DEVELOPMENTAL BIOLOGY 156, 399-408 (1993)

Conservation of Mammalian Secondary Sperm Receptor Genes Enables the Promoter of the Human Gene to Function in Mouse Oocytes

LI-FANG LIANG AND JURRIEN DEAN

Laboratory of Cellular and Developmental Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

Accepted December 14, 1992

The human zona pellucida is an extracellular sheath composed of three major proteins (ZP1, ZP2, and ZP3) which surround the ovulated egg and mediate the initial interactions with sperm. Although fertilization is relatively species-specific and human sperm will not bind to mouse zona, there is a high degree of conservation between the coding regions of human *ZP3* and mouse *Zp-3* (the primary sperm receptor) genes. We now report the characterization of the human *ZP2* gene and demonstrate that the sequences of its coding regions are 70% identical with those of the mouse *Zp-2* (the secondary sperm receptor) gene. In addition, the first 300 bp of the 5′ flanking regions of human *ZP2* and mouse *Zp-2* are highly conserved. This region of 5′ flanking DNA contains a previously described 12-bp DNA sequence (element IV) that forms an oocyte-specific DNA–protein complex important for mouse *Zp-2* and *Zp-3* promoter activity. Human element IV forms a DNA–protein complex in gel mobility shift assays when incubated with human or mouse ovarian extracts. The formation of this complex is inhibited with molar excess of either human or mouse element IV sequences and is not present in extracts of testes, uterus, spleen, lung, or kidney. The human promoter region (0.3 kbp), coupled to a luciferase reporter gene and microinjected into the nuclei of 50-μm-diameter mouse oocytes, results in reporter gene activity at a level comparable to that of the homologous mouse promoter. Clustered point mutations in element IV in either the mouse or the human sequence dramatically decrease reporter gene activity. These results indicate that the similarity between mouse *Zp-2* and human *ZP2* genes enables the human promoter to utilize the heterologous transcription machinery in mouse oocytes. The observed transcription may involve the recognition of promoter sequences in element IV by conserved transcription factor(s).

INTRODUCTION

The zona pellucida surrounding mammalian oocytes mediates the initial binding of sperm to oocytes, confers species specificity at fertilization, and plays an important role in the postfertilization block to polyspermy (Wassarman, 1988; Dean, 1992). The mouse zona pellucida is composed of three major sulfated glycoproteins, ZP1, ZP2, and ZP3 (Bleil and Wassarman, 1980; Shimizu *et al.*, 1983). The mouse *Zp-2* and *Zp-3* genes encode two of these proteins and are coordinately transcribed in oocytes where their expression is limited to the growth phase of oogenesis (Liang *et al.*, 1990; Ringuette *et al.*, 1986, 1988; Philpott *et al.*, 1987; Kinloch *et al.*, 1988; Chamberlin and Dean, 1989; Roller *et al.*, 1989). Because *Zp-2* and *Zp-3* are the only genes whose expression is known to be restricted to the mouse female germline, they represent a paradigm for investigating mechanisms that control oocyte-specific gene expression. The presence of an extracellular zona pellucida matrix in all mammals (including monotremes, marsupials, and eutherians) suggests that zona genes are conserved and that a comparison of promoter regions in different mammals might provide insight into mechanisms that direct oocyte-specific gene expression.

Consistent with this hypothesis, the human zona pellucida (like that of the mouse) is composed of three glycoproteins (Shabanowitz and O'Rand, 1988), and the exon map and coding regions of the mouse *Zp-3* and human *ZP3* genes are well conserved (Chamberlin and Dean, 1990). Furthermore, comparison of the 5′ flanking sequences of mouse *Zp-2*, mouse *Zp-3*, human *ZP2*, and human *ZP3* genes identified five conserved DNA elements (Millar *et al.*, 1991). One, element IV, contains 12 bp that are 75% identical among the four genes. Mutating the core 6 bp of this element results in a dramatic decrease in expression of reporter genes driven either by mouse *Zp-3* or *Zp-2* 5′ flanking regions when microinjected into the nuclei of growing mouse oocytes. An identical, oocyte-specific protein-DNA complex is formed with oligonucleotides containing element IV from *Zp-2* or *Zp-3*. These data suggest that a common transcription factor(s) binds to element IV and may be involved in the coordinate expression of the two mouse zona genes.

We now describe the isolation and characterization of an additional zona gene, human *ZP2*, which is the homo-

0012-1606/93 $5.00

log of the mouse *Zp-2* gene (Liang *et al.*, 1990). We present evidence that sequence conservation between the human and mouse promoters enables the human promoter to be transcriptionally active in mouse oocytes. The data suggest that human and mouse oocytes may utilize a conserved transcription factor(s) to regulate expression of the zona genes.

MATERIALS AND METHODS

Screening and Isolation of Human ZP2 Genomic DNA and cDNA Clones

A mouse ZP2 cDNA fragment, pZP2.4 (Liang *et al.*, 1990), was used to screen a human genomic DNA library in Charon 4A (Lawn *et al.*, 1978). Three overlapping genomic clones, λHuZP2G.1 (14.4 kbp), λHuZP2G.2 (14.5 kbp), and λHuZP2G.3 (17.4 kbp) were isolated and contained the entire human *ZP2* gene.

Human ovarian mRNA, isolated from an ovary of a 14-year-old female, was used in the synthesis of first-strand cDNA using Superscript RNase H^- reverse transcriptase (BRL). Two sets of oligonucleotide primers based on human exon sequence were used in polymerase chain reactions (PCR) to generate two human cDNA fragments (Fig. 1, nt 22 to 480 and nt 343 to 1189). These two fragments were subcloned into Bluescript vectors, resulting in pHuZP2.1 and pHuZP2.2, respectively. In addition, human ovarian mRNA was used in the construction of an ovarian cDNA library using the Uni-ZAP cDNA library construction system (Stratagene). The library was probed with pHuZP2.2, and a clone containing a 1191-bp insert plus 19 adenosines (Fig. 1, nt 1077 to 2267) was isolated, pHuZP2.3. DNA sequence was determined by dideoxy sequencing (Sanger *et al.*, 1977) using Sequenase (U.S. Biochemical Corp.) and synthetic oligonucleotide primers based on human *ZP2* genomic and cDNA sequences. Sequence analysis was performed using PC/Gene (IntelliGenetics, Inc.) and the 5′ flanking sequences were aligned using MACAW software (Schuler *et al.*, 1991).

