A Sperm Membrane Protein That Binds in a Species-specific Manner to the Egg Extracellular Matrix Is Homologous to von Willebrand Factor*

Daniel M. Hardy† and David L. Garbers§

From the Department of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical School, Dallas, Texas 75235-9050

We have purified a sperm membrane protein, designated zonadhesin, that binds in a species-specific manner to the extracellular matrix (zona pellucida) of the egg, and cloned its cDNA. The cDNA encodes a novel protein with a single transmembrane segment representing a 36 amino acid, highly basic intracellular C terminus from a 2418-amino acid extracellular region. The extracellular sequence specifies a mosaic protein comprising a unique N-terminal domain, a mucin-like domain, and five tandem domains proximal to the membrane that are homologous to prepro von Willebrand factor. The N-terminal and mucin-like domains were absent from zonadhesin that bound to the egg extracellular matrix, suggesting that processing occurs during sperm maturation and/or capacitation. By Northern blotting and in situ hybridization, zonadhesin mRNA was detected only within the testis, where it was expressed primarily in haploid spermatids. The unique domain structure of zonadhesin suggests multiple functions, one of which is to mediate sperm adhesion to the zona pellucida.

Fertilization in both invertebrates and vertebrates is species-specific, and events such as motility activation, gamete adhesion, and induction of the acrosome reaction show absolute or relative species specificity (1–5). Thus, sperm membrane proteins that interact with the egg in a species-specific manner are likely to be primary gamete recognition and/or signaling components. Efforts to reach a molecular understanding of mammalian fertilization have been hampered by the apparent asynchrony of the sperm population; at any given time only a fraction of spermatozoa possess the ability to fertilize the egg. To overcome this problem we used the porcine egg extracellular matrix (zona pellucida) as an affinity medium to isolate proteins from large quantities of pig sperm membranes (6). One of these proteins bound in a species-specific manner to the egg extracellular matrix and migrated in SDS-PAGE at M, 150,000 under nonreducing conditions and as M, 105,000 and 45,000 subunits (p105 and p45) after disulfide bond reduction (6). We have now purified this sperm membrane protein (named zonadhesin), obtained partial amino acid sequence, and isolated its cDNA. Zonadhesin is expressed by the haploid sperm and is homologous to both von Willebrand factor and mucins.

MATERIALS AND METHODS

Zonadhesin Purification—Zonadhesin was purified by batchwise binding to native, particulate zona pellucida (ZP).1 ZP were incubated with detergent-solubilized, biotinylated proteins of pig sperm membranes, then washed consecutively by centrifugation with 1% (w/v) CHAPS in 20 mM NaHEPES, 130 mM NaCl, 1 mM EDTA, pH 7.5, and then with 1% (w/v) Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS in 25 mM NaHEPES, 0.5 mM NaCl, 1 mM EDTA, pH 7.5 (6). Biotinylated zonadhesin was then purified from the ZP-zonadhesin complex by streptavidin-agarose chromatography (6). Western blotting and tryptic peptide isolation and sequencing were as described previously (6, 7). RNA Isolation—Total RNA was isolated by homogenizing tissues in guanidinium thiocyanate and N-lauryl sarcosine, extracting with acidic phenol/CHCl3, and precipitating with isopropyl alcohol (8). The RNA was dissolved in 10 mM Tris-HCl, pH 8.0, 0.2% SDS, 0.1% sodium deoxycholate, 0.1% sodium dodecyl sulfate, and random hexamer-primed pig testis cDNA libraries constructed in λZAP II (10). Stringency washes were 3 × 15 min at 65°C in 1.0 × SSC, 0.2% SDS, 0.1% Na2PO4. Phage clones were rescued as pBluescript plasmids, and inserts were sequenced with or without prior subcloning into M13 filamentous phages. Approximately 1.3 kb of 5′ end cDNA was cloned by 5′-RACE (5′ Amplifier Kit, Clontech) using nested antisense primers (reverse transcription primer: TGGAAAGGT-GGTGCTTTTAAAGAC; downstream PCR primer: GACAGGAAT-TCGAAGGTCTTCTCACTG; upstream PCR primer: GAGCACAATTTCAATACGCGATGCGT; PCR for 5′-RACE (denatured at 94°C for 0 s, annealed at 55–65°C for 0 s, extended at 72°C for 30 s) was done using an air-driven thermal cycler (Idaho Technologies). Primary amplification products (35 cycles with the downstream PCR primer) were excised from low melting point agarose gels, and 0.25% of each was used as template for re-amplification for 30–35 cycles with the upstream PCR primer. RACE products were cloned in pBluescript and sequenced by double-stranded cycle sequencing using an Applied Biosystems automated sequencer. Both strands of cloned cDNAs were sequenced at least once, and the composite sequence was assembled and analyzed using DNASTAR software. The presence of a polyadenylation signal and a poly(A) tail in the 60-base 3′-untranslated sequence and of in-frame stop codons in the 297 base 5′-untranslated sequence confirmed that the cloned cDNAs spanned the entire zonadhesin mRNA.

