**Human Reproduction Vol.19, No.7 pp. 1580–1586, 2004**
Advance Access publication May 13, 2004
DOI: 10.1093/humrep/deh301

# Four zona pellucida glycoproteins are expressed in the human*[1]

**L.Lefièvre[1,2], S.J.Conner[1,2], A.Salpekar[3], O.Olufowobi[2], P.Ashton[4], B.Pavlovic[2], W.Lenton[2], M.Afnan[2], I.A.Brewis[1,2], M.Monk[3], D.C.Hughes[5] and C.L.R.Barratt[1,2,6]**

[1]Reproductive Biology and Genetics Group, Division of Medical Sciences, University of Birmingham, Birmingham B15 2TT, [2]Assisted Conception Unit, Birmingham Women's Hospital, Birmingham B15 2TG, [3]Molecular Embryology Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, [4]School of Chemical Sciences, University of Birmingham, Birmingham B15 2TT and [5]Division of Biology, School of Environmental and Applied Sciences, University of Derby, Derby DE22 1GB, UK

[6]To whom correspondence should be addressed. E-mail: c.l.barratt@bham.ac.uk

The first two authors contributed equally to this work.

**BACKGROUND: The zona pellucida (ZP) is an extracellular glycoprotein matrix which surrounds all mammalian oocytes. Recent data have shown the presence of four human zona genes (ZP1, ZP2, ZP3 and ZPB). The aim of the study was to determine if all four ZP proteins are expressed and present in the human. METHODS: cDNA derived from human oocytes were used to amplify by PCR the four ZP genes. In addition, isolated native human ZP were heat-solubilized, trypsin-digested and subjected to tandem mass spectrometry (MS/MS). RESULTS: All four genes were expressed and the respective proteins present in the human ZP. Moreover, a bioinformatics approach showed that the mouse ZPB gene, although present, is likely to encode a non-functional protein. CONCLUSIONS: Four ZP genes are expressed in human oocytes (ZP1, ZP2, ZP3 and ZPB) and preliminary data show that the four corresponding ZP proteins are present in the human ZP. Therefore, this is a fundamental difference with the mouse model**

*Key words:* human/mouse ZPB/oocyte/proteomics/zona pellucida

## Introduction

Fertilization is a complex process which, to be successful, requires several steps. Critical early events include the binding of sperm to the zona pellucida (ZP) and subsequent induction of the acrosome reaction. The ZP is a matrix composed of glycoproteins which surrounds all mammalian oocytes and mediates several important roles such as binding sperm in a species-specific manner, inducing the acrosome reaction, preventing polyspermy and protecting the embryo prior to implantation. Experiments, primarily in the mouse, have led to the conclusion that the ZP is composed of three proteins (ZP1, ZP2 and ZP3). The murine ZP is the accepted model for ZP structure in higher vertebrates and is composed of repeating units of ZP2–ZP3 heterodimers arranged in filaments cross-linked by ZP1 dimers (Greve and Wassarman, 1985). Experiments in the mouse, using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE)-purified zona proteins showed that ZP3 was the primary sperm ligand (Bleil and Wassarman, 1980a, 1983) and ZP2 was the secondary ligand binding to acrosome-reacted sperm and triggering events that are important for the prevention of polyspermy (Bleil and Wassarman, 1980b; Bleil *et al*., 1981). ZP1 is proposed to contribute to the structural integrity of the ZP matrix (Greve and Wassarman, 1985; Green, 1997; Wassarman, 1999).

Cloning of the ZP genes from a number of species has resulted in the introduction of an alternative nomenclature of ZPA (ZP2), ZPB (ZP1) and ZPC (ZP3) based on gene size (Harris *et al*., 1994). However, it has now become apparent that human ZP1 and ZPB genes are, in fact, paralogues (Hughes and Barratt, 1999), and thus the human genome contains four ZP genes and not three. This is consistent with the identification of both ZP1 and ZPB genes in chicken (Bausek *et al*., 2000) and rat (accession number AF456325), and with recent molecular phylogenetic analysis of vertebrate ZP genes (Conner and Hughes, 2003; Spargo and Hope, 2003). These recent findings show a greater complexity in the number of zona genes across vertebrates than previously anticipated (Conner and Hughes, 2003; Lefièvre *et al*., 2003). In this context it is surprising that the presence of the mouse ZPB gene has not been reported.

Current models for the structure of mammalian zona pellucida are based upon the existence of three ZP proteins

*Part of this work has been presented at ESHRE 2003 held in Madrid [L.Lefièvre *et al*. (2003), Hum. Reprod., 18 (Suppl. 1), 0–266: 91].

 *© European Society of Human Reproduction and Embryology 2004; all rights reserved*

Downloaded from https://academic.oup.com/humrep/article/19/7/1580/2356591 by U.S. Department of Justice user on 16 August 2022

**Table I.** PCR primers used to amplify ZP1, ZP2, ZP3, ZPB and housekeeping genes β-*actin* and *HPRT*

| Gene | Primer sequences 5′–3′[a] | References |
|---|---|---|
| ZP1 | F1 GAAGCCTGTCAGCAGGCTGG | |
| | R1 GCTGTCCCGCGTGATGGAAC | |
| ZP2 | F1 GCCTCCCAGGACCCATTCTC | |
| | R1 CAGGTAGCAGATGGAGCCTA | |
| ZP3 | F1 GAGGCAGCCTCATGTCATG | |
| | R1 AGGCAAAGCCCACTGCTC | |
| ZP3 | F1 CTGCTGCTCTGCAGGTACCATG | |
| | R1 TTTATTCGGAAGCAGACAGAGG | |
| β-*actin* | F1 CGGATGTCCACGTCACACTT | Ponte *et al.* (1984) |
| | R1 GTTGCTATCCAGGCTGTGCT | |
| *HPRT* | F1 AATTATGGACAGGACTGAACGTC | Gibbs *et al.* (1989) |
| | R1 GGCGATGTCAATAGGACTCCAGATG | |
| | R2 CGTGGGGTCCTTTTCACCAGCAAG | |

[a]F1 = first round forward primers.
R1 and R2 = first and second round reverse primers; *HPRT* = hypoxanthine phosphoribosyl transferase.

