ANALYSIS OF PROTEIN-LINKED GLYCOSYLATION IN A SPERM-SOMATIC CELL ADHESION SYSTEM

Mark Sutton-Smith, Nyet Kui Wong, Kay-Hooi Khoo, Sz-Wei Wu, Shin-Yi Yu, Manish S. Patankar, Richard Easton, Frank A. Lattanzio, Howard R. Morris, Anne Dell, and Gary F. Clark

1Division of Molecular Biosciences, Imperial College London, SW7 2AZ, UK
2Present address: School of Science and Technology, University Malaysia Sabah, Locked Bag 2073, 88999 Kota Kinabalu, Sabah, Malaysia
3Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan
4Department of Obstetrics and Gynecology, University of Wisconsin Medical School, Madison, WI 53792-6177
5Present address: M-SCAN Ltd., 3 Millars Business Centre, Fishponds Close, Wokingham, Berkshire RG41 2TZ, UK
6Department of Physiological Sciences, Eastern Virginia Medical School, 700 Olney Road, Norfolk, VA 23501
7M-SCAN Mass Spectrometry Research and Training Centre, Silwood Park, Ascot SL5 7PZ, UK
8Department of Obstetrics, Gynecology and Women’s Health, Division of Reproductive and Perinatal Research, School of Medicine, University of Missouri-Columbia, Columbia, MO 65212

Keywords: sperm-egg binding; mouse, zona pellucida, N-glycans, O-glycans

Running Title: Glycosylation in a Somatic Cell Adhesion Model

9To whom correspondence should be addressed: a.dell@ic.ac.uk; clarkgf@health.missouri.edu

© The Author 2007. Published by Oxford University Press. All rights reserved.

For permissions, please e-mail: journals.permissions@oxfordjournals.org
Murine sperm initiate fertilization by binding to the specialized extracellular matrix of their complementary eggs, known as the zona pellucida. Based on data reported in this study, mouse sperm also bind to rabbit erythrocytes with higher affinity than they do to murine eggs. This unusual interaction between a germ cell and a somatic cell ("sperm-somatic cell adhesion system") is also carbohydrate dependent based on its sensitivity to mild periodate oxidation. To determine what types of carbohydrate sequences could be involved in this interaction, the protein-linked oligosaccharides of rabbit erythrocytes were sequenced using novel matrix assisted laser desorption ionization time-of-flight tandem mass spectrometry methods that enabled the analysis of individual components up to \( m/z \) 9000. The N-glycans are primarily complex biantennary and triantennary types terminated with Gal\( \alpha_1 \)-3Gal sequences. The majority of these oligosaccharides also possess one antenna consisting of a highly branched polylactosamine type sequence that is also associated with many glycosphingolipids that coat rabbit erythrocytes. These erythrocytes also express Core 1 and Core 2 O-glycans terminated primarily with Gal\( \alpha_1 \)-3Gal sequences and to a lesser extent sialic acid. These results confirm that rabbit erythrocytes and mouse eggs present very different types of carbohydrate sequences on their surfaces. However, oligosaccharides terminated with \( \beta_1 \)-6 linked N-acetyllactosamine or its \( \alpha_1 \)-3 galactosylated analogue are expressed on both the mouse zona pellucida and this somatic cell type. The far more abundant presentation of such sequences on rabbit erythrocytes compared to murine eggs could explain why mouse sperm display such exceptional affinity for this somatic cell type.
Sexual reproduction is initiated in metazoans when sperm undergo initial species-specific binding to their eggs. There is now substantial understanding about the molecular basis of this interaction in many lower species (Mengerink and Vacquier, 2001; Ohlendieck and Lennarz, 1996; Vacquier, 1998). However, there exists a great deal of controversy about the precise receptors and ligands that mediate initial gamete binding in mammals (Jungnickel et al., 2003). Nowhere is this debate more evident than in the mouse, the predominant model employed to study mammalian fertilization (reviewed, Clark and Dell, 2006).

Murine eggs are coated with a specialized extracellular matrix known as the mZP. The mZP is composed of three major glycoproteins, designated mZP1, mZP2, and mZP3. A classical model suggests that initial murine sperm-egg binding requires the recognition of O-glycans positioned at two specific serine residues (Ser-332, Ser-334) of mZP3 by egg binding proteins located on the sperm plasma membrane (Wassarman et al., 2001). Another paradigm suggests that a supramolecular complex of the three major mZP glycoproteins forms a matrix that enables sperm binding (Rankin et al., 2003; Rankin et al., 1998). In another recently proposed model, a combination of two different egg binding proteins (a β1-4 galactosyltransferase and a murine homologue of the porcine p47 protein designated SED1) mediates initial murine gamete binding by interacting with both mZP2 and mZP3 (Ensslin and Shur, 2003). In addition, Shur and colleagues recently proposed that a 250 kDa basic glycoprotein associates with the mZP during ovulation and acts as an additional mZP3 independent ligand for murine sperm binding (Rodeheffer and Shur, 2003). However, there is an extremely active debate about these different adhesion systems in mediating initial sperm-egg binding (Jungnickel et al., 2003). Thus far, no model has been validated at the molecular level.
One of the most unusual cell-cell interactions that occurs in vitro is the extremely tight binding of murine sperm to rabbit erythrocytes (Clark et al., 1996). Like initial murine gamete binding, the molecular basis for this interaction, first observed over two decades ago (Yamagata et al., 1983), remains an enigma. The implication of this association is that the rabbit erythrocytes express ligands that are either very similar or identical to those coating the murine zona pellucida. However, establishing this relationship requires intimate knowledge about the carbohydrate sequences coating both the mZP and rabbit erythrocytes. Until recently, such a comparison could not be made because MS strategies were either not sufficiently sensitive or incapable of dealing with the structural complexity.

Because of the efforts of two investigative groups, the sequences of both the N- and the O-glycans coating the mZP and mZP3 have now been characterized (Dell et al., 2003; Easton et al., 2000; Noguchi and Nakano, 1993). However, rabbit erythrocytes are coated with oligosaccharides linked to both glycosphingolipids and glycoproteins. The glycosphingolipids associated with rabbit erythrocytes have previously been characterized (Dabrowski et al., 1988; Dabrowski and Hanfland, 1982; Dabrowski et al., 1984; Egge et al., 1985; Hanfland et al., 1981; Hanfland et al., 1988; Honma et al., 1981). However, only very recently have the MS methods been developed that could unambiguously define the sequences of the complex polylactosamine type N-glycans presented by rabbit erythrocytes. This structural analysis is presented in this report.

Results

Characterization of murine sperm binding to rabbit erythrocytes

Capacitated murine sperm were incubated with erythrocytes from several mammalian species under optimal conditions for sperm binding to eggs. There was no interaction observed with erythrocytes obtained from many other mammalian species (human, mouse, rat, goat, sheep,
cow, horse, dog, and cat) (data not shown). However, murine sperm underwent immediate and very robust binding to rabbit erythrocytes, as reported earlier (Yamagata et al., 1983). As shown in Fig. 1 (panel A), sperm would often form aggregates with the smaller rabbit erythrocytes. These aggregates were stable in vitro for up to 3 h. Previous electron microscopy studies indicate that there is no induction of the acrosome reaction during this incubation period (Clark et al., 1996).