Northern Blots—Poly(A)+ RNAs from various pig tissues were separated on 1% formaldehyde-agarose gels and blotted overnight to nylon membranes (11). Blots were hybridized with an [α-32P]dCTP-labeled probe from the cloned cDNA.
Fig. 1. Purification of zonadhesin. Lanes 1–3, SDS-PAGE/Western blot of sperm biotinylated proteins (disulfides reduced), detected with streptavidin-peroxidase. Lane 1, 0.2 μg of total membrane protein (starting material for purification of zonadhesin). Lanes 2 and 3, proteins from 20 μg of membrane that bound to native porcine and bovine zona pellucida, respectively. After extensive washing with detergent solutions, only zonadhesin remained bound to the porcine zona pellucida (lane 2). Zonadhesin did not bind to the bovine (lane 3) or mouse zona pellucida (6), nor to the Xenopus laevis egg envelope (6) under these conditions. Lane 4, SDS-PAGE (disulfides reduced) of purified zonadhesin, stained with Coomassie Brilliant Blue R-250.

Table I

| Protein | Peptide sequence | Location |
|---------|------------------|----------|
| p45     | VTYILAQP/LFYYP   | 823–830/920–925 |
|         | YVTLPSTVTLK      | 859–872 |
|         | XLYSSTYOT        | 960–967 |
| p105    | GGNLEAKYVR       | 1235–1244 |
|         | LGASW            | 1349–1354 |
|         | GSYHPVEWDTN      | 1518–1532 |
|         | EQGPPAFYLR       | 1624–1633 |
|         | QVYVD FlynnTLKQGDVLXG    | 1634–1656 |
|         | VSLPATTQIR       | 1658–1667 |
|         | AQEOCAAFQAPAWANCAT | 1777–1795 |
|         | GTFLP VGR        | 1914–1921 |

900-bp EcoRI-Xbal fragment from a partial length cDNA clone. Stringency washes (65°C) were 1 × 10 min with 2 × 3 SSC, 0.1% SDS, and then 2 × 20 min with 0.5 × 5 SSC, 0.1% SDS.

In Situ Hybridization—Paraffin-embedded sections of pig testis (Novagen) were de-waxed and hybridized with digoxigenin-labeled RNA probes, washed at high stringency, and digoxigenin was then detected by incubation with alkaline phosphatase-conjugated anti-digoxigenin followed by color development with 5-bromo-4-chloro-3-indoylphosphate and nitro blue tetrazolium (12, 13). Sense and antisense probes derived from the 1566-bp BamHI-SalI end fragment of the zonadhesin cDNA were synthesized with T3 or T7 polymerases using templates of appropriately linearized pBluescript plasmids containing the 1566-bp insert.

