Differential O-Glycosylation of a Conserved Domain Expressed in Murine and Human ZP3[†]

Sara Chalabi,[‡] Maria Panico,[‡] Mark Sutton-Smith,[‡] Stuart M. Haslam,[‡] Manish S. Patankar,[§] Frank A. Lattanzio,[||] Howard R. Morris,[‡,⊥] Gary F. Clark,*[,@] and Anne Dell*[,‡]

Division of Molecular Biosciences, Imperial College London, London SW7 2AZ, U.K., Department of Obstetrics and Gynecology, University of Wisconsin Medical School, Madison, Wisconsin 53792-6177, Department of Physiological Sciences, Eastern Virginia Medical School, Norfolk, Virginia 23501-1980, M-SCAN Mass Spectrometry Research and Training Centre, Silwood Park, Ascot SL5 7PZ, U.K., and Department of Obstetrics, Gynecology and Women's Health, Division of Reproductive and Perinatal Research, School of Medicine, University of Missouri, Columbia, Missouri 65212

Received July 4, 2005; Revised Manuscript Received November 9, 2005

ABSTRACT: Murine sperm initiate fertilization by binding to the zona pellucida (mZP), the specialized extracellular matrix of their homologous eggs. O-Glycans occupying two highly conserved vicinal glycosylation sites (Ser-332 and Ser-334) on the mZP glycoprotein designated mZP3 were previously implicated in this interaction. However, recent biophysical analyses confirm that neither site is occupied, implying that an alternate O-glycosylation domain may be operational in native mZP3. Since human ZP3 (huZP3) can substitute for mZP3 in rescue mice to mediate sperm binding, the site specificity of O-glycosylation in both native mZP3 and huZP3 was analyzed using ultrasensitive mass spectrometric techniques. Two O-glycosylation sites in native mZP3, one at Thr-155 and the other within the glycopeptide at positions 161−168 (ATVSSEEK), are conserved in huZP3 derived from transgenic mice. Thus, there is a specific O-glycosylation domain within native mZP3 expressing two closely spaced O-glycans that is very well conserved in an evolutionarily related glycoprotein. In native mZP3, core 2 O-glycans predominate at both sites. However, in huZP3 derived from rescue mice, the O-glycans associated with Thr-156 (analogous to Thr-155 in mZP3) are exclusively core 1 and related Tn sequences, whereas core 2 O-glycans predominate at the other conserved site. This unique restriction of O-glycan expression suggests that sequence differences in the conserved O-glycosylation domains of mZP3 and huZP3 affect the ability of core 2 *N*-acetylglucosaminyltransferase(s) to extend the core 1 sequence. However, this difference in O-glycosylation at Thr-156 does not affect the fertility or the sperm binding phenotype of eggs derived from female huZP3 rescue mice.

Development in eutherian mammals begins when sperm undergo initial binding to the specialized extracellular matrix of the egg, known as the ZP.[1] The molecular basis for this interaction remains poorly understood, though exceptional knowledge exists about gamete binding in many lower species (*1*, *2*). The predominant animal model for eutherian sperm−egg binding is the mouse (*3*, *4*). The mZP consists of three major glycoproteins (mZP1−mZP3). mZP3 is responsible for initial sperm−egg binding and the induction of the acrosome reaction (*3*, *4*). Gene inactivation studies clearly demonstrate that the deletion of mZP3 leads to a complete loss of the mZP and therefore fertility (*5*).

The major model for initial murine gamete binding implicates mZP3-associated O-glycans in this interaction (*3*, *4*). Previous studies were carried out with biologically active recombinant mZP3 derived from F9 embryonal carcinoma cells (*6*) to map the exact O-glycosylation sites responsible for this interaction. Exon swapping studies suggested that a region (^{329}SNSSSS^{334}) encoded by exon 7 of the mZP3 gene was required for inhibition of murine sperm−egg binding in vitro (*7*). The biologically active O-glycosylation sites within this exon were mapped by replacing specific Ser residues with Ala. Conversion of either Ser-332 or Ser-334 to Ala resulted in mutant recombinant mZP3 molecules that do not inhibit murine sperm−ZP binding (*8*), suggesting that vicinal O-glycan presentation was required for this biological activity.

[†] This work was supported by grants from the National Institutes of Health (HD 35652 to G.F.C. via an institutional incentive fund), the Breeden-Adams Cancer Foundation (G.F.C. and M.S.P.), the Elsa U. Pardee Foundation (M.S.P.), and the Biotechnology and Biological Sciences Research Council (BBSRC) and the Wellcome Trust (to A.D. and H.R.M.). A.D. is a BBSRC Professorial Fellow, and S.C. was supported by a BBSRC Studentship.

* To whom correspondence should be addressed. G.F.C.: telephone, (573) 882-1725; fax, (573) 882-9010; e-mail, clarkgf@health.missouri.edu. A.D.: telephone, (+44-207) 594-5219; fax, (+44-207) 225-0458; e-mail, a.dell@imperial.ac.uk.

[‡] Imperial College London.

[§] University of Wisconsin Medical School.

[||] Eastern Virginia Medical School.

[⊥] M-SCAN Mass Spectrometry Research and Training Centre.

[@] University of Missouri.

[1] Abbreviations: CAD, collisional activated dissociation; huZP, human zona pellucida; MS, mass spectrometry; MS/MS, tandem mass spectrometry; LacdiNAc, di-*N*-acetyllactosamine; LacNAc, *N*-acetyllactosamine; mZP, murine zona pellucida; nano-LC ES-MS/MS, nano-liquid chromatography−electrospray-tandem mass spectrometry; PNGase F, peptide *N*-glycosidase F; TOF, time-of-flight; ZP, zona pellucida.

10.1021/bi0512804 CCC: $33.50
Published on Web 12/17/2005

A recent proteomic analysis indicates that Ser-332 and Ser-334 are not glycosylated in native mZP3 (*9*). These data exposed two major problems with the mapping studies (*7*, *8*). First, it was assumed that recombinant mZP3 is glycosylated in exactly the same manner as native mZP3. The possibility that O-glycosylation could be initiated at another mZP3 domain in F9 cells, whose polypeptide *N*-acetylgalactosaminyltransferase repertoire is likely to be different from that of the mouse (*10*), was not entertained. Second, compelling evidence indicates that polypeptide *N*-acetylgalactosaminyltransferases are also regulated by the amino acid sequence of the glycoprotein carrier (*11*). Thus, the mutations at Ser-332 and Ser-334 could have subtly affected the activity of the polypeptide *N*-acetylgalactosaminyltransferases or other glycosyltransferases that are involved in the addition of carbohydrate ligands for murine sperm binding.

Previous transgenesis studies indicate that huZP3 can substitute for mZP3 in eggs to mediate murine sperm−egg binding (*12*). In addition, native mZP3 and huZP3 derived from rescue mice also express identical O-glycans (*13*). Thus, huZP3 may be able to mimic the biological activity of native mZP3 by acquiring identical O-glycans. However, because the proper positioning of O-glycans was previously implicated in the binding interaction (*7*, *8*), it is essential to perform analysis of the O-glycosylation sites within huZP3 to determine if they are identical to native mZP3. We observed two conserved O-glycosylation sites close to each other in space in native mZP3 and huZP3 produced in rescue mice. However, the O-glycans linked at the first site within this domain differ between mZP3 and huZP3.

EXPERIMENTAL PROCEDURES

Materials. Purified ZP was isolated from flash-frozen mouse ovaries obtained from 12−14-week-old wild-type and huZP3 rescue mice as described previously (*13*). All other chemicals and reagents were purchased from Sigma unless otherwise indicated.

Reduction and Carboxymethylation. The ZP preparations were dissolved in 300 μL of tris(hydroxymethyl)aminomethane (Amresco) buffer (0.6 M, pH 8.0) previously degassed with N_2. Dithiothreitol was added in 1 molar equiv to a 4-fold molar excess over the number of disulfide bonds present in the sample and incubated for 1 h at 37 °C. The sample was carboxymethylated using a 5-fold molar excess of iodoacetic acid over the amount of reductant added. This derivatization was carried out in the dark for 1 h at room temperature. The sample was immediately transferred to Snakeskin pleated dialysis tubing (7 kDa M_r cutoff) (Pierce, Rockford, IL), dialyzed against 4.5 L of 50 mM ammonium hydrogen carbonate (pH 7.4) for 48 h at 4 °C with a buffer exchange every 12 h, and lyophilized.

Trypsin Digestion. The ZP preparations were dissolved in 200 μL of 50 mM ammonium hydrogen carbonate (pH 8.5) (AnalaR, Poole, U.K.) and incubated with TPCK-treated bovine pancreas trypsin at an enzyme:substrate ratio of 1:50 (w/w). The digestion was carried out at 37 °C for 5 h and terminated by boiling for 2 min prior to lyophilization.

PNGase F Digestion. The ZP glycopeptides were dissolved in 200 μL of ammonium hydrogen carbonate (50 mM, pH 8.5) and incubated with 0.5 unit of PNGase F (Roche) for 18 h at 37 °C to release N-glycans. The reaction was terminated by lyophilization.

C_{18} Sep-Pak Purification. Reverse-phase C_{18} Sep-Pak cartridges (Waters Corp.) were conditioned using successive 5 mL washes of methanol, 5% acetic acid, and 100% propanol and 10 mL of 5% acetic acid. The PNGase F-digested sample was then loaded onto the C_{18} Sep-Pak cartridge and eluted sequentially with 3 mL of 5% acetic acid and 2 mL each of 20, 40, and 60% propanol in 5% acetic acid followed by 2 mL of 100% propanol. The 5% acetic acid fraction was lyophilized. The other fractions containing the O-glycans and/or peptides were dried by a combination of vacuum centrifugation and lyophilization.

