Glycobiology vol. 19 no. 7 pp. 693–706, 2009
doi:10.1093/glycob/cwp025
Advance Access publication on February 24, 2009

Structural analysis of *N*-glycans from gull egg white glycoproteins and egg yolk IgG

Noriko Suzuki[1,2,3], Tseng-Hsiung Su[4], Sze-Wei Wu[5], Kazuo Yamamoto[3], Kay-Hooi Khoo[4,5], and Yuan C Lee[1,2]

[2]Department of Biology, The Johns Hopkins University, Baltimore, MD, 21218, USA; [3]Department of Integrated Biosciences, Graduate School of Frontier Sciences, University of Tokyo, Chiba, 277-8562, Japan; [4]Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan; and [5]NRPGM Core Facilities for Proteomics and Glycomics, Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei, 115, Taiwan

Received on January 7, 2009; revised on February 16, 2009; accepted on February 16, 2009

We previously showed that the expression of (Galα1-4Gal)-bearing glycoproteins among birds is related to their phylogeny. However, precise structures of (Galα1-4Gal)-containing *N*-glycans were only known for pigeon egg white glycoproteins and IgG. To compare structural features of (Galα1-4Gal)-containing *N*-glycans from other species, we analyzed *N*-glycans of gull egg white (GEW)-glycoproteins, ovomucoid, and ovotransferrin, and gull egg yolk IgG by HPLC, mass spectrometry (MS), and MS/MS analyses. GEW-glycoproteins included neutral, monosialyl, and disialyl *N*-glycans, and some of them contained Galα1-4Gal sequences. Bi-, tri-, and tetra-antennary oligosaccharides that lacked bisecting GlcNAc were the major core structures, and incomplete α-galactosylation and sialylation as well as the presence of diLacNAc on the branches generated microheterogeneity of the *N*-glycan structures. Moreover, unlike pigeon egg white glycoproteins, the major sialylation in GEW-glycoproteins is α2,3-, but not α2,6-linked sialic acids (NeuAc). In addition to the complex-type oligosaccharide, hybrid-type oligosaccharides that lack bisecting GlcNAc were also abundant in GEW-glycoproteins. Gull egg yolk IgG also contained Galα1-4Galβ1-4GlcNAcβ1- sequences, but unlike pigeon IgG, no Galα1-4Galβ1-4Galβ1-4GlcNAcβ1- sequence was detected. Bi- and tri-antennary complex-type oligosaccharides with bisecting GlcNAc and with core fucosylation as well as high-mannose-type oligosaccharides were the major structures in gull IgG. Our data indicated that some *N*-glycans from both GEW-glycoproteins and gull IgG contain the Galα1-4Galβ1-4GlcNAcβ1- sequence, but the ratio of α-Gal-capped residues to non-α-Gal-capped residues in the nonreducing termini of *N*-glycans is much lower than that in those of pigeon glycoproteins.

Keywords: egg yolk IgG/Galα1-4Galβ1-4GlcNAc/galabiose/glycan diversity/MS analysis

Introduction

Various carbohydrate structures attached to glycoproteins and/or glycolipids are often species-specific. While the glycans that are conserved among species are usually indispensable for the individual organism, the role of species-specific carbohydrates is not clearly understood. Several indirect investigations suggest that they may exist for defense against infection of bacteria, viruses, and parasites (Gagneux and Varki 1999; Hooper and Gordon 2001) since the carbohydrate-binding proteins of pathogens often target certain glycans expressed in animals during the first step of invasion. It is assumed that the surface glycans of oligosaccharide structures of some hosts might have changed their ability to evade infection by pathogens. This assumption is, however, based on limited information. Systematic investigations of species-specific glycan expression in nature are required to unravel the mechanisms underlying the generation, distribution, and biological roles of glycan differentiation acquired during animal evolution and biodiversity.

We previously investigated that the species-specific expression of Galα1-4Gal on glycoproteins is related to the phylogeny of birds (Suzuki, Laskowski, et al. 2004; Suzuki et al. 2006; Suzuki and Lee 2007). Modern birds (Neornithes) are monophyletic and split into three large taxa, namely Ratitae (traditionally called Palaeognathae, e.g., ostrich, emu), Galloanserae (e.g., chicken, quail, duck), and Neoaves (e.g., pigeon, swiftlet, gull), at an early stage in modern bird history (100–65 million years ago, or earlier). Our investigation revealed that Galα1-4Gal on egg white glycoproteins was absent in Ratitae and Galloanserae, and present in about two-thirds of Neoaves, based on the detection by *Griffonia simplicifolia* I (GS-I lectin, specific for terminal α-Gal/GalNAc) and anti-P_1 (specific for Galα1-4Galβ1-4GlcNAcβ1-) mAb. Since about 95% of avian species belong to Neoaves, the majority of the modern birds may express Galα1-4Gal on glycoproteins. In contrast, Galα1-4Gal is rarely found in mammalian glycoproteins, but occurs on the glycolipids of many mammals (Gahmberg and Hakomori 1975; Nakamura et al. 1984; Lyerla et al. 1986; Kotani et al. 1994; Yang et al. 1994). Although the physiological functions of Galα1-4Gal in mammals are unknown (Okuda et al. 2006), Galα1-4Gal is known as a minimum target of host cells for infections by some pathogenic microbes, such as uropathogenic *Escherichia coli* (Stapleton et al. 1998), and enterotoxins, such as Shiga toxin (Lindberg et al. 1987; Lingwood et al. 1987). Moreover, the cell surface expression of globotriaosylceramide (Gb3), which is one of the glycolipids containing Galα1-4Gal, is reported to protect against HIV infection (Lund et al. 2009; Ramkumar et al. 2009).

While Galα1-4Gal has been detected on avian glycoproteins by specific lectins and antibodies (Suzuki, Laskowski, et al. 2004), only a few documented structures of *N*-glycans containing Galα1-4Gal are available. One of the most exhaustive analyses of *N*-glycan was conducted on pigeon

[1]To whom correspondence should be addressed: Noriko Suzuki, Tel: +81-4-7136-3617; Fax: +81-4-7136-3619; e-mail: nrsuzuki@k.u-tokyo.ac.jp; Yuan C. Lee, Tel: +1-410-516-7041; Fax: +1-410-516-8716; e-mail: yclee@jhu.edu

(*Columba livia*) egg white (PEW) glycoproteins. All four major PEW-glycoproteins, namely ovotransferrin, two ovalbumins, and ovomucoid contain Galα1-4Gal (Suzuki et al. 2001), and all of their constituent *N*-glycans were found to possess Galα1-4Galβ1-4GlcNAc branches of tri-, tetra-, and penta-antennary structures (Takahashi et al. 2001). No sialylation is found on the branch that contains the Galα1-4Gal sequence. Analysis of the *N*-glycan structures of pigeon serum IgG demonstrated that Galα1-4Gal is not restricted to egg white glycoproteins. In addition to the presence of Galα1-4Galβ1-4GlcNAc branches, we identified novel *N*-glycan structures possessing Galα1-4Galβ1-4Galβ1-4GlcNAc branches (Suzuki et al. 2003) in pigeon IgG. Thus, a detailed structural analysis of oligosaccharides will sometimes provide unexpected discoveries of new structures and produce more useful information about the biosynthesis of oligosaccharides.

To further understand the characteristics of species-specific oligosaccharides, we analyzed *N*-glycans containing Galα1-4Gal from gull egg white (GEW) glycoproteins and gull egg yolk (GEY) IgG (also called IgY). According to the classification of avian species based on DNA–DNA hybridization, gulls belong to the order Ciconiiformes (Sibley and Ahlquist 1990; Sibley and Monroe 1990). In the traditional classification, gulls belong to the order Charadriiformes (Sibley and Ahlquist 1990; Chu 1995; Paton et al. 2003). Both Ciconiiformes and Charadriiformes belong to Neoaves, as do Columbiformes (which includes pigeons), but gull and pigeon are phylogenetically distant from each other. Here we report the *N*-glycan structures of GEW-glycoproteins and GEY-IgG from black-tailed gull (*Larus crassirostirs*), as determined by HPLC, mass spectrometry (MS), and MS/MS analysis. As expected, some structural features of these *N*-glycans from gull, such as the presence of Galα1-4Gal, are similar to those from pigeon. However, one of the significant differences is that the sialylations of the GEW-glycoproteins were found to be mostly α2,3-linkages, while α2,6-linkages of sialylation were the major in those from pigeon.

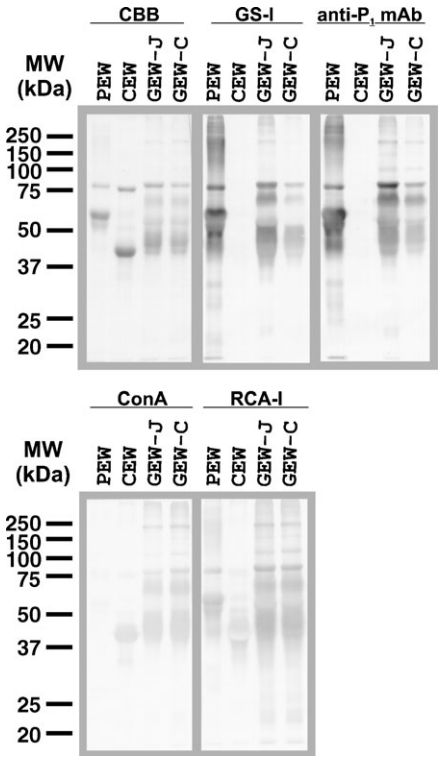

Fig. 1. Comparison of egg white glycoproteins from pigeon, chicken, and two species of gull by lectin blotting and immunoblotting. Egg white glycoproteins from pigeon (PEW, 2.5 μg/lane), chicken (CEW, 2.5 μg/lane), and two species of gull (GEW-J and GEW-C, 5 μg/lane each) were heat-denatured with the sample buffer containing 3% SDS and 5% 2-mercaptoethanol, and separated by SDS–PAGE (12.5%). After electrophoresis, the proteins were transferred to PVDF membranes and stained with Coomassie Brilliant Blue R-250 (CBB), GS-I lectin, anti-P_1 mAb, Con A or RCA-I.

Results

Detection of α-galactoside on GEW-glycoproteins by GS-I lectin and anti-P_1 mAb

Egg white glycoproteins from pigeon (*Columba livia*), chicken (*Gallus gallus*), and two different species of gulls were subjected to SDS–PAGE, transferred onto a PVDF membrane, and visualized by staining with Coomassie Brilliant Blue (CBB) (Figure 1, CBB staining). While two species of gulls (obtained from Japan (GEW-J) and Croatia (GEW-C)) exhibited almost identical SDS–PAGE profiles, their profiles did not resemble those of pigeon or chicken. The size and content of avian egg white proteins presumably differentiated in a species-specific manner during the diversification of birds. Indeed, by performing an SDS–PAGE analysis of 181 species of bird, we previously found that the patterns of egg white proteins varied among species (Suzuki, Laskowski, et al. 2004). The staining patterns of lectins and antibodies raised against GEW-glycoproteins were also compared with those of pigeon and chicken egg white glycoproteins. GEW-glycoproteins from both the two species of gull stained with GS-I lectin (specific for terminal α-Gal/GalNAc) and anti-P_1 mAb (specific for P_1 antigen, Galα1-4Galβ1-4GlcNAc-) (Figure 1), suggesting that they contain the Galα1-4Gal sequence on the glycoproteins as found in PEW. Chicken egg white (CEW)-glycoproteins were not stained with GS-I/anti-P_1 mAb because they lacked Galα1-4Gal.

