# Human Sperm Binding is Mediated by the Sialyl-Lewis$^x$ Oligosaccharide on the Zona Pellucida

Poh-Choo Pang[1,5], Philip C. N. Chiu[2,5], Cheuk-Lun Lee[2], Lan-Yi Chang[3], Maria Panico[1], Howard R. Morris[1], Stuart M. Haslam[1], Kay-Hooi Khoo[3], Gary F. Clark[4*], William S.B. Yeung[2], Anne Dell[1*]

[1]Division of Molecular Biosciences, Faculty of Natural Sciences, Imperial College London SW7 2AZ, UK

[2]Department of Obstetrics and Gynaecology, Centre for Reproduction, Development and Growth, University of Hong Kong, Pokfulam Road, Hong Kong, China

[3]Institute of Biological Chemistry, Academia Sinica, Taipei, 11529, Taiwan

[4]Department of Obstetrics, Gynecology, and Women's Health, School of Medicine, University of Missouri, Columbia, Missouri 65212, U.S.

[5] These authors contributed equally to the work

*To whom correspondence should be addressed: Email: a.dell@imperial.ac.uk or clarkgf@health.missouri.edu

**Human fertilization begins when spermatozoa bind to the extracellular matrix coating of the oocyte, known as the zona pellucida (ZP). One spermatozoan then penetrates this matrix and fuses with the egg cell, generating a zygote. Although carbohydrate sequences on the ZP have been implicated in sperm binding, the nature of the ligand was unknown. Here, ultrasensitive mass spectrometric analyses revealed that the sialyl-Lewis$^{x}$ sequence (NeuAcα2-3Galβ1-4(Fucα1-3)GlcNAc), a well-known selectin ligand, is the most abundant terminal sequence on the N- and O-glycans of human ZP. Sperm-ZP binding was largely inhibited by glycoconjugates terminated with sialyl-Lewis$^{x}$ sequences or by antibodies directed against this sequence. Thus, the sialyl-Lewis$^{x}$ sequence represents the major carbohydrate ligand for human sperm-egg binding.**

.

Mammalian sperm-egg binding is primarily mediated by the interaction of an egg binding protein (EBP) on the sperm plasma membrane with carbohydrate sequences expressed on glycoproteins of the egg's zona pellucida (ZP) (*1, 2*). Evidence that carbohydrate recognition plays a major role in human gamete binding was initially obtained when the polysaccharide fucoidan was shown to potently block this interaction (*3*). Fucoidan also inhibited leukocyte adhesion in the vascular and lymph systems in the same concentration range that it blocks human sperm-ZP interactions (*3, 4*). The binding of leukocytes to endothelial cells is mediated by C-type lectins known as selectins (*5, 6*). Therefore human sperm-ZP binding was hypothesized to involve a binding specificity that overlaps with the selectins (*7*). However, the structures of the human ZP glycans have remained enigmatic. Furthermore, characterization of these glycans is a challenge because of the scarcity of human eggs and their microscopic size. We have addressed this challenge by employing ultra-sensitive mass spectrometric methodologies for cell and tissue glycomics (*8, 9*). In N- and O-glycans, an antenna is defined as a branch emanating from a "core" structure (*10*). We have defined the structures of the majority of N-and O-glycans attached to human ZP and show that sialyl-Lewis$^x$ is the dominant antenna sequence. This epitope is present at densities that are at least two orders of magnitude higher than levels of selectin-ligand expressed on somatic cells (*11*).

ZP from 195 unfertilized human oocytes was isolated for glycan sequencing. Purity was assessed by proteomics analysis. Human ZP4 was identified as the top hit in the Mascot search (algorithm for Mass Spectral proteomic datasets (Table S1)). The other ZP glycoproteins ZP1, ZP2, and ZP3 were third, fifth, and sixth, respectively, on the Mascot list. At second and fourth position were haptoglobin and transthyretin, which are known constituents in follicular fluid (*12*).

The structures of the N- and O-glycans in the ZP sample were determined by glycomics analysis (*8*). Glycans were analyzed by Matrix Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) Mass Spectrometry (MS) as well as by collisionally activated dissociation (CAD) on a MALDI-TOF-TOF instrument (MS/MS). Mixtures of N- and O-glycans were released from tryptic digests by peptide N-glycosidase F and reductive elimination, respectively, and were permethylated before MS and MS/MS analyses.

The MALDI-TOF N-glycan fingerprint (Fig. 1) showed four families of bi-, tri- and tetra-antennary structures, three of which (shaded in yellow, green and pink) displayed an unusually high density of sialyl-Lewis$^x$ antennae (NeuAcα2-3Galβ1-4(Fucα1-3)GlcNAc). All members of the latter three families were core fucosylated, and each was fully sialylated. Heterogeneity was confined to differences in antenna fucosylation and length. Thus, bi-antennary glycans carried zero, one or two sialyl-Lewis$^x$ antennae (m/z 2966.6, 3140.7/3589.9/3764.0, and 3314.8/3764.0/3938.1, respectively); tri-antennary glycans carried zero, one, two or three sialyl-Lewis$^x$ antennae (m/z 3777.1, 3951.2, 4125.2/4574.8/4748.6, and 4299.3/4748.6/4922.6, respectively); and tetra-antennary glycans carried zero, one, two, three or four sialyl-Lewis$^x$ antennae (m/z 4587.5, 4761.6, 4935.8/5384.9/5559.0, 5109.4/5559.0/5733.1 (major portion), and 5733.1 (minor portion), respectively).

Antennae compositions were defined by CAD-MS/MS. As shown in Fig. 2 and Fig. S1, the most abundant fragment ions arose from cleavage of amino-sugar glycosidic bonds. CAD-MS/MS data from the highest molecular weight species observed in Fig. 1 (m/z 5733), together with its desialylated counterpart (m/z 4287.6, Fig. S2) are shown in the upper and lower panels, respectively, of Fig. 2. These data demonstrated that m/z 5733 was a mixture of tetra-antennary glycans having one extended antenna and three or four sialyl-Lewis$^x$ moieties. The extended antenna were largely composed of the sialyl Lewis$^x$-Lewis$^x$ sequence

(NeuAcα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAc), although a minority of glycans had only a single fucose on this antenna (right cartoon on Fig. 2, lower panel).

Fucose was confirmed to be 3-linked in the sialyl-Lewis$^x$ moiety via its diagnostic elimination in CAD-MS/MS experiments (Fig. 2 and Fig. S1). Linkages involving sialic acid were defined by MALDI analysis of an α2-3-specific neuraminidase digest of the N-glycans (Fig. S2), which showed that α2-6 sialylation was confined to the aforementioned fourth family (Fig. 1, shaded grey), that differs from the other families in having no core fucosylation. The human plasma glycome is characterized by the absence of core fucose plus high levels of α2-6 sialylation (*13, 14*). Therefore we concluded that this fourth glycan family was largely derived from the follicular fluid constituents that co-purify with ZP (Table S1).

An extended sialyl-Lewis$^x$-Lewis$^x$ antenna, and/or its monofucosylated counterpart, which was identified in the m/z 5733 component (Fig. 2), was additionally found in seven other members of the sialyl-Lewis$^x$-containing families (Fig. 1: m/z 3589.9, 3764.0, 3938.1, 4574.8, 4748.6, 5384.9, and 5559.0). These extended sequences were firmly established by MS/MS analyses; examples of diagnostic fragment ions are illustrated in Fig S1 (Panel C).

ZP N-glycans were also investigated to determine if any were sulfated, a modification that is known to play a key role in selectin-mediated leukocyte trafficking (*15, 16*). This study was done after desialylation by using ultrasensitive MS methodologies which have been optimized for sulfo-glycomics (*17*). Only trace levels of sulfated N-glycans were observed (Fig. S3). Their compositions correspond to sulfated counterparts of the core fucosylated glycans shown in Fig. 1.

ZP-associated O-glycans were released from the glycopeptides recovered from the peptide N-glycosidase F digestion, permethylated, and analyzed by MALDI-TOF-TOF. A

limited number of core 1 and core 2 O-glycans were observed (Fig. S4), the latter carrying a single sialyl-Lewis$^x$ epitope. No sulfated O-glycans were detected in the MS experiments. Potential O-sulfation was also investigated by using the MECA-79 antibody, which recognizes 6-sulfated GlcNAc on extended core 1 sequences including those terminated by 6-sulfo sialyl-Lewis$^x$. No immunoreactivity was observed (Fig. S5).

The high density of sialyl-Lewis$^x$ antennae observed on the ZP N-glycans was unusual. Examples of multivalent sialyl-Lewis$^x$ N-glycans are displayed in Fig. 3. This epitope is highly expressed in cells and tissues associated with many human cancers (*18*) and on orosomucoid in the sera of septic shock patients (*19*). However, in healthy humans, where sialyl-Lewis$^x$ plays a vital role in leukocyte trafficking, glycomics studies have suggested that fewer than 1% of the N-glycans carry sialyl-Lewis$^x$ and none has been found to carry more than one sialyl-Lewis$^x$ antenna (*11*). Another feature of the ZP N-glycome was the presence of extended antennae carrying an internal Lewis$^x$ sequence. This structure was found in members of all three families and was a particularly abundant constituent of the tetra-antennary family. This extended sialyl-Lewis$^x$-Lewis$^x$ sequence has previously been found on tumor cells, but not on normal somatic cells (*20*).

The hemizona assay was employed to determine the effect of sialyl-Lewis$^x$ terminated glycoconjugates on sperm-ZP binding (*21*). In this assay, nonliving human eggs were bisected by surgical manipulation, generating two equivalent hemispheres of ZP (hemizona). This test allows for an internally controlled comparison of sperm binding to a matching zona surface. Compared with the untreated controls, the number of sperm bound to the hemizona was significantly decreased after treatment with sialyl-Lewis$^x$-BSA at concentrations of 0.5 μM (p= 3.3e-4), 1 μM (p= 9.4e-6), 1.5 μM (p =3.7e-8) and 2.0 μM (p=1.2e-7). The sialyl-Lewis$^x$ oligosaccharide also significantly decreased binding at concentrations of 100 μM (p= 1.9e-3), 200 μM (p= 3.2e-5) and 500 μM (p= 6.9e-9) (Fig. 4A and Fig. S6). Except for sialyl-

N-acetyllactosamine oligosaccharide at the highest concentration, no significant inhibition was observed with Lewis$^x$, sialyl-N-acetyllactosamine oligosaccharide or BSA conjugates of these sequences (Fig. 4A and Fig. S6). Fluorescently-labeled sialyl-Lewis$^x$-BSA was bound to the head of capacitated spermatozoa (Fig. 4B) but not to the hemizona (Fig. S7). In contrast, Lewis$^x$-BSA, sialyl-N-acetyllactosamine-BSA and BSA were not bound to the sperm head **(Fig. S8—move these panels to the SOM).** These treatments did not affect the acrosomal status and motility of spermatozoa (Fig. S9). Consistently, anti-sialyl-Lewis$^x$ (Fig. 4C), but not anti-Lewis$^x$ antibody (Fig. S5), was bound strongly to ZP and suppressed sperm-ZP binding dose-dependently (Fig. 4D and Fig. S10). To confirm the importance of sialylation, the binding of fluorescence-labeled native and desialylated solubilized ZP to human spermatozoa was compared. Desialylation significantly reduced the binding of solubilized ZP to capacitated spermatozoa (Fig. 4E).