Reporter Gene Constructs

Reporter gene constructs were made by ligating portions of 5′ upstream DNA sequences from mouse *Zp-2* and human *ZP2* genes into the multiple cloning site of the luciferase vector pXP1 (Nordeen, 1988). pMZP2luc[1.5] and pHZP2luc[1.35] were generated by using restriction digestion fragments containing 1.50 and 1.35 kbp of 5′ flanking sequence of the mouse *Zp-2* and human *ZP2* genes, respectively. pMZP2luc[0.3] and pHZP2luc[0.3] contain 253 and 266 bp of PCR-generated fragments immediately upstream of mouse and human TATAA boxes, respectively. Mutations in mouse (pMZP-2luc[0.3]MtIV) and human (pHZP2luc[0.3]MtIV) element IV were made by PCR using primers containing the desired sequence changes. In each mutation, the 6-bp core, CACCTG, was changed to TCTAGA. pMZP2luc[TATAA], contained a 43-bp synthetic oligonucleotide of mouse 5′ flanking sequence that includes the TATAA box. The integrity of the plasmids containing PCR fragments and oligonucleotides was confirmed by dideoxy sequencing. The plasmid pCHRSV (Millar *et al.*, 1991), which contained the Rous sarcoma virus long terminal repeat upstream of the *lacZ* gene, was used as the internal control for all microinjection assays. The luciferase vector pXP1 was used as negative control.

Microinjection of Mouse Oocytes

Oocytes (50 μm in diameter) were isolated from 12-day-old NIH Swiss mice as described previously (Millar *et al.*, 1991). Each luciferase reporter construct (approximately 9×10^4 copies in a volume of 1.5 pl) was injected into the nuclei of 70–80 mouse oocytes. Half of the injected oocytes were assayed for luciferase activity using the Luciferase Assay System (Promega) with purified luciferase (Analytical Luminescence Laboratories) as a standard. The other half of the injected oocytes was assayed for β-galactosidase, using purified β-galactosidase (Boehringer-Mannheim) as standards (Millar *et al.*, 1991). The level of luciferase activity for each reporter construct was adjusted based on the level of the activity of the internal control β-galactosidase plasmid pCHRSV. Corrections were consistently less than 4%.

Gel Mobility Shift Assays

Ovarian extracts were made from 10–50 mg of: newborn mouse ovaries, a human ovary from a 17-year-old (National Disease Research Interchange, Philadelphia), or surgical specimens of human testes, uterus, lung, spleen, and kidney (Millar *et al.*, 1991). Double-stranded human *ZP2* element IV oligonucleotide, CACCAATTCACTCACCTGGAGCTGATTTCA, and mouse *Zp-2* element IV oligonucleotide, CACTAATTTACTCACCTGGAGCCAATTTTG, were gel purified and labeled with [α-^{32}P]dCTP and [α-^{32}P]dATP using the Klenow fragment of DNA polymerase I. After incubation with 5 μg of tissue extracts, the products were analyzed by polyacrylamide gel electrophoresis and autoradiography (Millar *et al.*, 1991).

Accession Number

The GenBank Accession Number for human *ZP2* is M90366.

RESULTS

Human ZP2 mRNA and Protein

The structure of the human ZP2 mRNA was deduced from the nucleic acid sequence of three overlapping human cDNA clones and the human genomic exons (Fig. 1). The close proximity of a TATAA (−57 bp) box upstream of the initiation codon ATG (see below) suggests that human ZP2 mRNA has a short 5′ untranslated region similar to that of murine ZP2 (30 nt) and ZP3 (29 nt). Likewise, the 31-nt 3′ untranslated region conforms to the short lengths observed in mouse ZP2 (32 nt), mouse ZP3 (16 nt), and human ZP3 (17 nt) (Liang *et al.*, 1990; Ringuette *et al.*, 1988; Chamberlin and Dean, 1990).

The initiator ATG is preceded by a purine at −3 bp and followed by a G residue, a context considered important for initiation of translation in eukaryotes (Kozak, 1991). The human ZP2 mRNA contains an open reading frame of 2235 nt that can code for a polypeptide of 82,356 Da containing 745 amino acids (10.2% acidic, 11.5% basic, 9.4% aromatic, and 50.3% hydrophobic). Human and mouse ZP2 amino acid sequences are 60.7% identical. Examination of human ZP2 protein revealed a potential signal peptidase cleavage site which contains amino acids at the −1 and −3 positions that are in accordance with the (−3, −1) rule proposed by von Heijne (Von Heijne, 1985, 1986). Cleavage at the presumptive signal peptidase site would give rise to a signal sequence of 38 amino acids (four residues longer than mouse ZP2) and a resultant protein with a predicted molecular mass of 78,200 Da.

The deduced polypeptide chain contains six potential N-linked glycosylation sites (Asn-X-Ser/Thr), four of which are conserved in the mouse ZP2 polypeptide (Fig. 1). The predicted hydropathicity of the human and mouse ZP2 proteins is quite similar, reflecting both amino acid identity and conservative amino acid substitutions (Fig. 2). The conservation of all 20 cysteine residues in the mature human and mouse proteins suggests that at least some of these residues participate in disulfide bonds important for tertiary structure. An additional exon found in human *ZP2* (see below) encodes a 28-amino-acid hydrophilic region (residues 671–698) near the carboxyl terminus. Whether this motif is involved with the species specificity of human fertilization remains to be determined.