Concanavalin A (Con A)-staining revealed a significant difference between PEW and GEWs. Whereas the glycoproteins from both GEWs and CEW were stained clearly with Con A, PEW-glycoproteins were stained only faintly. Con A can bind to CEW-glycoproteins since they contain high-mannose-type and hybrid-type oligosaccharides (Tai et al. 1975; Yamashita et al. 1983). In contrast, pigeon ovalbumins, ovotransferrin, and ovomucoid contain predominantly tri-, tetra-, and penta-antennary complex-type *N*-glycans (Suzuki et al. 2001; Takahashi et al. 2001), which are known to be undetectable by Con A. Thus, oligosaccharides of GEW-glycoproteins possess different features from those of PEW-glycoproteins, although both of them contain Galα1-4Gal termini. Glycoproteins of both GEW-J and GEW-C, as well as those of PEW and CEW, were stained with RCA-I (that detects terminal β-Gal, Figure 1), but not with PNA (that detects Galβ1-3GalNAc) and anti-(Galα1-3Gal) mAb (data not shown). The SDS–PAGE profile and lectin/antibody-staining pattern of GEW-J and GEW-C as shown in Figure 1 were almost the same as those from another species of gull, the herring gull (*Larus argentatus*), which was utilized in previous experiments (Suzuki, Laskowski, et al. 2004). Thus, clear Con A staining was consistent among three species of gull.

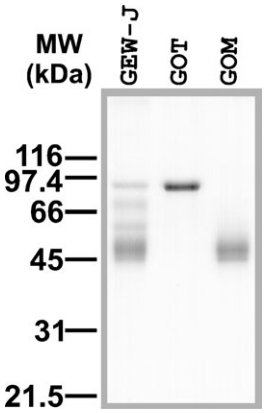

Fig. 2. SDS–PAGE analysis of gull egg white (GEW)-glycoproteins and isolated gull ovotransferrin (GOT) and ovomucoid (GOM). GOT and GOM were isolated from GEW-J, as described in *Material and methods*, and analyzed by SDS–PAGE.

Table I. Monosaccharide compositions of GEW-glycoproteins

	Content[a]		
Monosaccharide	GOM	GOT	GEW
GlcN	381 (6.42)	98.0 (6.53)	245 (5.57)
Man	178 (3)	45.0 (3)	132 (3)
Gal	204 (3.44)	40.1 (2.67)	145 (3.30)
α-Gal[b]	12.8 (0.216)	3.02 (0.201)	10.7 (0.243)
[% of total Gal]	[6.27]	[7.53]	[7.38]
NeuAc[b]	66.3 (1.12)	2.30 (0.153)	47.0 (1.07)

Amounts of protein were measured by the BCA assay, and the compositions of monosaccharides were analyzed by HPAEC. GlcN, D-glucosamine; Man, D-mannose; Gal, D-galactose; NeuAc, *N*-acetylneuraminic acid; GOM, gull ovomucoid; GOT, gull ovotransferrin; GEW, gull egg white.
[a]The results are expressed as nanomoles of sugar/mg of protein. The values in parentheses are based on setting the number of Man residues as three.
[b]Terminal α-galactosides and sialic acids were released from glycoproteins with green coffee bean α-galactosidase and neuraminidase from *A. ureafaciens*, respectively, and the released monosaccharides were measured with HPAEC.

Isolation of GEW-glycoproteins

To analyze oligosaccharide structures of GEW-glycoproteins, we first isolated two of the major egg white glycoproteins, namely ovotransferrin and ovomucoid, from GEW-J. Ovotransferrin is characterized by its metal-binding activity and can easily be isolated from other major proteins in egg white by a metal-chelate affinity column (Al-Mashikhi and Nakai 1987). We utilized Ni^{2+}-iminodiacetic acid (IDA)-agarose, which gave sufficient yields of ovotransferrin. High molecular mass proteins (>100 kDa) in the Ni^{2+}-IDA-agarose-bound fraction were removed with a DEAE-Sepharose column, and the purified gull ovotransferrin (GOT, molecular mass: 80 kDa) was obtained (Figure 2). Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF)-MS analysis revealed that the $[M+H]^+$ value of isolated GOT was 79,679, which is comparable with that of pigeon ovotransferrin (78,707) (Suzuki et al. 2001).

Ovomucoid is known to be a major serine-protease inhibitor in egg white and is often detected by its trypsin-inhibitory activity. To isolate gull ovomucoid (GOM), the Ni^{2+}-IDA-agarose-unbound fraction was further separated with a Butyl-Toyopearl 650 M column, based on hydrophobic interaction chromatography, and eluted using a linearly decreasing concentration of NaCl. The activity of GOM was detected as described in *Material and methods*, and the fractions with strong trypsin-inhibitory activity were collected. The GOM fraction produced broad bands on an SDS–PAGE gel (Figure 2) and revealed that 45 kDa and 47 kDa are the molecular masses that are close to that of GOM isolated by trichloroacetic acid-precipitation methods (Kato et al. 1987; Laskowski et al. 1987) (data not shown). MALDI-TOF-MS analysis of the isolated GOM indicated two broad peaks with the $[M+H]^+$ values of 45,278 and 47,212. As is the case in chicken and pigeon ovomucoids, GOM probably has a variable number of glycosylation sites, or a wide range of glycan sizes, which results in a broad range of molecular sizes upon MALDI-TOF-MS analysis.

Monosaccharide composition analysis of GEW-glycoproteins

The monosaccharide composition of GOM and GOT as well as of whole GEW-glycoproteins is shown in Table I. Both GOM and GOT contain GlcNAc (which is detected as GlcN after hydrolysis), Man, Gal, and NeuAc, but not Fuc and GalNAc. Whole GEW-glycoproteins had a similar ratio of monosaccharides as that of GOM, but they also contained GalNAc (6.16 nmol/mg of protein), which is presumably derived from *O*-glycans from certain glycoproteins, such as ovomucin. Whereas GOM and GOT had similar ratios of GlcN:Man:Gal, the relative NeuAc content in GOT was about 7-fold lower than in GOM. α-Galactosidase released 6–8% of the total Gal from GOM, GOT, and GEW, providing evidence that GEW-glycoproteins contain α-galactosides. However, the relative amounts of α-Gal in the total Gal (GOM, 6.27%; GOT, 7.53%) were significantly lower than those of PEW-glycoproteins (POM, 18.7%; POT, 37.5%) (Suzuki et al. 2001).

Fractionation and analysis of PA-derivatized N-glycans from GEW-glycoproteins

The *N*-glycans from GOM and GOT, as well as from whole GEW-glycoproteins, were derivatized with 2-aminopyridine (PA), fractionated first into neutral, monosialylated, and disialylated fractions with a DEAE column (Figure 3A), and then further fractionated on an ODS column (Figure 3B). The molar ratios of GEW *N*-glycans calculated from the peak areas of PA-oligosaccharides were neutral 53%, monosialylated 33%, and disialylated 14%, and those of GOM were almost the same. The elution profile of GEW and GOM on both DEAE and ODS columns was very similar, probably because the major sources of *N*-glycans in GEW-glycoproteins are ovomucoid and ovalbumin, which are generally rich in the *N*-glycans of avian egg whites. In contrast, *N*-glycans of GOT were mostly neutral, and only a trace amount of sialylated *N*-glycans was detected (Figure 3A). Furthermore, the neutral *N*-glycans of GOT revealed an elution profile on an ODS column different from those of GEW and GOM (Figure 3B). When neutral PA-glycans from GOT, GOM, and GEW were analyzed on the ODS column, only small peaks were detected at elution times of around 4–12 min (Figure 3B), suggesting that high-mannose-type *N*-glycans are not major components of GEW *N*-glycans (Tomiya et al. 1988). To obtain more precise maps of their respective major *N*-glycans, the PA-derivatized neutral, monosialylated, and

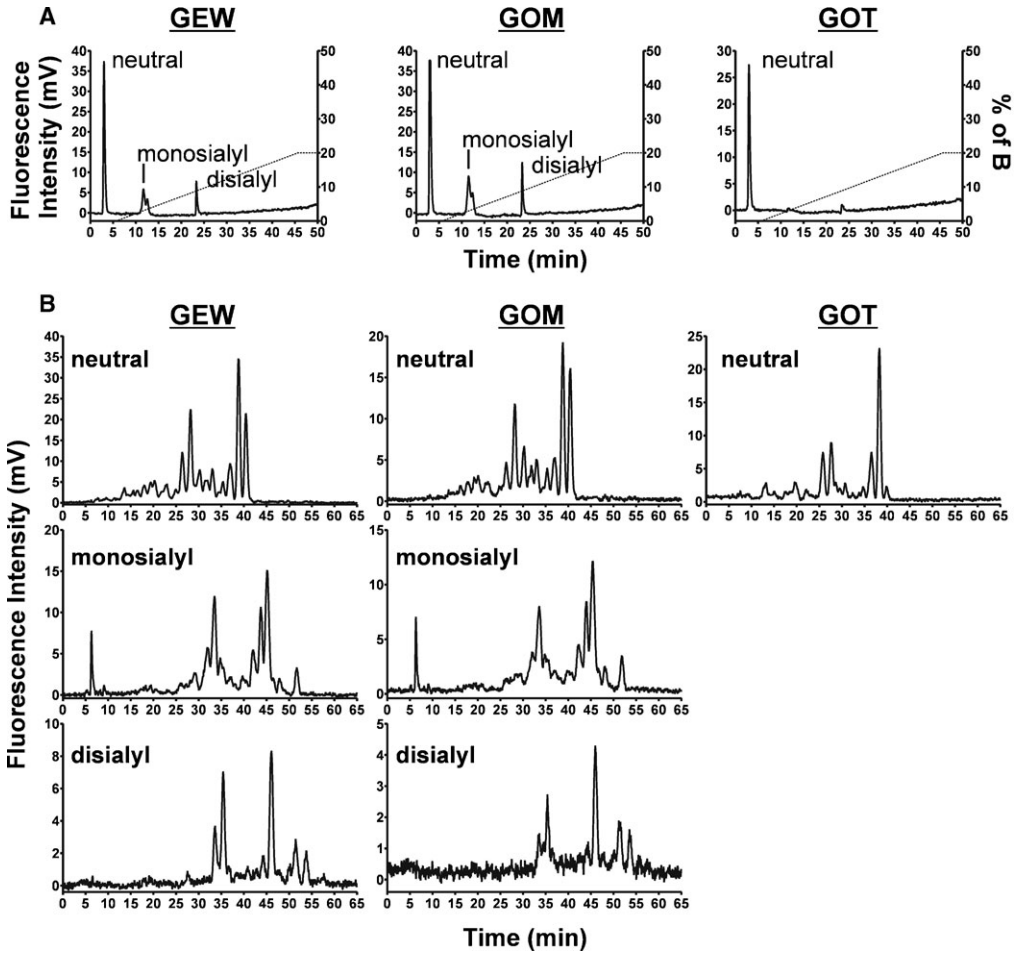

Fig. 3. Separation of PA-oligosaccharides from GEW, GOM, and GOT by HPLC. (**A**) Total PA-oligosaccharides from GEW, GOM, and GOT were separated into neutral, monosialyl, and disialyl oligosaccharides on the DEAE column. (**B**) Elution profiles of the neutral, monosialyl, and disialyl PA-oligosaccharides on the ODS column.

disialylated samples were subjected to MALDI-MS profiling and selective MS/MS analyses after permethylation.