Previous studies indicated that antibodies directed against sialyl-Lewis$^a$, sialyl-Lewis$^x$, and Lewis$^b$ epitopes react with human ZP (*22, 23*). The anti-Lewis$^b$ antibody blocked human sperm-ZP binding in the hemizona assay (*22*). Erythroagglutinating phytohemagglutin also binds to human ZP, indicating that bisecting type N-glycans are expressed on this matrix (*24*). However, the current results have established that only the sialyl-Lewis$^x$ antigen is expressed at physico-chemically confirmable levels on ZP. Based on assessments of signal to noise for detected molecular and fragment ions, it was estimated that other antigens must be substantially less than 1% of the glycome. The biophysical analyses here confirm the expression of carbohydrate sequences. Antibodies and lectins are useful for detecting carbohydrate ligands once their existence is confirmed, but they would not detect the multivalent presentations of sialyl-Lewis$^x$ and sialyl-Lewis$^x$-Lewis$^x$ sequences.

Little is known about how human spermatozoa bind to eggs. The work described here provides insight to the key binding interactions that are essential for natural human

fertilization, supporting the hypothesis that human gamete binding primarily involves the participation of the selectin ligand sialyl-Lewis$^{x}$. Since human spermatozoa do not express selectins (*25*), the major egg binding protein is very likely a lectin with a binding specificity that overlaps with the selectins.

**References and Notes**


1. K. J. Mengerink, V. D. Vacquier, *Glycobiology* **11**, 37R (2001).
2. P. M. Wassarman, L. Jovine, E. S. Litscher, *Nat. Cell Biol.* **3**, E59 (2001).
3. T. T. F. Huang, E. Ohzu, R. Yanagimachi, *Gamete Res.* **5**, 355 (1982).
4. L. M. Stoolman, S. D. Rosen, *J. Cell Biol.* **96**, 722 (1983).
5. L. A. Lasky, *Science* **258**, 964 (1992).
6. M. Fukuda, N. Hiraoka, J. C. Yeh, *J. Cell Biol.* **147**, 467 (1999).
7. M. S. Patankar, S. Oehninger, T. Barnett, R. L. Williams, G. F. Clark, *J. Biol. Chem.* **268**, 21770 (1993).
8. S. J. North *et al.*, *Methods Enzymol.* **478**, 27 (2010).
9. P. C. Pang *et al.*, *J. Biol. Chem.* **282**, 36593 (2007).
10. A. Varki *et al.*, Eds., *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2009)
11. P. Babu et al., *Glycoconj. J.* 26, 975 (2009).
12. F. J. Schweigert, B. Gericke, W. Wolfram, U. Kaisers, J. W. Dudenhausen, *Hum. Reprod.* **21**, 2960 (2006).
13. M. Ferens-Sieczkowska, M. Olczak, *Z. Naturforsch. [C]* **56**, 122 (2001).
14. Y. Mechref *et al.*, *J. Proteome Res.* **8**, 2656 (2009).
15. Y. Imai, L. A. Lasky, S. D. Rosen, *Nature* **361**, 555 (1993).
16. S. D. Rosen, *Annu. Rev. Immunol.* **22**, 129 (2004).
17. K. H. Khoo, S. Y. Yu, *Methods Enzymol* . 478, 3 (2010).
18. J. L. Magnani, *Glycobiology* **1**, 318 (1991).
19. E. C. Brinkman-van der Linden, E. C. van Ommen, W. van Dijk, *Glycoconj. J.* 13, 27 (1996).
20. M. Fukuda et al., *J. Biol. Chem.* 260, 12957 (1985).

21. P. C. Chiu et al., *Biol. Reprod*. 79, 869 (2008).

22. H. Lucas et al., *Hum. Reprod.* 9, 1532 (1994).

23. M. Jimenez-Movilla et al., *Hum. Reprod.* 19, 1842 (2004).

24. M. S. Patankar et al., *Mol. Hum. Reprod.* 3, 501 (1997).

25. G. F. Clark, M. S. Patankar, K. D. Hinsch, S. Oehninger, *Hum. Reprod.* 10 Suppl. 1, 31 (1995).

26. WHO, *World Health Organization Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction* (Cambridge University Press, Cambridge, UK, 1999).

27. P. C. Chiu *et al.*, *Biol. Reprod.* 69, 365 (2003).

28. J. Jang-Lee *et al.*, *Methods Enzymol.* 415, 59 (2006).

29. M. Sutton-Smith, A. Dell, in *Cell Biology: A Laboratory Handbook* J. E. Celis, Ed. (Academic Press, San Diego, 2006), vol. 4, pp. 415-425.

30. P. C. Chiu *et al.*, *J. Biol. Chem.* 280, 25580 (2005).

31. P. C. Chiu *et al.*, *Endocrinology* 151, 3336 (2010).

32. P. C. Chiu *et al.*, *J. Cell Sci.* 120, 33 (2007).

33. J. Mitoma *et al.*, *Glycoconj. J.* 26, 511 (2009).

34. C. L. Lee *et al.*, *J. Biol. Chem.* 284, 15084 (2009).

35. P. C. Chiu *et al.*, *Hum. Reprod.* 23, 1385 (2008).

36. This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC), grant BBF0083091 (to A.D., H.R.M. and S.M.H.); a Royal Society International Joint Project Award (to A.D. and K-H.K.); the Breeden-Adams Foundation and a Life Sciences Mission Enhancement Reproductive Biology Program funded by the State of Missouri (to G.F.C.); and the University Research Committee, University of Hong Kong (to P.C.N. and W.S.B.Y); the proteomics and

sulfoglycomics data were acquired at the NRPGM Core facilities for Proteomics and Glycomics funded by the Taiwan National Science Council (NSC99-3112-B-001-025). A patent for an invention related to this report has been filed.

**Fig. 1**. MALDI-TOF profiling of human ZP N- and O-glycans. **(A)** MALDI-TOF mass spectrum of N-glycans of human ZP. The upper panel covers the m/z range from 2,700-4,500 and the lower, overlapping, panel spans m/z 4,200-5,800. Each panel was normalised so that the most abundant peak is 100% intensity. Structural assignments were based on compositions assigned from molecular weights, complemented by MS/MS information and the results of the neuraminidase digest (Fig. S2). Where more than one structure is shown, the upper structure is the more abundant. Grey structures have antennae carrying α2-6 sialic acid. All the other glycans are exclusively α2-3 sialylated (yellow = bi-antennary; green = tri-antennary; pink = tetra-antennary).

**Fig. 2.** Partial MALDI-TOF-TOF fragment ion spectra obtained after collisional activation of the molecular ion at m/z 5733 in Fig. 1 and, m/z 4288 in Fig. S1, (upper and lower panels, respectively), illustrating how the fragment ions arising from loss of antennae in sialylated and desialylated samples, respectively, defined the antennae sequences. A minus refers to the loss of the designated antenna from molecular ion, Isomeric glycans were identified, which differ in fucose location (see cartoons on lower panel).

**Fig. 3.** Structures of the core fucosylated, multivalent sialy-Lewis$^x$ members of each of the bi-, tri- and tetra-antennary families (shaded yellow, green and pink, respectively) that were found on human ZP.

**Fig. 4.** Sialyl-Lewis$^x$ is involved in sperm-ZP binding. **(A)** Comparison of hemizona binding index (HZI) of capacitated spermatozoa incubated in the presence of sialyl-Lewis$^x$ (SLEX), Lewis$^x$ (LEX) and sialyl-N-acetyllactosamine (SLN) oligosaccharide or their BSA conjugates with medium alone (control). Each point represents the mean±SEM of the results of 10

hemizona assays. *p<0.05 when compared to the corresponding untreated control. †p<0.05 when compared with the desialylated counterpart. **(B)** Representative fluorescent images of capacitated spermatozoa incubated with Alexa Fluor-594 labeled sialyl-Lewis$^x$-BSA. N=5. **(C)** Immunostaining of sialyl-Lewis$^x$ sequences on hemizona. Matching hemizona were incubated with anti-sialyl-Lewis$^x$ (CSLEX1) or anti-6-sulfo lacNAc sequences on extended core 1 O-glycans (MECA-79) or anti-Lewis$^x$ antibodies or preabsorbed antibodies. N=5**. (D)** Effect of CSLEX1 and anti-Lewis$^x$ antibody on sperm-ZP binding in hemizona binding assays. Matching hemizona were incubated with the antibodies or preabsorbed antibody as outlined in the Methods supplement. N=5. *p<0.05 when compared with the control. **(E) Left:** Effect of desialylation of solubilized ZP (1 μg/ml) on sperm-ZP binding in hemizona binding assays. N=5. **Right:** Representative photographs of the binding of native or desialylated ZP to spermatozoa.

**Fig. 1**

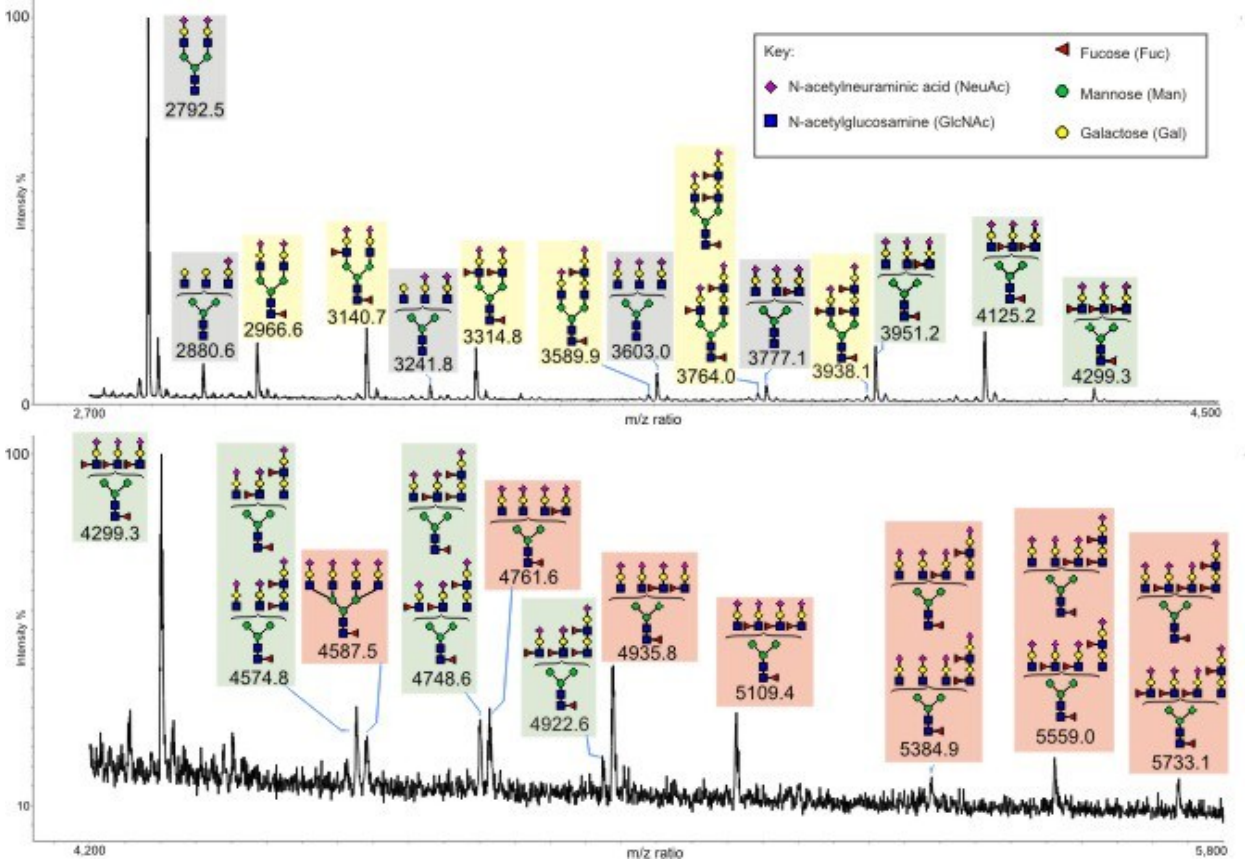
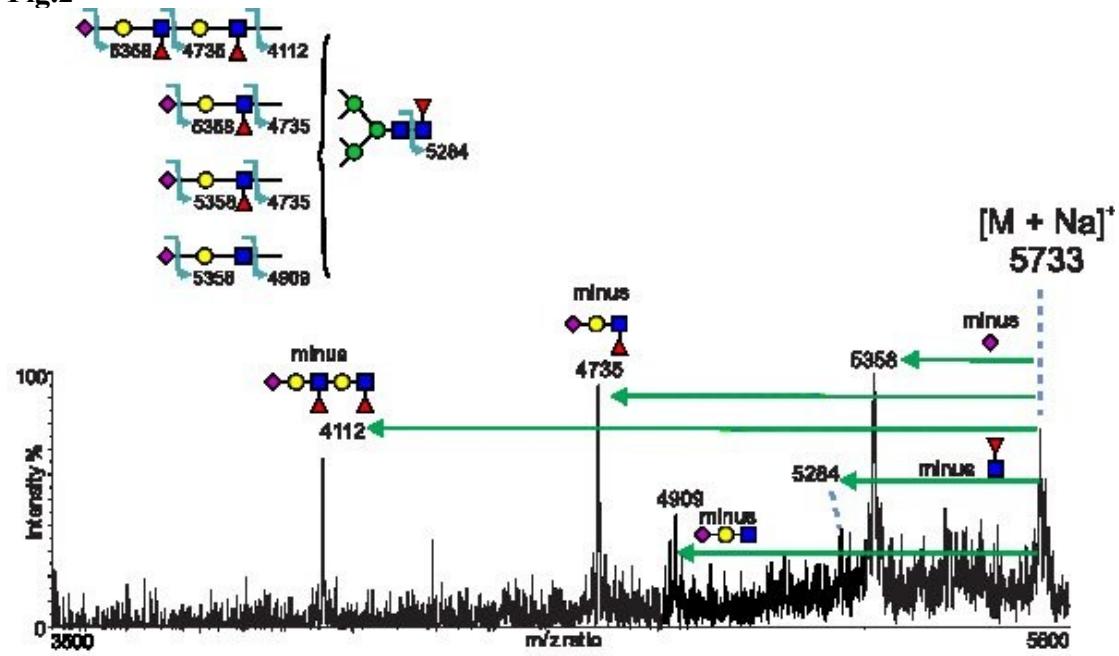

