Glycobiology vol. 16 no. 3 pp. 244–257, 2006
doi:10.1093/glycob/cwj062
Advance Access publication on November 29, 2005

Glycomic survey mapping of zebrafish identifies unique sialylation pattern

Yann Guérardel[1,2,3], Lan-Yi Chang[2,4], Emmanuel Maes[3], Chang-Jen Huang[2,4], and Kay-Hooi Khoo[1,2,4]

[2]Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei 11529, Taiwan; [3]Unité de Glycobiologie Structurale et Fonctionnelle, USTL, 59655 Villeneuve d'Ascq, France; and [4]Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan

Received on October 9, 2005; accepted on November 14, 2005

Functional genomics and proteomics studies of the developmental glycobiology of zebrafish are greatly hampered by the current lack of knowledge on its glycosylation profile. To furnish the requisite structural basis for a more insightful functional delineation and genetic manipulation, we have initiated a survey mapping of the possible expression of stage-specific glycoconjugates in zebrafish. High-sensitivity mass spectrometry (MS) analysis in conjunction with the usual array of enzymatic and chemical derivatization was employed as the principal method for rapid differential mapping of the glycolipids and sequentially liberated *N*- and *O*-glycans from the total extracts. We demonstrated that all developmental stages of the zebrafish under investigation, from fertilized eggs to hatched embryos, synthesize oligomannosyl types of *N*-glycans, as well as complex types with additionally β4-galactosylated, Neu5Ac/Neu5Gc monosialylated Lewis x termini. A combination of collision-induced dissociation (CID)-MS/MS and nuclear magnetic resonance (NMR) analyses led to the identification of an abundant and unusual mucin-type O-glycosylation, based on a novel sequence Fucα1-3GalNAcβ1-4(Neu5Ac/Neu5Gcα2-3)Galβ1-3GalNAc. This core structure may be further oligosialylated, but exclusively in the earlier development stages. Similarly, MS and MS/MS analyses of the extracted glycolipid fraction revealed the presence of a heterogeneous family of oligosialylated lactosylceramide compounds. In contrast to the *O*-glycans, these glycolipids only appear in the later development stages, suggesting a complex pattern of regulation for sialyltransferase activities during zebrafish embryogenesis.

Key words: glycomics/mass spectrometry/sialylation/structure analysis/zebrafish

Introduction

The zebrafish, *Danio rerio*, has emerged in recent years as an excellent model system to study the genetic underpinnings of vertebrate development. Large-scale genetic screens have identified thousands of mutant variants that allow *in vivo* dissection of developmental processes at single cell and molecular resolution. Among other advantages, it has a very short generation time which enables fast generation of transgenic lines. Its embryos are optically transparent, develop externally, and can absorb mutagens directly from surrounding water. The popularity of this model system has led to rapid accumulation of a large body of genetic data which provides a unique opportunity to follow the functional involvement of glycoconjugates in a vertebrate model throughout its complete embryogenesis. Such studies are increasingly accessible owing to recent development of tools designed to identify and classify enzymes involved in complex carbohydrate biogenesis based on sequence and folding similarities (Coutinho and Henrissat, 1999).

Strategies based on sequence similarities appear to be very successful in higher vertebrate models, including human, of which the glycosylation potentials are already well known. However, they show major shortcomings in the case of phylogenetically distant models because of the existence of unknown forms of glycosylation deriving from unpredictable enzymatic activities. In such cases, the definition of the fine specificity of glycosylation-related enzymes is rendered very delicate by the extreme variability of their activities. Such variability has been well illustrated by the modulation of FUT-3 substrate acceptor through the mutation of a single amino acid in its hypervariable stem (Dupuy *et al.*, 1999). The presence of unpredictable glycan structures has been identified in many animal models including those commonly used for developmental studies (Guerardel *et al.*, 2000, 2001, 2003; Haslam and Dell, 2003), which amply demonstrated the pertinence of integrated strategies combining genetic approach with direct observation of endogenous forms of glycosylation. Only then can the fine specificity of identified enzymes be truly assessed using endogenous substrates rather than commonly found glycan motifs.

In zebrafish, structural and functional studies of glycoconjugates have so far been focused only on very specific forms of glycosylation. In particular, the involvement of chitin oligosaccharides during zebrafish embryogenesis was strongly suggested by the inhibition of their biosynthesis (Bakkers *et al.*, 1997; Semino *et al.*, 1998; Semino and Allende, 2000). More recently, the existence of a specific receptor for chitin tetrasaccharide that would link its activity to the Raf, MEK, and ERK pathway in zebrafish cells was demonstrated (Snaar-Jagalska *et al.*, 2003). Other glycoconjugates such as the glycosaminoglycans (GAG) and polysialic acids were likewise shown to express a wide range of functions during zebrafish embryogenesis, including central nervous system (Bernhardt and Schachner, 2000; Marx *et al.*, 2001; Becker and Becker, 2002) and muscle development (Bink *et al.*, 2003). However, definitive structures of

[1]To whom correspondence should be addressed; e-mail: kkhoo@gate.sinica.edu.tw/yann.guerardel@univ-lille1.fr

glycolipids, N- or O-linked glycans have hitherto not been reported which precluded any functional studies on these essential components.

As a prelude to decipher the influence of glycosylation during zebrafish embryogenesis, we have initiated a systematic profiling of glycoconjugates at different developmental stages. The first aim was to define the structures of major glycoprotein-derived glycans and of glycolipids expressed by this organism, from which the endogenous activity of glycosylation-related enzymes, including glycosyltransferases, can be inferred. Mass spectrometry (MS)-based glycomics mapping were followed by more detailed analysis for the novel structural features thus identified. In parallel, comparative analyses were extended to the extracted glycomes from other developmental stages to critically evaluate their differential expression, especially in relation to the unique sialylation pattern.

Results

Our overall glycomic survey mapping strategy involved sequential extraction of glycolipids and glycoproteins and the subsequent sequential release of *N*- and *O*-glycans from the proteolytic-digested peptides/glycopeptides mixtures, for matrix-assisted laser-desorption ionization-MS (MALDI-MS) and MS/MS analyses. The released glycans were permethylated to allow more informative MS/MS sequencing, but native glycans were also analyzed where sample amount permitted. Such approach, in general, gives a good representative profile of the glycome but does not optimize for the yield of any particular class of glycoconjugates. It provides the first picture, uncovers any novel structural features, and facilitates subsequent more detailed investigations. Typically, the fertilized eggs at five distinctive developmental stages, 0.5, 8, 24, 45, and 48 h, were analyzed to allow a fair assessment of possible developmental regulation from first cell stage to hatching. Any significant differences were noted while common features were reported without distinguishing the origin of sample stage.

Identification of the major N*-glycans*

MALDI-MS profiling of the permethyl derivatives of *N*-glycans released from the total zebrafish embryo extracts afforded five major peaks at *m/z* 1579, 1783, 1988, 2193, and 2397, corresponding respectively, to sodiated molecular ions, $[M + Na]^+$, of the composition $Hex_{5\text{-}9}HexNAc_2$ (Figure 1A). Further collision-induced dissociation (CID)-MS/MS analysis and treatment with α-mannosidase (data not shown) demonstrated that these major signals are indeed the common high–mannose-type structures. In addition, several signals of lower intensity were visibly present among which a prominent cluster at *m/z* 3551, 3581, and 3611 could be tentatively assigned as $[M + Na]^+$ of $Neu5Ac_2Fuc_2Hex_7HexNAc_4$, $Neu5Ac_1Neu5Gc_1Fuc_2Hex_7HexNAc_4$, and $Neu5Gc_2Fuc_2Hex_7HexNAc_4$, respectively. These unusual compositions were shown by MALDI-MS/MS analyses to be biantennary complex-type structures with monosialylated Hex_2(Fuc)HexNAc sequence on both antennae (Figure 2). Both parent ions at *m/z* 3551 and 3611 afforded similar consecutive losses of terminal sialic acid residues and monosialylated antennae. Importantly, after losing both sialylated antennae, a common fragment ion at *m/z* 1143 was produced which corresponds to the sodiated trimannosyl core, $Man_3GlcNAc_2$, containing two free OH groups and thus confirming their biantennary nature.

For the $Neu5Ac_2$-containing parent (*m/z* 3551), a primary sodiated b ion, $Neu5Ac_1Hex_2$(Fuc)HexNAc, was detected at *m/z* 1225, accompanied by a $Neu5Ac^+$ oxonium ion at *m/z* 376 and a sodiated c ion, Neu5Ac-Hex_2-OH, at *m/z* 824 (Figure 2B). In comparison, the $Neu5Gc_2$-containing parent (*m/z* 3611) afforded a sodiated b ion, $Neu5Gc_1Hex_2$(Fuc)HexNAc, at *m/z* 1255, a $Neu5Gc^+$ ion at *m/z* 406, and a sodiated c ion, Neu5Gc-Hex_2-OH, at *m/z* 854 (Figure 2C). Further loss of the distinguishing Neu5Ac/Neu5Gc residue from the respective primary b and c ions degenerated the mass difference and yielded the common secondary fragment ions at *m/z* 850 and 449, corresponding to $(HO)_1Hex_2$-$(Fuc)_1$HexNAc and $(HO)_1Hex_2$-OH, respectively. The mass difference of 60 u between the two parents could thus be unambiguously attributed to a Neu5Ac and Neu5Gc difference (30 u) on each of the two monosialylated antennae. It could be further deduced that the third molecular ion signal (*m/z* 3581) in between the $Neu5Ac_2$- and $Neu5Gc_2$-containing parents corresponds to a similar biantennary complex-type structure carrying a Neu5Ac- and a Neu5Gc-sialylated antennae.

Higher in mass (Figure 1A) and of even lower abundance was another cluster of molecular ion signals which could be assigned as trisialylated triantennary complex-type structures with similar monosialylated terminal sequence carrying the Neu5Gc/Neu5Ac heterogeneity. Thus the signal at *m/z* 4830 corresponds to species with all three Hex_2(Fuc)HexNAc antennae sialylated by Neu5Gc, whereas the one at *m/z* 4740 carries only Neu5Ac-Hex_2(Fuc)HexNAc antenna. Supporting data were obtained when after desialylation by neuraminidase, a peak at *m/z* 3656 was detected which corresponds to $[M + Na]^+$ of a triantennary complex-type structure with three Hex_2-$(Fuc)_1$HexNAc antennae. The corresponding desialylated biantennary structure was observed as a major sodiated molecular ion at *m/z* 2827 (Figure 1B). Interestingly, after an overnight (>12 h) digestion, Neu5Ac desialylation appeared to be more complete than removal of Neu5Gc. Additional molecular ions corresponding to incompletely digested mono-Neu5Gc sialylated bi- and triantennary structures were detected at *m/z* 3218 and 4048, respectively (Figure 1B), but not their mono-Neu5Ac-sialylated counterparts which could only be observed if the neuraminidase digestion was kept to a shorter period (data not shown).