Nano-LC−ES-MS/MS Analysis. Tryptic digests were analyzed by nano-LC−ES-MS/MS using a reverse-phase nano-HPLC system (Dionex, Sunnyvale, CA) connected to a quadrupole TOF mass spectrometer (Q-STAR Pulsar I, MDS Sciex). The digests were separated by a binary nano-HPLC gradient generated by an Ultimate pump fitted with a Famos autosampler and a Switchos microcolumn switching module (LC Packings, Amsterdam, The Netherlands). An analytical C_{18} nanocapillary (75 μm inside diameter × 15 cm, PepMap) and a micro precolumn C_{18} cartridge were employed for on-line peptide separation. The digest was first loaded onto the precolumn and eluted with 0.1% formic acid (Sigma) in water (HPLC grade, Purite) for 4 min. The eluant was then transferred onto an analytical C_{18} nanocapillary HPLC column and eluted at a flow rate of 150 nL/min using the following gradient of solvent A [0.05% (v/v) formic acid in a 95:5 (v/v) water/acetonitrile mixture] and solvent B [0.04% formic acid in a 95:5 (v/v) acetonitrile/water mixture]: 99% A from 0 to 5 min, 99 to 90% A from 5 to 10 min, 90 to 60% A from 10 to 70 min, 60 to 50% A from 70 to 71 min, 50 to 5% A from 71 to 75 min, 5% A from 75 to 85 min, 5 to 95% A from 85 to 86 min, and 95% A from 86 to 90 min. Data acquisition was performed using Analyst QS software with an automatic information-dependent-acquisition (IDA) function.

RESULTS

Analytical Strategy. mZP3 is the major carrier of O-glycans in the mZP (*13*). Using this knowledge, sensitivity was maximized by analyzing proteolytic digests of mZP rather than purified mZP3. To search for O-linked glycosylation sites on ZP3, on-line nano-LC−ES-MS/MS was performed on tryptic digests of ZP depleted of N-glycans by PNGase F digestion. The MS data across the elution profile were then manually searched for the presence of low-mass signals corresponding to carbohydrate fragments [e.g., *m*/*z* 204 (HexNAc) and *m*/*z* 366 (HexHexNAc)] indicative of O-glycosylation. All spectra exhibiting putative carbohydrate fragment ions were scrutinized for the presence of molecular ions corresponding to glycopeptides. The existing database of O-glycans linked to mZP3 was employed to inform these searches (*13*). Ions separated from their neighboring ions by increments corresponding to sugar *m*/*z* differences were specifically flagged for subsequent analysis. For example, since NeuAc and NeuGc differ by 16 Da, pairs of doubly and triply charged signals separated by *m*/*z* 8^{2+} and *m*/*z* 5.3^{3+}, respectively, are indicative of sialylated glycopeptides. The putative glycopeptide assignments were then confirmed by MS/MS fragmentation data.

Exclusive Expression of Core 1 O-Glycans and Related Tn Sequences on Thr-156 of huZP3. Carbohydrate fragment

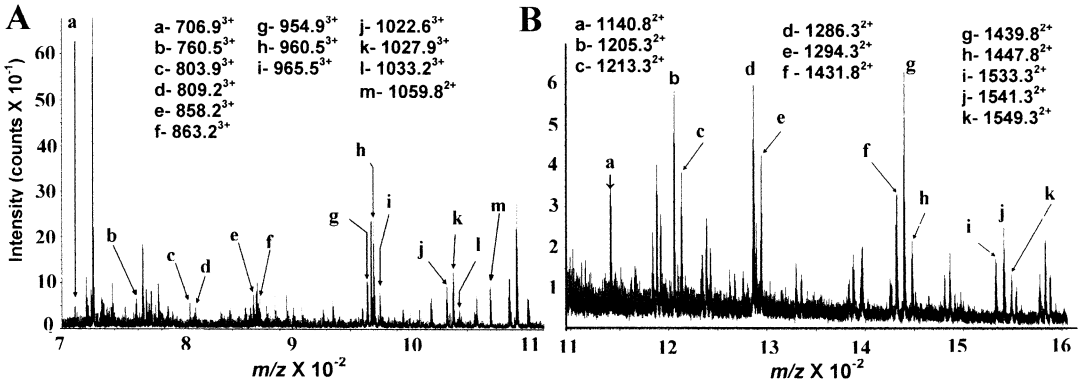
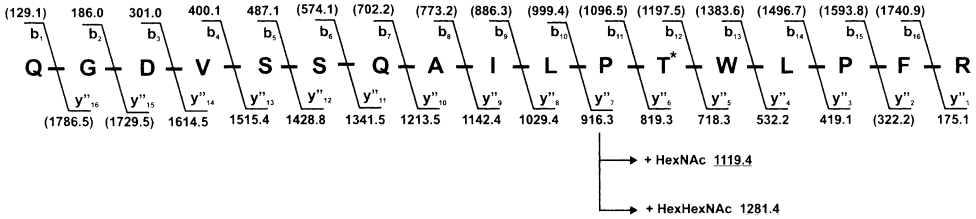


FIGURE 1: MS data derived from components eluting at 68−69 min from on-line nano-LC−ES-MS of the 20% propanol Sep-Pak fraction obtained following trypsin and PNGase F digestion of huZP3. Panels A and B correspond to low- and high-mass regions of the spectrum, respectively. Multiply charged ions correspond to glycans attached to the huZP3 peptide ^{145}QGDVSSQAILPTWLPFR^{161}. A comprehensive list of the O-glycans at this site and the *m*/*z* values observed are given in Table 1.



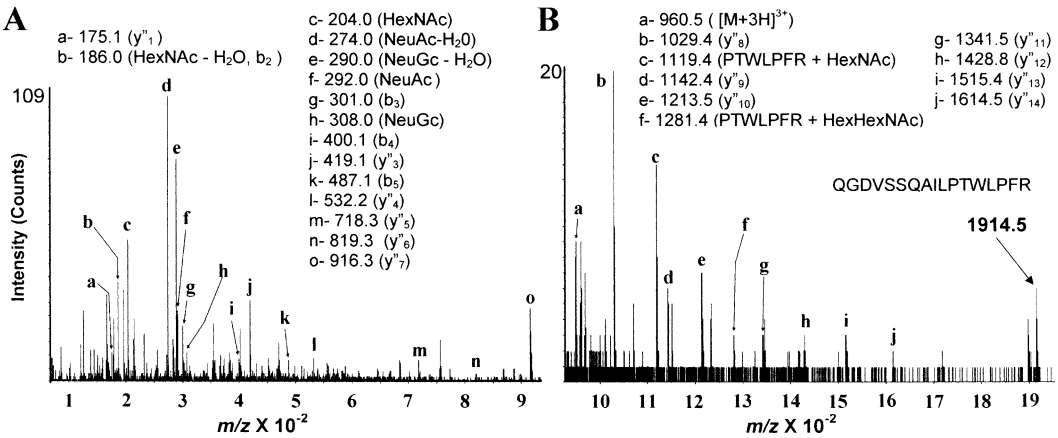


FIGURE 2: CAD ES-MS/MS spectrum of *m*/*z* 960.5^{3+} acquired during the on-line nano-LC−ES-MS/MS experiment of the 20% propanol Sep-Pak fraction obtained following trypsin and PNGase F digestion of huZP3. Panels A and B correspond to low- and high-mass regions of the spectrum, respectively. The characteristic sugar ions at *m*/*z* 204.0, 292.0, and 308.0 correspond to HexNAc, NeuAc, and NeuGc, respectively. These are accompanied by signals at *m*/*z* 186.0, 274.0, and 290.0, respectively, that correspond to the loss of water from these sugars. The signal at *m*/*z* 186.0 also corresponds to a peptide fragment. The peptide sequence assignments of ^{145}QGDVSSQAILPTWLPFR^{161} are provided above the spectra. The asparagine at position 147 was converted to aspartic acid following the release of N-glycans by PNGase F treatment. The HexNAc and HexHexNAc sugars were observed to be attached to the peptide fragment ^{155}PTWLPFR^{161}. The bracketed numbers on the figure correspond to the *m*/*z* values expected but not observed in the CAD ES-MS/MS spectrum of *m*/*z* 960.5^{3+}. The asterisk denotes the site of O-glycosylation at Thr-156.

ions indicative of O-glycopeptides were observed in the MS spectra of components eluting at 68−69 min in the nano-LC−ES-MS analysis of the 20% propanol Sep-Pak fraction of huZP. Detailed examination of the MS data from this portion of the chromatogram revealed the presence of clusters of doubly and triply charged ions showing mass intervals that were consistent with glycopeptides (Figure 1). Thus, the cluster of triply charged signals observed at *m*/*z* 1022.6^{3+}, 1027.9^{3+}, and 1033.2^{3+} has a characteristic *m*/*z* difference of *m*/*z* 5.3^{3+}, suggestive of disialylated glycopeptides differing in their sialic acid content: capped with NeuAc (*m*/*z* 1022.6^{3+}, M_r = 3064.6), NeuAc and NeuGc (*m*/*z* 1027.9^{3+}, M_r = 3080.6), and NeuGc (*m*/*z* 1033.2^{3+}, M_r = 3096.6). Using this cluster as a starting point, structural relationships to neighboring signals can be elucidated. For example, a related signal cluster observed at *m*/*z* 954.9^{3+}, 960.5^{3+}, and 965.5^{3+} corresponds to one fewer HexNAc residue (203 Da) than the differentially sialylated glycopeptides described above. The same strategy was employed to assign the remaining signals present in the MS spectrum (Table 1). The spectrum of one of the triply charged signals (*m*/*z* 960.5^{3+}) that was selected for CAD ES-MS/MS analysis is displayed in Figure 2. The fragment ions map onto the human ZP3 sequence ^{145}QGDVSSQAILPTWLPFR^{161}. There are three possible O-glycosylation sites within this sequence (Ser-149, Ser-150, and Thr-156). The signals at *m*/*z* 1119.4 and 1281.4