Key:
N-acetylneuraminic acid (NeuAc)
N-acetylglucosamine (GlcNAc)
Fucose (Fuc)
Mannose (Man)
Galactose (Gal)
Intensity %
m/z ratio
2792.5
2880.6
2966.6
3140.7
3241.8
3314.8
3589.9
3603.0
3764.0
3777.1
3938.1
3951.2
4125.2
4299.3
4574.8
4587.5
4748.6
4761.6
4922.6
4935.8
5109.4
5384.9
5559.0
5733.1

**Fig.2**


5358
4735
4112
5358
4735
5284
5358
4735
5358
4909
[M + Na]+
5733
minus
minus
minus
minus
minus
minus
4735
4112
5358
5284
4909
100
Intensity %
0
3600
m/z ratio
5800

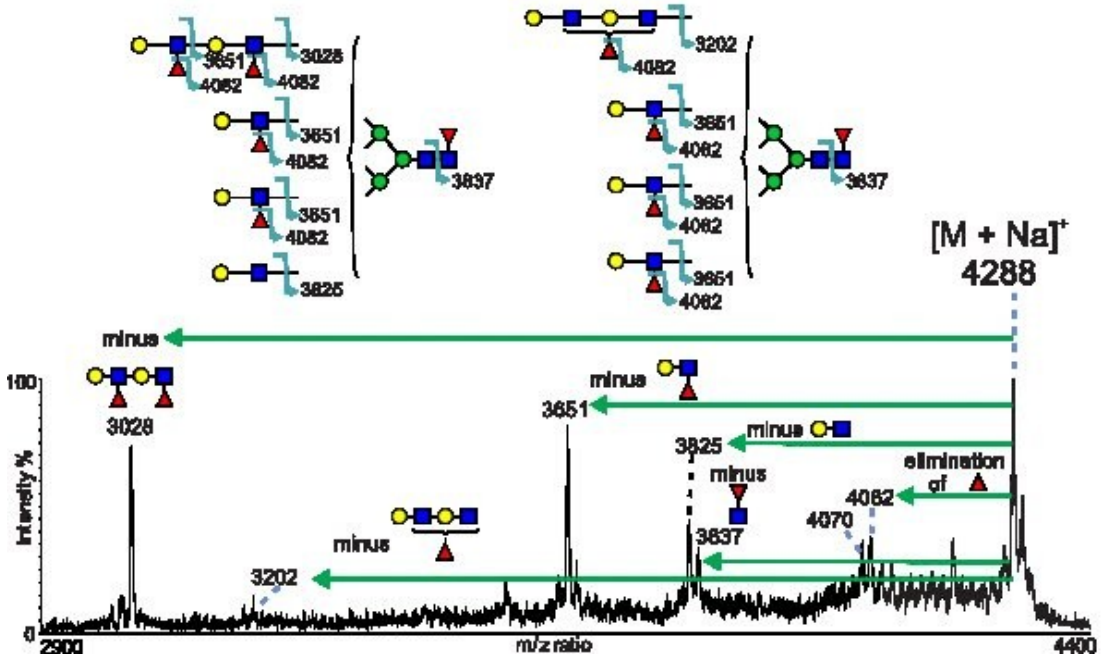

3651
4082
3028
4082
3651
4082
3651
4082
3825
3837
4082
3202
3651
4082
3651
4082
3651
4082
3837
[M + Na]+
4288
minus
minus
minus
minus
minus
minus
3028
3651
3825
3837
3202
elimination
of
4082
4070
100
Intensity %
0
2900
m/z ratio
4400

**Fig 3**

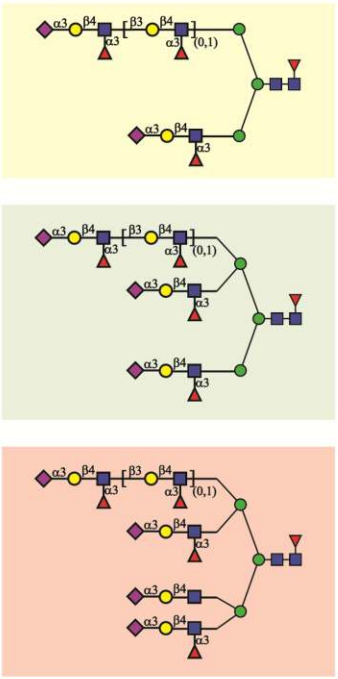
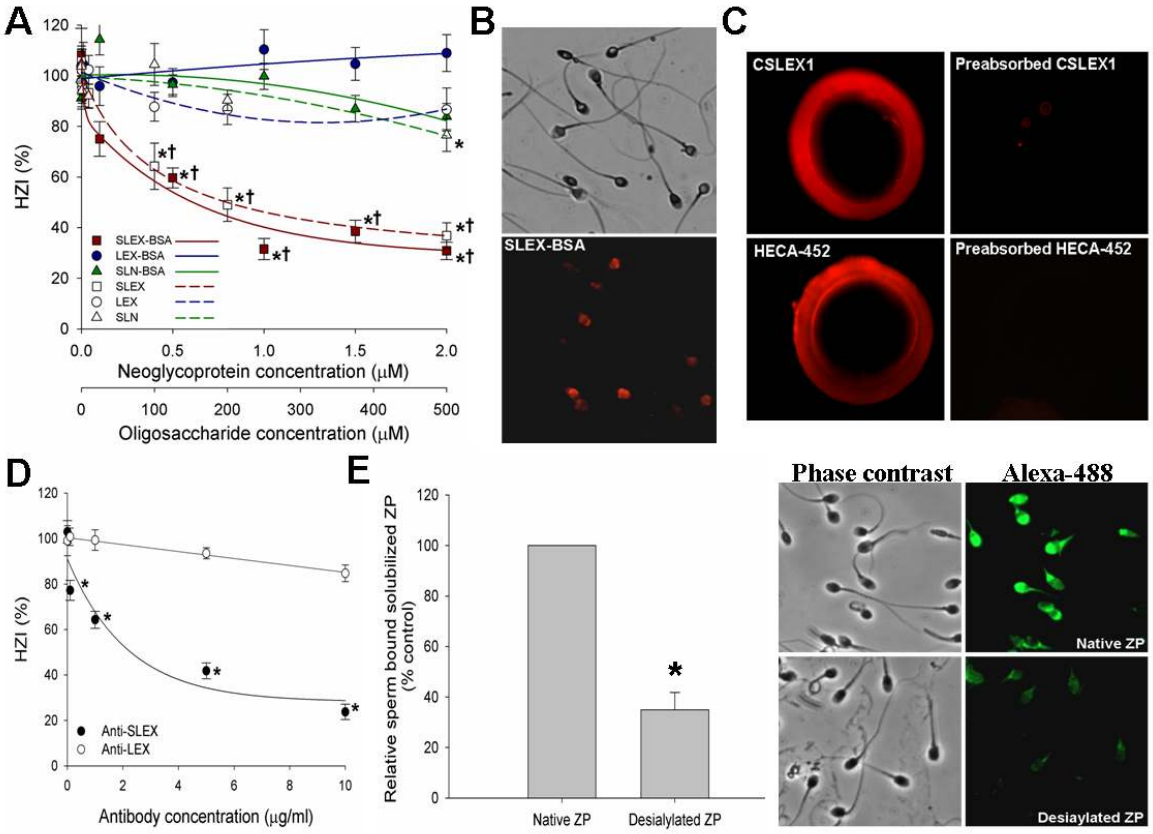

α3
β4
β3
β4
α3
α3
(0,1)
α3
β4
α3
α3
β4
α3
(0,1)
α3
β4
α3
α3
β4
α3
α3
β4
α3
(0,1)
α3
β4
α3
α3
β4
α3
α3
β4
α3

**Fig 4**


A
HZI (%)
SLEX-BSA
LEX-BSA
SLN-BSA
SLEX
LEX
SLN
Neoglycoprotein concentration (μM)
Oligosaccharide concentration (μM)
B
SLEX-BSA
C
CSLEX1
Preabsorbed CSLEX1
HECA-452
Preabsorbed HECA-452
D
HZI (%)
Anti-SLEX
Anti-LEX
Antibody concentration (μg/ml)
E
Relative sperm bound solubilized ZP (% control)
Native ZP
Desialylated ZP
Phase contrast
Alexa-488
Native ZP
Desiaylated ZP


Supporting Online Material for

# Human Sperm Binding is Mediated by the Sialyl-Lewis$^x$ Oligosaccharide on the Zona Pellucida

Poh-Choo Pang, Philip C. N. Chiu, Cheuk-Lun Lee, Lan-Yi Chang, Maria Panico, Howard R. Morris, Stuart M. Haslam, Kay-Hooi Khoo, Gary F. Clark*, William S.B. Yeung, Anne Dell*.

*To whom correspondence should be addressed: a.dell@imperial.ac.uk, clarkgf@health.missouri.edu

**This PDF file includes:**

## Materials and Methods

### Purification of solubilized zona pellucida

Unfertilized oocytes were obtained from the assisted reproduction program at Queen Mary Hospital, Hong Kong. The protocol of the study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster. Informed consent was obtained from patients donated their oocytes for the study. The purification of solubilized ZP was performed as described *(21)*. Briefly, the purification involved the isolation of the ZP from the oocytes under a dissection microscope. The ZPs were then washed and heat-solubilized at 70°C in 5 mM $NaH_2PO_4$ buffer (pH 2.5) for 90 minutes.

