
(2) Helenius, A.; Aebi, M. Intracellular functions of N-linked glycans. *Science* **2001**, *291*, 2364–2369.
(3) Zak, I.; Lewandowska, E.; Gnyp, W. Selectin glycoprotein ligands. *Acta Biochim. Pol.* **2000**, *47*, 393–412.
(4) Feizi, T.; Gooi, H. C.; Childs, R. A.; Picard, J. K.; Uemura, K.; Loomes, L. M.; Thorpe, S. J.; Hounsell, E. F. Tumour-associated and differentiation antigens on the carbohydrate moieties of mucin-type glycoproteins. *Biochem. Soc. Trans.* **1984**, *12*, 591–596.

(5) Kannagi, R.; Izawa, M.; Koike, T.; Miyazaki, K.; Kimura, N. Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. *Cancer Sci.* **2004**, *95*, 377–384.
(6) Hada, T.; Kondo, M.; Yasukawa, K.; Amuro, Y.; Higashino, K. Discrimination of liver cirrhosis from chronic hepatitis by measuring the ratio of Aleuria aurantia lectin-reactive serum cholinesterase to immunoreactive protein. *Clin. Chim. Acta* **1999**, *281*, 37–46.
(7) Sutton-Smith, M.; Morris, H. R.; Grewal, P. K.; Hewitt, J. E.; Bittner, R. E.; Goldin, E.; Schiffmann, R.; Dell, A. MS screening strategies: investigating the glycomes of knockout and myodystrophic mice and leukodystrophic human brains. *Biochem. Soc. Symp.* **2002**, 105–115.
(8) Taniguchi, N.; Ekuni, A.; Ko, J. H.; Miyoshi, E.; Ikeda, Y.; Ihara, Y.; Nishikawa, A.; Honke, K.; Takahashi, M. A glycomic approach to the identification and characterization of glycoprotein function in cells transfected with glycosyltransferase genes. *Proteomics* **2001**, *1*, 239–247.
(9) Morelle, W.; Michalski, J. C. Glycomics and mass spectrometry. *Curr. Pharm. Des.* **2005**, *11*, 2615–2645.
(10) Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J. C. The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics* **2006**, *6*, 3993–4015.
(11) Hagglund, P.; Matthiesen, R.; Elortza, F.; Hojrup, P.; Roepstorff, P.; Jensen, O. N.; Bunkenborg, J. An enzymatic deglycosylation scheme enabling identification of core fucosylated N-glycans and O-glycosylation site mapping of human plasma proteins. *J. Proteome Res.* **2007**, *6*, 3021–3031.
(12) Alvarez-Manilla, G.; Atwood, J., 3rd; Guo, Y.; Warren, N. L.; Orlando, R.; Pierce, M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. *J. Proteome Res.* **2006**, *5*, 701–708.
(13) Zhao, J.; Qiu, W.; Simeone, D. M.; Lubman, D. M. N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. *J. Proteome Res.* **2007**, *6*, 1126–1138.
(14) Kyselova, Z.; Mechref, Y.; Al Bataineh, M. M.; Dobrolecki, L. E.; Hickey, R. J.; Vinson, J.; Sweeney, C. J.; Novotny, M. V. Alterations in the serum glycome due to metastatic prostate cancer. *J. Proteome Res.* **2007**, *6*, 1822–1832.
(15) Zamfir, A.; Seidler, D. G.; Schonherr, E.; Kresse, H.; Peter-Katalinic, J. On-line sheathless capillary electrophoresis/nanoelectrospray ionization-tandem mass spectrometry for the analysis of glycosaminoglycan oligosaccharides. *Electrophoresis* **2004**, *25*, 2010–2016.
(16) Karlsson, N. G.; Wilson, N. L.; Wirth, H. J.; Dawes, P.; Joshi, H.; Packer, N. H. Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2282–2292.
(17) Kawasaki, N.; Ohta, M.; Itoh, S.; Hyuga, M.; Hyuga, S.; Hayakawa, T. Usefulness of sugar mapping by liquid chromatography/mass spectrometry in comparability assessments of glycoprotein products. *Biologicals* **2002**, *30*, 113–123.
(18) Kremmer, T.; Szollosi, E.; Boldizsar, M.; Vincze, B.; Ludanyi, K.; Imre, T.; Schlosser, G.; Vekey, K. Liquid chromatographic and mass spectrometric analysis of human serum acid alpha-1-glycoprotein. *Biomed. Chromatogr.* **2004**, *18*, 323–329.
(19) Thomsson, K. A.; Karlsson, H.; Hansson, G. C. Sequencing of sulfated oligosaccharides from mucins by liquid chromatography and electrospray ionization tandem mass spectrometry. *Anal. Chem.* **2000**, *72*, 4543–4549.
(20) Kawasaki, N.; Ohta, M.; Hyuga, S.; Hyuga, M.; Hayakawa, T. Application of liquid chromatography/mass spectrometry and liquid chromatography with tandem mass spectrometry to the analysis of the site-specific carbohydrate heterogeneity in erythropoietin. *Anal. Biochem.* **2000**, *285*, 82–91.
(21) Schmid, D.; Behnke, B.; Metzger, J.; Kuhn, R. Nano-HPLC-mass spectrometry and MEKC for the analysis of oligosaccharides from human milk. *Biomed. Chromatogr.* **2002**, *16*, 151–156.
(22) Maslen, S.; Sadowski, P.; Adam, A.; Lilley, K.; Stephens, E. Differentiation of isomeric N-glycan structures by normal-phase liquid chromatography-MALDI-TOF/TOF tandem mass spectrometry. *Anal. Chem.* **2006**, *78*, 8491–8498.
(23) Yuan, J.; Hashii, N.; Kawasaki, N.; Itoh, S.; Kawanishi, T.; Hayakawa, T. Isotope tag method for quantitative analysis of carbohydrates by liquid chromatography-mass spectrometry. *J. Chromatogr. A* **2005**, *1067*, 145–152.
(24) Lawrence, R.; Olson, S. K.; Steele, R. E.; Wang, L.; Warrior, R.; Cummings, R. D.; Esko, J. D. Evolutionary differences in glycosaminoglycan fine structure detected by quantitative glycan reductive isotope labeling. *J. Biol. Chem.* **2008**, *283*, 33674–33684.
(25) Ridlova, G.; Mortimer, J. C.; Maslen, S. L.; Dupree, P.; Stephens, E. Oligosaccharide relative quantitation using isotope tagging and normal-phase liquid chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 2723–2730.
(26) Harazono, A.; Kawasaki, N.; Kawanishi, T.; Hayakawa, T. Site-specific glycosylation analysis of human apolipoprotein B100 using LC/ESI MS/MS. *Glycobiology* **2005**, *15*, 447–462.
(27) Itoh, S.; Kawasaki, N.; Harazono, A.; Hashii, N.; Matsuishi, Y.; Kawanishi, T.; Hayakawa, T. Characterization of a gel-separated unknown glycoprotein by liquid chromatography/multistage tandem mass spectrometry: analysis of rat brain Thy-1 separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Chromatogr. A* **2005**, *1094*, 105–117.
(28) Satomi, Y.; Shimonishi, Y.; Hase, T.; Takao, T. Site-specific carbohydrate profiling of human transferrin by nano-flow liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2983–2988.
(29) Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H. Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J. Chromatogr., B: Analyt. Technol. Biomed. Life Sci.* **2007**, *849*, 115–128.
(30) Stadlmann, J.; Pabst, M.; Kolarich, D.; Kunert, R.; Altmann, F. Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics* **2008**, *8*, 2858–2871.
(31) Picariello, G.; Ferranti, P.; Mamone, G.; Roepstorff, P.; Addeo, F. Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics* **2008**, *8*, 3833–3847.
(32) Hirabayashi, J. Lectin-based structural glycomics: glycoproteomics and glycan profiling. *Glycoconj. J.* **2004**, *21*, 35–40.
(33) Taketa, K. Characterization of sugar chain structures of human alpha-fetoprotein by lectin affinity electrophoresis. ibarahp@oka.urban.ne.jp. *Electrophoresis* **1998**, *19*, 2595–2602.
(34) Kuno, A.; Kato, Y.; Matsuda, A.; Kaneko, M. K.; Ito, H.; Amano, K.; Chiba, Y.; Narimatsu, H.; Hirabayashi, J. Focused Differential Glycan Analysis with the Platform Antibody-assisted Lectin Profiling for Glycan-related Biomarker Verification. *Mol. Cell. Proteomics* **2009**, *8*, 99–108.
(35) Madera, M.; Mechref, Y.; Klouckova, I.; Novotny, M. V. Semiautomated high-sensitivity profiling of human blood serum glycoproteins through lectin preconcentration and multidimensional chromatography/tandem mass spectrometry. *J. Proteome Res.* **2006**, *5*, 2348–2363.
(36) Yang, Z.; Hancock, W. S.; Chew, T. R.; Bonilla, L. A study of glycoproteins in human serum and plasma reference standards (HUPO) using multilectin affinity chromatography coupled with RPLC-MS/MS. *Proteomics* **2005**, *5*, 3353–3366.
(37) Zhao, J.; Simeone, D. M.; Heidt, D.; Anderson, M. A.; Lubman, D. M. Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. *J. Proteome Res.* **2006**, *5*, 1792–1802.
(38) Hashii, N.; Kawasaki, N.; Itoh, S.; Hyuga, M.; Kawanishi, T.; Hayakawa, T. Glycomic/glycoproteomic analysis by liquid chromatography/mass spectrometry: analysis of glycan structural alteration in cells. *Proteomics* **2005**, *5*, 4665–4672.
(39) Hashii, N.; Kawasaki, N.; Itoh, S.; Harazono, A.; Matsuishi, Y.; Hayakawa, T.; Kawanishi, T. Specific detection of Lewis x-carbohydrates in biological samples using liquid chromatography/multiple-stage tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 3315–3321.
(40) Carr, S. A.; Huddleston, M. J.; Bean, M. F. Selective identification and differentiation of N- and O-linked oligosaccharides in glycoproteins by liquid chromatography-mass spectrometry. *Protein Sci.* **1993**, *2*, 183–196.
(41) Harazono, A.; Kawasaki, N.; Itoh, S.; Hashii, N.; Ishii-Watabe, A.; Kawanishi, T.; Hayakawa, T. Site-specific N-glycosylation analysis of human plasma ceruloplasmin using liquid chromatography with electrospray ionization tandem mass spectrometry. *Anal. Biochem.* **2006**, *348*, 259–268.
(42) Ueda, T.; Shimada, E.; Urakawa, T. The clinicopathologic features of serum CA 19–9-positive colorectal cancers. *Surg. Today* **1994**, *24*, 518–525.
(43) Nakagoe, T.; Sawai, T.; Tsuji, T.; Jibiki, M.; Ohbatake, M.; Nanashima, A.; Yamaguchi, H.; Yasutake, T.; Ayabe, H.; Tagawa, Y. Differences in release mechanisms and distributions for sialyl Le(a) and sialyl Le(x) antigens in colorectal cancer. *Ann. Surg. Oncol.* **2000**, *7*, 289–295.

