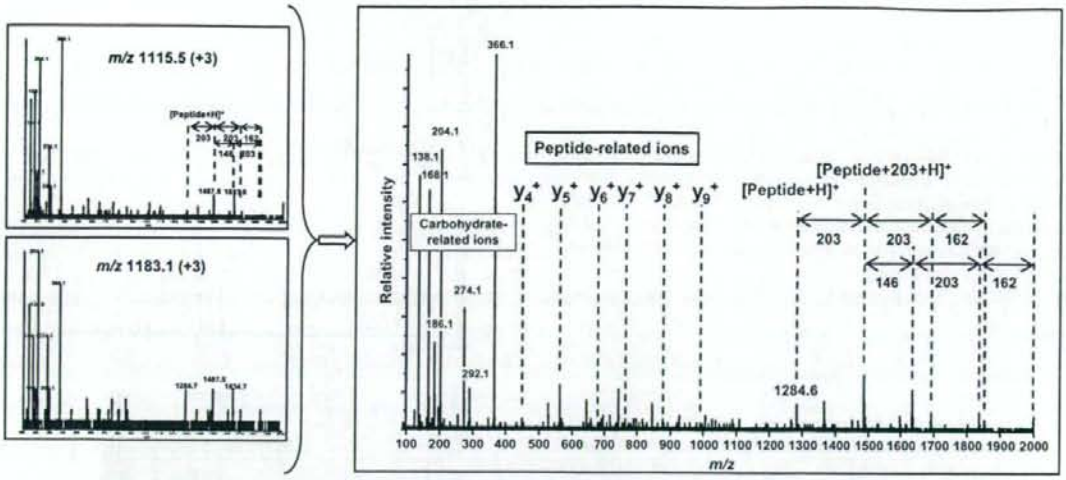



**Fig. 4.** Integrated MS/MS spectrum of *m/z* 1115.5 (3+) and 1183.1 (3+) at 52–53 min that show similar fragment patterns. Mascot database search using *m/z* 1284.5 (1+) of peptide and fragment ions (*m/z* 370–1280) suggested YKN[46]NSDISSTR in Ig mu chain C region (P01871).

the peptide was calculated to be 2678.4. For the peptide identification, a database search requires the *m/z* values and charge state of the peptide precursor ions and fragment ions but not of the carbohydrate- and glycopeptide-related ions. We deleted the carbohydrate-related ions in the lower *m/z* region (under *m/z* 370) and the glycopeptide-related ions in the higher *m/z* region (over *m/z* 1340) from the peak-list text files, and then submitted the modified peak-list text files for a Mascot database search of the human Swiss–Prot database with 1 missed cleavage, peptide tolerance of 1.2 Da, fragment ion tolerance of 0.8 Da and variable modifications of cysteine (carboxymethylation). The peptide suggested was MVSHHN[184]LTTGATLINEQWLLTTAK in human haptoglobin (P00738). As shown in Fig. 3, many ions were consistent with *b*- and *y*-series peptide fragment ions derived form MVSHHNLTTGATLINEQWLLTTAK. The molecular mass of the carbohydrate moiety was calculated to be 2204.7, which suggests the carbohydrate composition of $HexNAc_4Hex_5NeuAc_2$.

### 3.3. Assignment of glycopeptide peaks by a database search with integrated spectra

Glycopeptides that have the same peptide backbone show quite similar fragment patterns in the case of Qq-TOF MS. When glycopeptides showed insufficient peptide fragment ions in the CID-MS/MS spectra due to low peak intensity, we integrated the similar MS/MS spectra into one spectrum, and the integrated spectrum was submitted for a database search. As a representative example, the spectrum obtained by integrating two spectra of *m/z* 1115.5 (3+) and 1183.1 (3+) acquired around 60 min is shown in Fig. 4. Mascot database search using the information of *m/z* 1284.5 (1+) of peptide which was deduced by sequential degradation pattern at N-glycan core structure, and modified peak-list text files between *m/z* 370 and 1250 suggests that the peptide moiety is YKN[46]NSDISSTR in Ig mu chain C region (P01871).

By elucidating MS/MS spectra, 19 tryptic glycopeptides (20 N-glycosylation sites) in 14 glycoproteins were determined (Table 1). The ions, which were confirmed as glycopeptides by data-dependent MS/MS, were underlined. Other glycoforms, whose MS/MS spectra were not acquired, were assigned based on their mass difference of saccharide units from characterized glycopeptides. Since high intensity glycopeptide ions showed high quality of MS/MS spectra and were subjected to data-dependent MS/MS several times, many of them could be assigned. Low intensity glycopeptide ions showed poor MS/MS spectra for detection of peptide fragment ions, about 20% of MS/MS spectra of glycopeptides could not be assigned (data not shown).

### 3.4. Confirmation of glycopeptide peaks using commercially available glycoproteins

We conducted peptide mapping of commercially available polyclonal IgG and haptoglobin, and then *m/z* values and charge states of the glycopeptides were used for confirmation of assignment of glycopeptides and assignment of undetected glycopeptides. Glycosylation data of ceruloplasmin in previous report [28] was also utilized.

Tryptic digest (0.2 μg and 0.4 μg) of commercially available human polyclonal IgG was analyzed by LC/ESI MS/MS at *m/z* 400–2000 and 1000–2000 with a gradient of 5–90% of B in 85 min. The MS data were submitted for database searching against the human Swiss–Prot database using the computer program Mascot. Polypeptides of IgG heavy chain C region of IgG1 (P01857), IgG2 (P01859), IgG3 (P01860) and IgG4 (P01861) and light chain C region of Kappa (P01834) and Lambda (P01842) chain and other proteins were identified (data not shown). Fig. 5A and A′ show TIC of LC/MS/MS at *m/z* 1000–2000 of polyclonal IgG and EIC of data-dependent MS/MS at *m/z* 204.05–205.15, respectively. It was found that glycopeptide ions were eluted at 7–12 min (fraction A), 15–17 min (fraction C) and 18–21 min (fraction B) based on the presence of the oligosaccharide-related ions in their MS/MS spectra and mass differences of saccharide units. MS/MS spectra after 25 min were not of glycopeptides. The glycopeptide peaks from fraction A and fraction B were assigned as the glycopeptides containing Fc-glycosylation site in IgG1 (EEQYNSTYR) and IgG2 (EEQFNSTFR) based on data-dependent MS/MS spectra, respectively (data not shown). Data-dependent MS/MS spectra from fraction C suggested molecular mass of 1171.5 Da for the peptide, but could not suggest amino acid sequence due to low abundance of peptide fragment ions (data not shown). Based on the molecular mass of the peptide, the glycopeptide peaks from fraction C would be EEQYNSTFR from IgG3 (CAA67886) and/or EEQFNSTYR from IgG4 (P01861), which are attached to core-fucosylated agalacto-

**Table 1**
Summary of analysis of serum glycoproteome with higher ion intensities

| Glycopeptide | | | | | Oligosaccharide | Protein (Protein ID) | Theoretical MW |
|---|---|---|---|---|---|---|---|
| Retention time (min) | m/z[a] | Charge | Observed MW | Relative peak intensity[b] | Observed MW | Glycopeptide<br>Peptide sequence<br>Deduced oligosaccharide composition[c] | Peptide<br>Oligosaccharide |
| | | | | | | *Ig gamma-1 chain C region (P01857)* | 1[d] |
| | | | | | | EEQYNSTYR[A] | 1188.50 |
| 67.3 | 1479.6 | 2+ | 2957.1 | 13.1 | 1768.6 | [HexNAc]4[Hex]5[Fuc]1 | 1768.64 |
| 67.3 | 1297.0 | 2+ | 2592.0 | 3.0 | 1403.5 | [HexNAc]3[Hex]4[Fuc]1 | 1403.51 |
| 67.4, 67.6 | 1398.5 | 2+ | 2795.1 | 33.1 | 1606.5 | [HexNAc]4[Hex]4[Fuc]1 | 1606.59 |
| 67.4, 67.7 | 1216.0 | 2+ | 2429.9 | 4.3 | 1241.4 | [HexNAc]3[Hex]3[Fuc]1 | 1241.45 |
| 67.7 | 1317.5 | 2+ | 2633.0 | 27.8 | 1444.5 | [HexNAc]4[Hex]3[Fuc]1 | 1444.53 |
| 67.9 | 1500.1 | 2+ | 2998.1 | 2.3 | 1809.6 | [HexNAc]5[Hex]4[Fuc]1 | 1809.67 |
| | 1000.4 | 3+ | 2998.0 | | | | |
| 67.9 | 1581.1 | 2+ | 3160.2 | 0.2 | 1971.7 | [HexNAc]5[Hex]5[Fuc]1 | 1971.72 |
| 68.0 | 1406.6 | 2+ | 2811.1 | 0.9 | 1622.6 | [HexNAc]4[Hex]5 | 1622.58 |
| 68.2, 68.4 | 1325.5 | 2+ | 2649.0 | 2.8 | 1460.5 | [HexNAc]4[Hex]4 | 1460.53 |
| 68.2 | 1419.1 | 2+ | 2836.1 | 2.9 | 1647.6 | [HexNAc]5[Hex]3[Fuc]1 | 1647.61 |
| 68.5 | 1244.5 | 2+ | 2486.9 | 1.4 | 1298.4 | [HexNAc]4[Hex]3 | 1298.48 |
| 69.1 | 1625.1 | 2+ | 3248.1 | 2.2 | 2059.6 | [HexNAc]4[Hex]5[NeuAc]1[Fuc]1 | 2059.73 |
| | 1083.7 | 3+ | 3248.1 | | | | |
| 69.6 | 1544.1 | 2+ | 3086.2 | 0.5 | 1897.6 | [HexNAc]4[Hex]4[NeuAc]1[Fuc]1 | 1897.68 |
| | | | | | | EEQYNSTYRVVSVLTVLHQDWLNGK[F] | 2977.49 |
| 161.9, 162.2 | 1147.0 | 4+ | 4584.1 | | 1606.6 | [HexNAc]4[Hex]4[Fuc]1 | 1606.59 |
| 162.4 | 1106.5 | 4+ | 4422.1 | | 1444.6 | [HexNAc]4[Hex]3[Fuc]1 | 1444.53 |
| | | | | | | EEQYNSTYRVVSVLTVLHQDWLNGKEYK[F] | 3397.69 |
| 156.4 | 1034.3 | 5+ | 5166.4 | | 1768.7 | [HexNAc]4[Hex]5[Fuc]1 | 1768.64 |
| 156.6, 157.1 | 1001.8 | 5+ | 5004.0 | | 1606.3 | [HexNAc]4[Hex]4[Fuc]1 | 1606.59 |
| | | | | | | *Ig gamma-2 chain C region (P01859)* | 1[d] |
| | | | | | | EEQFNSTFR[B] | 1156.51 |
| 85.4 | 1463.6 | 2+ | 2925.1 | 6.3 | 1768.6 | [HexNAc]4[Hex]5[Fuc]1 | 1768.64 |
| 85.5 | 1281.0 | 2+ | 2560.0 | 1.4 | 1403.5 | [HexNAc]3[Hex]4[Fuc]1 | 1403.51 |
| 85.7, 86.3 | 1382.5 | 2+ | 2763.1 | 19.4 | 1606.5 | [HexNAc]4[Hex]4[Fuc]1 | 1606.59 |
| 85.7, 86.4 | 1200.0 | 2+ | 2397.9 | 2.8 | 1241.4 | [HexNAc]3[Hex]3[Fuc]1 | 1241.45 |
| 85.7 | 1565.1 | 2+ | 3128.2 | 0.0 | 1971.7 | [HexNAc]5[Hex]5[Fuc]1 | 1971.72 |
| 86.0 | 1484.1 | 2+ | 2966.2 | 1.0 | 1809.6 | [HexNAc]5[Hex]4[Fuc]1 | 1809.67 |
| 86.5 | 1301.5 | 2+ | 2601.0 | 21.8 | 1444.5 | [HexNAc]4[Hex]3[Fuc]1 | 1444.53 |
| 86.5 | 1390.5 | 2+ | 2779.0 | 0.0 | 1622.5 | [HexNAc]4[Hex]5 | 1622.58 |
| 86.9 | 1403.0 | 2+ | 2804.1 | 1.8 | 1647.5 | [HexNAc]5[Hex]3[Fuc]1 | 1647.61 |
| 87.0, 87.5 | 1309.5 | 2+ | 2617.0 | 0.1 | 1460.5 | [HexNAc]4[Hex]4 | 1460.53 |
| 87.6 | 1228.5 | 2+ | 2454.9 | 0.2 | 1298.4 | [HexNAc]4[Hex]3 | 1298.48 |
| 89.4 | 1609.1 | 2+ | 3216.2 | 1.6 | 2059.7 | [HexNAc]4[Hex]5[NeuAc]1[Fuc]1 | 2059.73 |
| | 1073.1 | 3+ | 3216.2 | | | | |
| 90.0 | 1528.1 | 2+ | 3054.2 | 1.5 | 1897.6 | [HexNAc]4[Hex]4[NeuAc]1[Fuc]1 | 1897.68 |
| | 1019.1 | 3+ | 3054.1 | | | | |
| | | | | | | *Gamma 3 immunoglobulin constant heavy chain (CAA67886)* | 2[d] |
| | | | | | | EEQYNSTFR[C] | 1172.51 |
| | | | | | | *Ig gamma-4 chain C region (P01861)* | 1[d] |
| | | | | | | EEQFNSTYR[C] | 1172.51 |
| 76.4 | 1471.6 | 2+ | 2941.1 | 1.0 | 1768.6 | [HexNAc]4[Hex]5[Fuc]1 | 1768.64 |
| 76.5 | 1289.0 | 2+ | 2576.0 | 0.2 | 1403.5 | [HexNAc]3[Hex]4[Fuc]1 | 1403.51 |
| 76.6, 76.8 | 1390.6 | 2+ | 2779.1 | 3.4 | 1606.6 | [HexNAc]4[Hex]4[Fuc]1 | 1606.59 |
| 76.5, 76.8 | 1208.0 | 2+ | 2413.9 | 0.4 | 1241.4 | [HexNAc]3[Hex]3[Fuc]1 | 1241.45 |
| 76.7 | 1492.1 | 2+ | 2982.1 | 0.2 | 1809.6 | [HexNAc]5[Hex]4[Fuc]1 | 1809.67 |
| 76.9 | 1309.5 | 2+ | 2617.0 | 3.6 | 1444.5 | [HexNAc]4[Hex]3[Fuc]1 | 1444.53 |
| 77.0 | 1398.6 | 2+ | 2795.1 | 0.1 | 1622.6 | [HexNAc]4[Hex]5 | 1622.58 |
| 76.9 | 1317.5 | 2+ | 2633.0 | 0.1 | 1460.5 | [HexNAc]4[Hex]4 | 1460.53 |
| 77.0 | 1411.1 | 2+ | 2820.1 | 0.3 | 1647.6 | [HexNAc]5[Hex]3[Fuc]1 | 1647.61 |
| 77.4 | 1236.4 | 2+ | 2470.8 | 0.1 | 1298.3 | [HexNAc]4[Hex]3 | 1298.48 |
| 78.5 | 1617.1 | 2+ | 3232.1 | 0.3 | 2059.6 | [HexNAc]4[Hex]5[NeuAc]1[Fuc]1 | 2059.73 |
| 79.1 | 1536.1 | 2+ | 3070.1 | 0.1 | 1897.6 | [HexNAc]4[Hex]4[NeuAc]1[Fuc]1 | 1897.68 |
| | | | | | | *Haptoglobin (P00738)* | 4[d] |
| | | | | | | MVSHHNLTTGATLINEQWLLTTAK[D] | 2678.39 |
| 137.9 | 1531.7 | 3+ | 4592.1 | | 1913.7 | [HexNAc]4[Hex]5[NeuAc]1 | 1913.68 |
| | 1149.0 | 4+ | 4592.0 | 30.1 | 1913.6 | | |