### Semen samples

Informed consent was obtained from male donors for semen collection. The protocol for this collection was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster. Spermatozoa from normal semen were processed by density gradient centrifugation on Percoll (Pharmacia, Uppsala, Sweden) *(26, 27)*. After capacitation in Earle's balanced salt solution (EBSS; Flow Laboratories, Irvine, UK) supplemented with sodium pyruvate, penicillin-G, streptomycin sulfate, and 3% bovine serum albumin for 3 hours, the spermatozoa were resuspended in EBSS containing 0.3% BSA (EBSS/BSA).

**Proteomics analysis**

NanoLC was performed on an nanoACQUITY UPLC System (Waters, Milford, USA) coupled to an LTQ-Orbitrap Velos hybrid mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a PicoView nanospray interface (New Objective, Woburn, USA). Peptide mixtures were loaded onto a 75-μm × 250-mm nanoACQUITY UPLC BEH130 column packed with C18 resin (Waters, Milford, USA) and were separated at a flow rate of 300 nl/min using a linear gradient of 5 to 40% solvent B (95% acetonitrile with 0.1% formic acid) in 30 min, followed by a sharp increase to 85% B in 1 min and held at 85% B for another 10 min. Solvent A was 0.1% formic acid in water. The mass spectrometer was operated in the data-dependent mode. Briefly, survey full-scan MS spectra were acquired in the Orbitrap (m/z 350–1600) with the resolution set to 60,000 at m/z 400 and automatic gain control target at 106. The 20 most intense ions were sequentially isolated for CID MS/MS fragmentation and detection in the linear ion trap (automatic gain control target at 5000) with previously selected ions dynamically excluded for 90 s. Ions with single and unrecognized charge states were also excluded. All the measurements in the Orbitrap were performed with the lock mass option for internal calibration.

All MS and MS/MS raw data were processed by Raw2MSM and searched against all entries in Swissprot database (vr 2010_11), or only the human subset in taxonomy, using the Mascot Daemon 2.2 server, with the target-decoy database search option enabled. Search criteria used were: trypsin digestion; variable modifications set as carbamidomethylation (Cys), oxidation (Met) and deamidation (NQ); up to two missed

cleavages allowed; and mass accuracy of 10 ppm for the parent ion and 0.60 Da for the fragment ions. Returned peptide hits were further filtered by the built-in Percolator scoring option with significant threshold set at $p < 0.01$ (peptide ion score > 20), which resulted in a zero peptide false discovery rate.

**Glycan sequencing by MALDI-TOF and MALDI-TOF/TOF**

Purified human ZP were digested using trypsin (Sigma) and purified by reverse-phase Sep-Pak C18 cartridge (Waters Corp) as described *(28)*. The N-glycans were then released by N-glycosidase F (Roche Applied Science) and purified on a Sep-Pak C18 cartridge. The purified native N-glycans were permethylated as described *(29)*, purified using a Sep-Pak C18 cartridge, dissolved in methanol and mixed with 20 mg/mL 2,5-dihydrobenzoic acid in 70% methanol at a 1:1 ratio (v/v). The glycan-matrix mixture (1 μL) was spotted on a stainless steel target plate and dried in vacuum. MALDI-TOF and -TOF/TOF data were obtained using a 4800 MALDI-TOF/TOF mass spectrometer (AB Sciex UK Limited). Argon was used as the collision gas with collision energy of 1 kV. The MS and MS/MS data obtained were analyzed using Data Explorer 4.9. The assignment of glycan sequence was done by manual annotation informed by knowledge of human biosynthetic pathways.

**MALDI-MS screening of sulfated glycans**

A portion of the released native N-glycans was additionally permethylated using the NaOH/dimethyl sulfoxide slurry method for 3 h at 4˚C, followed by careful neutralization

with 5% aqueous acetic acid on ice and then applied directly to a pre-washed and equilibrated C18 Sep-Pak cartridge (Waters), as described *(29)*. For MALDI-MS analyses, the permethylated sample was redissolved in acetonitrile and mixed 1:1 with a 3,4-diaminobenzophenone matrix solution (10 mg/ml in 75% acetonitrile/0.1% trifluoroacetic acid) (Acros Organics) for spotting onto the MALDI target plate. MALDI-TOF MS analyses in negative ion mode were performed on a 4700 Proteomics Analyzer (Applied Biosystems), operated in the reflectron mode.

**Determination of acrosomal status and motility of spermatozoa**

Fluorescein isothiocyanate labeled peanut (*Pisum sativum*) agglutinin (FITC-PSA; Sigma) and Hoechst staining techniques were used to determine the acrosome reaction of spermatozoa *(30)*. The fluorescence patterns of 300 spermatozoa in randomly selected fields were determined under a fluorescence microscope (Zeiss) with 400x magnification. Hobson Sperm Tracker System (Hobson Tracking Systems Ltd) was used to determine the motility of spermatozoa. The procedures and the set-up parameters of the system were described elsewhere *(31)*.

**Hemizona binding assay**

The hemizona binding assay was performed as described previously *(32)*. Unfertilized oocytes were micro-bisected into two identical hemizonae by a micromanipulator. Each hemizona was incubated with $2\times10^6$ capacitated spermatozoa/ml in a 100 μl droplet of EBSS/BSA for 3 hours at 37°C in an atmosphere of 5% $CO_2$ in air

under mineral oil. The numbers of tightly bound spermatozoa on the outer surface of the hemizonae were counted. The hemizona binding index (HZI) was defined as the ratio of the number of bound spermatozoa in the test droplet to that in the control droplet times 100.

**Effects of sialyl-Lewis$^x$-BSA/Lewis$^x$-BSA/sialyl-N-acetyllactosamine neoglycoprotein and sialyl-Lewis$^x$/Lewis$^x$/sialyl-N-acetyllactosamine oligosaccharide**

Hemizona binding assay were performed as described *(31)* in the presence of different concentrations of sialyl-Lewis$^x$/Lewis$^x$/sialyl-N-acetyllactosamine neoglycoprotein (0.01-2 μM) or sialyl-Lewis$^x$/Lewis$^x$/sialyl-N-acetyllactosamine oligosaccharide (Dextra; 0.1-500 μM) to determine their effects on the ZP binding capacity of capacitated spermatozoa. The effect of the neoglycoprotein and oligosaccharide on the acrosomal status, motility and viability of spermatozoa were also determined as described above.

The binding of sialyl-Lewis$^x$-BSA/Lewis$^x$-BSA/sialyl-N-acetyllactosamine-BSA to capacitated spermatozoa and hemizona was visualized by cytochemical staining. Sialyl-Lewis$^x$-BSA/Lewis$^x$-BSA/sialyl-N-acetyllactosamine-BSA (Dextra) was fluorescently labeled with Alexa Fluor-594 microscale fluorescence labeling kit (Invitrogen) according to the manufacturer's protocol. Motile processed spermatozoa ($2\times10^6$ spermatozoa/ml) or hemizona were incubated with 0.5 μM Alexa Fluor-594-labeled sialyl-Lewis$^x$-BSA/Lewis$^x$-BSA/sialyl-N-acetyllactosamine-BSA in an atmosphere of 5% $CO_2$ in air at 37°C for 240 minutes. The treated spermatozoa or hemizona were washed with PBS

containing 0.1% Triton-X 100 and examined under a phase-contrast microscope. Spermatozoa or hemizona incubated with labeled BSA were used as control. Image analysis was performed using Image-Pro Plus (Media Cybernetics).

**Effects of anti-sialyl-Lewis$^x$ and Lewis$^x$ antibodies**

For immunostaining, hemizona were incubated with 0.2 µg/ml of mouse monoclonal anti-sialyl-Lewis$^x$ or anti-Lewis$^x$ antibody (BD) for 3 hours at 37°C. Three anti-sialyl-Lewis$^x$ antibodies with different specificities were used: CSLEX1 binds sialyl-Lewis$^x$ but not 6-sulfo sialyl-Lewis$^x$, MECA-79 only binds to 6-sulfo lacNAc on extended core 1 O-glycans and HECA-452 binds both sialyl-Lewis$^x$ and 6-sulfo sialyl-Lewis$^x$ *(33)*. Matching hemizona treated with irrelevant antibody or antibody preabsorbed by the addition of 1:100 sialyl-Lewis$^x$-BSA or Lewis$^x$-BSA were used as controls. Bound antibodies were detected by Alexa Fluor-594-conjugated goat anti-mouse IgG or anti-rat IgM (Invitrogen).

To determine the effect of antibodies on the ZP binding capacity of capacitated spermatozoa, matching hemizona were pre-incubated either in various concentrations (0.1-10 µg/ml) of anti-sialyl-Lewis$^x$ (CSLEX1)/Lewis$^x$ antibody or preabsorbed antibody at 37°C for 3 hours. The hemizona were then washed with fresh EBSS/BSA. The hemizona binding assays were performed on these treated hemizona as described *(32)*.

**Binding of solubilized zona pellucida to spermatozoa**

Solubilized ZP was desialylated by incubation with sialidase coated agarose beads (Sigma) in 1M Tris-HCl (pH 7) at 37 °C for 18 hours *(34)*. The free sialic acid produced was removed by dialysis with 2 mM Tris-HCl, pH 7.5 at 4°C. The success of desialylation was verified by the decreased binding of the treated solubilized ZP to wheat germ agglutinin which binds strongly to sialylated glycans and weakly to other glycoconjugates. Both native and desialylated solubilized ZP were fluorescently labeled with Alexa Fluor-488 microscale fluorescence labeling kit (Invitrogen).

The binding of native or desialylated solubilized ZP to spermatozoa was performed as described *(35)*. Capacitated spermatozoa ($2\times10^6$/ml) were mildly fixed in 0.5% paraformaldehyde for 10 minutes at room temperature and washed followed by the incubation with 1 μg/ml solubilized ZP at 4°C with slow shaking. After 24 hours, the spermatozoa were washed and the fluorescent signals were quantified using a microplate reader (Dynatech MR5000, Dynatech Laboratories). The results were expressed as percentage of fluorescence intensity relative to the control using native ZP.

**Data Analysis**

All the data were expressed as mean ± standard error of the mean (SEM). The data were analyzed by statistical software packages (SigmaPlot 8.02 and SigmaStat 2.03, Jandel Scientific). For all experiments, the non-parametric repeated measures ANOVA on Rank test for multiple comparisons were used. If the data were normally distributed, Tukey

Test or Parametric Student t-test was used where appropriate as the post-test. A probability value $p<0.05$ was considered to be statistically significant.

**References**


26. WHO, *World Health Organization Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction* (Cambridge University Press, Cambridge, UK, 1999).

27. P. C. Chiu *et al.*, *Biol. Reprod.* 69, 365 (2003).

28. J. Jang-Lee *et al.*, *Methods Enzymol.* 415, 59 (2006).

29. M. Sutton-Smith, A. Dell, in *Cell Biology: A Laboratory Handbook* J. E. Celis, Ed. (Academic Press, San Diego, 2006), vol. 4, pp. 415-425.