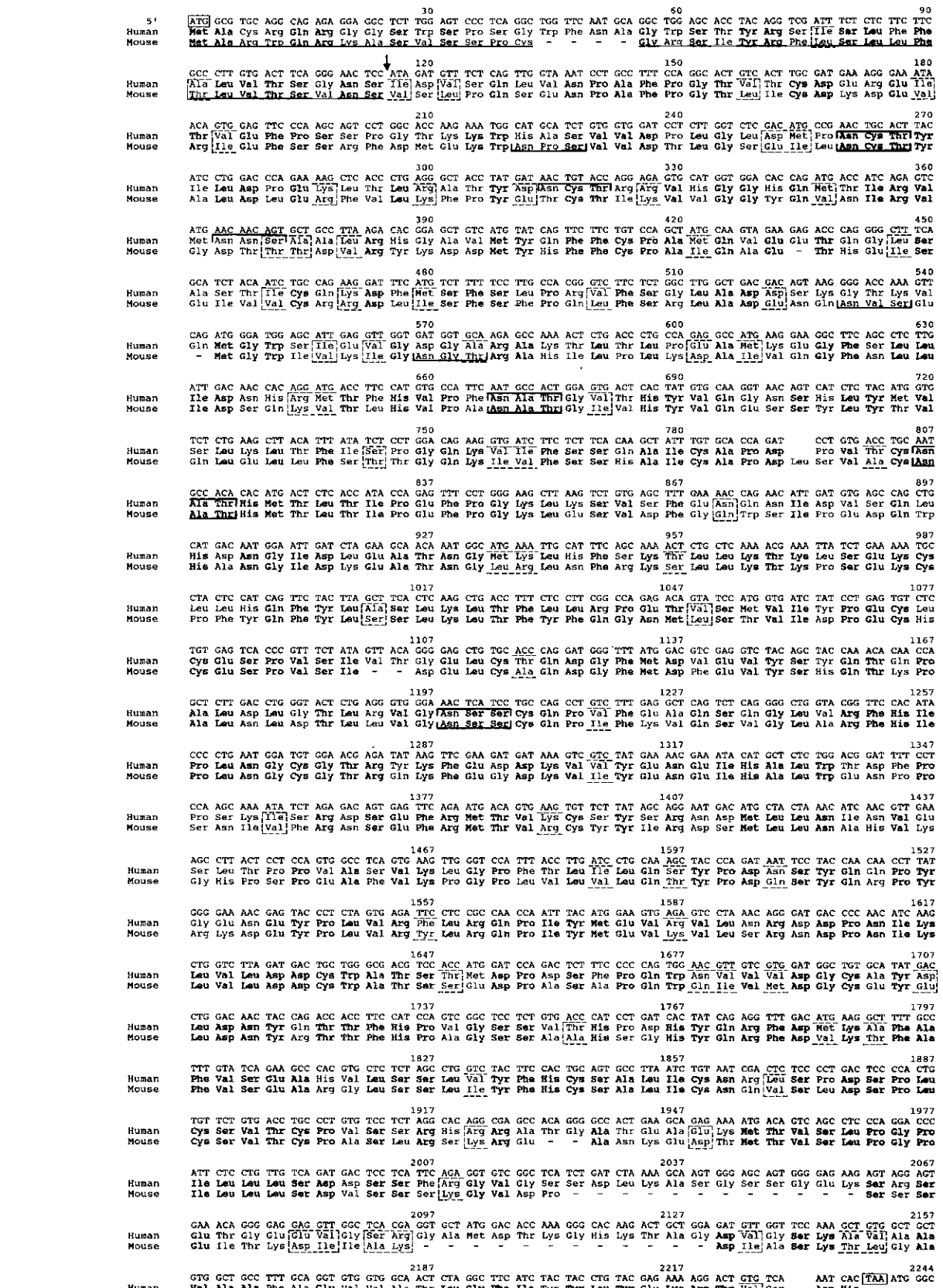
Human ZP2 Genomic Locus

The human *ZP2* genomic locus is 14.0 kbp in length, consisting of 19 exons ranging from 45 to 190 bp. (Fig. 3A, Table 1). The nucleotide sequences of the human *ZP2* exons are identical to those determined from the human ZP2 cDNA. The sizes of the introns, ranging from 74 to 3000 bp, were determined either by direct sequencing or by analyzing polymerase chain reaction products primed with synthetic oligonucleotides that mapped to regions flanking the introns (Table 1). The exons have canonical splice acceptor and splice donor consensus sequences (Breathnach and Chambon, 1981), and the spatial arrangement of human *ZP2* exons is similar to that of the mouse *Zp-2* gene. Overall, the coding regions are 70% identical to those of mouse ZP2. However, human *ZP2* contains an extra exon of 84 bp (exon 18) that is not found in mouse ZP2 cDNA. Sequence analysis of mouse *Zp-2* intron 17 revealed a region of 76 bp that shares a 70% sequence homology with human *ZP2* exon 18 (data not shown), indicating that the DNA sequence has been used quite differently since the evolutionary divergence of primates and rodents (ca. 60–80 million years ago).

5′ Flanking Sequence of the Human ZP2 Gene

The minimal 5′ limit of human *ZP2* exon 1 was defined by comparing human genomic sequences with a near full-length human ZP2 cDNA clone. Sequence analysis of 1.8 kbp upstream of the putative initiation ATG revealed the presence of a TATAA box and a CAAT box at −57 and −94 bp, respectively. Two human *Alu* repeats were identified at approximately −700 and −1700 bp upstream of the ATG. Comparison of the human *ZP2* (1.35 kbp) and mouse *Zp-2* (1.5 kbp) 5′ flanking sequences using the MACAW alignment algorithm (Schuler *et al.*, 1991) revealed that, although overall the sequences are quite different, there are several regions with greater than 60% sequence identity (Fig. 3B).

Most notably, the first 300 bp upstream of the two genes are 70% identical (Fig. 3B). This high level of sequence conservation is not observed between human *ZP3* and mouse *Zp-3* 5′ flanking regions (Chamberlin and Dean, 1990). The high degree of identity between *ZP2* and *Zp-2* 5′ sequences extends through the first intron (69%) and the first 42 bp of intron 2 (71%) for a total of 575 bp. This 5′ flanking region contains a previously described 12-bp DNA sequence (element IV) at −247 bp to the initiator ATG that is present at comparable distances upstream of mouse *Zp-2*, mouse *Zp-3*, and human *ZP3* genes (Millar *et al.*, 1991). The 12-bp element IV is identical in human *ZP2* and mouse *Zp-2*, and a 30-bp region centered at element IV is 80% the same between the two species (Fig 3B). Results of earlier studies involving gel mobility shift assays and microinjection assays with mouse *Zp-2* and *Zp-3* promoter-luciferase reporter gene constructs suggested that element IV may have an important role in the regulation of zona pellucida gene expression in growing mouse oocytes (Millar *et al.*, 1991).


5'
Human
Mouse
TTC TAA ATA AAG CAG TCA AAA T 3'

Gel Mobility Shift Assays of Element IV with Human Ovarian Extracts

Oligonucleotides containing element IV from mouse *Zp-2* and *Zp-3* genes form DNA-protein complexes that have identical mobility in gel mobility shift assays and are oocyte-specific (Millar *et al.*, 1991). Both human *ZP2* and mouse *Zp-2* element IV are located approximately the same distance upstream of their respective TATAA boxes (−190 and −185 bp, respectively), placing them within the region of high sequence conservation (Fig. 3B).

Gel mobility shift assays were performed using human and mouse oligonucleotides (30 bp, centered on element IV) incubated with ovarian extract from either human or mouse. The human and mouse element IV oligonucleotides form DNA-protein complexes with identical mobility after incubation with mouse ovary extract (Fig. 4A, top arrow, lanes 2 and 6, respectively). The other bands observed in the gel are DNA-protein complexes that are not oocyte-specific (as previously described (Millar *et al.*, 1991)). Likewise, human and mouse element IV oligonucleotides form DNA-protein complexes with identical mobility after incubation with human ovary extract (Fig. 4B, bottom arrow, lanes 3 and 5, respectively). The DNA-protein complex formed with human ovarian extract has a slightly faster mobility than that observed with mouse ovarian extract, suggesting that the protein(s) binding to element IV differs somewhat between human and mouse.