(44) Harvey, D. J.; Mattu, T. S.; Wormald, M. R.; Royle, L.; Dwek, R. A.; Rudd, P. M. "Internal residue loss": rearrangements occurring during the fragmentation of carbohydrates derivatized at the reducing terminus. *Anal. Chem.* **2002**, *74*, 734–740.

(45) Mattu, T. S.; Royle, L.; Langridge, J.; Wormald, M. R.; Van den Steen, P. E.; Van Damme, J.; Opdenakker, G.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M. O-glycan analysis of natural human neutrophil gelatinase B using a combination of normal phase-HPLC and online tandem mass spectrometry: implications for the domain organization of the enzyme. *Biochemistry* **2000**, *39*, 15695–15704.

(46) Brull, L. P.; Kovacik, V.; Thomas-Oates, J. E.; Heerma, W.; Haverkamp, J. Sodium-cationized oligosaccharides do not appear to undergo 'internal residue loss' rearrangement processes on tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1520–1532.

(47) Franz, A. H.; Lebrilla, C. B. Evidence for long-range glycosyl transfer reactions in the gas phase. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 325–337.

(48) Ma, Y. L.; Vedernikova, I.; Van den Heuvel, H.; Claeys, M. Internal glucose residue loss in protonated O-diglycosyl flavonoids upon low-energy collision-induced dissociation. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 136–144.

(49) Wuhrer, M.; Koeleman, C. A.; Hokke, C. H.; Deelder, A. M. Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1747–1754.

(50) Christensen, E. I.; Birn, H. Megalin and cubilin: synergistic endocytic receptors in renal proximal tubule. *Am. J. Physiol. Renal Physiol.* **2001**, *280*, F562–573.

(51) Yamashita, K.; Hitoi, A.; Tateishi, N.; Higashi, T.; Sakamoto, Y.; Kobata, A. The structures of the carbohydrate moieties of mouse kidney gamma-glutamyltranspeptidase: occurrence of X-antigenic determinants and bisecting N-acetylglucosamine residues. *Arch. Biochem. Biophys.* **1985**, *240*, 573–582.

(52) Habib, G. M.; Shi, Z. Z.; Cuevas, A. A.; Lieberman, M. W. Identification of two additional members of the membrane-bound dipeptidase family. *Faseb J.* **2003**, *17*, 1313–1315.

(53) Morelle, W.; Haslam, S. M.; Ziak, M.; Roth, J.; Morris, H. R.; Dell, A. Characterization of the N-linked oligosaccharides of megalin (gp330) from rat kidney. *Glycobiology* **2000**, *10*, 295–304.

(54) Wendeler, M. W.; Praus, M.; Jung, R.; Hecking, M.; Metzig, C.; Gessner, R. Ksp-cadherin is a functional cell-cell adhesion molecule related to LI-cadherin. *Exp. Cell Res.* **2004**, *294*, 345–355.

(55) Barnes, K.; Walkden, B. J.; Wilkinson, T. C.; Turner, A. J. Expression of endothelin-converting enzyme in both neuroblastoma and glial cell lines and its localization in rat hippocampus. *J. Neurochem.* **1997**, *68*, 570–577.

(56) Lucius, R.; Sievers, J.; Mentlein, R. Enkephalin metabolism by microglial aminopeptidase N (CD13). *J. Neurochem.* **1995**, *64*, 1841–1847.

(57) Luan, Y.; Xu, W. The structure and main functions of aminopeptidase N. *Curr. Med. Chem.* **2007**, *14*, 639–647.

PR9000527

Current Topics

# New Era of Glycoscience: Intrinsic and Extrinsic Functions Performed by Glycans

## The Significance of Glycosylation Analysis in Development of Biopharmaceuticals

Nana KAWASAKI,* Satsuki ITOH, Noritaka HASHII, Daisuke TAKAKURA, Yan QIN, Xiaoyu HUANG, and Teruhide YAMAGUCHII

*Division of Biological Chemistry & Biologicals, National Institute of Health Sciences; 1–18–1 Kamiyoga, Setagaya-ku, Tokyo 158–8501, Japan.*

Received October 24, 2008

**Many glycoproteins and glycosaminoglycans are approved for clinical use. Carbohydrate moieties in biopharmaceuticals affect not only their physicochemical properties and thermal stability, but also their reactivity with their receptors and circulating half-life. Modification of glycans is one target of drug design for enhancement of efficacy. Meanwhile, there have been reports of serious adverse events caused by some carbohydrates. It is crucial to maintain the constancy of carbohydrate moieties for the efficient and safe use of glycosylated biopharmaceuticals. On the other hand, for scientific, safety-related, and economic reasons, changes in the manufacturing process are frequently made either during the development or after the approval of new biopharmaceuticals. Furthermore, the development of biosimilar glycoprotein products has been attempted by different manufacturers. Changes in pharmaceutical manufacturing processes possibly cause alteration of glycosylation and raise concerns about alteration of their quality, safety, and efficacy. In this review we provide some current topics of glycosylated biopharmaceuticals from the viewpoints of efficacy, safety, and the manufacturing process and discuss the significance of glycosylation analysis for development of biopharmaceuticals.**

Key words glycoprotein; glycosaminoglycan; biopharmaceutical; efficacy; safety; manufacturing process

## 1. INTRODUCTION

Many biological molecules containing carbohydrates are approved for clinical use in Japan (Table 1). In the beginning, most therapeutic glycoconjugates were naturally occurring glycoproteins and glycosaminoglycans (GAGs) derived from human and healthy animal tissues, such as gonadotropins from human urine, and heparins from the porcine intestine. Recombinant glycoproteins, including erythropoietin and tissue plasminogen activator (tPA), have been developed as copies of native human glycoproteins since the early 1990s in Japan. In many cases their carbohydrate moieties were different from the original human ones. Carbohydrate moieties in glycoproteins affects not only their physicochemical properties and thermal stability but also their reactivity with their receptors and circulating half-life.[1] Modification of carbohydrate moieties is one target of drug design to enhance the efficacy of the original ones. Meanwhile, there have been reports from around the world of serious adverse events caused by carbohydrate-related drugs. For the efficient and safe use of glycoprotein/GAG products, it is necessary to maintain the structures and heterogeneity of carbohydrate moieties in glycosylated biopharmaceuticals.

On the other hand, carbohydrate moieties in biotechnology-derived drugs are variable when the manufacturing process is changed. For scientific, safety-related and economic reasons, it has become common for companies to change the manufacturing process for their approved products. Furthermore, biosimilar glycoprotein products, which are manufactured by different companies, have been developing in many regions.[2,3] One of the main issues for the devel-

Table 1. Glycosylated Biopharmaceuticals in Japan

| Origin | Japanese accepted name |
|---|---|
| Granulocyte-colony stimulating factor | Lenograstim |
| Granulocyte macrophage colony-stimulating factor | Mirimostim |
| Interferon | Interferon Alfa (NAMALWA), Interferon Alfa (BALL-1), Interferon Beta, Interferon Beta-1a, Interferon Gamma-n1 |
| Erythropoetin | Epoetin Alfa, Epoetin Beta, Darbepoetin Alfa |
| Monoclonal antibody | Ibritumomab Tiuxetan, Basiliximab, Infliximab, Rituximab, Cetuximab, Gemtuzumab Ozogamicin, Palivizumab, Tocilizumab, Trastuzumab , Bevacizumab, Adalimumab |
| Receptor | Etanercept |
| Follicle stimulating hormone | Follitropin Alfa, Follitoropin Beta |
| Gonadotropin | Human Menopausal Gonadotropin, Human Chorionic Gonadotropin, Serum Gonadotropin |
| Factor VII | Eptacog Alfa (Activated) |
| Factor VIII | Octocog Alfa, Rurioctocog Alfa |
| Thrombomodulin | Thrombomodulin Alfa |
| Urokinase | Urokinase |
| Pro-urokinase | Nasaruplase |
| Tissue-plasminogen activator | Alteplase, Monteplase, Pamiteplase |
| Enzymes | Kallidinogenase, Agalsidase Alfa, Agalsidase Beta, Alglucosidase Alfa, Alglucerase, Idursulfase, Imiglucerase, Laronidase, Galsulfase |
| Heparins | Heparin Sodium, Heparin Calcium, Parnaparin Sodium, Dalteparin Sodium, Enoxaparin Sodium, Reviparin Sodium |
| Hyaruronate | Sodium Hyaruronate |
| Chondroitin sulfate | Chondroitin sulfate |

* To whom correspondence should be addressed. e-mail: nana@nihs.go.jp

opment of biosimilar products is how to assure the similarity of carbohydrate moieties between the biosimilar products and the reference products.