30. P. C. Chiu *et al.*, *J. Biol. Chem.* 280, 25580 (2005).

31. P. C. Chiu *et al.*, *Endocrinology* 151, 3336 (2010).

32. P. C. Chiu *et al.*, *J. Cell Sci.* 120, 33 (2007).

32. J. Mitoma *et al.*, *Glycoconj. J.* 26, 511 (2009).

34. C. L. Lee *et al.*, *J. Biol. Chem.* 284, 15084 (2009).

35. P. C. Chiu *et al.*, *Hum. Reprod.* 23, 1385 (2008).

**Fig. S1.**

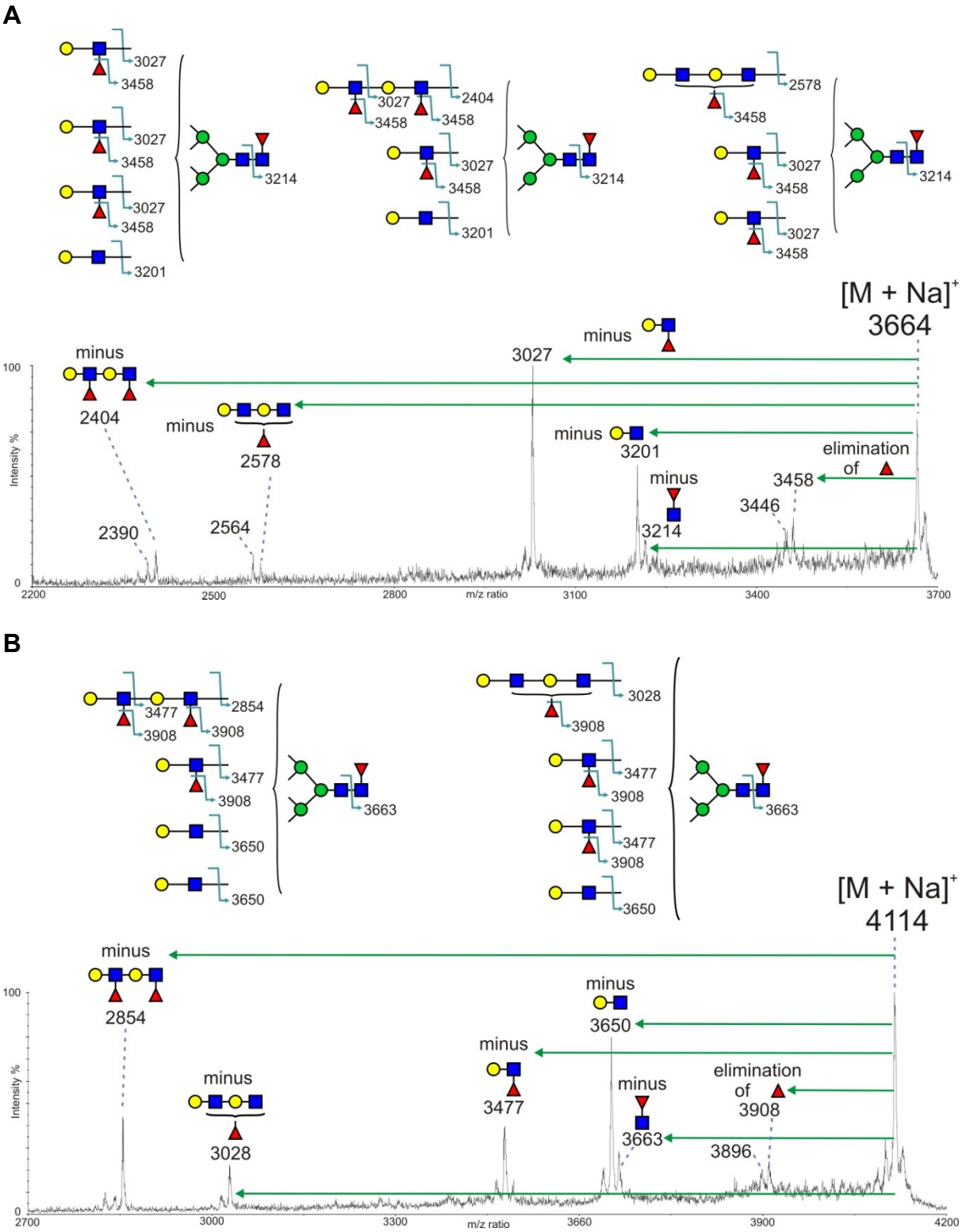

A
3027
3458
3027
3458
3027
3458
3201
3214
3027
3458
2404
3458
3027
3458
3201
3214
2578
3458
3027
3458
3027
3458
3214
[M + Na]+
3664
minus
3027
minus
minus
2404
minus
2578
minus
3201
elimination
of
minus
3458
3446
3214
2390
2564
100
Intensity %
0
2200
2500
2800
m/z ratio
3100
3400
3700
B
3477
3908
2854
3908
3477
3908
3650
3650
3663
3028
3908
3477
3908
3477
3908
3650
3663
[M + Na]+
4114
minus
2854
minus
3650
minus
3477
minus
3028
minus
3663
elimination
of
3908
3896
100
Intensity %
0
2700
3000
3300
m/z ratio
3660
3900
4200

**Fig. S1. (Continued)**

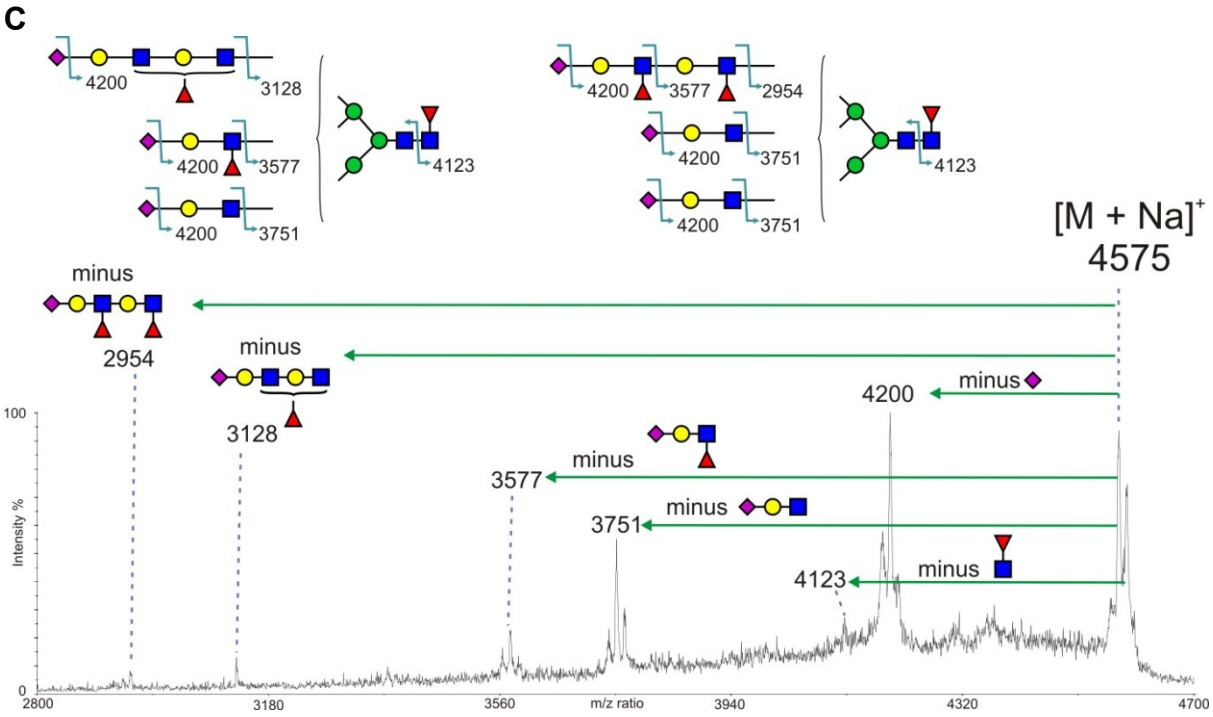



**Fig. S1.** MALDI-TOF-TOF fragment ion spectra obtained after collisional activation of (A) m/z 3664 and (B) m/z 4114 from the desialylated N-glycan sample (see Fig. S2 for MS data), and (C) m/z 4575 from the sialylated N-glycan sample (see Fig. 1a for MS data). These data complement the spectra shown in Fig. 2, further illustrating how the fragment ions define the antennae sequences. The cartoons show the sites of fragmentation within the antennae and the horizontal green arrows show the moieties which are liberated from the molecular ion when each of the fragment ions is formed. Note the diagnostic fragment ions at m/z 3458 (panel A) and 3908 (panel B) which arise specifically from elimination of the 3-linked fucose.

**Fig. S2**

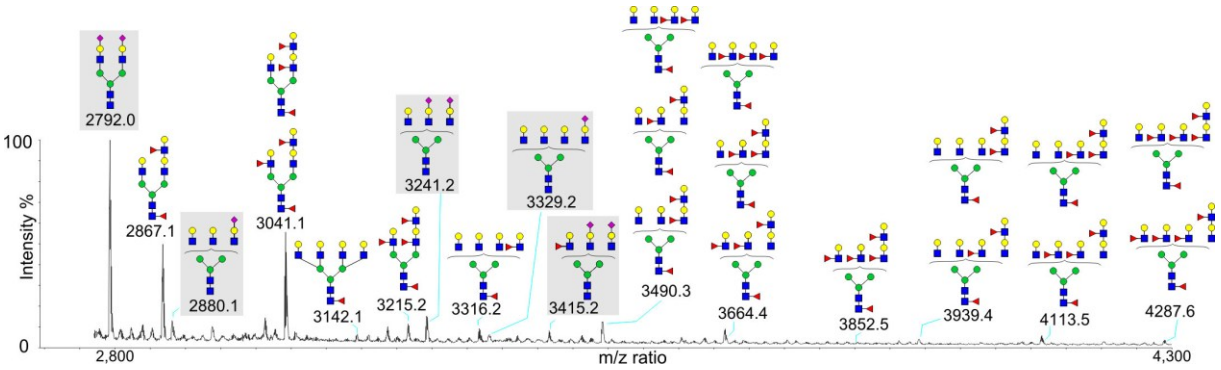



**Fig. S2** MALDI-TOF mass spectrum of N-glycans after α2-3-specific sialidase digestion. Glycans shaded grey have α2-6 sialylated antennae. All sialyl-Lewis$^X$ antennae observed in Fig. 1 have been converted to Lewis$^X$ by the sialidase (see annotations).

**Fig. S3**

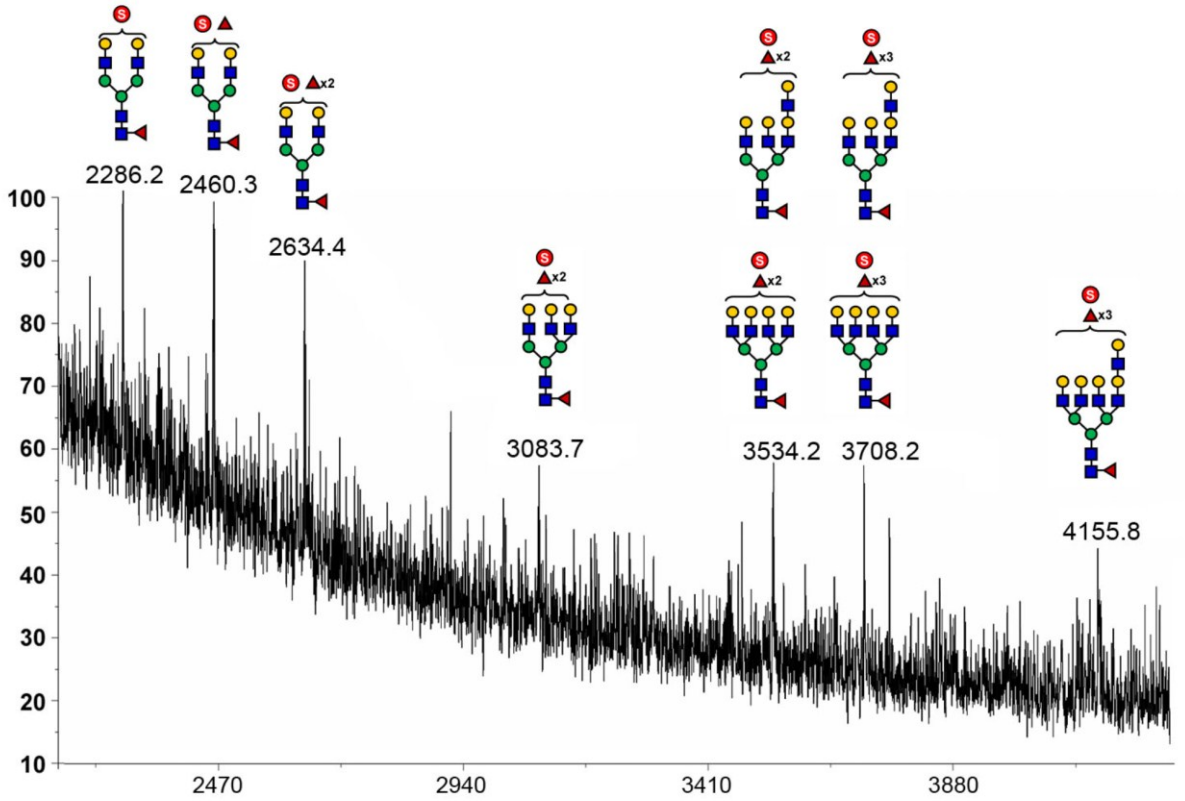



**Fig. S3.** Identification of sulfated N-glycans in the human ZP sample. A portion of the released N-glycans was taken through permethylation and screened by MALDI-MS in negative ion mode for the presence of sulfated glycans as described (15). No significant peak was observed initially but several weak signals were detected and could be assigned as annotated if the sample was first desialylated to reduce the heterogeneity. The desialylated, sulfated N-glycans thus identified correspond in composition to the desialylated, sulfated counterparts of those multiantennary structures carrying multiple sialyl Lewis$^x$ epitopes which are shown in Fig. 1.