**Analysis of carbohydrate dependence and binding affinity**

Rabbit erythrocytes were initially subjected to mild periodate oxidation followed by reduction. For controls, cells were treated in exactly the same manner except that ethylene glycol was added immediately at the initiation of the reaction to inactivate the oxidizing agent. As shown in Fig. 1 (panel B), rabbit erythrocytes in the control reaction bind to murine sperm as well as the native reaction (panel A). However, virtually all sperm binding was eliminated when these erythrocytes were oxidized in the absence of ethylene glycol (Panel C). This result suggests that this interaction between rabbit erythrocytes and murine sperm is carbohydrate dependent.

Additional studies were performed to determine if this interaction relied completely upon terminal α-linked or subterminal β1-4 linked Gal residues. Rabbit erythrocytes were digested with coffee bean α-galactosidase under conditions routinely employed for this enzyme, except that the saline content was adjusted to prevent hypotonic lysis. Loss of terminal α1-3 linked galactose residues was monitored using FITC labeled *Griffonia simplicifolia*-I lectin (Ozgur et al., 1998). There was >90% loss in fluorescence compared to controls incubated under the same conditions in the absence of enzyme. However, there was no discernable difference in the ability of these α-galactosidase treated erythrocytes to bind to mouse sperm compared to control erythrocytes (data not shown). Studies were undertaken to remove subterminal β-linked Gal
residues from these degalactosylated erythrocytes using commercially available $\beta$-galactosidases. However, both intact and $\alpha$-galactosidase treated erythrocytes were extremely sensitive to lysis in isotonic acidic buffers below pH 6.0 obligatory for $\beta$-galactosidase activity. These $\alpha$-galactosidase treated erythrocytes were also sensitive to lysis under conditions employed for mild periodate oxidation. In summary, the results of both chemical and enzymatic modifications suggest that, though an intact terminal monosaccharide is absolutely required for binding, oligosaccharides capped with either $\alpha_{1-3}$ or $\beta_{1-4}$ linked Gal can mediate binding to murine sperm.

The adhesive strength of this interaction was also investigated. Competitive sperm-egg binding assays are usually conducted by incubating capacitated sperm with ten eggs and 3-5 mouse embryos under optimal conditions for sperm binding, usually for 15 min (Johnston et al., 1998). The mixture is then micropipetted until no more than 2-5 sperm remain bound to the embryos. The remaining sperm bound to the eggs are considered to be tightly bound. They are fixed and then counted manually (Johnston et al., 1998). However, it is possible to dislodge even these “bound sperm” from eggs by continued cycles of micropipetting. We therefore performed the same binding analysis in parallel with murine sperm bound to rabbit erythrocytes. When subjected to sufficient cycles of micropipetting that completely dislodge sperm from murine eggs, the number of sperm bound to at least one erythrocyte did not change. These studies indicate that the affinity of murine sperm for rabbit erythrocytes is both highly species-specific and exceeds natural murine sperm-egg binding.

**Structural Analysis Strategy**

The overall structural strategy employed to characterize rabbit erythrocyte glycosylation is summarized in Scheme 1. N-Glycans were released from reduced and carboxymethylated
trypsinized detergent extracts by PNGase F digestion and were separated from peptides and O-linked glycopeptides by passing the digest through a Sep-Pak C18 cartridge. The void volume fraction containing the free N-glycans was analyzed by MALDI-TOF-MS, CAD-MS/MS and linkage analysis after permethylation and Sep-Pak purification. Data from these experiments were complemented by sequential exoglycosidase digestions which were similarly analyzed after permethylation. O-glycans were released from the peptide/glycopeptide fraction by reductive elimination and were analyzed by MALDI-TOF-MS and CAD-MS/MS after permethylation.

**MALDI-TOF Screening Defines N-Glycan Compositions**

MALDI-TOF profiling of the permethylated N-glycans afforded spectra rich in \([M+Na]^+\) molecular ion signals up to \(m/z\) 9000. A minor population of high mannose structures, constituting less than 5\% of the sample, were observed at \(m/z\) 1580, 1784, 1988, 2102 and 2396 corresponding to Man5 to Man9 respectively (data not shown), but the vast majority of glycans were found to be of the complex type (Fig. 2). The latter afforded a distinctive series of signal clusters which are clearly related by a unique mass increment of 1102 u, corresponding to a Hex3HexNAc2 difference (Fig. 2). Two prominent oxonium type fragment ions were detected at \(m/z\) 668 and 1770 (data not shown), corresponding to Hex2HexNAc+ and Hex3HexNAc2HexNAc+ respectively. Thus the non-reducing termini of the majority of the N-glycans detected can be inferred to be terminated with a Hex-HexNAc epitope, which was identified as Galα1-3Galβ1-4GlcNAc by further experimentation (see below). Interestingly, this terminal sequence is not carried on a simple linear Galβ1-4GlcNAc (LacNAc) repeat, since non-reducing terminal fragment ions corresponding to Hex2HexNAc-(Hex-HexNAc)n+ were not afforded. Instead, the repeating unit is apparently defined by a Hex3HexNAc2 increment, consistent with an α-Gal-capped branched LacNAc moiety: -[(Gal-Gal-GlcNAc)Gal-GlcNAc]-.
CAD-MS/MS Defines Sequences and Branching of N-glycans

CAD-MS/MS experiments were carried out using both MALDI-Q/TOF and MALDI-TOF-TOF instrumentation in order to optimize the quality and quantity of fragment ion data. The Q/TOF instrumentation has an upper m/z limit of about 4000 while the TOF-TOF instrument has no such restrictions. Collisional energies are different for the two instruments so fragmentation pathways may differ in the two experiments. We found that the disadvantage of not being able to analyze high mass components on the Q/TOF was compensated for by a richer variety of fragment ions afforded by the smaller glycans when compared with TOF-TOF data. Thus, MALDI-Q/TOF MS/MS analyses of the two major molecular ions in the series, namely m/z 2652 and 3755, clearly demonstrated that both are core fucosylated biantennary complex type structures (Fig. 3). The former afforded direct losses of two terminal Hex2HexNAc moieties to give a sodiated trimannosyl core y ion at m/z 1317, carrying two free OH groups (Fig. 3, panel A). The latter afforded a direct loss of both terminal Hex2HexNAc and Hex3HexNAc3, while the same trimannosyl core y ion at m/z 1317 was produced after losing both antenna (Fig. 3, panel B). The sodiated b ion at m/z 690 corresponding to terminal Hex2HexNAc could be detected in the MS/MS of both parents but m/z 3755 additionally afforded a sodiated b ion at m/z 1793, corresponding to the aforementioned Hex3HexNAc3, or (Hex3HexNAc)3Hex-Hex-HexNAc, as depicted in Fig. 3. In comparison, MS/MS on the sodiated parent ion at m/z 3306 gave only the b ion at m/z 690, complemented by a series of y ions corresponding to losses of 1 to 3 terminal Hex2HexNAc epitopes (Fig. 3, panel C). The triantennary nature of this structure is thus defined by the trimannosyl core ion at m/z 1303 which carries a total of three free OH groups. Further loss of the Man on the 3-arm gave the ion at m/z 1113 with only two free OH groups and is therefore indicative of a 2,4-substituted Man on the 3-arm.
The bi- and triantennary structures which give the sodiated molecular ions at m/z 2652, 3755 and 3306 (Fig. 2) define the characteristics of rabbit erythrocyte N-glycans. The most abundant complex type structure at m/z 2652 corresponds to a biantennary, core-fucosylated structure that is substituted with a Hex$_2$-HexNAc epitope on both antennae. Considerable heterogeneity exists, as represented by (i) incomplete core fucosylation (m/z 2478); (ii) incomplete $\alpha$-galactosylation (m/z 2448, 2244); and (iii) replacing one or both terminal $\alpha$-Gal with NeuAc or NeuGc (m/z 2809, 2839; 2966, 2996, 3026). The corresponding change in the terminal fragment ions afforded by these structures has been verified by MS/MS analyses (data not shown). A full substitution of an additional antenna gives the triantennary structure at m/z 3306 which is itself accompanied by similar heterogeneity including incomplete $\alpha$-galactosylation (minus 204 u), incomplete core fucosylation (minus 174 u) and NeuAc/NeuGc sialylation instead of $\alpha$-galactosylation (+157 u or +187 u for NeuAc and NeuGc respectively). The full complexity of this heterogeneity is apparent from the labeled peaks in the first segment of Fig. 2.