**Table 1** (*Continued*)

| Glycopeptide | | | | | Oligosaccharide | Protein (Protein ID) | Theoretical MW |
|---|---|---|---|---|---|---|---|
| Retention time (min) | *m/z*[a] | Charge | Observed MW | Relative peak intensity[b] | Observed MW | Glycopeptide<br>Peptide sequence<br>Deduced oligosaccharide composition[f] | Peptide<br>Oligosaccharide |
| 141.4 | 1221.8 | 4+ | 4883.1 | 88.7 | 2204.7 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | M(O)VSHHNLTTGATLINEQWLLTTAK | 2694.38 |
| 137.3 | 1153.0 | 4+ | 4608.1 | 28.9 | 1913.7 | [HexNAc]4[Hex]5[NeuAc]1 | 1913.68 |
| 140.5, 141.1 | 1225.8 | 4+ | 4899.1 | 64.3 | 2204.7 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | 1634.1 | 3+ | 4899.1 | | 2204.8 | | |
| | | | | | | NLFLNHSENATAK[E] | 1457.73 |
| 86.3 | 1395.0 | 4+ | 5576.0 | 0.1 | 4118.3 | [HexNAc]8[Hex]10[NeuAc]3 | 4118.45 |
| 87.0 | 1650.3 | 4+ | 6597.3 | 0.3 | 5139.6 | [HexNAc]10[Hex]12[NeuAc]4 | 5139.81 |
| 87.6 | 1595.6 | 4+ | 6378.3 | 0.9 | 4920.6 | [HexNAc]9[Hex]11[NeuAc]4[Fuc]1 | 4920.73 |
| 87.9 | 1559.1 | 4+ | 6232.3 | 2.7 | 4774.5 | [HexNAc]9[Hex]11[NeuAc]4 | 4774.68 |
| 88.6 | 1504.3 | 4+ | 6013.0 | 0.1 | 4555.3 | [HexNAc]8[Hex]10[NeuAc]4[Fuc]1 | 4555.60 |
| 88.9 | 1467.8 | 4+ | 5867.1 | 4.1 | 4409.4 | [HexNAc]8[Hex]10[NeuAc]4 | 4409.54 |
| 90.4 | 1759.6 | 4+ | 7034.5 | 0.2 | 5576.8 | [HexNAc]10[Hex]12[NeuAc]5[Fuc]1 | 5576.96 |
| 90.7 | 1723.1 | 4+ | 6888.5 | 0.5 | 5430.8 | [HexNAc]10[Hex]12[NeuAc]5 | 5430.90 |
| 91.5 | 1668.3 | 4+ | 6669.4 | 0.3 | 5211.6 | [HexNAc]9[Hex]11[NeuAc]5[Fuc]1 | 5211.83 |
| 91.7 | 1631.8 | 4+ | 6523.3 | 0.4 | 5065.6 | [HexNAc]9[Hex]11[NeuAc]5 | 5065.77 |
| 87.8 | 1124.7 | 3+ | 3371.2 | - | 1913.5 | [HexNAc]4[Hex]5[NeuAc]1 | 1913.68 |
| 91.6 | 1221.7 | 3+ | 3662.2 | - | 2204.5 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | VVLHPNYSQVDIGLIK[F] | 1794.00 |
| 127 | 1358.6 | 3+ | 4072.8 | 2.2 | 2278.8 | [HexNAc]5[Hex]6[NeuAc]1 | 2278.81 |
| 128.2 | 1236.9 | 3+ | 3707.7 | 3.1 | 1913.7 | [HexNAc]4[Hex]5[NeuAc]1 | 1913.68 |
| 131.4 | 1455.6 | 3+ | 4363.8 | 4.8 | 2569.8 | [HexNAc]5[Hex]6[NeuAc]2 | 2569.90 |
| | 1092.0 | 4+ | 4363.8 | | | | |
| 131.8 | 1333.9 | 3+ | 3998.7 | 89.2 | 2204.7 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | 1000.7 | 4+ | 3998.7 | | 2204.7 | | |
| 134.1 | 1552.7 | 3+ | 4655.0 | 7.5 | 2860.9 | [HexNAc]5[Hex]6[NeuAc]3 | 2861.00 |
| 134.1 | 1164.7 | 4+ | 4654.9 | | 2860.9 | | |
| | | | | | | *Transferrin (P02787)* | 2[d] |
| | | | | | | CGLVPVLAENYNK[G] | 1476.73 |
| 126.1 | 1252.8 | 3+ | 3755.5 | 0.9 | 2278.8 | [HexNAc]5[Hex]6[NeuAc]1 | 2278.81 |
| 127.0 | 1131.1 | 3+ | 3390.4 | 1.7 | 1913.7 | [HexNAc]4[Hex]5[NeuAc]1 | 1913.68 |
| 129.8 | 1349.9 | 3+ | 4046.7 | 1.6 | 2569.9 | [HexNAc]5[Hex]6[NeuAc]2 | 2569.90 |
| 130.6 | 1228.2 | 3+ | 3681.5 | 46.8 | 2204.7 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| 133.1 | 1446.9 | 3+ | 4337.8 | 0.8 | 2861.0 | [HexNAc]5[Hex]6[NeuAc]3 | 2861.00 |
| | | | | | | QQQHLFGSNVTDCSGNFCLFR[H] | 2516.08 |
| 143.8 | 1623.3 | 3+ | 4866.9 | 4.1 | 2350.8 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| 143.9 | 1181.2 | 4+ | 4720.9 | | 2204.8 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | 1574.6 | 3+ | 4720.9 | 50.5 | 2204.8 | | |
| 146.0 | 1842.0 | 3+ | 5523.0 | 0.9 | 3007.0 | [HexNAc]5[Hex]6[NeuAc]3[Fuc]1 | 3007.06 |
| 146.2 | 1793.4 | 3+ | 5377.3 | 1.6 | 2861.2 | [HexNAc]5[Hex]6[NeuAc]3 | 2861.00 |
| | | | | | | *Ceruloplasmin (P00450)* | 4[d] |
| | | | | | | EHEGAIYPDNTTDFQR[I] | 1891.83 |
| 95.6 | 1415.2 | 3+ | 4242.7 | 0.1 | 2350.8 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| | 1061.7 | 4+ | 4242.6 | | 2350.8 | | |
| 96.3 | 1366.5 | 3+ | 4096.6 | 6.0 | 2204.8 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | 1025.2 | 4+ | 4096.7 | | 2204.9 | | |
| 98.1, 98.5 | 1633.9 | 3+ | 4898.7 | 0.4 | 3006.9 | [HexNAc]5[Hex]6[NeuAc]3[Fuc]1 | 3007.06 |
| | 1226.0 | 4+ | 4899.9 | | 3008.0 | | |
| 98.8 | 1585.2 | 3+ | 4752.7 | 0.4 | 2860.9 | [HexNAc]5[Hex]6[NeuAc]3 | 2861.00 |
| | 1189.2 | 4+ | 4752.6 | | 2860.8 | | |
| | | | | | | ENLTAPGSDSAVFFEQGTTR[J] | 2125.99 |
| 127.0 | 1493.2 | 3+ | 4476.6 | 0.0 | 2350.6 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| 127.4 | 1444.5 | 3+ | 4330.6 | 2.8 | 2204.6 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| 129.3 | 1663.3 | 3+ | 4987.0 | 0.2 | 2861.0 | [HexNAc]5[Hex]6[NeuAc]3 | 2861.00 |
| | | | | | | ELHHLQEQNVSNAFLDK[K] | 2021.00 |
| 104.3 | 1093.9 | 4+ | 4371.7 | 1.4 | 2350.7 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| | 1458.3 | 3+ | 4371.8 | 0.2 | 2350.8 | | |
| 105.2 | 1057.4 | 4+ | 4225.7 | 4.5 | 2204.7 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | 1409.6 | 3+ | 4225.6 | 1.2 | 2204.6 | | |

**Table 1** (*Continued*)

| Glycopeptide | | | | | Oligosaccharide | Protein (Protein ID) | Theoretical MW |
|---|---|---|---|---|---|---|---|
| Retention time (min) | m/z[a] | Charge | Observed MW | Relative peak intensity[b] | Observed MW | Glycopeptide<br>Peptide sequence<br>Deduced oligosaccharide composition[c] | Peptide<br>Oligosaccharide |
| 106.6 | 1294.6 | 4+ | 5174.2 | 1.2 | 3153.2 | [HexNAc]5[Hex]6[NeuAc]3[Fuc]2 | 3153.12 |
| 106.8, 107.4 | 1258.0 | 4+ | 5027.9 | 2.0 | 3006.9 | [HexNAc]5[Hex]6[NeuAc]3[Fuc]1 | 3007.06 |
| 107.7 | 1221.5 | 4+ | 4881.8 | 1.9 | 2860.8 | [HexNAc]5[Hex]6[NeuAc]3 | 2861.00 |
| | | | | | | *Alpha-1-antitrypsin (P01009)* | 3[d] |
| | | | | | | YLGNATAIFFLPDEGK | 1754.89 |
| 154.6 | 1369.6 | 3+ | 4105.7 | 2.2 | 2350.8 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| 154.8 | 1320.9 | 3+ | 3959.7 | 140.6 | 2204.8 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | *Alpha-2-HS-glycoprotein (P02765)* | 2[d] |
| | | | | | | VCQDCPLLAPLNDTR | 1772.81 |
| 136.9 | 1326.9 | 3+ | 3977.7 | | 2204.8 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | *Alpha-2-macroglobulin (P01023)* | 8[d] |
| | | | | | | VSNQTLSLFFTVLQDVPVR | 2162.17 |
| 187.9, 188.8 | 1505.3 | 3+ | 4512.7 | 5.1 | 2350.6 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| 188.3 | 1456.7 | 3+ | 4367.0 | 22.5 | 2204.8 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | *Beta-2-glycoprotein 1 (P02749)* | 4[d] |
| | | | | | | VYKPSAGNNSLYR | 1467.75 |
| 83.5 | 1273.8 | 3+ | 3818.5 | 1.5 | 2350.8 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| 83.6 | 1225.2 | 3+ | 3672.5 | 6.9 | 2204.8 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| 85.2 | 1492.6 | 3+ | 4474.6 | 0.3 | 3006.9 | [HexNAc]5[Hex]6[NeuAc]3[Fuc]1 | 3007.06 |
| 85.4 | 1443.9 | 3+ | 4328.6 | 0.5 | 2860.8 | [HexNAc]5[Hex]6[NeuAc]3 | 2861.00 |
| | | | | | | LGNWSAMPSCK | 1250.54 |
| 109.7 | 1152.8 | 3+ | 3455.3 | | 2204.7 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | *Complement C3 (P01024)* | 3[d] |
| | | | | | | TVLTPATNHMGNVTFTIPANR | 2254.15 |
| 121.0 | 1265.9 | 3+ | 3794.6 | 5.4 | 1540.4 | [HexNAc]2[Hex]7 | 1540.53 |
| 121.2 | 1211.8 | 3+ | 3632.5 | 47.8 | 1378.4 | [HexNAc]2[Hex]6 | 1378.48 |
| 121.6 | 1157.8 | 3+ | 3470.4 | 10.2 | 1216.3 | [HexNAc]2[Hex]5 | 1216.42 |
| | | | | | | *Hemopexin (P02790)* | 5[d] |
| | | | | | | SWPAVGNCSSALR | 1404.65 |
| 115.3 | 1252.8 | 3+ | 3755.5 | 0.7 | 2350.8 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| 115.8 | 1204.1 | 3+ | 3609.4 | 10.3 | 2204.7 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | ALPQPQNVTSLLGCTH | 1735.86 |
| 115.8 | 1314.5 | 3+ | 3940.4 | 10.3 | 2204.5 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | *Ig alpha-1 chain C region (P01876)* | 2[d] |
| | | | | | | *Ig alpha-2 chain C region (P01877)* | 4[d] |
| | | | | | | LSLHRPALEDLLLGSEANLTCTLTGLR | 2963.58 |
| 165.2, 165.7 | 1157.8 | 4+ | 4627.2 | 8.0 | 1663.6 | [HexNAc]5[Hex]4 | 1663.61 |
| 165.8 | 1117.3 | 4+ | 4465.0 | 15.4 | 1501.4 | [HexNAc]5[Hex]3 | 1501.56 |
| 165.9 | 1046.0 | 4+ | 4180.0 | 5.6 | 1216.4 | [HexNAc]2[Hex]5 | 1216.42 |
| 169.2 | 1220.3 | 4+ | 4877.2 | 48.8 | 1913.6 | [HexNAc]4[Hex]5[NeuAc]1 | 1913.68 |
| 168.8, 169.4 | 1256.9 | 4+ | 5023.5 | 1.3 | 2059.9 | [HexNAc]4[Hex]5[NeuAc]1[Fuc]1 | 2059.73 |
| 169.9 | 1179.8 | 4+ | 4715.3 | 4.0 | 1751.7 | [HexNAc]4[Hex]4[NeuAc]1 | 1751.62 |
| 170.0 | 1169.6 | 4+ | 4674.2 | 5.1[f] | 1710.6 | [HexNAc]3[Hex]5[NeuAc]1 | 1710.60 |
| 169.0 | 1271.2 | 4+ | 5080.8 | 21.0[f] | 2117.2 | [HexNAc]5[Hex]5[NeuAc]1 | 2116.76 |
| | 1017.2 | 5+ | 5081.0 | | 2117.4 | | |
| 169.9 | 1230.4 | 4+ | 4917.6 | 11.8 | 1954.0 | [HexNAc]5[Hex]4[NeuAc]1 | 1954.70 |
| 173.1 | 1293.2 | 4+ | 5168.8 | 9.0 | 2205.2 | [HexNAc]4[Hex]5[NeuAc]2 | 2204.77 |
| | | | | | | PALEDLLLGSEANLTCTLTGLR[e] | 2357.21 |
| 174.4 | 1287.2 | 3+ | 3858.6 | 14.3 | 1501.4 | [HexNAc]5[Hex]3 | 1501.56 |
| 176.7 | 1424.6 | 3+ | 4270.9 | 65.8 | 1913.7 | [HexNAc]4[Hex]5[NeuAc]1 | 1913.68 |
| 176.4 | 1492.5 | 3+ | 4474.5 | 20.5 | 2117.3 | [HexNAc]5[Hex]5[NeuAc]1 | 2116.76 |
| | | | | | | *Ig alpha-2 chain C region (P01877)* | |
| | | | | | | TPLTANITK | 957.55 |
| 84.1 | 1006.8 | 3+ | 3017.2 | 3.7 | 2059.7 | [HexNAc]4[Hex]5[NeuAc]1[Fuc]1 | 2059.73 |
| | 1509.6 | 2+ | 3017.2 | | 2059.6 | | |
| 84.1 | 1074.4 | 3+ | 3220.3 | 4.8 | 2262.8 | [HexNAc]5[Hex]5[NeuAc]1[Fuc]1 | 2262.81 |
| | 1611.2 | 2+ | 3220.3 | | 2262.8 | | |
| 87.2 | 1103.8 | 3+ | 3308.3 | 1.3 | 2350.8 | [HexNAc]4[Hex]5[NeuAc]2[Fuc]1 | 2350.83 |
| | | | | | | *Ig mu chain C region (P01871)* | 5[d] |
| | | | | | | YKNNSDISSTR[e] | 1283.61 |