(Wassarman, 1999). If all four ZP genes are transcribed and translated in the human oocyte, a re-evaluation would be required, both of the structure of the ZP and, potentially, of the mechanisms of sperm–zona interaction. Thus, the aim of the present study was to investigate the ZP gene expression in the human oocyte and the protein composition of the human ZP.

## Materials and methods

### *mRNA purification and preparation of amplified oocyte cDNA in the human*

Amplified cDNA preparations were prepared from human oocytes (four oocytes per sample). The oocytes were lysed in 3 μl of ice-cold lysis buffer (0.8% Igepal, Sigma, UK; 1 IU/μl RNase inhibitor, Gibco BRL, UK; and 5 mmol/l dithiothreitol, Gibco BRL, UK) and the released mRNA molecules bound to oligo-(dT)-linked magnetic beads (Dynabeads; Dynal UK). Synthesis of cDNA on the mRNA bound to the beads, and PCR amplification of the cDNA, were carried out using the SMART™ cDNA library construction kit (Clontech, UK). Further details of the methods are given in Holding *et al*. (2000).

### *PCR amplification of specific gene sequences in amplified cDNA preparations*

Expression of four zona pellucida genes (ZP1, ZP2, ZP3 and ZPB), in the cDNA preparations derived from human oocytes was detected by PCR with the primers given in Table I. Primers for detection of the ubiquitously expressed genes β-*actin* and hypoxanthine phosphoribosyl transferase (*HPRT*) as control genes are also given in Table I. PCR amplification was carried out in total volume of 25 μl of reaction mixture, containing 1 μl of cDNA (concentrations were adjusted by ethidium bromide staining of diluted aliquots compared with standard DNA), 200 μmol/l each of dNTP (Pharmacia), 1 μg each of primers, 1×PCR buffer (Perkin Elmer), using 1.25 IU of AmpliTaq DNA polymerase (Perkin Elmer). With the exception of ZP3, the parameters for the PCR for the zona pellucida genes were one cycle of 95°C for 5 min, followed by 30 cycles of 95°C for 1 min, 57°C for 1 min and 72°C for 1 min. For the detection of expression of the ZP3 gene, the PCR cycling parameters were one cycle of 95°C for 5 min, followed by 30 cycles of 95°C for 1 min, 55°C for 1 min and 72°C for 1 min. PCR cycles for the detection of the housekeeping genes were as follows: β-*actin*, 35 cycles of 95°C for 1 min, 62°C for 1 min and 72°C for 1 min and *HPRT*, 35 cycles of 95°C for 1 min, 60°C for 1 min and 72°C for 1 min. For the hemi-nested PCR for the expression of *HPRT*, 1 μl of the first round product was then transferred to 24 μl of reaction mixture for second round amplification, using the same conditions except for the substitution of the appropriate second round primers (Table I).

### *Gel electrophoresis*

Ten microlitres of PCR product were mixed with 2 μl of loading buffer and electrophoresed for 90 min at 120 V on a 1 or 2% agarose gel and visualized under UV light.

### *Isolation of human ZP*

Metaphase II human oocytes which failed to fertilize after ICSI were donated by patients attending the Assisted Conception Unit, Birmingham Women's Hospital (HFEA centre number 0119). All patients provided informed consent and the research was approved by the local research ethics committee (LREC 0554). Human ZP were mechanically isolated from oocytes. Briefly, several slits were made in the ZP using 10 μm subzonal insemination (SUZI) micropipettes (Conception Technologies, USA). The oocytes were then flushed out from one of the slits by rapidly introducing medium into the perivitelline space from behind the oocyte. The ZP were then washed five times with 50 mmol/l $NH_4HCO_3$, pH 7.2.

### *Solubilization of human ZP*

Five isolated human ZP were solubilized in 300 μl of 50 mmol/l $NH_4HCO_3$, pH 7.2, for 1 h at 70°C. The sample was subsequently concentrated on a 10 kDa cut-off filter, transferred to a siliconized enzyme-linked immunosorbent assay plate and air-dried at 56°C. The dry spot was resuspended in trypsin solution (trypsin solution: 20 μg of trypsin was resuspended in 20 μl of 50 mmol/l acetic acid and then diluted to a final concentration of 12.5 ng/μl with 25 mmol/l $NH_4HCO_3$; sequencing grade modified trypsin, V5111, Promega UK Ltd) and incubated overnight at 37°C. The trypsin-digested human ZP were then resuspended in 6 μl of 1% formic acid and subjected to tandem mass spectrometry (MS/MS).

### *Mass spectrometry*

Direct MS/MS analysis was used to identify the human ZP proteins. This allows direct identification of individual proteins from complex mixtures in very small samples. Lefièvre *et al*. (2003) previously identified the presence of tryptic peptides corresponding to the three known pig ZP proteins (ZP2, ZPB and ZP3) using only 10 isolated porcine ZP.

Tryptic peptides obtained from five isolated human ZP were subjected to MS/MS on a nanoESI Q-Tof mass spectrometer (Q-Tof Ultima GLOBAL; Micromass UK Ltd) following separation of peptides using capillary liquid chromatography (Waters Ltd, UK) with a 15 cm C18 PepMap column (75 μm i.d.×15 cm; Cat. No. 160396, Dionex Ltd, UK). Following MS/MS the raw data were processed using MassLynx 3.5 (Micromass). The resulting tryptic peptide *de novo* sequences data were then compared with non-identical protein sequence databases using MASCOT software (Matrix Science Ltd, UK). Finally, a BLAST search of the GenBank database was performed for each of the peptides obtained to confirm that the proteins identified by MASCOT was the unique match in the non-redundant protein database for a particular peptide sequence.