In this review we provide some current topics of glycosylated biopharmaceuticals in terms of efficacy, safety, and manufacturing process and discuss the significance of glycosylation analysis in the development of biopharmaceuticals.

## 2. ROLE OF CARBOHYDRATES ON EFFICACY

Glycosylation in some biopharmaceuticals is crucial for their biological activity. Lysosomal storage diseases are characterized by deficiencies of lysosomal enzymes that degrade the glycoconjugates, such as mucopolysaccharides and glycolipids, and consequent cellular damages by their accumulated metabolites.[4] For use in enzyme-replacement therapy against lysosomal storage diseases, several recombinant glycoprotein products have been approved, namely *agalsidase alfa* and *agalsidase beta* for Fabry's disease, *alglucosidase alfa* for Pompe's disease, and *laronidase*, *idursulfase*, and *galsulfase* for mucopolysaccharidosis I, II, and VI, respectively.[4-8] These drugs contain *N*-linked oligosaccharides attached to mannose 6-phosphate (M-6-P), and the secreted enzymes are transported to an acidified prelysosomal compartment through the M-6-P receptor.[9,10] The carbohydrate residue is essential for lysosomal targeting and to exhibit complete enzyme activity in lysosomes. *Imiglucerase* is an analog of human β-glucocerebrosidase, which is used for the treatment of Gaucher's disease. This glycoprotein is produced by recombinant DNA technology and exoglycosidase treatment to expose mannose residues in *N*-linked oligosaccharides. The carbohydrate modification facilitates incorporation of this drug into macrophages through mannose-binding receptors.[11] These recombinant lysosomal enzymes have achieved dramatic therapeutic effects against the lysosomal storage diseases.

Several human glycoprotein analogs whose carbohydrates are modified to enhance their efficacy have been developed in recent years. Erythropoietin is a glycoprotein containing three *N*-glycans and one *O*-glycan, and sialylation on its nonreducing ends is closely associated with its circulating half-life.[12] *Darbepoetin alfa* is an erythropoietin analog to which two additional *N*-glycans are attached by replacement of five amino acid residues.[13] The modification of glycosylation prolongs the half-time of this analog compared to its native form. Similar genetic engineering for improvement of half lives has been successfully attempted in antithrombolytic drugs. T-PA consists of finger, epidermal growth factor (EGF), kringle1, kringle2 and catalytic domains, and three *N*-glycans. High-mannose type oligosaccharides at the kringle1 domain and EGF domain are related to the short circulation half-life of t-PA.[14] Extension of half-life in blood has been achieved by eliminating the kringle1 domain from human t-PA and replacing one amino acid residue (*pamiteplase*).[15] These improvements have contributed to reducing the frequency and dose of administration.

Recombinant monoclonal antibodies, which have been successfully used in the treatment of cancers and immune diseases, might be the next target of drug design by glyco-engineering. Several anti-tumor monoclonal antibodies that contain a constant region derived from immunoglobulin (Ig) G have antibody-dependent cellular cytotoxicity (ADCC), and removal of a fucose (Fuc) residue from *N*-linked oligosaccharides at the constant region causes the enhancement of ADCC.[16,17] A non-fucosylated IgG-derived antibody is expected to improve the therapeutic effects of these antitumor pharmaceuticals.

Glycosylation has also been used for the site-selective modification of proteins with polyethylene glycol (PEGylation).[18] In the GlycoPEGylation method, sialic acid covalently substituted with polyethylene glycol (PEG) can be enzymatically transferred to *O*-glycans at serine or threonine positions in proteins produced in *Escherichia coli*. This strategy has overcome the problems of the previous PEGylation, which provided a heterogeneous mixture of PEG positional isomers. There is increasing interest in utilization and modification of glycans in the development of biopharmaceuticals.

## 3. IMPACT OF CARBOHYDRATES ON SAFETY

Some glycans have caused serious adverse events in clinical stages. Heparin is a highly sulfated GAG composed of a disaccharide unit, β1-4 linked α-D-glucosamine and α-D-iduronic acid or β-D-glucuronic acid (averaging 2.5 sulfate groups per disaccharide). In 2007—2008, a serious adverse event associated with *heparin sodium*, including over eighty deaths, occurred in the United States (US).[19] Over-sulfated chondroitin sulfate (OSCS), which consists of fully sulfated β1-3 linked α-D-*N*-acetylgalactosamine (GalNAc) and a β-D-glucuronic acid unit, was identified as the contaminant in the heparin sodium that had caused hypersensitivity reactions.[20,21] It has been reported that OSCS activated the kinin-kallikrein system and induced generation of C3a and C5a, potent anaphylatoxins derived from complement proteins. The contaminated heparin sodium was distributed to at least twelve countries and raised concern about a shortage of heparin products. For ensuring the safety of heparin products, pharmacopoeias in Japan, the US and EU immediately amended their heparin sodium monograph to confirm the absence of OSCS by $^1$H-NMR and/or capillary electrophoresis (Fig. 1A). This adverse event has left concerns about the safety of GAGs and motivated the development of a sensitive and selective analytical method for GAGs products all over the globe (Fig. 1B).

An alternative concern for a carbohydrate-related adverse event is immunogenicity of nonhuman glycan motifs. The galactose-α1,3-galactose (Gal(α1-3)Gal) motif is known as one of the major problems in the transplantation of organs from pigs to humans.[22] Mouse myeloma cell lines, which are often used for production of recombinant glycoproteins, also expresses the Gal(α1-3)Gal motif in *N*-linked oligosaccharides. Recently, a high prevalence of hypersensitivity reactions was reported in patients who had been injected with *cetuximab*, which is produced in mouse myeloma cells.[23] This chimeric mouse-human antibody is attached to *N*-linked oligosaccharide containing the Gal(α1-3)Gal motif at the Fab region.[24] It is reported that IgE antibodies against the Gal(α1-3)Gal motif had been present in serum before therapy, and these antibodies caused hypersensitive reactions after cetuximab treatment. Another immunogenic nonhuman glycan is *N*-glycolylneuraminic acid (NeuGc), a sialic acid.[25]

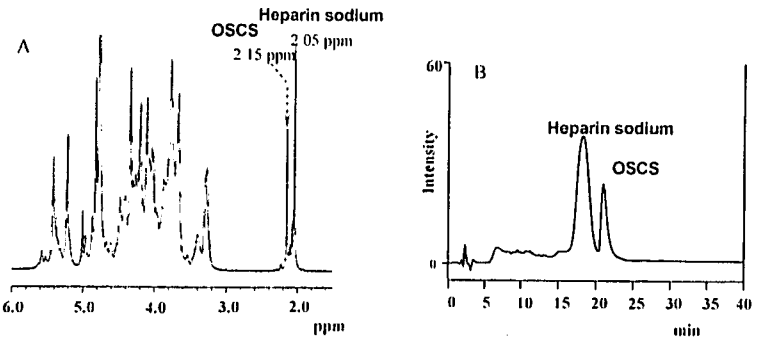


Fig. 1. $^1$H-NMR Spectra (A) and Anion-Exchange Chromatogram (B) of Heparin Sodium Containing 10% OSCS

$^1$H-NMR spectrum was obtained at 298 K using a 500 MHz JEOL JNM-ECA500 instrument equipped with a 5-mm filed gradient tunable probe with standard JEOL software. Chemical shifts were referenced to the signal of 3-(trimethylsilyl)propionic-2,2,3,3-$d_4$ acid sodium salt (TSP) as internal standard (0.00 ppm). Anion-exchange chromatogram was obtained using TSKgel DEAE-5PW (21.5 mm I.D.×15 cm, 10 $\mu$m) at a flow rate of 1.0 ml/min at 40 °C. The UV detector was set at 230 nm. GAGs were eluted by a linear gradient of 0→100% B buffer (buffer: A, 20 mM Tris-HCl (pH 8.0); B, 20 mM Tris-HCl (pH 8.0) containing 2 M NaCl) within 40 min.

Since Varki *et al.* reported the incorporation of NeuGc into human embryonic stem (ES) cells through the animal serum and the feeder layer, contamination of cell therapy products with NeuGc has become a serious issue in clinics.[26] These reports imply the significance of glycan analysis in recombinant glycoprotein products and risk assessment of nonhuman glycans.