As the molecular size increases, the minor components, especially those with sialylation, become less apparent. Nevertheless, the two series, as represented by [m/z 3755 + (Hex$_3$HexNAc$_2$)$_n$], and [m/z 3306 + (Hex$_3$HexNAc$_2$)$_n$] remain distinctive, each accompanied by one to two degrees of incomplete galactosylation. The first of these series is about three times more abundant than the second.

Each of the molecular ions detected in the first series, i.e. [m/z 3755 + (Hex$_3$HexNAc$_2$)$_n$], was further subjected to MALDI-TOF-TOF MS/MS (Fig. 4) which allowed selection of singly charged parent ions as large as m/z 8166 for direct CAD MS/MS sequencing. A distinguishing feature of the MS/MS fragmentation pattern afforded by MALDI-TOF-TOF is that mostly single cleavage events are registered. At lower collision cell energy (1 kV) and particularly when using
α-CHCA as matrix, the dominant fragment ions produced can be further restricted to b and y ions. Thus, as the parent ions increase in size, the Hex$_3$HexNAc$_2$ increment is reflected by the detection of the complementary sets of b and y ions due to single cleavages. The sodiated b ions at m/z 690, 1793, 2896, 3999 correspond to Hex$_2$HexNAc-[(Hex$_2$HexNAc)Hex-HexNAc]$_n$$^+$ where n = 0, 1, 2, and 3 respectively. From the other end, y ions derived from direct losses of terminal Hex$_2$HexNAc-[(Hex$_2$HexNAc)Hex-HexNAc]$_n$ were registered, fully supportive of a branched polyLacNAc sequence extending from one arm of the biantennary N-glycans. A common y ion was detected at m/z 1985, corresponding to loss of the entire branched polyLacNAc antennae but retaining the terminal Hex$_2$HexNAc- epitope on the other arm.

This primary y ion (m/z 1985) was however not detected for the second series, namely those represented by [m/z 3306 + (Hex$_3$HexNAc$_2$)$_n$] (Fig. 3, panel C and Fig. 5). Instead, both parent ions at m/z 4408 and 5510 afforded a sodiated y ion at m/z 2639 which is another Hex$_2$HexNAc increment from m/z 1985. In conjunction with data obtained from MALDI-Q/TOF MS/MS analysis (Fig. 3, panel C), the TOF/TOF MS/MS data is consistent with this second series being predominantly triantennary structures. The structure represented by the parent ion at m/z 4408 can be assigned as carrying two Hex$_2$HexNAc termini with a third one extended by the branched Hex$_2$HexNAc$_3$ epitope, as illustrated in Fig. 5 (panel A). For the parent at m/z 5511, at least two major isomeric structures exist. The first is homologous to that of m/z 4408 except that the branched polyLacNAc sequence carried on one of the antenna is further extended by a Hex$_3$HexNAc$_2$ unit, giving an antennae of the composition Hex$_8$HexNAc$_5$. A direct loss of this terminal epitope yielded the y ion at m/z 2639. The second major isomeric structure can be assigned as carrying two Hex$_4$HexNAc$_3$ and one Hex$_2$HexNAc epitope on its antenna, supported by the presence of the y ion at m/z 3741 (Fig. 5, panel B). Higher in the series, the general pattern
of substitution on predominantly triantennary structures are observed but the possible isomeric combination increases.

**Exoglycosidase Digestions Define Antennae Branching and Anomeric Stereochemistry**

To obtain information on terminal sequences, antennae branching and anomeric stereochemistry, the N-glycans were sequentially digested with α- and β-galactosidase, and β-N-acetylhexosaminidase. Aliquots were taken after each digestion, permethylated, and examined by MALDI-TOF after Sep-Pak purification (Fig. 6, panels A, B, and C). As shown in the annotations on Fig. 6 (panel A), α-galactosidase digestion truncated the structures observed in Fig. 2 by the predicted number of hexose residues assuming that each Hex2HexNAc epitope carries a terminal α-Gal residue. Similarly, further digestion with β-galactosidase followed by β-N-acetylhexosaminidase gave the mass shifts expected for the branching patterns suggested by the MS/MS experiments (see annotations on Fig. 6, panels B and C, respectively). The truncated glycans produced from the above three exoglycosidase digestions were then digested with a mixture of α-mannosidases (see Experimental Procedures) to address the issue of which arm carries the polylactosaminoglycan antenna in the biantennary glycans. The aim of this experiment was to remove the exposed mannose arising from digestion of the short antenna and to establish from the MALDI profiling that this elimination had been achieved. If mannose removal was successful, linkage analysis could then be exploited to establish whether the remaining antenna is linked at the 6- or the 3-position of the β-mannose of the core. Not unexpectedly, bearing in mind the presence of the polylactosamine antenna on the second arm, efficient removal of the mannose residue was very difficult. Fortunately, after exhaustive digestion, we successfully obtained a Sep-Pak fraction after permethylation in which the diagnostic glycans were of
significant abundance (see annotations of m/z 1581 and 2030 in Fig. 6, panel D). Linkage data from this sample are reported below.

**Linkage Analysis Confirms Branching Patterns**

Linkage analysis data for the N-glycans, and for the products of exoglycosidase digestions, are shown in Table 1. These data are consistent with the MALDI-based evidence that the major components are biantennary complex structures, with triantennary glycans being significantly less abundant and tetraantennary structures being largely absent. The data also support the other main conclusions of the MALDI experiments, namely that the glycans are mostly core fucosylated, and have highly branched polylactosamine structures whose non-reducing sequences are predominantly terminated with Galα1-3Gal. The key pieces of evidence supporting these conclusions are the following: (i) 2-Man is the most abundant linked mannose and is approximately twice as abundant as 3,6-Man, while 2,4-Man is relatively minor and 2,6-Man is present only at trace levels; (ii) 3-Gal is very abundant in the intact glycans and is dramatically reduced after digestion with α-galactosidase; (iii) 3,6-Gal is very abundant in the intact glycans and is dramatically reduced after sequential digestion with α-galactosidase, β-galactosidase and β-N-acetylhexosaminidase; (iv) 4,6-GlcNAc supports core fucosylation. The MS/MS experiments described earlier showed that the polylactosamine structures in the biantennary glycans are mostly confined to a single antenna (Fig. 4); exhaustive exoglycosidase digestion successfully removed a significant fraction of the short antenna from these glycans (Fig. 6, panel D). The linkage data from this experiment suggest that the short antenna is most likely attached to the 3-arm of the biantennary glycan because 6-Man but not 3-Man is detected after its removal (Table 1). Finally, the presence of a minor amount of 6-linked Gal, which is
retained throughout the successive exoglycosidase digestions, suggests that the minor sialylated glycans observed in the MALDI experiments probably carry 6-linked sialic acid.