### *Mouse ZPB gene analysis*

We adopted a bioinformatics approach to identify the mouse orthologue of the rat and human ZPB gene using available sequence data. A BLAST search of the non-redundant database with the rat ZPB

Downloaded from https://academic.oup.com/humrep/article/19/7/1580/2356591 by U.S. Department of Justice user on 16 August 2022

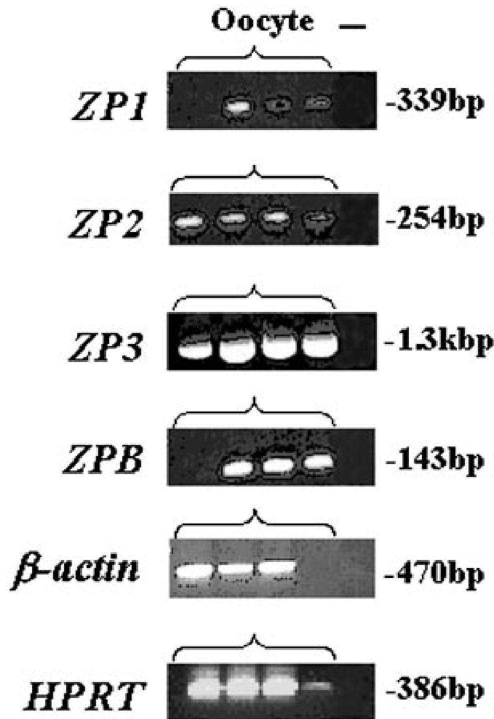


**Figure 1.** Analysis of expression by PCR, of human ZP genes ZP1, ZP2, ZP3, ZPB, and housekeeping genes β-*actin* and *HPRT* in amplified cDNA preparations from four samples of human oocytes. Sizes of expected PCR products are given to the right of the Figure in base pairs (bp). Blank (–): omitting DNA sample to the PCR reaction.

sequence (NM_172330) was used to identify murine genomic sequence with significant sequence similarity.

## Results

### *ZP gene expression in human oocytes*

To determine whether all four ZP genes were expressed, we amplified by PCR cDNA derived from four different samples of four human oocytes. PCR was performed with primers for human ZP1, ZP2, ZP3 and ZPB (Table I) on cDNA preparations from human oocytes. Figure 1 shows that, as expected, all four ZP genes are expressed in the human oocyte. This is the first demonstration that the ZP1 (Hughes and Barratt, 1999) and ZPB genes (Harris *et al*., 1994) are expressed in the human oocyte. Previously only expression of ZP2 and ZP3 had been demonstrated by PCR and serial analysis of gene expression (SAGE) of human oocytes (Neilson *et al*., 2000).

### *Human zona proteomics*

In the present study we have used five human oocytes to investigate the number of distinct glycoproteins in the human ZP. The ZP were mechanically isolated from the oocytes, heat-solubilized, trypsin-digested and analysed using direct MS/MS. Results obtained from peptide sequencing and comparison with non-redundant protein sequence databases confirmed the presence of peptides from all four ZP proteins (ZP1, ZP2, ZP3 and ZPB) (Figure 2). Several tryptic peptide sequences were obtained for ZP2, ZP3 and ZPB (Figure 2), and one for

**Human ZP1**

1 MAGGSATTWGYPVALLLLVATLGLGRWLQPDPGLPGLRHSYDCGIKGMQL
51 LVFPRPGQTLRFKVVDEFGNRFDVNNCSICYHWVTSRPQEPAVFSADYRG
101 CHVLEKDGRFHLRVFMEAVLPNGRVDVAQDATLICPKPDPSRTLDSQLAP
151 PAMFSVSIPQTLSFLPTSGHTSQGSGHAFPSPLDPGHSSVHPTPALPSPG
201 PGPTLATLAQPHWGTLEHWDVNKRDYIGTHLSQEQCQVASGHLPCIVRRT
251 SKEACQQAGCCYDNTREVPCYYGNTATVQCFRDGYFVLVVSQEMALTHRI
301 TLANIHLAYAPTSCSPTQHTEAFVVFYFPLTHCGTTMQVAGDQLIYENWL
351 VSGIHIQKGPQGSITRDSTFQLHVRCVFNASDFLPIQASIFPPPSPAPMT
401 QPGPLRLELRIAKDETFSSYYGEDDYPIVRLLREPVHVEVRLLQRTDPNL
451 VLLLHQCWGAPSANPFQQPQWPILSDGCPFKGDSYRTQMVALDGATPFQS
501 HYQR**FTVATFALLDSGSQR**ALRGLVYLFCSTSACHTSGLETCSTACSTGT
551 TRQRRSSGHRNDTARPQDIVSSPGPVGFEDSYGQEPTLGPTDSNGNSSLR
601 PLLWAVLLLPAVALVLGFGVFVGLSQTWAQKLWESNRQ

**Human ZP2**

1 MACRQRGGSWSPSGWFNAGWSTYRSISLFFALVTSGNSIDVSQLVNPAFP
51 GTVTCDER**EITVEFPSSPGTK**KWHASVVDPLGLDMPNCTYILDPEKLTLR
101 ATYDNCTRRVHGGHQMTIRVMNNSAALRHGAVMYQFFCPAMQVEETQGLS
151 ASTICQK**DFMSFSLPRVFSGLADDSK**GTK**VQMGWSIEVGDGAR**AK**TLTLP**
201 **EAMKEGFSLLIDNHR**MTFHVPFNATGVTHYVQGNSHLYMVSLKLTFISPG
251 QKVIFSSQAICAPDPVTCNATHMTLTIPEFPGKLK**SVSFENQNIDVSQLH**
301 **DNGIDLEATNGMK**LHFSKTLLKTKLSEKCLLHQFYLASLKLTFLLRPETV
351 SMVIYPECLCESPVSIVTGELCTQDGFMDVEVYSYQTQPALDLGTLRVGN
401 SSCQPVFEAQSQGLVRFHIPLNGCGTRYKFEDDK**VVYENEIHALWTDFPP**
451 **SK**ISRDSEFRMTVKCSYSR**NDMLLNINVESLTPPVASVKLGPFTLILQSY**
501 **PDNSYQQPYGENEYPLVR**FLR**QPIYMEVR**VLNRDDPNIKLVLDDCWATST
551 MDPDSFPQWNVVVDGCAYDLDNYQTTFHPVGSSVTHPDHYQRFDMKAFAF
601 VSEAHVLSSLVYFHCSALICNRLSPDSPLCSVTCPVSSRHRRATGATEAE
651 KMTVSLPGPILLLSDDSSFRGVGSSDLKASGSSGEKSRSETGEEVGSRGA
701 MDTKGHKTAGDVGSKAVAAVAAFAGVVATLGFIYYLYEKRTVSNH