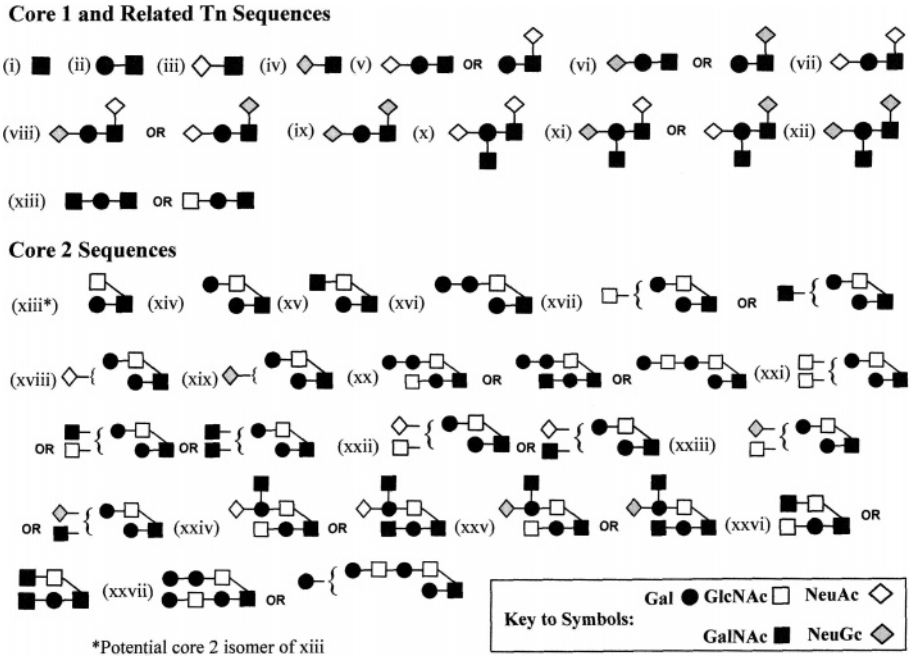


FIGURE 3: O-Glycans associated with the conserved O-glycosylation domain within huZP3 and mZP3. Structure i is commonly called the Tn antigen, and structures iii and iv are sialyl-Tn. Attachment of Gal to the 3-position of the Tn antigen generates core 1 structures. Branching of core 1 by the addition of GlcNAc to the 6-position of GalNAc in the core 1 structure produces the core 2 sequence. Thus, structures ii and v−xiii are core 1 O-glycans, while structures xiii*−xxvii are core 2 O-glycans. Only core 1 O-glycans and the related Tn sequences (i, iii, and iv) are found at Thr-156 in huZP3. In contrast, core 2 O-glycans (structures xiii*−xxv and xxvii) predominate on the huZP3 peptide ^{162}TTVFSEEK^{169}, although some core 1 and related Tn O-glycans are also present (structures i, ii, v, and vi). Structures i, ii, v, and vi and core 2 O-glycans (structures xiii*−xx and xxii−xxvii) are also linked to Thr-155 in mZP3. The second mZP3 glycosylation site (^{161}ATVSSEEK^{168}**)** also contains core 1 and related Tn structures (structures i, ii, v, and vi) and core 2 O-glycans (structures xiii*−xx, xxii−xxv, and xxvii). A comprehensive list of the O-glycans at each site and the *m*/*z* values observed are given in Tables 1−5.

in the MS/MS spectrum (Figure 2) correspond to increments of HexNAc and HexHexNAc, respectively, on the peptide fragment ^{155}PTWLPFR^{161}, confirming O-glycosylation at Thr-156. In total, 12 exclusively core 1 and Tn-related sequences were determined to be linked at this site (Figure 3, structures i−xii). A principal feature of these O-glycans was sialylation, with the majority carrying either one or two NeuAc or NeuGc sugars. The Sd^a epitope (NeuAc/NeuGcα2−3[GalNAcβ1−4]Gal) was also observed as a component of some of these glycans. Additional evidence for these assignments was obtained via nano-LC−ES-MS of the 40% propanol Sep-Pak fraction which yielded data that could be assigned to the peptide ^{145}QGDVSSQAILPTWLPF^{160} carrying glycan structures i−xii in Figure 3. This sequence is the same peptide described above, except that cleavage has occurred C-terminal to Phe-160 instead of Arg-161, presumably due to contaminating chymotryptic activity.

Core 1, Core 2, and Related Tn Sequences Are Expressed on Thr-155 of mZP3. The domain of mZP3 corresponding to the Thr-156 glycopeptide in huZP3 has the sequence ^{144}QGDVSSHPIQPTWVPFR^{160}. This murine glycopeptide was observed at an elution time of ∼33 min in the nano-LC−ES-MS experiment on the 20% propanol Sep-Pak fraction of mZP (Figure 4 and Table 2). The most intense peak assigned in Figure 4 at *m*/*z* 908.3^{3+} corresponds to $HexHexNAc_3$ attached to the mZP3 peptide ^{144}QGDVSSHPIQPTWVPFR^{160}. Further examination of the spectrum revealed the presence of related signals; for instance, *m*/*z* 894.6^{3+} corresponds to the substitution of a HexNAc residue for a hexose ($Hex_2HexNAc_2$ attached to ^{144}QGDVSSHPIQPTWVPFR^{160}). The most notable feature of the data is that the O-glycans identified at this site (Figure 3, structures i, ii, xiii−xvii, xx, xxvi, xxvii) were predominantly core 2 sequences, in contrast to those found at the equivalent site on huZP3. Assignments were confirmed by CAD ES-MS/MS experiments. For example, collisional activation of the triply charged signal at *m*/*z* 948.6^{3+} yielded fragment ion data establishing both the peptide sequence and the attach-

Table 1: Glycans Attached at Thr-156 within the Peptide Backbone ^{145}QGDVSSQAILPTWLPFR^{161} on huZP3^a

m/z Observed	M_r	Glycopeptide minus Peptide (M_r)	Predicted Carbohydrate Structure
958.3^{2+}	1914.6	0	No Carbohydrate Attached
1059.8^{2+} 706.9^{3+}	2117.6	203.0	■
1140.8^{2+} 760.5^{3+}	2279.6	365.0	●■
1205.3^{2+} 803.9^{3+}	2408.6	494.0	◇■
1213.3^{2+} 809.2^{3+}	2424.6	510.0	◆■
1286.3^{2+} 858.2^{3+}	2570.6	656.0	◇●■ OR ●■◇
1294.3^{2+} 863.2^{3+}	2586.6	672.0	◆●■ OR ●■◆
1431.8^{2+} 954.9^{3+}	2861.6	947.0	◇●■◇
1439.8^{2+} 960.5^{3+}	2877.6	963.0	◇●■◇ OR ◇●■◆
1447.8^{2+} 965.5^{3+}	2893.6	979.0	◆●■◆
1533.3^{2+} 1022.6^{3+}	3064.6	1150.0	◇●■◇
1541.3^{2+} 1027.9^{3+}	3080.6	1166.0	◇●■◇ OR ◇●■◆
1549.3^{2+} 1033.2^{3+}	3096.6	1182.0	◆●■◆

^{a} The M_r values correspond to nonprotonated species. Symbols are as follows: (●) galactose, (■) GalNAc, (◇) NeuAc, and (◆) NeuGc.