N-glycan Structures— Taking into consideration the MS, CAD-MS/MS, linkage and exoglycosidase data, we conclude that the major N-glycans on the rabbit erythrocytes have the structures shown in Fig. 7

Characterization of O-Glycans

Reductively eliminated O-glycans were permethylated and analyzed by MALDI-TOF and selected components were sequenced by CAD-ES-MS/MS. The molecular ions observed in the MALDI spectrum (Fig. 8) have compositions consistent with both Core 1 (m/z 534, 779, 895, 1256 and 1316) and Core 2 (m/z 983, 1187, 1549 and 1579) glycans. The majority of these are capped with either sialic acid (NeuAc and NeuGc) and/or α-Gal. The location of these residues was established by MS/MS which readily distinguishes the Core 1 arm from the Core 2 arm (data not shown) (Wong et al., 2003). The proposed structures for the O-glycans are shown in the cartoons in Fig. 8.

Discussion

Previous studies confirm that only acrosome intact murine sperm bind to rabbit erythrocytes (Clark et al., 1996), indicating that the interaction is between the plasma membranes of the two cell types. In addition, it is evident that this association does not induce the acrosome reaction (Clark et al., 1996), unlike natural gamete binding. The current results imply that the affinity of this sperm-somatic cell adhesion system exceeds native mouse sperm-egg binding. Because murine sperm-egg binding is hypothesized to be carbohydrate dependent, this study was undertaken to determine the precise degree of overlap in the protein-linked glycosylation of the mZP and rabbit erythrocytes.
A major finding arising from the analysis of the N-glycans associated with rabbit erythrocytes is their similarity to the previously determined polyglycosylceramides from the same source (Dabrowski et al., 1988; Dabrowski and Hanfland, 1982; Dabrowski et al., 1984; Egge et al., 1985; Hanfland et al., 1981; Hanfland et al., 1988; Honma et al., 1981). Both are characterized by α3-Gal-capped, extended highly branched polyLacNAc chains which can be described as \( \text{Gal}\alpha_1-3\text{Gal}\beta_1-\{4\text{GlcNAc}\beta_1-3(\text{Gal}\alpha_1-3\text{Gal}\beta_1-4\text{GlcNAc}\beta_1-6)\text{Gal}\beta_1-\}_n-4\text{GlcNAc}\beta_1-4\text{Gal}\beta_1-4\text{Glc-Cer} \). In the case of the polyglycosylceramides, \( n \) up to 7 has been studied in detail by NMR (Dabrowski et al., 1988) while FAB-MS analysis has been applied to the permethyl derivatives as large as carrying \( n = 4 \) repeating units (Hanfland et al., 1988), with an \( m/z \) value for the monosodiated molecular ion just above 6000.

Several recent technical advances enabled the present analysis to be performed. In particular, higher sensitivity for both MS and MS/MS is now afforded by the advent of instruments with MALDI sources and TOF/TOF tandem analyzers which allow high efficiency MS/MS. The polyLacNAc chains on the N-glycans are carried on a trimannosyl core instead of the ceramide. Compared with the glycosphingolipids, the N-glycans are much less abundant, more heterogeneous and more difficult to isolate. While FAB-MS sequencing relied mostly on direct fragmentation of purified components, true MS/MS can now be effected on MS-isolated parent ions from the analysis of mixtures at a sensitivity not previously feasible. Using the same formulation, the permethyl derivatives of N-glycan structures carrying a polyLacNAc chain of \( n = 0 \) up to 5 on one antennae were detected and successfully sequenced. MALDI-MS/MS on the sodiated parent ion readily affords both series of sodiated B and Y ions while higher collision energies can further induce ring cleavages to allow linkage specific assignments. This modern approach represents a natural extension of the capability afforded by FAB-MS sequencing which
relied mostly on the non-reducing terminal oxonium type ions produced at each HexNAc site along a polyLacNAc chain. We demonstrate here the first such application of MALDI-TOF/TOF MS/MS sequencing on a biologically derived sample up to \( m/z \) 9000.

In conjunction with sequential exoglycosidase digestion, and coupling each step to MS and conventional linkage analysis, we have succeeded in delineating the heterogeneity associated with an array of large polyLacNAc-containing N-glycans of the rabbit erythrocytes. In most instances the structural assignments were further corroborated by the low energy CAD MS/MS profile as obtained on the Q/TOF and the high energy CAD MS/MS (up to 3kV) on a TOF/TOF.

Taken together, the salient features of the N-glycans, apart from carrying the characteristic branched polyLacNAc chains, can be noted as: (i) comprising a small amount of hybrid and high mannose type glycans; (ii) predominantly biantennary complex type, with much lower amounts of triantennary analogues and almost complete absence of tetraantennary structures; (iii) a small subset of the antennae are terminated with NeuAc/NeuGcα2-3Galβ1-4GlcNAc and Galβ1-4GlcNAc sequences instead of the more common Galα1-3Galβ1-4GlcNAc sequences; (iv) the terminal LacNAc unit on each of the antennae can be further extended, branched and capped. However, in the case of biantennary glycans, there seems to be a preferential extension only on one antenna while the other is maintained as simple, short α-Gal capped LacNAc sequence. The issue of which arm of the trimannosyl core constitutes the preferred site of elongation remains equivocal at this stage. The high energy CAD MS/MS consistently produced a concerted double cleavage D ion which would indicate the presence of the short sequence on the 6-arm (data not shown). On the other hand, sequential enzyme digestion followed by linkage analysis indicated that the 3-arm of at least a portion of the N-glycans could be trimmed off and thus implicating a non-extended LacNAc sequence on it. This ambiguity notwithstanding, our
work represents the best effort to date to rigorously define the large N-glycans carrying polyLacNAc chains by direct MS/MS sequencing of intact molecular ions.

It is apparent that there are several significant differences between the N-glycosylation of rabbit erythrocytes and the mZP (Dell et al., 2003; Easton et al., 2000; Noguchi and Nakano, 1993). Previous analyses of the mZP N-glycans reveals the following: (i) about 5-10% of the total mZP oligosaccharides are high mannose type sequences; (ii) the majority of the N-glycans include biantennary, triantennary and tetraantennary complex type N-glycans; (iii) the antennae linked to the core mannose residues of complex type N-glycans include NeuAc/NeuGcα2-3Galβ1-4GlcNAc, Galβ1-4GlcNAc, terminal GlcNAc, and the Sd antigen (NeuAc/NeuGcα2-3[GalNAcβ1-4]Galβ1-4GlcNAc); (iv) terminal Galα1-3Galβ1-4GlcNAc sequences are also present but expressed at very low levels; and (v) linear polyLacNAc sequences could be detected at low abundance but not localized to any structure as yet. It should be noted in particular that such linear extension hardly exists in the rabbit erythrocyte N-glycans, as demonstrated in this study. All polyLacNAc sequences are found to be branched and capped by α-Gal.

The specific glycoprotein implicated in murine sperm-egg binding is mZP3 (Wassarman et al., 2001). Virtually all the O-glycans presented by the mZP are coupled to this glycoprotein (Dell et al., 2003). These oligosaccharides are primarily Core 2 sequences, with lower amounts of Core 1 sequences. These Core 2 sequences are primarily terminated with the same non-reducing ends as the N-glycans, except for the presence of lacdiNAc sequences on the Core 2 structures (Dell et al., 2003). In comparison from the results shown in Fig. 8, the sequence of the O-glycans linked to rabbit erythrocytes are summarized as follows: (i) both Core 1 and Core 2 oligosaccharides are present; (ii) sialic acid either in the form of NeuGe or NeuAc is linked to Core 1 O-glycans via α2-3 or α2-6 linkages; (iii) Core 2 O-glycans are terminated with either
β1-6 linked LacNAc or its α-galactosylated analogue; and (iv) some α-galactosylated Core 2 O-glycans express NeuAc/NeuGcα2-3Gal sequences on their Core 1 arm.