MS and MS/MS analyses of PA-derivatized N-glycans from GEW, GOM, and GOT

General Aspects. The MALDI-MS profiles of permethylated neutral PA-*N*-glycans from GEW, GOM, and GOT are shown in Figure 4. Relatively large size of neutral *N*-glycans were not detected for GOT (Figure 4C), whereas there was additional heterogeneity in GEW (Figure 4A) and GOM (Figure 4B), due to the presence of one and two additional Hex-HexNAc units (449 mass units), with variable numbers of additional Hex residues (204 mass units per Hex residue). Otherwise, the MS profiles afforded by the three samples were fairly similar. All the major peaks detected in the spectrum of neutral *N*-glycans from GEW, GOM, and GOT were subjected to MALDI MS/MS analysis.

To determine the *N*-glycan structures, (i) low-energy collision-induced dissociation (CID) MS/MS on a MALDI Q/TOF and (ii) high-energy CID MS/MS on a MALDI TOF/TOF analyses were performed as described previously (Yu et al. 2006). Representative examples of MALDI MS/MS spectra and the nomenclature of fragment ions are shown in Figure 5. As described in previous reports (Yu et al. 2006, 2008), the predominant fragment ions produced at low-energy CID are the B and Y ion pairs directed at HexNAc (e.g., the profile in Figure 5C), and they are most informative in mapping the core structures and the complement of nonreducing terminal epitopes. In contrast, the high-energy CID affords, in addition to B and Y ions, many other cleavage ions, including the ring cleavage ions (e.g., ^{3,5}A and ^{1,5}X ions in Figure 5A, B, and D) and other satellite ions resulting from concerted elimination of substituents around the pyranose ring (e.g., D, E, G, and H ions in Figure 5A, B, and D), which often provide linkage information. Based on both compositions and MS/MS data, the major structures could be deduced, as annotated in Figure 4A. In full agreement with the results, described above, of HPLC analysis on an ODS column, high-mannose-type structures were not detected among the major neutral *N*-glycans. Instead, the major neutral *N*-glycans were identified as hybrid and complex-types.

Importantly, the concerted low and high-energy CID MS/MS analyses confirmed the presence of a terminal Galα1-4Galβ1-4GlcNAc epitope, as inferred from the antibody staining (Figure 1). By low-energy CID MS/MS, the expected neutral loss of the terminal $Hex_2HexNAc$ moiety via glycosidic cleavage would give rise to specific Y ions, along with a prominent

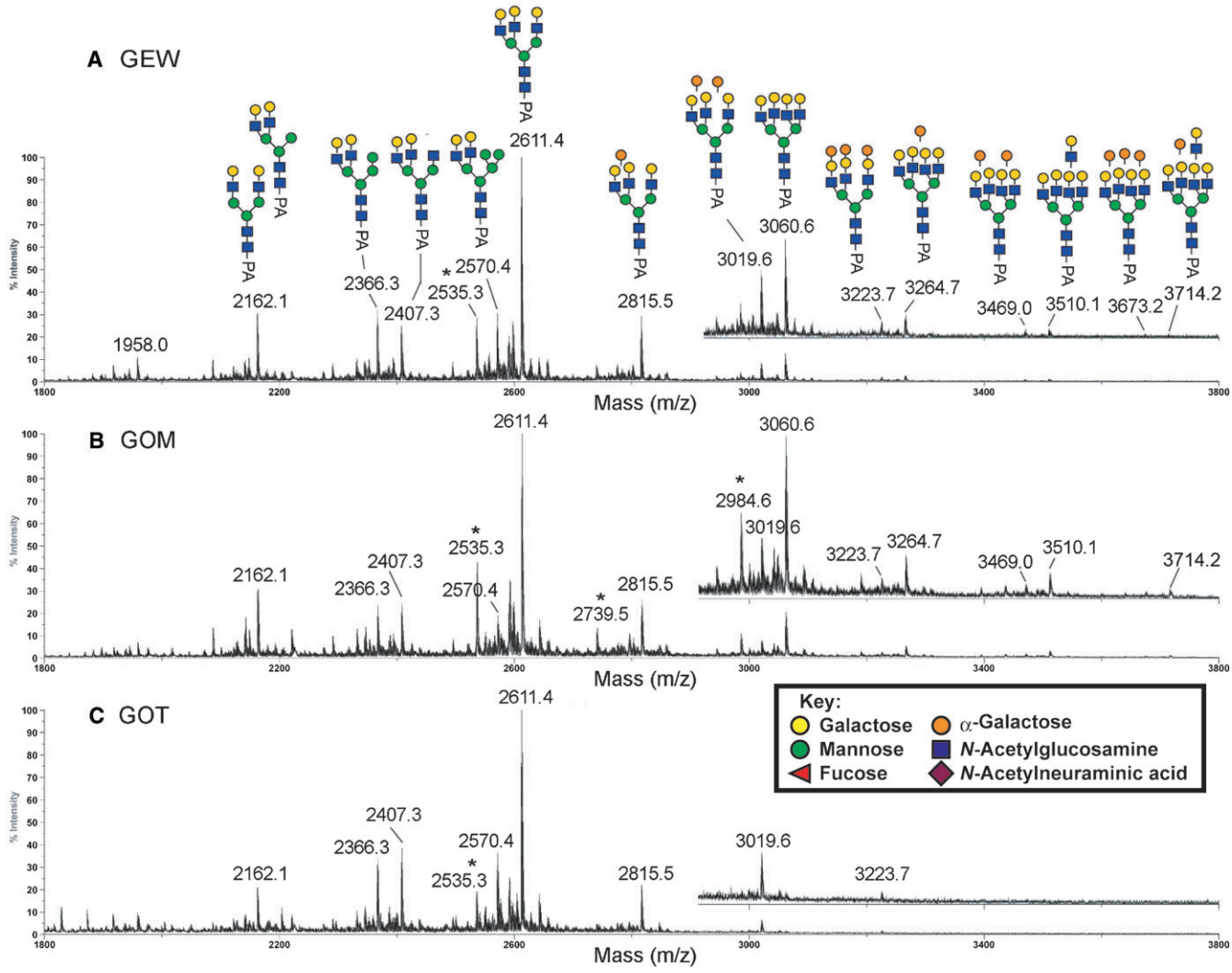

Fig. 4. Comparative MALDI-MS mapping of the neutral PA-glycans derived from GEW (**A**), GOM (**B**), and GOT (**C**). All major molecular ion signals afforded by the permethylated samples were monosodiated and the assigned structures, as deduced from both composition and MS/MS data, were annotated along with their monoisotopic *m/z* values. The mass regions from *m/z* 3000 to *m/z* 3800 were additionally magnified and presented as offsets above the main spectral traces. Satellite peaks at 76 mass units below that of major signals, most notably *m/z* 2535, are marked with an asterisk and correspond to non-PA tagged, but reduced and similarly permethylated glycans. Symbols used conform to the recommendations of the Consortium for Functional Glycomics, but the orange spheres represent α-Gal.

sodiated B ion for $Hex_2HexNAc$ at *m/z* 690 (Suzuki et al. 2003). By high-energy CID MS/MS, as shown in Figure 5A, the neutral loss would preferably occur via cross-ring cleavage to yield an $^{1,5}X$ ion, which is further supported by the satellite G and H ions, and corroborated by the same sodiated B ion at *m/z* 690, which is often also accompanied by the sodiated E ion at *m/z* 619 (Yu et al. 2006). Moreover, we confirmed that all peaks established as carrying α-Gal disappeared upon α-galactosidase digestion. This systematic analysis showed that the degree of α-Gal capping on both neutral and sialylated GEW *N*-glycans is not as extensive as that on the PEW *N*-glycans (Suzuki et al. 2001; Takahashi et al. 2001).

Neutral Oligosaccharides. MALDI-MS profiles indicated a dominant peak at *m/z* 2611 (Figure 4), and up to three additional Hex units were added to this structure to produce signals at *m/z* 2815, 3019, and 3223. MALDI MS/MS analysis revealed that these four peaks are the sodiated molecular ion signals of complex-type glycans with nonfucosylated, nonbisected, triantennary structures having 0–3 Hex-Hex-HexNAc termini, based on characteristic fragment ions. For example, the high-energy CID MS/MS data suggested that the peak at *m/z* 2815 is the triantennary structure carrying a single Hex-Hex-HexNAc terminus. This peak comprised a mixture of isomers, with antennae that can be α-galactosylated without any apparent site preference (Figure 5A). The D ions at the β-Man (*m/z* 1084 and 880) suggest the presence of two major isomeric forms, as illustrated in Figure 5A (the left and right structures of PA-*N*-glycans), with the single LacNAc antenna on the 6-arm being capped and not capped by α-Gal, respectively. The set of D, $^{1,5}X_3$, and C″/Y ions of each isomeric form, i.e., *m/z* 1084, 1726, and 692 of the left structure and *m/z* 880, 1522, and 896 of the *right* structure, respectively, define the triantennary structures unambiguously. The characteristic D ions detected at *m/z* 1084 and 880 were formed at the 3,6-linked β-Man and were not accompanied by a fragment ion at a position 321 mass units