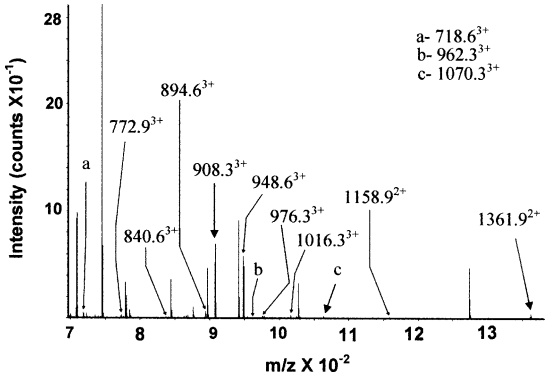


FIGURE 4: MS data derived from nonsialylated components eluting at 33 min from on-line nano-LC ES-MS of the 20% propanol Sep-Pak fraction obtained following trypsin and PNGase F digestion of mZP3. Multiply charged ions correspond to nonsialylated glycans attached to the mZP3 peptide ^{144}QGDVSSHPIQPTWVPFR^{160}. A comprehensive list of the O-glycans at this site and the *m*/*z* values observed are given in Table 2. Masses of all ions are given in Table 2.

ment site at Thr-155 (Figure 5). The MS/MS spectrum of *m*/*z* 948.6^{3+} displays definitive evidence of glycosylation via the presence of major low-mass signals at *m*/*z* 168.0 (HexNAc − 2H$_2$O), *m*/*z* 186.0 (HexNAc − H$_2$O), *m*/*z* 204.1 (HexNAc), *m*/*z* 366.1 (HexHexNAc), and *m*/*z* 528.1 (Hex$_2$HexNAc). The occurrence of the diagnostic Hex$_2$HexNAc fragment suggests the presence of a Gal-Gal-GlcNAc-containing glycopeptide, i.e., Figure 3, structures xvi (Hex$_3$HexNAc$_2$), xx (Hex$_3$HexNAc$_3$), and xxvii (Hex$_4$HexNAc$_3$). Subtracting the masses of each of these structures from the total molecular weight of the glycopeptide revealed that structure xvi (Figure 3) is attached to the mZP3 peptide ^{144}QGDVSSHPIQPTWVPFR^{160}. The site of O-glycosylation was deduced to be Thr-155 due to the presence of signals at *m*/*z* 1105.4 and 1267.4, corresponding to the peptide fragment ^{154}PTWVPFR^{160} carrying HexNAc and HexHexNAc sugars, respectively, and signals at *m*/*z* 1443.5 and 1605.6 which correspond to the addition of HexNAc and HexHexNAc, respectively, to the peptide fragment ^{151}PIQPTWVPFR^{160}. Interestingly, all the O-glycans observed at this elution time were nonsialylated, and the most abundant components were capped with galactose or contained a LacdiNAc (GalNAcβ1−4GlcNAc) antenna. This is in accord with previous MALDI screening of the O-glycan repertoire of mZP (*13*). Sialylated glycopeptides (Figure 3, structures v, vi, xviii, xix, and xxii−xxv) were found to elute later, at approximately 40 min (Figure 6 and Table 3). These data were assigned by utilizing our strategy, described earlier, of searching for characteristic *m*/*z* differences between peaks corresponding to differential sialylation. For instance, the triply charged peak at *m*/*z* 997.0 was found to correspond to NeuGcHex$_2$HexNAc$_2$ attached to the ^{144}QGDVSSHPIQPTWVPFR^{160} peptide. Further inspection of the data revealed a smaller triply charged signal at *m*/*z* 991.6, indicating the presence of the same glycopeptide containing NeuAc instead of NeuGc. Families of multiply charged signals were observed at 45 and 52 min in the nano-LC run (data not shown). These signals correspond to the nonsialylated and sialylated glycoforms, respectively, of the peptide lacking the C-terminal Arg residue, as described previously for the huZP3 peptide.

Core 2 O-Glycans Predominate on the ^{162}TTVFSEEK^{169} Sequence in huZP3. Analysis of MS data from the sample

Table 2: Nonsialylated Glycans Attached to the Peptide Backbone ^{144}QGDVSSHPIQPTWVPFR^{160} on mZP3[a]

m/z Observed	M_r	Glycopeptide minus Peptide (M_r)	Predicted Carbohydrate Structure
-	1950.0	0	No Carbohydrate Attached
718.6^{3+}	2152.8	202.8	■
1158.9^{2+} 772.9^{3+}	2315.8	365.8	●-■
840.6^{3+}	2518.8	568.8	●-■ OR ■-●-■ OR □-●-■
894.6^{3+}	2680.8	730.8	●-●-■
1361.9^{2+} 908.3^{3+}	2721.8	771.8	■-●-■
948.6^{3+}	2842.8	892.8	●-●-■-■
962.3^{3+}	2883.9	933.9	□-{●-■ OR ■-●-■}
976.3^{3+}	2925.9	975.9	■-●-■ OR ■-□-●-■
1016.3^{3+}	3045.9	1095.9	●-□-■ OR ●-■-■ OR □-●-■-■
1070.3^{3+}	3207.9	1257.9	●-●-□-■ OR ●-{□-●-■-■}

[a] The M_r values correspond to nonprotonated species. Symbols are as follows: (●) galactose, (■) GalNAc, and (□) GlcNAc.

Table 3: Sialylated Glycans Attached to the Peptide Backbone ^{144}QGDVSSHPIQPTWVPFR^{160} on mZP3[a]

m/z Observed	M_r	Glycopeptide minus Peptide (M_r)	Predicted Carbohydrate Structure
-	1950.0	0	No Carbohydrate Attached
1304.4^{2+} 869.9^{3+}	2606.8	656.8	◇-●-■ OR ●-■-◇
1312.4^{2+} 875.3^{3+}	2622.8	672.8	◇-●-■ OR ●-■-◇
991.6^{3+}	2971.8	1021.8	◇-{●-□-■}
997.0^{3+}	2988.0	1038.0	◆-{●-□-■}
1059.3^{3+}	3174.9	1224.9	◇-{●-□-■} OR ◇-{●-□-■}
1064.6^{3+}	3190.8	1240.8	◆-{●-□-■} OR ◆-{●-□-■}
1127.3^{3+}	3378.9	1428.9	◇-■-●-□-■ OR ◇-■-●-□-■
1132.6^{3+}	3394.8	1444.8	◆-■-●-□-■ OR ◆-■-●-□-■

[a] The M_r values correspond to nonprotonated species. Symbols are as follows: (●) galactose, (■) GalNAc, (□) GlcNAc, (◇) NeuAc, and (◆) NeuGc.

eluting from 13 to 15 min in the nano-LC−ES-MS experiment revealed the presence of a series of singly and doubly charged signals consistent with glycopeptides (Figure 7). Compositions were deduced using exactly the same logic as described earlier for the first glycopeptide, and assignments are given in Table 4. For example, CAD ES-MS/MS analysis of the doubly charged signal at *m*/*z* 981.2^{2+} gave singly charged peptide fragment ions that were consistent with the huZP3 peptide ^{162}TTVFSEEK^{169} (Figure 8). In addition, loss of part or all of the glycan chain without cleavage of the peptide backbone yielded singly charged ions at *m*/*z* 1143.3, 1305.3, 1346.4, 1508.2, and 1670.3, corresponding to the peptide plus increments of HexNAc, HexHexNAc, HexNAc$_2$, HexHexNAc$_2$, and Hex$_2$HexNAc$_2$, respectively. Taking into account the molecular weight of the intact glycopeptide, we deduced the O-glycan attached to the *m*/*z* 981.2^{2+} component to have the composition NeuAcHex$_2$HexNAc$_2$ (Figure 3,