## 4. ALTERATION OF CARBOHYDRATES BY CHANGES IN EXPRESSION SYSTEM AND MANUFACTURING PROCESS

In addition to recombinant technology using mammalian cells, a transgenic technique has been also focusing on large-scale production of therapeutic glycoproteins.[27,28] Antithrombin III is an anticoagulant factor, and a freeze-dried preparation of human antithrombin III has been approved for treatment of disseminated intravascular coagulation and thrombogenic tendencies. In 2006, the European Medicine Agency (EMEA) authorized the marketing of *antithrombin alfa*, a recombinant human antithrombin III produced from the milk of transgenic goats. Detailed structural analysis revealed that these two glycoproteins were structurally identical except for a difference in glycosylation.[27] An alternative approach for the large-scale and low-cost production of biopharmaceuticals is the use of transgenic plants.[29] Plant cells express nonhuman glycans, such as Fuc ($\alpha$1-3) *N*-acetylglucosamine (GlcNAc), xylose (Xyl) ($\beta$1-2) GlcNAc and Gal ($\beta$1-3) GlcNAc at the reducing-end of complex-type *N*-linked oligosaccharides. Currently, plant glyco-engineering, including modification of glycosyltransferases, is developing to overcome limitations in the production of glycoprotein products.[30]

For various scientific, safety-related and economic reasons, changes in manufacturing processes for biopharmaceuticals are often attempted during the development phase and after marketing authorization. Meanwhile recombinant glycoprotein products that have been claimed to be similar to a reference medical product already authorized (biogeneric/biosimilar/follow-on biologics) have been approved by different manufacturers. The changes in the manufacturing process possibly cause the alteration of glycosylation in the glycoprotein products and consequent changes in quality, efficacy and safety. The first biosimilar glycosylated pharmaceutical, *epoetin alfa*, was approved in the EU recently. According to the Japanese guidelines for the biosimilar products, the applicants have to submit some efficacy and safety study data but not all if they can show the data on structural properties and physicochemical/biological similarity between the authorized and biosimilar products. One of the challenging issues in the development of biosimilar glycoprotein products is a comparison of glycosylation between the authorized products and the new entry biosimilar products.

Using some *epoetin* products, we have studied the possibility of several analytical methods for comparison of the glycosylation between closely related biopharmaceuticals. Commercially available *epoetin* products (products A—D) were electrophoresed, and *N*-linked oligosaccharides were released from bands at 30 kDa by an in-gel glycopeptidase F digestion (Fig. 2A). The resulting oligosaccharides were reduced with $NaBH_4$ and subjected to LC/MS. In Fig. 2, *International nonproprietary names* (*INN*) of *epoetins* contained in the products A—C and D are tentatively named as *epoetin* $\alpha$ and $\beta$, respectively. Products A and B are marketed in country X, while products C and D are manufactured and distributed in country Y. Figure 2B shows the total ion chromatograms of *N*-linked oligosaccharides released from four *epoetin* products, and the mass spectra acquired from the most intense peaks (peak $z_{1\_4}$) are shown in Fig. 2C. The *m/z* values of the most intense ions (*m/z* 1226.8) and a series of triply charged ions with *m/z* 14 spacing pattern reveal that the most abundant glycan in *epoetin* products are a tetrasialyl fucosylated-tetraantennary oligosaccharide in common but there are tangible differences in acetylation of sialic acids between products A—C (*epoetin* $\alpha$) and product D (*epoetin* $\beta$) (Fig. 2C). Even among *epoetin* $\alpha$ products there were some significant differences, such as the non-fucosylated oligosaccharides that were found in *epoetin* $\alpha$ products in country X (products A and B) but not in country Y (product C). The latest analytical technology allows us to evaluate the similarity of the glycosylation of closely related biopharmaceuticals.[31,32]

## 5. CONCLUSION

As described above, glycosylation in most biopharmaceuticals affects their efficacy and safety, and the glycosylation is dependent on the manufacturing process and the expres-

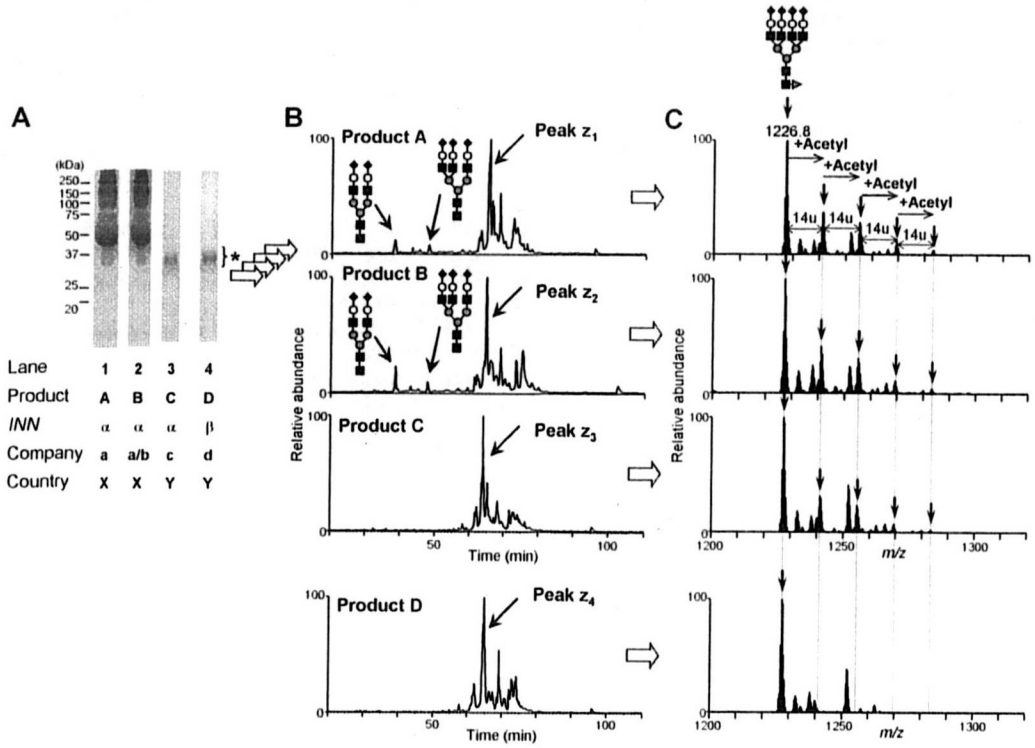


Fig. 2. Glycosylation Analysis of *Epoetin* Products

(A) SDS-PAGE images. Sample: lane 1, product A; lane 2, product B; lane 3, product C; lane 4, product D. (B) *N*-glycan profiles of products A—D acquired by LC/MS in the negative ion mode. (C) Mass spectra of peaks $z_{1-4}$. Product: A, an *epoetin* α product manufactured/distributed by company a in country X; B, an *epoetin* α product manufactured by company a and distributed by company b in country X; C, an *epoetin* α product manufactured/distributed by company c in country Y; D, an *epoetin* β product manufactured/distributed by company d in country Y. *INN* of products A—C, and D are tentatively named as *epoetin* α and *epoetin* β, respectively. Symbols: ●, Man; ○, Gal; ■, GlcNAc; ▲, Fuc; ◆, NeuAc. * Bands of *epoetins*. LC: instrument, nanoFrontier nLC system (Hitachi High-Technologies Corporation); column, graphitized carbon (0.075×150 mm, ThermoFisher Scientific); flow rate, 200 nl/min; buffer A, 5 mM ammonium acetate with 2% acetonitrile (pH 9.6); buffer B, 5 mM ammonium acetate with 80% acetonitrile (pH 9.6); gradient condition, 5—35% B (110 min). MS: instrument, LTQ-TF (ThermoFisher Scientific); electron voltage, 2.0 kV (negative ion mode).

sion system. Physicochemical and biological characterization of such glycosylation is crucial at various stages, namely the development of new biotherapeutic glycoproteins, the establishment of changes in the manufacturing process, and the development of biosimilar products. Appropriate glycan testing must be adopted if the carbohydrate moiety influences safety and efficacy of the pharmaceutical. Furthermore, an *in vivo* assay could be replaced by the glycan test if the glycan profile is strongly associated with *in vivo* activity. Advances in analytical techniques for carbohydrate moieties are expected to facilitate the development of biopharmaceuticals.

**Acknowledgments** The work was supported in part by a Grant-in-Aid from the Japanese Ministry of Health, Labor, and Welfare.