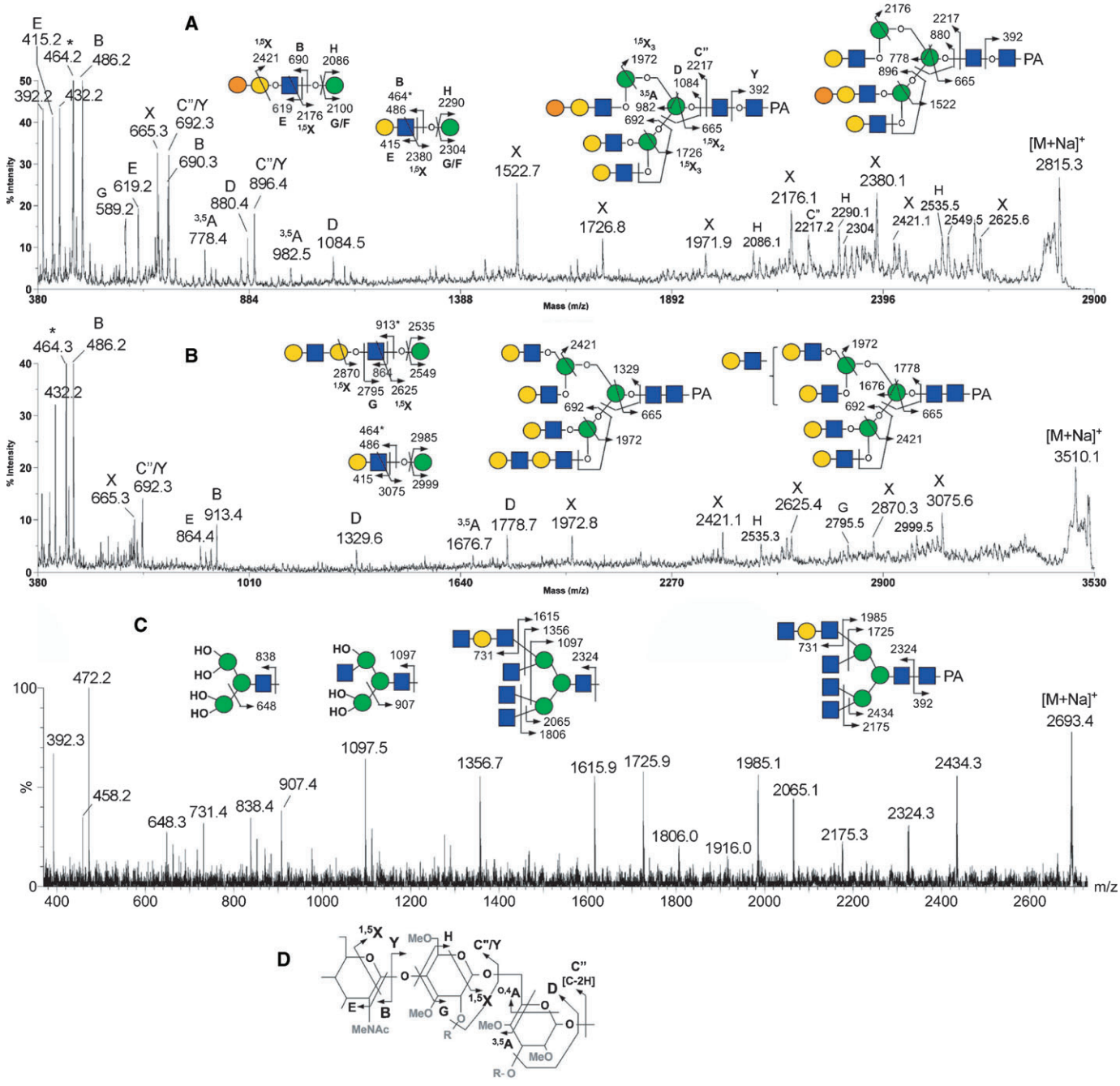

Fig. 5. Representative MALDI MS/MS spectra of permethylated neutral PA-glycans from GEW carrying a Gal-Gal-GlcNAc antenna (**A**), or a diLacNAc antenna before (**B**) and after (**C**) β-galactosidase digestion. Assignments of the high-energy CID MS/MS fragment ions afforded by analysis on MALDI-TOF/TOF (**A**, **B**) or low-energy CID MS/MS on MALDI Q/TOF (**C**) are schematically illustrated. In both (**A**) and (**B**), alternative isomeric forms that can be inferred from the presence of a few characteristic fragment ions (D and ^{3,5}A ion pairs formed at the β-Man and the $^{1,5}X_3$ ions) are given, along with their common terminal structures. Ion types for each of the major fragment ions detected are further labeled according to the expanded nomenclature adapted and described in Yu et al. (2006) and further illustrated in the schematic drawing shown in (**D**). For the nonreducing terminal Hex-HexNAc epitope, both oxonium-type (*m/z* 464) and sodiated B (*m/z* 486) ions were commonly detected, and the former is marked with an asterisk. For simplicity, the $^{1,5}X_3$ ions were labeled as X ions. In (**C**), the multiple glycosidic cleavages afforded by low-energy CID MS/MS (B and Y ions) do not allow for distinguishing the location of the diLacNAc antennae. The left three partial structures illustrate further losses of glycosyl residues from the B ions formed by glycosidic cleavage at the chitobiose core. OH represents exposed free hydroxyl groups at cleavage sites.

lower, suggesting the absence of bisecting GlcNAc (Chen et al. 2008). A structure with two LacNAc antennae on the 6-arm is not likely to be the major isomeric form, since D ions at *m/z* 1329 or 1533 were not detected. The predominance of such a triantennary structure with two antennae on the 3-arm is also supported by the MS/MS analysis of the non-α-Gal capped triantennary structure at *m/z* 2611 that shares the common biantennary 3-arm (data not shown, but see annotation on Figure 4A). The presence of an $^{1,5}X_2$ ion at the β-Man (*m/z* 665) and the Y_1 ion derived from cleavage between the core chitobiose (*m/z* 392) without

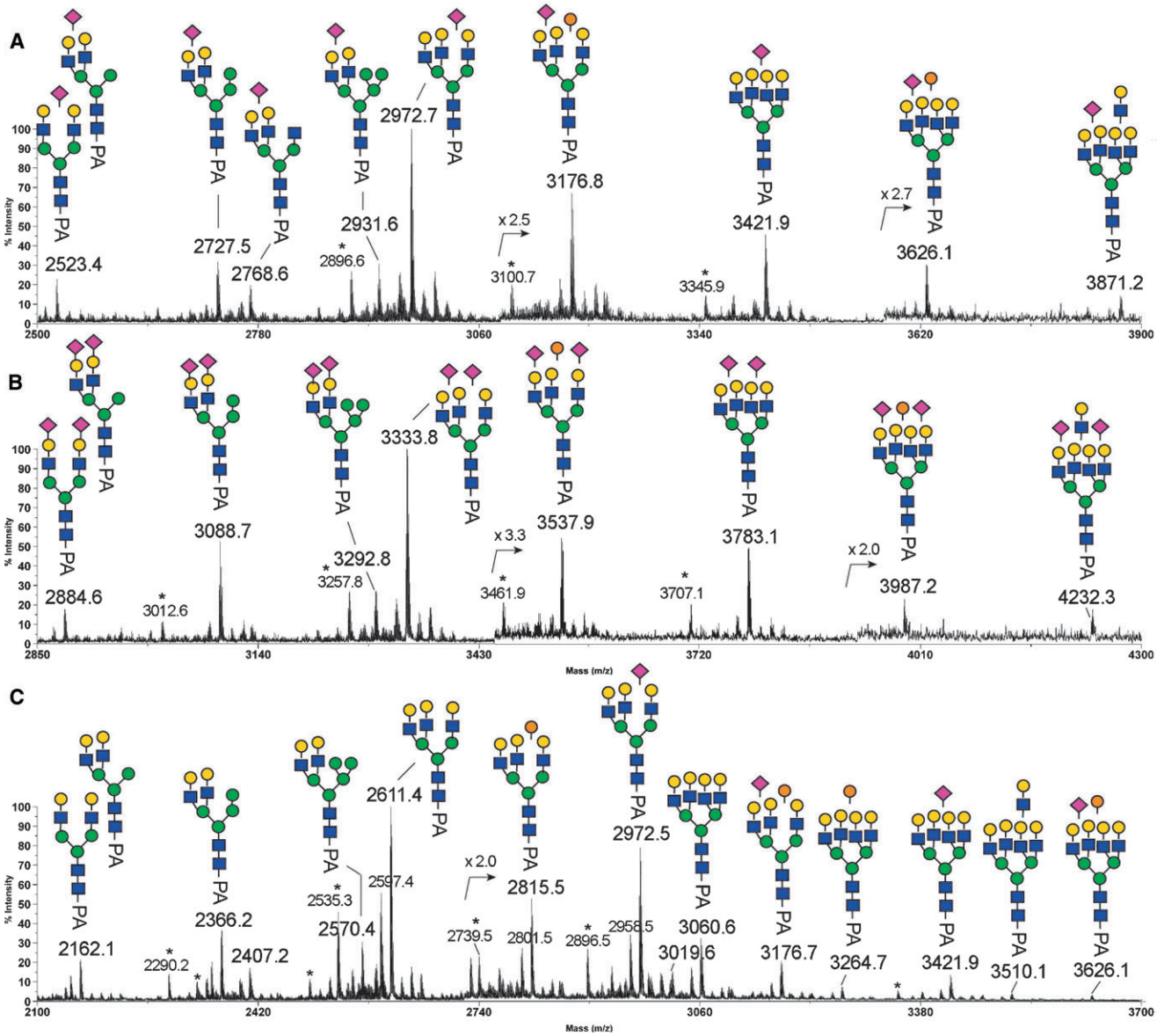

Fig. 6. MALDI-MS mapping of the monosialylated (**A**), disialylated (**B**), and α2,3-desialylated (**C**) PA-glycans derived from GEW. All major molecular ion signals afforded by the permethylated samples are monosodiated and the assigned structures, as deduced from both composition and MS/MS data, are annotated along with their monoisotopic *m/z* values. Satellite peaks (marked with an asterisk) at 76 mass units below that of major signals, correspond to non-PA tagged but reduced and similarly permethylated glycans. Symbols used conform to the recommendations of the Consortium for Functional Glycomics, but with orange spheres representing α-Gal, as shown in Figure 4.

an additional dHex unit (174 mass units) indicates that the core structures are nonfucosylated.

In neutral GEW and GOM *N*-glycans, the tetraantennary structures with 0–3 additional Hex residues could be detected at *m/z* 3060, 3264, 3469 and 3673, but a fully α-Gal-capped (+4 Hex) structure was apparently absent (Figure 4A and B). Instead, another LacNAc unit could be added to produce peaks at *m/z* 3510 and 3714. MS/MS analyses clearly showed that these are tetraantennary, and not pentaantennary, structures. The presence of an extended diLacNAc terminal epitope was firmly established by the ^{1,5}X ion at *m/z* 2625 and the nonsodiated, oxonium-type B ion at *m/z* 913, accompanied by the E ion at *m/z* 864 (Figure 5B). Further confirmation was obtained by β-galactosidase digestion, which succeeded in removing only four, and not five, terminal β-Gal residues from the parent at *m/z* 3510, to produce the product at *m/z* 2693, the structure of which was then identified by MS/MS analysis (Figure 5C).

The MS/MS data also indicated the presence of a series of hybrid-type *N*-glycans with two LacNAc antennae on the 3-arm and additional Hex residues, likely to be Man, on the 6-arm, as annotated in Figure 4A (*m/z* 2162, 2366, and 2570). For example, the ^{1,5}X ion at *m/z* 1972, which arises from the loss of three Hex residues via a single cleavage, and the ^{3,5}A ion at *m/z* 737 and the D ion at *m/z* 839 at the β-Man (Yu et al. 2008), were the characteristic fragment ions of hybrid-type *N*-glycans detected from the peak at *m/z* 2570 (data not shown). In fact, when we isolated Con A-bound neutral PA-*N*-glycans from GEW using Con A affinity chromatography, only those

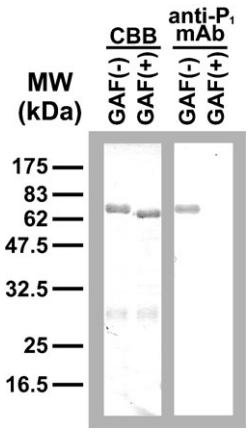

Fig. 7. Enzymatic de-*N*-glycosylation of gull egg yolk (GEY) IgG. The isolated glycoprotein was heat-denatured with 0.4% SDS and 100 mM 2-mercaptoethanol, and incubated with (+) or without (−) glycoamidease F (GAF) at 37°C overnight in the presence of 1% NP-40. Each sample (1 μg/lane) was separated by SDS–PAGE (12.5%), transferred to PVDF membranes, and stained with Coomassie Brilliant Blue R-250 (CBB, left) or anti-P_1 mAb (right).

assigned as hybrid types were detected by MALDI-MS analysis as major peaks (data not shown). This observation is consistent with the positive staining with Con A (Figure 1) despite the absence of high mannose structures. In accordance with the assignment supported by MS/MS analyses, all peaks established as carrying hybrid structures remained visibly unaffected upon α-galactosidase digestion (data not shown).

Sialic Acid Attachment. Having demonstrated that the underlying structures for GOM and GOT *N*-glycans broadly resemble those of the total GEW, further MS analyses of the sialylated structures were performed, mostly on the monosialylated and disialylated GEW *N*-glycans (Figure 6). Collectively, the MS and MS/MS data indicated that Neu5Ac-sialylation was found only on non-α-Gal-capped LacNAc termini because B ion fragments for NeuAc-Hex-HexNAc at *m/z* 847 but not for NeuAc-Hex-Hex-HexNAc at *m/z* 1051 were detected. The Neu5Ac-sialylation was evenly distributed among the hybrid, bi-, tri-, and tetraantennary structures, as annotated in Figure 6. Upon α2,3-neuraminidase digestion, the majority of the terminal Neu5Ac residues were removed, yielding an MS profile (Figure 6C) very similar to that of the neutral GEW *N*-glycans (Figure 4A), except for the additional presence of a few peaks corresponding to incompletely digested and/or α2,3-neuraminidase-resistant sialylated components at *m/z* 2972, 3176, 3421, and 3626. Further MALDI-TOF/TOF MS/MS analyses of these peaks yielded the characteristic D ion at *m/z* 588, corresponding to the sodiated fragment ion of Neu5Ac-Gal-OH having further eliminated substituent at C3 of the Gal and thus indicative of a 2,6-Neu5Ac linkage but not 2,3-Neu5Ac (Yu et al. 2006, 2008). Although not quantitative, the neuraminidase digestion and MS/MS data suggest that the GEW *N*-glycans are mostly α2,3-sialylated, but that a significant amount of α2,6-sialylation is also common.