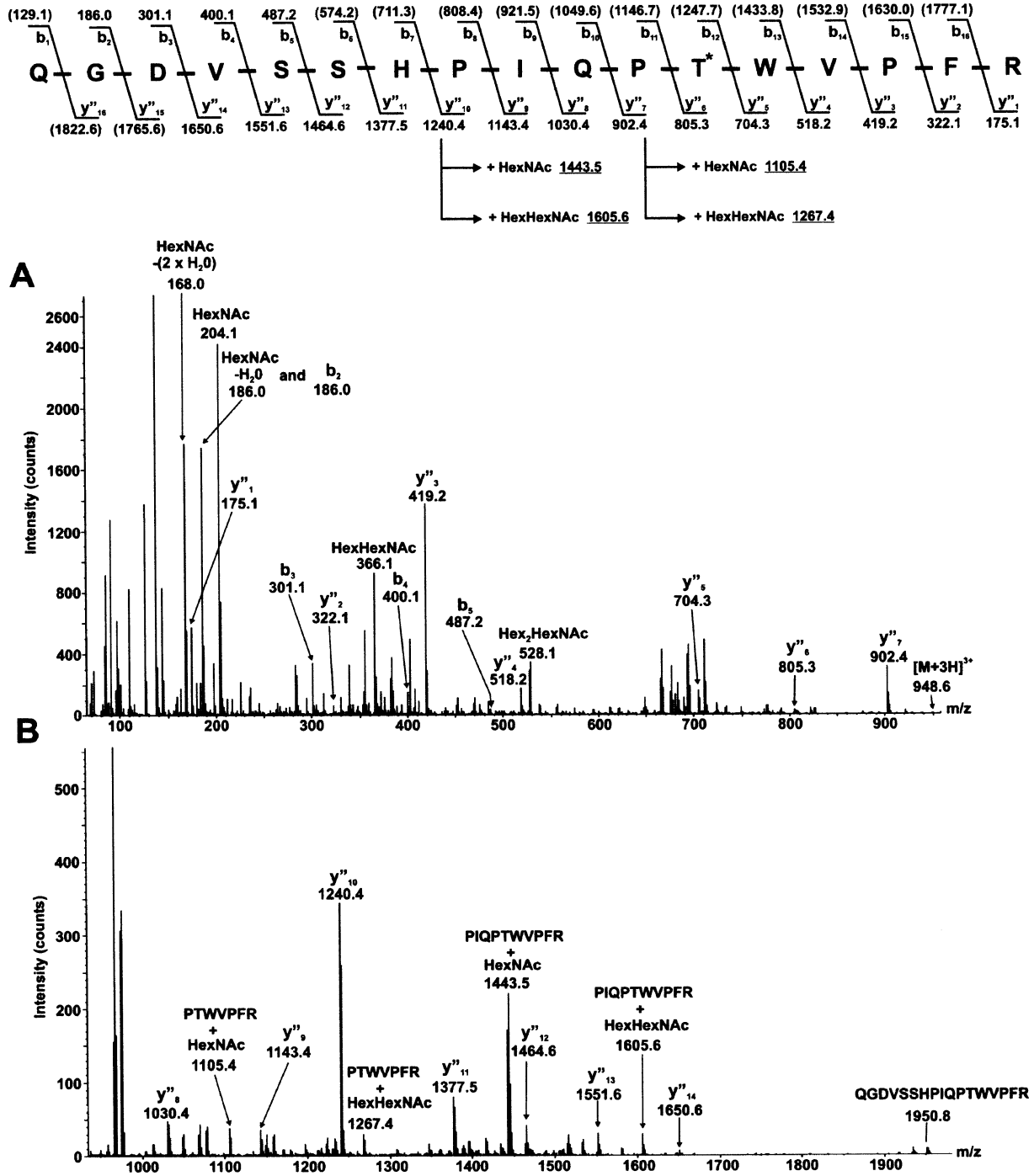


FIGURE 5: CAD ES-MS/MS spectrum of *m/z* 948.6^{3+} acquired during the on-line nano-LC−ES-MS/MS experiment of the 20% propanol Sep-Pak fraction obtained following trypsin and PNGase F digestion of mZP3. Panels A and B correspond to low- and high-mass regions of the spectrum, respectively. The signals at *m/z* 204.1, 366.1, and 528.1 are characteristic of sugars HexNAc, HexHexNAc, and Hex_2-HexNAc, respectively. These are accompanied by signals at *m/z* 168.0 and 186.0 which correspond to the loss of two water molecules and one water molecule, respectively, from HexNAc. The signal at *m/z* 186.0 also corresponds to a peptide fragment. The peptide sequence assignments of ^{144}QGDVSSHPIQPTWVPFR^{160} are shown above the spectra. The mZP3 sequence ^{144}QGDVSSHPIQPTWVPFR^{160} was derived from the CAD ES-MS/MS analysis of *m/z* 948.6^{3+} (Figure 6). The asparagine at position 146 was converted to aspartic acid following the release of N-glycans by PNGase F treatment. HexNAc and HexHexNAc sugars were observed to be attached to the peptide fragments ^{151}PIQPTWVPFR^{160} and ^{154}PTWVPFR^{160}. The numbers in brackets correspond to the *m/z* values expected but not observed in the CAD ES-MS/MS spectrum of *m/z* 948.6^{3+}. The asterisk denotes the site of O-glycosylation at Thr-155.

structure xviii). This glycan was one of 18 sequences located at this site (Figure 3, structures i, ii, v, vi, xiii−xxv, and xxvii; Table 4). Interestingly, in contrast to the restricted expression of core 1 and related Tn structures at Thr-156, the majority of sugars attached to ^{162}TTVFSEEK^{169} are core 2 O-glycans. There are three potential sites for O-glycosylation within this peptide at Thr-162, Thr-163, and Ser-166. Unfortunately, the exact attachment site could not be determined from the CAD ES-MS/MS data.

Core 2 O-Glycans Predominate on ^{160}RATVSSEEK^{168} in mZP3. A search of the data arising from nano-LC−ES-MS analysis of the 20% propanol Sep-Pak fraction of mZP failed to detect the ^{161}ATVSSEEK^{168} glycopeptide. Since the homologous glycopeptide from huZP3 eluted very early in

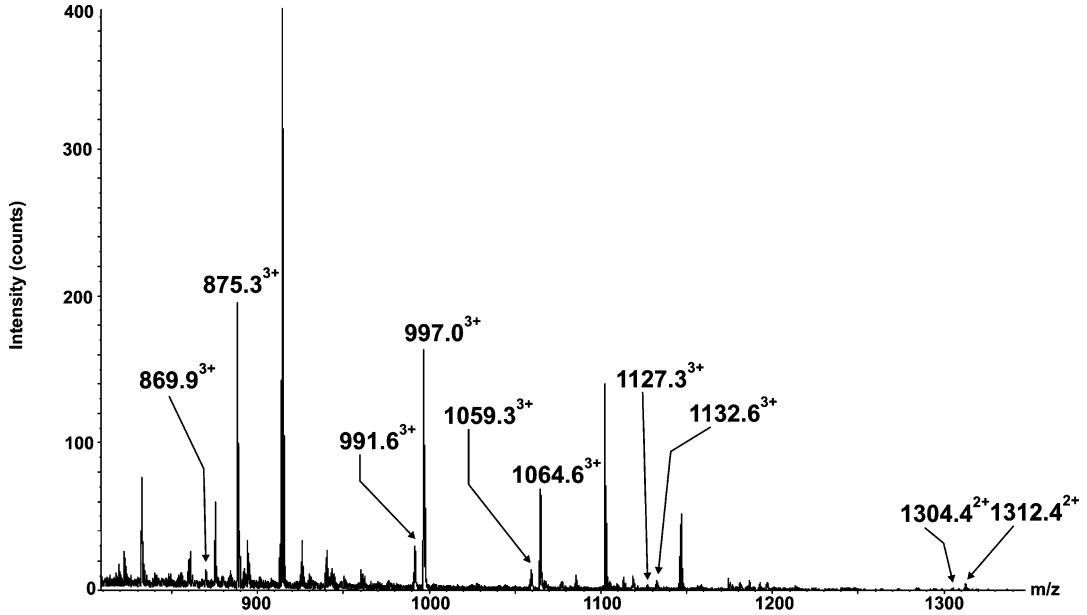


FIGURE 6: MS data derived from sialylated components eluting at 40 min from on-line nano-LC−ES-MS of the 20% propanol Sep-Pak fraction obtained following trypsin and PNGase F digestion of mZP3. Multiply charged ions correspond to sialylated glycans attached to the mZP3 peptide ^{144}QGDVSSHPIQPTWVPFR^{160}. A comprehensive list of the O-glycans at this site and the *m*/*z* values observed are given in Table 3.

the nano-LC run (see above), and the murine glycopeptide is predicted to be considerably more hydrophilic, we searched for its presence in the 5% acetic acid Sep-Pak fraction. On-line nano-LC−ES-MS of this fraction revealed the presence of glycopeptides eluting at approximately 7 min (Figure 9 and Table 5). CAD ES-MS/MS analyses of *m*/*z* 686.3^{2+} (Figure 10) confirmed the presence of the extended peptide ^{160}RATVSSEEK^{168} carrying primarily core 2 O-glycans (Figure 3, structures i, ii, v, vi, xiii−xx, xxii−xxv, and xxvii). This glycopeptide arises from chymotryptic cleavage at Phe-159 instead of tryptic cleavage at Arg-160. The glycans observed were the same as those on the equivalent huZP3 peptide (^{162}TTVFSEEK^{169}).

A Conserved O-Glycosylation Domain Is Present in both mZP3 and huZP3 Produced in Mice. A sequence alignment of the complete shared O-glycosylation domain is presented in Figure 11. Clearly, there are significant similarities and differences in the amino acid sequence within this region. Both share a conserved N-glycosylation site N-terminal to the O-glycosylation domain. As we have noted, the positioning of site 1 in each domain (Thr-155/156) is well-conserved, although it is apparent from our data that there are significant differences in O-glycosylation at this site. However, the O-glycans linked to site 2 are very similar in both huZP3 and mZP3.

DISCUSSION

There is a very substantial literature suggesting that O-glycans play a crucial role in murine sperm−egg binding (*3*, *4*). Wassarman and colleagues originally reported that O-glycans derived from mZP3 inhibit this binding (*14*). A subsequent study by Bleil and Wassarman suggested that O-glycans terminated with Galα1−3Gal sequences also inhibit murine sperm−egg binding (*15*). However, more recent knockout studies indicate that oligosaccharides terminated with α1−3-linked Gal are not required for either fertility (*16*) or sperm binding (*17*). Recombinant mZP3 that

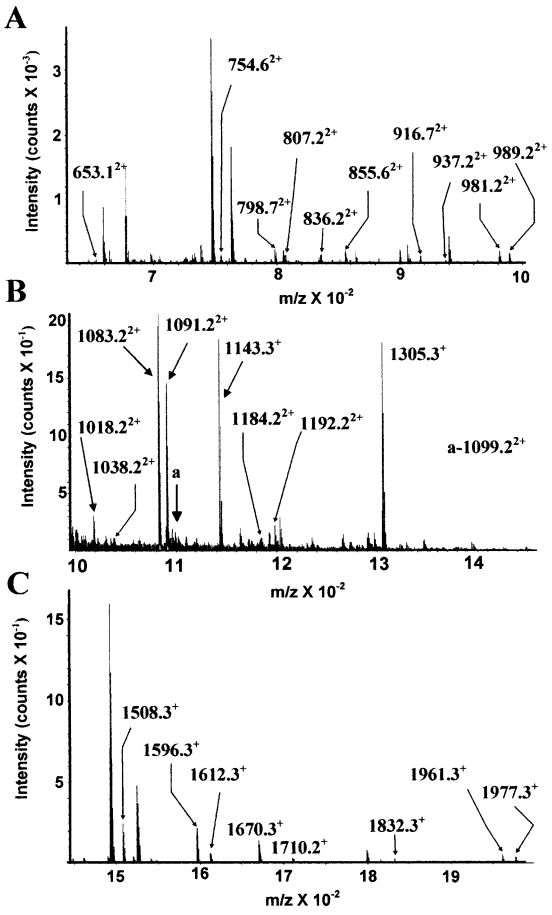


FIGURE 7: MS data derived from components eluting at 13−15 min from on-line nano-LC−ES-MS of the 20% propanol Sep-Pak fraction obtained following trypsin and PNGase F digestion of huZP3. Panels A−C correspond to low-, intermediate-, and high-mass regions of the spectrum, respectively. Multiply charged ions correspond to glycans attached to the huZP3 peptide ^{162}TTVFSEEK^{169}. A comprehensive list of the O-glycans at this site and the *m*/*z* values observed are given in Table 4.