To characterize further the nature of the element IV DNA-protein complex, competition studies were performed. Although mouse and human element IV are identical and the immediately surrounding DNA is 80% the same, DNA-protein complex formation between mouse element IV and mouse oocyte protein is competed more efficiently by mouse oligonucleotides containing element IV (Fig. 4B, lanes 2–4) than by human oligonucleotides (Fig. 4B, lanes 5–7). However, when labeled human element IV is incubated with human ovarian extract and competed with increasing amounts (10- to 100-fold molar excess) of unlabeled human (Fig. 4C, lanes 2–4) or mouse (Fig. 4C, lanes 5–7) element IV, both human and mouse element IV oligonucleotides compete to the same degree. These results indicate that the putative transcription factor(s) in the mouse ovarian extracts that binds to mouse element IV can also bind human element IV with relatively high efficiency.

To determine whether the protein-DNA complex formed with ovarian extract and element IV from the human *ZP2* gene can also be formed with protein from other tissues, gel mobility shift assays were performed using extracts of human testes, uterus, spleen, lung, and kidney (Fig. 4D). The complex formed with human ovary extract and the human *ZP2* oligonucleotide is not observed with extracts of the other human tissue examined. These data suggest that the formation of the DNA-protein complex is sex-specific in the gonad and, at least, relatively tissue-specific.

Human Promoter Activities in Microinjected Mouse Oocytes

The high degree of sequence identity in the immediate region upstream of mouse and human *ZP2* genes, and the ability of human element IV to form identical DNA-protein complexes in a gel mobility shift assay, suggests that the transcriptional machinery for *ZP2* expression may be conserved between mouse and human. To test this possibility, luciferase reporter gene constructs containing the DNA sequences upstream of mouse *ZP-2* and human *ZP2* genes were microinjected into the nuclei of 70–80 growing mouse oocytes and assayed for luciferase activity. Each construct was co-injected with pCHRSV (containing the RSV long terminal repeat upstream of the β-galactosidase reporter gene) as an internal control for efficiency of injection and plasmid expression. The plasmid pXP1, a promoterless luciferase construct, was used as the negative control, and plasmid pSV2L (de Wet *et al.*, 1987), which contains the simian virus 40 early region promoter and enhancer, was used as a positive control (Millar *et al.*, 1991).

Microinjection of a mouse promoter containing 1.5 kbp of *Zp-2* upstream flanking sequence (pMZP2luc[1.5]) resulted in reporter gene (average ± SD) activity of $1.4 \pm 0.5 \times 10^3$ luciferase units (Fig. 5). This is approximately 10 times greater than that obtained with a comparable mouse *Zp-3* construct microinjected into mouse oocytes (Millar *et al.*, 1991) and roughly comparable to that estimated for oocytes in transgenic mice with a 6.5-kbp *Zp-3* promoter coupled to luciferase (Lira *et al.*, 1990). A minimal mouse promoter (pMZP2luc[TATAA]) containing

FIG. 1. Primary structure of the human ZP2 mRNA and protein. The nucleic acid sequence of the near-full-length cDNAs and exon 1 of the *ZP2* gene were used to deduce the structure of the ZP2 mRNA and resultant protein. The initiation and termination codons are boxed, and the polyadenylation signal is overlined. The single 2235-nucleotide open reading frame is translated into a 745-amino-acid protein in line 2 and aligned in line 3 with the 713-amino-acid mouse ZP2 (Liang *et al.*, 1990). The putative 38-amino-acid signal peptide is underlined, and the arrow points to the predicted signal peptidase cut site. Identical amino acid residues between human and mouse ZP2 are shaded; conserved changes (Dayhoff and Orcutt, 1979) are enclosed in boxes with dotted lines. The potential N-linked glycosylation sites (Asn-X-Thr/Ser) are marked in bold brackets.

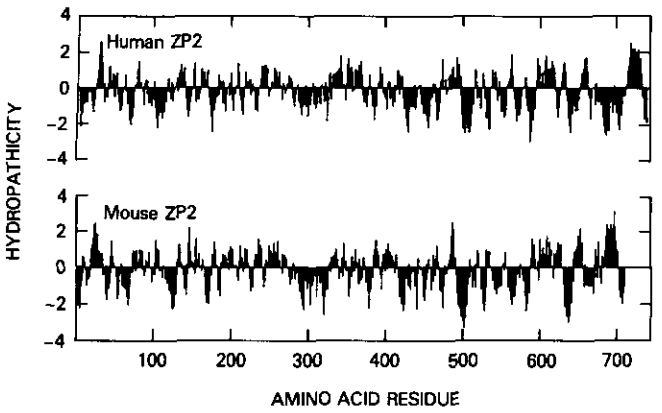


FIG. 2. Comparison of the secondary structure of the deduced human and mouse ZP2 proteins. The hydropathicity of human and mouse ZP2 (Kyte and Doolittle, 1982) indicates the overall similarity of the two proteins. Both have major hydropathic peaks in their signal peptides and near their carboxyl termini.

43 bp of upstream sequences including the TATAA box had 8.0 ± 11 luciferase units; the promoterless construct, pXP1, had 1.0 ± 2.2 luciferase units. The *Zp-2* mouse promoter with 0.3 kbp (pMZP2luc[0.3]) resulted in 1.1 ± 0.4×10^2 luciferase units, and a 6-bp cluster mutation in element IV (pMZP2luc[0.3]MtIV) reduced that activity to 4.5 ± 1.3 luciferase units which represent 4% of the reporter gene activity observed with the 0.3-kbp parenteral construct. These data confirm our earlier observation (Millar *et al.*, 1991) that the presence and integrity of element IV are important for reporter gene activity driven by the *Zp-2* mouse promoter in mouse oocytes.