Isolation of gull egg yolk IgG

Avian egg yolk is a rich source of IgG, which arises from the transport of maternal antibodies by the receptor-specific process (Patterson et al. 1962). Avian IgG either in serum or egg yolk is also called IgY because of its distinct structural difference from mammalian IgG (Leslie and Benedict 1970; Warr et al. 1995). GEY IgG was isolated, as described in *Material and methods*, and analyzed with SDS–PAGE and Western blotting. As shown in Figure 7, the isolated gull IgG consisted of heavy (H) chains (67 kDa) and light (L) chains (27 kDa), which were of typical sizes for these components in avian IgG (Leslie and Benedict 1970; Warr et al. 1995; Suzuki et al. 2003). MALDI-TOF-MS analysis revealed that the $[M+H]^+$ value of the isolated gull IgG was 170,345, which is slightly higher than that of chicken IgG (Suzuki and Lee 2004). The H chain of gull IgG was stained with anti-P_1 mAb, suggesting the presence of Galα1-4Galβ1-4GlcNAc- sequences (Figure 7). A decrease in molecular size after glycoamidase F (GAF) digestion was apparent in the H chains. The de-*N*-glycosylated gull IgG-H chains were no longer stained with anti-P_1 mAb. This indicates that oligosaccharides with the Galα1-4Galβ1-4GlcNAc- sequence on the H chain were completely removed by GAF.

MS and MS/MS analysis of N-glycans from gull egg yolk IgG

Without further PA-derivatization and HPLC fractionation, the GEY IgG *N*-glycans were directly permethylated for MS and MS/MS analyses. As expected, the resulting MALDI-MS profile (Figure 8A) is more complicated than those of GEW neutral and sialylated *N*-glycans. Nonetheless, using the same MALDI-MS/MS sequencing technique, all the major peaks could readily be assigned and annotated accordingly. However, it was apparent that among the less abundant components, delineation of the MS/MS spectra afforded by isomeric and isobaric parents did not allow the unambiguous identification of the presence of terminal Gal-Gal-GlcNAc epitope. Thus, an independent total ion mapping (TIM) analysis based on offline nanoESI-MS^2 (Aoki et al. 2007) was performed. In brief, successive automated MS^2 scans were acquired using the linear ion trap across the entire mass range (*m/z* 500–2000) at steps of 2 mass units, with an overlapping precursor ion isolation window. A pseudo-precursor ion scan can then be reconstructed by plotting the occurrence (signal intensity) in the MS^2 spectra of a particular product ion, in this case *m/z* 690 for Hex-Hex-HexNAc (Figure 8C). The accurate *m/z* values of parent ions that carry this terminal epitope can then be inferred from the corresponding high resolution, high accuracy nanoESI-MS profile acquired independently with the Orbitrap either before and/or after the TIM experiments (Figure 8B), and false positives can be discriminated by manual inspection of the actual MS^2 data for the candidate precursors (data not shown). Using this method, all peaks assigned as containing the Gal-Gal-GlcNAc epitope by MALDI-MS/MS, as annotated in Figure 8A, can be identified and confirmed by TIM (Figure 8C) at the MS^2 level. Moreover, we confirmed that the peaks that were established as carrying α-Gal were removed upon α-galactosidase digestion (data not shown). In addition, the peak at *m/z* 2489 assigned by MALDI-MS/MS as $[M+Na]^+$ of a biantennary, bisected, core fucosylated structure (Figure 8A) was shown by TIM to additionally include a relatively minor structural isomer (*m/z* 1256.6, $[M+2Na]^{2+}$, Figure 8C), which also carries the Gal-Gal-GlcNAc epitope.

Taken together, it can be concluded that, like the GEW *N*-glycans, the GEY IgG *N*-glycans also comprise a population of complex-type *N*-glycans that carry the Galα1-4Galβ1-4GlcNAc

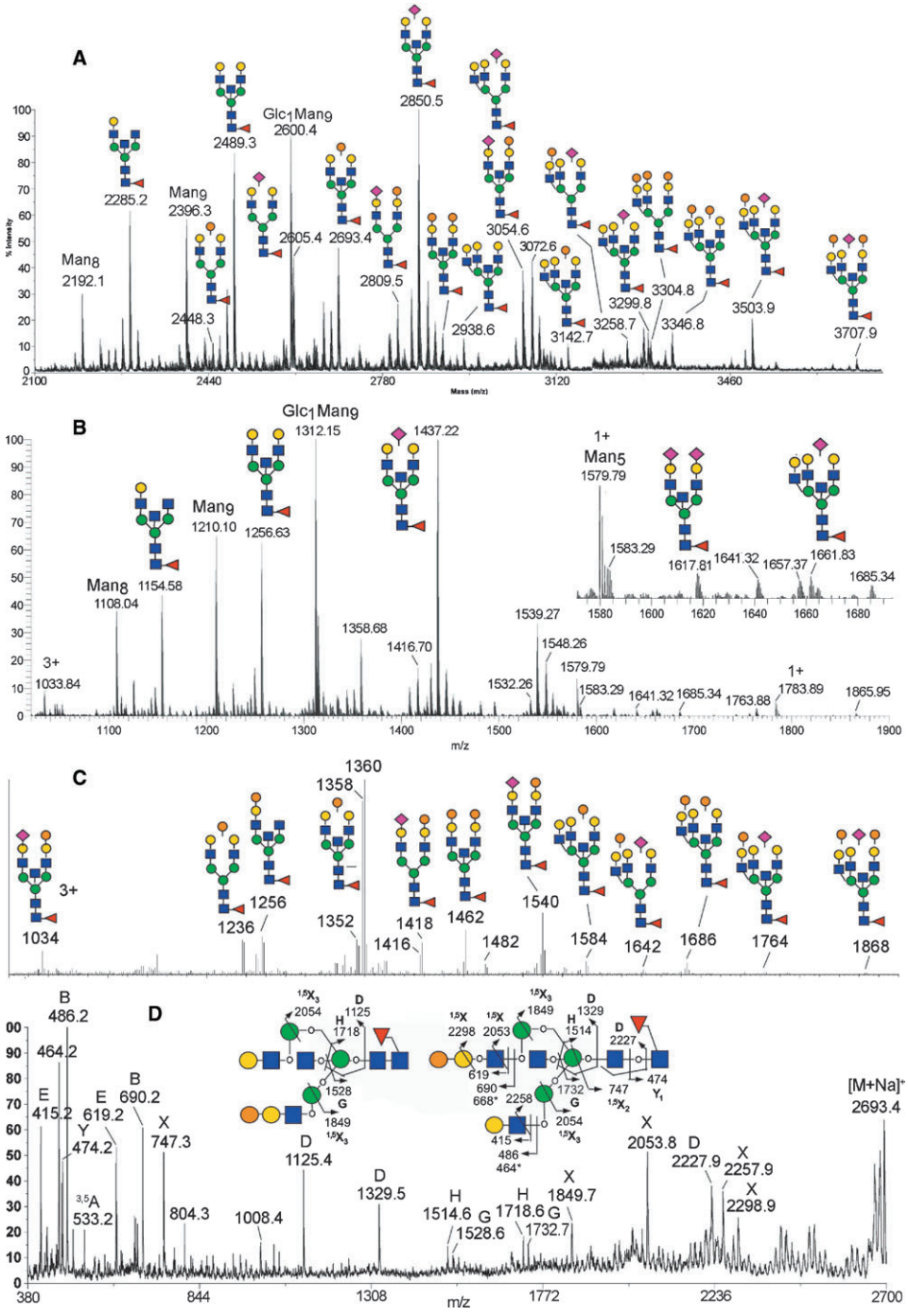

Fig. 8. MS and MS/MS analyses of GEY IgG *N*-glycans. MALDI-MS (**A**) and nanoESI-MS (**B**) profiles of permethylated total GEY IgG *N*-glycans. All major molecular ion signals afforded by MALDI-MS analysis of the permethylated samples in (**A**) are monosodiated, and the assigned structures as deduced from both composition and MS/MS data are annotated (only the major isomeric forms) along with their monoisotopic *m/z* values. The molecular ions afforded by nanoESI-MS in (**B**) are either mono-, di-, or trisodiated for singly, doubly, and triply charged species, respectively. Unless otherwise annotated, the majorities of ions are doubly charged. The high resolution and accuracy measurement using the Orbitrap MS allows unambiguous charge and composition assignments, which can readily be matched with all the major components detected by MALDI-MS. For simplicity, only those peaks assigned as being non-α-Gal-containing components were annotated, whereas those carrying α-Gal were annotated in the corresponding total ion mapping (TIM) profile (**C**). The TIM profile in (**C**) was generated by plotting the occurrence of a singly charged product ion, *m/z* 690 (sodiated B ion of Hex-Hex-HexNAc^{+}), during nanoESI-MS^{2} acquisition at steps of 2 mass units across the entire mass range, using the linear ion trap. Labeled peaks correspond to the *m/z* values of the selected precursor ion step that yielded the targeted fragment ion signals. Structures were assigned based on further examination of the collected MS^{2} spectra, and by referring to the high resolution Orbitrap MS profile in (**B**). A representative MALDI-TOF/TOF MS/MS spectrum of a permethylated GEY IgG *N*-glycan (*m/z* 2693) containing core fucosylation, bisecting GlcNAc and the Gal-Gal-GlcNAc epitope, is shown in (**D**) along with a schematic illustration of the fragment ion assignments. The symbols and nomenclature used are as described in the legend to Figure 5. The occurrence of bisecting GlcNAc was identified by the respective D ions (*m/z* 1125, 1329) formed at the 3,4,6-linked β-Man, which were accompanied by a characteristic fragment ion 321 mass units beneath them (804 and 1008, respectively), as well as the G/H ion pairs (*m/z* 1528/1718 and 1732/1514, respectively), as previously noted (Chen et al. 2008). These sets of fragment ions also indicated the presence of two structural isomers as illustrated, with the Gal-Gal-GlcNAc epitope located on either the 3- or 6-arm.

epitope. This α-Gal capping is similarly noncomprehensive, with many of the LacNAc termini remaining noncapped or Neu5Ac-sialylated. Either by manual MALDI-MS/MS or automated ESI-MS^2 in TIM mode, other terminal epitopes, such as Neu5Ac-Gal-Gal-GlcNAc (*m/z* 1051) or Gal-Gal-Gal-GlcNAc (*m/z* 894), were not detected. Apart from these similarities, there are, however, significant differences in structural features between the GEW and GEY IgG *N*-glycans. First, the sialylated GEY IgG *N*-glycans were mostly resistant to α2,3-neuraminidase digestion. Together with the aforementioned diagnostic D ion at *m/z* 588 afforded by high-energy CID MALDI-MS/MS analysis of these sialylated structures (data not shown), it can be inferred that the GEY IgG *N*-glycans are mostly α2,6-sialylated, although the presence of a small degree of coexisting α2,3-sialylation cannot be ruled out. Second, the GEY IgG *N*-glycans comprise a significant amount of high mannose structures, including a $Hex_{10}HexNAc_2$ structure, most likely corresponding to the $Glc_1Man_9GlcNAc_2$ structure commonly found in avian IgG (e.g., see (Suzuki et al. 2003; Suzuki and Lee 2004)), whereas no significant amounts of hybrid-type structures were detected. Third, the GEY IgG complex-type *N*-glycans are mainly limited to biantennary and triantennary sizes, and the majority of these complex-type structures are bisected and core fucosylated, as readily confirmed by MALDI MS/MS analysis (Figure 8D).