Thus the repertoire of O-glycans expressed on rabbit erythrocytes is also quite different from those associated with the mZP. However, previous studies suggest that rabbit erythrocytes are poorly O-glycosylated, likely due to their relatively low expression of glycophorin, the major carrier of O-glycans in mammalian erythrocytes (Fukuda et al., 1987). Thus O-glycans linked to rabbit erythrocytes likely play a very marginal role in mediating sperm binding.

There are two reasonable questions that arise from these structural studies: (i) how can a germ cell and somatic cell presenting very different carbohydrate sequences on their surfaces mediate sperm binding; and (ii) why do murine sperm bind with higher affinity to rabbit erythrocytes than they do to their homologous eggs?

In a previous study, Renkonen and associates synthesized the artificial oligosaccharide constructs shown in Table 2 for testing in the competitive murine sperm-egg binding assay (Seppo et al., 1995). Construct S1 is terminated with two identical branched polylactosamine type sequences capped at each end with α1-3 linked Gal (dashed boxes). Based on previous studies (16-22) and this report, this exact sequence also terminates most of the N-glycans and glycosphingolipids expressed on rabbit erythrocytes. Construct S1 inhibits murine sperm-egg binding by 70-75% at 4 μM concentration (Litscher et al., 1995). Construct S2 lacking terminal α1-3 linked Gal is as inhibitory as Construct S1 at the same concentration, indicating that terminal α1-3 linked Gal are not required for binding, a result consistent with a previous gene inactivation studies in mice (Liu et al., 1997). Construct S3, an analogue of Construct S2 lacking terminal β-linked Gal is not inhibitory, indicating a requirement for terminal β1-4 linked Gal (Litscher et al., 1995). However, an analogue of Construct S1 bearing only a single branched
polylactosamine type sequence (Construct S4) was poorly inhibitory, as was Construct S5, a monovalent analogue of Construct S2.

These results suggest that the egg binding proteins on the surface of murine sperm mediate binding by reacting with the branched polylactosamine sequence at the terminal ends of the N-glycans and glycosphingolipids presented by rabbit erythrocytes. It is evident from previous studies and more recent analysis that the mZP completely lacks branched polylactosamine sequences of this type (Dell et al., 2003; Easton et al., 2000). mZP also does not express β1-3 linked LacNAc sequences or its α-galactosylated analogue present at the terminal ends of Constructs S2 and S1, respectively (Dell et al., 2003; Easton et al., 2000). However, this matrix does present β1-6 linked LacNAc sequences (Construct S2, dashed box). These terminal sequences are specifically associated with core 2 O-glycans and the tri- and tetraantennary N-glycans within this matrix. It is therefore significant in this context to note that mZP3, the putative sperm receptor glycoprotein, carries virtually all of the core 2 O-glycans associated with the mZP (Dell et al., 2003) and also displays enhanced expression of tri- and tetraantennary N-glycans (Noguchi and Nakano, 1993) Therefore mZP3 is highly enriched in β1-6 linked LacNAc sequences compared to mZP1 and mZP2.

It is also obvious that rabbit erythrocytes present more of these potential recognition sequences than the mZP. Most of the N-glycans on rabbit erythrocytes are terminated with only a single branched polylactosamine type sequence. This restriction also applies to all polyglycosylceramides associated with this cell type (Dabrowski et al., 1984). Only a small subset of the larger biantennary N-glycans presents two branched polylactosamine sequences at their terminal ends. However, one branch is at the end of an extended polylactosamine chain, while the other is closer to the trimannosyl core (Fig. 7). Thus the spatial arrangement of the two
terminal branched polylactosamine sequences on this subset of N-glycans is not the same as Construct S1 (Table 2). Construct S4 is poorly inhibitory in the competitive murine sperm-egg binding assay (Litscher et al., 1995), suggesting that both protein and lipid-linked glycans from rabbit erythrocytes expressing only one terminal branched polylactosamine sequence could also be poorly inhibitory. Thus this somatic cell binding to murine sperm could occur primarily via multivalent interactions with oligosaccharides bearing one or more terminal branched polylactosamine sequences. The other possibility is that this interaction is mediated only by this small subset of N-glycans bearing two terminal branched polylactosamine sequences. However, testing of this latter hypothesis would require the selective isolation of this subset from a very complex mixture of N-glycans from this cell type, a very technically challenging task using available methods.

The current findings are consistent with the hypothesis that the majority of murine sperm-egg binding is carbohydrate dependent. However, carbohydrate inhibitors (Litscher et al., 1995) or glycosidase treatments (Mori et al., 1997) maximally inhibit murine sperm binding by no more than 70-75%. In addition, only 79% of the human sperm binding sites on the human zona pellucida are sensitive to mild periodate oxidation under the same conditions employed in this study (Ozgur et al., 1998). These findings collectively suggest that a carbohydrate independent system for sperm-egg binding coexists with the primary carbohydrate dependent pathway in both humans and mice.

The binding of murine sperm to rabbit erythrocytes does not induce the acrosome reaction (Clark et al., 1996). Data obtained in other laboratories also indicate that carbohydrate inhibitors of binding do not induce the acrosome reaction in murine sperm. It is quite possible that the binding of mZP3 to the sperm signaling complex that induces the acrosome reaction could
require a protein-protein interaction that could also be responsible for this residual carbohydrate independent binding. However, the possibility cannot be ruled out at this time that the acrosome reaction is also initiated by redundant protein-carbohydrate and protein-protein interactions.

Genetic manipulations have proven very useful for testing the requirements for carbohydrate mediation of murine sperm-egg binding. In an early study, terminal α-linked Gal was implicated in murine sperm-egg binding (Bleil and Wassarman, 1988). However, mice deficient in the α-galactosyltransferase that adds terminal Galα1-3Gal sequences to mZP glycoproteins remain fertile (Thall et al., 1995) and also bind equivalent numbers of sperm (Liu et al., 1997). It is apparent from the current study that the enzymatic removal of terminal α-linked Gal from rabbit erythrocytes did not affect their interaction with murine sperm, a result consistent with the findings of these previous knockout studies.

Genetic manipulations have proven very useful for testing the requirements for carbohydrate mediation of murine sperm-egg binding. Wassarman and colleagues initially suggested that murine sperm-egg binding required terminal α-linked Gal residues (Bleil and Wassarman, 1988). However, mice deficient in the α-galactosyltransferase that adds terminal Galα1-3Gal sequences to mZP glycoproteins remain fertile (Thall et al., 1995) and also bind equivalent numbers of sperm (Liu et al., 1997). It is apparent from the current study that the removal of terminal α1-3 linked Gal sequences from rabbit erythrocytes does not affect their interaction with murine sperm, a result consistent with these knockout experiments.

Recent studies involving the inactivation of glycosyltransferase genes also suggest that compensatory protein-protein interactions exist in the murine gamete binding system. An obligatory enzyme required for the synthesis of Core 2 O-glycans is the β1-6 N-acetylglucosaminyltransferase (Core 2 enzyme) that adds GlcNAc to the GalNAc in the Core 1
sequence (Galβ1-3GalNAc-Ser/Thr). There are three known Core 2 enzyme isoforms, designated I-III, but only isoform I is expressed in the ovary (Ellies et al., 1998). Mice with a genetic deletion of isoform I retain their fertility (Ellies et al., 1998).