In a parallel set of experiments (Fig. 5), a human promoter containing 1.35 kbp of *ZP2* upstream sequence (pHZP2luc[1.35]) resulted in 2.1 ± 1.1 × 10^2 luciferase units and was comparable to that (2.2 ± 0.9 × 10^2 units) obtained with only 0.3 kbp of human 5′ flanking sequence (pHZP2luc[0.3]). A 6-bp mutation in the center of element IV in the human *ZP2* 0.3-kbp construct (pHZP2luc[0.3]MtIV) dramatically reduced the luciferase activity to 3.7 ± 3.5 units or 2% of the parenteral construct. The results of these transient expression assays indicate that the 0.3-kbp human flanking sequence is recognized by mouse oocyte transcriptional machinery and indicate the importance of element IV for hu-

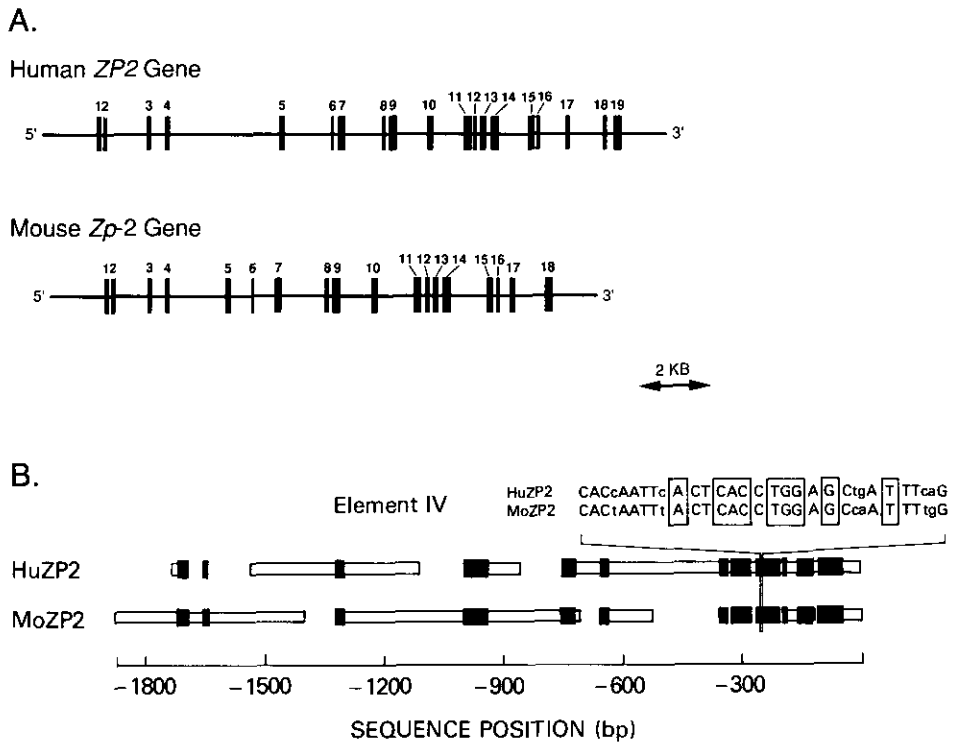


FIG. 3. Promoter region of human *ZP2* and mouse *Zp-2* genes. (A) Schematic representation of the exon-intron map of human *ZP2* and mouse *Zp-2* (Liang *et al.*, 1990). Dark bars represent exons of which there are 19 in the human and 18 in the mouse genes. The entire human *ZP2* gene is 14 kbp in length. (B) Alignment of the 5′ flanking sequence of human *ZP2* (HuZP2) and mouse *Zp-2* (MoZP2) genes using the MACAW algorithm (Schuler *et al.*, 1991). Regions that have sequence identity greater than 60% are shaded in black. Overall the first 300 bp 5′ flanking sequences are 70% conserved. The 30-bp sequences of the human and mouse genes containing element IV that are used in gel mobility shift assays are presented above the alignment. Capitalized nucleotides of the 12 bp of element IV (ACTCACCTGGAG) and the flanking sequences indicate identity between the two genes. Boxed sequences are identical among four zona genes: mouse *Zp-2* and *Zp-3*; and human *ZP2* and *ZP3*.

TABLE 1
EXON AND INTRON SIZES OF HUMAN *ZP2* GENE

Exon	Position	Length (bp)	Intron	Length (bp)
1	1-62[b]	62	1	90[a]
2	63-151	89	2	1150
3	152-235	84	3	380
4	236-330	95	4	3000
5	331-483	153	5	1200
6	484-528	45	6	146[a]
7	529-693	165	7	1023[a]
8	694-790	97	8	96[a]
9	791-972	182	9	783[a]
10	973-1099	127	10	900
11	1100-1287	188	11	80[a]
12	1288-1379	92	12	101[a]
13	1380-1504	125	13	147[a]
14	1505-1694	190	14	800
15	1695-1830	136	15	74[a]
16	1831-1927	97	16	706[a]
17	1928-2011	84	17	927[a]
18	2012-2095	84	18	142[a]
19	2096-2266	171		

[a] Exact length determined by sequencing.
[b] From the initiator ATG to the end of the first exon.

man promoter function in heterologous mouse oocytes. The lack of further increases in luciferase activity (Fig. 5) with the addition of more upstream human 5′ sequences (pHZP2luc[1.35]) suggests that the mouse upstream region (not well conserved in the human, see Fig. 3B) contains additional regulatory sequences important for zona gene expression in mouse oocytes.

DISCUSSION

The human *ZP2* and mouse *Zp-2* genes are well conserved and contain 575-bp segments at their 5′ ends that are 70% identical and extend from intron 2 upstream to include 300 bp of the promoter region. When coupled to luciferase coding sequences and microinjected into mouse oocytes, the human promoter (0.3 kbp) expressed levels of reporter gene activity comparable to those obtained with the mouse promoter (0.3 kbp). Thus, it appears that the conservation of sequences between the first 300 bp of the mouse and human promoters enables the human promoter to utilize the transcriptional machinery of the mouse oocyte. Additional mouse sequence (1.5 kbp) increased reporter gene activity 10- to 12-fold, whereas additional human sequence (1.35 kbp) had no such effect. Presumably, the greater divergence of these more upstream sequences between the two species prevents productive interactions between this region of human DNA and relevant transcription factor(s) in mouse oocytes.

Element IV, a previously described 12-bp sequence that is necessary and sufficient for expression of reporter genes microinjected into the nuclei of growing oocytes (Millar *et al.*, 1991), is present at position −190 and −185 upstream of the human *ZP2* and mouse *Zp-2* TATAA boxes, respectively. A 6-bp cluster mutation in the center of both human and mouse element IV dramatically decreased reporter gene activity in mouse oocytes microinjected with luciferase constructs containing 0.3 kbp of 5′ flanking sequence. Oligonucleotides (30 bp) centered on element IV and composed of either mouse or human sequence form complexes identical to each other in gel mobility shift assays after incubation with mouse ovarian extracts. A similar complex with a slightly faster mobility is formed in human ovarian extracts, indicating that the mouse and human transcription factors have some structural differences. Taken together, these data suggest that regulatory mechanisms that utilize element IV may be conserved between human and mouse.