**Table 1** (*Continued*)

| Glycopeptide | | | | | Oligosaccharide | Protein (Protein ID) | Theoretical MW |
|---|---|---|---|---|---|---|---|
| Retention time (min) | *m/z*[a] | Charge | Observed MW | Relative peak intensity[b] | Observed MW | Glycopeptide<br>Peptide sequence<br>Deduced oligosaccharide composition[c] | Peptide<br>Oligosaccharide |
| 59.7 | 1115.5 | 3+ | 3343.4 | 40.1 | 2059.7 | [HexNAc]4[Hex]5[NeuAc]1[Fuc]1 | 2059.73 |
| 60.3 | 1183.1 | 3+ | 3546.4 | 16.4 | 2262.8 | [HexNAc]5[Hex]5[NeuAc]1[Fuc]1 | 2262.81 |

[a] Underlines indicated that these ions were assigned by elucidating data-dependent MS/MS of LC/ESI MS/MS of human serum digest.
[b] Centroid peak intensity (count per sec) in integrated MS spectra during glycopeptide eluting period.
[c] Oligosaccharide compositions were deduced from molecular weights.
[d] Number of potential N-glycosylation sites.
[e] Missed cleavage or unexpected digestion.
[f] Other ions with same *m/z* overlapped.
[A–K] Mass spectra were shown in Fig. 7A–K.
All masses are monoisotopic. Cysteine residue was carboxymethylated. Potential N-glycosylation sites were underlined. M(O), oxidized methionine; Fuc, fucose; Hex, hexose; HexNAc, N-acetylhexosamine; NeuAc, N-acetylneuraminic acid

or mono/diagalacto-biantennary complex-type glycan. Integrated mass spectra of fraction A, B and C were shown in Fig. 5B–D.

Haptoglobin has four potential N-glycosylation sites. We performed peptide mapping using a tryptic digest of haptoglobin under a chromatographic condition similar to that of human serum. Fig. 6A and A′ show TIC obtained by LC/MS/MS with mass range *m/z* 1000–2000 and EIC of data-dependent MS/MS at *m/z* 204.05–205.15, respectively. Glycopeptides for four potential glycosylation sites were assigned by elucidating MS/MS spectra (spectra were not shown). Glycopeptides of NLFLN$^{207}$HSEN$^{211}$ATAK containing two N-glycosylation sites were eluted in fraction E as two glycosylated forms (Fig. 6B) and fraction E′ as one glycosylated forms (Fig. 6C). The former glycosylated form was more abundant than the later form. These glycosylation sites could not be characterized separately by trypsin digestion. Glycopeptides of VVLHPN$^{241}$YSQVDIGLIK and MVSHHN$^{184}$LTTGATLINEQWLLTTAK were eluted in fractions F and D, respectively (Fig. 6D and E). From the molecular masses of oligosaccharides we inferred that a majority of oligosaccharides in haptoglobin are di-, tri-, and tetra-antennary forms and that some oligosaccharides were not fully saturated with NeuAc, and few glycans were fucosylated.

Using the data of relative retention times, accurate *m/z* values and charge states obtained by peptide mapping of commercially available glycoproteins, we confirmed already assigned glycopeptides and further assigned undetected glycopeptides (IgG3/IgG4 and two sites of ceruloplasmin), with the exceptions of one of the glycopeptides from ceruloplasmin, intensity of which was only noise levels.

### 3.5. Site-specific glycosylation analysis

To analyze the heterogeneity of glycosylation at each site, we performed an additional LC/MS in which switching to MS/MS was not allowed (Fig. 2C). Utilizing the information of retention time, accurate *m/z* and charge state of assigned glycopeptides by LC/MS/MS, corresponding glycopeptides were assigned in LC/MS data by mass chromatogram. When two or more glycoforms were detected, mass spectrometric heterogeneity was calculated using

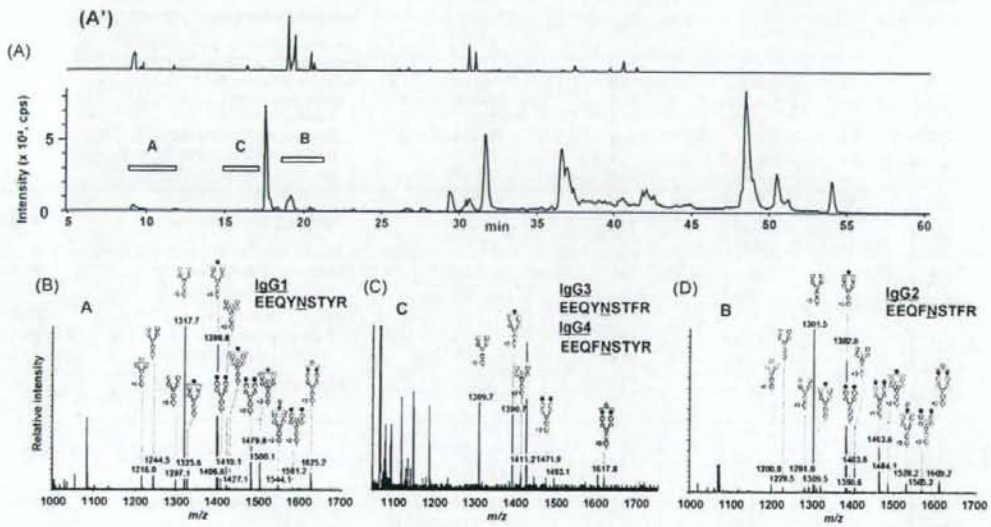



**Fig. 5.** Peptide map of commercially available human polyclonal IgG. (A) TIC (*m/z* 1000–2000) obtained by LC/MS/MS of trypsin-digested IgG. (A′) EIC (*m/z* 204.05–204.15) obtained by data-dependent MS/MS. (B) Mass spectrum of peak A, which was assigned as glycopeptides of EEQYNSTYR of IgG1 (P01857). (C) Mass spectrum of peak C, which would be glycopeptides of EEQYNSTFR of IgG3 (CAA67886) and/or EEQFNSTYR of IgG4 (P01861). (D) Mass spectrum of peak B, which was assigned as glycopeptides of EEQFNSTFR of IgG2 (P01859).

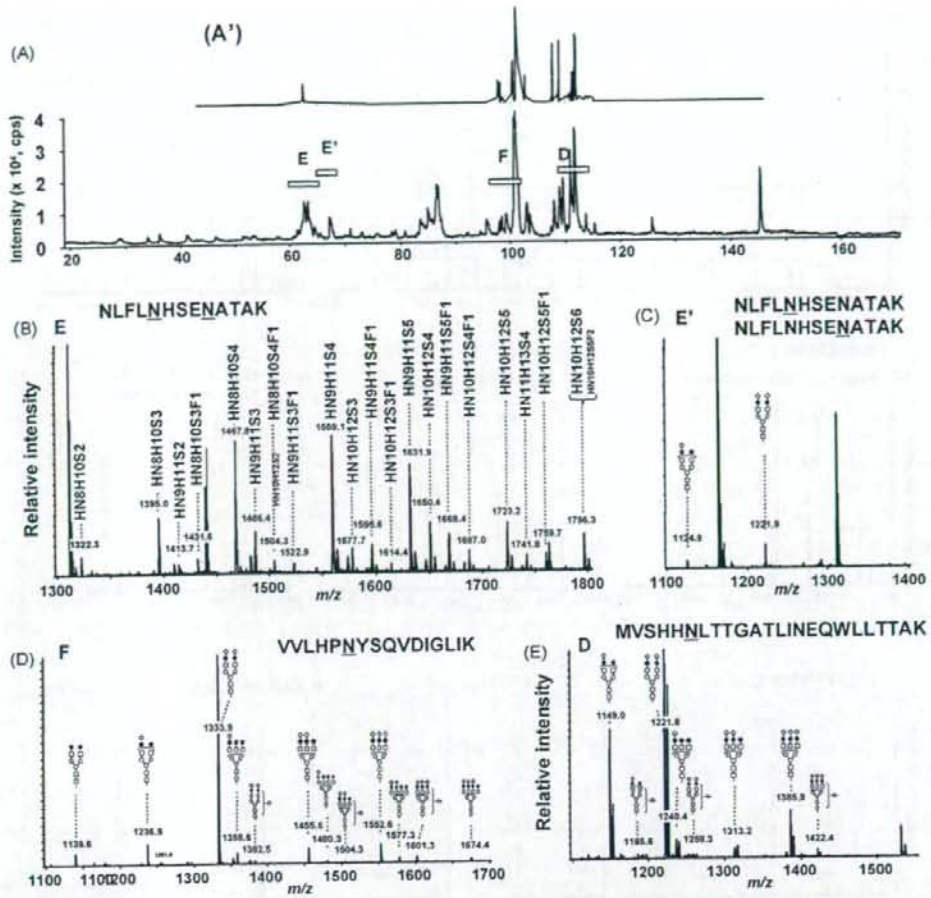



**Fig. 6.** Peptide map of commercially available human haptoglobin. (A) TIC (*m/z* 1000–2000) obtained by LC/MS/MS of trypsin-digested haptoglobin. (A') EIC (*m/z* 204.05–204.15) obtained by data-dependent MS/MS. (B) Mass spectrum of peak E, which was assigned as glycopeptides of NLFLN$^{207}$HSEN$^{211}$ATAK attached to two N-glycan. (C) Mass spectrum of peak E', which was assigned as glycopeptides of NLFLN$^{207}$HSEN$^{211}$ATAK attached to one N-glycan. (D) Mass spectrum of peak F, which was identified as VVLHPN$^{241}$YSQVDIGLIK. (E) Mass spectrum of peak D, which was identified as MVSHHN$^{184}$LTTGATLINEQWLLTTAK. H, hexose; HN, *N*-acetylhexosamine; S, *N*-acetylneuraminic acid; F, fucose.

integrated mass spectra during the periods eluting the glycopeptides with same peptide. In Fig. 7, we show integrated mass spectra of fraction A–K (Fig. 2C) as the mass spectrometric heterogeneity of glycosylation in IgG1 (Fig. 7A), IgG2 (Fig. 7B), IgG3/IgG4 (Fig. 7C), haptoglobin (Fig. 7D–F), transferrin (Fig. 7G and H) and ceruloplasmin (Fig. 7I–K). Centroid ion intensity (count/sec) of each glycopeptide at the most intense isotope distribution was used as relative peak intensity. The mass spectrometric heterogeneity of the Fc-glycosylation sites of IgG1 (Fig. 7A) and IgG2 (Fig. 7B) was consistent with those of the commercially available polyclonal IgG (Fig. 5B and D) and previous reports [29]. The glycosylation pattern of haptoglobin at each site was similar to that of the commercially available haptoglobin except that peak intensities of minor glycoforms were noise level (Figs. 6B–E and 7D–F). The glycosylation of transferrin (Fig. 7G and H) at each site was consistent with previous reports [29]. Three glycopeptides of the four expected ones derived from ceruloplasmin could be assigned on the chromatogram of the serum sample (Fig. 7I–K), and their glycosylation patterns were in agreement with those in our previous reports [28]. Table 1 summarized LC retention time, *m/z* and charge, relative peak intensities of assigned glycopeptides in LC/MS. No O-glycosylated peptides were detected in this study. It would be due to low amount of O-glycosylation in serum and huge sample complexity.