Results and Discussion

Zonadhesin was purified to apparent homogeneity based on its zona pellucida binding activity (Fig. 1). Approximately 30 mg of p105 and 10–15 mg of p45 were obtained from 880 mg of sperm membrane protein, using 44 mg of native, particulate zona pellucida as the affinity matrix. Amino acid sequences of eight p105 and five p45 tryptic peptides (Table I) were not present in existing protein sequence data bases, suggesting that zonadhesin was a novel protein.

Degenerate oligonucleotide primers based on the sequences of two p105 tryptic peptides were used to clone cDNAs encompassing the zonadhesin coding sequence by a combination of PCR, cDNA library screening, and 5′-RACE. The 7785-base composite sequence of the cDNAs contains a single major open reading frame, and the 2476-amino acid sequence deduced from it includes the sequences of the eight p105 and five p45 peptides in its C-terminal two-thirds (Table I and Fig. 2).

A29-amino acid putative signal peptide (14) is present at the N terminus of the deduced sequence (Fig. 2a). The predicted sequence of the tandem D-domains (designated D0–D4) of zonadhesin. Domain boundaries are the same as those defined for the D-domains of prepro-von Willebrand factor (21). Note the characteristic alignment of nearly all of the cysteine residues and presence of the CGLCG motif in D1 and D2, C-terminal sequence. The transmembrane segment is underlined. In b, alignment was by the dystal method using the PAM 250 alignment matrix. Pairwise alignment parameters were window size = 5, gap penalty = 3. Multiple alignment parameters were: gap penalty = 10, gap length penalty = 10.
molecular mass is larger than the sum (M, 150,000) of the apparent sizes of the zonadhesin subunits. The cDNA therefore encodes a precursor protein. Processing of this precursor apparently involves proteolytic generation of p105 and p45 as well as removal of approximately the N-terminal one-third of the protein. Spermatozoa acquire the capacity to fertilize the egg during maturation in the epididymis (1) and capacitation in the female reproductive tract (15); the changes that occur in either environment could involve proteolysis of zonadhesin. A requirement for sperm surface proteolytic activity to facilitate sperm adhesion to the zona pellucida has been reported (16).

A region of highly repetitive sequence is present between amino acids 300 and 700 of zonadhesin; it comprises 53 imperfect repeats of the consensus sequence PTE(K/R)(P/T)T(V/I). Repetitive sequences rich in proline and threonine are characteristic of mucins, which are O-glycosylated on numerous serines and threonines and have extended structures owing to their high proline content (17). These structural properties reflect the functions of mucins in regulating cellular interactions; the large, carbohydrate-rich domains extend beyond most other cell-surface glycoproteins and thereby either inhibit or promote cell adhesion (17, 18). The mucin-like domain of zonadhesin could function similarly during sperm migration through the male or female reproductive tracts, inhibiting inappropriate trapping of spermatozoa or promoting adhesion to the oviductal isthmus, which serves as a sperm reservoir in several species (19, 20).

Five homologous domains, in tandem, follow the mucin-like domain (Fig. 2b). Sequence comparisons (PIR data base) revealed that these domains are homologous to the D-domains of prepro-von Willebrand factor (vWF) (21). Conservation of nearly all of the cysteines within the five domains is readily apparent, and several other amino acids are also conserved at various positions in the sequences. The D0 domain of zonadhesin is about one-fourth the length of the D1–D4 domains, and it truncates at the same point in its sequence as a partial D-domain present in prepro-vWF. The sequence CGLCG of prepro-vWF D1–D3 domains is conserved in the zonadhesin D1 and D2 domains. The vicinal cysteines in this sequence may mediate covalent oligomerization of prepro-vWF monomers (22). Four vWF-like D-domains were recently identified in the human intestinal mucin MUC2 (23), which also exists as disulfide-bonded oligomers and possesses the CGLCG motif in its D1 and D3 domains. We previously observed high M, sperm proteins (SDS-PAGE, nonreducing conditions) that bound to the porcine zona pellucida and appeared to be covalent oligomers of zonadhesin (6). Thus, the CGLCG motif in the D1 and D2 domains of zonadhesin may mediate covalent oligomerization, which in turn could be important in sperm adhesion to the zona pellucida because of the increased binding avidity of multivalent interactions. The zona pellucida is itself a repetitive structure, comprising strands of noncovalently associated multimers of ZP2–ZP3 heterodimers (24), that could interact productively with an oligomeric receptor. Oligomerization could also promote membrane protein aggregation that may be important for induction of the sperm acrosome reaction (25, 26).