In accordance with MS/MS sequencing of the sialylated counterparts described above, MALDI-MS/MS on the desialylated biantennary structure yielded prominent nonreducing terminal primary fragment ions at *m/z* 463 and 864 (Figure 2A), corresponding to sodiated c ion, Hex_2-OH, and b ion, Hex_2-(Fuc)HexNAc, respectively. The Fuc substitution could be deduced as 3-linked to the HexNAc based on the detection of the secondary ions produced through the elimination of the Fuc (minus 206 mass units from the parent and other major primary fragment ions), whereas the

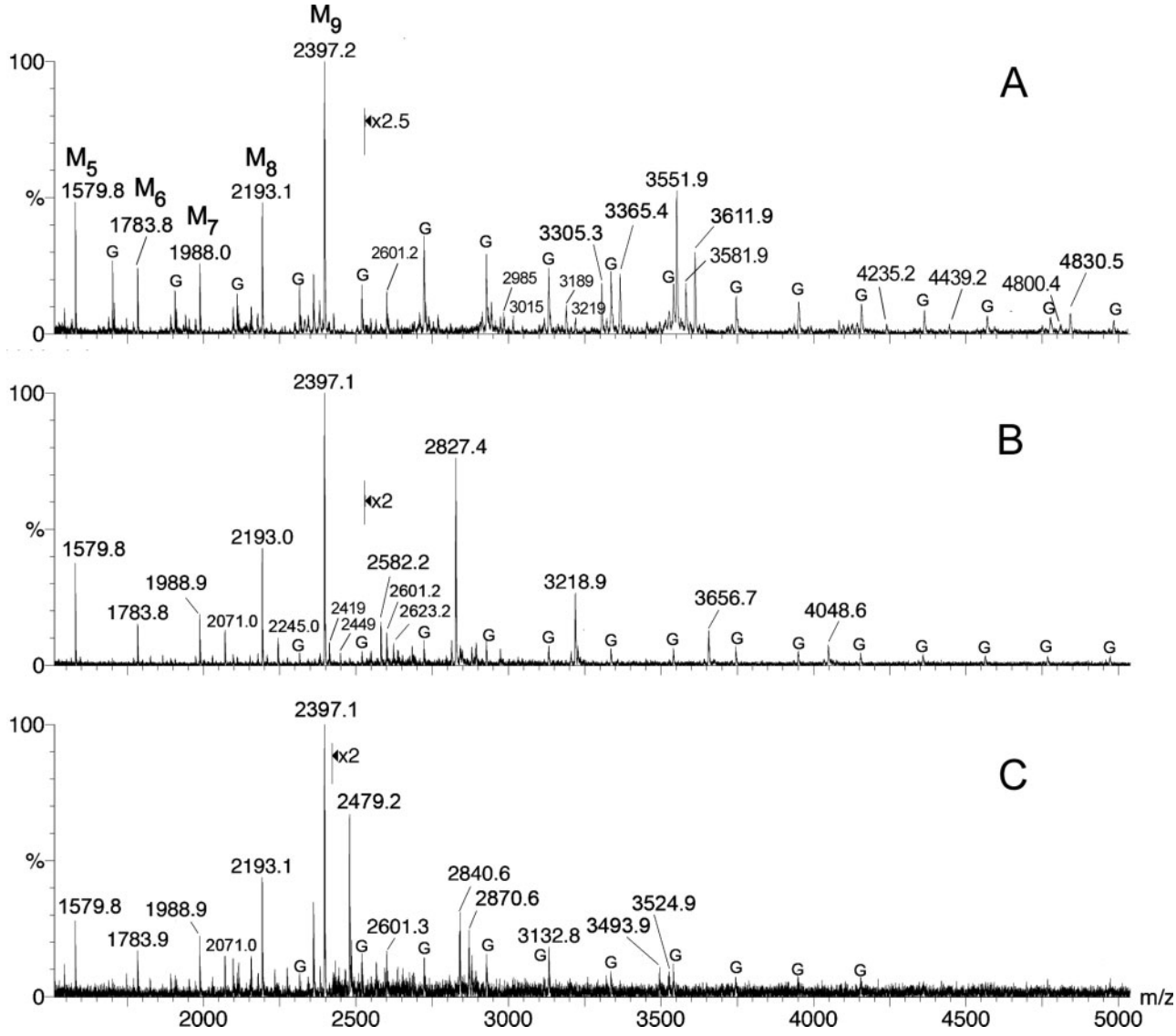

Fig. 1. MALDI-MS profiles of the permethylated *N*-glycans from zebrafish embryos before (**A**) and after neuraminidase (**B**) or after aqueous hydrofluoric acid (**C**) treatment. High–mannose-type structures were labeled M5–M9 in (**A**), representing $Man_{5\text{-}9}GlcNAc_2$ structures. Signal at *m/z* 2601 most likely corresponds to $Glc_1Man_9GlcNAc_2$. Glucose oligomer contaminants were labeled G. In panel **B**, additional minor signals corresponding to nonsialylated biantennary structures with incomplete fucosylation and/or galactosylation were also detected (*m/z* 2071, 2245, 2419, 2449, 2623), the smallest of which at *m/z* 2071 could be assigned as nonfucosylated biantennary *N*-glycan with simple Gal-GlcNAc termini. In panel **C**, complete defucosylation and desialylation produced the bi- (*m/z* 2479) and triantennary (*m/z* 3132) structures carrying $Hex_2HexNAc$ termini, accompanied by their mono-Neu5Ac/Neu5Gc sialylated counterparts at *m/z* 2840/2870 and 3492/3524, respectively. Under the conditions employed, aqueous hydrofluoric acid would remove α2,3,4-Fuc almost completely but sialic acid only partially. For simplicity, other minor products corresponding to a combination of incomplete galactosylation and desialylation were not labeled.

elimination of ±Neu5Ac/Neu5Gc-Hex_2 was not observed. Further confirmation was obtained when the desialylated structures were digested with β4-galactosidase. MALDI-MS and MS/MS analyses demonstrated that one Hex was removed from each of the nonreducing termini, whereas prior defucosylation with aqueous hydrofluoric acid afforded bi- and triantennary structures with Hex_2-HexNAc termini (Figure 1C) which could then be completely degalactosylated by β4-galactosidase (data not shown). The failure to remove the internal Gal attached to a fucosylated GlcNAc is consistent with the well-known selectivity of the β4-galactosidase acting on a Gal-(Fuc)GlcNAc unit. Linkage analysis on the isolated sialylated structures (see *Materials and methods*) further showed that the amount of terminal Gal relative to 2-linked Man or 3,6-linked Man did not change significantly before and after desialylation. Moreover, mono-substituted Gal residue was not detected. Instead, 3,4-linked Gal was quantitatively converted to 4-linked Gal after desialylation, therefore indicating that the sialic acid was attached to the 3-position of an internal 4-linked Gal and not to the terminal Gal.

Taken together, the data unambiguously defined the common monosialylated terminal sequence on each antennae of the major bi- and triantennary complex-type

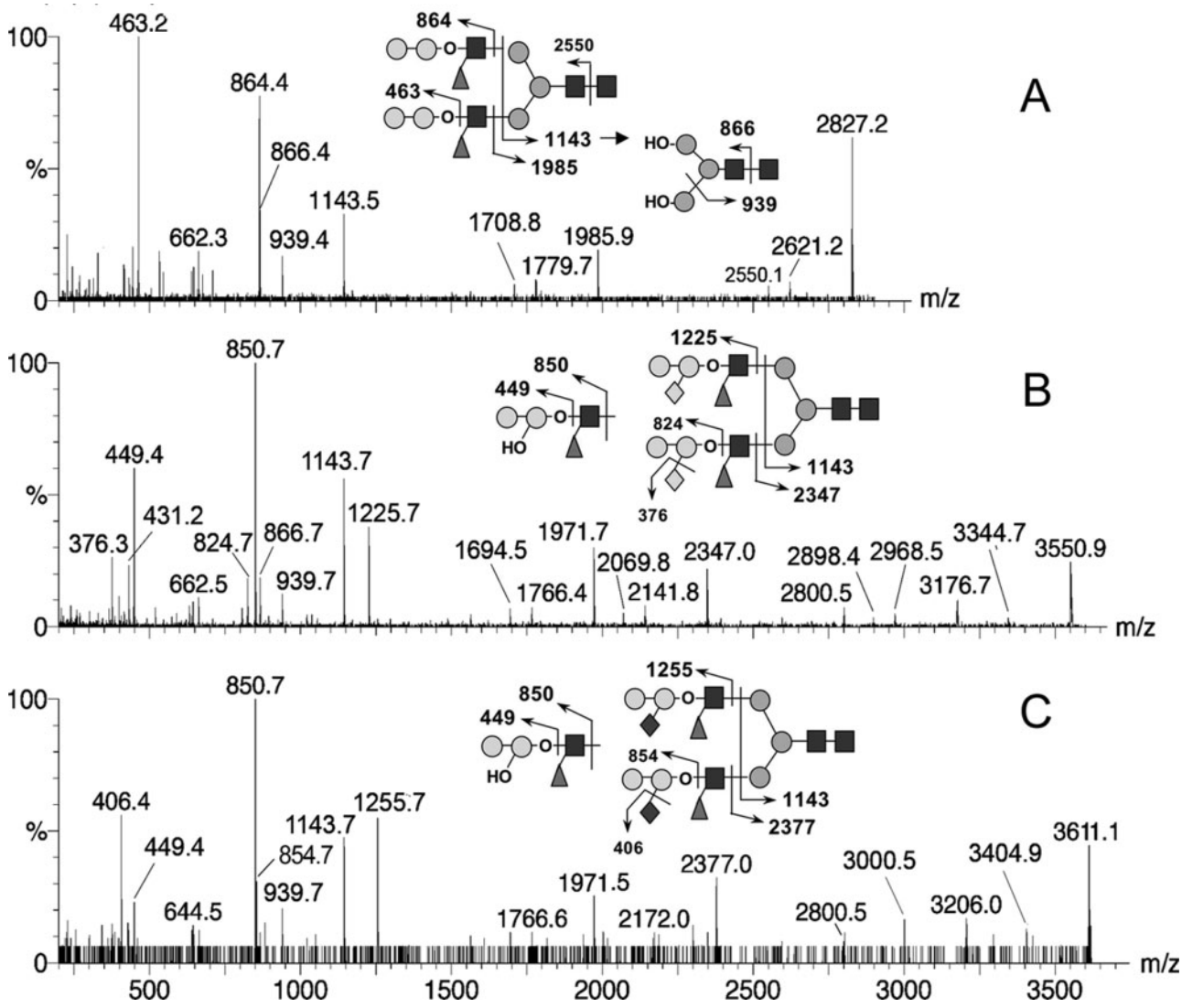

Fig. 2. MALDI-MS/MS sequencing of the biantennary complex-type *N*-glycans from zebrafish embryos. (**A**) MS/MS on the desialylated parent at *m/z* 2827; (**B**) MS/MS on the Neu5Ac-disialylated parent ion at *m/z* 3551; (**C**) MS/MS on the Neu5Gc-disialylated parent ion at *m/z* 3611. Major fragmentation pattern is as indicated schematically. Cleavage ions are mostly b and y ions except when indicated with an oxygen atom which correspond to c ions. All three gave common trimannosyl core ions at *m/z* 1143 as shown in panel **A** which yielded additional common fragment ions (*m/z* 939, 866, 662). Elimination of Fuc (minus 206 u) from parent and primary fragment ions are commonly observed, whereas loss of terminal Neu5Ac yielded signals at *m/z* 3176, 2800 (from 3550), and 1971 (from 2347) in panel **B**; loss of terminal Neu5Gc gave signals at *m/z* 3206, 2800 (from 3611), and 1971 (from 2377) in panel **C**. Symbols used are square, HexNAc; circle, Hex; diamond, Neu5Gc (dark) and Neu5Ac (light); triangle, Fuc. OH denotes exposed hydroxyl group because of cleavage on the permethyl derivatives.