## 4. Discussion

Alteration of glycans in several serum glycoproteins is a potential marker for several diseases. Several glycomic approaches to the diagnosis using mass spectrometric techniques have been proposed. The most common procedure involves analyzing the liberated glycans by MALDI-TOF MS or LC/ESI-MS, but this method provides no information on the glycosylation sites or protein sources. Another approach involves mass spectrometric analysis of glycopeptides resulting from proteolytic digestion. The enrichment of glycopeptides is useful due to their low ionization efficiency, but loss of glycopeptides cannot be avoidable. In the present study, we performed LC/MS(/MS) with high resolution separation to obtain

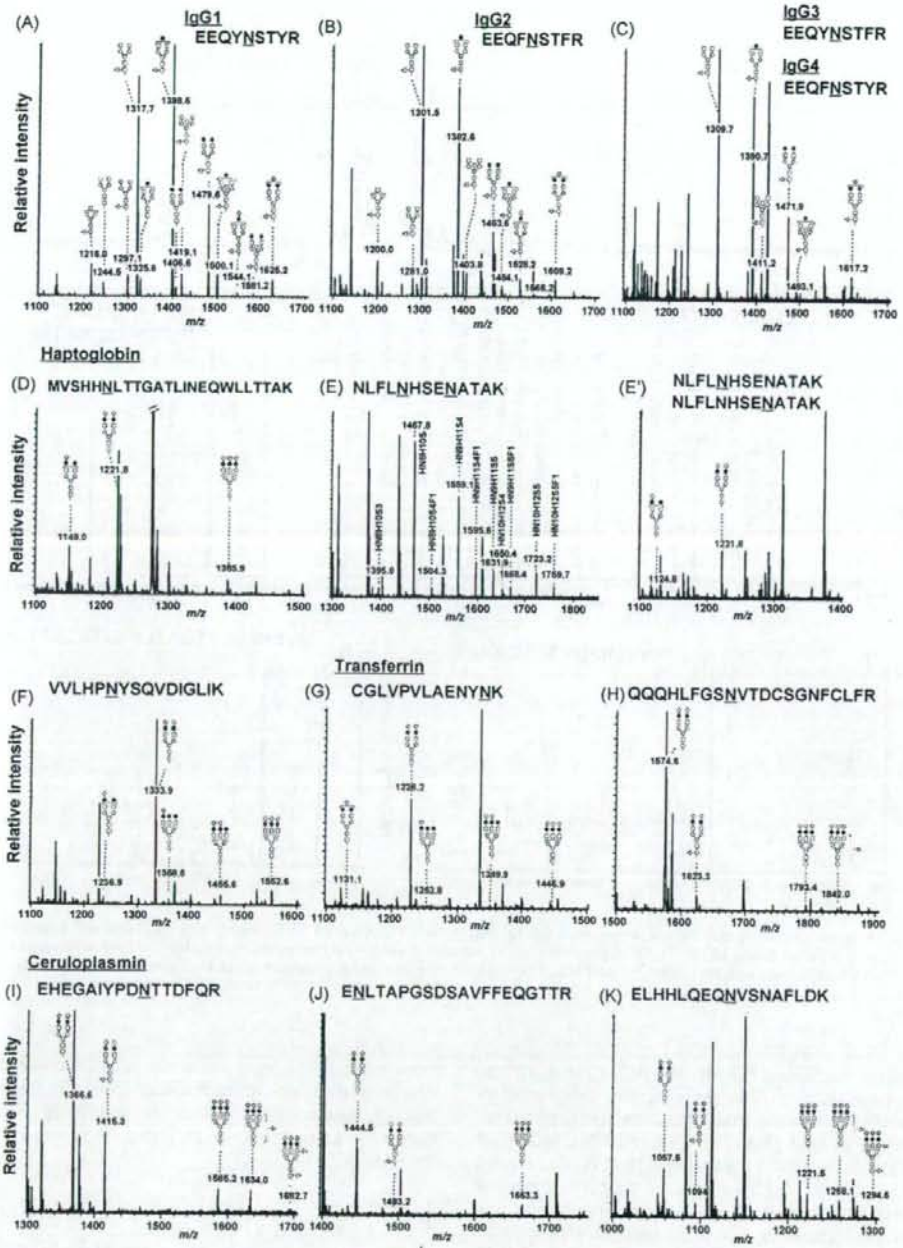



**Fig. 7.** Mass spectrometric heterogeneity of glycosylation in abundant serum glycoproteins. Integrated mass spectra obtained by the additional LC/MS of human serum digest. (A) IgG1; (B) IgG2; (C) IgG3/IgG4; (D–F) haptoglobin ((E) diglycosylated; (E'), monoglycosylated); (G and H) transferrin; (I–K) ceruloplasmin.

mass spectrometric glycosylation profiles at each glycosylation site of abundant glycoproteins in human serum.

MS/MS spectra are useful for detection and assignment of glycopeptide ions. Because MS/MS spectra of glycopeptide precursor ions have abundant carbohydrate B-ions, such as $m/z$ 204 ([HexNAc+H]$^+$), and 366 ([HexHexNAc+H]$^+$), presence of these ions is a useful indication of the selection of glycopeptide precursor ions. MS/MS spectra of glycopeptide also contain ions of peptide and peptide plus glycans and several *b*- and *y*-series fragment ions of peptide backbone when using Qq-TOF mass spectrometer. These allow us to differentiate the glycopeptide ions with different peptide backbone and further to deduce peptide containing

N-linked glycosylation sites by database search. When MS/MS spectra of the glycopeptides obtained in a data-dependent manner were poor for detection of peptide fragment ions, improvement of MS/MS spectra quality by integrating several similar MS/MS spectra into one spectrum was effective. Composition of attached glycan can be deduced from molecular weight of glycan. Utilizing the data of site-specific glycosylation analysis of commercial glycoproteins (IgG, haptoglobin and ceruloplasmin) allowed us assign the corresponding glycopeptides in complex LC/MS(/MS) chromatogram.

We preliminary performed LC/MS/MS of serum tryptic digest to locate glycopeptides and assign by data-dependent MS/MS. Using LC retention time, accurate *m/z* and charge state of assigned glycopeptides, we successfully determined mass spectrometric heterogeneity of 23 glycosylation sites in 15 glycoproteins by LC/MS analyze using digest corresponding 0.3 μl of serum. Although there have been many reports on the analysis of human serum digest to show the glycosylation sites of abundant serum glycoproteins [30–33], less has been reported on their glycosylation. Glycopeptides detected in this study were those derived from glycoproteins which are present at about 0.2–5 mg/ml in human serum, and only glycopeptides with higher ionization efficiency were detected. Thus, it was suggested that detection limit of our method without sample enrichment procedure would be >0.2 mg/ml. It was thought that sample comlexicity, ionization suppression of low abundant glycopeptides and necessity of high quality of MS/MS spectrum for database searching reduced sensitivity. In order to characterize more glycosylation sites, combination of glycopeptide enrichment and depletion of abundant serum proteins is needed.

## Acknowledgements

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

## References

[1] A. Varki, Glycobiology 3 (1993) 97.
[2] Z. Yang, W.S. Hancock, J. Chromatogr. A 1053 (2004) 79.
[3] R.B. Parekh, R.A. Dwek, B.J. Sutton, D.L. Fernandes, A. Leung, D. Stanworth, T.W. Rademacher, T. Mizuochi, T. Taniguchi, K. Matsuta, et al., Nature 316 (1985) 452.
[4] L.A. Omtvedt, L. Royle, G. Husby, K. Sletten, C.M. Radcliffe, D.J. Harvey, R.A. Dwek, P.M. Rudd, Arthr. Rheum 54 (2006) 3433.
[5] R. Dube, G.A. Rook, J. Steele, R. Brealey, R. Dwek, T. Rademacher, J. Lennard-Jones, Gut 31 (1990) 431.
[6] H.H. Freeze, Glycobiology 11 (2001) 129R.
[7] M. Ferens-Sieczkowska, K. Zwierz, A. Midro, I. Katnik-Prastowska, Arch. Immunol. Ther. Exp. (Warsz) 50 (2002) 67.
[8] S. Thompson, E. Dargan, G.A. Turner, Cancer Lett. 66 (1992) 43.
[9] R. Saldova, L. Royle, C.M. Radcliffe, U.M. Abd Hamid, R. Evans, J.N. Arnold, R.E. Banks, R. Hutson, D.J. Harvey, R. Antrobus, S.M. Petrescu, R.A. Dwek, P.M. Rudd, Glycobiology 17 (2007) 1344.
[10] Y. Otake, I. Fujimoto, F. Tanaka, T. Nakagawa, T. Ikeda, K.K. Menon, S. Hase, H. Wada, K. Ikenaka, J. Biochem. (Tokyo) 129 (2001) 537.
[11] B. Kossowska, M. Ferens-Sieczkowska, R. Gancarz, E. Passowicz-Muszynska, R. Jankowska, Clin. Chem. Lab. Med. 43 (2005) 361.
[12] G.A. Turner, Adv. Exp. Med. Biol. 376 (1995) 231.
[13] N. Okuyama, Y. Ide, M. Nakano, T. Nakagawa, K. Yamanaka, K. Moriwaki, K. Murata, H. Ohigashi, S. Yokoyama, H. Eguchi, O. Ishikawa, T. Ito, M. Kato, A. Kasahara, S. Kawano, J. Gu, N. Taniguchi, E. Miyoshi, Int. J. Cancer 118 (2006) 2803.
[14] M.A. Comunale, M. Lowman, R.E. Long, J. Krakover, R. Philip, S. Seeholzer, A.A. Evans, H.W. Hann, T.M. Block, A.S. Mehta, J. Proteome Res. 5 (2006) 308.
[15] J.E. Hansen, J. Iversen, A. Lihme, T.C. Bog-Hansen, Cancer 60 (1987) 1630.
[16] R.R. Drake, E.E. Schwegler, G. Malik, J. Diaz, T. Block, A. Mehta, O.J. Semmes, Mol. Cell Proteomics 5 (2006) 1957.
[17] Y. Wada, M. Tajiri, S. Yoshida, Anal. Chem. 76 (2004) 6560.
[18] M. Tajiri, S. Yoshida, Y. Wada, Glycobiology 15 (2005) 1332.
[19] P. Hagglund, J. Bunkenborg, F. Elortza, O.N. Jensen, P. Roepstorff, J. Proteome Res. 3 (2004) 556.
[20] J.F. Nemeth, G.P. Hochgesang Jr., L.J. Marnett, R.M. Caprioli, Biochemistry 40 (2001) 3109.
[21] J.P. Hui, T.C. White, P. Thibault, Glycobiology 12 (2002) 837.
[22] K. Hakansson, H.J. Cooper, M.R. Emmett, C.E. Costello, A.G. Marshall, C.L. Nilsson, Anal. Chem. 73 (2001) 4530.
[23] O. Krokhin, W. Ens, K.G. Standing, J. Wilkins, H. Perreault, Rapid Commun. Mass Spectrom. 18 (2004) 2020.
[24] U.M. Demelbauer, M. Zehl, A. Plematl, G. Allmaier, A. Rizzi, Rapid Commun. Mass Spectrom. 18 (2004) 1575.
[25] M. Wuhrer, C.H. Hokke, A.M. Deelder, Rapid Commun. Mass Spectrom. 18 (2004) 1741.
[26] A. Harazono, N. Kawasaki, T. Kawanishi, T. Hayakawa, Glycobiology 15 (2005) 447.
[27] S. Itoh, N. Kawasaki, A. Harazono, N. Hashii, Y. Matsuishi, T. Kawanishi, T. Hayakawa, J. Chromatogr. A 1094 (2005) 105.
[28] A. Harazono, N. Kawasaki, S. Itoh, N. Hashii, A. Ishii-Watabe, T. Kawanishi, T. Hayakawa, Anal. Biochem. 348 (2006) 259.
[29] Y. Wada, P. Azadi, C.E. Costello, A. Dell, R.A. Dwek, H. Geyer, R. Geyer, K. Kakehi, N.G. Karlsson, K. Kato, N. Kawasaki, K.H. Khoo, S. Kim, A. Kondo, E. Lattova, Y. Mechref, E. Miyoshi, K. Nakamura, H. Narimatsu, M.V. Novotny, N.H. Packer, H. Perreault, J. Peter-Katalinic, G. Pohlentz, V.N. Reinhold, P.M. Rudd, A. Suzuki, N. Taniguchi, Glycobiology 17 (2007) 411.
[30] J. Bunkenborg, B.J. Pilch, A.V. Podtelejnikov, J.R. Wisniewski, Proteomics 4 (2004) 454.
[31] T. Liu, W.J. Qian, M.A. Gritsenko, D.G. Camp 2nd, M.E. Monroe, R.J. Moore, R.D. Smith, J. Proteome Res. 4 (2005) 2070.
[32] Y. Wang, S.L. Wu, W.S. Hancock, Glycobiology 16 (2006) 514.
[33] P. Hagglund, R. Matthiesen, F. Elortza, P. Hojrup, P. Roepstorff, O.N. Jensen, J. Bunkenborg, J. Proteome Res. 6 (2007) 3021.