**Human ZP3**

1 MELSYRLFICLLLWGSTELCYPQPLWLLQGGASHPETSVQPVLVECQEAT
51 LMVMVSKDLFGTGKLIRAADLTLGPEACEPLVSMDTEDVVRFEVGLHECG
101 NSMQVTDDALVYSTFLLHDPRPVGNLSIVRTNRAEIPIECRYPRQGNVSS
151 QAILPTWLPFRTTVFSEEKLTFSLR**LMEENWNAEK**R**SPTFHLGDAAHLQA**
201 **EIHTGSHVPLR**LFVDHCVATPTPDQNASPYHTIVDFHGCLVDGLTDASSA
251 FKVPRPGPDTLQFTVDVFHFANDSRNMIYITCHLK**VTLAEQDPDELNK**AC
301 SFSKPSNSWFPVEGPADICQCCNKGDCGTPSHSRRQPHVMSQWSRSASRN
351 RRHVTEEADVTVGPLIFLDRRGDHEVEQWALPSDTSVVLLGVGLAVVVSL
401 TLTAVILVLTRRCRTASHPVSASE

**Human ZPB**

1 MWLLRCVLLCVSLSLAVSGQHKPEAPDYSSVLHCGPWSFQFAVNLNQEAT
51 SPPVLIAWDNQGLLHELQNDSDCGTWIRKGPGSSVVLEATYSSCYVTEWD
101 SHYIMPVGVEGAGAAEHKVVTERKLLKCPMDLLARDAPDTDWCDSIPARD
151 RLPCAPSPISRGDCEGLGCCYSSEEVNSCYYGNTVTLHCTREGHFSIAVS
201 RNVTSPPLLLDSVRLALRNDSACNPVMATQAFVLFQFPFTSCGTTRQITG
251 DR**AVYENELVATR**DVKNGSRGSVTRDSIFRLHVSCSYSVSSNSLPINVQV
301 FTLPPPFPETQPGPLTLELQIAKDK**NYGSYYGVGDYPVVK**LLR**DPIYVEV**
351 **SILHR**TDPYLGLLLQQCWATPSTDPLSQPQWPILVKGCPYIGDNYQTQLI
401 PVQKALDLPFPSHHQR**FSIFTFSFVNPTVEK**QALRGPVHLHCSVSVCQPA
451 ETPSCVVTCPDLSRRRNFDNSSQNTTASVSSKGPMILLQATKDPPEKLRV
501 PVDSKVLWVAGLSGTLILGALLVSYLAVKKQKSCPDQMCQ

**Figure 2.** Human ZP1 (Hughes and Barratt, 1999), ZP2 (AAA61335), ZP3 (AAA61336) and ZPB (NP_067009) amino acid sequences. The bold underlined sequences represent the tryptic peptides obtained from the MS/MS analysis.

ZP1 ($^{505}$FTVATFALLDSGSQR$^{519}$) (Figures 2 and 3). This latter tryptic peptide was compared with the human ZP1 peptide sequence identified by Hughes and Barratt (1999). The quindecapeptide amino acid sequence is a perfect match, which indicates the presence of ZP1 in the human ZP.

ZP1 is well recognized as being a low abundance protein (Bleil and Wassarman, 1980b; Epifano *et al*., 1995; Green, 1997). Using the equivalent of several thousand human ZP, a recent MS/MS study on the mouse confirmed the low abundance of ZP1 in that species (Boja *et al*., 2003). Moreover, semiquantitative data in Figure 1 suggest that human ZP1 is less abundant than ZP2, ZP3 or ZPB. Due to the inherent low abundance of ZP1, combined with the paucity of



Downloaded from https://academic.oup.com/humrep/article/19/7/1580/2356591 by U.S. Department of Justice user on 16 August 2022