**Fig. S4**

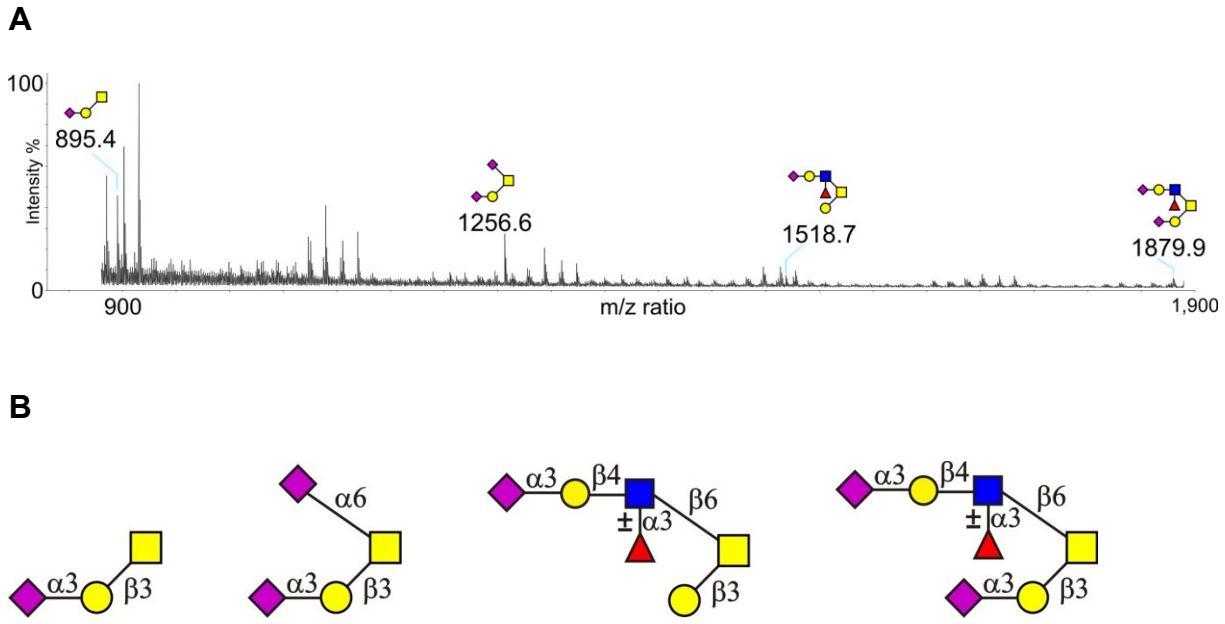



**Fig. S4.** O-glycans found on human ZP. **(**A**)** MALDI-TOF mass spectrum of O-glycans obtained from reductive elimination of residual glycopeptides remaining after N-glycan release from trypsinised human ZP. Glycans were permethylated prior to MS analysis. Structural assignments take into account compositions assigned from molecular weights together with MS/MS-derived sequence information. Unlabelled signals are from the matrix used in the MALDI experiment. **(**B**)** Structures of O-glycans found on human ZP.

**Fig. S5**

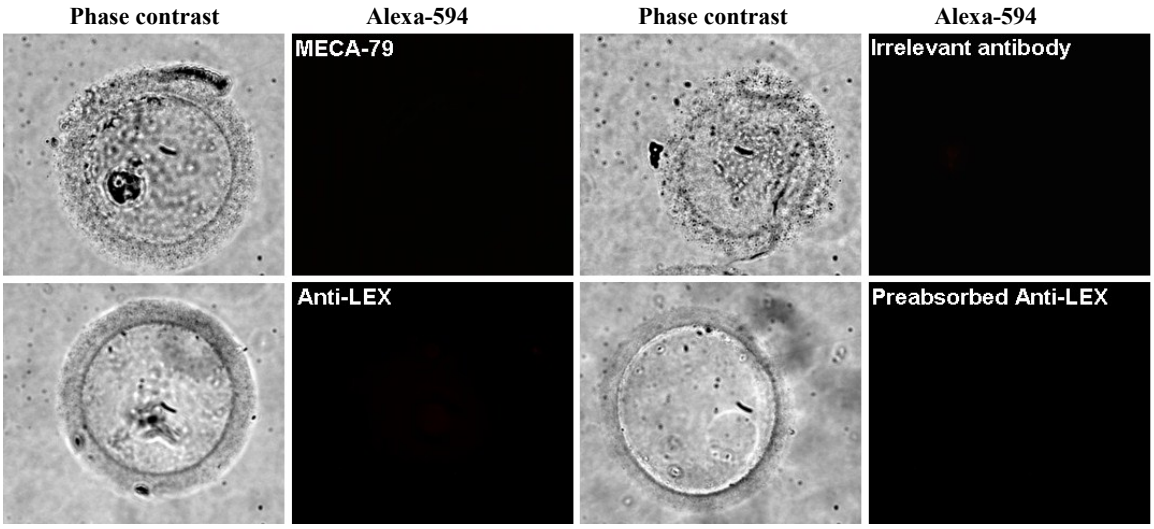



**Fig. S5.** No immunoreactivity was observed on ZP using MECA-79 and anti-Lewis$^x$ antibodies. Matching hemizona were incubated with 0.2 µg/ml anti-6-sulfo sialyl-Lewis$^x$ (MECA-79) or anti-Lewis$^x$ antibodies or antibodies preabsorbed with irrelevant antibody or Lewis$^x$-BSA for 3 hours at 37ºC. The immunoreactivites were visualized using Alexa Fluor 594-conjugated goat anti-mouse IgG or anti-rat IgM. The results shown are representative of 5 replicate experiments**.** LEX: Lewis$^x$

**Fig. S6**

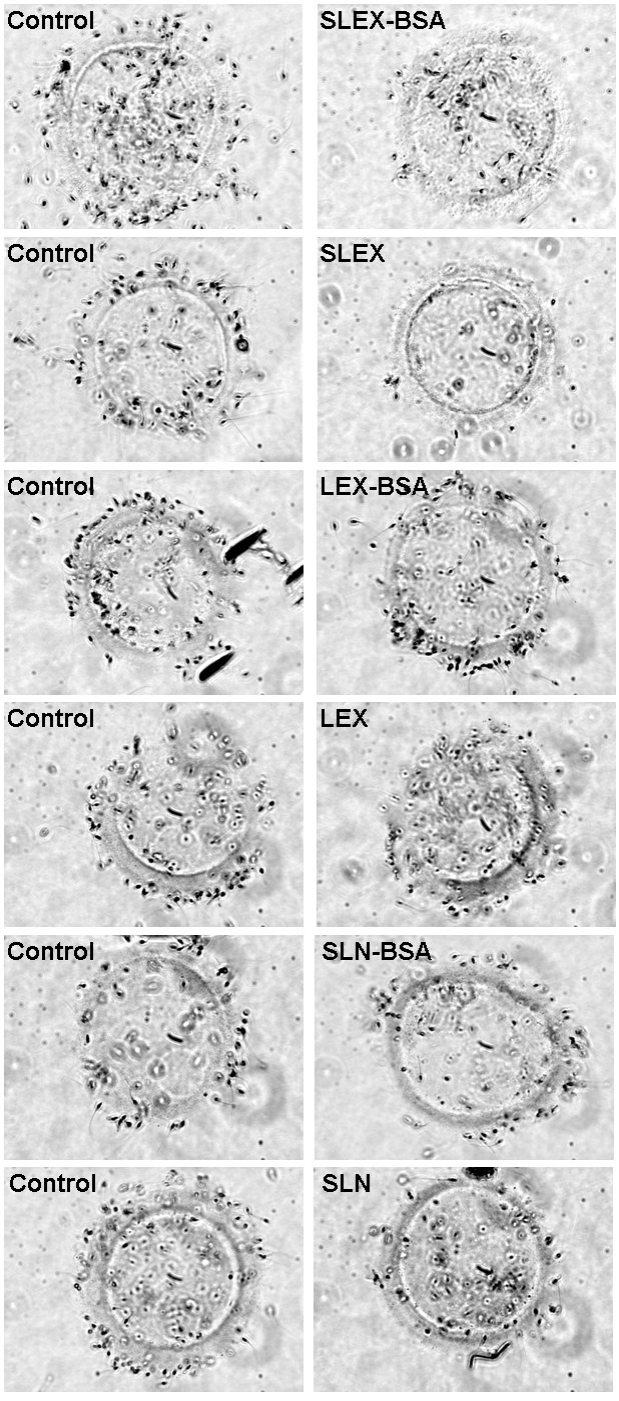

Control
SLEX-BSA
Control
SLEX
Control
LEX-BSA
Control
LEX
Control
SLN-BSA
Control
SLN

**Fig. S6.** Sialyl-Lewis$^x$, but not Lewis$^x$ and sialyl-N-acetyllactosamine, is involved in sperm-ZP binding. Representative photographs of ten replicate experiments showing the binding of capacitated spermatozoa to hemizona in the presence of Sialyl-Lewis$^x$-BSA, Lewis$^x$-BSA and sialyl-N-acetyllactosamine-BSA neoglycoprotein or alternatively Sialyl-Lewis$^x$, Lewis$^x$ and sialyl-N-acetyllactosamine-BSA oligosaccharide with medium alone (control). SLEX: Sialyl Lewis$^x$; LEX: Lewis$^x$; SLN: sialyl-N-acetyllactosamine.