J Pharmacol Sci 107, 285 – 294 (2008)

Journal of Pharmacological Sciences


*Full Paper*

# The Novel Compounds That Activate Farnesoid X Receptor: the Diversity of Their Effects on Gene Expression

Takuo Suzuki[1,2,*], Norimasa Tamehiro[3], Yoji Sato[4], Tetsu Kobayashi[1], Akiko Ishii-Watabe[1], Youichi Shinozaki[5], Tomoko Nishimaki-Mogami[3], Toshihiro Hashimoto[6], Yoshinori Asakawa[6], Kazuhide Inoue[7], Yasuo Ohno[5], Teruhide Yamaguchi[1], and Toru Kawanishi[1]

[1]*Division of Biological Chemistry and Biologicals,* [3]*Division of Biosignaling,* [4]*Division of Cellular and Gene Therapy Products,* [5]*Division of Pharmacology, National Institute of Health Sciences, Tokyo 158-8501, Japan*
[2]*Pharmaceuticals and Medical Device Agency, Tokyo 100-0013, Japan*
[6]*Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Tokushima 770-8514, Japan*
[7]*Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka 812-8582, Japan*

*Received January 8, 2008; Accepted May 16, 2008*

**Abstract.** Farnesoid X receptor (FXR) controls the expression of critical genes in bile acid and cholesterol homeostasis. To study FXR and to develop a regulator of cholesterol, some non-steroidal and steroidal ligands have been found in addition to endogenous ligands for FXR. In this study, we discovered five bile acid derivatives (methyl cholate, methyl deoxycholate, 5$\beta$-cholanic acid, 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol, and NIHS700) and two natural products (marchantin A and marchantin E) that activated FXR in the reporter assay. These compounds activated FXR to a high level comparable to the most potent endogenous bile acid, chenodeoxycholic acid, although it was not predicted from their structures; five of them were similar to the lower potency bile acids, and two were structurally much different from bile acids. The elevation levels of reporter gene expression by some of the screened compounds were varied in Cos-7, HepG2, HuH-7, and Caco-2 cells. These compounds also controlled the expression of genes regulated by FXR, and some of the compounds regulated these genes in a cell-type–specific and/or gene-selective fashion. Therefore, molecular design of the compounds can cause selective modulation of the expression of FXR target genes.

***Keywords***: farnesoid X receptor (FXR), reporter assay, ginkgolic acid, marchantin, cell-type–specific modulation

## Introduction

The farnesoid X receptor (FXR, *NR1H4*) is a member of the nuclear-receptor superfamily. Nuclear receptors are ligand-activated transcription factors that are involved in a variety of physiological, developmental, and toxicological processes. The nuclear-receptor superfamily includes receptors for thyroid and steroid hormones, retinoids, and vitamin D, as well as receptors for unknown ligands. These receptors share a highly conserved DNA-binding domain and a discrete ligand-binding domain; and they bind to hormone response elements (HRE) on the DNA during the formation of homodimers, heterodimers, or monomers. The ligand-binding to nuclear receptors leads to a conformational change of these receptors and the recruitment of coactivator complexes, resulting in transcriptional activation (1). Their ligand-dependent activity makes nuclear receptors good pharmacological targets.

FXR is a receptor for bile acids such as chenodeoxycholic acid (CDCA), deoxycholic acid, cholic acid, and their conjugates. Bile acids are synthesized in the liver and secreted into the intestine, where their physical properties facilitate the absorption of fats and vitamins through micelle formation. Cholesterol disposal from the liver is also dependent on the bile acid composition of the secreted bile. FXR is activated by bile acids and controls the expression of critical genes in bile

*Corresponding author (affiliation #1). tsuzuki@nihs.go.jp
Published online in J-STAGE on July 5, 2008 (in advance)
doi: 10.1254/jphs.08006FP

acid and cholesterol homeostasis such as the bile salt export pump (*BSEP*), small heterodimer partner (*SHP*), *CYP7A1*, ileal bile acid-binding protein (*IBABP*), and phospholipid transfer protein (*PLTP*) (2 – 4). FXR plays a critical role in lipid metabolism since FXR-null mice showed elevated serum cholesterol and triglyceride levels (5), and an FXR agonist has been shown to reduce serum triglyceride levels (6). Moreover, an FXR agonist has been reported to confer hepato-protection in a rat model of cholestasis (7). Recently, FXR has also been reported to mediate glucose metabolism and to protect the intestinal mucosa from bacterial overgrowth and inflammatory insults (8, 9). Therefore, the development of FXR agonists might prove useful for the treatment of a wide variety of diseases, including diabetes, cholesterol gallstones, and hepatic and intestinal toxicity.

In addition to bile acids, some compounds whose structures are much different from bile acids (e.g., GW4064) and several selective modulators (e.g., guggulsterone and AGN34) that regulate a subset of FXR-specific genes have been identified as FXR ligands (10, 11). These compounds other than bile acids are useful for analysis of the role of FXR in lipid and glucose metabolism because they may not have the FXR-independent property of bile acids (e.g., dietary lipid absorption) and are not metabolized to form harmful lithocholic acid. On the other hand, selective ligands have been studied in detail regarding the other nuclear receptors (e.g., selective estrogen-receptor modulators (SERMs)). These compounds exhibit variable effects (e.g., function as agonists or antagonists) depending on the cells and tissues, and they have been used in therapy (e.g., tamoxifen and raloxifen). Because FXR has been found to play many roles in addition to lipid metabolism, selective FXR modulators might be useful for therapy.

We previously reported the reporter assay system of FXR, RAR, and RXR using green fluorescent protein derivatives (12). We screened a compound library (NIHS library containing about 700 compounds) and found five bile acid derivatives (methyl cholate, methyl deoxycholate, 5$\beta$-cholanic acid, 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol, and NIHS700) and two natural products (marchantin A and marchantin E) as FXR activators. Concerning these seven compounds and ginkgolic acids that we previously showed as FXR activators, we investigated the FXR activation by reporter assay in four types of cells and the expression of the genes regulated by FXR. These compounds activated FXR comparably to the most potent bile acid, and some controlled the expression of genes regulated by FXR in a cell-type–specific and/or gene-selective fashion.

## Materials and Methods

### *Chemicals*

Methyl cholate, methyl deoxycholate, 5$\beta$-cholanic acid, and 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol were purchased from Steraloids, Inc. (Newport, RI, USA). NIHS700 was provided from Research Foundation Itsuu Laboratory (Tokyo). Ginkgolic acid 15:1 was purified *Ginkgo biloba L.* var. *diptera* as described previously (12). Ginkgolic acid 17:1 was purchased from Nagara Science (Gifu). Chenodeoxycholic acid was purchased from Sigma-Aldrich (St. Louis, MO, USA). Cholic acid, deoxycholic acid and lithocholic acid were purchased form Wako (Osaka).

### *Isolation of marchantins A and E from Marchantia paleacea var. diptera*

Fresh material (6.67 kg) of *Marchantia paleacea* var. *diptera* collected in Tokushima, Japan in 1993 was extracted with MeOH (10 L) for 1 month at room temperature. The extract was filtered and evaporated *in vacuo* to afford a brown residue (176.0 g), which was subjected repeatedly to column chromatography (CC) on silica gel (*n*-hexane-EtOAc, gradient) and Sephadex-LH-20 ($CHCl_3$-MeOH = 1:1) to afford marchantin A (79.5 g) and marchantin E (8.34 g) (13, 14).

### *Plasmid construction*

The construction of plasmids for the reporter assay using green fluorescent protein (GFP) derivatives has been described in a previous report (12). For expression of FXR and RXR$\alpha$, the ORF region of human FXR or human RXR$\alpha$ (accession number U68233, X52773) was inserted into pcDNA3.1 (Invitrogen, Carlsbad, CA, USA). For reporter plasmids, the FXR response element (4 copies of DR1: ggatccaaactgaGGGTCAgTGACCC aagtgaaaaactgaGGGTCAgTGACCCaagtgagaattcaaact gagGGTCAgTGACCCaagtgaaaaactgaGGGTCAgTGA CCCaagtgaagatct), the 3' region (201 bp) of cytomegalovirus (CMV201) promoter, and enhanced yellow fluorescent protein (EYFP) were ligated. As an internal control plasmid, the luciferase gene of pGL3-Control Vector (Promega, Madison, WI, USA) was replaced with enhanced cyan fluorescent protein (ECFP).

For a reporter assay using luciferase, FXRE was inserted into the *Mlu*I and *Bgl*II sites of pGL3-Control Vector, and the SV40 promoter was replaced with minimal CMV promoter. The phRL-CMV vector (Promega) was used as an internal control vector.

### *The reporter assay using GFP derivatives*

A monkey kidney cell line, COS-7, was kept in DMEM (Sigma-Aldrich) with penicillin (100 unit/ml),

streptomycin (100 μg/ml), and 10% FBS. Transfections were performed using Effectene transfection reagent (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The ratio of the reporter plasmid, FXR expression plasmid, RXR expression plasmid, and the internal control plasmid was 4:1:1:1. The culture medium was replaced with DMEM without phenol red (GIBCO, Carlsbad, CA, USA) supplemented with 10% charcoal-treated FBS (Hyclone, Logan, UT, USA) when the transfections were performed. At 15 h after transfection, the cells were treated with trypsin-EDTA (GIBCO) and divided among wells of a black 96-well plate with 100 μl of the culture medium. At 6 h after division among wells, the cells were treated with chemicals. After 40-h incubation, the medium was eliminated by decantation, the cells were washed twice with PBS, and the wells were filled with 200 μl PBS. Fluorescence was detected using a microplate reader (ARVO; Perkin Elmer, Fremont, CA, USA). The fluorescence of EYFP was detected with an excitation filter of 485 nm and an emission filter of 545 nm, and that of ECFP was detected with filters of 420 nm and 486 nm (Perkin Elmer), respectively. The autofluorescence in COS-7 cells was subtracted from each of the detected fluorescences, and the EYFP/ECFP ratio was calculated using the resulting values.

*The reporter assay using luciferase*

The human hepatocyte cell line HepG2 was kept in MEM (Sigma-Aldrich) with penicillin (100 unit/ml), streptomycin (100 μg/ml), and 10% FBS. The cells were transfected with 3 times more plasmids than recommended with salmon sperm DNA (200 ng for 1 well of a 6-well plate) using Effectene transfection reagent (Qiagen). In contrast, the human hepatocyte cell line HuH-7 was kept in DMEM with penicillin (100 unit /ml), streptomycin (100 μg/ml), and 10% FBS; and the human intestinal cell line Caco-2 was kept in DMEM with penicillin (100 unit/ml), streptomycin (100 μg /ml), 10% FBS, and 100 μM MEM Non-Essential Amino Acids Solution (GIBCO). HuH-7 and Caco-2 cells were transfected with plasmids using Effectene transfection reagent (Qiagen) according to the manufacturer's instructions. The ratio of the reporter plasmid using luciferase, the FXR expression plasmid, RXR expression plasmid, and the internal control plasmid using *renilla* luciferase was 4:1:1:1. FBS of the culture medium was replaced with charcoal-treated FBS (Hyclone) when the transfections were performed. The following manipulations were the same as those used in the reporter assay with GFP derivatives. After 40-h treatment with the compounds, measurement of luciferase and *renilla* luciferase was performed with the Dual-Glo™ Luciferase Assay System (Promega) according to the manufacturer's instructions.

*TaqMan primers and probes*

Oligonucleotide primers and probes for human *BSEP*, *SHP*, and *CYP7A1* were synthesized by Applied Biosystems (Foster City, CA, USA). These sequences (5' to 3') were as follows: Human *BSEP*, forward primer (GGGCCATTGTACGAGATCCTAA), probe (6FAM-TCTTGCTACTAGATGAAGCCACTTCTGCCTTAGA-TAMRA) and reverse primer (TGCACCGTCTTTT CACTTTCTG); Human *SHP*, forward primer (GGTG CAGTGGCTTCAATGC), probe (6FAM-TCTGGAG CCTGGAGCTTAGCCCCA-TAMRA), and reverse primer (GGTTGAAGAGGATGGTCCCTTT); Human *CYP7A1*, forward primer (GAGAAGGCAAACGGGT-GAAC), probe (6FAM-TGGATTAATTCCATACCTG GGCTGTGCTCT-TAMRA), and reverse primer (GGT ATGACAAGGGATTTGTGATGA). The primers and probe for 18S rRNA were also purchased from Applied Biosystems.

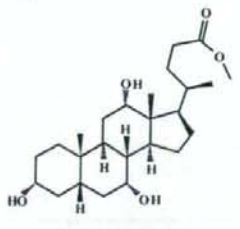
*RNA isolation and real-time quantitative PCR*

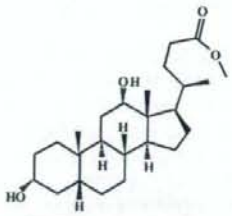
The culture medium of HepG2, HuH-7, and Caco-2 cells was replaced with the medium supplemented with 10% charcoal-treated FBS (Hyclone) at 24 h before treatment with the compounds. The cells were treated with the tested chemicals for 24 h, and total RNA was then prepared using the RNeasy purification system according to the manufacturer's instructions (Qiagen). Reverse transcription reactions and TaqMan-PCRs were performed using the High-Capacity cDNA Archive Kit (Applied Biosystems) and the TaqMan Universal PCR Master Mix (Applied Biosystems) according to the manufacturer's instructions. Sequence-specific amplification was quantified with the ABI Prism 7700 sequence detection system (Applied Biosystems), and values were normalized to 18S rRNA.