## REFERENCES

1) Varki A., *Glycobiology*, **3**, 97—130 (1993).
2) Schellekens H., *Nat. Biotechnol.*, **22**, 1357—1359 (2004).
3) Schneider C. K., Kalinke U., *Nat. Biotechnol.*, **26**, 985—990 (2008).
4) Hopwood J. J., Bate G., Kirkpatrick P., *Nat. Rev. Drug Discov.*, **5**, 101—102 (2006).
5) Lee K., Jin X., Zhang K., Copertino L., Andrews L., Baker-Malcolm J., Geagan L., Qiu H., Seiger K., Barngrover D., McPherson J. M., Edmunds, T., *Glycobiology*, **13**, 305—313 (2003).
6) Kishnani P. S., Nicolino M., Voit T., Rogers R. C., Tsai A. C., Waterson J., Herman G. E., Amalfitano A., Thurberg B. L., Richards S., Davison M., Corzo D., Chen Y. T., *J. Pediatr.*, **149**, 89—97 (2006).
7) Wraith J. E., Clarke L. A., Beck M., Kolodny E. H., Pastores G. M., Muenzer J., Rapoport D. M., Berger K. I., Swiedler S. J., Kakkis E. D., Braakman T., Chadbourne E., Walton-Bowen K., Cox G. F., *J. Pediatr.*, **144**, 581—588 (2004).
8) Wraith J. E., Scarpa M., Beck M., Bodamer O. A., De Meirleir L., Guffon N., Meldgaard Lund A., Malm G., Van der Ploeg A. T., Zeman J., *Eur. J. Pediatr.*, **167**, 267—277 (2008).
9) Dahms N. M., Lobel P., Kornfeld S., *J. Biol. Chem.*, **264**, 12115—12118 (1989).
10) Van der Ploeg A. T., Kroos M. A., Willemsen R., Brons N. H., Reuser A. J., *J. Clin. Invest.*, **87**, 513—518 (1991).
11) Van Patten S. M., Hughes H., Huff M. R., Piepenhagen P. A., Waire J., Qiu H., Ganesa C., Reczek D., Ward P. V., Kutzko J. P., Edmunds T., *Glycobiology*, **17**, 467—478 (2007).
12) Higuchi M., Oh-eda M., Kuboniwa H., Tomonoh K., Shimonaka Y., Ochi N., *J. Biol. Chem.*, **267**, 7703—7709 (1992).
13) Egrie J. C., Dwyer E., Browne J. K., Hitz A., Lykos M. A., *Exp. Hematol.*, **31**, 290—299 (2003).
14) Rijken D. C., Otter M., Kuiper J., van Berkel T. J., *Thromb. Res. Suppl.*, **10**, 63—71 (1990).
15) Oikawa K., Kamimura H., Watanabe T., Miyamoto I., Higuchi S., *Thromb. Res.*, **101**, 493—500 (2001).
16) Shinkawa T., Nakamura K., Yamane N., Shoji-Hosaka E., Kanda Y., Sakurada M., Uchida K., Anazawa H., Satoh M., Yamasaki M., Hanai N., Shitara K., *J. Biol. Chem.*, **278**, 3466—3473 (2003).
17) Kanda Y., Yamada T., Mori K., Okazaki A., Inoue M., Kitajima-Miyama K., Kuni-Kamochi R., Nakano R., Yano K., Kakita S., Shitara K., Satoh M., *Glycobiology*, **17**, 104—118 (2007).
18) DeFrees S., Wang Z. G., Xing R., Scott A. E., Wang J., Zopf D., Gouty D. L., Sjoberg E. R., Panneerselvam K., Brinkman-Van der Linden E. C., Bayer R. J., Tarp M. A., Clausen H., *Glycobiology*, **16**, 833—843 (2006).
19) *MMWR Morb. Mortal Wkly. Rep.*, **57**, 124—125 (2008).
20) Kishimoto T. K., Viswanathan K., Ganguly T., Elankumaran S., Smith S., Pelzer K., Lansing J. C., Sriranganathan N., Zhao G., Galcheva-Gargova Z., Al-Hakim A., Bailey G. S., Fraser B., Roy S., Rogers-

Cotrone T., Buhse L., Whary M., Fox J., Nasr M., Dal Pan G. J., Shriver Z., Langer R. S., Venkataraman G., Austen K. F., Woodcock J., Sasisekharan R., *N. Engl. J. Med.*, **358**, 2457—2467 (2008).
21) Guerrini M., Beccati D., Shriver Z., Naggi A., Viswanathan K., Bisio A., Capila I., Lansing J. C., Guglieri S., Fraser B., Al-Hakim A., Gunay N. S., Zhang Z., Robinson L., Buhse L., Nasr M., Woodcock J., Langer R., Venkataraman G., Linhardt R. J., Casu B., Torri G., Sasisekharan R., *Nat. Biotechnol.*, **26**, 669—675 (2008).
22) Sandrin M. S., Vaughan H. A., Dabkowski P. L., McKenzie I. F., *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 11391—11395 (1993).
23) Chung C. H., Mirakhur B., Chan E., Le Q. T., Berlin J., Morse M., Murphy B. A., Satinover S. M., Hosen J., Mauro D., Slebos R. J., Zhou Q., Gold D., Hatley T., Hicklin D. J., Platts-Mills T. A., *N. Engl. J. Med.*, **358**, 1109—1117 (2008).
24) Qian J., Liu T., Yang L., Daus A., Crowley R., Zhou Q., *Anal. Biochem.*, **364**, 8—18 (2007).
25) Higashi H., Naiki M., Matuo S., Okouchi K., *Biochem. Biophys. Res. Commun.*, **79**, 388—395 (1977).
26) Martin M. J., Muotri A., Gage F., Varki A., *Nat. Med.*, **11**, 228—232 (2005).
27) Edmunds T., Van Patten S. M., Pollock J., Hanson E., Bernasconi R., Higgins E., Manavalan P., Ziomek C., Meade H., McPherson J. M., Cole E. S., *Blood*, **91**, 4561—4571 (1998).
28) Jongen S. P., Gerwig G. J., Leeflang B. R., Koles K., Mannesse M. L., van Berkel P. H., Pieper F. R., Kroos M. A., Reuser A. J., Zhou Q., Jin X., Zhang K., Edmunds T., Kamerling J. P., *Glycobiology*, **17**, 600-619 (2007).
29) Reggi S., Marchetti S., Patti T., De Amicis F., Cariati R., Bembi B., Fogher C., *Plant Mol. Biol.*, **57**, 101—113 (2005).
30) Kisung Ko M.-H. A., Song M., Choo Y.-K., Kim H. S., Ko K., Joung H., *Mol. Cells*, **25**, 494—503 (2008).
31) Ohta M., Kawasaki N., Itoh S., Hayakawa T., *Biologicals*, **30**, 235 244 (2002).
32) Kawasaki N., Ohta M., Itoh S., Hyuga M., Hyuga S., Hayakawa T., *Biologicals*, **30**, 113—123 (2002).

Immunology
The Journal of cells, molecules, systems and technologies
British Society for immunology

IMMUNOLOGY ORIGINAL ARTICLE

# Alteration of *N*-glycosylation in the kidney in a mouse model of systemic lupus erythematosus: relative quantification of *N*-glycans using an isotope-tagging method

Noritaka Hashii,[1,2] Nana Kawasaki,[1,2] Satsuki Itoh,[1] Yukari Nakajima,[1,2] Toru Kawanishi[1] and Teruhide Yamaguchi[1]

[1]*Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, Setagaya-ku, Tokyo, Japan, and* [2]*Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Agency (JST), Kawaguchi City, Saitama, Japan*

doi:10.1111/j.1365-2567.2008.02898.x
Received 19 March 2008; revised 28 May 2008; accepted 2 June 2008.
Correspondence: N. Kawasaki, Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1 Kamiyouga, Setagaya-ku, Tokyo 158-8501, Japan. Email: nana@nihs.go.jp
Senior author: Teruhide Yamaguchi, email: yamaguch@nihs.go.jp

## Summary

Changes in the glycan structures of some glycoproteins have been observed in autoimmune diseases such as systemic lupus erythematosus (SLE) and rheumatoid arthritis. A deficiency of α-mannosidase II, which is associated with branching in *N*-glycans, has been found to induce SLE-like glomerular nephritis in a mouse model. These findings suggest that the alteration of the glycosylation has some link with the development of SLE. An analysis of glycan alteration in the disordered tissues in SLE may lead to the development of improved diagnostic methods and may help to clarify the carbohydrate-related pathogenic mechanism of inflammation in SLE. In this study, a comprehensive and differential analysis of *N*-glycans in kidneys from SLE-model mice and control mice was performed by using the quantitative glycan profiling method that we have developed previously. In this method, a mixture of deuterium-labelled *N*-glycans from the kidneys of SLE-model mice and non-labelled *N*-glycans from kidneys of control mice was analysed by liquid chromatography/mass spectrometry. It was revealed that the low-molecular-mass glycans with simple structures, including agalactobiantennary and paucimannose-type oligosaccharides, markedly increased in the SLE-model mouse. On the other hand, fucosylated and galactosylated complex type glycans with high branching were decreased in the SLE-model mouse. These results suggest that the changes occurring in the *N*-glycan synthesis pathway may cause the aberrant glycosylations on not only specific glycoproteins but also on most of the glycoproteins in the SLE-model mouse. The changes in glycosylation might be involved in autoimmune pathogenesis in the model mouse kidney.

**Keywords:** isotope-tagging method; liquid chromatography/multiple-stage mass spectrometry; systemic lupus erythematosus

## Introduction

Glycosylation is one of the most common post-translational modifications[1,2] and contributes to many biological processes, including protein folding, secretion, embryonic development and cell–cell interactions.[3] Alteration of glycosylation is associated with several diseases, including inflammatory responses and malignancies;[4–6] for instance, significant increases in fucosylation and branching are found in ovarian cancer and lung cancer.[7] Additionally, the carbohydrate structure changes from type I glycans (Galβ1-3GlaNAc) to type II glycans (Galβ1-4GalNAc) in carcinoembryonic antigen in colon cancer.[8] Furthermore, an increase in biantennary oligosaccharides lacking galactose (Gal) was found on immunoglobulin G (IgG) in systemic lupus erythematosus (SLE) and rheumatoid arthritis,[9–11] and agalactoglycans are used for the early diagnosis of rheumatoid arthritis.[12]

Systemic lupus erythematosus is an autoimmune disease characterized as chronic and as a systemic disease, with symptoms such as kidney failure, arthritis and erythema. In addition to the known changes in glycosylation on IgG, there have been several reports on the association between glycosylation and inflammation in SLE and rheumatoid


 

arthritis.[13–15] A deficiency of α-mannosidase II (αM-II), which is associated with branching in *N*-glycans, has been found to induce human SLE-like glomerular nephritis in a mouse model.[16] Green *et al.* reported that branching structures of *N*-glycan in mammals are involved in protection against immune responses in autoimmune disease pathogenesis.[17] Although there is no direct evidence that alteration of glycosylation is the upstream event in the pathogenesis of SLE, these findings suggest that changes in the glycan structure may be involved in the inflammatory-related autoimmune disorder. Glycosylation analysis may lead to the development of improved diagnostic methods and may help to clarify the carbohydrate-related pathogenic mechanism of inflammation in SLE.