The human *ZP2* element IV (12 bp) is 100% identical to that of mouse *Zp-2* and its core contains a sequence, CACCTG, that is critical for the binding of Myo D to the enhancer region of a creatine kinase gene (Lassar *et al.*, 1989) and of E12 and E47 to the κE2 site of the immunoglobulin κ-chain enhancer (Murre *et al.*, 1989). These previously reported transcription factors are related, and each contains a basic region that interacts with DNA and a helix-loop-helix protein motif important for the formation of homo- or heterodimers. The presence of this core sequence in element IV raises the possibility that helix-loop-helix factor(s) may be involved in the regulation of zona gene expression. It may be that nucleotides flanking the core sequence are also important for productive protein-DNA interactions and may have provided selective pressure for the conservation of the sequences surrounding element IV that result in their 80% identity in human and mouse.

We have previously reported that mutation of mouse element IV inhibits expression from mouse *Zp-2* and *Zp-3* promoters. These results, and the oocyte specificity of the element IV DNA-protein complex, led us to suggest that the factor(s) binding to element IV may be involved in the coordinate, oocyte-specific regulation of zona gene expression (Millar *et al.*, 1991). However, element

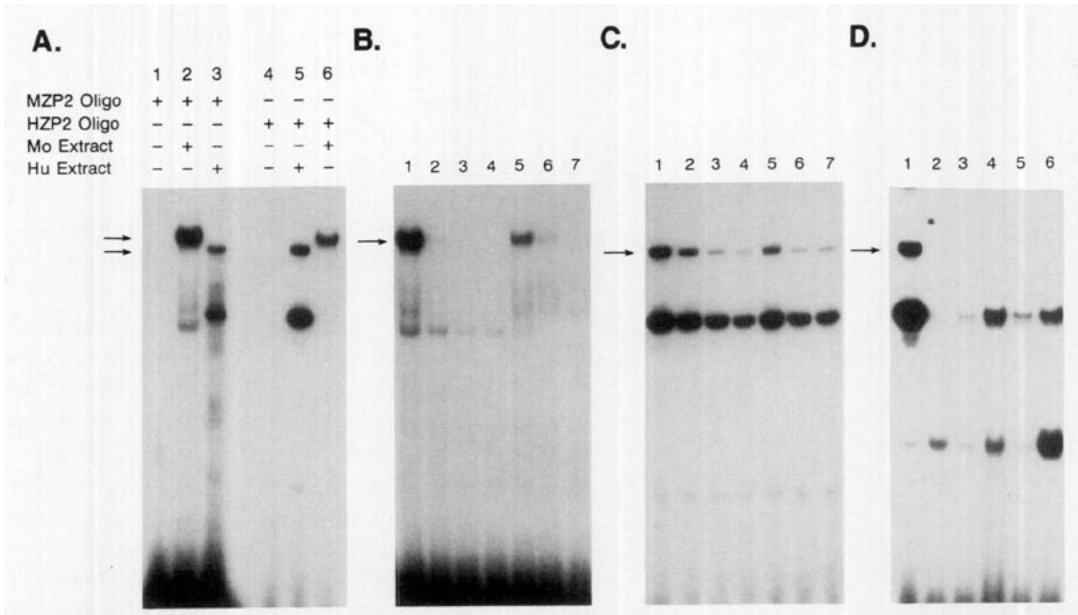


FIG. 4. Gel mobility shift competition assays. (A) Gel mobility shift assays using double-stranded human and mouse oligonucleotides containing the 12-bp conserved element IV. ^{32}P-labeled human or mouse element IV oligonucleotides (30 bp; 1.5×10^4 cpm) were incubated with either human or mouse ovarian extracts and analyzed by polyacrylamide gel electrophoresis. Arrows denote the tissue-specific DNA–protein complex bands; the faster mobility band is present in human extracts and the slower in mouse extracts. (B) ^{32}P-labeled mouse *Zp-2* element IV oligonucleotides were incubated with mouse ovarian extract in the presence of increasing amounts of unlabeled competitor oligonucleotides. Lane 1, ^{32}P-labeled mouse *Zp-2* element IV (30 bp; 1.5×10^4 cpm) incubated with mouse ovarian extract (5 μg); lanes 2–4, same as lane 1 but with 10-, 50-, and 100-fold molar excess of unlabeled mouse *Zp-2* element IV oligonucleotides, respectively; lanes 5–7, 10-, 50-, and 100-fold molar excess of unlabeled human *ZP2* element IV, respectively. Arrow indicates DNA-protein complex previously shown to be oocyte-specific (Millar *et al.*, 1991). (C) ^{32}P-labeled human *ZP2* element IV oligonucleotides were incubated with human ovarian extract in the presence of increasing amount of unlabeled competitor oligonucleotides. Lane 1, ^{32}P-labeled human *ZP2* element IV (30 bp; 1.5×10^4 cpm) incubated with human ovarian extract (5 μg); lanes 2–4, same as lane 1 but with 10-, 50-, and 100-fold molar excess of unlabeled human *ZP2* element IV oligonucleotides, respectively; lanes 5–7, 10-, 50-, and 100-fold molar excess of unlabeled mouse *Zp-2* element IV, respectively. Arrow denotes human *ZP2* element IV-specific DNA–protein complex. (D) ^{32}P-labeled human *ZP2* element IV oligonucleotides were incubated with human tissue extracts. Lane 1, ^{32}P-labeled human *ZP2* element IV (30 bp; 1.5×10^4 cpm) incubated with human ovarian extract (5 μg) ; lanes 2–6, same as lane 1 but with human extracts from testes, uterus, spleen, lung, and kidney.

IV represents only a small portion of a large conserved region identified in the 5′ flanking sequence of mouse *Zp-2* and human *ZP2* genes and it is likely that additional regulatory elements will be necessary for high level, *in vivo* expression of the mouse *Zp-2* and *Zp-3* genes. Several transcription factors have been identified in mouse oocytes (Rosner *et al.*, 1990; Scholer *et al.*, 1989, 1990), one of which binds to DNA sequences upstream of the *Zp-3* gene (Schickler *et al.*, 1992). If these proteins have a functional role in mouse zona gene expression, it is likely that homologous proteins will be important for human zona gene expression.