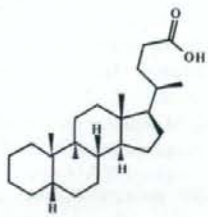
## Results

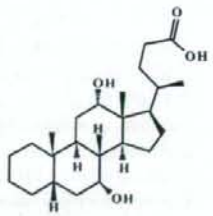
We found seven compounds that activate FXR (the chemical structures of which are shown in Fig. 1A) using the reporter assay system described by Suzuki et al. (12). In the reporter assay system, two fluorescent proteins, EYFP and ECFP, were used for detection of FXR activation and as an internal control, respectively. The activation of FXR by the seven compounds and some endogenous ligands [i.e., cholic acid (CA), deoxycholic acid (DCA), lithocholic acid (LCA), and chenodeoxycholic acid (CDCA)] is shown as the increased ratio of EYFP/ECFP fluorescence intensity in the upper panel of Fig. 2A. As a control, the reporter vector

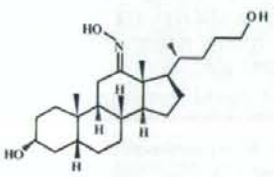
A

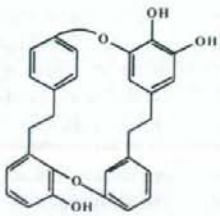


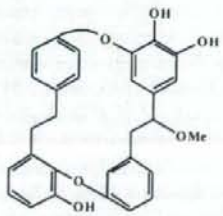



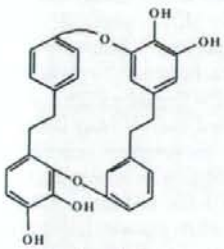



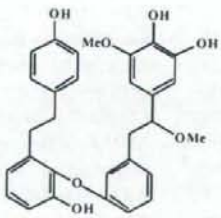



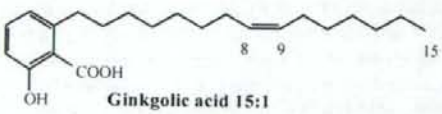




B




C



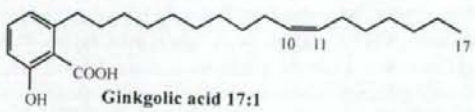




**Fig. 1.** Structure of the compounds that activated FXR and their related compounds. A: Structure of the compounds that activated FXR as determined by the reporter assay. B: The compounds similar to marchantin A and marchantin E. C: Ginkgolic acid 15:1 and ginkgolic acid 17:1 that highly activated FXR described in Ref. 12.

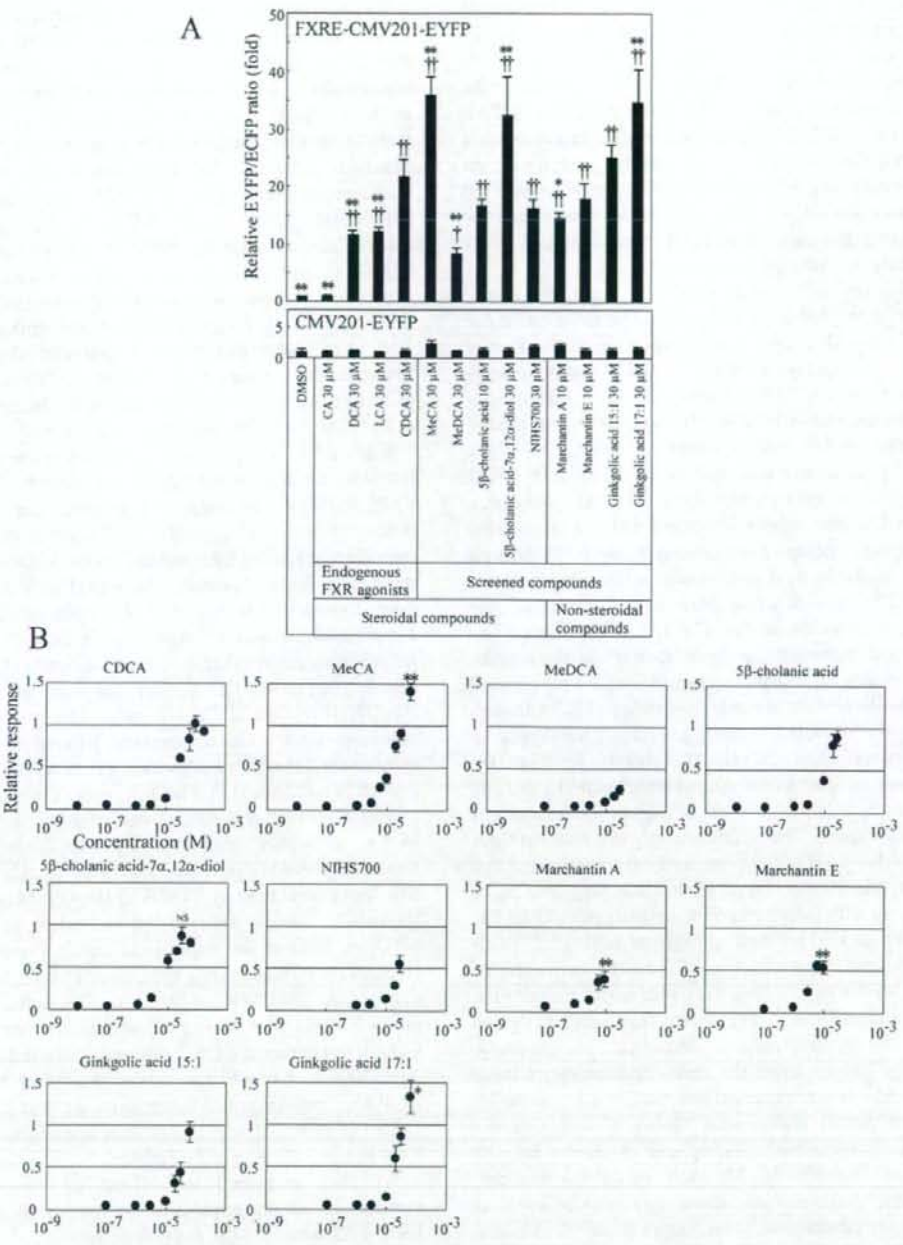



**Fig. 2.** Response in the reporter expression. A: The activation of FXR. COS-7 cells were transfected with the reporter plasmid containing FXRE, the expression plasmids of FXR and RXR$\alpha$, and the internal control plasmid. The transfected cells were treated with each compound. All values are means ± S.D., n = 4. $^{\dagger}P<0.05$ *vs* DMSO, $^{\dagger\dagger}P<0.01$ *vs* DMSO, $*P<0.05$ *vs* CDCA (30 $\mu$M), $**P<0.01$ *vs* CDCA (30 $\mu$M), according to Dunnett's test. B: Dose–response analyses on the reporter assay of FXR. The response was shown as a ratio, compared with that by CDCA at 100 $\mu$M. All values are means ± S.D., n = 4. $*P<0.05$, $**P<0.01$; NS, not a significant difference, compared with CDCA at 100 $\mu$M according to Student's *t*-test.

without FXR response element was used in place of the reporter vector to determine the response by the unexpected factors (the change in the transcriptional efficiency unrelated to FXR, the self-fluorescence of the tested chemicals, and so forth) (lower panel of Fig. 2A). Moreover, it was confirmed that these seven compounds activated the expression of the reporter gene EYFP via FXR because the induction of EYFP expression required cotransfection of an FXR expression vector and the compounds did not activate RXR homodimer and RAR-RXR heterodimer (data not shown).

The seven compounds could be separated into 2 groups: those that contained steroid skeletons and those that did not (Fig. 1A). The compounds in the former group were methyl cholate (MeCA), methyl deoxycholate (MeDCA), 5$\beta$-cholanic acid, 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol, and NIHS700. Methyl cholate and methyl deoxycholate are methyl esters of endogenous FXR ligands (i.e., cholic acid and deoxycholic acid). FXR was also activated by 5$\beta$-cholanic acid and 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol, whose structures differed only in the 3$\alpha$-hydroxyl group from endogenous FXR ligands (i.e., lithocholic acid and cholic acid). The structure of NIHS700 was different from that of lithocholic acid at the substituents of the 11- and 24-positions. The compounds in the latter group, non-steroidal compounds, were marchantin A and marchantin E that were isolated from the liverwort Marchantia species (13). Although these two compounds strongly activated FXR (Fig. 2A), marchantin B and paleatin B (shown in Fig. 1B), analogues of marchantin A and marchantin E, did not cause such activation, even at higher concentrations (data not shown). The difference between marchantin A and marchantin B is only one hydroxyl group, and that between marchantin E and paleatin B is ring-opening or not. Along with the seven compounds shown in Fig. 1A, ginkgolic acid 15:1 and ginkgolic acid 17:1 (their structures are shown in Fig. 1C and the activation of FXR is shown in Fig. 2A), the major constituents of the crude extracts from ginkgo leaves that highly activated FXR (12) were also studied in detail in the present study.

In Fig. 2B, we show the dose-dependency of these compounds as compared to the maximal response by the most potent endogenous ligand, CDCA. All the compounds were hydrophobic and could not be dissolved at over 30 – 60 $\mu$M, each, in culture medium. Moreover, responses by some compounds even at soluble concentrations (Marchantin E at 30 $\mu$M and MeDCA at 60 $\mu$M) could not be measured due to their toxicity. Therefore, we could not obtain the data at higher concentrations. Although the value of treatment with 30 $\mu$M 5$\beta$-cholanic acid is shown in Fig. 2B, this compound slightly separated out from the solution at this concentration. Therefore, the concentration applied in the other experiments was limited to 10 $\mu$M. Since the $EC_{50}$ values of 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol, marchantin A, and marchantin E could be estimated to be 3 – 10, 3 – 6, and 3 – 6 $\mu$M, respectively, the values were lower than that of CDCA (approximately 30 $\mu$M). Maximal activation by 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol was comparable to that by CDCA, but those by the marchantins were lower ($P$<0.01). The $EC_{50}$ and maximal values of MeCA, MeDCA, 5$\beta$-cholanic acid, NIHS700, and ginkgolic acids could not be assessed because of low solubility or toxicity. However, activations of FXR by 60 $\mu$M of MeCA and ginkgolic acid 17:1 were higher than maximal activation by CDCA. Since the $EC_{50}$ values of the 6 compounds could not be estimated, we analyzed activations by all the compounds at 30 or 10 $\mu$M in the following experiments.

Since FXR was primarily localized in the liver and intestine, we determined that the compounds also act as FXR activators in cultured hepatoma and intestinal cancer cells (HepG2, HuH-7, and Caco-2 cells). The reporter assay using GFP derivatives could be applied to these cell lines. However, the signal of the internal control was less than in COS-7 cells because the internal control plasmid was not replicated in the cells without large T antigen. We therefore used luciferase (LUC) and *renilla* luciferase (RLUC) for the reporter assay in HepG2, HuH-7, and Caco-2 cells. The extent of the induction with some compounds differed among the cells on the basis of the activation extent by endogenous ligands (Figs. 2A and 3). The differences are enumerated below. Reporter expression was strongly induced by MeCA to higher level than by CDCA in COS-7, moderately induced in HuH-7, and only weakly induced to a lower level than by CDCA in HepG2. Marchantins (A and E) scarcely induced the expression in HepG2. MeDCA induced the expression higher than LCA in HepG2 ($P$<0.05, according to Student's $t$-test), but lower than LCA in COS-7 ($P$<0.01). The induction by ginkgolic acid 17:1 in HepG2 was lower than that by CDCA, but higher in COS-7. Moreover, the induction by NIHS700 in Caco-2 was lower than that by LCA ($P$<0.05, according to Student's $t$-test), but higher in COS-7 ($P$<0.01). These compounds seemed to activate FXR in a cell-type–specific fashion.

We then examined the effects of the screened compounds on the expression of genes regulated by FXR. FXR controls the expression of critical genes in bile acid and cholesterol homeostasis (2 – 4). In this experiment, we detected the expression of three genes, bile salt export pump (*BSEP*), small heterodimer partner (*SHP*), and *CYP7A1*. By FXR activation, the expression of *BSEP* and *SHP* genes are directly upregulated, and

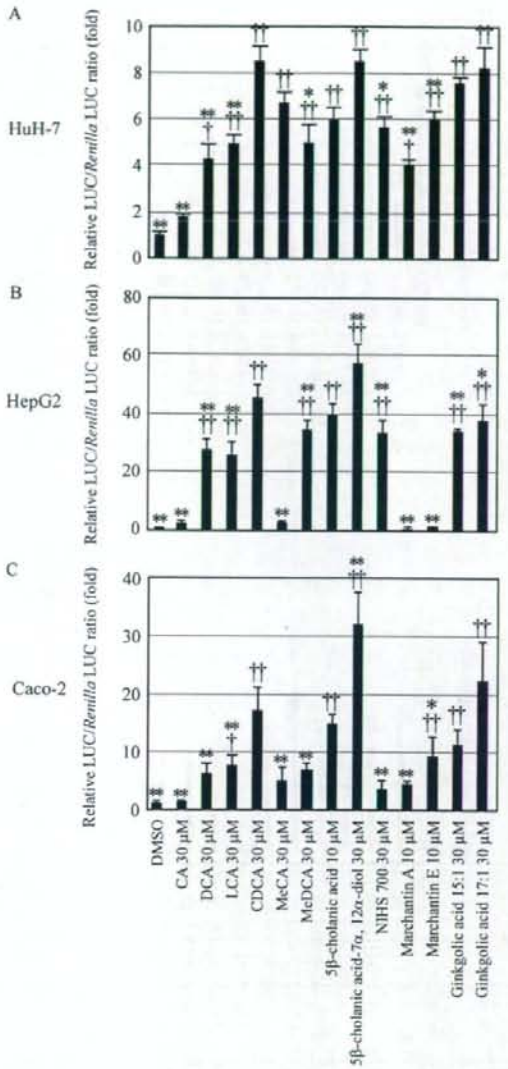



**Fig. 3.** The difference of FXR-activation in the reporter assay. A: The activation of FXR in HuH-7 cells is shown. HuH-7 cells were transfected with the reporter plasmid containing FXRE and luciferase, the expression plasmids of FXR and RXR$\alpha$, and the internal control plasmid containing *renilla* luciferase. The transfected cells were treated with each compound. Data are shown as the means ± S.D. derived from four wells. †$P$<0.05 *vs* DMSO, ††$P$<0.01 *vs* DMSO, *$P$<0.05 *vs* CDCA (30 $\mu$M), **$P$<0.01 *vs* CDCA (30 $\mu$M), according to Dunnett's test. The activation of FXR in HepG2 and Caco-2 cells is shown in the middle and lower panel, respectively. Except for the cell type, all manipulations were the same as those shown above. B: HepG2 cells, n = 4. C: Caco-2 cells, n = 4 – 5.

that of *CYP7A1* gene is indirectly down-regulated. HepG2 cells without transfection of artificial genes were treated with the compounds, and the amount of the mRNAs was measured by real-time quantitative PCR. The expression of the genes in HepG2 cells is shown in Fig. 4 (A – C). The compounds other than MeCA, 5$\beta$-cholanic acid, marchantin A, and marchantin E significantly induced the expression of *SHP* mRNA ($P$<0.05 *vs* DMSO, according to Dunnett's test). Because *BSEP* was expressed only at low levels in HepG2 cells, Fig. 4C showed scattered results. However, the change in *BSEP* expression was similar to the change in *SHP* expression. On the other hand, the reduction levels of *CYP7A1* mRNA by some compounds differed from the change of *SHP* mRNA levels. The reduction in *CYP7A1* mRNA accumulation by MeDCA was lower than that by CDCA, although the compound induced *SHP* mRNA accumulation to a higher level than CDCA. Moreover, marchantin A and E reduced *CYP7A1* mRNA accumulation ($P$<0.05 *vs* DMSO), although the regulation of *SHP* mRNA accumulation was not detected.