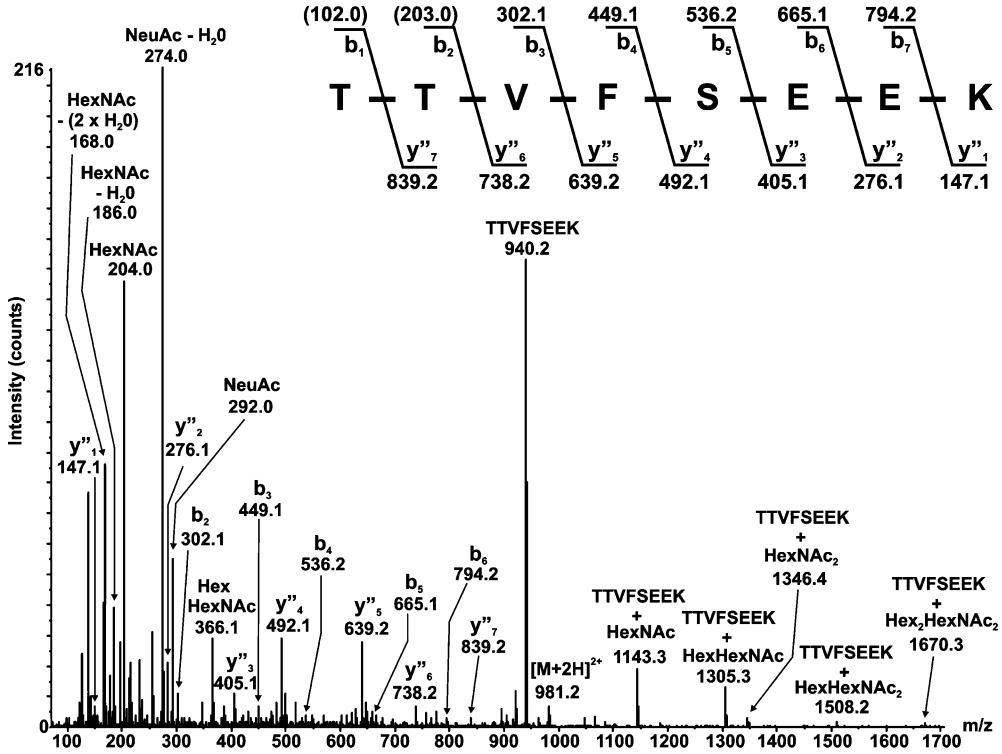


FIGURE 8: CAD ES-MS/MS spectrum of *m*/*z* 981.2^{2+} acquired during the on-line nano-LC−ES-MS/MS experiment of the 20% propanol Sep-Pak fraction obtained following trypsin and PNGase F digestion of huZP3. The signals at *m*/*z* 204.0, 292.0, and 366.1 are characteristic of sugars HexNAc, NeuAc, and HexHexNAc, respectively. These are accompanied by signals at *m*/*z* 168.0 (corresponding to the loss of two water molecules from HexNAc) and *m*/*z* 186.0 and 274.0 which are consistent with the loss of one water molecule from HexNAc and NeuAc, respectively. The peptide sequence assignments are shown in the inset. The numbers in brackets correspond to the *m*/*z* values expected but not observed in the CAD ES-MS/MS spectrum of *m*/*z* 981.2^{2+}.

potently inhibited murine sperm−egg binding was also isolated from F9 embryonal carcinoma cells (*6*). This mZP3 isoform was employed to identify O-glycans linked to Ser-332 and Ser-334 as crucial for its ability to inhibit murine sperm−egg binding (*7*, *8*), though these sites are not occupied in native mZP3. One possible explanation for this discrepancy is a shift in the biologically active sperm binding domain to the region of recombinant mZP3 containing Ser-332 and Ser-334. The mutations could also have induced a major change in the glycosylation of the bioactive domain of recombinant mZP3, resulting in the loss of its physiological carbohydrate ligands.

Previous biophysical analyses have defined the potential glycosylation sites in mZP3 from mouse ovaries and recombinant huZP3 (*9*, *18*), and the overall population of O-glycans (*13*), but did not establish the repertoires of glycans at each glycosylation site. By exploiting ultra-high-sensitivity nano-LC−ES-MS methodologies, we have now delineated the O-glycosylation profiles at two highly conserved sites in mZP3 and in huZP3 expressed in murine ovaries. Site-specific presentation of O-glycans in vivo could play a crucial role in mediating binding. Thus, defining glycosylation in both wild-type and huZP3 rescue mice is an essential first step in establishing the requirements for carbohydrate-dependent binding. Despite the paucity of available material, we were successful in acquiring very high-quality MS and MS/MS data that enabled unambiguous analysis of site-specific glycosylation.

Using this approach, both mZP3 and huZP3 produced in mice were shown to express a shared O-glycosylation domain containing two O-glycans spaced five to eight amino acids apart (Figure 11). However, though native mZP3 and huZP3 produced in rescue mice show positional conservation in this region, the profile of O-glycans expressed at the first homologous site (Thr-155/Thr-156) is quite different. Thus, in mZP3, core 2 sequences are primarily expressed at Thr-155, whereas in huZP3, only core 1 and related Tn sequences are present at Thr-156. To our knowledge, this type of site specificity has not been observed when the glycoprotein carrier also expresses core 2 O-glycans at another O-glycosylation site.

Significantly, some of the core 1 O-glycans at Thr-156 of huZP3 have sialic acid in α2−6 linkages to the GalNAc attached to Thr, indicating that the GalNAc is accessible to an α2−6 sialyltransferase but apparently not to the core 2 *N*-acetylglucosaminyltransferase that attaches β1−6-linked GlcNAc to make core 2 sequences. Thus, it is apparent that simple accessibility may not be a factor in this differential processing, but that some other type of regulatory event must occur to restrict core 2 addition at this site. If we bear in mind that there are three sequence differences just upstream of the glycosylation site (Gln/His, Ala/Pro, and Leu/Gln; see Figure 11), one explanation for the positional specificity is that the activity of the core 2 *N*-acetylglucosaminyltransferases that modify mZP3 or huZP3 in the murine egg is influenced by the hydrophobicity and/or conformation of the peptide in the vicinity of the glycosylation site. An alternate explanation is that the sialyltransferase which modifies the 6-position of the GalNAc residue might be influenced by the local peptide sequence. Since this enzyme competes with the core 2 glycosyltransferase, an increase in its activity is likely to inhibit production of core 2 structures.

Table 4: Glycans Attached to the Peptide Backbone ^{162}TTVFSEEK^{169} on huZP3[a]

m/z Observed	M_r	Glycopeptide minus Peptide (M_r)	Predicted Carbohydrate Structure
-	939.5	0	No Carbohydrate Attached
1143.3^{+}	1142.3	202.8	
1305.3^{+} 653.1^{2+}	1304.3	364.8	
1508.3^{+} 754.6^{2+}	1507.3	567.8	
1596.3^{+} 798.7^{2+}	1595.3	655.8	
1612.3^{+} 807.2^{2+}	1611.3	671.8	
1670.3^{+} 836.2^{2+}	1669.3	729.8	
1710.2^{+} 855.6^{2+}	1709.2	769.7	
1832.3^{+} 916.7^{2+}	1831.3	891.8	
937.2^{2+}	1872.4	932.9	
1961.3^{+} 981.2^{2+}	1960.3	1020.8	
1977.3^{+} 989.2^{2+}	1976.3	1036.8	
1018.2^{2+}	2034.4	1094.9	
1038.2^{2+}	2074.4	1134.9	
1083.2^{2+}	2164.4	1224.9	
1091.2^{2+}	2180.4	1240.9	
1099.2^{2+}	2196.4	1256.9	
1184.2^{2+}	2366.4	1426.9	
1192.2^{2+}	2382.4	1442.9	

[a] The M_r values correspond to nonprotonated species. Symbols are as follows: (●) galactose, (■) GalNAc, (□) GlcNAc, (◊) NeuAc, and (♦) NeuGc.

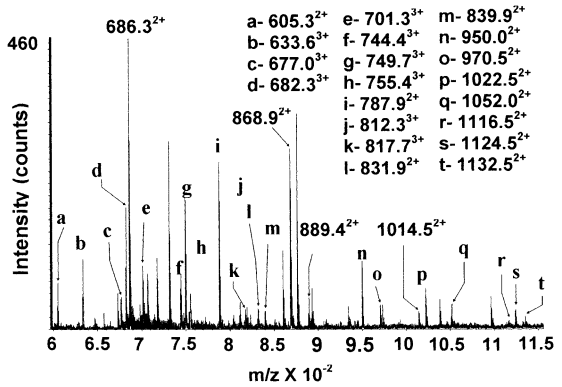


FIGURE 9: MS data derived from components eluting at 7 min from on-line nano-LC−ES-MS of the 5% acetic acid Sep-Pak fraction obtained following trypsin and PNGase F digestion of mZP3. Multiply charged ions correspond to glycans attached to the mZP3 peptide ^{160}RATVSSEEK^{168}. A comprehensive list of the O-glycans at this site and the *m/z* values observed are given in Table 5.

Similar regulatory mechanisms may apply to the huZP glycoproteins.