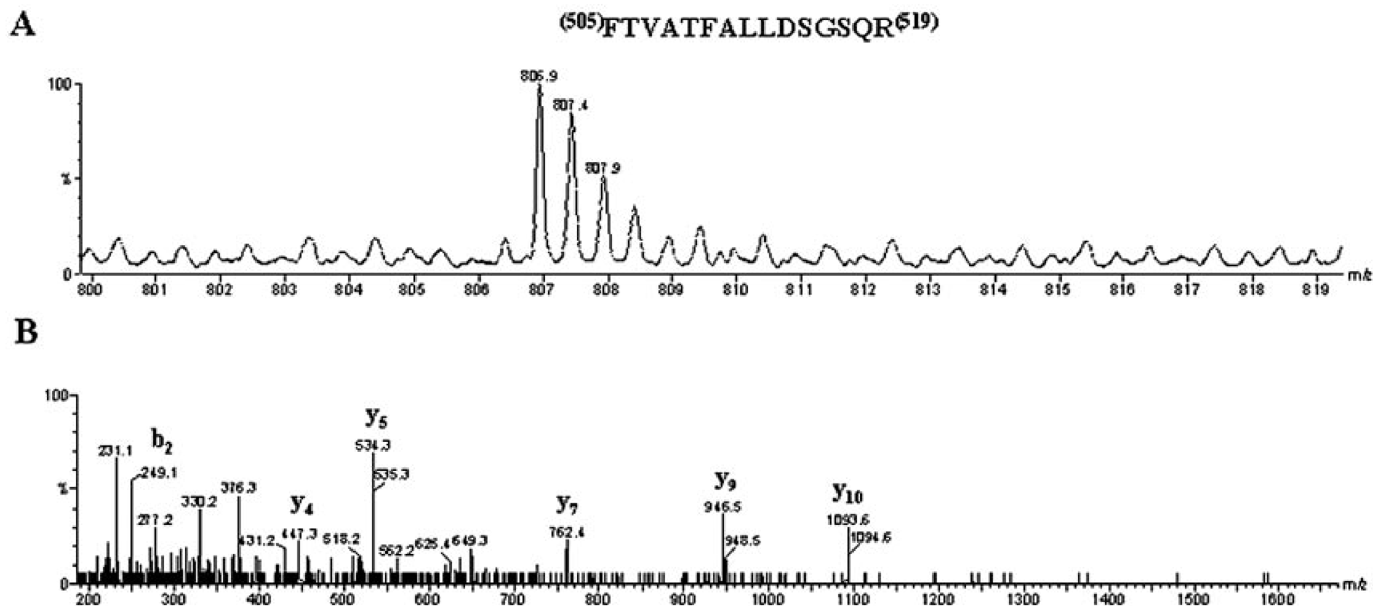


**Figure 3.** Tandem mass spectrometry (MS/MS) spectra of a C-terminal peptide of MH+ 1611.88 Da corresponding to the sequence $^{505}$FTVATFALDSGSQR$^{519}$. (**A**) Reconstructed ion chromatogram MS/MS for ZP1 peptide as a 2+ charged ion at *m/z* 806.95 Da and (**B**) collision-induced dissociation (CID) spectrum which confirms the sequence identity of this peptide through b- and y-ions series of peptide fragments.

human oocytes for research purposes, we have been unable to obtain further peptides for ZP1. Additional experiments have been performed using a variety of preparation methods including enzymatic deglycosylation of the ZP prior to the trypsin digestion as described by Boja *et al*. (2003). Although ZP2, ZP3 and ZPB were identified on each occasion, no peptides for ZP1 were obtained.

### *Existence of mouse ZPB*

The current model of the structure of mouse ZP incorporates three glycoproteins, consistent with the detection, to date, of only three mouse ZP genes (ZP1, ZP2 and ZP3). Having established that a fourth ZP gene is present in the human genome (Hughes and Barratt, 1999) and that this gene is expressed (Figure 1) and protein present (Figure 2) in the human ZP, we undertook analysis of the mouse genome to determine whether a ZPB gene is present.

Using a bioinformatics approach, a putative mouse ZPB cDNA sequence was generated by comparison with rat ZPB sequence, coupled with manual identification of splice sites in the genomic sequence. This mapped to chromosome 13, a region of conserved synteny with human chromosome 1 (1q43) where human ZPB is found and rat chromosome 17 (17q12.1) where rat ZPB is found. However, this hypothetical sequence did not encode a continuous open reading frame. Alignment with the rat ZPB cDNA sequence revealed a number of insertions and deletions which resulted in shifts in the open reading frame and the generation of premature termination codons (Figure 4). To confirm the accuracy of the mouse genomic sequence, individual exons of the putative mouse ZPB gene were used to search the trace sequence database (http://www.ncbi.nlm.nih.gov/blast/tracemb). For each exon screened, the sequence was confirmed, with multiple independent trace sequences of both DNA strands (data not shown).

## Discussion

This paper presents the first evidence for the existence of four glycoproteins in the human ZP. The exact match of the quindecapeptide with the deduced ZP1 coding sequence (Hughes and Barratt, 1999) in addition to the presence of mRNA transcript for ZP1 in four independent experiments strongly supports the premise that four ZP glycoproteins ZP1, ZP2, ZP3 and ZPB are expressed and present in the human. This study also represents the first protein analysis of the human ZP relying on protein sequencing rather than antibody-based approaches.

The difficulty in obtaining a definitive protein identification of ZP1 is primarily due to the low abundance of ZP1 in the mammalian ZP. For example, in the mouse, levels of ZP1 mRNA were four times lower compared to ZP2 and ZP3 (Epifano *et al*, 1995) and consequently only 56% of the ZP1 polypeptide chain was identified by direct MS/MS, compared to 96 and 100% for ZP2 and ZP3 respectively (Boja *et al*., 2003). However, despite its low abundance, experiments using ZP1 null mice demonstrated that ZP1 is an essential element of the ZP primarily required for the structural integrity of the zona matrix, although not involved directly in sperm binding (Rankin *et al*., 1999).

The demonstration of the existence of four human ZP proteins requires a re-interpretation of the numerous electrophoretic studies on the ZP. Although only three diffuse bands were observed on one-dimensional (1D) electrophoresis, the increased resolution provided by two-dimensional (2D)