Mass spectrometry (MS) and liquid chromatography/mass spectrometry (LC/MS) are the most prevalent strategies for identifying disease-related glycans in glycomics.[18–20] Aberrant glycosylations in some disease samples have been found by comparing mass spectra or chromatograms between normal and disease samples; however, because of the tremendous heterogeneities of the sugar moiety in glycoprotein as well as the low reproducibility of LC/MS, accurate quantitative analysis is difficult using MS and LC/MS alone. To overcome these problems, we previously developed the stable isotope-tagging method for the quantitative profiling of glycans using 2-aminopyridine (AP).[21] After the glycans are released from sample and the reference glycoproteins are derivatized to pyridyl amino ($d_0$-PA) glycans and to tetra-deuterium-labelled pyridyl amino ($d_4$-PA) glycans, respectively, a mixture of both $d_0$-PA and $d_4$-PA glycans was subjected to LC/MS, and the levels of individual glycans were calculated from the intensity ratios of $d_0$-glycan and $d_4$-glycan molecular ions (Fig. 1a). Recently, alternative isotope-tagging methods using deuterium-labelled compounds, such as 2-aminobenzoic acid its derivatives, and permethylation, have been proposed by other groups.[22–24] All of these studies prove the utility of isotope-tagging methods for the quantitative analysis of glycosylation.

In the present study, we used the isotope-tagging method to analyse changes in *N*-glycosylation in the disordered kidney in an SLE mouse model. We used an MRL/MpJ-lpr/lpr (MRL-lpr) mouse which lacks the Fas antigen gene.[25–27] The MRL-lpr mouse is known to naturally develop SLE-like glomerular nephritis and is widely used in SLE studies. MRL/MpJ-+/+ (MRL-+/+) mice were used as controls.

## Materials and methods

### *Materials*

The kidneys of the SLE-model mice (MRL-lpr) and control mice (MRL-+/+) ($n$ = 3) were purchased from Japan SLC, Inc. (Hamamatsu, Japan). Thermolysin (EC 3.4.24.27), originating from *Bacillus thermoproteolyticus*

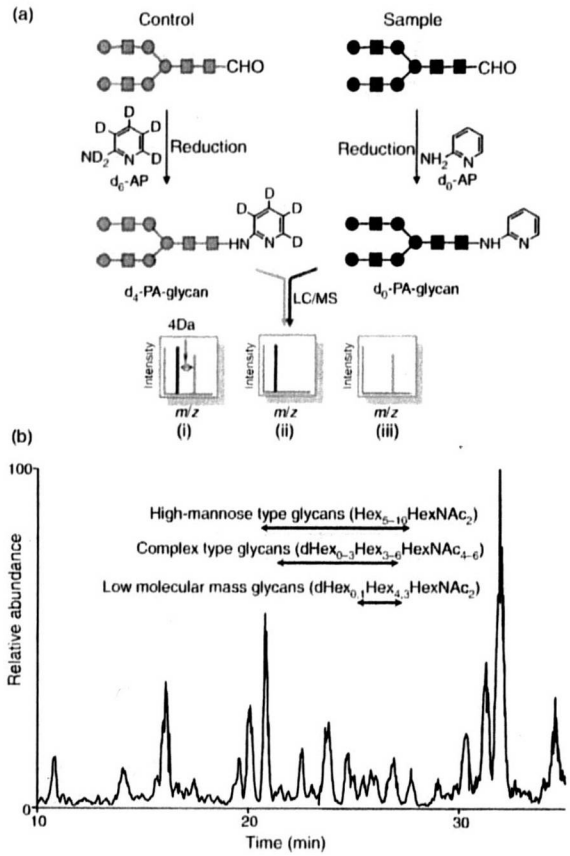


Figure 1. (a) Quantitative glycan profiling using the stable isotope-tagging method and liquid chromatography/mass spectrometry (LC/MS). (i) sample = control, (ii) sample > control, (iii) sample < control. (b) Total ion chromatogram obtained by a single scan ($m/z$ 700–2000) of the $d_0$-glycan and $d_4$-glycan mixture.

Rokko, was purchased from Daiwa Kasei (Shiga, Japan). Glycopeptidase A (PNGase A) was obtained from Seikagaku Kogyo Corporation (Tokyo, Japan). Non-deuterium-labelled 2-aminopyridine ($d_0$-AP) and deuterium-labelled 2-aminopyridine ($d_6$-AP) were purchased from Takara Bio (Otsu, Japan) and Cambridge Isotope Laboratories (Andover, MA), respectively.

### *Sample preparation*

Mouse kidneys were filtered using a cell strainer (70 μm; BD Biosciences, San Jose, CA) and contaminating blood cells in the kidney cells were burst in 140 mM $NH_4Cl$–Tris buffer (pH 7·2). The surviving kidney cells were washed three times with phosphate-buffered saline containing a mixture of protease inhibitors (Wako, Tokyo, Japan) and dissolved in guanidine–HCl buffer (8 M guanidine–HCl, 0·5 M Tris–HCl, pH 8·6) containing a mixture of protease inhibitors by vortexing at 4°. The protein concentration was measured using a 2-D Quant Kit (GE Healthcare

Bio-Sciences, Uppsala, Sweden). The protein solution (200 μg proteins) was incubated with 40 mM dithiothreitol at 65° for 30 min. Freshly prepared sodium iodoacetate (final concentration, 96 mM) was added to the sample solution, and the mixture was incubated at room temperature for 40 min in the dark. The reaction was stopped by adding cystine (6 mg/ml in 2 M HCl) in an amount equal to the amount of dithiothreitol. The solution containing carboxymethylated proteins was diluted in four times its volume of $H_2O$, and the mixture was incubated with 0·1 μg of thermolysin at 65° for 1 hr. After terminating the reaction by boiling, the reaction mixture was diluted in four times its volume of 0·2 M acetate buffer. The *N*-linked glycans were released by treatment with PNGase A (1 mU) at 37° for 16 hr and were desalted using an EnviCarb C cartridge (Supelco, Bellefonte, PA).

### *Labelling of N-glycans with $d_0$-AP and $d_6$-AP*

Glycans released from the SLE-model mouse cells were incubated in acetic acid (20 μl) with 12·5 M $d_0$-AP at 90° for 1 hr. Next, 3·3 M borane–dimethylamine complex reducing reagent in acetic acid (20 μl) was added to the solution and the mixture was incubated at 80° for 1 hr. Excess reagent was removed by evaporation, and $d_0$-PA glycans were desalted using an EnviCarb C cartridge, concentrated in a SpeedVac and reconstituted in 20 μl of 5 mM ammonium acetate (pH 9·6). Glycans released from the control mouse were labelled with $d_6$-AP in a similar manner. The resulting $d_4$-PA glycans were combined with $d_0$-PA glycans, which were prepared from an equal amount of proteins.

### *On-line liquid chromatography/mass spectrometry*

The sample solution (4 μl) was injected into the LC/MS system through a 5-μl capillary loop. The $d_0$-PA and $d_4$-PA glycans were separated in a graphitized carbon column (Hypercarb, 150 × 0·2 mm, 5 μm; Thermo Fisher Scientific, Waltham, MA) at a flow rate of 2 μl/min in a Magic 2002 LC system (Michrom Bioresources, Auburn, CA). The mobile phases were 5 mM ammonium acetate containing 2% acetonitrile (pH 9·6, A buffer) and 5 mM ammonium acetate containing 90% acetonitrile (pH 9·6, B buffer). The PA-glycans were eluted with a linear gradient of 5–45% of B buffer for 90 min.

Mass spectrometric analysis of PA glycans was performed using a Fourier transform ion cyclotron resonance/ion trap mass spectrometer (FT-ICR-MS, LTQ-FT; Thermo Fisher Scientific) equipped with a nanoelectrospray ion source (AMR, Tokyo, Japan). For MS, the electrospray voltage was 2·0 kV in the positive ion mode, the capillary temperature was 200°, the collision energy was 25% for $MS^n$ experiment, and the maximum injection times for FT-ICR-MS and $MS^n$ were 1250 and 50 milliseconds, respectively. The resolution of FT-ICR-MS was 50 000, the scan time (*m/z* 700–2000) was approximately 0·2 seconds, dynamic exclusion was 18 seconds, and the isolation width was 3·0 U (range of precursor ions ± 1·5).