N-glycan structures as Galβ1-4(Neu5Gc/Neu5Acα2-3)Galβ1-4(Fucα1-3)GlcNAc, namely an internal Lewis x unit which was further galactosylated and sialylated with either Neu5Gc or Neu5Ac. A very small amount of incomplete sialylation could be detected (Figure 1A) as monosialylated bi- (*m/z* 3189/3219) and disialylated triantennary structures (*m/z* 4439), as well as species that lack both sialic acid and Gal residues on one of the antennae (*m/z* 2985/3015; 4235). Curiously, the major disialylated biantennary structures were found to occur also as minor species lacking the reducing terminal GlcNAc, giving sodiated molecular ion signals at *m/z* 3305/3335/3365 (Figure 1A). CID MS/MS analysis of its desialylated counterpart (*m/z* 2582; Figure 1B) firmly established that the same antennary sequence is carried on the implicated $Hex_3HexNAc_1$ core in place of the usual $Man_3GlcNAc_2$ for *N*-glycans (data not shown).

Identification of the major O*-glycans*

MALDI-MS analysis of the permethylated *O*-glycans, released from de-N-glycosylated peptides as oligoglycosyl alditols through reductive elimination, afforded two predominant molecular ion signals at *m/z* 1315 and 1345. Other barely detectable weak signals at higher *m/z* values became more apparent only after enrichment by stepwise elution on an anion exchange column (Dowex 1 × 2 anionic resin). Thus, an early eluting fraction was found to contain only the two major signals, whereas a higher salt-eluted

fraction carried additional minor signals at higher mass range (Figure 3A), including the pair at *m/z* 1706/1736 which was related to the *m/z* 1315/1345 pair by a Neu5Gc increment. Assignment of the corresponding compositions and sequences was afforded by MALDI CID-MS/MS analyses, as schematically shown in Figure 3B–D.

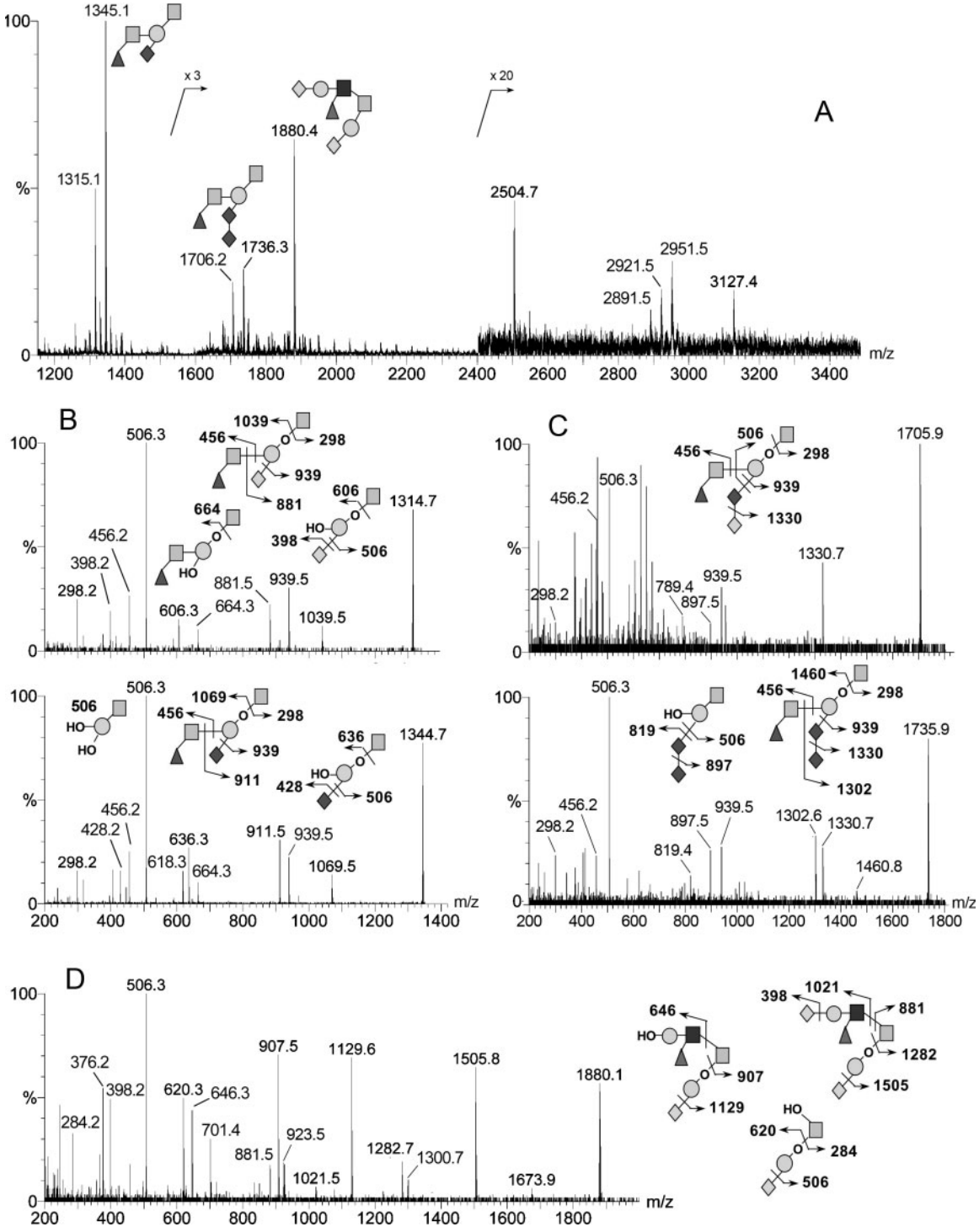

Fig. 3. MALDI-MS profile of permethylated *O*-glycans from zebrafish embryos (**A**) and MS/MS analyses on the major signals detected (**B**–**D**). MS profile shown in panel **A** was that on higher salt-eluted fraction from anion exchange chromatography. Nonenriched and earlier eluting fraction both afforded only the two major peaks at *m/z* 1315/1345 (not shown). For clarity, fragmentation pattern and the origin of major fragment ions are schematically indicated on structures corresponding to both parent and primary fragment ions. As in Figure 2, cleavage ions are mostly b and y ions except when indicated with an oxygen atom which correspond to c and z ions. OH denotes exposed hydroxyl group because of cleavage on the permethyl derivatives. All reducing end HexNAc of the released *O*-glycans is HexNAcitol by implication. The MS/MS spectrum of *m/z* 1705 (**C**, **upper** panel) contained several abundant low mass ions clearly not derived from the parent. Symbols used are square, HexNAc; circle, Hex; diamond, Neu5Gc (dark) and Neu5Ac (light); triangle, Fuc.

Notably for the parent ions at *m/z* 1315 and 1335 (Figure 3B), a core 1-type structure was indicated by the common z_1 ion at *m/z* 298, consistent with a mono-substituted reducing end HexNAcitol. The presence of a nonreducing terminal Neu5Ac/Neu5Gc was supported by the respective b_1 ions at *m/z* 398 (Neu5Ac) and 428 (Neu5Gc), as well as their facile loss from either the parent ions or the c ions at *m/z* 1039/1069 which degenerated the 30 u mass difference, giving rise to the common ions at *m/z* 939 and 664, respectively. Another distinctive set of b and y ion pair at *m/z* 456 and 881/911 defined a nonreducing terminal Fuc-HexNAc moiety, thus completing a rather unique sequence of (Fuc-HexNAc)(Neu5Ac/Neu5Gc)Hex-HexNAcitol. Localization of the terminal Neu5Ac/Neu5Gc to an internal Hex was supported by the common y ion at *m/z* 506 which corresponds to $(OH)_2$Hex-HexNAcitol, and the ions at *m/z* 606/636 corresponding to $(OH)_1$(Neu5Ac/Neu5Gc)$_1$Hex-OH.

Both the parent ions at *m/z* 1706 and 1736 likewise afforded a prominent fragment ion at *m/z* 506, which together with the z_1 ion at *m/z* 298, clearly indicated a similar core 1-type structure branched at the internal Hex (Figure 3C). For the *m/z* 1736 parent which gave a better quality MS/MS data, direct loss of either a single terminal Neu5Gc or a dimeric Neu5Gc-Neu5Gc yielded the y ions at *m/z* 1330 and 939, respectively. This was accompanied by the corresponding b ion for Neu5Gc-Neu5Gc at *m/z* 819 which firmly established the presence of a Neu5Gc-disialylated sequence. Importantly, the characteristic b ion at *m/z* 456 defined a similar nonreducing terminal Fuc-HexNAc moiety, complemented by the pairing y ion at *m/z* 1302 which corresponds to $(OH)_1$(Neu5Gc-Neu5Gc)Hex-HexNAcitol. Further loss of a terminal Neu5Gc from the latter yielded the ion at *m/z* 897, whereas losing both Neu5Gc residues gave the aforementioned $(OH)_2$Hex-HexNAcitol at *m/z* 506. The MS/MS spectrum of the other parent at *m/z* 1706 was of inferior quality and did not afford a full range of fragment ions. Nevertheless, apart from the common ions at *m/z* 456 and 506 which established its basic core structure, direct loss of either a terminal Neu5Ac (*m/z* 1330) or a dimeric Neu5Ac-Neu5Gc (*m/z* 939), but not a terminal Neu5Gc, indicated that its internal Hex was extended by a Neu5Gc and terminating with a Neu5Ac, contrasting with the other disialylated structure which carries a Neu5Gc-Neu5Gc extension. The detection of the corresponding b ion at *m/z* 789 for Neu5Ac$_1$Neu5Gc$_1$ and not *m/z* 819 for Neu5Gc$_2$ further supported the assignment. Thus, the disialylated structures may be considered as carrying an extra Neu5Ac or Neu5Gc extension from a Neu5Gc monosialylated structure but not a Neu5Ac monosialylated structure, all of which share the same basic Fuc-HexNAc-Hex-HexNAcitol backbone.