The apparent conservation of transcriptional machinery for zona pellucida gene expression is not unexpected since zona structures are present around all mammalian eggs and have importance in fertilization and early development. However, since it has been demonstrated that the zona pellucida confers relative species specificity to sperm–egg interactions, the high degree of conservation of the encoded human and mouse zona proteins was somewhat surprising. The 745-amino-acid-human ZP2 protein is 60% identical to that of its mouse counterpart (713 amino acids) (Liang *et al.*, 1990). The relative position of all cysteine residues is conserved between the two species, and the similarities of predicted secondary structures (allowing for conservative amino acid substitutions) suggest that the overall three-dimensional structures of the ZP2 proteins are likely to be conserved. As previously reported, the human and mouse ZP3 proteins are also similar to each other, being 424 amino acids long each and 67% identical (Chamberlin and Dean, 1990). The very hydrophobic region observed near the carboxyl terminus of human ZP2 is also present in mouse ZP2 (Liang *et al.*, 1990), mouse ZP3 (Ringuette *et al.*, 1988; Kinloch *et al.*, 1988; Chamberlin and Dean, 1989), human ZP3 (Chamberlin and Dean, 1990), hamster ZP3 (Kinloch *et al.*, 1990), and rabbit rc55 (Schwoebel *et al.*, 1991). It seems likely to play a role in the biological function or structural integrity of the mammalian zona pellucida.

Mouse ZP3 has been shown to inhibit sperm binding to ovulated eggs (but not two-cell embryos) and is capa-

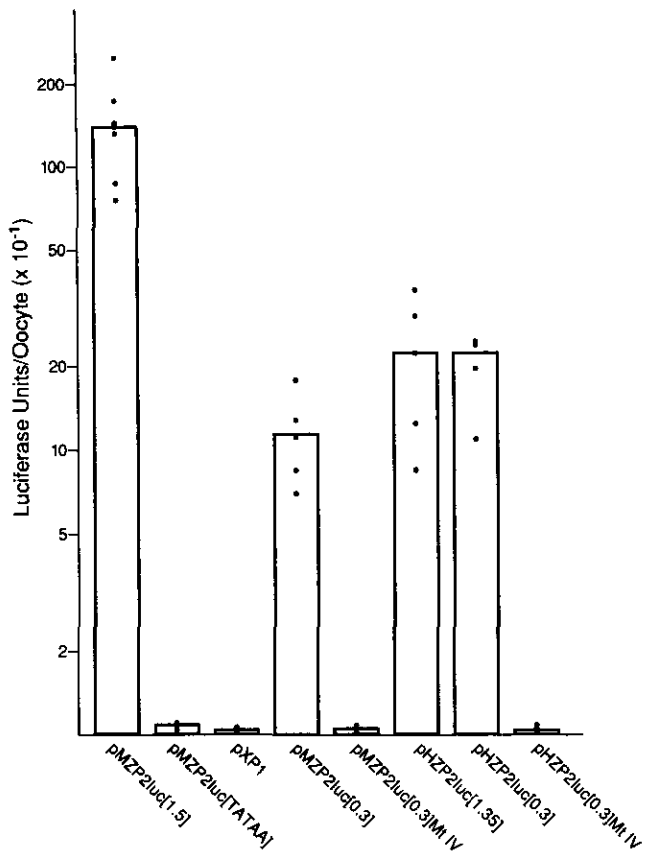


FIG. 5. Luciferase activity of the *ZP2* promoter-reporter genes in growing mouse oocytes. Reporter gene constructs were made with portions of DNA sequences 5′ of the mouse *Zp-2* and human *ZP2* genes inserted into the luciferase vector pXP1 (Nordeen, 1988). The following plasmids containing promoter fragments were microinjected into the nuclei of 50-μm oocytes: pMZP2luc[1.5], 1.5 kbp 5′ to mouse *Zp-2;* pMZP2luc[TATAA], 43 bp of mouse *Zp-2* 5′ flanking sequence including the TATAA box; pXP1, no promoter sequences (negative control); pMZP2luc[0.3], 0.3 kbp 5′ to mouse *Zp-2;* pMZP2luc[0.3]MtIV, the same 0.3-kbp mouse fragment containing a 6-bp clustered mutation in element IV; pHZP2luc[1.35], 1.35 kbp 5′ to human *ZP2;* pHZP2luc[0.3], 0.3 kbp 5′ to human *ZP2;* and pHZP2luc[0.3]MtIV, the same 0.3-kbp human fragment containing a 6-bp clustered mutation in element IV. Seventy to eighty oocytes were injected for each data point and luciferase activity was normalized according to the β-galactosidase activity from a co-injected control plasmid. Each experiment (●) was performed three to seven times and luciferase activity is reported on a log scale.

ble of inducing the sperm acrosome reaction, resulting in the release of lytic enzymes important for penetration of the zona pellucida. Mouse ZP2 acts as a secondary sperm receptor and, following fertilization, it undergoes a proteolytic cleavage thought to play a role in the post-fertilization block to polyspermy. Both proteins persist in the zona pellucida during early cleavage stages and are vital for passage of the embryo down the oviduct prior to implantation (Wassarman, 1988; Dean, 1992). Although the molecular details of fertilization have not been delineated in humans, it has been reported that human sperm–egg interaction is far more specific than mouse sperm–egg interaction. Spermatozoa from mouse, hamster, and rabbit can adhere to the surface of oocytes from other mammalian species including human (Bedford, 1977; Swenson and Dunbar, 1982). However, human spermatozoa have a very limited affinity for other mammalian oocytes and will only adhere to oocytes from Hominoidea primates (Bedford, 1977).

Differences in how human sperm interact with mouse and human oocytes may result from differences in the primary structure of the zona proteins (for instance, the 28-amino-acid hydrophilic domain encoded by the additional exon in human *ZP2*) or may be due to differences in post-translational modifications of these proteins. There is considerable evidence that glycosylation of the zona proteins plays an important role in their biological functions of mediating sperm binding and inducing the sperm acrosome reaction. Despite the similarity of the polypeptide chain length of human and mouse ZP2 and of human and mouse ZP3, they migrate quite differently in SDS–PAGE. The human zona pellucida glycoproteins have apparent molecular weights of 90–110, 64–76, and 57–73 kDa (Shabanowitz and O'Rand, 1988); these are significantly smaller than the molecular weights of their mouse counterparts (185–200, 120–140, and 83 kDa) (Bleil and Wassarman, 1980; Shimizu *et al.*, 1983). Presumably these differences result from post-translational modification of the zona proteins, and such modifications may also account for the species-specific nature of these glycoproteins. Now that full-length cDNAs of both human ZP2 and ZP3 have been isolated, it will be interesting to analyze the structure and biological function of these zona proteins following their expression in mouse oocytes.