## Results

### Quantitative profiling of kidney oligosaccharides in the SLE-model mouse

The recovery of oligosaccharides from whole tissues and cells is generally low because of the insolubility of the membrane fraction and possible degradation of the glycans. To improve the recovery of *N*-glycans from kidney cells, whole cells were dissolved in guanidine hydrochloride solution, and all proteins, including membrane proteins, were digested into peptides and glycopeptides with thermolysin. The *N*-glycans were then released from the glycopeptides with PNGase A, which is capable of liberating *N*-linked oligosaccharides even at the N- and/or C-terminals of peptides. The *N*-linked oligosaccharides from the SLE-model mice and control mice were labelled with $d_0$-AP and $d_6$-AP, respectively. The mixture of labelled glycans derived from an equal amount of proteins was subjected to quantitative glycan profiling using $LC/MS^n$.

Figure 1(b) shows the total ion chromatogram obtained by a single mass scan (*m/z* 700–2000) of the glycan mixture in the positive ion mode. Although the MS data contain many MS spectra derived from contaminating low-molecular-weight peptides, the MS/MS spectra of oligosaccharides could be sorted based on the existence of carbohydrate-distinctive ions, such as $HexHexNAc^+$ (*m/z* 366) and $Hex(dHex)HexNAc^+$ (*m/z* 512). The monosaccharide compositions of the precursor ions were calculated from accurate *m/z* values acquired by FT-ICR-MS. Oligosaccharides found at 25–27 min were assigned to low-molecular-mass glycans consisting of $dHex_{0,1}Hex_{4,3}HexNAc_2$ (dHex, deoxyhexose; Hex, hexose; HexNAc, *N*-acetylhexosamine). High-mannose-type glycans, which consist of $Hex_{5-10}HexNAc_2$, were located at 20–28 min; complex-type glycans ($dHex_{0-3}Hex_{3-6}HexNAc_{4-6}$) were found at 21–27 min. Figure 2(a) shows the relative intensities of the molecular ions of *N*-glycans in the SLE-model mouse, which may correspond roughly to the levels of individual *N*-glycans. More than half of all glycans were complex-type oligosaccharides, and the most prominent glycan was $dHex_3Hex_5HexNAc_5$. Man-9 ($Hex_9HexNAc_2$) was the second most common oligosaccharide. Nearly one-quarter of the glycans were low-molecular-mass glycans, and $dHex_1Hex_2HexNAc_2$ was the third most abundant glycan in the SLE-model mouse. The rate of percentage change in individual glycans between the SLE-model mice and control mice was calculated from the intensity ratio of $d_0$-glycan and $d_4$-glycan


 

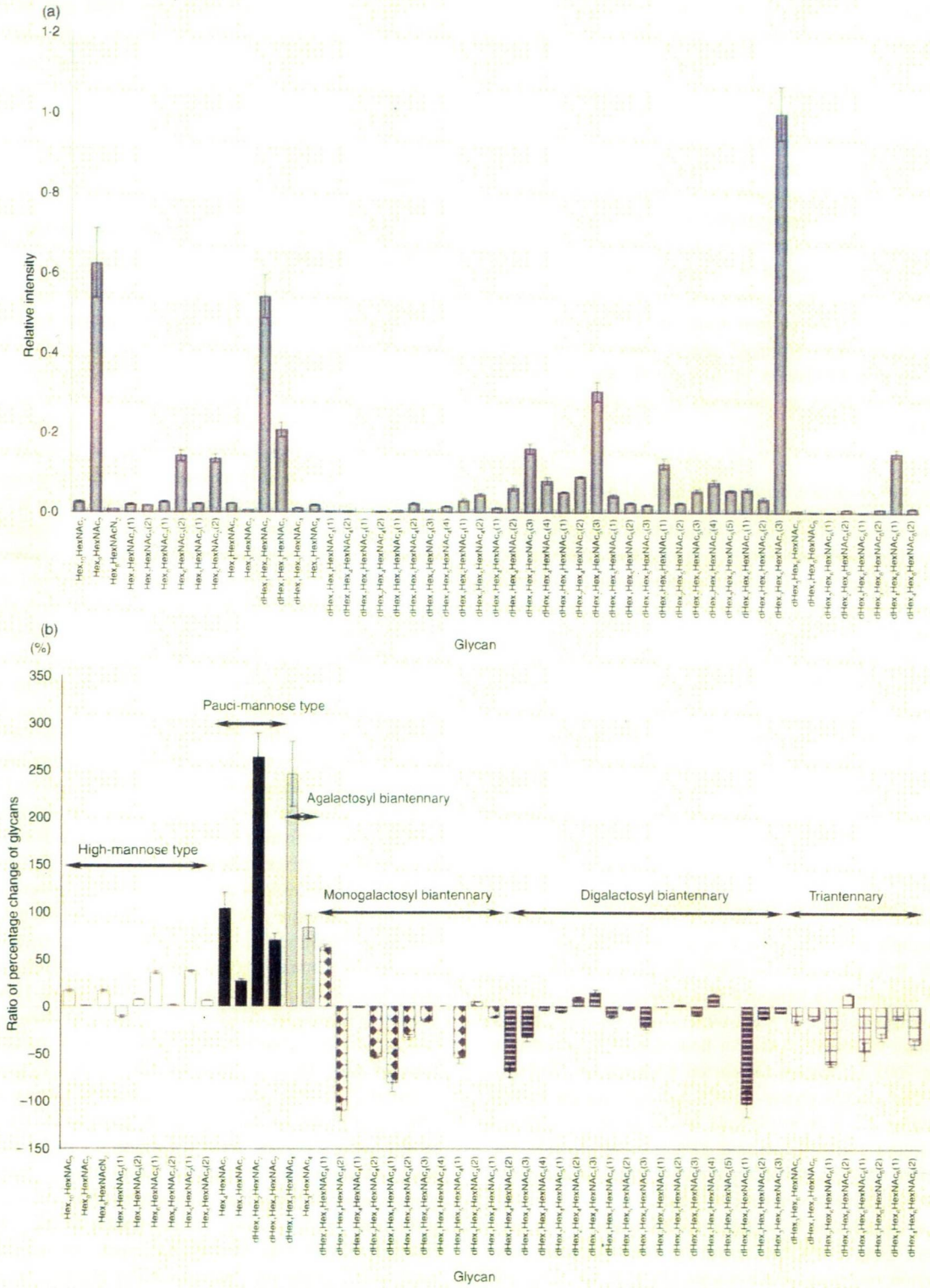


Figure 2. (a) Relative intensities of the molecular ions of $d_0$-pyridyl amino (PA) glycans from the systemic lupus erythematosus (SLE) model mouse. The intensity of the most intense ion ($[M + 2H]^{2+}$ of $d_4$-PA $dHex_3Hex_5HexNAc_3(3)$, *m/z* 1180·97) was taken as 1·0. (b) Rate of percentage change of $d_0/d_4$-glycans. Each value is the average of three biological repeats. Error bars correspond to the standard deviation. The numbers in parentheses show the isomers.

molecular ions (Fig. 2b). The significant changes found in many glycans are described below.

### Increased oligosaccharides in the SLE-model mouse

Figure 3(a,b) show the mass and MS/MS spectra of the most increased glycan, which showed a notable increase in the SLE-model mouse. Based on *m/z* values of molecular ions and differences of 1·00 U in *m/z* values among monoisotopic ions, the intense ion (*m/z* 973·40) and its neighbour ion (*m/z* 977·43) were assigned to $[M+H]^+$ of $d_0$-PA $dHex_1Hex_2HexNAc_2$, and $d_4$-PA $dHex_1Hex_2HexNAc_2$, respectively (Fig. 3a). The intensity ratio of these ions suggested that the level of $dHex_1Hex_2HexNAc_2$ increased 3·6-fold in the SLE-model mouse. The structure of this oligosaccharide was estimated to be a core-fucosylated trimannosyl core lacking a Man residue from the successive cleavages of Man ($Y_3$: *m/z* 815), Man ($Y_2$: *m/z* 653), GlcNAc ($Y_1$: *m/z* 450) and Fuc ($Y_{1/1'}$: *m/z* 304) (inset in Fig. 3b). Such a defective *N*-glycan is known as a paucimannose-type glycan, and is rarely found in vertebrates. All paucimannose-type glycans, such as $dHex_1Hex_3HexNAc_2$ (a core-fucosylated trimannosyl core) and $Hex_3HexNAc_2$ (a non-fucosylated trimannosyl core) were increased in the SLE-model mouse. Furthermore, a two-fold increase was found in $Hex_4HexNAc_2$ (Man-4).

Figure 4 shows the molecular ratios of individual *N*-glycans between the SLE-model mice and control mice. A remarkable increase (3·5-fold) was also found in $dHex_1Hex_3HexNAc_4$, which is assigned to a core-fucosylated biantennary oligosaccharide lacking two non-reducing terminal Gal residues; its non-fucosylated form ($Hex_3HexNAc_4$) was also increased 1·8-fold in the SLE-model mouse. In other complex-type glycans, $dHex_1Hex_4HexNAc_4$ (1), which is assigned to a biantennary oligosaccharide lacking one molecule of Gal, increased 1·6-fold. Interestingly, a significant decrease was found in $dHex_1Hex_4HexNAc_4$ (2), a positional isomer of $dHex_1Hex_4HexNAc_4$ (1); this might have been caused by galactosylation on either GlcNAc-Man$\alpha$1–3 or GlcNAc-Man$\alpha$1–6. In contrast, no change was found between fucosylated and non-fucosylated oligosaccharides, nor between bisected and non-bisected oligosaccharides.