Monosaccharide composition analysis of the purified monosialylated fraction revealed an almost equal relative amount of Gal, Fuc, GalNAc, GalNAcitol, and sialic acid, fully consistent with the assigned sequence. To elucidate the complete primary structures of the two oligoglycosyl alditols, the monosialylated fraction was analyzed as a mixture by 400 MHz ^{1}H-nuclear magnetic resonance (^{1}H-NMR) spectroscopy. Chemical shifts of the protons of individual constituents obtained from sequential ^{1}H-1H homonuclear correlation spectroscopy-90 (COSY-90) and total correlation spectroscopy (TOCSY) experiments were compiled in Table I. Spin systems of monosaccharides confirmed the results obtained from composition analysis. In particular, as shown on TOCSY spectrum (Figure 4A), the internal HexNAc residue was unambiguously identified as a GalNAc owing to its spin system, reminiscent of the galacto configuration. The presence of both Neu5Ac and Neu5Gc was easily assessed owing to the observation of intense *N*-acetyl and *N*-glycolyl protons as singlets at δ2.045 and 4.122 p.p.m., respectively. Differential integration of these signals afforded an estimated Neu5Gc/Neu5Ac ratio of about 8:2. However, chemical shifts of their respective H-3ax, H-3eq, and H-5 could not be distinguished from one another.

Table I. ^{1}H-NMR chemical shifts of monosialylated *O*-glycans

	Chemical shifts of ^{1}H (p.p.m.)						
Residue	1	2	3	4	5	6	NAc/NGc
GalNAc-ol I	3.79/3.76	4.38	4.064	3.531	4.148	3.64	–
Gal(β1-3) II	4.571	3.416	4.182	4.117	3.78	nd	–
GalNAc(β1-4) III	4.791	4.052	3.705	3.964	nd	nd	–
Fuc(α1-3) F	5.088	3.702	3.903	3.812	4.141	1.206	–
Neu5Gc(α2-3)	–	–	1.951/2.705	3.873	3.655	–	4.122
Neu5Ac(α2-3)	–	–	1.951/2.705	3.873	3.655	–	2.045

nd, not determined.

As expected, the chemical shifts of H-2 and H-5 of GalNAc-ol are clearly indicative of mono substitution of GalNAc-ol by Gal in β1-3 linkage (Kamerling and Vliegenthart, 1992), consistent with a core 1-type structure deduced by CID/MS-MS analyses. The chemical shift of GalNAc H-1 at δ4.791 p.p.m. in conjunction with its coupling constant $J_{1,2}$ of about 8.5 Hz established the GalNAc III residue as β-anomer. In accordance with previous work (Herkt *et al*., 1985; Coppin *et al*., 2002), the chemical shifts of Gal, GalNAc, and Neu5Ac/Neu5Gc residues are indicative of the sequence GalNAcβ1-4(Neu5Acα2-3)Galβ1-, known as the Cad determinant. In particular, the downfield position of H-3ax of Neu5Ac or Neu5Gc to δ1.951 p.p.m. is very specific to this motif (Mourad *et al*., 2001). However, the chemical shift values of GalNAc protons showed significant discrepancies from those of Cad determinant, in particular, its H-2 signal that was noticeably deshielded by ~0.15 to δ3.903 p.p.m., suggesting that the GalNAc residue was further substituted. Substitution of GalNAc residue was assessed by the observation of GalNAc ^{13}C chemical shifts owing to heteronuclear ^{13}C-1H heteronuclear multiple quantum coherence (HMQC) experiment (data not shown). In particular, the GalNAc C-3 signal was strongly deshielded to δ79.1 p.p.m. compared with C-4 signal at δ70.0 p.p.m., which indicated that only C-3 position bore further substitution. Accordingly, nuclear Overhauser effect spectroscopy (NOESY) experiment showed an intense NOE contact between Fuc H-1 at δ5.088 p.p.m. and GalNAc III H-3 at δ3.705 p.p.m. (Figure 4B), confirming the terminal Fucα1-3GalNAc sequence. It also corroborated

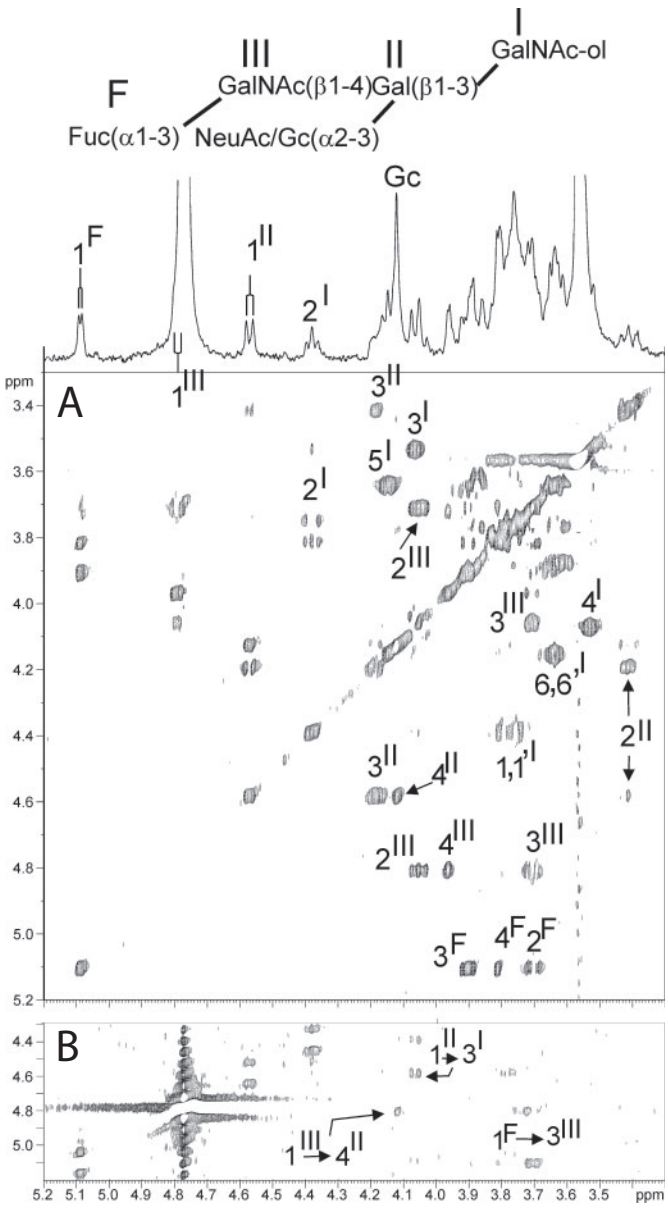

Fig. 4. Four hundred megahertz NMR analysis of monosialylated *O*-glycans purified from the total *O*-glycans by anion exchange chromatography. (**A**) Two-dimensional ^{1}H-1H TOCSY spectrum (δ^{1}H: 3.3–5.2 and 3.3–5.2) that allows the observation of the spin systems of residues II, III, and F. The presence of Neu5Gc and Neu5Ac are clearly established owing to intense peaks of *N*-glycolyl protons δ4.122 p.p.m. and *N*-acetyl protons at δ2.045 p.p.m. (not shown). (**B**) Two-dimensional ^{1}H-1H NOESY spectrum (δ^{1}H: 3.3–5.2 and 4.3–5.2) showing the connectivities between residues F, III, II, and I.

the assigned GalNAc1-4Gal1-3GalNAc-ol backbone sequence owing to clear NOE contacts between GalNAc H-1 and Gal H-4 and between Gal H-1 and GalNAc-ol H-3 at δ4.571 and 4.064 p.p.m., respectively.

In summary, the monosaccharide composition, MS/MS and NMR data collectively and unambiguously defined the two major monosialyated *O*-glycans as: Fucα1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-3GalNAc-itol and Fucα1-3GalNAcβ1-4(Neu5Gcα2-3)Galβ1-3GalNAc-itol. At a much lower level, the latter, but not the former, can be further extended by an additional Neu5Gc or Neu5Ac. At even higher mass, another cluster of extremely weak signals at *m/z* 2891, 2921, and 2951 apparently also exhibited the Neu5Gc/Neu5Ac heterogeneity of 30 u apart (Figure 3A), but no CID-MS/MS could be successfully obtained to establish their sequence. In contrast, the signals at *m/z* 1880, 2504, and 3127 did not afford the Neu5Gc/Neu5Ac sialylation pattern and apparently constitute a different series of *O*-glycan structures.

MALDI-MS/MS on the parent ion at *m/z* 1880 (Figure 3D) afforded the same fragment ion at *m/z* 506, corresponding to a disubstituted Hex-HexNAcitol core. However, instead of z_1 ion at *m/z* 298, *m/z* 284 was detected and therefore indicating that branching is at the reducing end HexNAcitol. Terminal Neu5Ac substitution was evident from the protonated and sodiated b_1 ions at *m/z* 376 and 398, respectively, whereas consecutive losses of two terminal Neu5Ac (*m/z* 1505 and 1129) from the parent suggested a disialylated structure with two different monosialylated termini. The c and z ion pair at *m/z* 620 and 1282, together with the other y_1 ion at *m/z* 881 (see the schematic drawing on Figure 3D) is consistent with a Neu5Ac-Hex motif attached to the 3-arm of a 3,6-branched HexNAcitol and a $Neu5Ac_1$-($Fuc_1Hex_1HexNAc_1$)- motif on the 6-arm. The latter was supported by observing the corresponding b ion at *m/z* 1021 and after losing the terminal Neu5Ac at *m/z* 646. In the absence of other data owing to lack of sample material, the exact structure for the fucosylated motif could not be established although facile elimination of a 3-linked Fuc (minus 206 u) from the parent and several other ions was strongly indicative of a Lewis x epitope. This is also consistent with it being extended by another one or two such -[Hex(Fuc)HexNAc]- repeats on the 6-arm to give the higher mass molecular ions detected at *m/z* 2504 and 3127 (Figure 3A), corresponding to a sialylated poly-Lewis x sequence.

As in the case with the *N*-glycans, the two major monosialylated *O*-glycans characterized were consistently detected across all five developmental stages of the zebrafish embryos. However, larger *O*-glycans, either based on the same core 1-type structure but disialylated on the Neu5Gc appendage or based on a distinct branched core 2-type structure with Neu5Ac monosialylation on both arms, were in general of very low abundance and only readily detectable in earlier stages (0.5 and 8 h). Similar enrichment on anion exchange column failed to yield any signal for samples derived from the 45 and 48 h embryonic stages while there was some batch-to-batch variation for detecting their presence in the 24 h samples. Thus although their low abundance and heterogeneity in structures precluded firm conclusion with respect to their exact structures and developmentally regulated expression, oligosialylation on *O*-glycans appeared to be preferentially associated with early development, before the completion of morphological differentiation.