**Fig. S7**

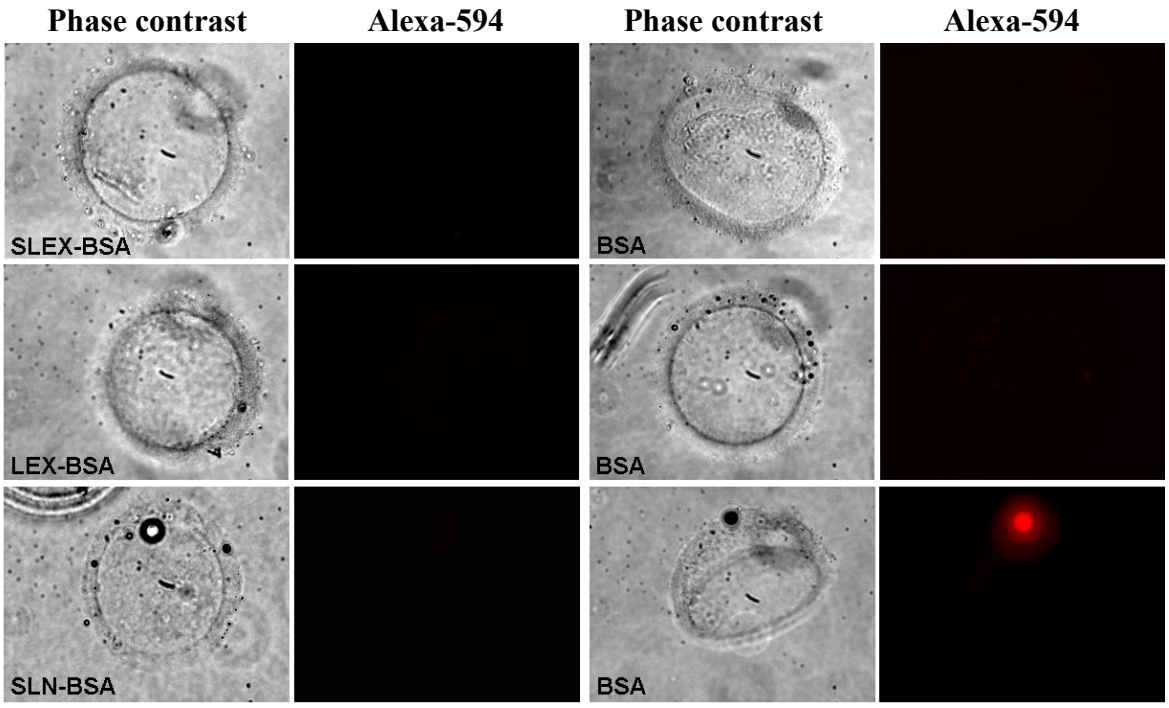
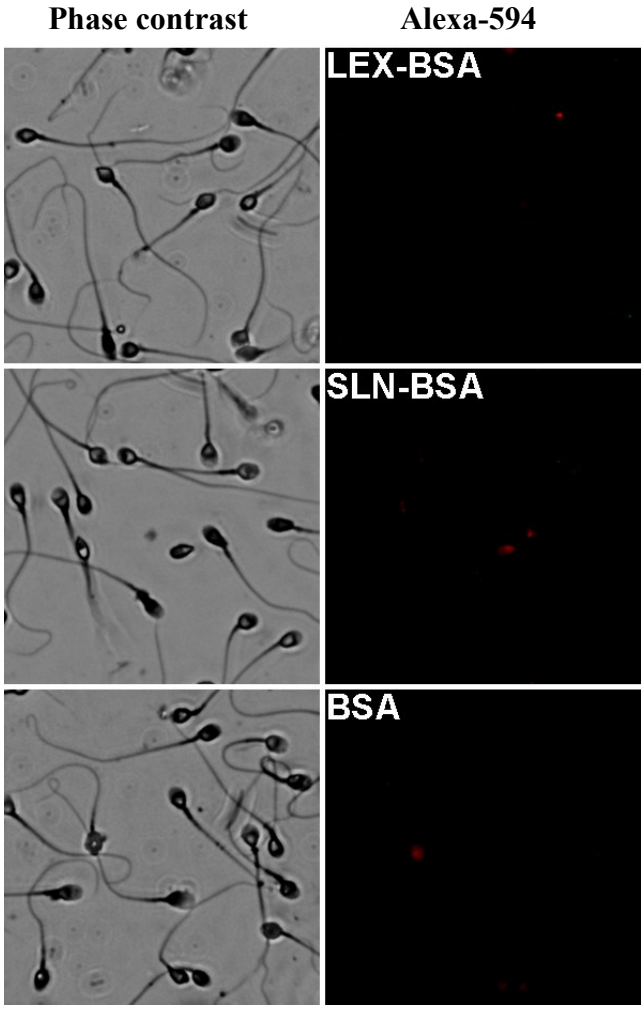



**Fig. S7.** Fluorescently labeled sialyl-Lewis$^x$-BSA/Lewis$^x$-BSA/sialyl-N-acetyllactosamine-BSA did not bind to ZP. Matching hemizona were incubated with 0.5 μM Alexa-594 conjugated sialyl-Lewis$^x$-BSA/Lewis$^x$-BSA/sialyl-N-acetyllactosamine-BSA or BSA for 240 minutes at 37ºC. Fluorescently-labeled sialyl Lewis$^x$/Lewis$^x$-BSA/sialyl-N-acetyllactosamine-BSA and BSA did not bind to the hemizona. The results shown are representative of 5 replicate experiments. SLEX: Sialyl Lewis$^x$; LEX: Lewis$^x$; SLN: sialyl-N-acetyllactosamine.

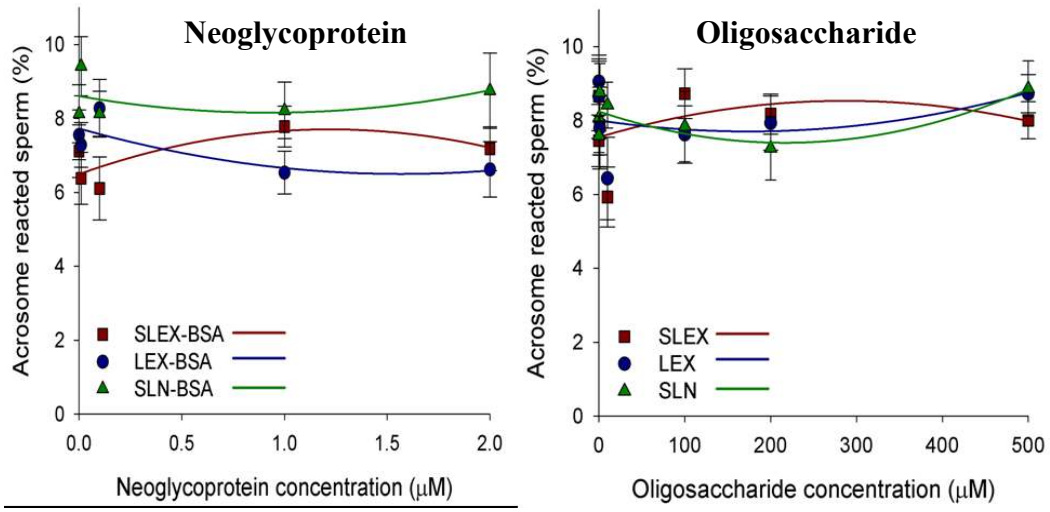
**Fig. S8**




**Supplementary Fig. S8.** Fluorescently labeled sialyl-Lewis$^x$-BSA/Lewis$^x$-BSA/sialyl-N-acetyllactosamine-BSA did not bind to capacitated spermatozoa. Representative fluorescent images of capacitated spermatozoa incubated with 0.5 μM Alexa Fluor-594 labeled sialyl-Lewis$^x$-BSA, Lewis$^x$-BSA, sialyl-N-acetyllactosamine-BSA or BSA for 240 minutes at 37ºC. The results shown are representative of 5 replicate experiments. SLEX: Sialyl Lewis$^x$; LEX: Lewis$^x$; SLN: sialyl-N-acetyllactosamine

**Fig. S9**

**A**




**B**

| Treatment | Sperm motility parameters | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | VAP (μm/s) | VCL (μm/s) | VSL (μm/s) | BCF (Hz) | ALH (μm) | LIN (%) | STR (%) | HYP (%) | Progressive Motility (%) |
| **Control** | 54.6 ± 2.3 | 80.1 ± 4.0 | 42.4 ± 2.7 | 10.4 ± 1.4 | 7.0 ± 1.0 | 53.4 ± 3.8 | 78.2 ± 6.2 | 6.6 ± 1.5 | 48.2 ± 4.3 |
| **Neoglycoprotein** | | | | | | | | | |
| **SLEX-BSA** | 57.6 ± 2.0 | 79.9 ± 5.7 | 40.4 ± 2.7 | 10.4 ± 0.8 | 7.0 ± 1.1 | 50.8 ± 2.0 | 70.4 ± 4.9 | 6.4 ± 1.2 | 55.6 ± 3.3 |
| **LEX-BSA** | 50.8 ± 3.0 | 79.6 ± 3.5 | 42.2 ± 3.4 | 10.6 ± 1.6 | 9.6 ± 1.4 | 53.6 ± 5.4 | 83.0 ± 4.3 | 7.8 ± 1.9 | 56.4 ± 3.9 |
| **SLN-BSA** | 50.2 ± 2.2 | 73.2 ± 3.8 | 41.0 ± 4.2 | 10.0 ± 0.7 | 7.6 ± 1.6 | 57.4 ± 8.3 | 81.8 ± 7.2 | 9.2 ± 2.4 | 47.9 ± 5.9 |
| **Oligosaccharide** | | | | | | | | | |
| **SLEX** | 55.6 ± 3.0 | 76.7 ± 4.5 | 39.0 ± 2.0 | 9.2 ± 1.5 | 8.6 ± 1.5 | 50.8 ± 1.6 | 71.1 ± 6.1 | 8.0 ± 3.0 | 56.7 ± 4.4 |
| **LEX** | 53.6 ± 3.0 | 78.6 ± 4.3 | 41.5 ± 3.7 | 9.0 ± 0.8 | 8.4 ± 1.2 | 53.4 ± 6.3 | 79.8 ± 12.0 | 8.0 ± 0.7 | 55.4 ± 4.1 |
| **SLN** | 54.0 ± 4.6 | 77.8 ± 5.9 | 40.8 ± 4.9 | 9.8 ± 0.8 | 8.6 ± 1.5 | 53.4 ± 7.8 | 79.4 ± 13.2 | 7.4 ± 0.8 | 50.6 ± 2.9 |

**Fig. S9.** Sialyl Lewis$^x$/Lewis$^x$/sialyl-N-acetyllactosamine-BSA and sialyl Lewis$^x$/Lewis$^x$/sialyl-N-acetyllactosamine oligosaccharides have no effect on the acrosomal status and motility parameters of human spermatozoa. (A) The effect of neoglycoproteins (0.01-2 μM) and oligosaccharides (0.1-500 μM) on the acrosomal status of spermatozoa. The fluorescence patterns of 300 fluorescein isothiocyanate-labeled *Pisum sativum* agglutinin-treated spermatozoa in randomly selected fields were determined. Data represent mean ± SEM of 5 separate experiments. (B) The effect of neoglycoproteins (2 μM) and oligosaccharides (500 μM) on motility were determined Hobson Sperm Tracker System. Parameters of spermatozoan motility measured: average path velocity (VAP), curvilinear velocity (VCL), straight line velocity (VSL), beat cross frequency (BCF), amplitude of lateral head displacement (ALH), linearity (LIN; VSL/VCL), straightness (STR; VSL/VAP), percentage hyperactivation (HYP; VCL ≥ 100 μm/s, LIN ≤ 60% and ALH ≥ 5.0 μm) and percentage progressive motility (VAP ≥ 25 μm/s). Data represent the mean ± SEM of five separate experiments using 5 different samples of spermatozoa. SLEX: Sialyl Lewis$^x$; LEX: Lewis$^x$; SLN: sialyl-N-acetyllactosamine.