Discussion

Oligosaccharides are biosynthesized by the concerted actions of several kinds of processing enzymes and glycosyltransferases. Depending on the expression levels and activities of the individual enzymes in cells or tissues, various structures of oligosaccharides are present despite the common structures and/or overall topologies of core proteins. The structural analysis of oligosaccharides can provide more accurate information of the biosynthetic mechanism that produces the various structures than lectin/antibody staining. Moreover, comparing the oligosaccharide structures from different species will provide a clue to the mechanism that gives rise to species-specific glycans and further our understanding of their biological roles.

When avian egg white glycoproteins from 181 species were analyzed by lectin/antibody staining (Suzuki, Laskowski, et al. 2004), we realized that Con A staining differed from species to species, regardless of the presence of Galα1-4Gal. For example, among the egg white glycoproteins from 104 species that were stained with GS-I/anti-P_1 mAb, 75 species were also stained with Con A. In addition, GEW-glycoproteins from two species were positive for both Con A and GS-I/anti-P_1 mAb staining, as shown in this work, while PEW-glycoproteins were stained with GS-I/anti-P_1 mAb, but not with Con A. Therefore, GEW-glycoproteins were expected to possess different oligosaccharide structures from those of PEW, although both GEW and PEW express Galα1-4Gal. Con A is known to bind to high-mannose-type, hybrid-type, and some biantennary complex-type oligosaccharides (Goldstein and Poretz 1986). Because of its broad specificity, we could not identify the actual structures recognized by Con A without the structural analysis of *N*-glycans from GEW-glycoproteins.

We detected five differences between the structures of *N*-glycans from GEW-glycoproteins and those from PEW-glycoproteins. (i) Hybrid-type oligosaccharides were detected in GEW-glycoproteins, while they were not found in PEW-glycoproteins. It is most likely that the hybrid-type oligosaccharides on GEW-glycoproteins were recognized by Con A. (ii) Bisecting GlcNAc was not detected on GEW-glycoproteins, while approximately 40% of *N*-glycans from PEW-glycoproteins contain bisecting GlcNAc (Takahashi et al. 2001). (iii) Not all of the complex-type *N*-glycans from GEW-glycoproteins possess the Galα1-4Gal sequence on their branches, and LacNAc or sialylated LacNAc on nonreducing termini was prominent. In contrast, all of the major *N*-glycans from PEW-glycoproteins possess the Galα1-4Gal sequence on 1-4 branches. (iv) Some branches from GEW-*N*-glycans were occupied with diLacNAc (LacNAc-LacNAc), which was not detected in the major PEW-*N*-glycans. (v) α2,3-NeuAc, and not α2,6-NeuAc, was the major sialylation on GEW-glycoproteins, while α2,6-NeuAc was the prominent sialylation in PEW-glycoproteins.

Based on these observations, we deduced that there are several differences between gull and pigeon glycoproteins, which are due to the expression and activity of glycosyltransferases in the oviduct, where biosynthesis and glycosyl modification of the major egg white glycoproteins of birds are carried out. First, hybrid-type oligosaccharides are most likely generated by the limiting action of Golgi α-mannosidase II. Chicken egg white glycoproteins also contain hybrid-type oligosaccharides, but the significant difference is the absence of bisecting GlcNAc in GEW-glycoproteins. Although the insertion of a bisecting GlcNAc residue by the action of *N*-acetylglucosaminyltransferase (GlcNAcT)-III was assumed to contribute to the production of hybrid-type oligosaccharides by preventing the action of α-mannosidase II in chicken oviducts (Allen et al. 1984; Shao and Wold 1989), this may not be the case in gull oviducts. In contrast, hybrid-type oligosaccharides are not the major component of pigeon *N*-glycans, and about 40% of the *N*-glycans are complex-type with a bisecting GlcNAc. α-Mannosidase II is presumably sufficient in the pigeon oviduct, and it was not interfered with the action of GlcNAcT-III. Second, activity of α1,4-galactosyltransferase (α1,4-GalT) in the gull oviduct appears to be lower than that in the pigeon oviduct since the LacNAc branches were not fully α-galactosylated. Instead, di-LacNAc was detected in some branches, although the LacNAc repeat was not prominent. This result indicates that GlcNAcT, which mediates the production of diLacNAc, is present in the gull oviduct. Since the nonreducing terminal residues of *N*-glycans are richer in LacNAc-branches than in α-galactosylated or α-sialylated LacNAc-branches, the GlcNAcT for diLacNAc may have a better chance of accessing the acceptor substrates. Third, α2,3-sialyltransferase (α2,3-SiaT), rather than α2,6-SiaT, is mainly expressed to produce sialylated oligosaccharides of GEW. Only α2,6-linked sialic acid had been detected among the sialylated *N*-glycans containing Galα1-4Gal, which were from pigeon egg white glycoproteins and IgG. Therefore, this is the first time that *N*-glycans possessing both Galα1-4Gal and NeuAcα2-3Gal at different nonreducing termini are observed. In GEW-glycoproteins, α2,3-linked sialic acid is found on the branch that does not contain the Galα1-4Gal sequence. It seems that the α2,3-sialylation and α4-galactosylation occurred by the competitive actions of α2,3-SiaT and α4GalT in the gull oviduct. The *N*-glycan structures also imply that the sialylation linkages can differ, regardless of the presence of the Galα1-4Gal sequence. The tissue distribution of α2,3- or α2,6-sialylation is

reported to differ in a species-specific manner in the case of mammals (Gagneux et al. 2003), and similar phenomena may occur among avian species. This subject should be extensively investigated considering that sialylation is crucial for infection with influenza and other viruses (Webster et al. 1992; Suzuki 2005).

Gull IgG also contains Galα1-4Gal (Figures 7 and 8). This fact indicates that the production of (Galα1-4Gal)-containing glycoproteins is not limited to the gull oviduct, but also at least occurs in antibody-producing cells. Structural features of gull IgG were similar to those of pigeon IgG (Suzuki et al. 2003), except that gull IgG lacks the Galα1-4Galβ1-4Galβ1-4GlcNAc and Galβ1-4Galβ1-4GlcNAc sequences. The presence of the Galα1-4Galβ1-4Galβ1-4GlcNAc sequence is also reported in *O*-linked glycans of salivary gland mucin glycoproteins of Chinese swiftlet (genus *Collocalia*) (Wieruszeski et al. 1987). Therefore, some avian species have the ability to produce Galβ1-4Gal, but its distribution among avian species remains to be investigated. Unlike egg white glycoproteins, most sialylations of *N*-glycans from gull IgG were α2,6-linkages, as are those from pigeon and chicken (Suzuki et al. 2003; Suzuki and Lee 2004).

The presence of monoglucosylated high-mannose-type oligosaccharides is well conserved among pigeon, chicken and quail IgG (Ohta et al. 1991; Matsuura et al. 1993; Raju et al. 2000; Suzuki et al. 2003; Suzuki and Lee 2004), and in this study, we found that monoglucosylated high-mannose-type oligosaccharides are also present in gull egg yolk IgG. Unlike mammalian IgG, avian IgG contains one additional domain in the constant region of its heavy (H) chains (Parvari et al. 1988; Warr et al. 1995). Previously, we demonstrated that *N*-glycosylation sites on the CH3 domains of pigeon and chicken IgG, which correspond to the CH2 domains of mammalian IgG, are occupied only by high-mannose-type oligosaccharides, including the monoglucosylated oligosaccharides. This site-specific *N*-glycosylation pattern might have resulted from the steric hindrance imposed by the unique conformational structures of avian IgG.

In conclusion, some *N*-glycans, from both GEW-glycoproteins and GEY-IgG, contain Galα1-4Gal at their nonreducing termini. However, the remarkable differences of *N*-glycans of gull glycoproteins from those of pigeon glycoproteins are due to the incomplete or competitive α-galactosylation, α-sialylation, and diLacNAc-elongation of LacNAc branches, which consequently generate microheterogeneity with various isoforms by the substitution on different branches. Although incomplete α-galactosylation may imply that extensive Galα1-4Gal is unessential at least for egg glycoproteins, the microheterogeneity may serve to prevent a strong interaction with some pathogens, and influence the tropism by reducing the binding between host's glycans and their specific receptors on microbes and viruses.

Material and methods

Materials

Eggs of black-tailed gull (*Larus crassirostirs*) were obtained from Kabu Island in Hachinohe, Japan, and those of yellow-legged gull (*Larus cachinnans*) were from VIS Island in Croatia. GEW were separated from egg yolks, lyophilized, and kept at −20°C until used. Alkaline phosphatase-conjugated goat anti-mouse IgM and iminodiacetic acid (IDA) immobilized agarose were purchased from Sigma (St. Louis, MO). Glycoamidase F (GAF, also called PNGase F, glycopeptide *N*-glycosidase F or *N*-glycanase) was from Roche Diagnostics GmbH (Penzberg, Germany). Trypsin and chymotrypsin were from Worthington Biochem, Co. (Lakewood, NJ). Alkaline-phosphatase-conjugated concanavalin A (Con A), *Ricinus communis* agglutinin I (RCA-I), GS-I, and peanut agglutinin (PNA) were purchased from EY Lab., Inc. (San Mateo, CA). Anti-P_1 mAb (mouse IgM) was from Gamma Biologicals, Inc. (Houston, TX). Anti-(Galα1-3Gal) mAb (M86, mouse IgM) (Galili et al. 1998) was a generous gift from Dr Galili (Rush University, Chicago, IL). The BCIP/NBT kit for alkaline phosphatase assays was purchased from Zymed Laboratories Inc. (South San Francisco, CA). Neuraminidase from *Arthrobacter ureafaciens* was a generous gift from Marukin Shoyu Co. BCA™ Protein Assay reagents and the Eggcellent™ Chicken IgY Purification Kit were from Pierce (Rockford, IL). Polyvinylidene difluoride (PVDF) membranes for blotting and ultrafiltration membrane (YM 10) were from Millipore (Bedford, MA). DEAE-Sepharose Fast Flow (HiTrap, 1 mL) and Superdex 200 (HiLoad 26/60) columns were from GE Healthcare UK Ltd. (Buckinghamshire, UK). A Shim-Pack CLC-ODS column (6.0 × 150 mm) was from Shimadzu Co. (Kyoto, Japan). Butyl-Toyopearl 650 M and a TSKgel DEAE-5PW (7.5 × 75 mm) column were from Tosoh Biosep LLC (Montgomeryville, PA). Carbograph tubes (25 mg) were from Alltech. A CarboPac PA-20 column (3 × 150 mm) and the BioLC with an ED50 electrochemical detector for high-performance anion-exchange chromatography (HPAEC) were from Dionex Corporation (Sunnyvale, CA).

Buffers and standard procedures

TBS contains 50 mM Tris-HCl (pH 7.4) and 150 mM NaCl. TBST contains 0.1% Tween 20 in TBS. The procedures used for SDS–PAGE and lectin/antibody blotting are as published (Suzuki and Lee 2004). Protein concentrations were measured using the BCA assay (Smith et al. 1985), with bovine serum albumin (BSA) as a standard.