MS profiling of the glycolipids

In contrast to oligosialylation on the *N*- and *O*-glycans, MALDI-MS profiling of the permethyl derivatives of glycolipids extracted from various developmental stages of zebrafish embryos revealed that potential molecular ion signals corresponding to oligosialylated glycosphingolipids were only present at the later stages (Figure 5). Some minor

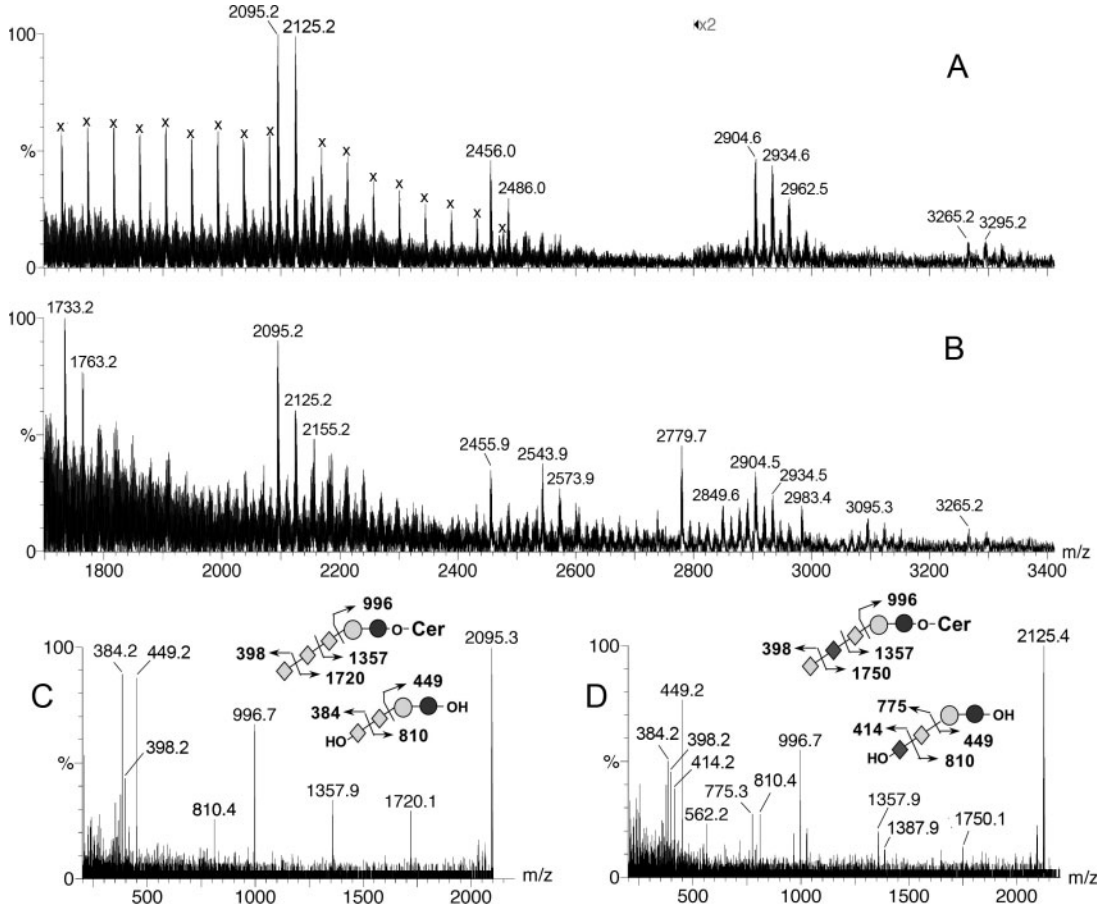

Fig. 5. MALDI-MS and MS/MS analyses of permethylated glycolipids from zebrafish embryos. Oligosialylated glycolipids related to the series at *m/z* 2095/2125 were only apparent in the profiles of samples from 24 h (**A**) and 45–48 h (**B**) after the fertilization. MS profiles of earlier stages (not shown) afforded other weak signals that may be attributed to glycan-containing species which were not further characterized in this study because of low abundance. MS profiles of the 45 and 48 h samples are almost identical, and only the latter was shown in panel **B**. Signals marked with x in panel **A** are contaminants. The two major sodiated molecular ions at *m/z* 2095 and 2125 from panel **A** were selected for MS/MS analyses, and the derived trisialylated sequences were schematically shown in panels **C** and **D**, respectively, along with the major fragmentation pattern observed. All fragment ions are of b and y ions. Signals at *m/z* 2543/2573 in panel **B** are a Hex-HexNAc increment from the trisialylated LacCer at *m/z* 2095/2125. Signals at *m/z* 2779 and 2983 are related by a Hex, but their exact sequences were not established. Symbols used are circle, Glc (dark) and Gal (light); diamond, Neu5Gc (dark) and Neu5Ac (light); Cer, ceramide; OH, exposed hydroxyl group because of primary cleavage.

variations between the profile of 24 h (Figure 5A) and those of 45–48 h (Figure 5B) were noted, but the salient features were mostly conserved. Among the weak signals, a most prominent cluster occurred at *m/z* 2095 and 2125. MALDI-MS/MS sequencing of the parent ion at *m/z* 2095 (Figure 5C) revealed a consecutive loss of Neu5Ac residues. Only the first Neu5Ac loss corresponds to loss of a fully methylated terminal residue, whereas subsequent losses of the other two Neu5Ac correspond to further cleavage of internal Neu5Ac residues, as distinguished by their distinctive residual mass values. In addition to the linear stretch of a $Neu5Ac_3$ sequence thus established, the fragment ions at *m/z* 449 and 810 could be assigned as sodiated $(OH)_1$Hex-Hex-OH and $(OH)_1$Neu5Ac-Hex-Hex-OH, respectively, consistent with a direct attachment of the $Neu5Ac_3$ terminal sequence to a lactosylceramide (LacCer). Based on the *m/z* values of the sodiated molecular ion and the deduced glycosyl sequence, the ceramide moiety could be calculated as corresponding to a d18:1 base with a C16:0 fatty acyl chain or their equivalent permutation thereof.

For the parent ion at *m/z* 2125, MS/MS sequencing (Figure 5D) demonstrated that one of the internal Neu5Ac in the $Neu5Ac_3$-LacCer was replaced by a Neu5Gc. Significantly, only loss of terminal Neu5Ac and not Neu5Gc was observed, giving the y ion at *m/z* 1750. This could be followed by further loss of an internal Neu5Gc (*m/z* 1357) and then an internal Neu5Ac to yield the sodiated $(OH)_1$LacCer ion at *m/z* 997. Additional sodiated ions at *m/z* 449 and 810 indicated an internal -Neu5Ac-Hex_2- sequence, therefore suggesting a unique sequence of Neu5Ac-Neu5Gc-Neu5Ac-LacCer. Alternative arrangement of the $Neu5Ac_2Neu5Gc_1$ sequence on the LacCer was, however, not ruled out nor stage and/or batch variations be investigated due to its low abundance. At one Neu5Ac residue smaller, the molecular ion signals at *m/z* 1733 and 1763 could be tentatively assigned as $Neu5Ac_2$-LacCer and $Neu5Ac_1Neu5Gc_1$LacCer, respectively, whereas signals at *m/z* 2456 and 2486 are consistent with $Neu5Ac_4$-LacCer and $Neu5Ac_3Neu5Gc_1$-LacCer, respectively. Other weak signals that may be tentatively assigned include the clusters

at *m/z* 2904/2934 which correspond to a Hex-HexNAc increment from the tetrasialylated LacCer, and signals at *m/z* 3265/3295 which carried an additional sialic acid. In all cases though, the low abundance amid contaminant noise peaks, compounded by additional heterogeneity which may be contributed by Neu5Ac/Neu5Gc differences and/or the lipid moiety, precluded firm definition of the entire spectra of the glycolipids.

Discussion

A glycomic analysis aims, in general. to define the glycosylation potential of a biological source under investigation. As applied to lower organism such as the zebrafish for which current knowledge on its glycobiology is scarce, a MS approach offers several distinctive advantages. Of prime consideration, the detection and tentative compositional assignment, including *de novo* sequencing, is not dependent on standard references and hence more conducive to identification of novel structures than any other methods. In this context, MALDI-MS mapping coupled with facile CID MS/MS sequencing on the permethyl derivatives is by far the most informative and sensitive analytical strategy although not without its limitations. As a first attempt, we have successfully derived an overall picture of the zebrafish glycome, as presented on both the glycoproteins and the glycolipids, but have excluded analysis on the GAG or the chitin oligosaccharides. Our collective results show that a most striking feature is the diverse oligosialylation pattern which appears to be developmentally regulated.

The high–mannose-type *N*-glycans are the only nonsialylated population of the glycome that occurs at any abundance. Otherwise, both the complex-type *N*-glycans and the *O*-glycans are each predominantly represented by a single unique terminal sequence, monosialylated with either Neu5Ac or Neu5Gc. For the *N*-glycans, the Galβ1-4(Neu5Gc/Neu5Acα2-3)Galβ1-4(Fucα1-3)GlcNAc nonreducing terminal sequence constitutes the antenna of the major bi- and triantennary structures, along with some minor degrees of incomplete sialylation and/or galactosylation on the internal Lewis x epitope. For the *O*-glycans, a core 1-type structure was identified which carries a Fucα1-3GalNAcβ1-4(Neu5Gc/Neu5Acα2-3)Galβ1-3GalNAc sequence. In both cases, α2-3-sialylation was found on an internal β-Gal. The zebrafish sequences can be distinguished from those of more commonly found mammalian type by virtue of either an additional β4-Gal extension on a sialyl Lewis x or an α3-fucosylated β4-GalNAc extension on a sialylated core 1 *O*-glycan. On the other hand, they bear much similarity to other characterized fish glycans (reviewed in Inoue and Inoue, 1997).

The Galβ1-4Galβ1-4GlcNAc motif was first identified on the complex-type free sialoglycans released from the glycophosphoproteins of unfertilized eggs of *Tribolodon hakonensis* and *Oryzia latipes* (Inoue *et al.*, 1989; Iwasaki *et al.*, 1992). α2-3-Neu5Ac monosialylation was found to occur on either the terminal or the internal Gal, giving monosialylated antenna for the predominantly bi- and triantennary structures. Such epitope was also identified on the bulky multiantennary *N*-glycans isolated from cortical alveolus glycoproteins (hyosophorins) of fertilized fish eggs which carry species-specific, highly branched poly-*N*-acetyllactosaminoglycans (Taguchi *et al.*, 1993, 1994, 1995, 1996). Of all the structural variants determined, only those presented by the hyosophorins of medaka fish, *O. latipes*, also contain the fucosylated version of the α2-3-Neu5Ac-sialylated Galβ1-4Galβ1-4GlcNAc motif, identical to that currently identified on the zebrafish *N*-glycans. Interestingly, the Fucα1-3GalNAcβ1 epitope as found on the zebrafish *O*-glycans has also been identified on the *N*-glycans of hysophorin of flounder (Seko *et al.*, 1989), but not sialylated and is attached to the same -3Galβ1-4Galβ1-4GlcNAc sequence. Thus, a Galβ1-4Galβ1-4GlcNAc unit, occurring either at the terminal or as an internal unit, with and without further α2-3Neu5Ac sialylation on the β4-Gal and/or α3-fucosylation on the β4-GlcNAc appears to be a shared feature among the *N*-glycans from several fish eggs characterized to date. A slightly different version with an additional α4-Gal capping instead of sialylation or fucosylation has also been recently identified on the pigeon serum immunoglobulin G (Suzuki *et al.*, 2003), suggesting a possible wider occurrence on nonmammalian vertebrates.