Northern blots of poly(A)^+ RNAs from several pig tissues detected expression of a 7.5–8 kb zonadhesin mRNA in testis, but not in heart, liver, lung, spleen, brain, kidney, or epididymis (Fig. 3). Because testis is a heterogeneous tissue containing several different cell types, we localized zonadhesin expression...
more precisely by in situ hybridization (Fig. 4). Some but not all of the seminiferous tubules in testicular sections hybridized strongly with the zonadhesin probe, and this hybridization was restricted to the gerninal epithelia of the tubules. The apparent expression of zonadhesin mRNA by only a fraction of the tubules in a given tissue section probably reflects the asynchrony of spermiogenesis among tubules. Within strongly hybridizing tubules, expression was detected only in the haploid spermatids. Spermatid-specific expression of zonadhesin supports the hypothesis that this protein mediates sperm-specific function(s).

The domain structures of zonadhesin, prepro-vWF, and MUC2 are illustrated schematically in Fig. 5. Each is a mosaic protein (27) and among them zonadhesin is smallest and is unique in having a putative transmembrane segment and its protein (27) and among them zonadhesin is smallest and is MUC2 are illustrated schematically in Fig. 5. Each is a mosaic

The domain structures of zonadhesin, prepro-vWF, and MUC2 are illustrated schematically in Fig. 5. Each is a mosaic protein (27) and among them zonadhesin is smallest and is unique in having a putative transmembrane segment and its protein (27) and among them zonadhesin is smallest and is MUC2 are illustrated schematically in Fig. 5. Each is a mosaic

In addition to their functions in oligomerization, vWF D-domains bind heparin (28). This property is of potential interest, because heparin and/or other glycosaminoglycans promote capitation and the acrosome reaction of bovine and hamster spermatozoa in vitro (29); however, the molecular target(s) for these agents have not been identified. It is possible that the D-domains of zonadhesin mediate these heparin/glycosaminoglycan effects. Like heparin, zona pellucida glycoproteins contain sulfated carbohydrate (30); this further suggests that zonadhesin's binding activity may derive in part from interaction of its D-domains with sulfated carbohydrates in the egg extracellular matrix.

Zonadhesin differs from other potential gamete recognition proteins of spermatozoa for which cDNAs have been done (31–34). The substrate specificity of sperm surface galactosyltransferase, a M, 60,000 plasma membrane-associated variant of the Golgi enzyme, is consistent with its hypothesized function in gamete adhesion (35). Monoclonal and polyclonal antibodies to PH-20, a M, 56,000–64,000 sperm surface hyaluronidase, inhibit adhesion of guinea pig spermatozoa to the zona pellucida (32, 36). Recently, two different proteins potentially involved in mouse (33) and human (34) sperm-egg adhesion (sp56 and ZRK, respectively) were also described. Bookbinder et al. (33) reported a correlation between detectability of sp56 in spermatozoa from various species and ability of the cells to adhere to mouse eggs; species specificity has not been addressed for the other molecules. Thus, zonadhesin is presently the only protein known to bind in a species-specific manner to the egg extracellular matrix.

Acknowledgments—We thank Anita Schultz, Jennifer Craig, Cecelia Green, Debbie Miller, and Lynda Doolittle for technical assistance and Joan Hsu, Carolyn Moonaw, and Prof. Clive Slaughter for peptide isolation and sequencing.

REFERENCES
[References provided in the document]