Another major question is whether this physiological O-glycosylation domain plays a role in initial sperm binding. Because of the data presented in other studies, alternative models for initial murine gamete binding independent of carbohydrate recognition have been presented. One model proposes that mZP1−mZP3 form a supramolecular complex

Table 5: Glycans Attached to the Peptide Backbone ^{160}RATVSSEEK^{168} on mZP3[a]

m/z Observed	M_r	Glycopeptide minus Peptide (M_r)	Predicted Carbohydrate Structure
503.8^{2+}	1005.6	0	No Carbohydrate Attached
605.3^{2+}	1208.6	203	
686.3^{2+}	1370.6	365	
787.9^{2+}	1573.8	568.2	
831.9^{2+}	1661.8	656.2	
839.9^{2+}	1677.8	672.2	
868.9^{2+}	1735.8	730.2	
889.4^{2+}	1776.8	771.2	
950.0^{2+} 633.6^{3+}	1898.0	892.4	
970.5^{2+}	1939.0	933.4	
1014.5^{2+} 677.0^{3+}	2027.0	1021.4	
1022.5^{2+} 682.3^{3+}	2043.0	1037.4	
1052.0^{2+} 701.3^{3+}	2102.0	1096.4	
1116.5^{2+} 744.4^{3+}	2231.0	1225.4	
1124.5^{2+} 749.7^{3+}	2247.0	1241.4	
1132.5^{2+} 755.4^{3+}	2263.0	1257.4	
812.3^{3+}	2433.9	1428.3	
817.7^{3+}	2450.1	1444.5	

[a] The M_r values correspond to nonprotonated species. Symbols are as follows: (●) galactose, (■) GalNAc, (□) GlcNAc, (◊) NeuAc, and (♦) NeuGc.

that mediates initial sperm−egg binding (*12*). Another model suggests that SED1, a sperm plasma membrane glycoprotein homologous to a small group of secreted cell matrix adhesive proteins, mediates binding via a protein−protein interaction (*19*). However, conclusive data supporting either of these newer models have not been obtained (*20*).

By contrast, there is some compelling data indicating that carbohydrate recognition plays a key role in this binding interaction. Although no carbohydrate inhibitor blocks all initial murine sperm−egg binding in vitro, oligosaccharides or artificial carbohydrate constructs terminated with LacNAc (Galβ1−4GlcNAc) inhibit binding of murine sperm to eggs by 70−80% (*21*). In addition, the extent of sperm−egg binding is decreased by 70−75% following the digestion of the mZP with a specific β1−4 galactosidase (*22*). These results indicate that the majority of the natural binding sites for murine sperm apparently require terminal LacNAc sequences.

There is also strong circumstantial evidence indicating that N- or O-glycans preferentially terminated with β1−6-linked LacNAc sequences act as ligands for initial gamete binding in the mouse (*23*). This sequence is expressed on core 2 but not core 1 O-glycans on the mZP. The possibility that vicinal core 2 O-glycans at Ser-332 and Ser-334 could act as a binding domain for murine sperm−egg binding was proposed in a previous study (*23*), but can now be eliminated on the basis of subsequent structural analyses (*9*). In addition, only core 1 O-glycans are linked at Thr-156 of huZP3, yet huZP3 rescue mice retain their fertility and sperm binding phenotype. Thus, the expression of core 2 O-glycans at both sites

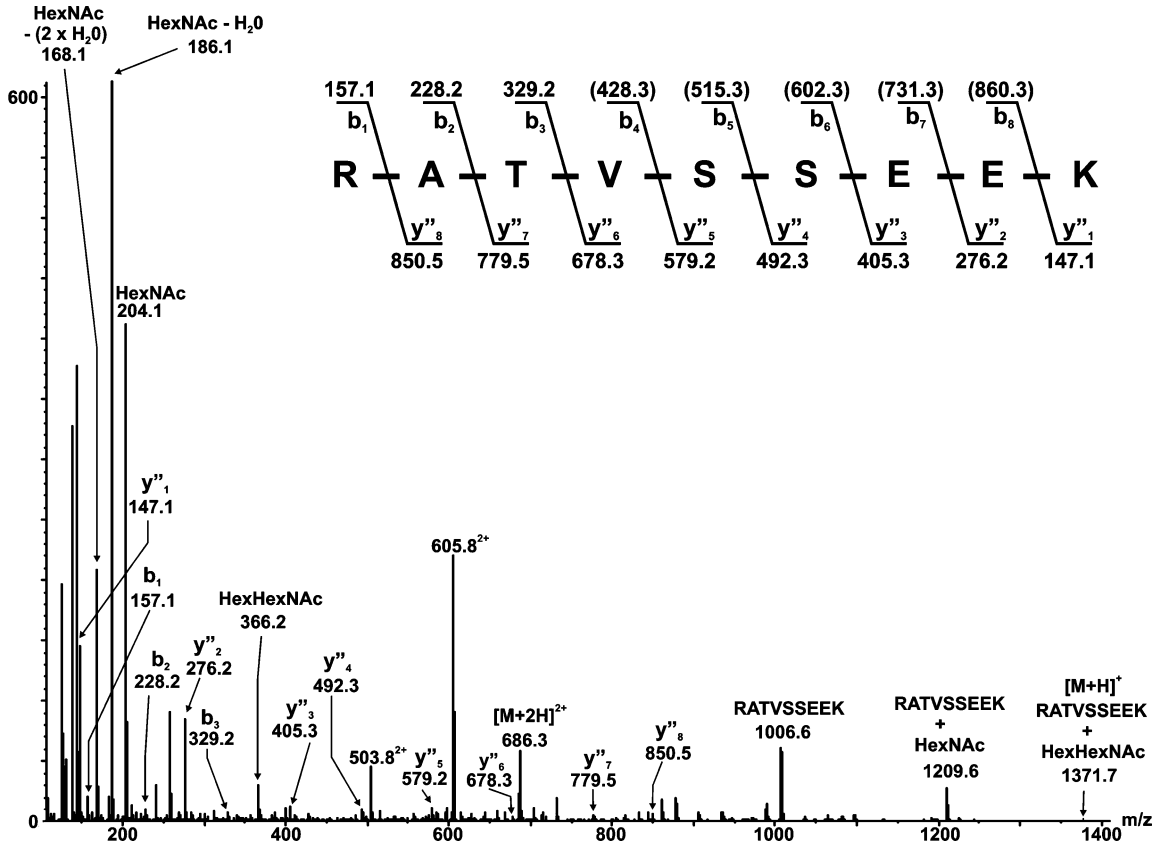


FIGURE 10: CAD ES-MS/MS spectrum of *m*/*z* 686.3^{2+} acquired during the on-line nano-LC−ES-MS/MS experiment of the 5% acetic acid Sep-Pak fraction obtained following trypsin and PNGase F digestion of mZP3. The signals at *m*/*z* 204.1 and 366.2 are characteristic of sugars HexNAc and HexHexNAc, respectively. These are accompanied by signals at *m*/*z* 168.1 and 186.1 which correspond to the loss of two water molecules and one water molecule, respectively, from HexNAc. In addition, the singly charged species at *m*/*z* 1006.6 (RATVSSEEK) and 1209.6 (RATVSSEEK + HexNAc) are also observed as the doubly charged signals at *m*/*z* 503.8^{2+} and 605.8^{2+}, respectively. The peptide sequence assignments are shown in the inset. The numbers in brackets on the assignments correspond to the *m*/*z* values expected but not observed in the CAD ES-MS/MS spectrum of *m*/*z* 686.3^{2+}.

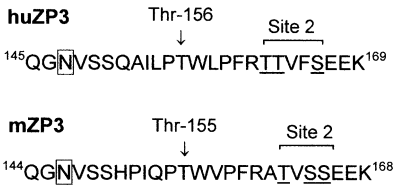


FIGURE 11: Sequences of the conserved O-glycosylation domain of huZP3 (top) and mZP3 (bottom) characterized in this study are shown. In each glycoform, a single O-glycan is linked at Thr-156/155 and another at one of three potential O-glycosylation sites (underlined) at site 2. The glycoforms of mZP3 carry both core 1 and core 2 structures at each site, while huZP3 exclusively carries core 1 glycans at Thr-156 and both core 1 and core 2 structures at site 2. A highly conserved N-glycosylation site N-terminal to this domain is indicated by the boxed asparagine residue. This asparagine is converted to aspartic acid following release of N-glycans from the glycopeptide by PNGase F digestion.

within this conserved O-glycosylation domain is apparently not obligatory for biological activity.

This result suggests several potential hypotheses concerning the role of core 2 O-glycans in murine sperm−egg binding. It is possible that there could be an absolute requirement for the expression of a core 2 O-glycan at only one site, with a core 1 or core 2 O-glycan filling the other site. A sperm binding domain could also involve the obligatory expression of a core 1 O-glycan at one site and a core 2 O-glycan at the other site. This exact specificity will have to be determined empirically.

Another possibility is that there may not be a defined O-glycosylation binding domain that mediates binding, but a lectin-like interaction that recognizes a carbohydrate sequence shared by both N- and O-glycans (*23*). In this model, the oligosaccharides linked to the mZP rather than their absolute positional arrangement may regulate binding. mZP3 may be a more effective inhibitor of binding because it carries more of the physiological ligands required for this interaction.