Moreover, regulation of gene expression in a cell-type–specific fashion was also detected (Fig. 4D). The induction of *BSEP* expression by MeCA in HuH-7 cells was higher than that of *BSEP* and *SHP* expression in HepG2 cells, and that by MeDCA was lower. These results indicate that the screened compounds could regulate the expression of critical genes in bile acid and cholesterol homeostasis and suggested that MeCA, MeDCA, marchantin A, and marchantin E possibly regulate expression of the genes in a cell-type–specific and/or gene-species selective fashion.

## Discussion

In this paper, we found five steroidal compounds and two non-steroidal compounds that activate FXR. These compounds possess some properties that differ from those of CDCA.

First, two of the steroidal compounds, 5$\beta$-cholanic acid and 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol, were found to be effective ligands for FXR in a reporter assay and in quantitative real-time PCR. The crystal structure of the FXR ligand binding domain was analyzed as a complex with 3-deoxyCDCA and revealed that the 3-hydroxyl group was not responsible for the activation of FXR (15). However, FXR was activated to lower levels by lithocholic acid and cholic acid, whose structures were different only in the 3$\alpha$-hydroxyl group from 5$\beta$-cholanic acid and 5$\beta$-cholanic acid-7$\alpha$,12$\alpha$-diol. In this case, elimination of the 3-hydroxyl group increased the potency of the activation.

The other three steroidal compounds were MeCA,

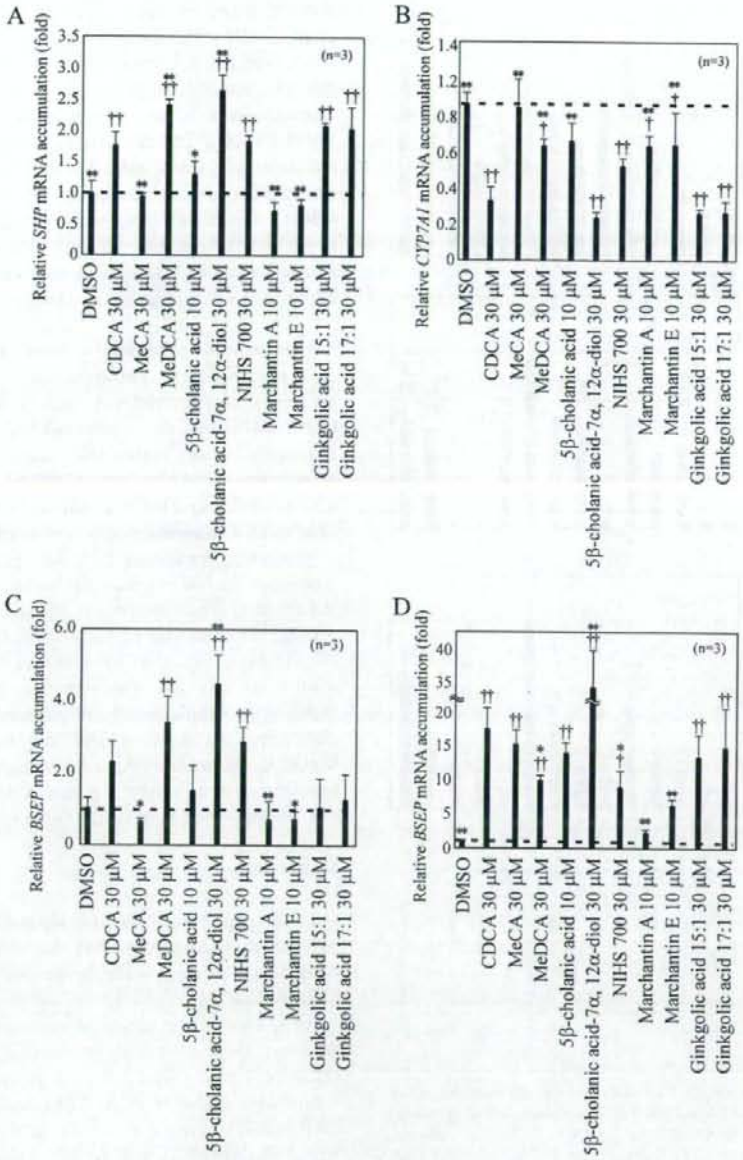



**Fig. 4.** Expression of *SHP*, *CYP7A1*, and *BSEP* genes in cells treated with the compounds. HepG2 cells were treated with each compound for 24 h. Accumulation of *SHP* (A), *CYP7A1* (B), and *BSEP* (C) mRNA in HepG2 cells was detected by real-time quantitative PCR. D: Accumulation of *BSEP* mRNA in HuH-7 cells treated with each compound for 24 h. All values are means ± S.D., n = 3. †$P<0.05$ *vs* DMSO, ††$P<0.01$ *vs* DMSO, *$P<0.05$ *vs* CDCA (30 μM), **$P<0.01$ *vs* CDCA (30 μM), according to Dunnett's test.

MeDCA, and NIHS700. Although the structures of these compounds were similar to those of the endogenous ligands, the slight modification of bile acids caused different regulation of FXR. It has been suggested that

the binding of bile acids with a slightly different structure (CA, DCA, UDCA, and CDCA) resulted in different FXR conformations, which in turn differentially regulated expression of individual FXR targets (16). Therefore, these compounds may also produce different FXR conformations than that produced by CDCA. In fact, MeDCA showed different properties in the regulation of the genes as compared to CDCA. DCA, CA, and UDCA have been reported to partially increase *BSEP* expression, but repress *CYP7A* mRNA with nearly equal effects as CDCA (16). Adversely, MeDCA strongly induced *BSEP* mRNA, but only weakly reduced *CYP7A1* mRNA (Fig. 4: B and C). It is possible that the *CYP7A1* expression was also influenced by other factors such as c-jun N-terminal kinase and xenobiotic receptor (17, 18). However, since MeDCA possibly induces the export of bile acids from hepatocytes without disturbing catabolization from cholesterol to bile acids, they might effectively improve the disorder of cholesterol and bile acids. NIHS700 seemed to have properties similar to those of MeDCA, although statistically significant differences could not be observed by Dunnett's test. Therefore, further analyses about MeDCA and NIH700 are needed for effective regulation of FXR.

Since MeCA and MeDCA are methyl esters of endogenous FXR ligands, there is a question about whether these compounds are hydrolyzed and act as CA or DCA. Considering that MeDCA showed different properties in the gene expression as compared to DCA, these compounds did not seem to act as hydrolyzed forms.

The compounds screened in this study contained several non-steroidal chemicals. Marchantin A was isolated from the liverwort *Marchantia* species, as its major component (13). It shows antifungal, antimicrobial, cytotoxic, muscle-relaxing, and 5-lipoxygenase, cyclooxygenase, and calmodulin inhibitor activities [reviewed in by Asakawa et al. (19)]. The compounds with a slight structural change from marchantin A and E (i.e., marchantin B and paleatin B) did not activate FXR. It may be interesting to examine the mechanism of binding to FXR and that of the subsequent regulation of the target genes. Since these two compounds and ginkgolic acids have a much different structure from bile acids, the structural change of FXR in response to these compounds might be different from that in response to bile acids, which might result in different regulation pattern of the genes of bile acid homeostasis. In fact, marchantin A and marchantin E activated FXR diversely in each cell, although ginkgolic acids did not show FXR activation in a cell-type specific or gene-selective fashion. Bile acids function in the dietary lipid absorption and were reported to activate mitogen-activated protein kinase pathways without FXR (17, 20). The non-steroidal compounds were thought not to have these functions, although it has not been clear whether these compounds have effects other than FXR activation. Moreover, the non-steroidal compounds are not metabolized to form harmful lithocholic acid (21, 22). Therefore, it is possible that these compounds can be used for studying the pharmacology of FXR.

As described above, some compounds demonstrated properties of cell-type–specific and/or gene-selective modulators. Until now, several compounds have been found to be cell-type–specific and/or gene-selective modulators of FXR (10, 11, 16). Although the mechanism of selective FXR modulation remains to be elucidated, it was suggested that differences in coregulator recruitment play a critical role in cell-type–specific and promoter-specific regulation by other nuclear receptors (23, 24). We therefore investigated the binding capacity of FXR with receptor-interacting domains of the activator for thyroid hormone and retinoid receptors (ACTR), vitamin D–interacting protein 205 (DRIP205), glucocorticoid receptor–interacting protein (GRIP), receptor-interacting protein 140 (RIP140), and steroid receptor coactivator-1 (SRC-1) with a mammalian two-hybrid assay. However, we could not detect enough obvious differences to explain the cell-type (and/or gene) specific modulation (data not shown).

There are the other possibilities about cell-type–specific modulation, although the gene specific modulation could not be explained. First, the compounds are metabolized in cultured cells and their metabolites bind to FXR as ligands. In case that the metabolism varies with cell-type, FXR is differentially-activated. Second, the compounds are inactivated by metabolism in some cells. Third, permeability of the compounds is different according to cell-type. Since all the compounds tested were hydrophobic and separated out in culture medium at lower concentrations as compared to CDCA, the third possibility is unlikely. To reveal the mechanism of selective FXR modulation and to produce a synthetic selective modulator, farther analyses are required.

Finally, FXR has pleiotropic therapeutic potential, but a simple FXR agonist will have undesired side-effects (reviewed in refs. 25 and 26). The compounds discussed in the present paper appear to be useful for studying FXR modulation leading to selective FXR modulation for therapy.

## Acknowledgments

This work was supported by a grant-in-aid (MF-16) from the Pharmaceuticals and Medical Device Agency, a

grant-in-aid for Research on Publicly Essential Drugs and Medical Devices from the Japan Health Science Foundation, and a grant-in-aid from the Ministry of Health, Labour, and Welfare of Japan.

## References


1 Khorasanizadeh S, Rastinejad F. Nuclear-receptor interactions on DNA-response elements. Trends Biochem Sci. 2001;26:384–390.

2 Makishima M, Okamoto AY, Repa JJ, Tu H, Learned RM, Luk A, et al. Identification of a nuclear receptor for bile acids. Science. 1999;284:1362–1365.

3 Parks DJ, Blanchard SG, Bledsoe RK, Chandra G, Consler TG, Kliewer SA, et al. Bile acids: natural ligands for an orphan nuclear receptor. Science. 1999;284:1365–1368.

4 Wang H, Chen J, Hollister K, Sowers LC, Forman BM. Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. Mol Cell. 1999;3:543–553.

5 Sinal CJ, Tohkin M, Miyata M, Ward JM, Lambert G, Gonzalez FJ. Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. Cell. 2000;102:731–744.

6 Maloney PR, Parks DJ, Haffner CD, Fivush AM, Chandra G, Plunket KD, et al. Identification of a chemical tool for the orphan nuclear receptor FXR. J Med Chem. 2000;43:2971–2974.

7 Liu Y, Binz J, Numerick MJ, Dennis S, Luo G, Desai B, et al. Hepatoprotection by the farnesoid X receptor agonist GW4064 in rat models of intra- and extrahepatic cholestasis. J Clin Invest. 2003;112:1678–1687.

8 Zhang Y, Lee FY, Barrera G, Lee H, Vales C, Gonzalez FJ, et al. Action of the nuclear receptor FXR improves hyperglycemia and hyperlipidemia in diabetic mice. Proc Natl Acad Sci U S A. 2006;103:1006–1011.

9 Inagaki T, Moschetta A, Lee Y-K, Peng L, Zhao G, Downes M, et al. Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor. Proc Natl Acad Sci U S A. 2006;103:3920–3925.

10 Cui J, Huang L, Zhao A, Lew J-L, Yu J, Sahoo S, et al. Guggulsterone is a farnesoid X receptor antagonist in coactivator association assays but acts to enhance transcription of bile salt export pump. J Biol Chem. 2003;278:10214–10220.

11 Dussault I, Beard R, Lin M, Hollister K, Chen J, Xiao JH, et al. Identification of gene-selective modulators of the bile acid receptor FXR. J Biol Chem. 2003;278:7027–7033.

12 Suzuki T, Nishimaki-Mogami T, Kawai H, Kobayashi T, Shinozaki Y, Sato Y, et al. Screening of novel nuclear receptor agonists by a convenient reporter gene assay system using green fluorescent protein derivatives. Phytomedicine. 2006;13:401–411.

13 Asakawa Y. Chemical constituents of hepaticae. In: Herz W, Grisebach H, Kirby WG, editors. Progress in the chemistry of organic natural products, Vol. 42, Wien: Spriger; 1982, p. 1–285.