Downloaded from https://academic.oup.com/humrep/article/19/7/1580/2356591 by U.S. Department of Justice user on 16 August 2022

```
mouse   ATGCTAGGCTAGTGAAGGAGAAAGAATGGCTAAGCAGGCTCTAAGGAGTACTCTGTGGCTTCTGCCAAGCA
rat     ATG-TGGGTTAGTGAAGGAGAAAGGATGGCTAGGCAGGCTCTAAGGAGCACTCTGTGGCTTCTACCAAGCA
        *** * ** *************** ******* *************** ************** *******

mouse   TCTTACTGTGTTTCCCATTCTGTCCTCCCTTGAGTGGCCAG------------------GTGTGCTCCAC
rat     TCTTACTGTGTTTCCCATTCTGTCTTCCTTTGAGTGGCCAGCATGTTACCGAGTTGCCAGGTGTGCTCCAC
        ************************ *** ************                  ************

mouse   TGTGGGTTACAGAGCTTCCAGTTTACTGTGAACCTCAGCCTGGAGGCAGAGAGTCCCATGCTAACAGCTTG
rat     TGTGGGTTACAGAGCTTCCAGTTTGCTGTGAACCTCAGCCTGGAGGCAGAGAGTCCTGTGCTAACAACTTG
        *************3********** ****************************** ' ************

mouse   GGATAGCCAAGGGCTGCCACACAGGCTTAAGAATGACTCTGACTG--GTACATGGGTGATGGACAGAACTG
rat     GGATAGCCAAGGGCTGCCACACAGGCTTAAGAATGACTCTGACTGTGGTACATGGGTGATGGACAGTCCTG
        *********************************************  ****************   ****

mouse   ATGGATTTTTGGTATTGGAAGCCACCCACA---------ATGTCACTCTGGAGGGCTCCCATTATGTCATG
rat     ATGGCTTTCTGGTCTTGGAAGCCAGCTACAGTGGCTGCTATGTCACTCTGGAGGGATCCCACTACATCATG
        **** *** **** ********** * ***         **************** ***** **  *****

mouse   ATGGTCGGCGTGCAAGAGGTAGATGTAGCTGGAAATATGAGAGGGACAAGA--GAGACTGCTTAAGTGCCC
rat     ACCGTTGGTGTGCAAGAGGCAGATGTAGCTGGACATGTAGCAGGAACAAGACAAAGACTGCTTACGTGCCC
        *  ** ** ********** ************ ** *   *** ******   ********** ******

mouse   TTTGGATCTTCACAGTAAGGCCCGAAATACAGCAAGTGCTGAAGTGTGCAGTCCTGTGC-----------
rat     TTTGGCTCTTCAGGGTAAAGCCCCAGATACACCAAATGCTAAAGTGTGCAGTCCTGTGCCAGTAAAGGAAA
        ***** ******  **** **** * ***** *** **** *****************

mouse   -------------TCCCTCGCCCATCTCCAGAGGAAACTGTGAAGAGGTGGTCTGCTGCTACAGCTCTGAA
rat     GGCTGCCCTGTGCTTCCTCGACCATCTCCAGAGGAGACTGTGAAGAGTTGGGGTGCTGCTACAGCTCTGAA
                     * ***** ************** **************   *****************

mouse   GAGGAAAAGGCAGGTTCCTGTTACTATGGAAACACAGTGACCTCCCGTTGTACCAGGGAAGGCTGCTTTTC
rat     GAGGAAGGGGCAGACTCCTGTTACTATGGAAACACAGTGACCTCCCATTGCACCAAGGAGGGCCACTTTTC
        ******  *****  ****************************** *** **** *** ***  ******

mouse   CATTGCTGTGTTCAGGAATGCAACCTCGCCACCCCTACGCTTGGATTCCCTATCCTTGGTCTTCAGGAACA
rat     CATTGCTGTGTCCAGGGACGTGACCTCGCCCCCTCTGCGCTTGGATTCACTTCGCTTAGGCTTCAGGAACA
        *************** * *  ******** * * ** ********** **  **** * ***********

mouse   GCAGT---GGGTGTGATCCTGTGATGATGACATCCACCTTTGTCCTGTTCCAATTTCCACTTACTTCCTGT
rat     TCACCACAGGGTGTGATCCTGTGATGAAAACATCCACCTTTGTCCTGTTCCAATTTCCACTTACTTCCTGT
         **     ******************* ******************************************

mouse   GGGACCACACGGCAGATCACTGGAGACCAGTCCGTGTACAAAAATGAGCTAGTAGTCATTCGGGATGTGCA
rat     GGGACTACACAGCGGATCACTGGAGACCAGGCCATGTATGAAAATGAGCTAGTGGCCATTCGGGATGTACA
        ***** **** ** *************** ** ****  ************* * ************ **

mouse   AGCTTGGGGCAGAAGCTCTATTACCCGATACAGCAACTTCAGTCTCTGAGTCAGCTGTACTTACTCTGCTC
rat     AGCATGGGGCAGAAGCTCTATTACTAGAGACAGCAACTTCAGGCTCCGAGTCAGCTGCACGTACTCCATTC
        **** ********************  ** ************ *** ************* *****   **

mouse   TCAGCAACACATCCCCAATTAACATGCAAGTGCTGGCTCTCCCACCACCCCTTTCTAAGACCCAGCCTGGG
rat     ACAGCATCATGTCCCCAGTTAACATGCAAGTGTGGACTCTCCCACCACCCCTTCCTAAGACCCAGCCCGGA
         ****** **  ****** **************  * ****************** **************

mouse   CCCCTCTCTCTGGAACTTCAGATTGCCAAGGATAAAGGCTATGGTTCTTACTATGGTTCTGATGCCTACCT
rat     CCCCTCTCTCTGGAACTTCAGATTGCCCAGGATAAAAACTACAGCTCTTACTACGGTACTGATGCCTACCC
        *************************** ********  ***  * ******** *** ************

mouse   ACTGGCAAAATTACTCCAGGATCCTTTTTATGTGGAGGTCTCCATCATTCACAGAACAGGCCCCTCACTGG
rat     ACTGGTAAAATTCCTCCAGGATCCCATTTACGTGGAGGTCTCCATCCTTCATAGAACAGACCCCTCGTTGA
        ***** ****** ***********  **** *************** **** ******* ******  **

mouse   TTTCTGCTGCTAGAACAATGTTGGGCCACACCTGGCTCTAATCCTTTTCATCAACCACAATGGCCAATCCT
rat     GC-CTGCTGCTAGAGCAATGTTGGGCTACACCTGGCTCTAATCCTTTTCATCAACCACAATGGCCAATCCT
           *********** *********** ********************************************

mouse   GGTGAAGGGGTGCCCATATGCCGGAGACAACTATCAGACCGAAAGGATCCCTGTCCAGAAAGCATCAAGTC
rat     AGTGAAGGGATGCCCATATGCTGGAGACAATTATCAGACCAAAAGGATCCCGGTTCAGAAAGCATCAGATG
         ******** *********** ******** ********* ********** ** ************  *

mouse   CCTTTCCGTCTCATCACCAGCACTTCAGCAGAGCTACCTTCAGCTTCATGAGTGCTGTAAGGCAGAAGCAG
rat     TCTTTCCATCTCATCATCAGCGCTTCAGTATCTCTACCTTCAGCTTCATGAGTGCTGGAAGGGAGAAGCAG
         ****** ******** **** ****** *   ************************ *************

mouse   GTTTTAAGTGGACAGGTGTACCTGCACTGCAGTGCATCAGTCTGCCAGCCTGCTGGGATGCCATCCTGTGT
rat     GTTTTGGGTGGACAGGTATACCTGCACTGCAGTGCATCGGTCTGCCAGCCTGCTGGGATGCCATCCTGCAC
        *****  ********** ******************** *****************************

mouse   GATAGTCTGCCGTGCTTCCAGGAGAAGAAGAAAATTTGTGCTTCATTTTGAGACCACCACCAGCATATCTA
rat     GGTTATCTGTCCTGCTTCCAGGAGAAGAAGAAAATCTGAGCTTTATTTTGATAACTCCACCAGCATATCTA
        * *  **** * ********************** ** **** ******* * * ***************

mouse   GCAAAGGTCCCATGATCCTCCTCCAAGCCTCTAAGGACTCTGAAGACATGCTTCCTAGACACTCGAGCACC
rat     GCAAAGGTCCCGTGATCCTCCTCCAAGCCACTAAGGACCCTGCAGTCATGCTTCATAAACACTCAGGCACC
        *********** ***************** ******** *** ** ******** ** ******  *****

mouse   CTGGTGGATTCTACTGCTCTGTGAGTAATGGGGCTTTCTGCAACCGTGATCATCATTGGAGTCTTGGTAGT
rat     CATGCAGATTCTCCCACTCTGTGGGTAATGGGACTTTCTGCAAGCATGGTCATCACTGGAGTCTTGGTAGT
        * *  ****** *  ******* ******** ********** * ** ****** ****************

mouse   ATCCTACTTGGCCATCAGAAAATGGAGATAA 1579
rat     ATCCTACTTGGCCACCAGGAAACAGAGATGA 1638
        ************** *** ***  ***** *
```