Isolation of ovotransferrin and ovomucoid from gull egg white

Lyophilized egg whites (5.5 g) from black-tailed gull were dissolved in 55 mL of distilled water, homogenized, dialyzed against water (12 kDa molecular cut-off), and then centrifuged to remove insoluble materials. The supernatant was filtered through a 0.45-μm membrane. An IDA-agarose column (1.5 × 12 cm, about 20 mL) was prewashed with 10 column volumes of 0.1 M Na_2CO_3, water, 0.1 M EDTA (pH 8.0), and then water. A Ni^{2+}-IDA column was generated by adding 40 mL of 0.1 M $NiCl_2$ to the prewashed IDA-agarose column, and washed with 200 mL of water, and then equilibrated with the binding buffer (50 mM Tris-HCl (pH 8.0), 0.5 M NaCl). GEW was diluted with the binding buffer and loaded onto the Ni^{2+}-IDA-column (1 g of total proteins/load). The column was washed with 150 mL of binding buffer, and the retained proteins were then eluted with 75 mL of elution buffer (0.1 M imidazole in the binding buffer). The Ni^{2+}-IDA-agarose-bound fraction was dialyzed against water and lyophilized. This fraction was further purified with DEAE-Sepharose using a gradient of 0–0.1 M NaCl in 10 mM Tris-HCl (pH 8.0), while monitoring the

absorbance at 280 nm. The major peak was collected and used as purified ovotransferrin.

The Ni^{2+}-IDA-agarose-unbound fraction was diluted with 4 M NaCl in 10 mM NaOAc (pH 5.0). A Butyl-Toyopearl 650 M column (2.5 × 16 cm, about 75 mL) was equilibrated with 4 M NaCl in 10 mM NaOAc (pH 5.0), and the Ni^{2+}-IDA-agarose-unbound fraction (1.2 g of proteins/load) was loaded onto the column. After the column was washed with 200 mL of 4 M NaCl in 10 mM NaOAc (pH 5.0) (flow rate: 0.5 mL/min), the concentration of NaCl was linearly decreased to reach 0 M within 400 mL of the mobile phase, and the column was washed with 230 mL of 10 mM NaOAc (pH 5.0). Proteins in the fractions were monitored with the absorbance at 280 nm. To detect the trypsin-inhibitory activity of ovomucoid-containing fractions, 10 μL of each fraction were preincubated with 10 μL of trypsin (0.1 mg/mL) in TBS with 2 mM $CaCl_2$ at 37°C for 10 min, and then 100 μL of 2 mM Nα-benzyoyl-D,L-arginine *p*-nitroanilide hydrochloride (BAPNA) was added to the reaction mixture. After incubation at 37°C for 30 min, the absorbance at 405 nm was measured to detect the product of trypsin-digestion. The ovomucoid fractions were collected, dialyzed against water, and then lyophilized.

Monosaccharide composition analysis

Glycoproteins (100 μg of GOM, 200 μg each of GOT and GEW) were hydrolyzed with 2 M trifluoroacetic acid at 100°C for 4 h to release neutral sugars or with 4 M HCl at 100°C for 6 h to release amino sugars (Fan et al. 1994). The released monosaccharides were analyzed with HPAEC using a CarboPac PA-20 (3 × 150 mm) column and isocratic elution with 10 mM NaOH at a flow rate of 0.4 mL/min. To release terminal α-linked galactose specifically, glycoproteins were incubated at 25°C overnight with green coffee bean α-galactosidase (50 mU for 10–20 μg of glycoproteins) in the 100 mM citrate-phosphate buffer (pH 6.0), and the released galactose was measured with HPAEC under the same conditions as described above. Sialic acids were released from glycoproteins using neuraminidase from *A. ureafaciens* (10 mU for 10–20 μg glycoproteins) by incubating the glycoproteins in 50 mM sodium acetate (pH 5.0) at 37°C overnight and were analyzed with HPAEC using a CarboPac PA-20 column and a gradient from 20 mM sodium acetate in 100 mM NaOH to 300 mM sodium acetate in 100 mM NaOH within 25 min at a flow rate of 0.4 mL/min.

Isolation of gull egg yolk IgG

Lyophilized egg yolk (4 g) from black-tailed gull was dissolved in 100 mL of distilled water, and centrifuged to remove insoluble materials. Egg yolk IgG was isolated with the Eggcellent™ Chicken IgY Purification Kit and further purified with gel-filtration using Superdex 200 (HiLoad 26/60, 2.6 × 60 cm) at a flow rate of 2.5 mL/min, with PBS as the mobile phase. Fractions containing egg yolk IgG were collected and concentrated with an Amicon Ultra-15 10 K (Millipore).

MALDI-TOF-MS analysis for glycoproteins

MALDI-TOF-MS analysis for glycoproteins isolated from GEW and egg yolk was performed using Voyager DE-STR (Applied Biosystems). An aliquot of each sample (0.5 μL) diluted with distilled water was mixed with 0.5 μL of 10 mg/mL sinapinic acid in 50% acetonitrile and 0.1% trifluoroacetic acid, and analyzed as described previously (Suzuki and Lee 2004).

Preparation of PA-derivatized oligosaccharides

N-Glycans were prepared as described previously (Suzuki et al. 2003). Briefly, glycoproteins were reduced and alkylated with dithiothreitol and iodoacetamide, respectively, and digested with trypsin/chymotrypsin mixtures in 50 mM NH_4HCO_3 (pH 8.0) at 37°C overnight. After inactivating the enzymes at 100°C for 10 min, oligosaccharides were released with GAF-treatment in 50 mM NH_4HCO_3, pH 8.0, at 37°C, overnight. The mixture was loaded onto Dowex 50 W × 2 (H^+ form, 50–100 mesh, Sigma), and the pass-through fraction was loaded onto a Carbograph tube (25 mg, Alltech). Oligosaccharides were eluted with 25% CH_3CN containing 0.05% trifluoroacetic acid. Lyophilized free oligosaccharide fractions were PA-derivatized as described previously (Takahashi et al. 1995). The mixture of PA-oligosaccharides was first separated by HPLC with a TSKgel DEAE-5PW column (7.5 × 75 mm) as described previously (Nakagawa et al. 1995), and the separated neutral, monosialylated, and disialylated fractions (monitored by fluorescence, Ex 320 nm, Em 400 nm) were collected separately and lyophilized. These fractions were individually dissolved in water and separated on a Shim-Pack CLC-ODS column (6.0 × 150 mm), as described previously (Tomiya et al. 1988).

Exo-glycosidase digestion

The *N*-glycans were digested with various exo-glycosidases under the following conditions: 2 mU of α2,3-specific neuraminidase from *Macrobdella decora* (Recombinant, Calbiochem) in 20 μL of 50 mM ammonium acetate buffer (pH 6.0) at 37°C for 24 h; 10 mU of α-galactosidase from green coffee bean (Calbiochem) in 20 μL of 50 mM ammonium acetate buffer (pH 6.0) at 37°C for 48 h; and 3 mU of β1,4-galactosidase from *Streptococcus pneumoniae* (Recombinant, Calbiochem) in 20 μL of 50 mM ammonium acetate buffer (pH 6.0) at 37°C for 24 h.

Permethylation and MS analyses

Permethylation of the released *N*-glycans, and subsequent MALDI-MS and MS/MS analyses on either a MALDI Q/TOF Ultima (Mircomass) or a MALDI TOF/TOF (ABI 4700 Proteomic Analyzer), were performed essentially as described previously (Yu et al. 2006). For additional nanoESI-MS and MS/MS analyses on an LTQ-Orbitrap XL hybrid FT mass spectrometer (Thermo Scientific), the permethylated glycans in 50% acetonitrile/1 mM NaOH were infused directly by static nanospray for data acquisition using either the linear ion trap or Orbitrap mass analyzer. Using the total ion mapping (TIM) functionality provided by the XCalibur instrument control software package (version 2.0.7), automated MS/MS spectra (at 35% collision energy) were obtained across the entire mass range *m/z* 500–2000 with a parent ion stepping of 2 mass units, in collection windows that were 2.8 mass units in isolation width (±1.4 amu). Two microscans, each 1000 ms, in maximum injection time with the AGC (auto gain control) on, were averaged for each collection window. The 0.8 mass unit overlap from step to step was set to allow minor parent ion signals occurring at the edge of an individual window to be sampled in a more representative fashion. For data analysis, the Qual Browser (Thermo

Fisher Scientific) was used in ion-map mode to filter out the parent ions containing specific terminal epitope, either by product ion or neutral loss sorting, to create pseudo-precursor ion or neutral loss scan profiles. These low resolution reconstructed TIM mass profiles were then matched against the high resolution, high accuracy MS profile obtained with the Orbitrap to facilitate manual peak assignment. False positives were manually eliminated by further examination of the respective ion trap MS^2 data.

Funding

National Institutes of Health Research (DK09970 to NS, YCL); Academia Sinica (to KHK); and the NRPGM Core Facilities for Proteomics and Glycomics (NSC97-3112-B-001-018; MALDI-MS and MS/MS data).

Acknowledgements

The authors thank Drs Katsuko Yamashita and Gordan Lauc for providing gull eggs. The Japanese Agency of Cultural Affairs and Aomori Prefectural Board of Education are acknowledged for their permission to collect Umineko (seagull) eggs. We further acknowledge the Academia Sinica Common Mass Spectrometry facilities located at the Institute of Biological Chemistry for the use of the LTQ-Orbitrap hybrid FT mass spectrometer.

Conflict of interest statement

None declared.

Abbreviations

CBB, Coomassie Brilliant Blue R-250; CEW, chicken egg white; CID, collision-induced dissociation; ESI, electrospray ionization; GAF, glycoamidase F; GalT, galactosyltransferase; GEW, gull egg white; GEY, gull egg yolk; GlcNAcT, *N*-acetylglucosaminyltransferase; GOM, gull ovomucoid; GOT, gull ovotransferrin; GS-I, *Griffonia simplicifolia* I; Hex, hexose; HexNAc, *N*-acetylhexosamine; HPAEC, high-performance anion-exchange chromatography; IDA, iminodiacetic acid; LacNAc, *N*-acetyllactosamine; MALDI, matrix-assisted laser desorption/ionization; MS, mass spectrometry; NeuAc, *N*-acetylneuraminic acid; ODS, octadecylsilica; PA, 2-aminopyridine; PBS, phosphate buffered saline; PEW, pigeon egg white; SiaT, sialyltransferase; TBS, Tris buffered saline; TIM, total ion mapping; TOF, time of flight.

References

Al-Mashikhi SA, Nakai S. 1987. Separation of ovotransferrin from egg white by immobilized metal affinity chromatography. *Agric Biol Chem.* 51:2881–2887.

Allen SD, Tsai D, Schachter H. 1984. Control of glycoprotein synthesis. The in vitro synthesis by hen oviduct membrane preparations of hybrid asparagine-linked oligosaccharides containing 5 mannose residues. *J Biol Chem.* 259:6984–6990.

Aoki K, Perlman M, Lim JM, Cantu R, Wells L, Tiemeyer M. 2007. Dynamic developmental elaboration of *N*-linked glycan complexity in the Drosophila melanogaster embryo. *J Biol Chem.* 282:9127–9142.

Chen HS, Chen JM, Lin CW, Khoo KH, Tsai IH. 2008. New insights into the functions and *N*-glycan structures of factor X activator from Russell's viper venom. *FEBS J.* 275:3944–3958.

Chu PC. 1995. Phylogenetic reanalysis of Strauch's osteological data set for the Charadriiformes. *Condor.* 97:174–196.