The presence of the Lewis x-type α3-fucosylation is in agreement with the characterization of two zebrafish α,1-3 fucosyltransferases capable of synthesizing Lewis x from lacto-*N*-neotetraose *in vitro* (Kageyama *et al.*, 1999). A recent communication (Natsuka *et al.*, 2005) has further reported the identification of such Lewis x carrying *N*-glycans which appeared from the segmentation period (18 h) onward. The occurrence of the sialylated version of the same structure or other *N*-glycans were, however, not investigated (Natsuka *et al.*, 2005; Takemoto *et al.*, 2005), which probably biased the overall glycomic representation. Our MS-based profiling did not reveal a significant increase in the relative abundance of nonsialylated complex-type *N*-glycans following segmentation. It should, however, be noted that the discrepancy may also arise from different starting materials used because we have attempted to start from total delipidated extracts, whereas Takemoto *et al.* (2005) have treated the lyophilized embryos, free of corion and yolk, directly with hydrazinolysis for *N*-glycan release. Interestingly, they have also identified a significant proportion of biantennary *N*-glycans with the reducing end GlcNAc missing, namely with a trimannosyl $GlcNAc_1$ core instead of the expected di-*N*-acetylchitobiose, and attributed the findings to elevated endo-β-*N*-acetylglucosaminidase activity (Natsuka *et al.*, 2005). Although the activity of peptide:*N*-glycanase or glycoamidase has been convincingly demonstrated in the early embryos of medaka fish (Seko *et al.*, 1991) and elsewhere in other animals, an endoglycosidase F or a chitobiase-type activity has not been previously identified in fish. It is nevertheless conceivable that mammalian-type stepwise action of the lysosomal aspartylglucosaminidase and chitobiase (Michalski *et al.*, 1977; Strecker *et al.*, 1988) could lead to intracellular generation of such free sialylated complex-type *N*-glycans in zebrafish embryos which were not completely removed from our glycoprotein sample preparation because of the use of a low molecular weight cutoff dialysis (3500 Da), coupled with subsequent omission of C18 Sep-Pak step after tryptic digestion to increase the yield of oligosialylated *N*-glycans. Chemical degradation during permethylation

was deemed unlikely as similar structures were also observed with peracetylation. Furthermore, none of the more abundant high–mannose-type *N*-glycans was found to exhibit similar degradation which seems to be restricted to the complex-type subset. The origin of these atypical "*N*-glycans" remains a moot point and may also represent a novel linkage or form of glycosylation merit further investigations.

Despite similarity to previously characterized glycan structures of fish eggs, the zebrafish glycans are unique and novel in many additional aspects. First, the hyosophorin *N*-glycans are bulky, and those of *O. latipes* have been shown to be exclusively pentaantennary, whereas the much simpler bi- and triantennary *N*-glycans from the fish glycophosphoproteins characterized to date do not have fucosylation. Second, whereas sialylated core 1 *O*-glycans with terminal Fucα1-3GalNAc epitope have also been identified in fish eggs (Inoue and Inoue, 1997), none carries the exact sequence as defined here. Oligo- or polysialylation, when occurs, extends from the C6 of the reducing end GalNAc or nonreducing terminal GalNAc. In contrast, the zebrafish Neu5Ac-Neu5Gc or Neu5Gc-Neu5Gc disialyl unit extends from an internal Gal of the *O*-glycans. Further work is needed to confirm the tentatively defined structures for the disialylated core 2-type *O*-glycans which appear to carry one to several units of Lewis x on its 6-arm.

Notably then, we have shown that Neu5Ac and Neu5Gc sialylation were not evenly distributed. Most Neu5Ac is located at the nonreducing terminal position and, mostly, if not, exclusively as monosialylated motif. Neu5Gc occurs as both terminal and internal residues. For the *O*-glycans, our MS/MS data indicated that only the species sialylated with Neu5Gc can be further sialylated with another Neu5Ac or Neu5Gc residue. Likewise, although both Neu5Ac and Neu5Gc monosialylated antennae could be detected for the *N*-glycans, a preliminary oligosialylation analysis with the more sensitive 1,2-diamino-4,5-methylenedioxybenzene high-pressure liquid chromatography (DMB-HPLC) fluorescent detection method indicated that only a Neu5Gcα2-8Neu5Gc-DMB derivative could be additionally derived from the *N*-glycans (Guerardel, unpublished data). These data strongly suggest the occurrence of a certain form of donor and acceptor substrate selectivity in the differential transfer of Neu5Ac and Neu5Gc on glycoprotein-type glycans of zebrafish or a strict intra-/extracellular compartmentalization of sialyltransferase activities.

The concentration of CMP-Neu5Gc in the cytosol has been suggested to play the most important role in regulating the level of Neu5Gc sialylation because neither the CMP-sialic acid antiporter nor the sialyltransferases examined so far seem to exhibit a preference for CMP-Neu5Ac or CMP-Neu5Gc (Higa and Paulson, 1985; Lepers *et al.*, 1989, 1990; Schauer and Kamerling, 1997). In contrast, different donor substrate specificities have been observed for enzymes involved in the elongation of oligo-/polysialylated chains. Thus, whereas rainbow trout polysialyltransferase (polyST) can use both CMP-Neu5Ac and CMP-Neu5Gc as activated sialyl donors, chick brain polyST was shown to not recognize CMP-Neu5Gc (Kitazume *et al.*, 1994; Sevigny *et al.*, 1998). However, to our knowledge, nothing is presently known on the possible specificity of these enzymes toward their acceptor substrates for Neu5Ac/Neu5Gc composition that may explain the absence of polysialyl elongation from Neu5Ac residues in zebrafish *O*-glycans.

Contrary to the glycoproteins, synthesis of oligosialyl sequences in glycolipids did not seem to be affected by the same biosynthetic restrictions. The major sialylated glycolipids detected conform to a family of lactosylceramides extended by up to four sialic acids which can be further elongated by a Hex-HexNAc unit to form either sialylated ganglio-tetraglycosylceramides or sialylated (neo)lactotetraglycosylceramides, with up to five sialic acids. Both series of glycolipids have previously been identified in other fishes (Ando and Yu, 1979; DeGasperi *et al.*, 1987; Nakamura *et al.*, 1997). However, to our knowledge, tetrasialylated lactosylceramides have not been observed previously in any model system. Such a compound, that would be named GQ3 according to used nomenclature, does not fit into accepted ganglioside synthesis pathway model in which GT3 is the biosynthetic precursor of the so called c-series (including GT2, GT1c, GQ1c, and GP1c) and is not further elongated by sialic acids (Freischutz *et al.*, 1995). The sialylated moieties of all observed glycolipids are made up by heterogeneous mixtures of Neu5Ac and Neu5Gc residues in all possible combinations. The presence of polymerized Neu5Ac sequences distinguishes their sialylation pattern from those of *N*- and *O*-glycans. Furthermore, homogeneously Neu5Ac-sialylated glycolipids are the major forms compared with Neu5Gc-containing glycolipids. Altogether, these data demonstrate that although the glycoprotein glycans and glycolipids are both highly sialylated, the biosynthesis of their respective oligosialylated moieties are differently regulated, and the sialylation pattern changes as the embryos develop.

In particular, the disialylated *O*-glycans were exclusively observed in the very first stages of development, before 24 h after the fertilization, whereas, surprisingly, the pattern of oligosialylation in glycolipids seems to follow the opposite trend with the oligosialylated glycolipids being exclusively observed in later developmental stages. The general low yield of the glycolipids relative to the major *N*- and *O*-glycans prevented more definitive structural characterization. We nevertheless could detect glycolipids of even higher degree of sialylation which collectively represent a complete shift in the glycolipid profile from the very early stage that contained a putative range of very different, highly heterogeneous neutral glycolipids. These findings strongly suggest the existence of a very complex, regulated expression pattern of sialylation according to the class of glycoconjugates and developmental stages. By furnishing the structural data pertaining to the glycome of *D. rerio*, our studies reported here provide a solid basis for further functional investigations into the specificity of glyco-related enzymes and, by extension, the role of glycosylation during development. In particular, we believe that *D. rerio* is a very promising model for the study of the fine regulation of sialylation events. Preliminary screening of gene data banks revealed that zebrafish genome not only contains orthologs of all identified human polyST-coding genes, but also several potentially new members of this family (Harduin-Lepers *et al.*, 2005). The structures reported also give the opportunity to identify other novel glycosyltransferase activities

such as the fucosyltransferase activity involved in the synthesis of Fuc(α1-3)GalNAc(β1- motif.

Materials and methods

Sample collection

Zebrafish (*D. rerio*) were maintained at 28°C on a 14-h light/10-h dark cycle. Embryos were incubated at 28°C, and different developmental stages were determined according to the description in the Zebrafish Book (Westerfield, 1995).

Extraction and preparation of glycoconjugates

Embryos were suspended in 200 µL of water and homogenized by sonication on ice. The resulting material was dried and then sequentially extracted three times by chloroform/methanol (2:1, v/v) and three times by chloroform/methanol/water (1:2:0.8, v/v/v). Supernatants from the latter extractions were pooled, dried, and subjected to a mild saponification in 0.1 M sodium hydroxide in methanol at 37°C for 3 h and then evaporated to dryness (Schnaar, 1994). Sample was reconstituted in methanol/water (1:1, v/v) and applied to a C18 Sep-Pak cartridge (Waters, Milford, MA) equilibrated in the same solvent system. After washing with five volumes of methanol/water (1:1, v/v), glycosylceramides were eluted by five volumes of methanol and five volumes of chloroform/methanol (2:1, v/v).

Delipidated pellet from chloroform/methanol/water extraction was resuspended in a solution of 6 M guanidinium chloride and 5 mM ethylenediaminetetraacetic acid (EDTA) in 0.1 M Tris/HCl, pH 8, and agitated for 4 h at 4°C. Dithiothreitol was then added to a final concentration of 20 mM and incubated for 5 h at 37°C, followed by the addition of iodoacetamide to a final concentration of 50 mM and further incubated overnight in the dark at room temperature. Reduced/alkylated sample was dialyzed against water at 4°C for 3 days and lyophilized. The recovered protein samples were then sequentially digested by (L-(tosylamido-2-phenyl) ethyl chloromethyl ketone)-treated trypsin for 5 h and chymotrypsin overnight at 37°C, in 50 mM ammonium bicarbonate buffer, pH 8.4. Crude peptide fraction was separated from hydrophilic components on a C18 Sep-Pak cartridge equilibrated in 5% acetic acid by extensive washing in the same solvent and eluted with a step gradient of 20, 40, and 60% propan-1-ol in 5% acetic acid. Pooled propan-1-ol fraction was dried and subjected to *N*-glycosidase F (Roche, Basel, Switzerland) digestion in 50 mM ammonium bicarbonate buffer, pH 8.4, overnight at 37°C. Alternatively, both the chloroform/methanol/water extraction and the C18 Sep-Pak purification step following tryptic digestion may be omitted to increase the yield of the sialylated *N*-glycans. Omission of the latter step would, however, increase the content of contaminant Hex polymers and possibly other free glycans.

The released *N*-glycans were separated from peptides using the same C18 Sep-Pak procedure as described above. To liberate *O*-glycans, retained peptide fraction from C18 Sep-Pak was submitted to alkaline-reductive elimination in 100 mM NaOH containing 1.0 M sodium borohydride at 37°C for 72 h. The reaction was stopped by addition of Dowex 50 × 8 cation-exchange resin (25–50 mesh, H^+ form) at 4°C until pH 6.5 and, after evaporation to dryness, boric acid was distilled as methyl ester in the presence of methanol. Total material was then submitted to cation-exchange chromatography on a Dowex 50 × 2 column (200–400 mesh, H^+ form) to remove residual peptides.