Regardless of these potential requirements, models involving carbohydrate recognition are more consistent with the results in transgenic mice expressing human glycoproteins. Murine but not human sperm bind to eggs from huZP3, huZP2, and huZP2/huZP3 rescue mice (*12*, *24*), clearly implying that these human glycoproteins acquire mZP glycans when produced in murine ova. Evidence that carbohydrate recognition plays a key role in defining binding specificity was recently obtained in another study utilizing defined mammalian sperm−egg recognition model systems. Recombinant porcine ZP glycoproteins expressing the N-glycan sequences implicated in initial bovine sperm−egg interaction bind to bovine but not porcine sperm (*25*).

This controversy about the role of carbohydrate recognition in murine gamete binding has recently been extended to the human model. huZP apparently consists of four major glycoproteins (huZP1−huZP3 and huZPB) (*26*). A newly developed paradigm suggests that human sperm have evolved

to interact with a matrix composed of all four huZP glycoproteins, an extrapolation of the murine supramolecular model (*12*). However, there is very convincing data indicating that human sperm−ZP binding is primarily carbohydrate dependent (*27*). In addition, it is also evident that rat zona also consists of four homologous glycoproteins, but that human sperm do not bind to rat zona pellucida (*28*). A compelling reason to determine the validity of the carbohydrate- versus non-carbohydrate-mediated models for initial gamete binding in both the murine and human systems now exists.

REFERENCES

1. Mengerink, K. J., and Vacquier, V. D. (2001) Glycobiology of sperm-egg interactions in deuteriostomes, *Glycobiology 11*, 37R−43R.
2. Ohlendieck, K., and Lennarz, W. J. (1996) Molecular mechanisms of gamete recognition in sea urchin fertilization, *Curr. Top. Dev. Biol. 32*, 39−58.
3. Wassarman, P. M. (1999) Mammalian fertilization: Molecular aspects of gamete adhesion, exocytosis, and fusion, *Cell 96*, 175−183.
4. Talbot, P., Shur, B. D., and Myles, D. G. (2003) Cell adhesion and fertilization: Steps in oocyte transport, sperm-zona pellucida interactions, and sperm-egg fusion, *Biol. Reprod. 68*, 1−9.
5. Rankin, T., Familari, M., Lee, E., Ginsberg, A., Dwyer, N., Blanchette-Mackie, J., Drago, J., Westphal, H., and Dean, J. (1996) Mice homozygous for an insertional mutation in the ZP3 gene lack a zona pellucida and are infertile, *Development 122*, 2903−2910.
6. Kinloch, R. A., Mortillo, S., Stewart, C. L., and Wassarman, P. M. (1991) Embryonal carcinoma cells transfected with ZP3 genes differentially glycosylate similar polypeptides and secrete active mouse sperm receptor, *J. Cell Biol. 115*, 655−664.
7. Kinloch, R. A., Sakai, Y., and Wassarman, P. M. (1995) Mapping the mouse ZP3 combining site for sperm by exon swapping and site-directed mutagenesis, *Proc. Natl. Acad. Sci. U.S.A. 92*, 263−267.
8. Chen, J., Litscher, E. S., and Wassarman, P. M. (1998) Inactivation of the mouse sperm receptor, mZP3, by site-directed mutagenesis of individual serine residues located at the combining site for sperm, *Proc. Natl. Acad. Sci. U.S.A. 95*, 6193−6197.
9. Boja, E. S., Hoodbhoy, T., Fales, H. M., and Dean, J. (2003) Structural characterization of native mouse zona pellucida proteins using mass spectrometry, *J. Biol. Chem. 278*, 34189−34202.
10. Schwientek, T., Bennett, E. P., Flores, C., Thacker, J., Hollmann, M., Reis, C. A., Behrens, J., Mandel, U., Keck, B., Schafer, M. A., Haselmann, K., Zubarev, R., Roepstorff, P., Burchell, J. M., Taylor-Papadimitriou, J., Hollingsworth, M. A., and Clausen, H. (2002) Functional conservation of subfamilies of putative UDP-*N*-acetylgalactosamine:polypeptide *N*-acetylgalactosaminyltransferases in *Drosophila*, *Caenorhabditis elegans*, and mammals. One subfamily composed of *l(2)35Aa* is essential in *Drosophila*, *J. Biol. Chem. 277*, 22623−22638.
11. Gerken, T. A., Tep, C., and Rarick, J. (2004) Role of peptide sequence and neighboring residue glycosylation on the substrate specificity of the uridine 5′-diphosphate-α-*N*-acetylgalactosamine: polypeptide *N*-acetylgalactosaminyl transferases T1 and T2: Kinetic modeling of the porcine and canine submaxillary gland mucin tandem repeats, *Biochemistry 43*, 9888−9900.
12. Rankin, T. L., Tong, Z. B., Castle, P. E., Lee, E., Gore-Langton, R., Nelson, L. M., and Dean, J. (1998) Human ZP3 restores fertility in ZP3 null mice without affecting order-specific sperm binding, *Development 125*, 2415−2424.
13. Dell, A., Chalabi, S., Easton, R. L., Haslam, S. M., Sutton-Smith, M., Patankar, M. S., Lattanzio, F., Panico, M., Morris, H. R., and Clark, G. F. (2003) Murine and human ZP3 derived from mouse eggs express identical O-glycans, *Proc. Natl. Acad. Sci. U.S.A. 100*, 15631−15636.
14. Florman, H. M., and Wassarman, P. M. (1985) O-Linked oligosaccharides of mouse egg ZP3 account for its sperm receptor activity, *Cell 41*, 313−324.
15. Bleil, J. D., and Wassarman, P. M. (1988) Galactose at the nonreducing terminus of O-linked oligosaccharides of mouse egg zona pellucida glycoprotein ZP3 is essential for the glycoprotein's sperm receptor activity, *Proc. Natl. Acad. Sci. U.S.A. 85*, 6778−6782.
16. Thall, A. D., Maly, P., and Lowe, J. B. (1995) Oocyte Galα1−3Gal epitopes implicated in sperm adhesion to the zona pellucida glycoprotein ZP3 are not required for fertilization in the mouse, *J. Biol. Chem. 270*, 21437−21440.
17. Liu, D. Y., Baker, H. W., Pearse, M. J., and d'Apice, A. J. (1997) Normal sperm-zona pellucida interaction and fertilization *in vitro* in α1−3-galactosyltransferase gene knockout mice, *Mol. Hum. Reprod. 3*, 1015−1016.
18. Zhao, M., Boja, E. S., Hoodbhoy, T., Nawrocki, J., Kaufman, J. B., Kresge, N., Ghirlando, R., Shiloach, J., Pannell, L., Levine, R. L., Fales, H. M., and Dean, J. (2004) Mass spectrometry analysis of recombinant human ZP3 expressed in recombinant human ZP3 expressed in glycosylation-deficient CHO cells, *Biochemistry 43*, 12090−12104.
19. Ensslin, M. A., and Shur, B. D. (2003) Identification of mouse sperm SED1, a bimotif EGF repeat and discoidin-domain protein involved in sperm-egg binding, *Cell 114*, 405−417.
20. Jungnickel, M. K., Sutton, K. A., and Florman, H. M. (2003) In the beginning. Lessons from fertilization in mice and worms, *Cell 114*, 401−404.
21. Litscher, E. S., Juntunen, K., Seppo, A., Penttila, L., Niemela, R., Renkonen, O., and Wassarman, P. M. (1995) Oligosaccharide constructs with defined structures that inhibit binding of mouse sperm to unfertilized eggs *in vitro*, *Biochemistry 34*, 4662−4669.
22. Mori, E., Mori, T., and Takasaki, S. (1997) Binding of mouse sperm to β-galactose residues on egg zona pellucida and asialofetuin-coupled beads, *Biochem. Biophys. Res. Commun. 238*, 95−99.
23. Easton, R. L., Patankar, M. S., Lattanzio, F. A., Leaven, T. H., Morris, H. R., Clark, G. F., and Dell, A. (2000) Structural analysis of murine zona pellucida glycans. Evidence for the expression of core 2-type O-glycans and the Sd^a antigen, *J. Biol. Chem. 275*, 7731−7742.
24. Rankin, T. L., Coleman, J. S., Epifano, O., Hoodbhoy, T., Turner, S. G., Castle, P. E., Lee, E., Gore-Langton, R., and Dean, J. (2003) Fertility and taxon-specific sperm binding persist after replacement of mouse sperm receptors with human homologs, *Dev. Cell 5*, 33−43.
25. Yonezawa, N., Kudo, K., Terauchi, H., Kanai, S., Yoda, N., Tanokura, M., Ito, K., Miura, K. I., Katsumata, T., and Nakano, M. (2005) Recombinant porcine zona pellucida glycoproteins expressed in Sf9 cells bind to bovine sperm but not to porcine sperm, *J. Biol. Chem. 280*, 20189−20196.
26. Lefievre, L., Conner, S. J., Salpekar, A., Olufowobi, O., Ashton, P., Pavlovic, B., Lenton, W., Afnan, M., Brewis, I. A., Monk, M., Hughes, D. C., and Barratt, C. L. (2004) Four zona pellucida glycoproteins are expressed in the human, *Hum. Reprod. 19*, 1580−1586.
27. Ozgur, K., Patankar, M. S., Oehninger, S., and Clark, G. F. (1998) Direct evidence for the involvement of carbohydrate sequences in human sperm-zona pellucida binding, *Mol. Hum. Reprod. 4*, 318−24.28.
28. Hoodbhoy, T., Joshi, S., Boja, E. S., Williams, S. A., Stanley, P., and Dean, J. (2005) Human sperm do not bind to rat zonae pellucidae despite the presence of four homologous glycoproteins, *J. Biol. Chem. 280*, 12721−12731.

BI0512804