Deletion of the N-acetylglucosaminyltransferase that adds β1-6 linked GlcNAc to the trisaccharide core (Mgat5 or GnT V) generates mutant mice that are deficient in tri- and tetraantennary N-glycans bearing β1-6 linked LacNAc sequence and extensions (Demetriou et al., 2001). However, these mice remain fertile. Stanley and coworkers recently inactivated the gene for N-acetylglucosaminyltransferase I (Mgat1) in oocytes using the ZP3Cre recombinase transgene system (Shi et al., 2004). These eggs are coated with mZP glycoproteins that completely lack complex and hybrid type N-glycans, yet are fertilized, yield embryos that implant, and generate heterozygotes that develop to birth (Shi et al., 2004). Thus this study is also consistent with the hypothesis that protein-protein interactions occur during mammalian sperm-egg binding (Ozgur et al., 1998).

Thus the results of carbohydrate inhibition, specific periodate oxidation and gene deletion studies indicate that because of potential protein-protein interactions, fertility may not be completely eliminated by inactivating glycosyltransferases. More investigation will be required to determine the potential carbohydrate dependent and independent mechanisms that mediate murine sperm-egg binding. In this context, the sperm-somatic cell adhesion system may provide useful insights into this physically straightforward but physiologically complex interaction involving the murine gametes.

Materials and Methods

Materials
Erythrocytes were obtained from New Zealand White rabbits. Whole blood samples from other animal species (mouse, rat, goat, sheep, cow, horse, dog, and cat) were obtained from Pel-Freeze (Rogers, Arkansas) or local slaughterhouses. Human blood samples were obtained from the local American Red Cross laboratories. All chemicals employed in this study were obtained from Sigma unless otherwise noted.

**Sperm-erythrocyte binding assays**

Sperm were isolated from the epididymis of male CD-1 mice using the procedure developed by Bleil and Wassarman (Bleil and Wassarman, 1983). Sperm were incubated for 1 h in medium 199 supplemented with 2 mg/ml BSA and 30 mg/ml sodium pyruvate (referred to as M199-M) to insure capacitation. They were then diluted in this medium to a final concentration of 1 X 10^6/ml.

Erythrocytes were isolated from whole blood samples using an established procedure (Yamagata et al., 1983). Briefly, whole blood collected in the presence of citrate was centrifuged in microfuge tubes (5000 x g, 5 min). Leukocytes in the buffy coat were aspirated off. The red blood cell pellet was resuspended in sterile phosphate buffered saline (pH 7.4) and centrifuged again (5000 X g, 5 min). After removal of the supernatant, this washing procedure of resuspension, centrifugation and supernatant removal was repeated twice more. The erythrocytes were then diluted to 1-2% (v/v) in sterile PBS (137 mM NaCl, 10 mM phosphate, 2.7 mM KCl, 1 mM CaCl₂, 1 mM MgCl₂, pH 7.4) and stored at 4°C. All erythrocyte preparations were used in binding studies within 48 h of their isolation.

The sperm suspension (20 μl) and erythrocyte suspension (2 μl) were added sequentially to a glass slide and gently mixed. The slides were covered with cover slips and immediately viewed with an Olympus BX41 microscope with a DP70 camera using an UplanFl40x objective under
bright field. Camera settings were a shutter speed of 1/700-1/400 with ISO 200. For each sperm/erythrocyte sample mixture at least 30 images of 1040x1360 pixels were rapidly collected and stored in TIF files for later analysis.

Procedures to test carbohydrate dependence

For studies involving mild periodate oxidation, the erythrocytes were centrifuged and resuspended in (i) sterile PBS (control); (ii) sterile PBS containing 10 mM sodium $m$-periodate (oxidized); or (iii) PBS containing 10 mM sodium $m$-periodate and 50 mM ethylene glycol (oxidation-blocked). All samples were incubated for 1 h at 25°C in the dark. The oxidized and oxidation-blocked erythrocyte samples were centrifuged for 5 min x 5000g. The supernatants were removed and resuspended in PBS containing 50 mM ethylene glycol. After 10 min at 25°C, each sample was centrifuged at 5 min x 5000g. The cell pellet was washed twice more in sterile PBS by repeated centrifugation and then resuspended in sterile PBS containing 50 mM sodium borohydride. After 1 h at 25°C, the erythrocyte subjected to each condition were washed by repeated centrifugation three times in sterile PBS and immediately analyzed for sperm binding using the standard assay.

For enzymatic digestions, rabbit erythrocytes were centrifuged at 5000 x g for 5 min. After removing the supernatant, 50 μl of packed cells were resuspended in 1 ml of 50 mM ammonium acetate buffer, pH 6.0 containing 100 mM NaCl and 1 unit/ml coffee bean $\alpha$-galactosidase (EC 3.2.1.22 [EC]) (Sigma). The cells were digested at 23°C room temperature for 16 h. This temperature and amount of time was found to be optimal for avoiding lysis and maximizing the loss of terminal $\alpha$1-3 linked Gal. To monitor the efficacy of this modification, both treated and control erythrocytes were incubated with 1 μg/ml of fluorescein isothiocyanate labeled *Griffonia simplicifolia*-I lectin. The loss of lectin binding sites was then estimated using an established
procedure (Ozgur et al., 1998). These erythrocytes were also incubated in isotonic buffers for β-galactosidases derived from jack beans (Glyko) and bovine testes (Sigma) (50 mM ammonium acetate, pH 5.5 containing 100 mM NaCl) to determine if underlying β-linked Gal residues could also be removed.

**Analysis of the adhesive strength of the sperm-somatic cell adhesion system**

Sperm were isolated from male mice as previously described (Bleil and Wassarman, 1983). Mouse eggs were obtained via an established procedure (Bleil and Wassarman, 1980). Sperm were mixed with eggs for 15 min exactly as described for the competitive murine sperm-egg binding assay (Johnston et al., 1998). The mixture was micropipetted until all the sperm were completely dislodged from the eggs as assessed by microscopic examination. Samples containing sperm and rabbit erythrocytes mixed as described in the standard binding assay were subjected to exactly the same amount of micropipetting employed to completely disrupt sperm-egg binding. The number of murine sperm bound to erythrocytes with and without micropipetting was then quantitated.

**Processing of erythrocyte glycans**

Erythrocytes from 10 ml of packed membranes were sonicated in extraction buffer (25mM Tris, 150 mM, 5 mM, EDTA, and 1% CHAPS at pH 7.4) and then dialyzed as described previously (Sutton-Smith and Dell, 2006). Reduction and carboxymethylation were carried out using an established procedure (Dell et al., 1993). Erythrocyte preparations were reduced in 50 mM Tris-HCl buffer, pH 8.5, containing dithiothreitol (2 mg/ml). Reduction was performed under a nitrogen atmosphere at 37 °C for 1 h. Carboxymethylation was carried out in iodoacetic acid (5-fold molar excess over dithiothreitol), and the reaction was allowed to proceed under a nitrogen atmosphere at 37 °C for 1 h. Carboxymethylation was terminated by dialysis against 4 x 2.5 l of
50 mM ammonium bicarbonate, pH 8.5, at 4 °C for 48 h. After dialysis, the sample was lyophilized. Reduced/carboxymethylated material was incubated with trypsin (EC 3.4.21.4 [EC]; Sigma) at a 50:1 ratio (w/w) in 50 mM ammonium bicarbonate, pH 8.5, for 5 h at 37 °C. The digestion was terminated by placing in boiling water for 3 min, followed by lyophilization. PNGase F digestion of the tryptic glycopeptides was carried out in 50 mM ammonium bicarbonate, pH 8.5, for 16 h at 37 °C with 3 units of enzyme. The reaction was terminated by lyophilization, and the released N-glycans (1.2 mg) were separated from peptides and O-glycopeptides by Sep-Pak C18 (Waters Corp.) as described (Dell et al., 1993). O-glycans (100 µg) were released by reductive elimination in 400 µl of sodium borohydride (1 mg/ml in 0.05 M sodium hydroxide) at 45 °C for 16 h. This reaction was terminated by dropwise addition of glacial acetic acid followed by Dowex chromatography and borate removal (Dell et al., 1993).