Chromatographic separation of glycans

The released *N*-glycans were either analyzed directly or after separation into neutral and sialylated fractions on a weak anion exchanger, DEAE Sephadex A-25 column (Amersham, Piscataway, NJ). Samples were dissolved in 20 mM Tris/HCl, pH 8, for loading onto a column equilibrated in the same buffer. Nonbinding neutral glycans were recovered in the washed through fractions, whereas sialylated glycans were eluted in a single fraction by a 0.8 M NaCl solution in 20 mM Tris/HCl, pH 8. Both fractions were desalted by passage through a Bio-Gel P2 column (Bio-rad, Hercules, CA) equilibrated in water.

To remove the contaminating neutral *N*-glycans and to enrich for the sialylated components, the *O*-glycans were dissolved in water and fractionated on a strong anion exchanger Dowex 1 × 2 (200–400 mesh, $HCOO^-$ form) column preequilibrated in water. Neutral glycans were washed off by water, whereas mono- and oligosialylated compounds were recovered by a stepwise elution at 0.1 and 2 M pyridine acetate, pH 5.5, respectively. High salt fractions were desalted by passage through a Bio-Gel P2 column equilibrated in water.

Exoglycosidase digestions

The *N*-glycans were digested with 20 mU of neuraminidase from *Arthrobacter ureafaciens* (Roche) in 100 µL of 50 mM sodium acetate buffer, pH 5.5, at 37°C for 16–18 h. Desialylated *N*-glycans were further treated with 3 mU of β1-4 galactosidase from *Streptococcus pneumoniae* (Calbiochem, Merck, Darmstadt, Germany) in 100 µL of 50 mM sodium acetate buffer, pH 5.5, at 37°C for 12 h, before and after chemical defucosylation by 48% aqueous hydrofluoric acid at 4°C for 48 h.

Chemical derivatization

Monosaccharide compositions were determined by gas chromatography (GC)-MS analysis as either per-heptafluorobutyryl (Zanetta *et al*., 1999) or alditol acetate derivatives. For alditol acetates analysis, glycan samples were hydrolyzed in 4 M trifluoroacetic acid (TFA) for 4 h at 100°C and then reduced with sodium borohydride in 0.05 M NaOH for 4 h. Reduction was stopped by dropwise addition of acetic acid until pH 6 was reached, and borate salts were codistilled by repetitive evaporation in dry methanol. Peracetylation was performed in acetic anhydride at 100°C for 2 h. To determine the chemical nature of sialic acids, intact sialic acids were liberated directly by mild hydrolysis in 0.01 N TFA at 50°C and reacted with a volume of DMB reagent at 50°C for 2 h 30 min (Hara *et al*., 1987). The monomeric DMB-sialic acid derivatives were separated isocratically on a C18 reverse phase (RP) HPLC column (250 × 4.6 mm, 5 micron, Vydac, Hesperia, CA) by

a solvent mixture of acetonitrile/methanol/water (7:9:84) and identified by referring to the elution positions of standard Neu5Ac and Neu5Gc derivatives.

For MALDI-MS analyses, the glycan samples were permethylated using the NaOH/dimethyl sulfoxide slurry method (Ciucanu and Kerek, 1984), as described by Dell *et al*. (1994). The permethyl derivatives were then extracted in chloroform and repeatedly washed with water. GC-MS linkage analysis was performed as described previously (Suzuki *et al*., 2003).

MS analyses of glycans and glycolipids

For MALDI-time-of-flight (MALDI-TOF) MS glycan profiling, the permethyl derivatives in acetonitrile were mixed 1:1 with 2,5-dihydroxybenzoic acid (DHB) matrix (10 mg/mL in acetonitrile), spotted on the target plate, air-dried, and recrystallized on-plate with ethanol whenever necessary. Data acquisition was performed manually on a benchtop M@LDI LR system (Micromass) operated in the reflectron mode. For DHB matrix, the coarse laser energy control was set at high and fine adjusted using the % slider according to sample amount and spectra quality. Laser shots (5 Hz, 10 shots/spectrum) were accumulated until a satisfactory signal to noise ratio was achieved when combined and smoothed. Glycan mass profiling was also performed on a dedicated Q-Tof Ultima MALDI instrument (Micromass, Manchester, UK), in which case the permethylated samples in acetonitrile were mixed 1:1 with α-cyano-4-hydrocinnamic acid matrix (in acetonitrile : 0.1% TFA, 99:1 v:v) for spotting onto the target plate. The nitrogen UV laser (337 nm wavelength) was operated at a repetition rate of 10 Hz under full power (300 μJ/pulse). MS survey data were manually acquired, and the decision to switch over to CID MS/MS acquisition mode for a particular parent ion was made on the fly upon the examination of the summed spectra. Argon was used as the collision gas with a collision energy manually adjusted (between 50 and 200 V) to achieve optimum degree of fragmentation for the parent ions under investigation.

NMR analyses

Before NMR spectroscopic analysis, sample was repeatedly exchanged in $^{2}H_2O$ (99.97% purity, Euriso-top, CEA Saclay, France) with intermediate freeze drying and then dissolved in 250 μL of Me_2SO-d_6 (Euriso-top). Chemical shifts were expressed in p.p.m. downfield from the signal of the methyl group of Me_2SO-d_6 ($\delta^{1}H$/tetramethyl-silyl ester [TMS] = 2.52 p.p.m., $\delta^{13}C$/TMS = 40.98 p.p.m. at 343 K). The sample was analyzed in 200 × 5 mm BMS-005-B Shigemi tubes on a Bruker ASX-400 spectrometer (Centre d'Analyses RMN, Villeneuve d'Ascq, France) ^{1}H, 400.33 MHz; ^{13}C, 100.66 MHz equipped with a double resonance (^{1}H/X) Broad Band Inverse z-gradient probe head. All NMR data were recorded without sample spinning.

The one-dimensional proton ^{1}H spectrum was measured using a 90° tipping angle for the pulse and 1.5 s as a recycle delay between each of 32 acquisitions of 2.4 s. The spectral width of 4006 Hz was collected in 16384 complex data points. Two-dimensional homonuclear (^{1}H-1H) spectra (COSY and TOCSY) were measured using standard Bruker pulse programs. Rotating frame Overhauser enhancement spectroscopy (ROESY) spectra were acquired with 400 ms mixing times and acquired in the States mode, according to Bax and Davis (1985). Moreover, the two-dimensional TOCSY spectrum was recorded using a MLEV-17 mixing sequence of 120 ms. The spin lock field strength corresponded to a 90° pulse width of 62 μs. The spectral width was 2402 Hz in both dimensions. About 256 spectra of 4096 data points with 32 scans per t1 increment were recorded giving a spectral resolution of 0.6 Hz/pt in F2 and ~9.4 Hz/pt in F1. Heteronuclear HMQC ^{1}H-13C spectrum was obtained using with standard Bruker inv4tp pulse sequence. The spectral width was 2403 Hz in F2 and 12080 Hz in ^{13}C dimension, giving a spectral resolution of 0.6 Hz/pt and 47.2 Hz/pt, respectively.

Acknowledgments

This work was supported by Academia Sinica, Taiwan (to K.H.K. and C.J.H.), and an Academia Sinica Postdoctoral Fellowship (to Y.G.). Mass spectrometry analyses were performed at the Core Facilities for Proteomics Research located at the Institute of Biological Chemistry, Academia Sinica, supported by a Taiwan National Science Council grant (NSC 93-3112-B-001-010-Y) and the Academia Sinica.

Abbreviations

CID, collision-induced dissociation; DMB, methylenedioxybenzene; Hex, hexose; HexNAc, *N*-acetyl hexosamine; HexNAcitol, reduced *N*-acetyl hexosaminitol; LacCer, lactosylceramide; MALDI, matrix-assisted laser-desorption ionization; MS, mass spectrometry; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; polyST, polysialyltransferase; TFA, trifluoroacetic acid; TMS, tetramethyl-silyl ester; TOCSY, total correlation spectroscopy.

References

Ando, S. and Yu, R.K. (1979) Isolation and characterization of two isomers of brain tetrasialogangliosides. *J. Biol. Chem.*, **254**, 12224–12229.

Bakkers, J., Semino, C.E., Stroband, H., Kijne, J.W., Robbins, P.W., and Spaink, H.P. (1997) An important developmental role for oligosaccharides during early embryogenesis of cyprinid fish. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 7982–7986.

Bax, A.D. and Davis, J. (1985) Practical aspect of two-dimentional transvers NOE spectroscopy. *J. Magn. Reson.*, **63**, 207–213.

Becker, C.G. and Becker, T. (2002) Repellent guidance of regenerating optic axons by chondroitin sulfate glycosaminoglycans in zebrafish. *J. Neurosci.*, **22**, 842–853.

Bernhardt, R.R. and Schachner, M. (2000) Chondroitin sulfates affect the formation of the segmental motor nerves in zebrafish embryos. *Dev. Biol.*, **221**, 206–219.

Bink, R.J., Habuchi, H., Lele, Z., Dolk, E., Joore, J., Rauch, G.J., Geisler, R., Wilson, S.W., den Hertog, J., Kimata, K., and Zivkovic, D. (2003) Heparan sulfate 6-O-sulfotransferase is essential for muscle development in zebrafish. *J. Biol. Chem.*, **278**, 31118–31127.

Ciucanu, I. and Kerek, F. (1984) A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.*, **131**, 209–217.

Coppin, A., Maes, E., and Strecker, G. (2002) Species-specificity of amphibia carbohydrate chains: the Bufo viridis case study. *Carbohydr. Res.*, **337**, 121–132.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzymes. Available at: http://afmb.cnrs-mrs.fr/CAZY/. Accessed December 12, 2005.

DeGasperi, R., Koerner, T.A., Quarles, R.H., Ilyas, A.A., Ishikawa, Y., Li, S.C., and Li, Y.T. (1987) Isolation and characterization of gangliosides with hybrid neolacto-ganglio-type sugar chains. *J. Biol. Chem.*, **262**, 17149–17155.

Dell, A., Reason, A.J., Khoo, K.H., Panico, M., McDowell, R.A., and Morris, H.R. (1994) Mass spectrometry of carbohydrate-containing biopolymers. *Methods Enzymol.*, **230**, 108–132.

Dupuy, F., Petit, J.M., Mollicone, R., Oriol, R., Julien, R., and Maftah, A. (1999) A single amino acid in the hypervariable stem domain of vertebrate alpha1,3/1,4-fucosyltransferases determines the type 1/type 2 transfer. Characterization of acceptor substrate specificity of the Lewis enzyme by site-directed mutagenesis. *J. Biol. Chem.*, **274**, 12257–12262.

Freischutz, B., Saito, M., Rahmann, H., and Yu, R.K. (1995) Characterization of sialyltransferase-IV activity and its involvement in the c-pathway of brain ganglioside metabolism. *J. Neurochem.*, **64**, 385–393.