**Fig. S10**

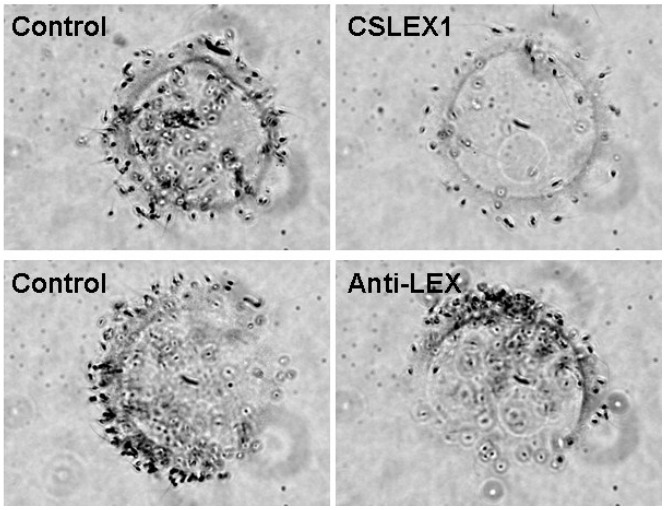



**Fig. S10.** Representative photographs of binding of spermatozoa to hemizona following incubation with anti-sialyl-Lewis$^x$ or Lewis$^x$ antibody. Matching hemizona were incubated with the 10 μg/ml anti-sialyl-Lewis$^x$ (CSLEX1)/Lewis$^x$ antibody or preabsorbed antibody antibodies at 37°C for 3 hours.. The results shown are representative of 5 replicate experiments.

**Table S1.**

| | Protein description | protein score | % protein coverage | charge state | peptide mw (theoretical) | peptide score | peptide sequence | peptide modifications |
|---|---|---|---|---|---|---|---|---|
| 1 | Zona pellucida sperm-binding protein 4 | 1023 | 32.8 | 2 | 1096.5699 | 68.73 | LPCAPSPISR | Carbamidomethyl (C) |
| | | | | 2 | 1101.5567 | 38.33 | EGHFSIAVSR | |
| | | | | 2 | 1263.6459 | 152.61 | AVYENELVATR | |
| | | | | 2 | 1410.7718 | 48.33 | NVTSPPLLLDSVR | Deamidated (NQ) |
| | | | | 2 | 1439.7773 | 140.53 | DPIYVEVSILHR | |
| | | | | 3 | 1439.7773 | 44.45 | DPIYVEVSILHR | |
| | | | | 2 | 1617.6729 | 53.82 | DAPDTDWCDSIPAR | Carbamidomethyl (C) |
| | | | | 2 | 1679.7831 | 70.34 | NYGSYYGVGDYPVVK | |
| | | | | 2 | 1761.8978 | 152.61 | FSIFTFSFVNPTVEK | |
| | | | | 2 | 2093.0252 | 152.61 | GCPYIGDNYQTQLIPVQK | Carbamidomethyl (C) |
| | | | | 4 | 3248.4836 | 77.07 | GPVHLHCSVSVCQPAETPSCVVTCPDLSR | 4 Carbamidomethyl (C) |
| | | | | 3 | 3248.4836 | 152.61 | GPVHLHCSVSVCQPAETPSCVVTCPDLSR | 4 Carbamidomethyl (C) |
| | | | | 3 | 3547.3844 | 44.05 | GDCEGLGCCYSSEEVNSCYYGNTVTLHCTR | 5 Carbamidomethyl (C) |
| 2 | Haptoglobin | 815 | 38.4 | 2 | 919.4552 | 61.88 | GSFPWQAK | |
| | | | | 2 | 979.4876 | 145.46 | VGYVSGWGR | |
| | | | | 2 | 1202.6295 | 152.61 | VTSIQDWVQK | |
| | | | | 2 | 1344.6384 | 56.31 | SCAVAEYGVYVK | Carbamidomethyl (C) |
| | | | | 2 | 1438.6576 | 46.17 | TEGDGVYTLNNEK | |
| | | | | 2 | 1459.6943 | 26.55 | NLFLNHSENATAK | 2 Deamidated (NQ) |
| | | | | 2 | 1722.8069 | 26.04 | YVMLPVADQDQCIR | Carbamidomethyl (C) Oxidation (M) |
| | | | | 3 | 1794.0040 | 51.14 | VVLHPNYSQVDIGLIK | |
| | | | | 3 | 1856.9124 | 144.1 | AVGDKLPECEAVCGKPK | 2 Carbamidomethyl (C) |
| | | | | 2 | 2187.0453 | 152.61 | SPVGVQPILNEHTFCAGMSK | Carbamidomethyl (C) Oxidation (M) |
| | | | | 3 | 2695.3639 | 152.61 | MVSHHNLTTGATLINEQWLLTTAK | Deamidated (NQ) Oxidation (M) |
| 3 | Zona pellucida sperm-binding protein 2 | 780 | 19.9 | 2 | 992.4597 | 32.62 | VMNNSAALR | 2 Deamidated (NQ) Oxidation (M) |
| | | | | 2 | 1114.5117 | 56.69 | DFMSFSLPR | Oxidation (M) |

| | Protein description | protein score | % protein coverage | charge state | peptide mw (theoretical) | peptide score | peptide sequence | peptide modifications |
|---|---|---|---|---|---|---|---|---|
| | | | | 3 | 1271.6081 | 57.08 | FHIPLNGCGTR | Carbamidomethyl (C)<br>Deamidated (NQ) |
| | | | | 2 | 1390.6980 | 34.6 | EITVEFPSSPGTK | |
| | | | | 2 | 1519.7089 | 53.94 | VQMGWSIEVGDGAR | Oxidation (M) |
| | | | | 2 | 1860.8710 | 152.61 | LSPDSPLCSVTCPVSSR | 2 Carbamidomethyl (C) |
| | | | | 2 | 2169.1351 | 152.61 | NDMLLNINVESLTPPVASVK | Oxidation (M) |
| | | | | 3 | 2786.2932 | 152.61 | WHASVVDPLGLDMPNCTYILDPEK | Carbamidomethyl (C)<br>Deamidated (NQ)<br>Oxidation (M) |
| | | | | 3 | 3402.6298 | 115.11 | VIFSSQAICAPDPVTCNATHMTLTIPEFPGK | 2 Carbamidomethyl (C)<br>Deamidated (NQ) |
| 4 | Transthyretin | 658 | 55.1 | 2 | 1365.7517 | 152.61 | GSPAINVAVHVFR | |
| | | | | 2 | 2359.2311 | 152.61 | YTIAALLSPYSYSTTAVVTNPK | |
| | | | | 3 | 2450.1979 | 26 | ALGISPFHEHAEVVFTANDSGPR | |
| | | | | 2 | 2454.1438 | 152.61 | TSESGELHGLTTEEEFVEGIYK | |
| | | | | 2 | 2488.2737 | 152.61 | YTIAALLSPYSYSTTAVVTNPKE | |
| 5 | Zona pellucida sperm-binding protein 3 | 586 | 17.2 | 2 | 1278.5550 | 133.46 | LMEENWNAEK | Oxidation (M) |
| | | | | 2 | 1470.7202 | 152.61 | VTLAEQDPDELNK | |
| | | | | 2 | 2603.2095 | 152.61 | AADLTLGPEACEPLVSMDTEDVVR | Carbamidomethyl (C)<br>Oxidation (M) |
| | | | | 2 | 2603.2095 | 74.5 | AADLTLGPEACEPLVSMDTEDVVR | Carbamidomethyl (C)<br>Oxidation (M) |
| | | | | 3 | 3035.2670 | 152.61 | ACSFSKPSNSWFPVEGSADICQCCNK | 4 Carbamidomethyl (C) |
| 6 | Zona pellucida sperm-binding protein 1 | 570 | 14.1 | 2 | 1611.8257 | 152.61 | FTVATFALLDSGSQR | |
| | | | | 2 | 1963.8557 | 143.56 | EVPCYYGNTATVQCFR | 2 Carbamidomethyl (C) |
| | | | | 3 | 1980.9939 | 50.64 | VDVAQDATLICPKPDPSR | Carbamidomethyl (C) |
| | | | | 2 | 2054.8745 | 152.61 | DETFSSYYGEDDYPIVR | |
| | | | | 3 | 2767.3170 | 150.92 | DYIGTHLSQEQCQVASGHLPCIVR | 2 Carbamidomethyl (C) |
| 7 | Hemopexin | 483 | 15.6 | 2 | 1404.6456 | 143.71 | SWPAVGNCSSALR | Carbamidomethyl (C)<br>Deamidated (NQ) |
| | | | | 2 | 1494.6714 | 109.15 | YYCFQGNQFLR | Carbamidomethyl (C) |
| | | | | 2 | 1499.6755 | 100.67 | EWFWDLATGTMK | Oxidation (M) |
| | | | | 2 | 1711.7624 | 72.32 | GECQAEGVLFFQGDR | Carbamidomethyl (C) |
| | | | | 2 | 2363.1580 | 136.94 | LLQDEFPGIPSPLDAAVECHR | Carbamidomethyl (C) |

| | Protein description | protein score | % protein coverage | charge state | peptide mw (theoretical) | peptide score | peptide sequence | peptide modifications |
|---|---|---|---|---|---|---|---|---|
| 8 | Afamin | 455 | 10 | 2 | 1589.8202 | 84.13 | ESLLNHFLYEVAR | |
| | | | | 2 | 1621.6930 | 152.61 | AESPEVCFNEESPK | Carbamidomethyl (C) |
| | | | | 2 | 1712.9196 | 152.61 | IAPQLSTEELVSLGEK | |
| | | | | 2 | 1886.9302 | 73.35 | SDVGFLPPFPTLDPEEK | |
| 9 | Zinc-alpha-2-glycoprotein | 174 | 9.4 | 2 | 1450.6762 | 81.3 | AYLEEECPATLR | Carbamidomethyl (C) |
| | | | | 2 | 1781.9352 | 112.24 | EIPAWVPFDPAAQITK | |
| 10 | Ferritin light chain | 153 | 8.6 | 2 | 1606.7991 | 152.61 | LGGPEAGLGEYLFER | |
| 11 | Alpha-1-acid glycoprotein 1 | 153 | 7 | 2 | 1741.7981 | 152.61 | EQLGEFYEALDCLR | Carbamidomethyl (C) |
| 12 | Alpha-1B-glycoprotein | 153 | 3.4 | 2 | 1874.9924 | 152.61 | VTLTCVAPLSGVDFQLR | Carbamidomethyl (C) |
| 13 | Major vault protein | 134 | 1.8 | 2 | 1814.9454 | 133.87 | LAQDPFPLYPGEVLEK | |
| 14 | Alpha-1-antitrypsin | 123 | 4.3 | 2 | 1833.8996 | 122.61 | VFSNGADLSGVTEEAPLK | Deamidated (NQ) |
| 15 | N-acetylmuramoyl-L-alanine amidase | 117 | 2.3 | 2 | 1491.7028 | 117.27 | TDCPGDALFDLLR | Carbamidomethyl (C) |
| 16 | Prostaglandin-H2 D-isomerase | 105 | 10 | 2 | 1919.0000 | 105.34 | SVVAPATDGGLNLTSTFLR | Deamidated (NQ) |
| 17 | Polyubiquitin-B | 105 | 7 | 2 | 1786.9200 | 105.17 | TITLEVEPSDTIENVK | |
| 18 | MAP kinase-activating death domain protein | 50 | 3.3 | 4 | 6132.2975 | 49.71 | VDIEVLPQELQPALTFALPDPSRFTLVDFPLHLP LELLGVDACLQVLTCILLEHK | Deamidated (NQ) |
| 19 | Trypsin-1 | 38 | 4 | 2 | 1191.6057 | 38.37 | TLNNDIMLIK | 2 Deamidated (NQ) Oxidation (M) |
| 20 | Peripheral plasma membrane protein CASK | 36 | 1.2 | 2 | 1298.6223 | 36.27 | MNELNHCIVAR | |
| 21 | Hemoglobin subunit beta | 26 | 6.8 | 2 | 1273.7183 | 26.15 | LLVVYPWTQR | |

**Table S1. Shotgun proteomic analysis of the purified human ZP sample**. The protein band corresponding to the purified human ZP sample was subjected to in-gel tryptic digestion and extracted peptides were further de-*N*-glycosylated by PNGase F prior to LC-MS/MS analysis under data dependent acquisition mode. Full experimental conditions and peptide identification criteria were as described in Methods.