**Sequential Exoglycosidase Digestion**

The released glycans were incubated with the following enzymes and conditions: α-galactosidase (green coffee bean, EC 3.2.1.22 [EC]) 0.8 U in 100 µl of 50 mM ammonium acetate buffer, pH 6.0; β-galactosidase (bovine testes, EC 3.2.1.23 [EC]) 20 mU in 100 µl of 50 mM ammonium acetate buffer, pH 4.6; β1-2,3,4,6-N-Acetylglucosaminidase (Streptococcus pneumoniae, Recombinant, E. coli, EC 3.2.1.30 [EC]; Calbiochem) 0.2 U in 100 µl of 50 mM ammonium acetate buffer, pH 4.6; α1-6-mannosidase (Xanthomonas sp., EC 3.2.1.24 [EC]; Calbiochem), 0.5 U in 100 µl of 50 mM ammonium acetate buffer, pH 4.6; 10 mU in 100 µl of 50 mM ammonium formate buffer, pH 4.6; and β-N-acetylhexosaminidase (jack bean, EC 3.2.1.30 [EC]; Calbiochem), 0.2 U in 100 µl of 50 mM ammonium formate buffer, pH 4.6. All of the enzyme digestions were carried out at 37 °C for 48 h with a fresh aliquot of enzyme added after 24 h. Each digestion was terminated by boiling for 3 min before lyophilization.
Derivatizations

Permethylation was performed using the sodium hydroxide procedure as described previously (Dell et al., 1993). Briefly, sodium hydroxide pellets were crushed with dimethyl sulfoxide to form a slurry. An aliquot of this slurry was added to dried glycans along with 1 ml of methyl iodide. The reaction was terminated by the addition of water, and permethylated glycans were recovered by chloroform extraction. The chloroform layer was washed several times with water to remove any impurities. Partially methylated alditol acetates were prepared from permethylated samples for GC-MS linkage analysis as described (Dell et al., 1993). Briefly, the permethylated glycans were hydrolyzed with 2 M trifluoroacetic acid for 2 h at 121 °C, reduced with 10 mg/ml sodium borodeuteride in 2 M aqueous ammonium hydroxide at room temperature for 2 h, and then acetylated with acetic anhydride at 100 °C for 1 h.

MALDI-MS Profiling

MALDI data were acquired using either a Perseptive Biosystems Voyager-DETM STR instrument or a 4700 Proteomics Analyzer (Applied Biosystems, Framingham, MA), both operated in reflectron positive ion mode with delayed extraction. Permethylated samples were dissolved in 10 µl of methanol or acetonitrile, and 1 µl of the dissolved sample was premixed with 1 µl of matrix (2,5-dihydrobenzoic acid (DHB), 10 mg/ml in water) before loading onto a metal plate.

CAD-MALDI-MS/MS Analysis

These studies were performed on either a Q-TOF Ultima™ MALDI (Micromass, Manchester, UK) or a 4700 Proteomics Analyzer (Applied Biosystems, Framingham, MA), both operated in reflectron positive ion mode. For acquisition on the Q-TOF, the permethylated samples in acetonitrile were mixed 1:1 with α-cyano-4-hydrocinnamic acid (CHCA) matrix (5 mg/ml in
50% acetonitrile/0.1% TFA) for spotting onto the target plate. MS survey and CAD MS/MS data were manually acquired. Argon was used as the collision gas with a collision energy manually adjusted (between 100–200 V) to achieve an optimum degree of fragmentation for the parent ions under investigation. For MALDI CAD-MS/MS on the TOF/TOF, both DHB and CHCA matrices were used in conjunction with setting the potential difference between the source acceleration voltage and the collision cell at either 1 kV or 2 kV to obtain different degrees and patterns of fragmentation. The indicated collision cell pressure was normally increased from 3.0x10^{-8} torr (no collision gas) to 5.0x10^{-7} torr (argon).

*CAD-ES-MS/MS Analysis*

CAD-ES-MS/MS spectra were acquired using a Q-STAR (Applied Biosystems, Framingham, MA) mass spectrometer. The permethylated glycans were dissolved in methanol before loading into a spray capillary coated with a thin layer of gold/palladium, inner diameter 2 µl (Proxeon, Odense, Denmark). A potential of 1.5 kV was applied to a nanoflow tip to produce a flow rate of 10–30 nl/min. The drying gas used was N₂ and the collision gas was argon, with the collision gas pressure maintained at 10^{-4} millibar. Collision energies varied depending on the size of the carbohydrate, typically between 30 and 90 eV.

*GC-MS Linkage Analysis*

Partially methylated alditol acetates were analyzed using a PerkinElmer Clarus 500 instrument, fitted with a RTX-5 column (30 m x 0.25-mm internal diameter, Restek Corp.) The sample was dissolved in hexanes and injected onto the column at 65 °C. The column was maintained at this temperature for 1 min and then heated to 290 °C at a rate of 8 °C per min.
Acknowledgements

The authors would like to thank Drs. Kathy Sharpe-Timms and Breton Barrier for editing this manuscript. This study was supported by the Wellcome Trust and the Biotechnology and Biological Sciences Research Council (BBSRC) (A.D., H.R.M.), a BBSRC Professorial Fellowship (A.D.), a Malaysian Government Scholarship (N.K.W.), Academia Sinica, Taiwan (K.H.K.), and a Mission Enhancement Program in Reproductive Biology and Medicine (C8783) funded by the state of Missouri (G.F.C.). The MALDI-MS/MS analyses were performed at the National Core Facilities for Proteomics Research funded by a Taiwan National Science Council grant (NSC 93-3112-B-001-010-Y).

Abbreviations

CAD-ES-MS, collisionally activated nanoelectrospray tandem mass spectrometry; GC, gas chromatography; LacNAc, Galβ1-4GlcNAc; MALDI-TOF, matrix assisted laser desorption/ionization time of flight; MS, mass spectrometry; MS/MS, tandem mass spectrometry; mZP, murine zona pellucida.

Supplementary Material

Figures 3-8 are also presented in color in the supplementary materials online at http://glycob.oxfordjournals.org/.
References


FIGURE LEGENDS

**Fig. 1.** Evidence that sperm binding to rabbit erythrocytes is carbohydrate dependent. Murine sperm bind to rabbit erythrocytes to form aggregates (panel A). Rabbit erythrocytes were subjected to mild periodate oxidation as described in the text in the presence (panel B) or absence (panel C) of excess ethylene glycol to determine carbohydrate dependence.