**Figure 4.** Alignment of rat ZPB mRNA and putative mouse ZPB mRNA produced using CLUSTAL W (Higgins *et al.*, 1994). Predicted termination codons are italicized, underlined and in bold. The published initiation codon for rat ZPB is in bold (NP_758833).

electrophoresis has previously suggested additional levels of complexity to the human ZP (Bercegeay *et al.*, 1995; Moos *et al.*, 1995). This is further complicated by the variable patterns of glycosylation observed. With hindsight, it is apparent that the ZP3-containing band on 1D electrophoresis most likely resolves into two proteins which correspond to ZP3 and ZPB [Yurewicz *et al.*, 1987 (pig); Topper *et al.*, 1997 (cow); Gupta *et al.*, 1998 (human)].



Downloaded from https://academic.oup.com/humrep/article/19/7/1580/2356591 by U.S. Department of Justice user on 16 August 2022

The presence of the ZPB gene in a large number of species (cat, chick, cow, human, macaque, marmoset, pig, possum, rat and rabbit; see Lefièvre *et al.*, 2003) raises the question: does the mouse have a ZPB orthologue? Our bioinformatic analysis demonstrates that, like humans and rats, the mouse has four ZP genes. However, comparative sequence analysis reveals that the mouse ZPB gene has acquired a number of changes making it unlikely that functional ZPB protein will be expressed (Figure 4). This *in silico* evidence is supported by recent data using MS analysis which failed to identify ZPB (Boja *et al.*, 2003). Numerous peptides from ZP1, ZP2 and ZP3 were identified but the authors did not find any unassigned peptides which could correspond to ZPB.

Knockout experiments, in which murine ZP proteins were replaced by human equivalents (Rankin *et al.*, 1998, 2003), showed that, although mouse sperm continued to bind to mouse oocytes engineered to express human ZP2 and ZP3 proteins, human sperm did not bind. One explanation for this result is that mouse oocytes express mouse ZP O-glycans on both mouse and human ZP3 (Dell *et al.*, 2003) and that it is these O-glycans that are key to sperm binding (Wassarman, 1999). Another complementary explanation, derived from our results, is that human sperm have evolved to interact with a ZP composed of four, not three, proteins, with ZPB as well as ZP3 being required for sperm zona binding. This explanation is supported by data from other species, e.g. rabbit, bovine, porcine and macaque, which show that ZPB has sperm-binding properties (Prasad *et al.*, 1996; Topper *et al.*, 1997; Yurewicz *et al.*, 1998; Govind *et al.*, 2000). Thus, although the mouse ZP may support the hypothesis that the mammalian ZP has three proteins of which ZP3 is the primary sperm receptor, it may fall short as a hypothesis for sperm–zona interaction in other species. Consequently it appears that the model species for the study of mammalian fertilization may only serve as a good model for itself.

## Acknowledgements

We would like to express sincere thanks to Drs Bayard Storey and Stephen Publicover for critical reading of this manuscript. The authors also wish to thank the patients who participated in this investigation. This study was supported by Fonds québécois de la recherche sur la nature et les technologies (L.L.), Medical Research Council ROPA grant (C.L.R.B.), The Lord Dowding Fund for Humane Research (C.L.R.B. and I.A.B.), the Birmingham Women's Hospital (C.L.R.B.), Biotechnology and Biological Sciences Research Council (M.M., A.S. and the Functional Genomics Laboratory of the University of Birmingham).

## References

Bausek N, Waclawek M, Schneider WJ and Wohlrab F (2000) The major chicken egg envelope protein ZP1 is different from ZPB and is synthesized in the liver. J Biol Chem 275,28866–28872.

Bercegeay S, Jean M, Lucas H and Barriere P (1995) Composition of human zona pellucida as revealed by SDS–PAGE after silver staining. Mol Reprod Dev 41,355–359.

Bleil JD and Wassarman PM (1980a) Mammalian sperm–egg interaction: identification of a glycoprotein in mouse egg zonae pellucidae possessing receptor activity for sperm. Cell 20,873–882.

Bleil JD and Wassarman PM (1980b) Structure and function of the zona pellucida: identification and characterization of the proteins of the mouse oocyte's zona pellucida. Dev Biol 76,185–202.