We thank Drs. A. M. Ginsberg, R. McIsaac, S. Millar, R. H. Morse, and R. T. Simpson for their valuable comments and suggestions on the manuscript.

REFERENCES

Bedford, J. M. (1977). Sperm/egg interaction: The specificity of human spermatozoa. *Anat. Rec.* **188,** 477–488.

Bleil, J. D., and Wassarman, P. M. (1980). 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.

Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50,** 349–383.

Chamberlin, M. E., and Dean, J. (1989). Genomic organization of a sex specific gene: The primary sperm receptor of the mouse zona pellucida. *Dev. Biol.* **131,** 207–214.

Chamberlin, M. E., and Dean, J. (1990). Human homolog of the mouse sperm receptor. *Proc. Natl. Acad. Sci. USA* **87,** 6014–6018.

Dayhoff, M. O., and Orcutt, B. C. (1979). Methods for identifying proteins by using partial sequences. *Proc. Natl. Acad. Sci. USA* **76,** 2170–2174.

de Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R., and Subramani, S. (1987). Firefly luciferase gene: Structure and expression in mammalian cells. *Mol. Cell. Biol.* **7,** 725–737.

Dean, J. (1992). Biology of mammalian fertilization: Role of the zona pellucida. *J. Clin. Invest.* **89,** 1055–1059.

Kinloch, R. A., Roller, R. J., Fimiani, C. M., Wassarman, D. A., and Wassarman, P. M. (1988). Primary structure of the mouse sperm receptor polypeptide determined by genomic cloning. *Proc. Natl. Acad. Sci. USA* **85,** 6409–6413.

Kinloch, R. A., Ruiz-Seiler, B., and Wassarman, P. M. (1990). Genomic organization and polypeptide primary structure of zona pellucida glycoprotein hZP3, the hamster sperm receptor. *Dev. Biol.* **142,** 414–421.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266,** 19867–19870.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132.

Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., and Weintraub, H. (1989). MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. *Cell* **58,** 823–831.

Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G., and Maniatis, T. (1978). The isolation and chacterization of a linked δ and β-globin gene from a cloned library of human DNA. *Cell* **15,** 1157–1174.

Liang, L-F., Chamow, S. M., and Dean, J. (1990). Oocyte-specific expression of mouse Zp-2: Developmental regulation of the zona pellucida genes. *Mol. Cell. Biol.* **10,** 1507–1515.

Lira, S. A., Kinloch, R. A., Mortillo, S., and Wassarman, P. M. (1990). An upstream region of the mouse ZP3 gene directs expression of firefly luciferase specifically to growing oocytes in transgenic mice. *Proc. Natl. Acad. Sci. USA* **87,** 7215–7219.

Millar, S. E., Lader, E., Liang, L-F., and Dean, J. (1991). Oocyte-specific factors bind a conserved upstream sequence required for mouse zona pellucida promoter activity. *Mol. Cell. Biol.* **12,** 6197–6204.

Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., and *et al.* (1989). Interactions between heterologous helix–loop–helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58,** 537–544.

Nordeen, S. K. (1988). Luciferase reporter gene vectors for analysis of promoters and enhancers. *Biotechniques* **6,** 454–458.

Philpott, C. C., Ringuette, M. J., and Dean, J. (1987). Oocyte-specific expression and developmental regulation of ZP3, the sperm receptor of the mouse zona pellucida. *Dev. Biol.* **121,** 568–575.

Ringuette, M. J., Chamberlin, M. E., Baur, A. W., Sobieski, D. A., and Dean, J. (1988). Molecular analysis of cDNA coding for ZP3, a sperm binding protein of the mouse zona pellucida. *Dev. Biol.* **127,** 287–295.

Ringuette, M. J., Sobieski, D. A., Chamow, S. M., and Dean, J. (1986). Oocyte-specific gene expression: Molecular characterization of a cDNA coding for ZP-3, the sperm receptor of the mouse zona pellucida. *Proc. Natl. Acad. Sci. USA* **83,** 4341–4345.

Roller, R. J., Kinloch, R. A., Hiraoka, B. Y., Li, S. S., and Wassarman, P. M. (1989). Gene expression during mammalian oogenesis and early embryogenesis: Quantification of three messenger RNAs abundant in fully grown mouse oocytes. *Development* **106,** 251–261.

Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W., and Staudt, L. M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345,** 686–692.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

Schickler, M., Lira, S. A., Kinloch, R. A., and Wassarman, P. M. (1992). A mouse oocyte-specific protein that binds to a region of mZP3 promoter responsible for oocyte-specific mZP3 gene expression. *Mol. Cell. Biol.* **12,** 120–127.

Scholer, H. R., Dressler, G. R., Balling, R., Rohdewohld, H., and Gruss, P. (1990). Oct-4: A germline-specific transcription factor mapping to the mouse t-complex. *EMBO J.* **9,** 2185–2195.

Scholer, H. R., Hatzopoulos, A. K., Balling, R., Suzuki, N., and Gruss, P. (1989). A family of octamer-specific proteins present during mouse embryogenesis: Evidence for germline-specific expression of an Oct factor. *EMBO J.* **8,** 2543–2550.

Schuler, G. D., Altschul, S. F., and Lipman, D. J. (1991). A workbench for multiple alignment construction and analysis. *Proteins Struct. Funct. Genet.* **9,** 180–190.

Schwoebel, E., Prasad, S., Timmons, T. M., Cook, R., Kimura, H., Niu, E-M., Cheung, P., Skinner, S., Avery, S. E., Wilkins, B., and Dunbar, B. S. (1991). Isolation and characterization of a full-length cDNA encoding the 55 kDa rabbit zona pellucida protein. *J. Biol. Chem.* **266,** 7214–7219.

Shabanowitz, R. B., and O'Rand, M. G. (1988). Characterization of the human zona pellucida from fertilized and unfertilized eggs. *J. Reprod. Fertil.* **82,** 151–161.

Shimizu, S., Tsuji, M., and Dean, J. (1983). *In vitro* biosynthesis of three sulfated glycoproteins of murine zonae pellucidae by oocytes grown in follicle culture. *J. Biol. Chem.* **258,** 5858–5863.

Swenson, C. E., and Dunbar, B. S. (1982). Specificity of sperm–zona interaction. *J. Exp. Zool.* **219,** 97–104.

Von Heijne, G. (1985). Signal sequences: The limits of variation. *J. Mol. Biol.* **184,** 99–105.

Von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14,** 4683–4690.

Wassarman, P. M. (1988). Zona pellucida glycoproteins. *Annu. Rev. Biochem.* **57,** 415–442.