**Fig. 2.** MALDI-MS analysis of the permethylated N-glycans. The major [M+Na]+ molecular ion signals can be grouped into clusters, which are related to the next one up in series by a Hex3HexNAc2 (1102 u) increment, shown here as vertically aligned spectral segments. Within each signal cluster, the major biantennary and triantennary complex type structures, e.g. m/z 2652 and 3306 for the first segment, is accompanied by several satellite signals corresponding to incomplete fucosylation, galactosylation, or NeuAc/NeuGe sialylation replacing terminal ±-galactosylation, as described in text.

**Fig. 3.** MALDI Q/TOF MS/MS analyses of the major N-glycans. The fragment ions produced are mostly of the b and y ion series derived from the preferred cleavages at HexNAc residues, as schematically illustrated on the deduced structures for the parents at m/z 2652 (A), 3755 (B) and 3306 (C). Two series of y ions can in general be detected, i.e. one derived from the parent ions and the other from the primary b ions resulting from cleavage at the chitobiose core. Key to symbols for sugars: open circles, Gal; closed circles, Man; closed squares, GlcNAc; open triangles, Fuc.

**Fig. 4.** MALDI TOF/TOF MS/MS analyses of the major N-glycans. All sodiated molecular ions in the series defined by \([m/z \text{ 2652} + (\text{Hex}_3\text{HexNAc}_2)n]\) were selected for MS/MS. Major
fragment ions afforded are of the sodiated y and b ion series which collectively establish the branched polyLacNAc sequence, as illustrated schematically. Key to symbols is the same as in Fig. 3.

**Fig. 5.** MALDI TOF/TOF MS/MS analyses of the major triantennary N-glycans. Major fragment ions afforded are of the sodiated y and b ion series. One major isomeric structure for (A) and two major ones for (B), as deduced by the MS/MS analyses, are depicted. The b ions are similar to those detected for the other series (Fig. 4). The TOF/TOF MS/MS spectra of the smallest parent ion in the series, i.e. m/z 3306, is similar to that afforded by Q/TOF (Fig. 3, panel C) but with less multiple cleavage ions. These data collectively establish the branched polylactosamine sequence, as illustrated schematically. Key to symbols for sugars is the same as shown in Fig. 3.

**Fig. 6.** MALDI-MS profiles of permethylated N-glycans from rabbit erythrocytes after successive exoglycosidase digestions. Prior to permethylation, N-glycans from rabbit erythrocytes were digested with α-galactosidase (Panel A), β-galactosidase (Panel B), β1-2,3,4,6-N-Acetylglucosaminidase (Panel C) and α-mannosidases (Panel D). Key to symbols for sugars is the same as shown in Fig. 3.

**Fig. 7.** Schematic representation of the major N-glycans detected in rabbit erythrocytes. The most prominent series is shaded in the left panel and consists of biantennary structures which are preferentially extended and branched on one arm. The increment in size corresponds to a -GlcNAc-(Gal-Gal-GlcNAc)-Gal- repeating unit. As the structures become larger, additional branching on the other arm is also supported by the MS data, as represented by the structure at the bottom of the boxed series. The notation n refers to the number of -GlcNAc-(Gal-Gal-GlcNAc)-Gal- repeats, as drawn. For the major biantennary series, structures corresponding to n...
= 0 up to at least \( n = 5 \) were detected. The triantennary structures are less abundant relative to the biantennary glycans and are related similarly by one -GlcNAc-(Gal-Gal-GlcNAc)-Gal-increment. Key to symbols is the same as shown in Fig. 3.

**Fig. 8.** MALDI-MS analysis of permethylated O-glycans released from rabbit erythrocytes by reductive elimination. Core 1 O-glycans gave signals at \( m/z \) 534, 779, 895, 925, 1256 and 1316. Core 2 O-glycans gave signals at \( m/z \) 983, 1187, 1549 and 1579. Key to symbols for sugars: open circles, Gal; closed squares, GlcNAc; open squares, GalNAc; open diamonds, NeuGc; closed diamonds, NeuAc.
Table 1 GC-MS analysis of partially methylated alditol acetates obtained from N-glycosidase F released N-glycans from rabbit erythrocytes

<table>
<thead>
<tr>
<th>Elution Time (Mins)</th>
<th>Characteristic Fragment Ions</th>
<th>Assignment</th>
<th>Relative Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.59</td>
<td>102, 115, 118, 131, 175</td>
<td>terminal fucose</td>
<td>0.12</td>
</tr>
<tr>
<td>19.07</td>
<td>102, 118, 129, 145, 161, 162, 205</td>
<td>terminal mannose</td>
<td>0.11</td>
</tr>
<tr>
<td>19.33</td>
<td>102, 118, 129, 145, 161, 162, 205</td>
<td>terminal galactose</td>
<td>0.69</td>
</tr>
<tr>
<td>20.22</td>
<td>129, 130, 161, 190, 234</td>
<td>2-linked mannose</td>
<td>0.55</td>
</tr>
<tr>
<td>20.51</td>
<td>101, 118, 129, 234</td>
<td>3-linked galactose</td>
<td>0.59</td>
</tr>
<tr>
<td>20.68</td>
<td>102, 118, 129, 162, 189, 233</td>
<td>6-linked mannose</td>
<td>Not Detected</td>
</tr>
<tr>
<td>21.04</td>
<td>102, 118, 129, 162, 189, 233</td>
<td>6-linked galactose</td>
<td>0.07</td>
</tr>
<tr>
<td>21.40</td>
<td>129, 130, 190, 233</td>
<td>2,4-linked mannose</td>
<td>0.09</td>
</tr>
<tr>
<td>21.82</td>
<td>129, 130, 189, 190</td>
<td>2,6-linked mannose</td>
<td>0.02</td>
</tr>
<tr>
<td>21.96</td>
<td>118, 129, 189, 234</td>
<td>3,6-linked mannose</td>
<td>0.24</td>
</tr>
<tr>
<td>22.14</td>
<td>118, 129, 189, 234</td>
<td>3,6-linked galactose</td>
<td>0.25</td>
</tr>
<tr>
<td>22.92</td>
<td>117, 143, 145, 159, 203, 205</td>
<td>terminal GlcNAc</td>
<td>0.02</td>
</tr>
<tr>
<td>23.79</td>
<td>117, 159, 233</td>
<td>4-linked GlcNAc</td>
<td>1.00</td>
</tr>
<tr>
<td>25.10</td>
<td>117, 159, 261</td>
<td>4,6-linked GlcNAc</td>
<td>0.15</td>
</tr>
</tbody>
</table>

*Signs significantly reduced after α-galactosidase digestion

*Signs significantly more intense after β-galactosidase digestion

*Signs significantly reduced after β-galactosidase digestion

*Signs significantly reduced after α-1-2,3,4,6-N-acetylglucosaminidase and α-mannosidase digestions

*Signs detected after α-1-2,3,4,6-N-acetylglucosaminidase and α-mannosidase digestions
Table 2: Oligosaccharide constructs previously employed in the competitive murine sperm-egg binding assays.

Key to symbols: closed circle, Gal; open squares, GlcNAc. Linkage information is inside the symbol assuming that the monosaccharide is linked via C1 (for example “β6” inside an open square is β1-6 linked GlcNAc).
Scheme 1

Rabbit erythrocytes

Reduction & Carboxymethylation

Trypsin Digestion

PNGase F Digestion/Sep-Pak Fractionation

Released N-glycans

MALDI TOF/TOF MS/MS Analysis

Permethylation

MALDI-TOF & GC-MS Linkage Analysis

Reductive Elimination

Permethylation

ESI MS/MS Analysis

MALDI-TOF Analysis
Figure 2
Figure 3
Figure 8