Guerardel, Y., Kol, O., Maes, E., Lefebvre, T., Boilly, B., Davril, M., and Strecker, G. (2000) O-glycan variability of egg-jelly mucins from *Xenopus laevis*: characterization of four phenotypes that differ by the terminal glycosylation of their mucins. *Biochem. J.*, **352**, 449–463.

Guerardel, Y., Balanzino, L., Maes, E., Leroy, Y., Coddeville, B., Oriol, R., and Strecker, G. (2001) The nematode Caenorhabditis elegans synthesizes unusual O-linked glycans: identification of glucose-substituted mucin-type O-glycans and short chondroitin-like oligosaccharides. *Biochem. J.*, **357**, 167–182.

Guerardel, Y., Petit, D., Madigou, T., Guillet, B., Maes, E., Maftah, A., Boujard, D., Strecker, G., and Kol, O. (2003) Identification of the blood group Lewis (a) determinant in the oviducal mucins of *Xenopus tropicalis*. *FEBS. Lett.*, **554**, 330–336.

Hara, S., Takemori, Y., Yamaguchi, M., Nakamura, M., and Ohkura, Y. (1987) Fluorometric high-performance liquid chromatography of N-acetyl- and N-glycolylneuraminic acids and its application to their microdetermination in human and animal sera, glycoproteins, and glycolipids. *Anal. Biochem.*, **164**, 138–145.

Harduin-Lepers, A., Mollicone, R., Delannoy, P., and Oriol, R. (2005) The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*, **15**, 805–817.

Haslam, S.M. and Dell, A. (2003) Hallmarks of Caenorhabditis elegans N-glycosylation: complexity and controversy. *Biochimie*, **85**, 25–32.

Herkt, F., Parente, J.P., Leroy, Y., Fournet, B., Blanchard, D., Cartron, J.P., van Halbeek, H., and Vliegenthart, J.F. (1985) Structure determination of oligosaccharides isolated from Cad erythrocyte membranes by permethylation analysis and 500-MHZ 1H-NMR spectroscopy. *Eur. J. Biochem.*, **146**, 125–129.

Higa, H.H. and Paulson, J.C. (1985) Sialylation of glycoprotein oligosaccharides with N-acetyl-, N-glycolyl-, and N-O-diacetylneuraminic acids. *J. Biol. Chem.*, **260**, 8838–8849.

Inoue, S. and Inoue, Y. (1997) Fish glycoproteins. In Montreuil, J., Vliegenthart, J.F.G., and Schachter, H. (eds), *Glycoproteins II*. Elsevier Sceince B.V., Amsterdam, The Netherlands, pp. 143–161.

Inoue, S., Iwasaki, M., Ishii, K., Kitajima, K., and Inoue, Y. (1989) Isolation and structures of glycoprotein-derived free sialooligosaccharides from the unfertilized eggs of Tribolodon hakonensis, a dace. Intracellular accumulation of a novel class of biantennary disialooligosaccharides. *J. Biol. Chem.*, **264**, 18520–18526.

Iwasaki, M., Seko, A., Kitajima, K., Inoue, Y., and Inoue, S. (1992) Fish egg glycophosphoproteins have species-specific N-linked glycan units previously found in a storage pool of free glycan chains. *J. Biol. Chem.*, **267**, 24287–24296.

Kageyama, N., Natsuka, S., and Hase, S. (1999) Molecular cloning and characterization of two zebrafish alpha (1,3) fucosyltransferase genes developmentally regulated in embryogenesis. *J. Biochem. (Tokyo)*, **125**, 838–845.

Kamerling, J.P. and Vliegenthart, J.F.G. (1992) High-resolution ^{1}H-nuclear magnetic resonance spectroscopy of oligosaccharide-alditols released from mucin-type O-glycoproteins. In Berliner, L.J. and Reuben, J. (eds), *Carbohydrates and Nucleic Acids*. Plenum Press, New York, pp. 1–194.

Kitazume, S., Kitajima, K., Inoue, S., Troy, F.A., 2nd, Cho, J.W., Lennarz, W.J., and Inoue, Y. (1994) Identification of polysialic acid-containing glycoprotein in the jelly coat of sea urchin eggs. Occurrence of a novel type of polysialic acid structure. *J. Biol. Chem.*, **269**, 22712–22718.

Lepers, A., Shaw, L., Cacan, R., Schauer, R., Montreuil, J., and Verbert, A. (1989) Transport of CMP-N-glycoloylneuraminic acid into mouse liver Golgi vesicles. *FEBS. Lett.*, **250**, 245–250.

Lepers, A., Shaw, L., Schneckenburger, P., Cacan, R., Verbert, A., and Schauer, R. (1990) A study on the regulation of N-glycoloylneuraminic acid biosynthesis and utilization in rat and mouse liver. *Eur. J. Biochem.*, **193**, 715–723.

Marx, M., Rutishauser, U., and Bastmeyer, M. (2001) Dual function of polysialic acid during zebrafish central nervous system development. *Development*, **128**, 4949–4958.

Michalski, J.C., Strecker, G., and Fournet, B. (1977) Structures of sialyloligosaccharides excreted in the urine of a patient with mucolipidosis I. *FEBS. Lett.*, **79**, 101–104.

Mourad, R., Morelle, W., Neveu, A., and Strecker, G. (2001) Diversity of O-linked glycosylation patterns between species. Characterization of 25 carbohydrate chains from oviducal mucins of Rana ridibunda. *Eur. J. Biochem.*, **268**, 1990–2003.

Nakamura, K., Tamai, Y., and Kasama, T. (1997) Gangliosides of dogfish (*Squalus acanthias*) brain. *Neurochem. Int.*, **30**, 593–604.

Natsuka, S., Takemoto, T., Moriguchi, K., Aoki, T., Ito, N., and Hase, S. (2005) Lewis X-oligosaccharides appeared at segmentation period of zebrafish embryo. *Glycoconj. J.*, **22**, 338. Abstract for Glyco XVIII.

Schauer, R. and Kamerling, J.P. (1997) Chemistry, biochemistry and biology of sialic acids. In Montreuil, J., Vliegenthart, J.F.G., and Schachter, H., eds. *Glycoproteins II*. Elsevier Science B.V., Amsterdam, The Netherlands, pp. 243–402.

Schnaar, R.L. (1994) Isolation of glycosphingolipids. *Methods Enzymol.*, **230**, 348–370.

Seko, A., Kitajima, K., Iwasaki, M., Inoue, S., and Inoue, Y. (1989) Structural studies of fertilization-associated carbohydrate-rich glycoproteins (hyosophorin) isolated from the fertilized and unfertilized eggs of flounder, Paralichthys olivaceus. Presence of a novel penta-antennary N-linked glycan chain in the tandem repeating glycopeptide unit of hyosophorin. *J. Biol. Chem.*, **264**, 15922–15929.

Seko, A., Kitajima, K., Inoue, S., and Inoue, Y. (1991) Identification of free glycan chain liberated by de-N-glycosylation of the cortical alveolar glycopolyprotein (hyosophorin) during early embryogenesis of the Medaka fish, *Oryzias latipes*. *Biochem. Biophys. Res. Commun.*, **180**, 1165–1171.

Semino, C.E. and Allende, M.L. (2000) Chitin oligosaccharides as candidate patterning agents in zebrafish embryogenesis. *Int. J. Dev. Biol.*, **44**, 183–193.

Semino, C.E., Allende, M.L., Bakkers, J., Spaink, H.P., and Robbins, P.P. (1998) Expression of Rhizobium chitin oligosaccharide fucosyltransferase in zebrafish embryos disrupts normal development. *Ann. N. Y. Acad. Sci.*, **842**, 49–54.

Sevigny, M.B., Ye, J., Kitazume-Kawaguchi, S., and Troy, F.A., 2nd. (1998) Developmental expression and characterization of the alpha2,8-polysialyltransferase activity in embryonic chick brain. *Glycobiology*, **8**, 857–867.

Snaar-Jagalska, B.E., Krens, S.F., Robina, I., Wang, L.X., and Spaink, H.P. (2003) Specific activation of ERK pathways by chitin oligosaccharides in embryonic zebrafish cell lines. *Glycobiology*, **13**, 725–732.

Strecker, G., Michalski, J.C., and Montreuil, J. (1988) Lysosomal catabolic pathway of N-glycosylprotein glycans. *Biochimie*, **70**, 1505–1510.

Suzuki, N., Khoo, K.H., Chen, C.M., Chen, H.C., and Lee, Y.C. (2003) N-glycan structures of pigeon IgG: a major serum glycoprotein containing Galalpha1-4 Gal termini. *J. Biol. Chem.*, **278**, 46293–46306.

Taguchi, T., Seko, A., Kitajima, K., Inoue, S., Iwamatsu, T., Khoo, K.H., Morris, H.R., Dell, A., and Inoue, Y. (1993) Structural studies of a novel type of tetraantennary sialoglycan unit in a carbohydrate-rich glycopeptide isolated from the fertilized eggs of Indian Medaka fish, Oryzias melastigma. *J. Biol. Chem.*, **268**, 2353–2362.

Taguchi, T., Seko, A., Kitajima, K., Muto, Y., Inoue, S., Khoo, K.H., Morris, H.R., Dell, A., and Inoue, Y. (1994) Structural studies of a

novel type of pentaantennary large glycan unit in the fertilization-associated carbohydrate-rich glycopeptide isolated from the fertilized eggs of *Oryzias latipes*. *J. Biol. Chem.*, **269**, 8762–8771.

Taguchi, T., Kitajima, K., Muto, Y., Inoue, S., Khoo, K.H., Morris, H.R., Dell, A., Wallace, R.A., Selman, K., and Inoue, Y. (1995) A precise structural analysis of a fertilization-associated carbohydrate-rich glycopeptide isolated from the fertilized eggs of euryhaline killi fish (*Fundulus heteroclitus*). Novel penta-antennary N-glycan chains with a bisecting N-acetylglucosaminyl residue. *Glycobiology*, **5**, 611–624.

Taguchi, T., Iwasaki, M., Muto, Y., Kitajima, K., Inoue, S., Khoo, K.H., Morris, H.R., Dell, A., and Inoue, Y. (1996) Occurrence and structural analysis of highly sulfated multiantennary N-linked glycan chains derived from a fertilization-associated carbohydrate-rich glycoprotein in unfertilized eggs of *Tribolodon hakonensis*. *Eur. J. Biochem.*, **238**, 357–367.

Takemoto, T., Natsuka, S., Nakakita, S., and Hase, S. (2005) Expression of complex-type N-glycans in developmental periods of zebrafish embryo. *Glycoconj. J.*, **22**, 21–26.

Westerfield, M. (ed.) (1995) *The Zebrafish Book*, 3rd edn. University of Oregan Press, Eugene, OR.

Zanetta, J.P., Timmerman, P., and Leroy, Y. (1999) Gas-liquid chromatography of the heptafluorobutyrate derivatives of the O-methyl-glycosides on capillary columns: a method for the quantitative determination of the monosaccharide composition of glycoproteins and glycolipids. *Glycobiology*, **9**, 255–266.