14 Asakawa Y. Biologically active terpenoids and aromatic compounds from liverworts and inedible mushroom *Cryptoporus volvatus*. In: Colegate SM, Molyneux RJ, editors. Bioactive natural products: detection, isolation, and structural determination. Florida: CRC Press; 1993, p. 319–347.

15 Mi L-Z, Devarakonda S, Harp JM, Han Q, Pellicciari R, Willson TM, et al. Structural basis for bile acid binding and activation of the nuclear receptor FXR. Mol Cell. 2003;11:1093–1100.

16 Lew JL, Zhao A, Yu J, De Pedro N, Peláez F, Wright SD, et al. The farnesoid X receptor controls gene expression in a ligand- and promoter-selective fashion. J Biol Chem. 2004;279:8856–8861.

17 Gupta S, Stravitz RT, Dent P, Hylemon PB. Down-regulation of cholesterol 7α-hydroxylase (*CYP7A1*) gene expression by bile acids in primary rat hepatocytes is mediated by the c-Jun N-terminal kinase pathway. J Biol Chem. 2001;276:15816–15822.

18 Staudinger JL, Goodwin B, Jones SA, Hawkins-Brown D, MacKenzie KI, LaTour A, et al. The nuclear receptor PXR is a lithocholic acid sensor that protects against liver toxicity. Proc Natl Acad Sci U S A. 2001;98:3369–3374.

19 Asakawa Y. Recent advances in phytochemistry of bryophytes-acetogenins, terpenoids and bis(bibenzyl)s from selected Japanese, Taiwanese, New Zealand, Argentinean and European liverworts. Phytochemistry. 2001;56:297–312.

20 Qiao L, Han SI, Fang Y, Park JS, Gupta S, Gilfor D, et al. Bile acid regulation of C/EBPβ, CREB, and c-Jun function, via the extracellular signal-regulated kinase and c-Jun $NH_2$-terminal kinase pathways, modulates the apoptotic response of hepatocytes. Mol Cell Biol. 2003;23:3052–3066.

21 Fischer S, Beuers U, Spengler U, Zwiebel FM, Koebe HG. Hepatic levels of bile acids in end-stage chronic cholestatic liver disease. Clin Chim Acta. 1996;251:173–186.

22 Javitt NB. Cholestasis in rats induced by taurolithocholate. Nature. 1966;210:1262–1263.

23 Kodera Y, Takeyama K, Murayama A, Suzuwa M, Masuhiro Y, Kato S. Ligand type-specific interactions of peroxisome proliferator-activated receptor gamma with transcriptional coactivators. J Biol Chem. 2000;275:33201–33204.

24 Shang Y, Brown M. Molecular determinants for the tissue specificity of SERMs. Science. 2002;295:2465–2468.

25 Modica S, Moschetta A. Nuclear bile acid receptor FXR as pharmacological target: are we there yet? FEBS Lett. 2006;580:5492–5499.

26 Cariou B, Staels B. FXR: A promising target for the metabolic syndrome? Trends Pharmacol Sci. 2007;28:236–243.

J. Biochem. 143, 537–545 (2008)
doi:10.1093/jb/mvm248

# Annexin A3 Expression Increases in Hepatocytes and is Regulated by Hepatocyte Growth Factor in Rat Liver Regeneration

**Mizuho Harashima[1], Kayo Harada[1], Yoshimasa Ito[1], Masashi Hyuga[2], Taiichiro Seki[1], Toyohiko Ariga[1], Teruhide Yamaguchi[2] and Shingo Niimi[2,*]**

*[1]Department of Nutrition and Physiology, Nihon University College of Bioresource Sciences, Kameino Fujisawa 252-8510; and [2]Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan*

Received November 14, 2007; accepted December 18, 2007; published online January 7, 2008

**Annexin (Anx) A3 increases and plays important roles in the signalling cascade in hepatocyte growth in cultured hepatocytes. However, no information is available on its expression and role in rat liver regeneration. In the present study, AnxA3 expression was investigated to determine whether it also plays a role in the signalling cascade in rat liver regeneration. AnxA3 protein and mRNA level both increase in liver after administration of carbon tetrachloride ($CCl_4$) or 70% partial hepatectomy. AnxA3 protein level increases in isolated parenchymal hepatocytes, but not in non-parenchymal liver cells, in these rat liver regeneration models. AnxA3 mRNA increases in hepatocytes after $CCl_4$ administration. Anti-hepatocyte growth factor antibody suppresses this increase in AnxA3 mRNA level. These results demonstrate that AnxA3 expression increases in hepatocytes through a hepatocyte growth factor-mediated pathway in rat liver regeneration models, suggesting that AnxA3 plays an important role in the signalling cascade in rat liver regeneration.**

**Key words: annexin A3, carbon tetrachloride, hepatocyte growth factor, parenchymal hepatocytes, partial hepatectomy.**

Abbreviations: Anx, Annexin; $CCl_4$, carbon tetrachloride; HGF, hepatocyte growth factor.

Annexin (Anx) A3 is a member of the Anx family, which binds to phospholipids and membranes in a $Ca^{2+}$-dependent manner (*1–4*). AnxA3 has been shown to have anti-coagulant and anti-phospholipase $A_2$ properties *in vitro* (*5*, *6*), plus to promote $Ca^{2+}$-dependent aggregation of isolated specific granules from human neutrophils (*5*, *6*). Some reports describe its regulation and role in cultured cells (*7–11*); however, there are no reports describing these characteristics *in vivo*.

We recently reported that AnxA3 is expressed in cultured rat hepatocytes, but not in isolated hepatocytes and that inhibition of AnxA3 expression by RNA interference results in a significant inhibition of hepatocyte growth (*10*, *12*, *13*). These findings indicate that AnxA3 plays an important role in the signalling cascade in hepatocyte growth in cultured hepatocytes, although the mechanism remains to be elucidated. The significance of AnxA3 in hepatocyte growth is also supported by the finding that known stimulatory or inhibitory actions of various factors to hepatocyte growth correlated well with the increase or decrease in AnxA3 expression (*14*).

These findings indicate that AnxA3 increases and is likely to play an important role in the signalling cascade in rat liver regeneration. AnxA1 increases in rat and mouse liver regeneration models, *e.g.* after administration of carbon tetracholoride ($CCl_4$) and 70% partial hepatectomy (*15*, *16*). Suppression of AnxA1 expression using anti-sense technology inhibits proliferation in a mouse hepatocyte cell line (*15*). Therefore, AnxA1 is also likely to play an important role in the signalling cascade in rat liver regeneration.

In the present study, AnxA3 expression in rat liver regeneration models was investigated to explore the possibility that AnxA3 plays important roles in the signalling cascade in rat liver regeneration.

## MATERIALS AND METHODS

*Animals and Experimental Conditions*—Adult male Wistar rats (180–200 g) were purchased from Japan SLC Co., Ltd. (Shizuoka, Japan) and used for all studies. They were maintained in a 12 h light/dark cycle, allowed food and water *ad libitium*. All animal care and procedures were approved by the institutional care committee and carried out in accordance with the guidelines established by the National Institute of Health.

For studies of liver regeneration after toxic injury, rats received $CCl_4$ intraperitoneally (2 ml/kg body weight of 50% solution of $CCl_4$ in olive oil). Control rats received olive oil intraperitoneally (1 ml/kg body weight of olive oil). Animals given $CCl_4$ or olive oil were sacrificed at 3–24 h after administration.

A 70% partial hepatectomy was performed according to Higgins and Anderson (*17*). In the sham operation, livers were exposed and manipulated but not removed. These procedures were performed under anaesthesia with Nembutal (Abbot, Chicago, IL, USA). Animals subjected to

*To whom correspondence should be addressed. Tel: +81-3-3700-9347, Fax: +81-3-3700-9084, E-mail: niimi@nihs.go.jp

partial hepatectomy or sham operation were sacrificed at 2.5–20 h after the operation.

For infusion of anti-human hepatocyte growth factor (HGF) antibody, rats were intravenously injected with 0.2 ml goat anti-human HGF IgG (Sigma–Aldrich, St Louis, MO, USA) (1.25 mg/kg body weight) diluted in phosphate-buffered saline (PBS) through the tail vein, then received $CCl_4$ intraperitoneally, as described earlier. Control rats were injected with the same volume and amount of control goat IgG, and then received $CCl_4$ intraperitoneally in a similar manner. Parenchymal hepatocytes were prepared from the rats after 6 h, as described subsequently.

*Preparation of Liver Lysate*—The procedures were performed at low temperature, unless described otherwise. Liver was *in situ* perfused with PBS *via* the portal vein, then removed from the body. Liver was homogenized with a Potter-Elvehjem homogenizer in 4× (v/w) buffer A [50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 10 mM EDTA and 2.5% (v/v) Triton-X 100] containing 1 mM benzylsulphonyl fluoride, 0.3 mM leupeptin and 0.5 mM aprotinin. The homogenate was shaken for 15 min at room temperature, then sonicated four times for 15 s each time. After centrifugation at 100,000*g*, the cytosolic fraction was stored at −70°C until use.

*Cell Isolation*—Parenchymal hepatocytes were isolated from rats by *in situ* perfusion of the liver with collagenase (*18*). Non-parenchymal liver cells were isolated from the supernatant of parenchymal cells by differential centrifugation, as described by Shimaoka *et al.* (*19*). In this article, hepatocytes are also referred to as parenchymal hepatocytes to distinguish between hepatocytes and non-parenchymal liver cells.

*Preparation of Cell Lysate*—Cell lysates were prepared by a modification of the reported by Römisch *et al.* (*20*). Procedures were performed at low temperature, unless described otherwise. Cells were resuspended in three volumes of buffer A containing 1/100 (v/v) protease inhibitor cocktail (Sigma–Aldrich, St Louis, MO, USA). They were then shaken for 15 min at room temperature and sonicated four times for 15 s each time. After centrifugation at 100,000*g*, the cytosolic fraction was stored at −70°C until use.

*Western Blot Analysis*—An equal amount of cytosolic protein from each experiment was subjected to SDS–PAGE on a 10% gel and electroblotted to PVDF membrane (GVHP; Millipore, Bedford, MA, USA). After blocking the membrane with 5% skimmed milk, a western blot analysis was performed using rabbit anti-human AnxA3 antibody serum (1: 5,250) (a gift from Drs F. Russo-Marie and C. Ragueness-Nicol), mouse anti-human GAPDH monoclonal antibody (1: 5,000) (Abcam, Cambridge, UK), or rabbit anti-beta-actin polyclonal antibody (1: 500) (BioLegend, San Diego, CA, USA). Detection was performed using the ECL detection system (GE Health care Bioscience, Buckinghamshire, UK). Housekeeping protein, GAPDH and beta-actin, were selected based on results of preliminary studies. Intensity of each band was measured over a proportional range. A computer-assisted analyser was used to quantitatively analyse intensity, with intensity of the AnxA3 band normalized to the intensity of the appropriate housekeeping protein. Protein amount from liver and cell lysate was measured using a previously described method (*21*), with bovine serum albumin used as a standard.

*Total RNA Extraction and Real-Time Quantitative PCR*—Total RNA was extracted from liver by a modification of guanidine thiocyanate–phenol–chloroform extraction method (*22, 23*). Total RNA was extracted from cells using Trizol® reagent (Invitrogen, Cergy Pontoise, France) in accordance with the manufacture's protocol. Equal amounts of RNA (~1 μg) from each experiments were reverse-transcribed using a THERMOSCRIPT™ RT–PCR System (Invitrogen, Cergy Pontoise, France) and oligo(dT)$_{20}$ in a final volume of 40 μl, in accordance with the manufacturer's protocol. Subsequently, 2 μl of cDNA was used as templates for real-time PCR analysis using a LightCycler system (Rosche Diagnostics, Tokyo, Japan) according to the manufacture's instructions. For AnxA3 and 28S rRNA, the PCR programme consisted of 40 cycles of 10 s at 94°C, 10 s at 60°C and 12 s at 72°C. Primer sequences for AnxA3 were 5′-CAA ATT CAC CGA GAT CCT GT-3′ and 5′-TGC TGG AGT GCT GTA CGA AA-3′ (*14*) and for 28S rRNA 5′-CCA GAG CGA AAG CAT TTG CCA-3′ and 5′-GGC ATC ACA GAC CTG TTA TTG CTC-3′ (*14*). AnxA3 levels were normalized to the levels of 28S rRNA.

*Statistical Analysis*—Data were analysed using Student's *t*-test, and *P*-values <0.05 were considered to be statistically significant.

*Immunohistochemical Examination*—Serial liver sections cut at 3 μm thick from the paraformaldehyde-fixed and paraffin-embedded blocks. De-paraffinated and re-hydrated sections were heated for 5 min at 100°C in 10 mM citrate buffer (pH 6.0) followed by the treatment with 10 μg/ml Proteinase K (TAKARA BIO Inc., Shiga, Japan) for 5 min at room temperature. These activated sections were then subjected to blocking with 10% bovine serum albumin for 1 h at room temperature. After washing with PBS, sections were simultaneously incubated for 2 h with antibodies, *e.g.* anti-rat hepatic sinusoidal endothelial cells mouse IgG (SE-1, Immuno-Biological Laboratories Co., Ltd. Gunma, Japan) 1:20 and rabbit anti-human AnxA3 antibody serum 1:200. The fluorescence-labelled secondary antibodies were AMCA-labelled sheep anti-mouse IgG (Jackson Immuno Research Laboratories, Inc., PA, USA) 1:200 and FITC-labelled sheep anti-rabbit IgG (MP Biomedicals Inc., Ohio, USA) 1:200. The liver sections were thus mounted on a cover glass with a mounting medium, Vectashield (Vector Laboratories, CA, USA), and subjected to microscopic observation.

## RESULTS

*AnxA3 Expression in Liver Following $CCl_4$ Treatment*—AnxA3 protein level increased ~3-fold at 6 h after administration of $CCl_4$ and this increased level was maintained to 24 h (Fig. 1). AnxA3 mRNA level started to increase at 3 h after administration, reaching an ~17-fold increase at 24 h (Fig. 2).