Bleil JD and Wassarman PM (1983) Sperm–egg interactions in the mouse: sequence of events and induction of the acrosome reaction by a zona pellucida glycoprotein. Dev Biol 95,317–324.

Bleil JD, Beall CF and Wassarman PM (1981) Mammalian sperm–egg interaction: fertilization of mouse eggs triggers modification of the major zona pellucida glycoprotein, ZP2. Dev Biol 86,189–197.

Boja ES, Hoodbhoy T, Fales HM and Dean J (2003) Structural characterization of native mouse zona pellucida proteins using mass spectrometry. J Biol Chem 278,34189–34202.

Conner SJ and Hughes DC (2003) Analysis of fish ZP1/ZPB homologous genes-evidence for both genome duplication and species-specific amplification models of evolution. Reproduction 126,347–352.

Dell A, Chalabi S, Easton RL, Haslam SM, Sutton-Smith M, Patankar MS, Lattanzio F, Panico M, Morris HR and Clark GF (2003) Murine and human zona pellucida 3 derived from mouse eggs express identical O-glycan. Proc Natl Acad Sci USA 100,15631–15636.

Epifano O, Liang L, Familari M, Moos MC and Dean J (1995) Coordinate expression of the three zona pellucida genes during mouse oogenesis. Development 121,1947–1956.

Gibbs RA, Nguyen P-N, McBride LJ, Koepf SM and Caskey CT (1989) Identification of mutations leading to Lesch–Nyhan syndrome by automated direct DNA sequencing of in-vitro amplified cDNA. Proc Natl Acad Sci USA 86,1919–1923.

Govind CK, Hasegawa A, Koyama K and Gupta SK (2000) Delineation of a conserved B cell epitope on bonnet monkey (Macaca radiata) and human zona pellucida glycoprotein-B by monoclonal antibodies demonstrating inhibition of sperm–egg binding. Biol Reprod 62,67–75.

Green DP (1997) Three-dimensional structure of the zona pellucida. Rev Reprod 2,147–156.

Greve JM and Wassarman PM (1985) Mouse egg extracellular coat is a matrix of interconnected filaments possessing a structural repeat. J Mol Biol 181,253–264.

Gupta SK, Yurewicz EC, Sacco AG, Kaul R, Jethanandani P and Govind CK (1998) Human zona pellucida glycoproteins: characterization using antibodies against recombinant non-human primate ZP1, ZP2 and ZP3. Mol Hum Reprod 4,1058–1064.

Harris JD, Hibler DW, Fontenot GK, Hsu KT, Yurewicz EC and Sacco AG (1994) Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families. DNA Seq 4,361–393.

Higgins D., Thompson J, Gibson T, Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22,4673–4680.

Holding C, Bolton V and Monk M (2000) Detection of human novel developmental genes in cDNA derived from replicate individual preimplantation embryos. Mol Hum Reprod 6,801–809.

Hughes DC and Barratt CLR (1999) Identification of the true human orthologue of the mouse Zp1 gene: evidence for greater complexity in the mammalian zona pellucida? Biochim Biophys Acta 1447,303–306.

Lefièvre L, Barratt CLR, Harper CV, Conner SJ, Flesch FM, Deeks E, Moseley FLC, Pixton KL, Brewis IA and Publicover SJ (2003) Physiological and proteomic approaches to study pre fertilisation events in the human. Reprod BioMed Online 7,419–427.

Moos J, Faundes D, Kopf GS and Schultz RM (1995) Composition of the human zona pellucida and modifications following fertilization. Mol Hum Reprod 10,2467–2471.

Neilson L, Andalibi A, Kang D, Coutifaris C, Strauss JF, Stanton JA and Green DP (2000) Molecular phenotype of the human oocyte by PCR–SAGE. Genomics 63,13–24.

Ponte P, Ng SY, Engel J, Gunning P and Kedes L (1984) Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. Nucl Acid Res 12,1687–1696.

Prasad SV, Wilkins B, Skinner SM and Dunbar BS (1996) Evaluating zona pellucida structure and function using antibodies to rabbit 55 kDa ZP protein expressed in baculovirus expression system. Mol Reprod Dev 43,519–529.

Rankin TL, Tong ZB, Castle PE, Lee E, Gore-Langton R, Nelson LM and Dean J (1998) Human ZP3 restores fertility in Zp3 null mice without affecting order-specific sperm binding. Development 125,2415–2424.

Rankin T, Talbot P, Lee E and Dean J (1999) Abnormal zonae pellucidae in

Downloaded from https://academic.oup.com/humrep/article/19/7/1580/2356591 by U.S. Department of Justice user on 16 August 2022

mice lacking ZP1 result in early embryonic loss. Development 126,3847–3855.

Rankin TL, Coleman JS, Epifano O, Hoodbhoy T, Turner SG, Castle PE, 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.

Spargo SC and Hope RM (2003) Evolution and nomenclature of the zona pellucida gene family. Biol Reprod 68,358–362.

Topper EK, Kruijt L, Calvete J, Mann K, Topfer-Petersen E and Woelders H (1997) Identification of bovine zona pellucida glycoproteins. Mol Reprod Dev 46,344–350.

Wassarman PM (1999) Mammalian fertilization: molecular aspects of gamete adhesion, exocytosis, and fusion. Cell 96,175–183.

Yurewicz EC, Sacco AG and Subramanian MG (1987) Structural characterization of the Mr = 55,000 antigen (ZP3) of porcine oocyte zona pellucida. Purification and characterization of alpha- and beta-glycoproteins following digestion of lactosaminoglycan with endo-beta-galactosidase. J Biol Chem 262,564–571.

Yurewicz EC, Sacco AG, Gupta SK, Xu N and Gage DA (1998) Hetero-oligomerization-dependent binding of pig oocyte zona pellucida glycoproteins ZPB and ZPC to boar sperm membrane vesicles. J Biol Chem 273,7488–7494.

*Submitted on August 6, 2003; accepted on April 14, 2004*

Downloaded from https://academic.oup.com/humrep/article/19/7/1580/2356591 by U.S. Department of Justice user on 16 August 2022