Fan J-Q, Namiki Y, Matsuoka K, Lee YC. 1994. Comparison of acid hydrolytic conditions for Asn-linked oligosaccharides. *Anal Biochem.* 219:375–378.

Gagneux P, Cheriyan M, Hurtado-Ziola N, van der Linden EC, Anderson D, McClure H, Varki A, Varki NM. 2003. Human-specific regulation of α2-6-linked sialic acids. *J Biol Chem.* 278:48245–48250.

Gagneux P, Varki A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology.* 9:747–755.

Gahmberg CG, Hakomori S. 1975. Surface carbohydrates of hamster fibroblasts: I. Chemical characterization of surface-labeled glycosphingolipids and aspecific ceramide tetrasaccharide for transformants. *J Biol Chem.* 250:2438–2446.

Galili U, LaTemple DC, Radic MZ. 1998. A sensitive assay for measuring α-Gal epitope expression on cells by a monoclonal anti-Gal antibody. *Transplantation.* 65:1129–1132.

Goldstein IJ, Poretz RD. 1986. Isolation, physicochemical characterization, and carbohydrate-binding specificity of lectin. In: Liener IE, Sharon N, Goldstein IJ, editors. *The Lectins: Properties, Functions, and Applications in Biology and Medicine*. New York: Academic: 35–247.

Hooper LV, Gordon JI. 2001. Glycans as legislators of host–microbial interactions: Spanning the spectrum from symbiosis to pathogenicity. *Glycobiology.* 11:1R–10R.

Kato I, Schrode J, Kohr WJ, Laskowski M Jr. 1987. Chicken ovomucoid: Determination of its amino acid sequence, determination of the trypsin reactive site, and preparation of all three of its domains. *Biochemistry* 26:193–201.

Kotani M, Kawashima I, Ozawa H, Ogura K, Ariga T, Tai T. 1994. Generation of one set of murine monoclonal antibodies specific for globo-series glycolipids: Evidence for differential distribution of the glycolipids in rat small intestine. *Arch Biochem Biophys.* 310:89–96.

Laskowski M Jr, Kato I, Ardelt W, Cook J, Denton A, Empie MW, Kohr WJ, Park SJ, Parks K, Schatzley BL, et al. 1987. Ovomucoid third domains from 100 avian species: Isolation, sequences, and hypervariability of enzyme-inhibitor contact residues. *Biochemistry.* 26:202–221.

Leslie GA, Benedict AA. 1970. Structural and antigenic relationships between avian immunoglobulins: II. Properties of papain- and pepsin-digested chicken, pheasant and quail IgG-immunoglobulins. *J Immunol.* 104:810–817.

Lindberg AA, Brown JE, Stromberg N, Westling-Ryd M, Schultz JE, Karlsson KA. 1987. Identification of the carbohydrate receptor for Shiga toxin produced by Shigella dysenteriae type 1. *J Biol Chem.* 262:1779–1785.

Lingwood CA, Law H, Richardson S, Petric M, Brunton JL, De Grandis S, Karmali M. 1987. Glycolipid binding of purified and recombinant *Escherichia coli* produced verotoxin in vitro. *J Biol Chem.* 262:8834–8839.

Lund N, Olsson ML, Ramkumar S, Sakac D, Yahalom V, Levene C, Hellberg A, Ma XZ, Binnington B, Jung D, et al. Forthcoming. The human Pk histo-blood group antigen provides protection against HIV-1 infection. *Blood*.

Lyerla TA, Gross SK, McCluer RH. 1986. Glycosphingolipid patterns in primary mouse kidney cultures. *J Cell Physiol.* 129:390–394.

Matsuura F, Ohta M, Murakami K, Matsuki Y. 1993. Structures of asparagine linked oligosaccharides of immunoglobulins (IgY) isolated from egg-yolk of Japanese quail. *Glycoconj J.* 10:202–213.

Nakagawa H, Kawamura Y, Kato K, Shimada I, Arata Y, Takahashi N. 1995. Identification of neutral and sialyl *N*-linked oligosaccharide structures from human serum glycoproteins using three kinds of high-performance liquid chromatography. *Anal Biochem.* 226:130–138.

Nakamura K, Fujita R, Ueno K, Handa S. 1984. Glycosphingolipids of porcine pancreas. *J Biochem (Tokyo).* 95:1137–1144.

Ohta M, Hamako J, Yamamoto S, Hatta H, Kim M, Yamamoto T, Oka S, Mizuochi T, Matsuura F. 1991. Structures of asparagine-linked oligosaccharides from hen egg-yolk antibody (IgY). Occurrence of unusual glucosylated oligo-mannose type oligosaccharides in a mature glycoprotein. *Glycoconj J.* 8:400–413.

Okuda T, Tokuda N, Numata S, Ito M, Ohta M, Kawamura K, Wiels J, Urano T, Tajima O, Furukawa K, et al. 2006. Targeted disruption of Gb3/CD77 synthase gene resulted in the complete deletion of globo-series glycosphingolipids and loss of sensitivity to verotoxins. *J Biol Chem.* 281:10230–10235.

Parvari R, Avivi A, Lentner F, Ziv E, Tel-Or S, Burstein Y, Schechter I. 1988. Chicken immunoglobulin γ-heavy chains: Limited VH gene repertoire, combinatorial diversification by D gene segments and evolution of the heavy chain locus. *EMBO J.* 7:739–744.

Paton TA, Baker AJ, Groth JG, Barrowclough GF. 2003. RAG-1 sequences resolve phylogenetic relationships within Charadriiform birds. *Mol Phylogenet Evol.* 29:268–278.

Patterson R, Younger JS, Weigle WO, Dixon F. 1962. Antibody production and transfer to egg yolk in chickens. *J Immunol.* 89:272–278.

Raju TS, Briggs JB, Borge SM, Jones AJ. 2000. Species-specific variation in glycosylation of IgG: Evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology.* 10:477–486.

Ramkumar S, Sakac D, Binnington B, Branch DR, Lingwood CA. 2009. Induction of HIV-1 resistance: Cell susceptibility to infection is an inverse function of globotriaosyl ceramide levels. *Glycobiology.* 19:76–82.

Shao MC, Wold F. 1989. The effect of the protein matrix on glycoprotein processing by oviduct Golgi enzymes. *J Biol Chem.* 264:6245–6251.

Sibley CG, Ahlquist JE. 1990. *Phylogeny and Classification of Birds: A Study in Molecular Evolution*. New Haven, CT: Yale University Press.

Sibley CG, Monroe BL. 1990. *Distribution and Taxonomy of Birds of the World*. New Haven, CT: Yale University Press.

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC. 1985. Measurement of protein using bicinchoninic acid. *Anal Biochem.* 150:76–85.

Stapleton AE, Stroud MR, Hakomori SI, Stamm WE. 1998. The globoseries glycosphingolipid sialosyl galactosyl globoside is found in urinary tract tissues and is a preferred binding receptor in vitro for uropathogenic *Escherichia coli* expressing *pap*-encoded adhesins. *Infect Immun.* 66:3856–3861.

Suzuki N, Khoo KH, Chen CM, Chen HC, Lee YC. 2003. *N*-Glycan structures of pigeon IgG: A major serum glycoprotein containing Galα1-4Gal termini. *J Biol Chem.* 278:46293–46306.

Suzuki N, Khoo KH, Chen HC, Johnson JR, Lee YC. 2001. Isolation and characterizaion of major glycoproteins of pigeon egg white: Ubiquitous presence of unique *N*-glycans containing Galα1-4Gal. *J Biol Chem.* 276:23221–23229.

Suzuki N, Laskowski M Jr, Lee YC. 2004. Phylogenetic expression of Galα1-4Gal on avian glycoproteins: Glycan differentiation inscribed in the early history of modern birds. *Proc Natl Acad Sci USA.* 101:9023–9028.

Suzuki N, Laskowski M Jr, Lee YC. 2006. Tracing the history of Galα1-4Gal on glycoproteins in modern birds. *Biochim Biophys Acta.* 1760:538–546.

Suzuki N, Lee YC. 2004. Site-specific *N*-glycosyaltion of chicken serum IgG. *Glycobiology.* 14:275–292.

Suzuki N, Lee YC. 2007. Glycophylogeny of Galα1-4Gal in avian egg glycoproteins. In: Kamerling JP, Boons G-J, Lee YC, Suzuki A, Taniguchi N, Voragen AGJ, editors. *Comprehensive Glycoscience: From Chemistry to Systems Biology*. Amsterdam: Elsevier: 237–251.

Suzuki Y. 2005. Sialobiology of influenza: Molecular mechanism of host range variation of influenza viruses. *Biol Pharm Bull.* 28:399–408.

Tai T, Yamashita K, Ogata-Arakawa M, Koide N, Muramatsu T, Iwashita S, Inoue Y, Kobata A. 1975. Structural studies of two ovalbumin glycopeptides in relation to the endo-β-*N*-acetylglucosaminidase specificity. *J Biol Chem.* 250:8569–8575.

Takahashi N, Khoo KH, Suzuki N, Johnson JR, Lee YC. 2001. *N*-Glycan structures from the major glycoproteins of pigeon egg white: Predominance of terminal Galα(1-4)Gal. *J Biol Chem.* 276:23230–23239.

Takahashi N, Nakagawa H, Fujikawa K, Kawamura Y, Tomiya N. 1995. Three-dimensional elution mapping of pyridylaminated *N*-linked neutral and sialyl oligosaccharides. *Anal Biochem.* 226:139–146.

Tomiya N, Awaya J, Kurono M, Endo S, Arata Y, Takahashi N. 1988. Analyses of *N*-linked oligosaccharides using a two-dimensional mapping technique. *Anal Biochem.* 171:73–90.

Warr GW, Magor KE, Higgins DA. 1995. IgY: Clues to the origins of modern antibodies. *Immunol Today.* 16:392–398.

Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. 1992. Evolution and ecology of influenza A viruses. *Microbiol Rev.* 56:152–179.

Wieruszeski JM, Michalski JC, Montreuil J, Strecker G, Peter-Katalinic J, Egge H, van Halbeek H, Mutsaers JH, Vliegenthart JF. 1987. Structure of the monosialyl oligosaccharides derived from salivary gland mucin glycoproteins of the Chinese swiftlet (genus *Collocalia*). Characterization of novel types of extended core structure, Galβ(1-3)[GlcNAcβ(1-6)] GalNAcα(1-3)GalNAc(-ol), and of chain termination, $[Gal\alpha(1\text{-}4)]_{0-1}[Gal\beta(1\text{-}4)]_2GlcNAc\beta(1\text{-}.)$. *J Biol Chem.* 262:6650–6657.

Yamashita K, Kamerling JP, Kobata A. 1983. Structural studies of the sugar chains of hen ovomucoid. Evidence indicating that they are formed mainly by the alternate biosynthetic pathway of asparagine-linked sugar chains. *J Biol Chem.* 258:3099–3106.

Yang Z, Bergstrom J, Karlsson KA. 1994. Glycoproteins with Galα4Gal are absent from human erythrocyte membranes, indicating that glycolipids are the sole carriers of blood group P activities. *J Biol Chem.* 269:14620–14624.

Yu SY, Khoo KH, Yang Z, Herp A, Wu AM. 2008. Glycomic mapping of *O*- and *N*-linked glycans from major rat sublingual mucin. *Glycoconj J.* 25:199–212.

Yu SY, Wu SW, Khoo KH. 2006. Distinctive characteristics of MALDI-Q/TOF and TOF/TOF tandem mass spectrometry for sequencing of permethylated complex type *N*-glycans. *Glycoconj J.* 23:355–369.