A significant increase was found in some high-mannose-type oligosaccharides, such as $Hex_5HexNAc_2$ (Man-5; + 137%) and $Hex_6HexNAc_2$ (1) (Man-6; + 136%), while $Hex_7HexNAc_2$ (1,2) (Man-7) and a positional isomer of $Hex_6HexNAc_2$ (1) [$Hex_6HexNAc_2$ (2)] remained unchanged in the SLE-model mouse. A slight increase was found in $Hex_8HexNAc_2$ (Man-8; + 116%) and $Hex_{10}HexNAc_2$ (possibly assigned to Man-9 plus Glc; + 116%).

### Decreased oligosaccharides in the SLE-model mouse

The mass spectrum of the most decreased glycan is shown in Fig. 5(a). Based on differences of 0·5 U in *m/z* values among monoisotopic ions, molecular ions at *m/z* 1180·97

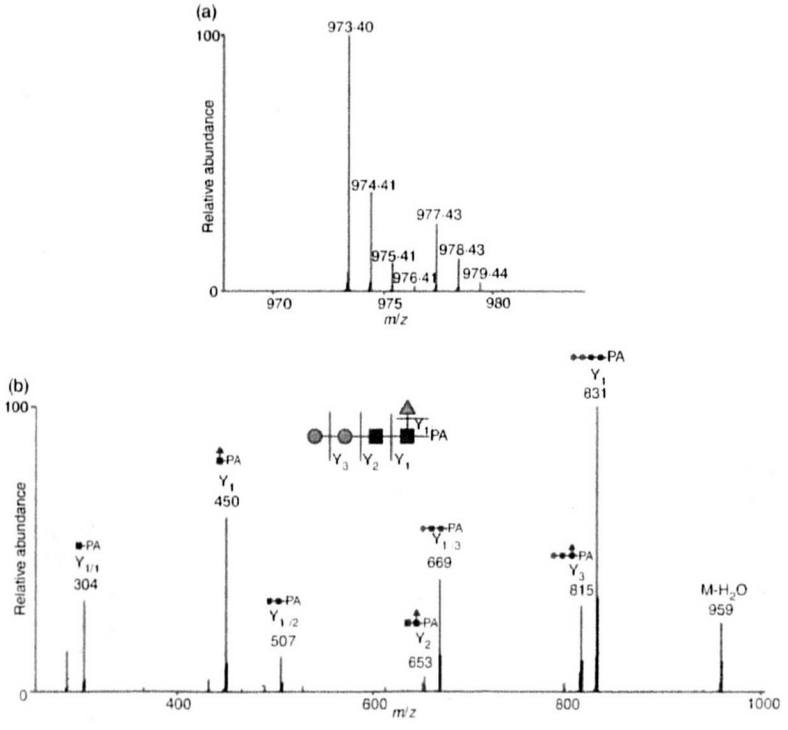


Figure 3. Mass (a) and mass spectrometry (MS)/MS (b) spectra of the most increased glycan ($dHex_1Hex_2HexNAc_2$). Precursor ion, *m/z* 973·4; grey circle, mannose; grey triangle, fucose; black square, *N*-acetylglucosamine.


 

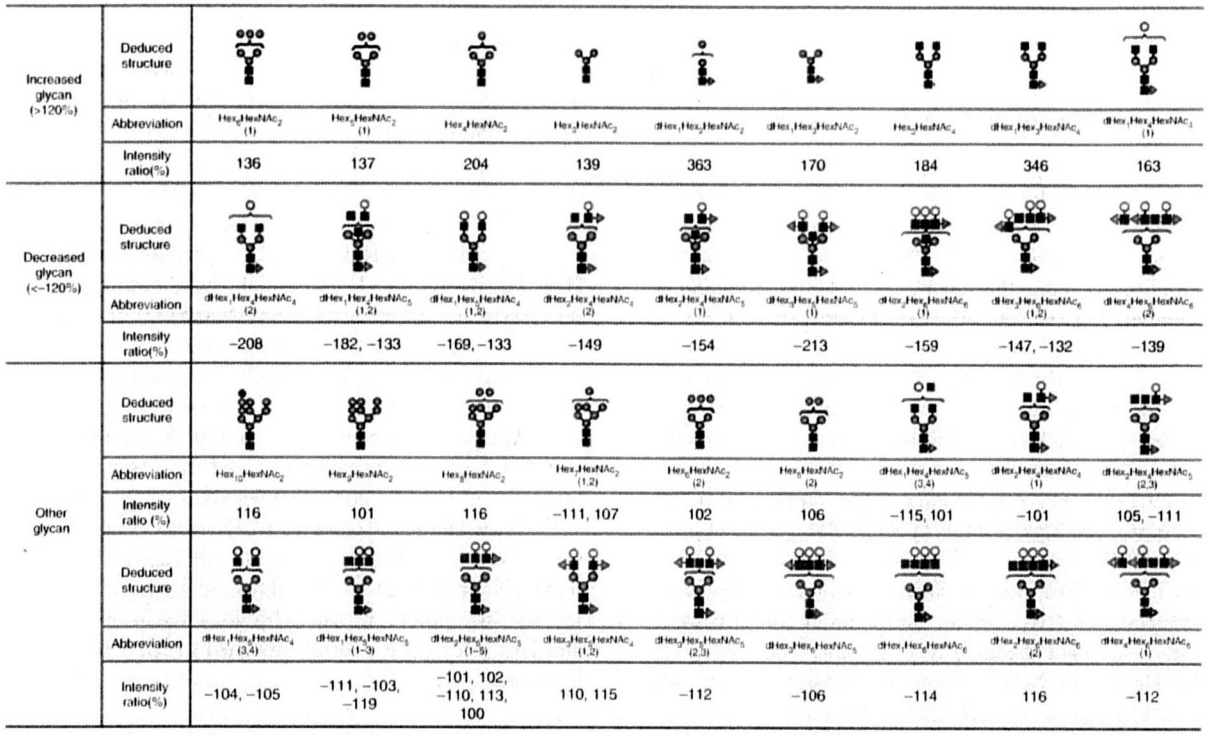


**Figure 4.** Summary of quantitative analysis of the systemic lupus erythematosus (SLE) model mouse against control mice. Values of relative ratios are the averages of three biological repeats. Grey circle, mannose; white circle, galactose; grey triangle, fucose; black square, *N*-acetylglucosamine.

and 1182·98 are estimated to be $[M + 2H]^{2+}$ of $d_0$-PA and $d_4$-PA $dHex_3Hex_5HexNAc_5$ (1), respectively. The intensity ratio of $d_0$ : $d_4$ glycans suggests that this glycan in the SLE-model mouse was decreased to 47% of the amount found in the control mouse. Figure 5(b) shows the $MS^{2-4}$ spectra of $d_0$-PA $dHex_3Hex_5HexNAc_5$ (1) (precursor ion, *m/z* 1180·97). The fragment ion at *m/z* 512 in MS/MS (i) and MS/MS/MS (ii) spectra, which corresponds to $dHex_1Hex_1HexNAc_1^+$, suggests the attachment of two Lewis motifs on the side chains of the glycan. The presence of $dHex_1HexNAc_1PA^+$ (*m/z* 446) and $dHex_1Hex_1HexNAc_3PA^+$ (*m/z* 1015) reveals the linkages of a core fucose and a bisecting GlcNAc. Based on these fragments, this decreased glycan is estimated to be a Lewis-motif-modified, core-fucosylated and bisected biantennary oligosaccharide (inset in Fig. 5).

As shown in Figs 2(b) and 4, oligosaccharides lacking one molecule of Gal with and without bisecting GlcNAc [$dHex_1Hex_4HexNAc_4$ (2) and $dHex_1Hex_4HexNAc_5$ (1)] were decreased to 48% and 55%, respectively. A significant decrease was also found in other monogalactobiantennary oligosaccharides, such as $dHex_2Hex_4HexNAc_4$ (2) (a Lewis-motif-modified, core-fucosylated monogalacto-biantennary) and $dHex_2Hex_4HexNAc_5$ (1) (a Lewis-motif-modified core-fucosylated and bisected monogalacto-biantennary).

The oligosaccharides, non-reducing ends of which are fully galactosylated, were decreased in the SLE-model mouse. For example, monofucosyl biantennary $dHex_1Hex_5HexNAc_4$ (1) and (2) were decreased 59% and 75%, respectively. The di-, tri- and tetra-fucosylated oligosaccharides, $dHex_2Hex_6HexNAc_6$ (1), $dHex_3Hex_6HexNAc_6$ (1,2) and $dHex_4Hex_6HexNAc_6$ (1,2), which were estimated to be tri- and tetraantennary forms, were also significantly decreased. These results show that oligosaccharides with a complicated structure, such as high branching oligosaccharides and di- and tri-fucosylated oligosaccharides, were decreased in the SLE-model mouse.

## Discussion

Using the isotope-tagging method, we demonstrated aberrant *N*-glycosylation on the kidney proteins of a SLE-model mouse. We found increases in low-molecular-mass glycans with simple structures, including paucimannose-type glycans, agalacto-biantennary oligosaccharides, Man-5 and Man-6, and decreases in glycans which have a complicated and diverse structure, such as digalacto-biantennary oligosaccharides and highly fucosylated glycans (Fig. 4). An increase in agalacto-biantennary oligosaccharides on IgG has been reported in the sera of patients with autoimmune diseases, including SLE, rheumatoid arthritis and IgA